UNIVERSITY OF CALIFORNIA
RIVERSIDE

AND

SAN DIEGO STATE UNIVERSITY

Lineage Diversification of Lizards (Phrynosomatidae) in Southwestern North America: Integrating Genomics and Geology

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Evolutionary Biology

by

Andrew David Gottscho

August 2015

Dissertation Committee:
Dr. Tod Reeder, Co-Chairperson
Dr. John Gatesy, Co-Chairperson
Dr. Kevin Burns
Dr. Cheryl Hayashi
Dr. Exequiel Ezcurra

The Dissertation of Andrew David Gottscho is approved:

_____

_____

_____

_____

Committee Co-Chairperson

_____

Committee Co-Chairperson

University of California, Riverside
San Diego State University

ACKNOWLEDGEMENTS

Producing this dissertation took me half a decade, and it would not have been possible without the help of numerous colleagues, friends, and family. For Chapter 1, I am indebted to Sean Reilly, Tod Reeder, Scott Westenberger, Sean Harrington, John Andermann, David Wake, and Richard Hey for critically reviewing earlier drafts of this manuscript, and to Brad Hollingsworth, Lee Grismer, Ezequiel Ezcurra, John Gatesy, Kevin Burns, Cheryl Hayashi, Andy Bohonak, Casey Richart, and Dean Leavitt for useful discussion of ideas related to this paper. As of August 2015, this chapter is currently in press (Biological Reviews, doi:10.1111/brv.12167) and copyright permissions have been obtained from the publisher. For Chapters 2 and 3, I would like to thank my dissertation committee (Tod Reeder, John Gatesy, Kevin Burns, Cheryl Hayashi, and Exequiel Ezcurra) and my lab members (Sean Harrington, John Andermann, Dean Leavitt) for critiquing the manuscripts. For tissue loans I thank Amy Vandergast and Dustin Wood (U.S. Geological Survey) for the *Uma* tissues in Chapter 2, and Brad Hollingsworth (Department of Herpetology, San Diego Natural History Museum) for the tissues used in Chapter 3. For help collecting specimens in the field, I am indebted to Jimmy Rabbers, Amelia Meier, Tim Higham, Clint Collins, Sam Murray, Heather Heinz, Sean Harrington, John Andermann, Alexander Stubbs, Eric Stiner, Paul Maier, Kate Rebecca O'Hara, and Sandra Correa. For help with laboratory work and data analyses, I thank Adam Leaché, Jared Grummer, Paul Maier, Evan McCartney-Melstad, and Mark Phuong. I am

indebted to Julio Lemos-Espinal who obtained collecting and export permits in México and helped coordinate the logistics of field research.

For Chapter 2, permits to collect animals were provided by the California Department of Fish and Wildlife (SC-9768), Bureau of Land Management (6500 CA-610-21), the Arizona Department of Game and Fish (SP604B45 CLS), the National Park Service (JOTR-2008-SCI-0004, DEVA-2008-SCI-0013, and MOJA-2008-SCI-0015), and SEMARNAT (SGPA/DGVS/01239/12). For Chapter 3, permits to collect animals were provided by the California Department of Fish and Wildlife (SC-9768) and SEMARNAT (SGPA/DGVS/01239/12, SGPA/DGVS/00916/13, SGPA/DGVS/00976/14). Export permits were provided by PROFEPA (2012-30461, 2013-32393, 2014-34440), and importations were approved by the U.S. Fish and Wildlife Service. Animal welfare protocols were approved by SDSU (APF# 12-04-010R, 14-03-008R). This research was generously funded by UC MEXUS (Dissertation Research Grant), the Anza-Borrego Foundation (Howie Wier Memorial Conservation Grant), the Community Foundation (Desert Legacy Fund), the American Philosophical Society (Lewis and Clark Fund For Exploration and Field Research), and the National Science Foundation (Doctoral Dissertation Improvement Grant, Award Number 1406589).

Beyond the direct roles that the above individuals and organizations played in my research, some people went above and beyond for me during my time as a doctoral student. I owe a special thank you to Tod Reeder, my advisor at SDSU, for taking me under his wing and providing me the opportunity to

pursue my career goals, while giving me intellectual space to grow. I would like to thank all members, past and present, of the SDSU "Squamate Squad" and the communities of Biology graduate students at SDSU and UCR for providing intellectual companionship and camaraderie. I thank Annalisa Berta, Medora Bratlien, Melissa Gomez, and Cheryl Hayashi for helping me navigate the complexities of academic bureaucracy. Finally, I thank my parents (Katherine Paszkolovits and Richard Gottscho) for directly supporting my research in a myriad of ways.

DEDICATION

I dedicate this dissertation to my family, who encouraged me to pursue my passions and dreams.

ABSTRACT OF THE DISSERTATION

Lineage Diversification of Lizards (Phrynosomatidae) in Southwestern North America: Integrating Genomics and Geology

by

Andrew David Gottscho

Doctor of Philosophy, Graduate Program in Evolutionary Biology
University of California, Riverside, August 2015
Dr. Tod Reeder and Dr. John Gatesy, Co-Chairpersons

In the first chapter of my dissertation, I provide an ultimate tectonic hypothesis for several well-studied zoogeographic boundaries along the west coast of North America, specifically along the San Andreas Fault system. Reviewing the literature, I demonstrate that four Great Pacific Fracture Zones correspond with spatially concordant phylogeographic breaks for a variety of marine and terrestrial animals. I hypothesize that the four zoogeographic boundaries reviewed here ultimately originated via the same tectonic process (triple junction evolution along the San Andreas Fault system), and I suggest how a comparative phylogeographic approach can be used to test this hypothesis. In the second chapter, I investigate the systematics and species delimitation of fringe-toed lizards of the *Uma notata* complex in the deserts of southwestern North America. Ten nuclear loci were Sanger sequenced and genome-wide sequence and single-nucleotide polymorphism (SNP) data were collected using restriction-associated DNA (RAD) sequencing. I validated five

species-level lineages within the *U. notata* complex, three of which were previously described as full species, one originally described as a subspecies but later synonymized, and one previously documented yet undescribed species from Mohawk Dunes, Arizona, USA. The results support the hypothesis that Pleistocene glacial cycles promoted allopatric speciation via dispersal across a landscape matrix shaped by older tectonic events, but I also recovered evidence for a vicariant role of the Colorado River during the Pleistocene epoch. In the third chapter, I studied the comparative phylogeography of lizards (Phrynosomatidae) of the Baja California Peninsula, including the genera *Callisaurus*, *Petrosaurus*, *Urosaurus*, and *Sceloporus*. I collected sequence/SNP data from 228 individual lizards from the peninsula and eight islands using RADseq. The estimated divergence dates across co-distributed clades revealed that for each of the five regions where comparisons are possible, at least two or three episodes of divergence are required to explain the observed patterns. I tested for range expansions or bottlenecks associated with Pleistocene glacial cycles or continent-to-peninsula invasions. The results are consistent with the hypothesis that plate tectonic events established a complex landscape matrix with numerous barriers to dispersal that ultimately facilitated divergence, generating the idiosyncratic patterns observed among co-distributed lineages.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Southwestern North America is an ideal place to study evolutionary biology, particularly historical biogeography, due to the topographic complexity of the landscape, which has ultimately been shaped by the region's dynamic geological history. The San Andreas Fault (SAF) system, extending from Cape Mendocino in northern California southeast along the California coast through the Gulf of California to Cabo Corrientes in central México, is the defining geological feature of the region, an active tectonic boundary that separates the Pacific and North American plates – although as the reader shall learn in Chapter 1, this is a recent arrangement. The SAF system (as broadly defined, see Appendix A), extending from approximately 40° N to 20° N, encompasses a diverse array of terrestrial and marine ecosystems, including temperate and tropical forests, woodlands, grasslands, chaparral, thorn scrub, deserts, rivers, lakes, wetlands, sand dunes, mountains, plains, volcanoes, beaches, bays, kelp forests, reefs, and hundreds of islands of varying ages, sizes, and geological types. Many of these complex landscape and seascape features are directly associated with tectonic activity along the SAF system – even oceanic currents are influenced by prominent capes and headlands. As such, this system represents a fascinating arena to test hypotheses regarding the interactions of geological history and biological evolution. Yet there is still much we do not understand about the complex history of this region.

In the first chapter, I start by reviewing the literature on the zoogeography of the SAF system. This chapter is not meant to be exhaustive – rather, it is meant to synthesize across disciplines to present a novel tectonic hypothesis regarding the role that the Great Pacific Fracture Zones (GPFZs) played in the tectonic evolution of the SAF system. I first review what is known about the geological history of the SAF system before describing the fundamental pattern of interest, that is, how four GPFZs in the Pacific plate correspond with phylogeographic boundaries or distributional limits for a variety of marine and terrestrial animals, mostly those with limited dispersal abilities (e.g. plethotondid salamanders and surfperches). I hypothesize that the evolution of triple junctions along the SAF has ultimately resulted in the complex topography of western North America, including distinct east-west irregularities, such as Point Conception and the Transverse Ranges, along the dominant northwest-southeast orientation of coastal features. It is clear that these major biogeographic boundaries have in turn influenced the diversification of the local fauna, however, additional types of data are needed to test the alternative proximate mechanisms of speciation (e.g. vicariance and dispersal). I conclude the first chapter by suggesting how large genetic/genomic datasets can be used to explicitly test these hypotheses using a comparative phylogeographic approach.

In the second chapter, I turn my attention to the systematics and species delimitation of fringe-toed lizards (*Uma notata* species complex) of the southwestern United States and northwestern México. These lizards are distributed across the SAF where it transitions from a strike-slip fault to an

oceanic spreading center. Rifting of the SAF fault system has created a wide, deep basin (the Salton Trough), which has accumulated sediment from the Colorado River and Gulf of California over the past five million years. Today, lizards of the *U. notata* complex occupy a disjunct series of sand dunes across the Colorado Desert region, including the Gran Desierto, the largest sand dune field in North America. Because these lizards have allopatric distributions, the classic "species problem" comes into play – what is a species, and how do we define them in practice? The results of this theoretical question may have important conservation implications. For example, the Coachella Valley fringe-toed lizard (*Uma inornata*) is listed as a federally threatened/state endangered species, having lost the majority of its habitat to development, and its legal protection and survival ultimately result from its recognition as a discrete evolutionary lineage. It is possible that the discovery and validation of additional putative species within the *U. notata* complex may result in changes in management practices. In this chapter, I used both Sanger sequence data and restriction-associated-DNA sequencing (RADseq) data to investigate species delimitation within this complex using Bayes factors. The results may have important implications for the conservation of this complex, as the best supported model suggested the existence of five species, only three of which are currently recognized.

The third chapter of my dissertation focuses on the comparative phylogeography of phrynosomatid lizards of the Baja California Peninsula (BCP). Rugged, arid, and sparsely populated, this peninsula has long been an

enigma for phylogeographers and biogeographers. Here, desert and sea meet along thousands of kilometers of corrugated coastline, with numerous islands scattered offshore. Phrynosomatid lizards are ideal candidates for phylogeographic studies of the BCP. These lizards are abundant, diverse, and widely distributed along the length of the peninsula, and they are generally good at dispersing across salt water. Furthermore, previous studies have found numerous cases of deep lineage splits along the length of the BCP, especially in the Vizcaíno Desert, where one or more controversial mid-peninsular seaways have been hypothesized. A key question revealed through previous studies using mitochondrial DNA concerns the existence of spatially concordant genetic breaks: do multiple species or clades that show spatial concordant breaks also show temporal concordance? That is, did they diverge at the same times in the same areas? If a single period of higher sea levels during the Miocene or Pliocene fragmented the BCP into an archipelago, we might expect to find this to be the case. Alternatively, if different lineages diversified independently across a landscape matrix previously shaped by tectonic events (such as triple junction evolution), we might find evidence of multiple episodes of divergence. I tested these alternative hypotheses using RADseq data collected for four genera of phrynosomatid lizards that are widely distributed and have endemic species or clades along the BCP: the zebra-tailed lizards (*Callisaurus*), banded rock lizards (*Petrosaurus*), brush lizards (*Urosaurus*), and spiny lizards (*Sceloporus*). The results of this chapter suggest that although many taxa share spatially concordant phylogeographic boundaries, especially the Vizcaíno Desert where the Molokai

and Shirley Fracture Zones are aligned with Punta Eugenia and a transverse belt of volcanoes, for the most part divergence dating analyses suggest that multiple waves of divergence are required. Thus, it appears that these four genera evolved in idiosyncratic ways over the same landscape matrix previously shaped by tectonic events, getting back to the proximate mechanisms suggested in Chapter 1.

It is my hope that these three chapters provide the reader a novel perspective on the interplay between biological evolution and geological evolution along the SAF system, and may ultimately have important implications for biodiversity management.

Chapter 1: Zoogeography of the San Andreas Fault System: Great Pacific Fracture Zones Correspond with Spatially Concordant Phylogeographic Boundaries in Western North America

The purpose of this article is to provide an ultimate tectonic explanation for several well-studied zoogeographic boundaries along the west coast of North America, specifically, along the boundary of the North American and Pacific plates (the San Andreas Fault system). By reviewing 177 references from the plate tectonics and zoogeography literature, I demonstrate that four Great Pacific Fracture Zones (GPFZs) in the Pacific plate correspond with distributional limits and spatially concordant phylogeographic breaks for a wide variety ofmarine and terrestrial animals, including invertebrates, fish, amphibians, reptiles, birds, and mammals. These boundaries are: (1) Cape Mendocino and the North Coast Divide, (2) Point Conception and the Transverse Ranges, (3) Punta Eugenia and the Vizcaíno Desert, and (4) Cabo Corrientes and the Sierra Transvolcanica. However, discussion of the GPFZs is mostly absent from the zoogeography and phylogeography literature likely due to a disconnect between biologists and geologists. I argue that the four zoogeographic boundaries reviewed here ultimately originated via the same geological process (triple junction evolution). Finally, I suggest how a comparative phylogeographic approach can be used to test the hypothesis presented here.

I. INTRODUCTION

"*One of the joys of science is that, on occasion, we see a pattern that reveals the order in what initially seems chaotic*." – Shubin (2009, p. 82)

The theory of plate tectonics has undoubtedly enabled historical biogeographers to explain many of the ultimate processes driving global biogeographic patterns (Riddle and Hafner 2010, Bagley and Johnson 2014). For example, Wallace (1876, p. 389) noted that the fauna of Celebes (Sulawesi), an oceanic Indonesian island, "presents the most puzzling relations, showing affinities to Java, to the Philippines, to the Moluccas, to New Guinea, to continental India, and even to Africa; so that it is almost impossible to decide whether to place it in the Oriental or the Australian region." Unbeknownst to Wallace, Sulawesi was sutured from several palaeo-islands as the Australian and Asian plates converged (R. Hall 2002), elegantly explaining the geographically concordant genetic breaks observed in anurans, primates, and squamate reptiles (Evans et al. 2003, McGuire et al. 2007a). Similarly, disjunct tropical lineages that occur in the southern cape region of Baja California, México posed difficulties before the discovery of plate tectonics. It was assumed that the tropical taxa must have dispersed to the cape from mainland México, either overland through the northern Colorado Desert or by rafting across the Gulf of California (Savage 1960). Plate tectonic models subsequently revealed that Baja California was attached to North America 5–6 million years ago (Mya) and rifted apart from the

continent along the San Andreas Fault (SAF) system, suggesting that vicariance was also important in structuring the disjunct tropical communities of the cape region (Moore and Buffington 1968, Stock and Hodges 1989, Grismer 2002). Likewise, the confusing array of locally endemic slender salamanders (*Batrachoseps pacificus* species complex) in coastal California defied interpretation until molecular phylogenetic analyses (Jockusch et al. 2001) combined with palinspastic reconstructions of California during the late Cenozoic (C. A. Hall 2002) revealed that microplate capture and strike-slip displacement along the SAF explains much of the historical biogeography of this group (Wake 2006). Despite the broad acceptance of plate tectonics, the globally unique SAF system, which resulted from the late Cenozoic interaction of an oceanic spreading centre with a coastal subduction zone (Atwater 1989), remains a challenge for interpretation.

The purpose of this synthesis is to describe a broad zoogeographic pattern along the SAF system and to propose a tectonic process that might explain it, relying on published biological and geological literature. The pattern concerns four Great Pacific Fracture Zones (GPFZs) in the Pacific plate, each of which is aligned with a major zoogeographic boundary or "break" along the SAF system. These include (1) Cape Mendocino and the North Coast Divide in northern California, (2) Point Conception and the Transverse Ranges of southern California, (3) Punta Eugenia and the Vizcaíno Desert of Baja California, and (4) Cabo Corrientes and the Sierra Transvolcanica of central México (Figure 1.1). The GPFZs are straight, parallel, and evenly spaced, implying a common mechanism

for their origin. Examining this pattern on a map immediately raises questions of ultimate causation. Why are the GPFZs aligned with dominant physiographic structures of western North America, which in turn correspond with well-known zoogeographic boundaries? In some cases, the proximate mechanisms promoting divergence across these boundaries are well understood. For example, Cape Mendocino, Point Conception, and Punta Eugenia divert the dominant California Current offshore, establishing upwelling jets and warmer eddies in their leeward waters, which act as filter barriers to planktonic larvae, promoting ecological and genetic divergence in nearshore communities (Kelly and Palumbi 2010, Haupt 2011). In other cases, the proximate mechanisms promoting divergence are hotly debated, especially for terrestrial taxa spanning the Transverse Ranges (Chatzimanolis and Caterino 2007) and the mid-peninsular break of Baja California (Leaché et al. 2007). However, the phylogeography literature largely has not addressed the relationship between the GPFZs and biogeographic boundaries in North America (but see Cope 2004, Reilly and Wake 2015).

Here, I attempt to fill this void by explaining how biological evolution has recapitulated tectonic evolution throughout the SAF system, specifically, how the migration of the Mendocino and Rivera triple junctions and the positioning of the GPFZs shaped the distribution of zoogeographic provinces along the west coast of North America. Due to the voluminous geological and biological literature relevant to this region, this review is not intended to be exhaustive. Instead, I selected what I consider to be the most critical plate tectonics studies regarding the GPFZs and SAF system, as well as exemplar zoogeographic studies

representing a diverse array of marine and terrestrial animals along the coast. Therefore, the reader should be forewarned that uncited counter-examples may exist. Despite this limitation, this synthesis is needed because although the GPFZs indicate a relatively simple ultimate process that may explain complex distributional patterns, they have received scant attention in the North American zoogeography literature.

II. THE GREAT PACIFIC FRACTURE ZONES AND THE SAN ANDREAS FAULT SYSTEM

Before describing the aforementioned pattern, it is first necessary to consider the scientific history of the GPFZs: their discovery, their relevance to the theory of plate tectonics, and their impact on the evolution of the SAF system. A glossary of geological terms and landscape features is provided (see Appendix A). I encourage the reader to supplement the figures provided here by using free software such as Google Earth (http://earth.google.com) interactively to examine oceanic and continental features in three dimensions and to identify the various landscape features in the glossary.

The GPFZs played an important role in the discovery of plate tectonics. Although Wegener (1912) originally proposed continental drift, there was a paucity of direct evidence supporting his hypothesis because little was known about the structure of the oceanic lithosphere at the time. However, sonar technology enabled oceanographers to accumulate data supporting the theory of

continental drift. Early surveys revealed the Mendocino Fracture Zone (FZ) offshore northwestern California, marking the northern end of the SAF (Murray 1939, Shepard and Emery 1941, Menard and Dietz 1952). Following additional expeditions, Menard (1955) described the Mendocino, Murray, Clarion, and Clipperton FZs, which assume the form of semi-circles ranging between 2,250 and 5,310 km in length and covering nearly 20.7 million km^2, "characterized by great seamounts, deep narrow troughs, asymmetrical ridges, and escarpments" (p. 1149). Menard (1955) observed that prominent geomorphic provinces of western North America, including the Cascades, SAF, California Coast Ranges, Great Central Valley, Sierra Nevada, Continental Borderland, Baja California Peninsular Ranges, Gulf of California and Guatemalan Trench, are bounded by GPFZs or their continental extensions, while the Transverse Ranges and Sierra Transvolcanica appear to be extensions of the Murray FZ and Clarion FZ, respectively (Figures 1.2 and 1.3). Menard (1955) also predicted the existence of the Molokai FZ, later confirmed by Smith and Menard (1965), who recognized that the Molokai FZ delimits the southern end of the Continental Borderland. Menard (1960, p. 1742) proposed a "genetic relationship" between the GPFZs and the East Pacific Rise (EPR), an area with 2–3 km of vertical relief stretching 21,000 km from México to New Zealand. He also suggested that the EPR may support continental drift, and proposed that the EPR might extend underneath North America between Cabo Corrientes and Cape Mendocino.

Assembling this new evidence, Dietz (1961, p. 855) proposed the seafloor-spreading mechanism wherein "the median rises mark the up-welling sites or

divergences; the trenches are associated with the convergences or down-welling sites; and the fracture zones mark shears between regions of slow and fast creep." Dietz was correct regarding the origin of rises and trenches, but incorrect regarding fracture zones, which can arise with identical creep rates (Wilson 1965). Mason and Raff (1961) discovered that magnetic anomalies were offset 1,111 km across the Mendocino FZ and 155 km across the Murray FZ; Vine and Matthews (1963) later related these magnetic anomalies to reversals of polarity of the earth's magnetic field in a newly formed sea floor at an oceanic ridge. This prediction finally provided a way to test Wegener's (1912) theory of continental drift, which was confirmed by detailed surveys of anomalies across the Pacific-Antarctic and Atlantic Ridges (Pitman and Heirtzler 1966, Vine 1966).

Wilson (1965) proposed the first qualitative model of plate tectonics, recognizing that the seismic regions of Earth's surface are concentrated in narrow, mobile belts that form a global network separating rigid plates. Wilson defined transform faults as the seismic sections of fracture zones (confined to the zone between offset ridges) where the plates slide past each other in opposite directions, and was the first correctly to interpret the SAF system as an unusual transform fault separating the Pacific and North American plates. Wilson (1965, p. 346) suggested that "the offsets in the magnetic displacements observed in the aseismic fracture zones off California may not be fault displacements as has usually been supposed, *but that they reflect the shape of a contemporary rift in the Pacific Ocean*" (my emphasis). Menard and Atwater (1969) confirmed that the GPFZs resulted from transform faulting along spreading centres, explaining the

large offsets in magnetic anomalies across them. Vine (1966, p. 1411) recognized that "the more recent geologic history and structures of the western United States can be ascribed to the progressive westward drift of the North American continent away from the spreading Atlantic Ridge, and to the fact that the continent has overridden and partially resorbed first the trench system and more recently the crest of the East Pacific Rise." This was a major conceptual shift in our understanding of the tectonic mechanisms shaping the topography of the globe.

By 1970, the basic tectonic history of North America and the SAF system had emerged (McKenzie and Parker 1967, Moore and Buffington 1968, McKenzie and Morgan 1969, Atwater 1970, Atwater and Menard 1970). In summary, from the Mesozoic to the late Oligocene, the west coast of North America was bordered by an unbroken subduction zone (the Farallon Trench) with an associated volcanic arc, similar to the modern Andes of South America (Figure 1.4A). The granitic batholiths of the modern Sierra Nevada and Baja California Peninsular Ranges represent the eroded core of this ancient cordillera, and the chaotic coastal Franciscan formations are the jumbled remains of islands, seamounts, sediments, and metamorphic rock accreted to the North American plate from the subducted Farallon plate (Furlong and Schwartz 2004). The Farallon plate was separated from the Pacific plate to the west by the EPR and from the North American plate by the Farallon Trench (McKenzie and Morgan 1969). Throughout the Cretaceous into the Cenozoic, the Farallon plate grew narrower as the trench steadily devoured it. During this period, the EPR "was

offset by an especially long transform fault along the Mendocino fracture zone so that the section of ridge between the Mendocino and Murray zones formed a prominent eastward projection" (Atwater and Menard 1970, p. 449). This eastern projection finally collided with the Farallon Trench ~27–28 Mya, allowing the North American and Pacific plates to contact directly for the first time, establishing the Mendocino and Rivera triple junctions (Atwater 1989), where three plate boundaries meet (Figures 1.1 and 1.4B). The northern Farallon plate fragment would eventually become the Juan de Fuca plate and the southern Farallon plate would eventually give rise to the Cocos plate, each with its own subduction zone (McKenzie and Morgan 1969). The collision of the EPR with the west coast of North America changed the direction of relative plate motion to oblique transtension and created the modern SAF system, the primary structural puzzle of California (Atwater 1970). As more of the Pacific plate came into contact with the North American plate, the triple junctions migrated away from each other, lengthening the SAF (Figure 1.4C). Atwater and Menard (1970, p. 449) noted, "the timing of the origin of the various sections of the San Andreas system may be related to the timing of the ridge-trench collisions" and calculated that the northern SAF system formed about 27 Mya, the southern California section between 13 and 22 Mya, and the Gulf of California segment approximately 11 Mya. Today, the Mendocino triple junction (MTJ) marks the northern limit of the SAF system (Figure 1.4D), while the Rivera triple junction (RTJ) marks the southern limit.

Yet at least one piece of the puzzle regarding the tectonic evolution of the SAF system remained unsolved at the time: why do the GPFZs correspond with major physiographic provinces on the west coast? In other words, why are the SAF, Coast Ranges, Central Valley, Sierra Nevada, Mojave Desert, Continental Borderland, Baja California Peninsula, and Gulf of California effectively bounded by the GPFZs, and why do the GPFZs correspond with major headlands and transverse mountain belts on the continental margin as first described by Menard (1955)? McKenzie and Morgan (1969, p. 131) recognized that their novel triple junction model alone "is no help in understanding how the Central Valley was formed, nor why it extends from the extension of the Murray [FZ] to that of the Mendocino [FZ]. It is particularly difficult to understand the relation of the Valley to the oceanic transform faults if the San Andreas in the coast ranges has a displacement of at least 300 km." Von Huene (1969, p. 497) concluded that "even though the existing continental margin has no geophysical indications of an intersection with the Murray, the alignment of the Murray with the unique embayment into the continental slope and the Transverse Ranges is a phenomenon difficult to attribute to chance. *A similar alignment of continental features with the Mendocino fracture zone and the Clarion fracture zone strengthens this argument*. The idea is offered that a more extensive pre-Miocene Murray fracture influenced the trend of subsequent deformation in the San Andreas fault system to form the present Transverse Ranges" (my emphasis). Specifically, he proposed (p. 495) "the Transverse Ranges developed in response to stress that produced the San Andreas fault system along a pre-existing zone of weakness which in this

case could have been a transform fault thrust under the continent." However, his paper received little attention, having been cited only three times as of March 2014 (the last time in 1977), in part because the true relationship between the Murray FZ and the Transverse Ranges was far more complex. As Norris and Webb (1990, pp. 463–464) pointed out, "many geologists have sought to relate the [Murray FZ] to the Transverse Ranges because of their similar east-west orientations. But to date no really persuasive evidence has been found to corroborate such a connection."

We now know that much of the EPR, including the Murray FZ, was not simply subducted by the Farallon Trench as Von Huene (1969) hypothesized. Detailed tectonic analyses have subsequently revealed a more complex history involving multiple propagating rifts and microplates rearranging the accretional plate boundary geometry, so that not all the ridge segments reached the trench (Nicholson et al. 1994, Hey 1998). Analysis of offshore magnetic anomaly patterns has shown that at least five or six microplates were calved from the subducting Farallon plate or its daughter plates (Figure 1.5). From oldest to youngest and from north to south, these microplates include the Monterey (independent from 30 to 19 Mya), Arguello (20–18 Mya), Guadalupe (20–14 Mya), Magdalena (14–12 Mya), and Rivera (5 Mya to present) microplates (Atwater 1989, Lonsdale 1991, Stock and Lee 1994). The Arguello microplate continued to subduct until its final demise at about 18 Mya (Nicholson et al. 1994). The Guadalupe and Magdalena microplates were subducted beneath Baja California until the mid-Miocene, after which time the Pacific plate captured

their remnants as oblique transtensional rifting was initiated in the Gulf of California (Dickinson 1997). The Monterey microplate continued to subduct beneath North America until the early Miocene, at which time the Murray FZ was aligned with the RTJ as the western Transverse Ranges began to rotate clockwise (Luyendyk et al. 1980, Luyendyk 1991). The rotation of the Transverse Ranges to its current position was related to the spatial and temporal alignment of the Murray FZ and the RTJ with the Monterey microplate when rotation began in the late Oligocene, although the exact mechanism is controversial (Nicholson et al. 1994, Hey 1998). Much less is known about the parallel situation with the Molokai/Shirley FZs and the Vizcaíno Peninsula in central Baja California (Stock and Lee 1994).

To conclude this section, although much has been discovered about the history of North America since the plate tectonics revolution began, and the evolution of the SAF system remains an extremely active topic in plate tectonics today, there remains a large disconnect between the geological literature (especially regarding the GPFZs) and the biological literature covering the region. In an effort to bridge this gap, I will review each of the four biogeographic breaks that are aligned with GPFZs before hypothesizing how biological evolution has recapitulated geological evolution in western North America.

III. THE MENDOCINO FRACTURE ZONE, CAPE MENDOCINO, AND THE NORTH COAST DIVIDE

(1) Geology and geography

The Mendocino FZ (Figures 1.1 and 1.2) extends ~4500 km across the North Pacific Ocean, trending ENE from the vicinity of Martin Seamount north of the Midway Islands (35º06′ N, 175º40′ W) to Cape Mendocino in Humboldt County, California (40º26′ N, 124º24′ W). The MTJ, where the Juan de Fuca, Pacific, and North American plates meet, lies just offshore Cape Mendocino (McKenzie and Morgan 1969), one of the most seismically active regions on Earth (Furlong and Schwartz 2004). The MTJ marks a fundamental shift in tectonic regime along western North America from the Cascadia subduction zone north of the triple junction to the SAF to the south (Lock et al. 2006). A line following the trend of the Mendocino FZ inshore defines the southern limit of the Cascades (Mount Lassen and Mount Shasta), part of the Cascadia subduction zone. The Mendocino FZ and its inward extension also define the northern limit of the Sierra Nevada, Central Valley, and California Coast Ranges. The effects of the northward migration of the MTJ are strongly imprinted in the geology of coastal California. Passage of the triple junction first produced a thickening of the crust in advance of the triple junction and then thinning of the crust after its passage, explaining why the vertical relief peaks at Cape Mendocino and diminishes gradually to the south (Furlong and Schwartz 2004). Passage of the MTJ and its associated wave of uplift also influenced the pattern of river drainages along the

Coast Ranges (Lock et al. 2006). The physiography of Cape Mendocino strongly influences coastal substrate structure for hundreds of kilometers on either side. North of the cape, numerous rivers deposit sediment along the coast, which faces northwest and thus readily captures sand transported south by the prevailing northwesterly ocean swells and currents. Thus, beaches and dunes dominate the topography for over 130 km to the north of Cape Mendocino, confining exposed reefs to a few areas. By contrast, south of Cape Mendocino to Sonoma County, the steep, rocky coastline faces southwest and the coastal ridges paralleling the SAF restrict coastal sediment input to that of a few short streams. Thus, vertical bluffs and reefs dominate the coastline, while sandy beaches are a minor feature of the topography.

(2) Marine zoogeography

Cape Mendocino has been recognized as a marine zoogeographic boundary for over a century, first described based on the distribution limits of intertidal mollusks (Dall 1899, Smith 1919). Recently, a number of molecular studies have discovered phylogeographic breaks in this region in a wide variety of taxa including fish, crustaceans, molluscs, and the zooxanthellae of anemones (Table 1.1). Kelly and Palumbi (2010) conducted a comparative phylogeographic analysis of six species of hermit crabs, shore crabs, limpets, and barnacles using mitochondrial DNA (mtDNA), and concluded that "the upwelling center of Cape Mendocino, historically neglected by genetic sampling, houses a number of interesting genetic shifts even for high dispersal species" (p. 2). The most

comprehensive survey to date of Cape Mendocino as a phylogeographic barrier for marine animals is part of an unpublished doctoral dissertation (Haupt 2011) that reviewed 45 animals with distributions that spanned Cape Mendocino, and found that 53% of the sampled species (10/13 arthropods, 8/14 fish, 4/14 mollusks, and 2/4 echinoderms) showed significant genetic structure associated with the cape. The proximate causes for this biogeographic pattern are well understood – Cape Mendocino is a barrier to planktonic dispersal by diverting the California Current offshore, forming a leeward eddy (Hayward and Mantyla 1990). Cape Mendocino also has among the strongest upwelling in the California Current, making dispersal around the cape difficult or impossible for planktonic larvae during the peak upwelling season (Parrish et al. 1981). Additionally, Cope (2004) and Sivasundar and Palumbi (2010) speculated that the Mendocino FZ itself may act as a direct barrier to dispersal for some rockfish species.

(3) Terrestrial zoogeography

A number of terrestrial species, mostly with poor dispersal abilities, have northern or southern range limits that correspond with the MTJ or its inland extension, which delimits the northern end of the Sierra Nevada and Central Valley and the southern limit of the Cascades (Table 1.1). In addition, several subspecies (a few of which have now been elevated to full species status) integrate across this zone, and molecular studies have revealed phylogeographic breaks within several species that span Cape Mendocino or further east near the junction of the Sierra Nevada and Cascades, along the trend of the Mendocino FZ

(Table 1.1). For this reason, this region has been referred to as the 'North Coast Divide' (Reilly et al. 2012). Perhaps the most clear-cut example of a phylogeographic break associated with the MTJ occurs in the black salamander (Reilly and Wake 2015). Genetic lineages north and south of the MTJ are strongly differentiated. Furthermore, a second genetic divide to the south in Mendocino County coincides spatially and temporally with the former position of the MTJ 2Mya (Lock et al. 2006). To my knowledge, this is also the only terrestrial phylogeographic study wherein the authors have directly implicated the Mendocino FZ in their interpretation of the genetic results.

IV. THE MURRAY FRACTURE ZONE, POINT CONCEPTION, AND THE TRANSVERSE RANGES

(1) Geology and geography

The Murray Fracture Zone (Figure 1.2) extends ~4230 km from just east of the Midway Islands (27°20′ N, 167°10′ W) to 270 km west of Point Conception in Santa Barbara County, California (34°27′ N, 120°27′ W), the western extension of the Transverse Ranges. Magnetic anomaly isochrons are offset about 155 km across the Murray Fracture Zone (Mason and Raff 1961). The Transverse Ranges delimit the southern extent of the "Alta California" geomorphic province (the Coast Ranges, Central Valley, Sierra Nevada), as well as the Mojave Desert and the Deep Province of the Pacific. The Murray FZ also defines the northern limit of the Continental Borderland (Southern California Bight), Baja California

Peninsular Ranges, and Gulf of California geomorphic provinces (Menard 1955). The Transverse Ranges are extremely complex, heavily faulted and consist of several distinct ranges on both sides of the SAF (Luyendyk et al. 1980). The western ranges include the Santa Ynez, Sierra Madre, San Rafael, Sierra Pelona, Santa Monica, and the San Gabriel ranges, while the San Bernardino and Little San Bernardino ranges extend east into the North American plate. The SAF bends sharply here before transitioning into a spreading centre near the Colorado River delta.

The Transverse Ranges have intrigued and puzzled geologists and biologists alike since before the advent of plate tectonics. Weaver and Doerner (1967, p. 13) noted: "The Transverse Ranges as a whole form an anomalous feature to the general north-south trending structures of western North America. The Murray Escarpment, trending westward 1900 miles out to sea, is believed to be a structural extension of this continental anomaly. This east-west trend terminates to the east against the great San Andreas fault system and the San Bernardino Range. Not only is it of interest for its obvious structural significance as it is related to one of the great features of the earth's crust, but this transverse feature also has strong biogeographical implications… These east-west trending mountain ranges, at times partly submerged, partly emerged, served as the limits on more than one occasion for new biogeographic provinces…Thus the Transverse Ranges serve as a focal point of considerable significance not only for structural but for biogeographic studies as well." Norris and Webb (1990, p. 301) likewise recognized their importance: "Unlike California's other geological

provinces, the Transverse Ranges form a conspicuous east-west trending unit… It is interesting to note that in North America as a whole, apart from Alaska, there are few prominent east-west mountain belts. This is no accident and when we have a better understanding of the cause of the anomalous east-west ranges, we will also have a more complete understanding of the tectonic history of the continent."

Palaeomagnetic data collected from Miocene igneous rocks indicate that the western Transverse Ranges have rotated clockwise up to 110° from their original orientation (Luyendyk et al. 1980, Luyendyk 1991). There are two alternative models that explain how rotation of the Monterey plate produced the current configuration of the Transverse Ranges, Point Conception, and the Continental Borderland. Nicholson et al. (1994) hypothesized that the unsubducted remnant of the Monterey microplate (bounded on the south by the Murray FZ) was captured by the Pacific plate and rotated clockwise throughout most of the Miocene, and that transtension in the wake of the rotating Transverse Ranges was responsible for the formation of the Continental Borderland including the Los Angeles Basin during the Miocene. Finally, rifting of the Baja California Peninsula resulted in transpressional uplift during the Pliocene and Pleistocene. Alternatively, Hey (1998) speculated that propagating rifts (Hey et al. 1989) may have transferred subducted lithosphere to the Pacific plate during the evolution of the SAF system, pointing out that "that the Transverse Ranges and the Gulf of California occur specifically at the only two places along the margin that appear to have been intersected by left-offset propagators, and these

intersections correlate temporally with significant structural changes in these areas" (p. 248). Regardless of the exact mechanism, it is now apparent that the current structures of the Transverse Ranges are related to the spatial and temporal alignment of the Murray FZ, the RTJ, and the Monterey microplate in the late Oligocene when rotation began (Nicholson et al. 1994, Hey 1998).

(2) Marine zoogeography

Point Conception, where the southward-flowing California Current meets the warmer, less-productive Southern California Eddy (Parrish et al. 1981), has long been recognized as a zoogeographic boundary for marine invertebrates and fish, both extinct (Smith 1919, C. A. Hall 1964, 2002) and extant (Dall 1899, Steinbeck and Ricketts 1941, Valentine 1966). Horn and Allen (1978) analysed the distributions of 280 California non-bay fish species, and found that "102 species (41.3% of all species occurring south of Point Conception) occur exclusively south of Point Conception while only 32 species (18.1% of all species occurring north of Point Conception) occur exclusively north of Point Conception" (p. 30). Similarly, Burton (1998) found that of 21 species of gastropods, barnacles, and algae having range limits within 200km on either side of Point Conception, 11 had range end points within 10 km of the Point. Wares et al. (2001, p. 297) concluded that "Point Conception is a biogeographic boundary for species with pelagic larval development only' and that 'Point Conception has a disproportionately large number of northern range limits relative to southern range limits." Blanchette et al. (2008) investigated biogeographic patterns of 296

marine taxa in California and found that "127 were found exclusively north, and 47 exclusively south, of Point Conception" (p. 1603). C. A. Hall (2002), utilizing palinspastic reconstructions of California since the late Oligocene, recognized the Point Conception area as a molluscan provincial boundary between the Californian and the Oregonian provinces since the late Miocene (~5 Mya) when rotation and transpression of the Transverse Ranges separated warm temperate from temperate water masses.

Despite the long history of Point Conception as a major distributional boundary for marine animal species, the role of Point Conception as an intraspecific phylogeographic break has been contested (Table 1.2, Haupt 2011). Burton (1998, p. 734) argued that there are "no cases of intraspecific phylogeographic breaks associated with this well-recognized biogeographic boundary." Instead, he suggested (p. 743) that "Point Conception's strong impact on species distributions probably derives from its oceanographic position as a boundary between warm and cold water masses" and that (p. 734) the "concordance of phylogeography with biogeography will only be pronounced where the biogeographic boundary separates biotas that are phylogenetically related." Dawson (2001) found that three species of marine invertebrates and fish in the California Current system showed a break in the Los Angeles Region, and concluded that this region was more important than Point Conception. Pelé et al. (2009) found that three of six species with a planktonic larval phase demonstrated genetic breaks at Point Conception (consistent with Wares et al. 2001), while 5 of 10 direct developers showed breaks at the Los Angeles Region

instead. From the perspective of this review, both Point Conception and Los Angeles are associated with the Transverse Ranges and thus both are aligned with the Murray FZ, so this debate does not compromise the broader view of west coast biogeography presented here.

(3) Terrestrial zoogeography

As the Transverse Ranges are essentially the "structural knot" of California, where the Coast Ranges, Central Valley, Sierra Nevada, Mojave Desert, Colorado Desert, and Peninsular Ranges intersect, numerous species reach their northern or southern range limits along these ranges (Table 1.2). Phylogeographic breaks in the Transverse Ranges are well documented in the terrestrial literature (Calsbeek et al. 2003). Some species capable of either dispersing large distances or with poor geographic sampling exhibit generalized genetic breaks in the Transverse Ranges, but most studies that have found breaks within these ranges were able to pinpoint the break to a specific mountain pass or fault zone (Chatzimanolis and Caterino 2007). The Transverse Ranges are also an endemicity hotspot for small mammals (Davis et al. 2008) and have a high occurrence of irreplaceable lineages of amphibians and reptiles compared to most of California (Rissler et al. 2006). Chatzimanolis and Caterino (2007, p. 2128) noted that "few phylogeographic features in the California Floristic Province have garnered the attention of the 'Transverse Range Break'… there has been little attempt to synthesize a more general understanding of this significant phylogeographic break." Despite its possible role in the formation of the

Transverse Ranges (Nicholson et al. 1994, Hey 1998), the Murray FZ has not been implicated in any of the references cited in Table 1.2.

V. THE MOLOKAI AND SHIRLEY FRACTURE ZONES, PUNTA EUGENIA, AND THE VIZCAÍNO DESERT

(1) Geology and geography

The Molokai FZ combined with its eastern segment, the Shirley FZ (Figures 1.1 and 1.3), extends ~4200 km from Molokai, Hawaii (21°08′ N, 156°42′ W) to Punta Eugenia (27°50′ N, 115°04′ W) in Baja California Sur, México. The Shirley FZ directly corresponds with Punta Eugenia at the tip of the Vizcaíno Peninsula, the most prominent landform on the Pacific coast of the Baja California Peninsula. Inland of this point, a transverse belt of active volcanoes stretches across the peninsula, including the Sierra Vizcaíno, Sierra Santa Clara, Sierra Guadalupe, and Sierra San Francisco. Punta Eugenia delimits the southern end of the Continental Borderland and corresponds with a dramatic change in undersea topography in the Gulf of California, which becomes much deeper south of this latitude (Smith and Menard 1965). During most of the Miocene, well before the origin of the Gulf of California ~6 Mya, the west coast of Mexico (what is now Baja California) was bounded on the west by a subduction zone. About 20 Mya, when the RTJ was positioned at the eastern end of the Murray FZ, the South Farallon plate splintered along the Shirley FZ into the Guadalupe microplate and the Cocos plate (Figure 1.5). Later, the Cocos plate fragmented

again ~14 Mya, so that the Shirley FZ separated the Guadalupe and Magdalena microplates, as well as the Guadalupe ridge and Magdalena rise (Lonsdale 1991, Stock and Lee 1994). By ~12 Mya, seafloor spreading and subduction had largely ceased off the west coast of Baja California, leading to the death of these microplates, which were eventually transferred to the Pacific plate along with the Baja California Peninsula when rifting commenced ~6 Mya (Stock and Lee 1994). Nevertheless, the different tectonic histories on either side of the Shirley FZ are evident from examination of magnetic anomaly offsets, failed rifts, and the orientation of the coastline (Atwater 1989).

(2) Marine zoogeography

Like Cape Mendocino and Point Conception, Punta Eugenia diverts the cold California Current offshore, allowing the formation of a warmer eddy to the lee of the point, marking the southern limit of the temperate Californian ecosystem (Valentine 1966). Cooper (1895) was the first to propose the existence of a mid-peninsular seaway across central Baja California that would explain the disjunct temperate species present in the northern Gulf of California and Pacific Ocean north of Punta Eugenia, but absent from the tropical conditions of Baja California Sur. Smith (1919), Hall (1964), Valentine (1966) and Briggs (1974) defined Punta Eugenia (~28^{o} N) as the southern limit of the Californian province, while Steinbeck and Ricketts (1941) described this area as an overlap zone between the California and Panamic provinces. Community assemblages to the

north and south of Punta Eugenia differ for fishes, mollusks, and other marine taxa (Garth 1960, Soulé 1960, Hall 1964, Valentine 1966). Several pelagic fish populations south of Punta Eugenia are morphologically distinct from those to the north (e.g. Vrooman 1964). Phylogeographic studies have found genetic breaks in a variety of species that span Punta Eugenia or the Gulf of California at the same latitude and a few species demonstrate deep mid-peninsular phylogeographic breaks on both coasts (Table 1.3). Riginos (2005, p. 2678) concluded that the "simplest explanation for this concordant genetic division within both terrestrial and marine vertebrates is that the Baja Peninsula was fragmented by a Plio-Pleistocene marine seaway and that this seaway posed a substantial barrier to movement for near-shore fishes." However, direct geological evidence for a marine incursion connecting the Pacific Ocean to the Gulf of California is currently lacking, making the mid-peninsular seaway a controversial proximate explanation for the zoogeographic divide observed here (Ledesma-Vázquez and Carreño 2010).

(3) Terrestrial zoogeography

Molecular phylogeographic studies have found evidence for a deep mid-peninsular break in a variety of terrestrial species, including reptiles, birds, rodents, and spiders (Table 1.3). Several authors have attempted to synthesize a deeper understanding of the mid-peninsular break using a comparative approach (Lindell et al. 2006). Riddle et al. (2000) reviewed 12 phylogeographic studies of mammals, reptiles and birds, and found that 10 demonstrated a mid-

peninsular break. Zink (2002) examined 16 lineages of birds, reptiles, and mammals in Baja California, and found that 13 "show reciprocally monophyletic groups that meet today at 28–30° N" (p. 955). Lindell et al. (2006, p. 1329) argued, "it is possible that forces other than an ancient seaway have significantly affected individual genealogies among those harbouring deep genetic splits in the mid-peninsular region; however, the alternative explanations have no support and are countered by strong evidence. Consequently, a vicariance event affecting all species similarly remains the most parsimonious explanation for the congruent genealogical breaks around the mid-peninsular region." Leaché et al. (2007) analyzed mtDNA from 12 species of mammals and reptiles in a coalescent model, and found that their results "do not support the perception of a shared vicariant history among lineages exhibiting a north–south divergence at the mid-peninsula, and instead support two events differentially structuring genetic diversity in this region" (p. 646). Instead, the authors argued (p. 649) "if a geological explanation is sought for the mid-peninsular divergences, then it is possible that the combined forces of volcanism and marine incursions created the physical barriers necessary for population subdivision. Alternatively, a strong ecological and climatic gradient occurs at the mid-peninsula and limits the distribution of some species." There is no mention of the Molokai or Shirley FZs in any of the zoological papers cited in this section.

VI. THE CLARION FRACTURE ZONE, CABO CORRIENTES, AND THE SIERRA TRANSVOLCANICA

(1) Geology and geography

The Clarion FZ (Figures 1.1 and 1.3) extends over 5,600 km across the tropical North Pacific Ocean from the vicinity of Tethys Seamount (12º12′ N, 162º12′ W) to the western limit of the Rivera microplate ~380 km west (265º) of Cabo Corrientes (20º24′ N, 105º41′ W) in Jalisco, México. Cabo Corrientes marks the RTJ, the southern counterpart to the MTJ where the Pacific, North American, and Rivera plates meet. To the south of this junction, the EPR has not yet made contact with the western margin of North America, thus the Cocos plate (the southern sister of the Juan de Fuca plate) is currently being subducted beneath the Guatemalan Trench, which extends to Panama (Atwater 1989). Cabo Corrientes delimits the southern extent of the Gulf of California (Ledesma-Vázquez and Carreño 2010) where the cape of Baja California was attached before subduction of the EPR to the north resulted in the formation of the peninsula via rifting (Moore and Buffington 1968, Atwater 1989, Lonsdale 1989). The Sierra Transvolcanica, following the trend of the Clarion FZ east of Cabo Corrientes, defines the southern extent of the Central Mexican Plateau, Sierra Madre Occidental, and Sierra Madre Oriental, as well as the northern limit of the Mesoamerican subduction zone (Stock and Lee 1994). Although the modern RTJ is kinematically equivalent to a trench-trench-ridge triple junction, it has two limbs that are trenches and one limb that is an oblique extensional zone (Stock

and Lee 1994). This configuration is a recent development. The southern end of the SAF system, corresponding to the northern end of the Mesoamerican subduction zone, has occupied approximately the same position relative to the North American plate since 10 Mya because subduction of both plates has been nearly orthogonal to the trenches, but about 5 Mya the independently moving portion of the Cocos plate north of the Rivera transform fault broke off to become the modern Rivera plate (Stock and Lee 1994). The inferred age of origin of the Rivera microplate matches the age of the oldest recognized volcanism (5 Mya) within the Colima graben in the western Sierra Transvolcanica and initiation of oblique rifting in the Gulf of California, implying a direct relationship between the Rivera triple junction and continental deformation (Stock and Lee 1994).

(2) Marine zoogeography

In an oceanographic sense, Cabo Corrientes may be regarded as the southeastern boundary of the Gulf of California where the cape of Baja California formerly attached to the mainland (Ledesma-Vázquez and Carreño 2010). For example, 82 fish species (10% of total) are endemic to the Gulf of California as defined with Cabo Corrientes as its southern limit (Hastings et al. 2010). Although I only found two papers demonstrating intraspecific morphological or genetic breaks in marine animals across this barrier (Table 1.4), I suspect that this is due to a lack of phylogeographic research. I predict that as marine biologists further scrutinize this region, future studies will reveal more examples of phylogeographic breaks across Cabo Corrientes.

(3) Terrestrial zoogeography

Darwin (1839, p. 152) observed that "if we divide America, not by the Isthmus of Panama, but by the southern part of México in lat. 20º, where the great table-land presents an obstacle to the migration of species… we shall then have the two zoological provinces of North and South America strongly contrasted with each other." This "great table-land" is currently known as the Central Mexican Plateau, which is bounded on the south by the Sierra Transvolcanica at ~20º N. Moore (1945) analysed ornithological distributions in this region, and observed (p. 219) "hundreds of mighty cones, many extinct, massed and concentrated transversely across an enormous volcanic belt, four hundred miles long and sixty wide." Regarding the fauna, he noted (p. 219) "zoological life within them has been profoundly and continuously modified by the ever changing conditions for more than a million years." It later was revealed that the Sierra Transvolcanica did not act so much as a sharp boundary between the Nearctic and Neotropical flora and fauna, but more as the centre of a broader filter barrier, termed the Mexican Transition Zone (MTZ) (Ortega and Arita 1998, Morrone and Márquez 2001, Escalante et al. 2004). Marshall and Liebherr (2000) used a cladistic analysis to study biogeographic relationships among nine montane areas of endemism along the MTZ based on 33 clades of insects, fish, reptiles and plants. They found that "the Sierra Transvolcanica is home to an incredible array of endemic species" (p. 206) and that "there are two biogeographic assemblages in Mexico and Central America: one north of the Sierra Transvolcanica and one south… Perhaps novel is that the northern biotas

are limited to the regions north of the Sierra Transvolcanica, not the Isthmus of Tehuantepec" (p. 211). Phylogeographic breaks in a variety of aquatic and terrestrial animals including freshwater fish, amphibians, reptiles, and mammals are well documented across the MTZ and the Sierra Transvolcanica; in most of these cases, vicariance was implicated as a primary force of divergence, but dispersal-based hypotheses have also been proposed (Table 1.4).

VII. SYNTHESIS: BIOLOGICAL EVOLUTION RECAPITULATES TECTONIC EVOLUTION

The goal of this article is to bridge the disparate literatures of plate tectonics and biogeography in an attempt to link pattern with process, emphasizing how biological evolution recapitulates tectonic evolution. The pattern of interest is the alignment of GPFZs with major zoogeographic (phylogeographic) boundaries that correspond with anomalous transverse structures on the continental margin. Relying on geological literature largely uncited by most ecologists and evolutionary biologists, I have shown here that the evolution of the SAF system since the late Oligocene (triple junction migration, microplate formation, rotation, and accretion, propagating rifts, and strike-slip faulting) accounts for the ultimate origin of western North America's topography. The positions and orientations of prominent capes, points, islands, faults, volcanoes, mountain ranges, ridges, basins, seaways, and rivers have been constrained by the pre-existing tectonic boundaries set by the GPFZs, which in

turn are products of transform faults in the ancestral EPR long before the formation of the SAF system, as early as the Mesozoic. Furthermore, many studies have shown that these topographic features are manifested as important phylogeographic boundaries, revealed through genetic analyses of a phylogenetically broad sample of marine and terrestrial animals (Tables 1.1–1.4). I hope to convince the reader that the ultimate process at play is directly associated with triple junction evolution during the formation of the SAF system. A three-stage hypothetical model for the tectonic evolution of these zoogeographic boundaries based on the works reviewed earlier is summarized as follows:

(a) Stage 1 (discontinuity in cordillera): at this stage, the EPR had yet to contact the subduction zone (Figure 1.4A), but the eastern part of the FZ (in the Farallon plate or its descendent fragments) may have caused a minor constriction in the volcanic cordillera as it approached the subduction zone, possibly creating a filter barrier.Modern examples of Stage 1 may be observed currently in the Cocos plate off southern México, where the Tehuantepec FZ in the Cocos plate corresponds with the Isthmus of Tehuantepec that separates the Sierra Madre del Sur from the Guatemalan Highlands on the North American plate (Barrier et al. 1998), and the Orozco FZ which corresponds with the Colima graben in the Sierra Transvolcanica (Stock and Lee 1994). Stage 1 ended when the ridge met the subduction zone on one side of the FZ, creating a triple junction.

(b) Stage 2 (triple junction): after initial contact of the EPR and Farallon Trench, the modern Mendocino and Rivera triple junctions were formed (Figure

1.4B, McKenzie and Morgan 1969). The EPR met the subduction zone on only one side of the GPFZ, forming a transtensional transform fault, while on the other side of the FZ, the EPR and subduction zone had yet to meet, forming a triple junction (Figure 1.4C). Stage 2 ended when the EPR met the subduction zone on the other side of the FZ, and the triple junction migrated onwards.

(c) Stage 3 (fossilized triple junction): the reconfiguration from a spreading/subducting scheme to a transform boundary has been completed on both sides of the FZ, leaving an aseismic GPFZ that corresponds with a fossilized triple junction on the mainland (Figure 1.4D). Subsequent plate movement along strike-slip or spreading centres may have deformed, obscured, or destroyed any direct structural connection between the oceanic fracture zone and its continental counterpart via microplate rotation, transtention, and transpression, e.g. the Murray FZ, Point Conception and the Transverse Ranges (Nicholson et al. 1994).

The triple junction model may account for the ultimate origin of western North America's coastal topography since the late Oligocene, but what proximate mechanisms drive ecological divergence and speciation across these boundaries? The answer is complex and incomplete, as the proximate mechanisms are well known for some taxa (Kelly and Palumbi 2010), but poorly understood in others (Leaché et al. 2007). For most organisms, we can divide these boundaries into two possible categories: physical barriers and ecological barriers. In some cases, physical topography and oceanography may directly inhibit dispersal or promote vicariance. This has been demonstrated for sedentary species with narrow physiological limits that may only disperse tens of

metres in their lifetimes, such as plethodontid salamanders that show sharp, deep genetic breaks at plate boundaries and appear to simply ride the plates as they migrate over the eons (Wake 2006, Reilly and Wake 2015). Even the planktonic larvae of many marine species may be physically unable to overcome strong prevailing offshore currents associated with prominent headlands, especially during periods of heavy upwelling (Kelly and Palumbi 2010). Alternatively, for species capable of dispersal across these barriers, ecotones associated with prominent topographic features may still limit their distributions, as C. A. Hall (2002) has shown to be the case for nearshore molluscan communities since the late Miocene, or they may constrain hybrid zones due to strong selection overpowering ongoing gene flow, as in the case of *Ensatina* salamanders (Wake 2006). Oceanographic conditions vary on either side of a headland, different rock types with adapted plant communities often meet at fault lines, and rugged topography generates climatic ecotones of temperature and precipitation associated with elevation and slope, including the rain-shadow effect. These sharp gradients are likely to affect the establishment of species with narrow ecological niches in diverse ways, even if they are capable of dispersing across the physical barriers. Finally, wide-ranging species with broad niches and excellent dispersal abilities may not be affected by either physical or ecological boundaries (Bagley and Johnson, 2014). We should not expect that Cape Mendocino acts as a major barrier to most large cetaceans or raptors, thus lack of genetic divergence across this barrier for mobile, generalist species should not be taken as counter-evidence disputing the overall importance of these barriers to

more sedentary or specialized taxa. Although biologists will likely debate these details for decades to come, I believe most would not doubt that constant tectonic disturbance along the SAF system over millions of years has uniquely impacted the evolution and distribution of endemic species.

VIII. FUTURE DIRECTIONS

Historical biogeography typically advances via a three-step process: (*i*) the discovery and description of geospatial patterns of biodiversity, (*ii*) the development of mechanistic hypotheses to explain them, and (*iii*) rigorous testing of these hypotheses until the ones that are necessary and sufficient to account for the patterns are adopted (Lomolino et al. 2006). Here, I addressed the first two steps by describing a pattern (the alignment of the GPFZs with major zoogeographic boundaries) and proposing a tectonic process to explain it (triple junction evolution and the extension of the SAF system). The next step is to test this hypothesis in a quantitative framework. To tackle this challenge, collaboration among many biologists (marine and terrestrial) and geologists will be essential. Comparative phylogeography has much to offer. When multiple unrelated lineages show spatial and/or temporally congruent phylogenetic breaks across known tectonic barriers, this constitutes powerful evidence that a common underlying historical event affected all taxa similarly. The principle advantage of the comparative approach is that although every species has a unique evolutionary history influenced by stochastic events, making single-

species assessments unreliable, by studying larger numbers of co-distributed taxa we have more statistical power to identify the idiosyncrasies that mask patterns shared across taxa. Given the wide range in possible proximate mechanisms of divergence, we need not expect temporal concordance in divergence among all unrelated taxa across any given boundary, although we should expect to find at least some spatially and temporally congruent phylogeographic breaks for co-distributed species with similar life histories, which may yield important new insights into the mechanisms of tectonic and biological evolution.

Although the comparative phylogeographic approach presents a number of challenges, recent technological and computational advances are steadily reducing the obstacles. Following Bagley and Johnson (2014), who recently reviewed the phylogeography of lower Central America (another tectonically active biodiversity hotspot), I argue that phylogeographers should focus on four areas of advancement. First, better data (especially multiple unlinked markers) are needed to estimate demographic parameters with a high degree of statistical confidence. Although most of the molecular papers reviewed here have relied upon few genetic markers, most commonly mtDNA, this trend is rapidly changing as next-generation-sequencing strategies allow for the rapid, cost-effective collection of population genomic data. Two promising methods, restriction-associated DNA sequencing (Peterson et al. 2012) and target enrichment (Faircloth et al. 2012), now allow researchers to obtain population genomic-scale data sets for non-model organisms at a reasonable cost. Second,

phylogeographers should take advantage of the latest statistical methods that account for coalescent stochasticity (Knowles 2009, Bryant et al. 2012). These methods are needed to estimate accurately population divergence times, effective population sizes, migration rates, population growth, etc. Third, phylogeographers should attempt a priori to adopt geographically and taxonomically expanded sampling strategies. Many studies have been limited by geographic sampling, leaving readers to guess whether a specific fault or landscape feature is at play. Taxonomic breadth is also important, as animals with dramatically different life histories may interact with the landscape in different ways. Fourth, I urge authors to adopt hypothesis-driven approaches whenever possible (Knowles 2009). For example, Bayes factors can be used explicitly to test alternative species delimitation models (Grummer et al. 2014), and these species boundaries may correspond with tectonic features or historical events, such as those presented in Figure 1.5. Phylogenetic analyses of these data using relaxed molecular clocks could be used to test whether co-distributed taxa exhibit spatial and temporal concordance in their phylogeographic structure. If biogeographers meet these four criteria, they can rigorously test (and potentially reject) the GPFZ hypothesis on a case-by-case basis until a general consensus emerges. I predict that as with mitochondrial studies, large-scale comparative phylogeographic studies will find that models incorporating multiple temporal waves of diversification across these barriers (Leaché et al. 2007) are most consistent with the hypothesis presented here, but this remains to be seen.

IX. CONCLUSIONS

Four GPFZs in the Pacific plate correspond with zoogeographic and phylogeographic boundaries along the SAF system in western North America. The Mendocino FZ corresponds with the Mendocino triple junction, Cape Mendocino and the North Coast Divide, the Murray FZ corresponds with Point Conception and the Transverse Ranges, the Molokai/Shirley FZ corresponds with Punta Eugenia and the Vizcaíno Desert, and the Clarion FZ corresponds with the Rivera triple junction, Cabo Corrientes and the Sierra Transvolcanica. These four boundaries act as distributional limits or phylogeographic breaks between sister lineages for a wide variety of marine and terrestrial taxa, mostly those with limited dispersal abilities, including invertebrates, fish, amphibians, reptiles, birds, and mammals. The GPFZs are the sole remaining trace of ancient transform faults in the now-destroyed EPR before it met the Farallon Trench bordering North America in the late Oligocene. The GPFZs also constrained the positions and orientations of microplates and triple junctions during the evolution of the SAF system, ultimately resulting in zoogeographic boundaries that reflect tectonic boundaries. Thus, at least for some taxa, zoological evolution appears to have recapitulated tectonic evolution along the west coast of North America. A variety of proximate mechanisms may drive ecological divergence and speciation across these boundaries, although more research is needed in this area. These boundaries may act directly as barriers to dispersal, or indirectly by establishing ecological or climatic gradients. Species with poor dispersal abilities

or specialized ecological niches are more likely to be impacted by these barriers than are mobile, generalist species, which may show little to no phylogeographic structure, yet lack of divergence in these species should not be taken as counter-evidence against the importance of these tectonic boundaries for more specialized or sessile species. Despite being readily visible with internet software (e.g. Google Maps), discussion of the GPFZs is mostly absent from the zoogeography literature. Comparative phylogeographic analysis of co-distributed lineages with similar life-history traits might be the most promising approach for testing the hypotheses presented here. I urge more collaboration among marine and terrestrial phylogeographers with geologists to elucidate better the relationships between GPFZs and the biogeography of western North America.

X. REFERENCES

*References marked with asterisk have been cited within Appendix A only.

Aguirre LG, Morafka DJ, Murphy RW. 1999. The peninsular archipelago of Baja California: a thousand kilometers of tree lizard genetics. Herpetologica, 55:369–381.

Amman BR, Bradley RD. 2004. Molecular evolution in *Baiomys* (Rodentia: Sigmodontinae): evidence for a genetic subdivision in *B. musculus*. Journal of Mammalogy, 85: 162–166.

Arévalo E, Casas G, Davis SK, Lara G, Sites JW. 1993. Parapatric hybridization between chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae): structure of the Ajusco transect. Copeia, 1993:352–372.

Atwater T. 1970. Implications of plate tectonics for the Cenozoic tectonic evolution of western North America. Geological Society of America Bulletin, 81:3513–3536.

Atwater T. 1989. Plate tectonic history of the northeast Pacific and western North America. In: Winterer EL, Hussong DM, Decker RW editors. The Geology of North America, Volume N: the Eastern Pacific Ocean and Hawaii, The Geological Society of America, Boulder, p. 21–72.

Atwater T, Menard HW. 1970. Magnetic lineations in the northeast Pacific. Earth and Planetary Science Letters, 7:445–450.

Bagley JC, Johnson JB. 2014. Phylogeography and biogeography of the lower Central American Neotropics: diversification between two continents and between two seas. Biological Reviews, 89:767–790.

Barbour CD. 1973. A biogeographical history of *Chirostoma* (Pisces: Atherinidae): a species flock from the Mexican Plateau. Copeia, 1973:533–566.

Barrier E, Velasquillo L, Chavez M, Gaulon R. 1998. Neotectonic evolution of the Isthmus of Tehuantepec (southeastern México). Tectonophysics, 287:77–96.

Barrowclough GF, Gutierrez RJ, Groth JG. 1999. Phylogeography of spotted owl (*Strix occidentalis*) populations based on mitochondrial DNA sequences: gene

flow, genetic structure, and a novel biogeographic pattern. Evolution, 53:919–931.

Bernardi G. 2000. Barriers to gene flow in *Embiotoca jacksoni*, a marine fish lacking a pelagic larval stage. Evolution, 54:226–237.

Bernardi G, Findley L, Rocha–Olivares A. 2003. Vicariance and dispersal across Baja California in disjunct marine fish populations. Evolution, 57:1599–1609.

Bernardi G, Talley D. 2000. Genetic evidence for limited dispersal in the coastal California killifish, *Fundulus parvipinnis*. Journal of Experimental Marine Biology and Ecology, 255:187–199.

Blanchette CA, Miner CM, Raimondi PT, Lohse D, Heady KEK, Broitman BR. 2008. Biogeographical patterns of rocky intertidal communities along the Pacific coast of North America. Journal of Biogeography, 35:1593–1607.

Briggs JC. 1974. Marine Zoogeography. McGraw-Hill, New York.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Molecular Biology and Evolution, 29:1917–1932.

Bryson RW, García-Vázquez UO, Riddle BR. 2011a. Phylogeography of Middle American gophersnakes: mixed responses to biogeographical barriers across the Mexican Transition Zone. Journal of Biogeography, 38:1570–1584.

Bryson RW, Murphy RW, Lathrop A, Lazcano-Villareal D. 2011b. Evolutionary drivers of phylogeographical diversity in the highlands of Mexico: a case study of the *Crotalus triseriatus* species group of montane rattlesnakes. Journal of Biogeography, 38:697–710.

Bryson RW, Nieto-Montes de Oca A, Velasco JR. 2008. Phylogenetic position of *Porthidium hespere* (Viperidae: Crotalinae) and phylogeography of arid-adapted hognosed pitvipers based on mitochondrial DNA. Copeia, 2008:172–178.

Burford MO. 2009. Demographic history, geographical distribution and reproductive isolation of distinct lineages of blue rockfish (*Sebastes mystinus*),

a marine fish with a high dispersal potential. Journal of Evolutionary Biology, 22:1471–1486.

Burford MO, Bernardi G. 2008. Incipient speciation within a subgenus of rockfish (*Sebastosomus*) provides evidence of recent radiations within an ancient species flock. Marine Biology, 154:701–717.

Burton RS. 1998. Intraspecific phylogeography across the Point Conception biogeographic boundary. Evolution, 52:734–745.

Burton RS, Lee BN. 1994. Nuclear and mitochondrial gene genealogies and allozyme polymorphism across a major phylogeographic break in the copepod *Tigriopus californicus*. Proceedings of the National Academy of Sciences of the United States of America, 91:5197–5201.

Calsbeek R, Thompson JN, Richardson JE. 2003. Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. Molecular Ecology, 12:1021–1029.

Cassone BJ, Boulding EG. 2006. Genetic structure and phylogeography of the lined shore crab, *Pachygrapsus crassipes*, along the northeastern and western Pacific coasts. Marine Biology, 149:213–226.

Chatzimanolis S, Caterino MS. 2007. Toward a better understanding of the "Transverse Range break": lineage diversification in southern California. Evolution, 61:2127–2141.

Conroy CJ, Neuwald JL. 2008. Phylogeographic study of the California vole, *Microtus californicus*. Journal of Mammalogy, 89:755–767.

Cooper JG. 1895. Catalogue of marine shells, collected chiefly on the eastern shore of Lower California for the California Academy of Sciences during 1891–1892. Proceedings of the California Academy of Sciences, 5:34–38.

Cope JM. 2004. Population genetics and phylogeography of the blue rockfish (*Sebastes mystinus*) from Washington to California. Canadian Journal of Fisheries and Aquatic Sciences, 61:332–342.

Corbin KW. 1981. Genic heterozygosity in the White-Crowned Sparrow: a potential index to boundaries between subspecies. Auk, 98:669–680.

Crews SC, Hedin M. 2006. Studies of morphological and molecular phylogenetic divergence in spiders (Araneae: *Homalonychus*) from the American southwest, including divergence along the Baja California Peninsula. Molecular Phylogenetics and Evolution, 38:470–487.

Dall WH. 1899. The mollusk fauna of the Pribilof Islands. In: Jordan DS editor. The Fur Seals and Fur-seal Islands of the North Pacific Ocean, U.S. Government Printing Office, Washington, DC, p. 539–546.

Darda DM. 1994. Allozyme variation and morphological evolution among Mexican salamanders of the genus *Chiropterotriton* (Caudata: Plethodontidae). Herpetologica, 50:164–187.

Darwin C. 1839. Journal of Researches into the Natural History and Geology of the Countries Visited during the Voyage of H.M.S. Beagle. Henry Colburn, London.

Davis BJ, Demartini EE, McGee K. 1981. Gene flow among populations of a teleost (Painted Greenling, *Oxylebius pictus*) from Puget Sound to southern California. Marine Biology, 65:17–23.

Davis EB, Koo MS, Conroy C, Patton JL, Moritz C. 2008. The California hotspots project: identifying regions of rapid diversification of mammals. Molecular Ecology, 17:120–138.

Dawson MN. 2001. Phylogeography in coastal marine animals: a solution from California? Journal of Biogeography, 28:723–736.

Daza JM, Smith EN, Paez VP, Parkinson CL. 2009. Complex evolution in the Neotropics: the origin and diversification of the widespread genus *Leptodeira* (Serpentes: Colubridae). Molecular Phylogenetics and Evolution, 53:653–657.

Devitt TJ. 2006. Phylogeography of the Western Lyresnake (*Trimorphodon biscutatus*): testing aridland biogeographical hypotheses across the Nearctic–Neotropical transition. Molecular Ecology, 15:4387–4407.

Dickinson WR. 1997. Tectonic implications of Cenozoic volcanism in coastal California. Geological Society of America Bulletin, 109:936–954.

Dietz RS. 1961. Continent and ocean basin evolution by spreading of sea floor. Nature, 190:854–857.

Eberl R, Mateos M, Grosberg RK, Santamaria CA, Hurtado LA. 2013. Phylogeography of the supralittoral isopod *Ligia occidentalis* around the Point Conception marine biogeographical boundary. Journal of Biogeography, 40: 2361–2372.

Escalante T, Rodriguez G, Morrone JJ. 2004. The diversification of Nearctic mammals in the Mexican transition zone. Biological Journal of the Linnean Society, 83:327–339.

Evans BJ, Supriatna J, Andayani N, Setiadi MI, Cannatella DC, Melnick DJ. 2003. Monkeys and toads define areas of endemism on Sulawesi. Evolution, 57:1436–1443.

Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology, 61:717–726.

Feldman CR, Spicer GS. 2006. Comparative phylogeography of woodland reptiles in California: repeated patterns of cladogenesis and population expansion. Molecular Ecology, 15:2201–2222.

Forister ML, Fordyce JA, Shapiro AM. 2004. Geological barriers and restricted gene flow in the holarctic skipper *Hesperia comma* (Hesperiidae). Molecular Ecology, 13:3489–3499.

Furlong KP, Schwartz SY. 2004. Influence of the Mendocino triple junction on the tectonics of coastal California. Annual Review of Earth and Planetary Sciences, 32:403–433.

Garth JS. 1960. Distribution and affinities of the Brachyuran Crustacea. Systematic Zoology. 9:105–123.

Gottscho AD, Marks S, Jennings WB. 2014. Speciation, population structure, and demographic history of the Mojave Fringe-toed Lizard (*Uma scoparia*), a species of conservation concern. Ecology and Evolution, 4:2546–2562.

Grismer L. 2002. Amphibians and Reptiles of Baja California, Including its Pacific Islands and the Islands in the Sea of Cortés. University of California Press, Berkeley.

Grummer JA, Bryson RW, Reeder TW. 2014. Species delimitation using Bayes factors: simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). Systematic Biology, 63:119–133.

Hall CA Jr. 1964. Shallow-water marine climates and molluscan provinces. Ecology, 45:226–234.

Hall CA Jr. 2002. Nearshore marine paleoclimatic regions, increasing zoogeographic provinciality, molluscan extinctions, and paleoshorelines, California: late Oligocene (27 Ma) to late Pliocene (2.5 Ma). Geological Society of America Special Paper, 357:1–487.

Hall R. 2002. Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer–based reconstructions, model and animations. Journal of Asian Earth Sciences, 20:353–431.

Hastings PA, Findley LT, Van der Heiden AM. 2010. Fishes of the Gulf of California. In: Brusca RC editor. The Gulf of California: Biodiversity and Conservation, University of Arizona Press, Tucson, p. 96–118.

Haupt AJ. 2011. Historical and oceanographic influences on phylogeography in the California current ecosystem and application to management of marine species. PhD Dissertation, Department of Biology, Stanford University.

Hayward TL, Mantyla AW. 1990. Physical, chemical and biological structure of a coastal eddy near Cape Mendocino. Journal of Marine Research, 48:825–850.

Hedin M, Starrett J, Hayashi C. 2013. Crossing the uncrossable: novel trans-valley biogeographic patterns revealed in the genetic history of low-dispersal mygalomorph spiders (Antrodiaetidae, *Antrodiaetus*) from California. Molecular Ecology, 22:508–526.

Hey RN. 1998. Speculative propagating rift subduction zone interactions with possible consequences for continental margin evolution. Geology, 26:247–250.

Hey RN, Sinton JM, Duennebier FK. 1989. Propagating rifts and spreading centers. In: Winterer EL, Hussong DM, Decker RW editors. The Geology of North America, Volume N: the Eastern Pacific Ocean and Hawaii, The Geological Society of America, Boulder, p. 161–176.

Hollingsworth BD. 1998. The systematics of chuckwallas (*Sauromalus*) with a phylogenetic analysis of other iguanid lizards. Herpetological Monographs, 12:38–191.

Horn MH, Allen LG. 1978. Distributional analysis of California coastal marine fishes. Journal of Biogeography, 5:23–42.

Huang D, Bernardi G. 2001. Disjunct Sea of Cortez–Pacific Ocean *Gillichthys mirabilis* populations and the evolutionary origin of their Sea of Cortez endemic relative, *Gillichthys seta*. Marine Biology, 138:421–428.

Hulsey CD, de Leon FJG, Johnson YS, Hendrickson DA, Near TJ. 2004. Temporal diversification of Mesoamerican cichlid fishes across a major biogeographic boundary. Molecular Phylogenetics and Evolution, 31:754–764.

Hurtado LA, Mateos M, Santamaria CA. 2010. Phylogeography of supralittoral rocky intertidal *Ligia* isopods in the Pacific region from central California to central México. PLoS One, 5:e11633.

Jockusch EL, Yanev KP, Wake DB. 2001. Molecular phylogenetic analysis of slender salamanders, genus *Batrachoseps* (Amphibia: Plethodontidae), from central coastal California with descriptions of four new species. Herpetological Monographs, 15:54–99.

Kelly RP, Palumbi SR. 2010. Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. PLoS One, 5:e8594.

Knowles LL. 2009. Statistical phylogeography. Annual Review of Ecology, Evolution, and Systematics, 40:593–612.

Kuchta SR, Parks DS, Mueller RL, Wake DB. 2009. Closing the ring: historical biogeography of the salamander ring species *Ensatina eschscholtzii*. Journal of Biogeography, 36:982–995.

Kuchta SR, Tan AM. 2006. Lineage diversification on an evolving landscape: phylogeography of the California newt, *Taricha torosa* (Caudata: Salamandridae). Biological Journal of the Linnean Society, 89:213–239.

Law JH, Crespi BJ. 2002. The evolution of geographic parthenogenesis in *Timema* walking–sticks. Molecular Ecology, 11:1471–1489.

Leaché AD, Crews SC, Hickerson MJ. 2007. Two waves of diversification in mammals and reptiles of Baja California revealed by hierarchical Bayesian analysis. Biology Letters, 3:646–650.

Leaché AD, Koo MS, Spencer CL, Papenfuss TJ, Fisher RN, McGuire JA. 2009. Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). Proceedings of the National Academy of Sciences of the United States of America, 106: 12418–12423.

Leaché AD, Mulcahy DG. 2007. Phylogeny, divergence times and species limits of spiny lizards (*Sceloporus magister* species group) in western North American deserts and Baja California. Molecular Ecology, 16:5216–5233.

Leaché AD, Palacios JA, Minin VN, Bryson RW. 2013. Phylogeography of the Trans–Volcanic bunchgrass lizard (*Sceloporus bicanthalis*) across the highlands of south–eastern México. Biological Journal of the Linnean Society, 110:852–865.

Ledesma-Vázquez J, Carreño AL. 2010. Origin, age, and geological evolution of the Gulf of California. In: Brusca RC editor. The Gulf of California: Biodiversity and Conservation, University of Arizona Press, Tucson, p. 7–23.

Lindell J, Mendez–de la Cruz FR, Murphy RW. 2005. Deep genealogical history without population differentiation: discordance between mtDNA and allozyme divergence in the zebra–tailed lizard (*Callisaurus draconoides*). Molecular Phylogenetics and Evolution, 36:682–694.

Lindell J, Mendez–de la Cruz FR, Murphy RW. 2008. Deep biogeographical history and cytonuclear discordance in the black–tailed brush lizard (*Urosaurus nigricaudus*) of Baja California. Biological Journal of the Linnean Society, 94:89–104.

Lindell J, Ngo A, Murphy RW. 2006. Deep genealogies and the mid–peninsular seaway of Baja California. Journal of Biogeography, 33:1327–1331.

Lock J, Kelsey H, Furlong K, Woolace A. 2006. Late Neogene and quaternary landscape evolution of the northern California Coast Ranges: evidence for

Mendocino triple junction tectonics. Geological Society of America Bulletin, 118:1232–1246.

Lomolino MV, Riddle BR, Brown JH. 2006. Biogeography. Third Edition. Sinauer Associates, Inc., Sunderland.

Lonsdale P. 1989. Geology and tectonic history of the Gulf of California. In: Winterer EL, Hussong DM, Decker RW editors. The Geology of North America, Vol. N: the Eastern Pacific Ocean and Hawaii, The Geological Society of America, Boulder, p. 499–521.

Lonsdale P. 1991. Structural patterns of the Pacific floor offshore of Peninsular California. American Association of Petroleum Geologists Memoir, 47:87–125.

Luyendyk BP. 1991. A model for Neogene crustal rotations, transtension, and transpression in southern California. Geological Society of America Bulletin, 103:1528–1536.

Luyendyk BP, Kamerling MJ, Terres R. 1980. Geometric model for Neogene crustal rotations in southern California. Geological Society of America Bulletin, 91:211–217.

Lynch JF, Wake DB, Yang SY. 1983. Genic and morphological differentiation in Mexican *Pseudoeurycea* (Caudata: Plethodontidae), with a description of a new species. Copeia, 1983:884–894.

Maldonado JE, Vila C, Wayne RK. 2001. Tripartite genetic subdivisions in the ornate shrew (*Sorex ornatus*). Molecular Ecology, 10:127–147.

Marshall CJ, Liebherr JK. 2000. Cladistic biogeography of the Mexican transition zone. Journal of Biogeography, 27:203–216.

Mason RG, Raff AD. 1961. Magnetic survey off the west coast of North America, 32°N. latitude to 42°N. latitude. Geological Society of America Bulletin, 72: 1259–1265.

Mateos M. 2005. Comparative phylogeography of livebearing fishes in the genera *Poeciliopsis* and *Poecilia* (Poeciliidae: Cyprinodontiformes) in central México. Journal of Biogeography, 32:775–780.

Matocq MD. 2002. Phylogeographical structure and regional history of the dusky-footed woodrat, *Neotoma fuscipes*. Molecular Ecology, 11:229–242.

McGuire JA, Brown RM, Mumpuni AR, Andayani N. 2007a. The flying lizards of the *Draco lineatus* group (Squamata: Iguania: Agamidae): a taxonomic revision with descriptions of two new species. Herpetological Monographs, 21:179–212.

McGuire JA, Linkem CW, Koo MS, Hutchison DW, Lappin AK, Orange DI, Lemos-Espinal J, Riddle BR, Jaeger JR. 2007b. Mitochondrial introgression and incomplete lineage sorting through space and time: phylogenetics of crotaphytid lizards. Evolution, 61:2879–2897.

McHugh JL. 1951. Meristic variations and populations of northern anchovy. Bulletin of the Scripps Institution of Oceanography, 6:123–160.

McKenzie DP, Morgan WJ. 1969. Evolution of triple junctions. Nature, 224:125–133.

McKenzie DP, Parker RL. 1967. The North Pacific: an example of tectonics on a sphere. Nature, 216:1276–1280.

Menard HW. 1955. Deformation of the northeastern Pacific basin and the west coast of North America. Geological Society of America Bulletin, 66:1149–1199.

Menard HW. 1960. The East Pacific Rise. Science, 132:1737–1746.

*Menard HW. 1967. Extension of northeastern-Pacific fracture zones. Science, 155:72–74.

Menard HW, Atwater T. 1969. Origin of fracture zone topography. Nature, 222:1037–1040.

Menard HW, Dietz RS. 1952. Mendocino submarine escarpment. Journal of Geology, 60:266–278.

Montanucci RR. 2004. Geographic variation in *Phrynosoma coronatum* (Lacertilia, Phrynosomatidae): further evidence for a peninsular archipelago. Herpetologica, 60:117–139.

Moore RT. 1945. The transverse volcanic biotic province of central Mexico and its relationship to adjacent provinces. Transactions of the San Diego Society of Natural History, 10:217–236.

Moore DG, Buffington EC. 1968. Transform faulting and growth of the Gulf of California since the late Pliocene. Science, 161:1238–1241.

Morrone JJ, Márquez J. 2001. Halffter's Mexican transition zone, beetle generalized tracks, and geographical homology. Journal of Biogeography, 28:635–650.

Mulcahy DG, Macey JR. 2009. Vicariance and dispersal form a ring distribution in nightsnakes around the Gulf of California. Molecular Phylogenetics and Evolution, 53:537–546.

Mulcahy DG, Mendelson JR. 2000. Phylogeography and speciation of the morphologically variable, widespread species *Bufo valliceps*, based on molecular evidence from mtDNA. Molecular Phylogenetics and Evolution, 17:173–189.

Müller RD, Sdrolias M, Gaina C, Roest WR. 2008. Age, spreading rates, and spreading asymmetry of the world's ocean crust. Geochemistry, Geophysics, Geosystems, 9:Q04006.

Murphy RW. 1983. Paleobiogeography and genetic differentiation of the Baja California herpetofauna. Occasional Papers California Academy of Sciences, 137:1–48.

Murray HW. 1939. Submarine scarp off Mendocino, California. Field Engineers Bulletin, 13:27–33.

Myers EA, Rodriguez-Robles JA, Denardo DF, Staub RE, Stropoli A, Ruane S, Burbrink FT. 2013. Multilocus phylogeographic assessment of the California Mountain Kingsnake (*Lampropeltis zonata*) suggests alternative patterns of diversification for the California Floristic Province. Molecular Ecology, 22: 5418–5429.

Nicholson C, Sorlien CC, Atwater T, Crowell JC, Luyendyk BP. 1994. Microplate capture, rotation of the western Transverse Ranges, and initiation of the San Andreas transform as a low–angle fault system. Geology, 22:491–495.

Norris RM, Webb RW. 1990. Geology of California, Second Edition. Wiley, New York.

Ortega J, Arita HT. 1998. Neotropical–Nearctic limits in Middle America as determined by distributions of bats. Journal of Mammalogy, 79:772–783.

Parrish RH, Nelson CS, Bakun A. 1981. Transport mechanisms and reproductive success of fishes in the California current. Biological Oceanography, 1:175–203.

Pelé RA, Warner RR, Gaines SD. 2009. Geographical patterns of genetic structure in marine species with contrasting life histories. Journal of Biogeography, 36: 1881–1890.

Perryman WL, Westlake RL. 1998. A new geographic form of the spinner dolphin, *Stenella longirostris*, detected with aerial photogrammetry. Marine Mammal Science, 14:38–50.

Peterson DL, Kubow KB, Connolly MJ, Kaplan LR, Wetkowski MM, Leong W, Phillips BC, Edmands S. 2013. Reproductive and phylogenetic divergence of tidepool copepod populations across a narrow geographical boundary in Baja California. Journal of Biogeography, 40:1664–1675.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One, 7:e37135.

Petren K, Case TJ. 1997. A phylogenetic analysis of body size evolution and biogeography in chuckwallas (*Sauromalus*) and other iguanines. Evolution, 51:206–219.

Phillipsen IC, Metcalf AE. 2009. Phylogeography of a stream–dwelling frog (*Pseudacris cadaverina*) in southern California. Molecular Phylogenetics and Evolution, 53:152–170.

Pitman WC, Heirtzler JR. 1966. Magnetic anomalies over the Pacific–Antarctic Ridge. Science, 154:1164–1166.

Pook CE, Wuster W, Thorpe RS. 2000. Historical biogeography of the Western Rattlesnake (Serpentes: Viperidae: *Crotalus viridis*), inferred from mitochondrial DNA sequence information. Molecular Phylogenetics and Evolution, 15:269–282.

Radtkey RR, Fallon SM, Case TJ. 1997. Character displacement in some *Cnemidophorus* lizards revisited: a phylogenetic analysis. Proceedings of the National Academy of Sciences of the United States of America, 94:9740–9745.

Reilly SB, Marks SB, Jennings WB. 2012. Defining evolutionary boundaries across parapatric ecomorphs of Black Salamanders (*Aneides flavipunctatus*) with conservation implications. Molecular Ecology, 21:5745–5761.

Reilly SB, Wake DB. 2015. Cryptic diversity and biogeographic patterns within the black salamander (*Aneides flavipunctatus*) complex. Journal of Biogeography, 42:280–291.

Richmond JQ, Barr KR, Backlin AR, Vandergast AG, Fisher RN. 2013. Evolutionary dynamics of a rapidly receding southern range boundary in the threatened California Red-Legged Frog (*Rana draytonii*). Evolutionary Applications, 6:808–822.

Richmond JQ, Reeder TW. 2002. Evidence for parallel ecological speciation in scincid lizards of the *Eumeces skiltonianus* species group (Squamata: Scincidae). Evolution, 56:1498–1513.

Riddle BR, Hafner DJ. 2010. Integrating pattern with process at biogeographic boundaries: the legacy of Wallace. Ecography, 33:321–325.

Riddle BR, Hafner DJ, Alexander LF, Jaeger JR. 2000. Cryptic vicariance in the historical assembly of a Baja California peninsular desert biota. Proceedings of the National Academy of Sciences of the United States of America, 97:14438–14443.

Riginos C. 2005. Cryptic vicariance in Gulf of California fishes parallels vicariant patterns found in Baja California mammals and reptiles. Evolution, 59:2678–2690.

Rissler LJ, Hijman RJ, Graham CH, Moritz C, Wake DB. 2006. Phylogeographic lineages and species comparisons in conservation analyses: a case study of California herpetofauna. American Naturalist, 167:655–666.

Rodriguez-Robles JA, De Jesus-Escobar JM. 2000. Molecular systematics of new world gopher, bull, and pinesnakes (*Pituophis*: Colubridae), a transcontinental species complex. Molecular Phylogenetics and Evolution, 14:35–50.

Rodriguez-Robles JA, Denardo DF, Staub RE. 1999. Phylogeography of the California mountain kingsnake, *Lampropeltis zonata* (Colubridae). Molecular Ecology, 8:1923–1934.

Rodriguez-Robles JA, Stewart GR, Papenfuss TJ. 2001. Mitochondrial DNA-based phylogeography of North American rubber boas, *Charina bottae* (Serpentes: Boidae). Molecular Phylogenetics and Evolution, 18:227–237.

Ruane S, Bryson RW, Pyron AR, Burbrink FT. 2013. Coalescent species delimitation in milksnakes (genus *Lampropeltis*) and impacts on phylogenetic comparative analyses. Systematic Biology, 63:231–250.

Saavedra-Sotelo NC, Calderon-Aguilera LE, Reyes-Bonilla H, López-Pérez RA, Medina-Rosas P, Rocha-Olivares A. 2011. Limited genetic connectivity of *Pavona gigantea* in the Mexican Pacific. Coral Reefs, 30:677–686.

Sanders JG, Palumbi SR. 2011. Populations of *Symbiodinium muscatinei* show strong biogeographic structuring in the intertidal anemone *Anthopleura elegantissima*. Biological Bulletin, 220:199–208.

Satler JD, Starrett J, Hayashi CY, Hedin M. 2011. Inferring species trees from gene trees in a radiation of California trapdoor spiders (Araneae, Antrodiaetidae, *Aliatypus*). PLoS One, 6:e25355.

Savage JM. 1960. Evolution of a peninsular herpetofauna. Systematic Zoology, 9:184–212.

Sessions SK, Kezer J. 1987. Cytogenetic evolution in the plethodontid salamander genus *Aneides*. Chromosoma, 95:17–30.

Sgariglia EA, Burns KJ. 2003. Phylogeography of the California Thrasher (*Toxostoma redivivum*) based on nested-clade analysis of mitochondrial–DNA variation. Auk, 120:346–361.

Shepard FP, Emery KO. 1941. Submarine topography off the California coast. Geological Society of America Special Paper, 31:1–171.

Shubin N. 2009. Your Inner Fish: A Journey into the 3.5-Billion-Year History of the Human Body. Vintage Books, Random House Inc., New York.

Sivasundar A, Palumbi SR. 2010. Life history, ecology and the biogeography of strong genetic breaks among 15 species of Pacific rockfish, *Sebastes*. Marine Biology, 157:1433–1452.

Smith JP. 1919. Climatic relations of the tertiary and quaternary faunas of the California region. Proceedings of the California Academy of Sciences, 9:123–173.

Smith MF. 1979. Geographic variation in genic and morphological characters in *Peromyscus californicus*. Journal of Mammalogy, 60:705–722.

Smith SM, Menard HW. 1965. The Molokai Fracture Zone. Progress in Oceanography, 3:333–345.

Sotka EE, Wares JP, Barth JA, Grosberg RK, Palumbi SR. 2004. Strong genetic clines and geographical variation in gene flow in the rocky intertidal barnacle *Balanus glandula*. Molecular Ecology, 13:2143–2156.

Soulé JD. 1960. The distribution and affinities of the littoral marine Bryozoa (Ectoprocta). Systematic Zoology, 9:100–104.

Spinks PQ, Shaffer HB. 2005. Range-wide molecular analysis of the western pond turtle (*Emys marmorata*): cryptic variation, isolation by distance, and their conservation implications. Molecular Ecology, 14:2047–2064.

Stebbins RC. 2003. A Field Guide to Western Reptiles and Amphibians. Third Edition. Houghton Mifflin Co., Boston.

Steinbeck J, Ricketts EF. 1941. Sea of Cortez: A Leisurely Journal of Travel and Research. Viking Press, New York.

Stepien CA. 1999. Phylogeographical structure of the Dover Sole *Microstomus pacificus*: the larval retention hypothesis and genetic divergence along the deep continental slope of the northeastern Pacific Ocean. Molecular Ecology, 8:923–939.

Stepien CA, Rosenblatt RH, Bargmeyer BA. 2001. Phylogeography of the spotted sand bass, *Paralabrax maculatofasciatus*: divergence of Gulf of California and Pacific Coast populations. Evolution, 55:1852–1862.

Stock JM, Hodges KV. 1989. Pre-Pliocene extension around the Gulf of California and the transfer of Baja California to the Pacific Plate. Tectonics, 8:99–115.

Stock JM, Lee J. 1994. Do microplates in subduction zones leave a geological record? Tectonics, 13:1472–1487.

Terry A, Bucciarelli G, Bernardi G. 2000. Restricted gene flow and incipient speciation in disjunct Pacific Ocean and Sea of Cortez populations of a reef fish species, *Girella nigricans*. Evolution, 54:652–659.

Upton DE, Murphy RW. 1997. Phylogeny of the side-blotched lizards (Phrynosomatidae: *Uta*) based on mtDNA sequences: support for a midpeninsular seaway in Baja California. Molecular Phylogenetics and Evolution, 8:104–113.

Valentine JW. 1966. Numerical analysis of marine molluscan ranges on the extratropical northeastern Pacific shelf. Limnology and Oceanography, 11:198–211.

Vine FJ. 1966. Spreading of the ocean floor: new evidence. Science, 154:1405–1415.

Vine FJ, Matthews DH. 1963. Magnetic anomalies over oceanic ridges. Nature, 199:947–949.

Von Huene R. 1969. Geologic structure between the Murray fracture zone and the Transverse Ranges. Marine Geology, 7:475–499.

Vrooman AM. 1964. Serologically differentiated subpopulations of the Pacific sardine, *Sardinops caerulea*. Journal of the Fisheries Research Board of Canada, 21:691–701.

Wake DB. 2006. Problems with species: patterns and processes of species formation in salamanders. Annals of the Missouri Botanical Garden, 93:8–23.

Wallace AR. 1876. The Geographical Distribution of Animals. Macmillan and Co., London.

Wares JP, Gaines SD, Cunningham CW. 2001. A comparative study of asymmetric migration events across a marine biogeographic boundary. Evolution, 55:295–306.

Weaver DW, Doerner DP. 1967. Western Anacapia – a summary of the Cenozoic history of the northern Channel Islands. In: Philbrick RN, editor. Proceedings of the Symposium on the Biology of the California Islands, Santa Barbara Botanic Garden, Santa Barbara, p. 13–20.

Wegener A. 1912. Die Herausbildung der Grossformen der Erdrinde (Kontinente und Ozeane), auf geophysikalischer Grundlage. Petermanns Geographische Mitteilungen, 63:185–195, 253–256, 305–309.

Whorley JR, Alvarez-Castaneda ST, Kenagy GJ. 2004. Genetic structure of desert ground squirrels over a 20-degree-latitude transect from Oregon through the Baja California peninsula. Molecular Ecology, 13:2709–2720.

Wilson JT. 1965. A new class of faults and their bearing on continental drift. Nature, 207:343–347.

Zaldívar-Riverón A, Leon-Regagnon V, de Oca ANM. 2004. Phylogeny of the Mexican coastal leopard frogs of the *Rana berlandieri* group based on mtDNA sequences. Molecular Phylogenetics and Evolution, 30:38–49.

Zink RM. 2002. Methods in comparative phylogeography, and their application to studying evolution in the North American aridlands. Integrative and Comparative Biology, 42:953–959.

Zink RM, Blackwell RC. 1998. Molecular systematics and biogeography of aridland Gnatcatchers (Genus *Polioptila*) and evidence supporting species status of the California Gnatcatcher (*Polioptila californica*). Molecular Phylogenetics and Evolution, 9:26–32.

Zink RM, Blackwell RC, RojasSoto O. 1997. Species limits in the Le Conte's thrasher. Condor, 99:132–138.

Zink RM, Blackwell-Rago RC. 2000. Species limits and recent population history in the Curve–billed Thrasher. Condor, 102:881–886.

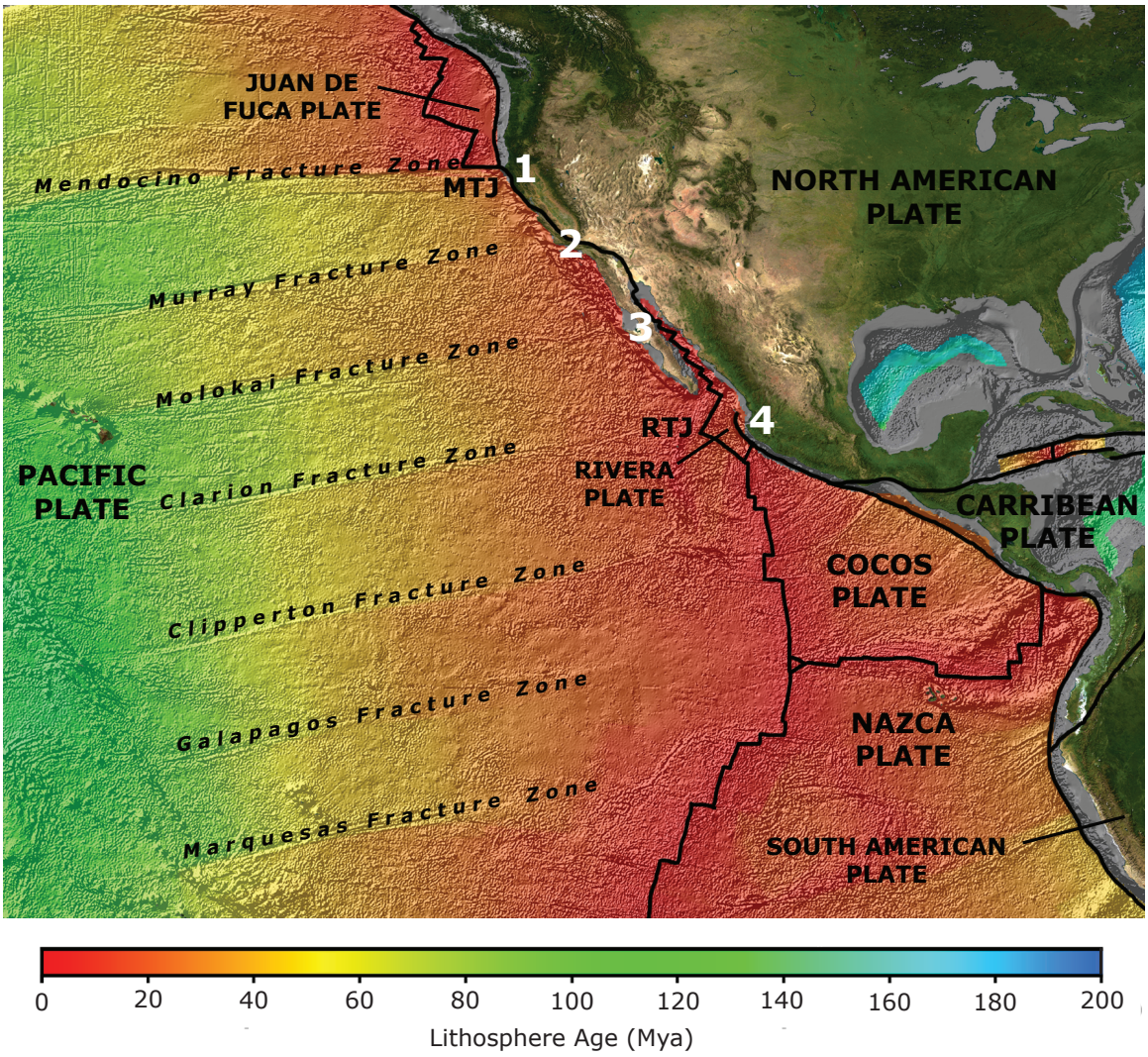

Figure 1.1. Map of the lithosphere of the northeast Pacific Ocean, modified from Müller et al. (2008). The lithosphere is colored by age (millions of years ago, Mya) and plate boundaries are outlined in black. For simplicity, microplates are not shown. The Great Pacific Fracture Zones (GPFZs) are parallel, linear, evenly spaced, and correspond with large offsets in lithosphere age and depth. Four GPFZs are aligned with prominent capes and zoogeographic boundaries where the Pacific plate contacts the western margin of the North American plate along the San Andreas Fault (SAF) system: (1) The Mendocino triple junction (MTJ), Cape Mendocino, and the North Coast Divide, (2) Point Conception and the Transverse Ranges, (3) Punta Eugenia and the Vizcaíno Desert, and (4) The Rivera triple junction (RTJ), Cabo Corrientes and the Sierra Transvolcanica. Computerized digital images and associated databases were obtained from the National Geophysical Data Center, National Oceanic and Atmospheric Administration, U.S. Department of Commerce (http://www.ngdc.noaa.gov). For a detailed map of time-calibrated isochrons, see Plate 3C accompanying Atwater (1989).

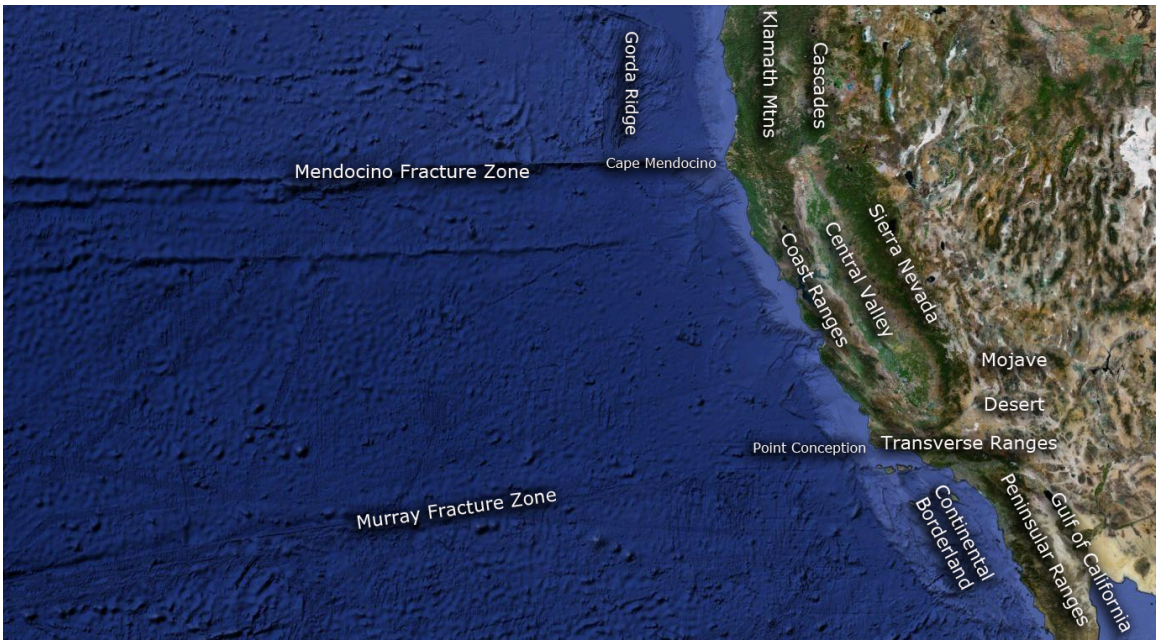

Figure 1.2. The Mendocino Fracture Zone (FZ) corresponds with Cape Mendocino and the North Coast Divide, which delimits the southern limit of the Cascadia subduction zone (the Gorda Ridge and Cascades) and the Klamath Mountains, and delimits the northern limit of the Coast Ranges, Central Valley, and Sierra Nevada (the 'Alta California' province). The Murray FZ, which corresponds with Point Conception and the Transverse Ranges, delimits the southern end of the Alta California province and the northern limit of the Continental Borderland, Peninsular Ranges, and Gulf of California provinces. Computerized digital images were obtained from Google, Inc. (http://maps.google.com).

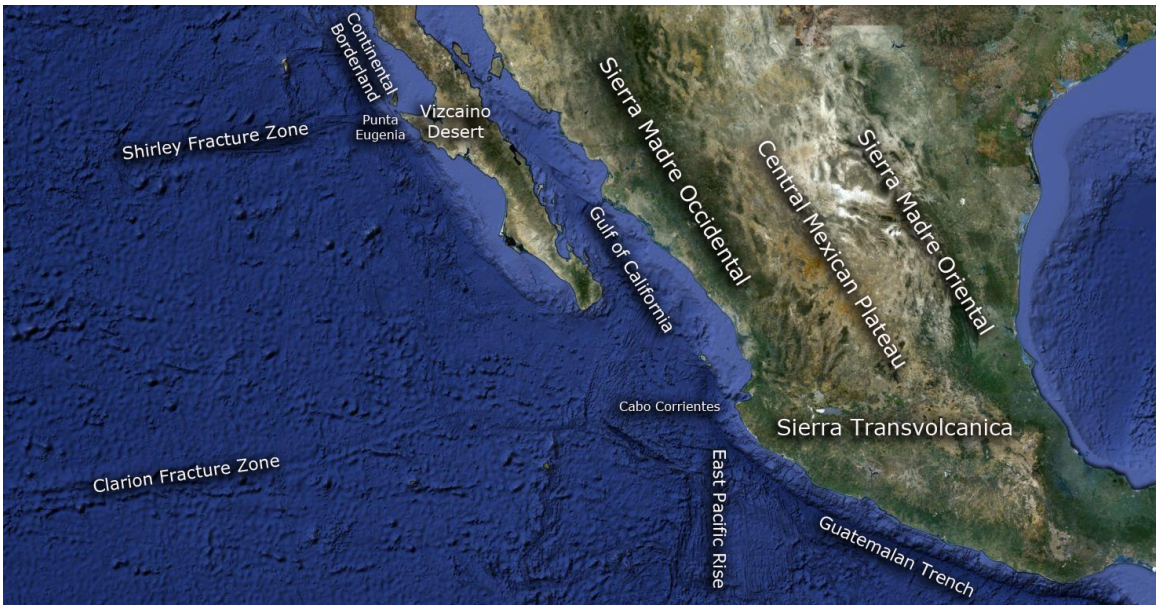

Figure 1.3. The Shirley Fracture Zone (FZ) (the eastern extension of the Molokai FZ) corresponds with Punta Eugenia and the Vizcaíno Desert (the mid-peninsular break), which delimits the southern limit of the Continental Borderland. The Clarion FZ corresponds with Cabo Corrientes and the Sierra Transvolcanica, which delimits the southern limit of the Gulf of California, Sierra Madre Occidental, Central Mexican Plateau, and the Sierra Madre Oriental, as well as the northern limit of the Guatemalan Trench and Cocos Plate. Computerized digital images were obtained from Google, Inc. (http://maps.google.com).

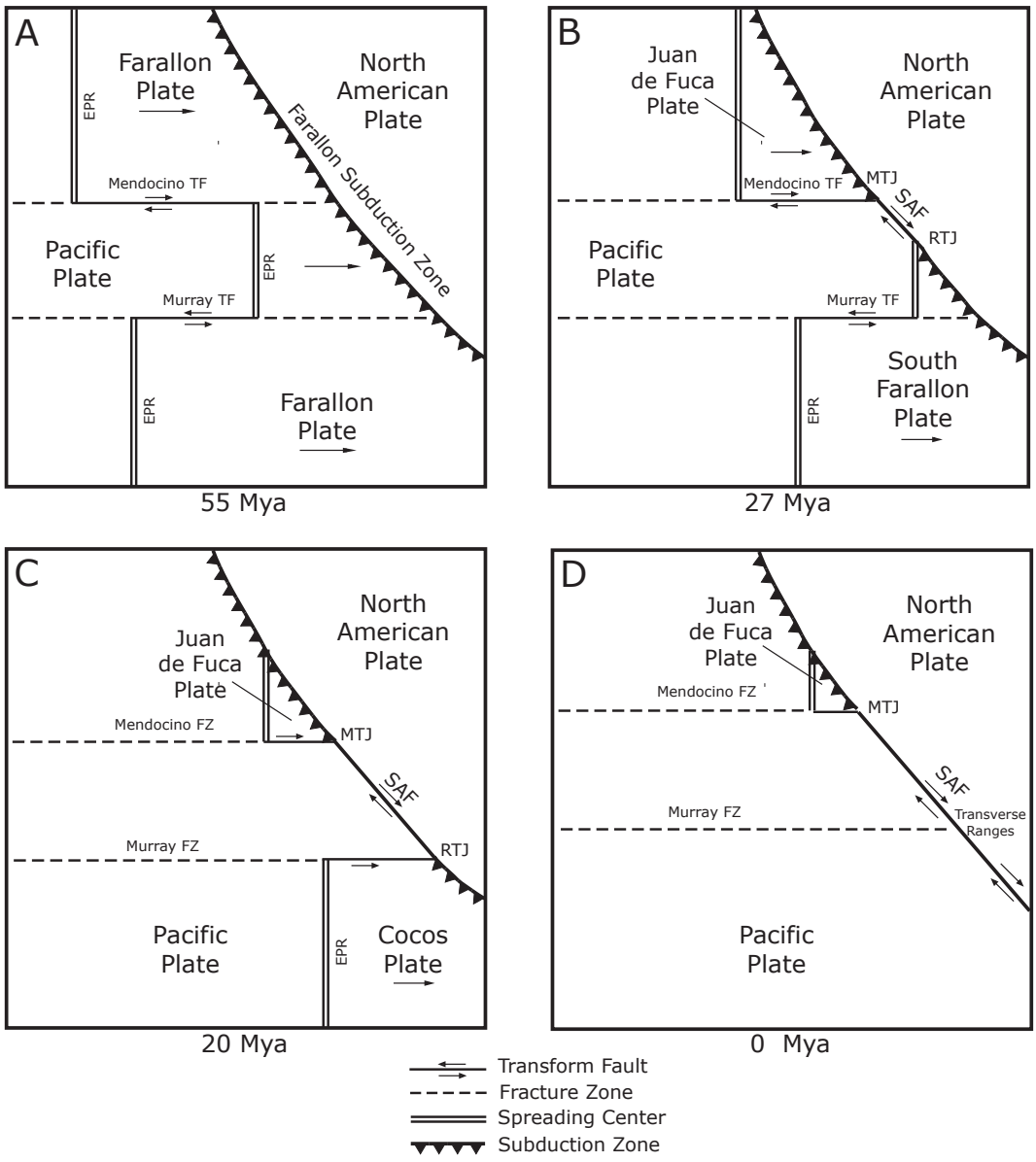
A
Farallon Plate
North American Plate
EPR
Farallon Subduction Zone
Mendocino TF
Pacific Plate
EPR
Murray TF
EPR
Farallon Plate
55 Mya
B
Juan de Fuca Plate
North American Plate
Mendocino TF
MTJ
SAF
RTJ
Pacific Plate
Murray TF
South Farallon Plate
EPR
27 Mya
C
North American Plate
Juan de Fuca Plate
Mendocino FZ
MTJ
SAF
Murray FZ
RTJ
Pacific Plate
EPR
Cocos Plate
20 Mya
D
Juan de Fuca Plate
North American Plate
Mendocino FZ
MTJ
SAF
Murray FZ
Transverse Ranges
Pacific Plate
0 Mya
Transform Fault
Fracture Zone
Spreading Center
Subduction Zone

Figure 1.4. Greatly simplified model of the Cenozoic formation of the northern San Andreas Fault (SAF) system, modified from McKenzie and Morgan (1969), Atwater and Menard (1970), and Atwater (1970). For clarity, all fracture zones (FZs) except the Mendocino and Murray are excluded, microplates are not shown, and diagrams are not drawn to scale. (A) At the time of formation of magnetic anomaly 24 (~55 Mya) in the early Eocene, the East Pacific Rise (EPR) separated the Pacific and Farallon plates while the latter was consumed beneath North America at the Farallon Subduction Zone (FSZ). Note that the spreading centres of the EPR were offset by active transform faults (TFs) aligned with aseismic FZs in the adjacent oceanic lithosphere. (B) In the late Oligocene at anomaly 7 (~27 Mya), the EPR contacted the FSZ, allowing the Pacific and North American plates to meet, creating the new SAF transform system, bounded by the Mendocino triple junction (MTJ) in the northwest and the Rivera triple junction (RTJ) in the southeast. Likewise, the Farallon plate gave rise to the Juan de Fuca plate and the South Farallon plate (the predecessor of the Cocos plate), each with its own subduction zone bordering North America. (C) At anomaly 6 (~20 Mya) in the early Miocene, the MTJ and RTJ migrated away from each other as the SAF lengthened. Note that the Murray FZ was aligned with the RTJ during this period, during which time the Transverse Ranges began their clockwise rotation (Nicholson et al., 1994). (D) In the present day, the MTJ is active in northwestern California, while the Murray FZ and the rotated Transverse Ranges of southern California mark the former site of the RTJ, which is currently located further southeast in mainland México (not shown). For more complex tectonic models that incorporate rifting in the Gulf of California, microplate capture and rotation, propagating rifts, and palaeoshorelines, see Lonsdale (1989), Nicholson et al. (1994), Stock and Lee (1994), Hey (1998), and C. A. Hall (2002).

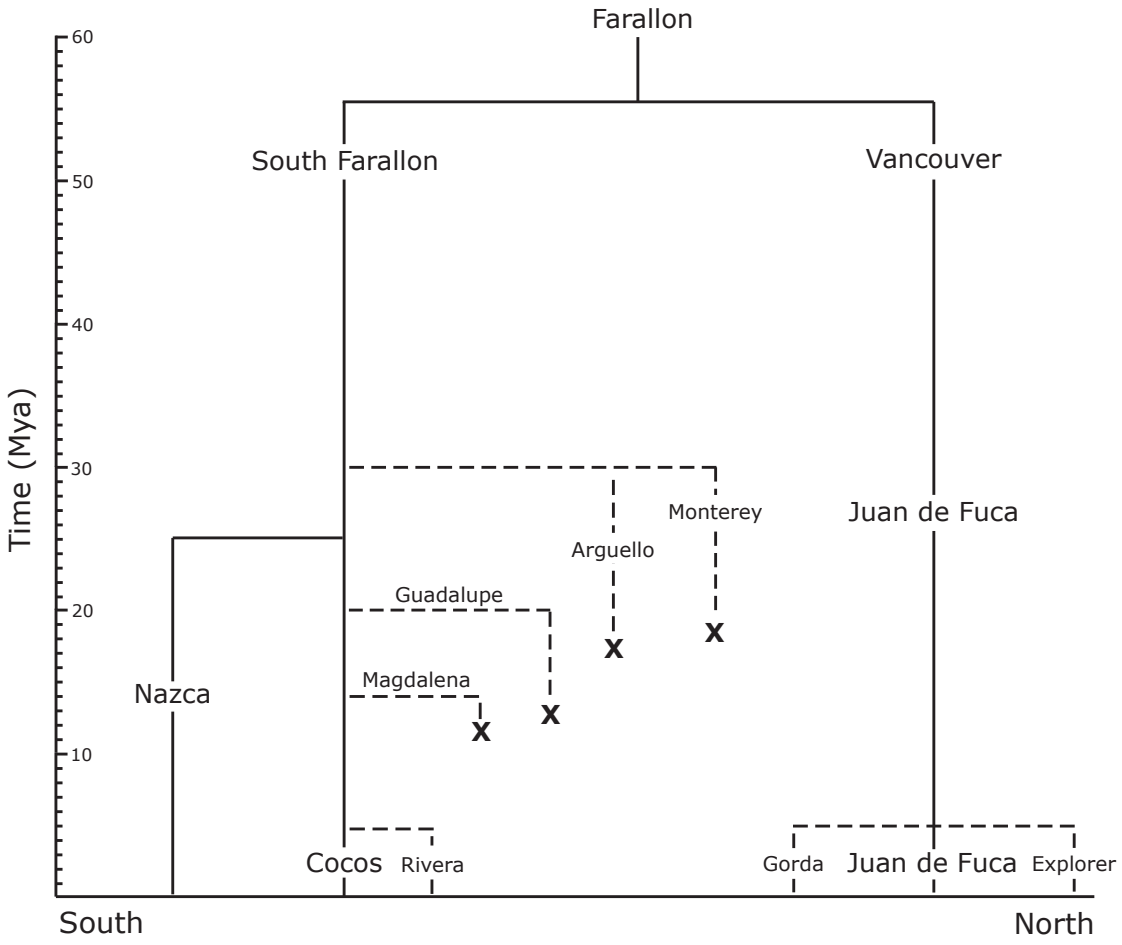

Figure 1.5. Time-calibrated genealogy of the breakup of the Farallon plate, following the microplate histories and dates detailed in Atwater (1989) and Lonsdale (1991), adapted from Stock and Lee (1994). Dashed lines with smaller font sizes indicate microplates, while solid lines with larger font sizes indicate major plate boundaries. 'X' symbols denote microplate deaths.

Table 1.1. Example taxa that recognize the Cape Mendocino (Mendocino triple junction) biogeographic boundary. Data types used to define these breaks include mitochondrial DNA (mtDNA) and nuclear DNA (nDNA).

Taxon	Common name	Break type	Citation
Aneides flavipunctatus	Black Salamander	Intraspecific (morphology, mtDNA, nDNA)	Reilly & Wake (2014)
Aneides lugubris	Arboreal salamander	Northern range limit; chromosomal differentiation at northern end of range	Sessions & Kezer (1987)
Anthopleura elegantissima	Aggregating anenome (zooxanthellae)	Intraspecific (mtDNA)	Sanders & Palumbi (2011)
Antrodiaetus riversi	Turret spider	Intraspecific (mtDNA, nDNA)	Hedin et al. (2013)
Aspidoscelis tigris	Tiger whiptail	Northern range limit	Stebbins (2003)
Balanus glandula	Acorn barnacle	Intraspecific (mtDNA, nDNA)	Sotka et al. (2004)
Charina bottae	Rubber boa	Intraspecific (mtDNA)	Rodriguez-Robles et al.(2001)
Elgaria coerulea	Northern alligator lizard	Intraspecific (morphology)	Stebbins (2003)
Elgaria multicarinata	Southern alligator lizard	Intraspecific (morphology)	Stebbins (2003)
Ensatina eschscholtzii	Ensatina salamander	Intraspecific (morphology, mtDNA)	Kuchta et al. (2009)
Hemigrapsus nudus	Shore crab	Intraspecific (mtDNA)	Kelly & Palumbi (2010)
Lottia digitalis	Ribbed limpet	Intraspecific (mtDNA)	Kelly & Palumbi (2010)
Microstomus pacificicus	Pacific sole	Intraspecific (mtDNA)	Stepien (1999)
Pagurus granosimanus	Granular hermit crab	Intraspecific (mtDNA)	Kelly & Palumbi (2010)
Pagurus hirsutiusculus	Hairy hermit crab	Intraspecific (mtDNA)	Kelly & Palumbi (2010)
Plethodon elongatus	Del Norte salamander	Southern range limit	Stebbins (2003)
Sebastes flavidus	Yellowtail rockfish	Intraspecific (mtDNA)	Sivasundar & Palumbi (2010)
Sebastes mystinus	Blue rockfish	Intraspecific (mtDNA, nDNA)	Cope (2004), Burford & Bernardi (2008) and Burford (2009)
Spea hammondii	Western spadefoot toad	Northern range limit	Stebbins (2003)
Strix occidentalis	Spotted owl	Intraspecific (mtDNA)	Barrowclough et al. (1999)
Taricha rivularis	Red-bellied newt	Northern range limit	Stebbins (2003)
Thamnophis ordinoides	Northwestern garter snake	Southern range limit	Stebbins (2003)
Zonotrichia leucophrys	White-crowned sparrow	Intraspecific (allozymes)	Corbin (1981)

Table 1.2. Example taxa that recognize the Point Conception/Transverse Ranges biogeographic boundary. Data types used to define these breaks include mitochondrial DNA (mtDNA) and nuclear DNA (nDNA).

Taxon	Common name	Break type	Citation
Aliatypus spp.	Trapdoor spiders	Interspecific (mtDNA, nDNA)	Satler et al. (2011)
Ambystoma californiense	California tiger salamander	Southern range limit	Stebbins (2003)
Batrachoseps gabrieli	San Gabriel slender salamander	Local endemic	Stebbins (2003)
Charina bottae	Rubber boa	Intraspecific (mtDNA)	Rodriguez-Robles et al. (2001)
Contia tenuis	Sharp-tailed snake	Southern range limit	Stebbins (2003)
Crotalus oreganus/helleri	Pacific rattlesnakes	Intraspecific (mtDNA)	Pook et al. (2000)
Crotalus ruber	Red diamond rattlesnake	Northern range limit	Stebbins (2003)
Crotaphytus vestigium	Baja California collared lizard	Northern range limit	Stebbins (2003)
Diadophis punctatus	Ringneck snake	Intraspecific (mtDNA)	Feldman & Spicer (2006)
Elgaria multicarinata	Southern alligator lizard	Intraspecific (mtDNA)	Feldman & Spicer (2006)
Embiotoca jacksoni	Black surfperch	Intraspecific (mtDNA)	Bernardi (2000)
Emys marmorata	Western pond turtle	Intraspecific (mtDNA)	Spinks & Shaffer (2005)
Ensatina eschscholtzii	Ensatina salamander	Intraspecific (morphology, mtDNA)	Kuchta et al. (2009)
Gambelia sila	Blunt-nosed leopard lizard	Southern range limit	Stebbins (2003)
Hesperia comma	Holartic skipper	Intraspecific (mtDNA)	Forister et al. (2004)
Lampropeltis zonata	California mountain kingsnake	Intraspecific (mtDNA, nDNA)	Rodriguez-Robles et al. (1999) and Myers et al. (2013)
Ligia occidentalis	Rock louse	Intraspecific (mtDNA, nDNA)	Eberl et al. (2013)
Microtus californicus	California vole	Intraspecific (mtDNA, nDNA)	Conroy & Neuwald (2008)
Neotoma fuscipes	Dusky-footed woodrat	Intraspecific (mtDNA)	Matocq (2002)
Oxylebius pictus	Painted greenling	Intraspecific (allozymes)	Davis et al. (1981)
Pachygrapsus crassipes	Lined crab	Intraspecific (mtDNA)	Cassone & Boulding (2006)
Peromyscus californicus	California deer mouse	Intraspecific (allozymes, morphology)	Smith (1979)

Taxon	Common name	Break type	Citation
Pleistiodon skiltonianus	Western skink	Intraspecific (mtDNA)	Richmond & Reeder (2002)
Pseudacris cadaverina	California tree frog	Intraspecific (mtDNA)	Phillipsen & Metcalf (2009)
Rana draytonii	California red-legged frog	Intraspecific (nDNA)	Richmond et al. (2013)
Sceloporus orcutti	Granite spiny lizard	Northern range limit	Stebbins (2003)
Sepedophilus castaneus	Rove beetle	Intraspecific (mtDNA)	Chatzimanolis & Caterino 2007
Sorex ornatus	Ornate shrew Intraspecific (mtDNA)	Intraspecific (mtDNA)	Maldonado et al. (2001)
Taricha tarosa	California newt	Intraspecific (allozymes, mtDNA)	Kuchta & Tan (2006)
Thamnophis couchii	Sierra garter snake	Southern range limit	Stebbins (2003)
Tigriopus californicus	Intertidal copepod	Intraspecific (mtDNA)	Burton & Lee (1994)
Timema spp.	Walking sticks	Interspecific (mtDNA)	Law & Crespi (2002)
Toxostoma redivivum	California thrasher	Intraspecific (mtDNA)	Sgariglia & Burns (2003)
Uma scoparia/inornata	Fringe-toed lizards	Interspecific (nDNA)	Gottscho et al. (2014)

Table 1.2 (continued).

Table 1.3. Example taxa that recognize the Punta Eugenia/Vizcaíno Desert biogeographic boundary. Data types used to define these breaks include mitochondrial DNA (mtDNA) and nuclear DNA (nDNA).

Taxon	Common name	Break type	Citation
Ammospermophilus leucurus	Ground squirrel	Intraspecific (mtDNA)	Whorley et al. (2004)
Anisotremus davidsonii	Sargo	Intraspecific (mtDNA)	Bernardi et al. (2003)
Aspidoscelis tigris	Tiger whiptail	Intraspecific (mtDNA)	Radtkey et al. (1997)
Auriparus flaviceps	Verdin	Intraspecific (mtDNA)	Zink (2002)
Axoclinus nigricaudus	Cortez triplefin	Intraspecific (mtDNA)	Riginos (2005)
Callisaurus draconoides	Zebra-tailed lizard	Intraspecific (mtDNA)	Lindell et al. (2005)
Campylorhynchus brunneicapillus	Cactus wren	Intraspecific (mtDNA)	Zink (2002)
Chaetodipus baileyi complex	Bailey's pocket mouse	Intraspecific (mtDNA)	Riddle et al. (2000)
Coralliozetus microbes	Scarletfin blenny	Intraspecific (mtDNA)	Riginos (2005)
Crotaphytus vestigium	Baja California collared lizard	Intraspecific (mtDNA)	McGuire et al. (2007b)
Dipodomys merriami	Merriam's Kangaroo Rat	Intraspecific (mtDNA)	Riddle et al. (2000)
Embiotoca jacksoni	Black surfperch	Intraspecific (mtDNA)	Bernardi (2000)
Engraulis mordax	Northern anchovy	Intraspecific (morphology)	McHugh (1951)
Fundulus parvipinnis	California killifish	Intraspecific (mtDNA)	Bernardi & Talley (2000)
Gillichthys mirabilis	Longjaw mudsucker	Intraspecific (mtDNA)	Huang & Bernardi (2001) and Bernardi et al. (2003)
Girella nigricans	Opaleye	Intraspecific (mtDNA)	Terry et al. (2000)
Homalonychus theologus	Spider	Intraspecific (mtDNA)	Crews & Hedin (2006)
Hypsiglena spp.	Night snakes	Interspecific (mtDNA, nDNA)	Mulcahy & Macey (2009)
Ligia spp.	Intertidal isopods	Interspecific (mtDNA)	Hurtado et al. (2010)
Malacoctenus hubbsi	Redsided blenny	Intraspecific (mtDNA)	Riginos (2005)
Paralabrax maculatofasciatus	Spotted sand bass	Intraspecific (mtDNA)	Riginos (2005), Stepien et al. (2001)
Parastichopus parvimensis	Warty sea cucumber	Intraspecific (mtDNA)	Haupt (2011)
Peromyscus eremicus complex	Cactus mouse	Intraspecific (mtDNA)	Riddle et al. (2000)
Phrynosoma coronatum complex	Coast horned lizards	Interspecific (morphology, mtDNA)	Montanucci (2004) and Leaché et al. (2009)
Pipilo crissalis	California towhee	Intraspecific (nDNA)	Zink (2002)
Pituophis vertebralis	Baja California gopher snake	Intraspecific (mtDNA)	Rodriguez-Robles & DeJesus-Escobar (2000)
Sardinops caerulea	Sardine	Intraspecific (morphology)	Vrooman (1964)
Sauromalus obesus/australis	Chuckwallas	Interspecific (morphology,mtDNA)	Hollingsworth (1998) and Petren & Case (1997)

Taxon	Common name	Break type	Citation
Sceloporus zosteromus	Baja California spiny lizard	Intraspecific (morphology, mtDNA)	Grismer (2002) and Leaché & Mulcahy (2007)
Tigriopus californicus	Intertidal copepod	Intraspecific (mtDNA)	Peterson et al. (2013)
Toxostoma lecontei	Leconte's thrasher	Intraspecific (mtDNA)	Zink et al. (1997)
Trimorphodon biscutatus	Lyre snake	Intraspecific (mtDNA)	Devitt (2006)
Urosaurus nigricaudus	Black-tailed brush lizard	Intraspecific (allozymes, mtDNA)	Aguirre et al. (1999) and Lindell et al. (2008)
Uta stansburiana	Side-blotched lizard	Intraspecific (allozymes, mtDNA)	Murphy (1983) and Upton & Murphy (1997)

Table 1.3 (continued).

Table 1.4. Example taxa that recognize the Cabo Corrientes/Sierra Transvolcanica (Rivera triple junction) biogeographic boundary. Data types used to define these breaks include mitochondrial DNA (mtDNA) and nuclear DNA (nDNA).

Taxon	Common name	Break type	Citation
Baiomys musculus	Pygmy mouse	Intraspecific (mtDNA)	Amman & Bradley (2004)
Chiropterotrion spp.	Salamanders	Interspecific (allozymes)	Darda (1994)
Chirostoma spp.	Silversides	Interspecific (morphology)	Barbour (1973)
Crotalus triseriatus complex	Dusky rattlesnakes	Interspecific (mtDNA)	Bryson et al. (2011b)
Herichthys spp.	Heroine cichlids	Interspecific (mtDNA)	Hulsey et al. (2004)
Lampropeltis trianglum complex	Milksnakes	Interspecific (mtDNA, nDNA)	Ruane et al. (2013)
Leptodeira spp.	Cat-eyed snakes	Interspecific (mtDNA, nDNA)	Daza et al. (2009)
Ollotis nebulifer/valliceps	Gulf Coast toads	Interspecific (mtDNA)	Mulcahy & Mendelson (2000)
Pavona gigantean	Coral	Intraspecific (nDNA)	Saavedra-Sotelo et al. (2011)
Pituophis deppei/lineaticollis	Gopher snakes	Interspecific (mtDNA)	Bryson et al. (2011a)
Poecilia butleri	Molly	Intraspecific (mtDNA)	Mateos (2005)
Poeciliopsis spp.	Minnows	Interspecific (mtDNA)	Mateos (2005)
Polioptila nigriceps/P. albiloris	Gnatcatchers	Interspecific (mtDNA)	Zink & Blackwell (1998)
Porthidium spp.	Hog-nosed pit vipers	Interspecific (mtDNA)	Bryson et al. (2008)
Pseudoeurycea spp.	Salamanders	Interspecific (allozymes)	Lynch et al. (1983)
Rana berlandieri complex	Leopard frogs	Interspecific (mtDNA)	Zaldívar-Riverón et al. (2004)
Sceloporus bicanthalis	Bunchgrass lizard	Intraspecific (mtDNA, nDNA)	Leaché et al. (2013)
Sceloporus grammicus	Mesquite lizard	Intraspecific (allozymes, mtDNA)	Arévalo et al. (1993)
Sceloporus scalaris complex	Bunchgrass lizard	Interspecific (mtDNA, nDNA)	Grummer et al. (2014)
Stenella longirostris	Spinner dolphin	Intraspecific (morphology)	Perryman & Westlake (1998)
Toxostoma curvirostre	Curve-billed thrasher	Intraspecific (mtDNA)	Zink & Blackwell-Rago (2000)
Trimorphodon biscutatus	Lyre snake	Intraspecific (mtDNA)	Devitt (2006)

Chapter 2: Species Delimitation and Phylogenetic Systematics of Fringe-toed Lizards (Squamata: Phrynosomatidae: *Uma notata* complex) in the Colorado Desert: Applications of Coalescent Theory to Conservation Biogeography

In this study, nuclear DNA sequence and genome-wide SNP data were used to investigate the systematics of fringe-toed lizards of the *Uma notata* complex, including the Federally Threatened/State Endangered Coachella Valley Fringe-toed Lizard (*U. inornata*), in the deserts of southwestern North America. I analyzed both ten nuclear loci collected using Sanger sequencing and genome-wide sequence and single-nucleotide polymorphism (SNP) data collected using restriction-associated DNA (RAD) sequencing. Concatenated phylogenetic methods, parametric and non-parametric clustering algorithms and coalescent species tree/species delimitation models were used to test hypotheses regarding genetic diversity, species delimitation, phylogeny, and historical biogeography. I validated five species-level lineages within this complex, three of which were previously described as full species, one originally described as a subspecies but later synonymized, and one previously documented yet undescribed species from Mohawk Dunes, Arizona, USA. All but one of these lineages were recovered as exclusive with strong support in the concatenated analyses and by clustering algorithms, and all were decisively supported as distinct species by ranking hierarchal nested speciational models with Bayes factors based on coalescent species tree methods. The time-calibrated and geospatial analyses support the hypothesis that Pleistocene glacial cycles promoted allopatric speciation via dispersal across a landscape matrix shaped by older tectonic

events, but I also recovered evidence for a vicariant role of the Colorado River during the Pleistocene. In all, it is recommended that the six lineages here (including *U. scoparia*) each be recognized as distinct species, and I outline possible conservation implications and recommendations of this study.

I. INTRODUCTION

"*In a world losing its biodiversity at an accelerating rate, systematics needs to be seen as a crucially relevant and important science in meeting the challenges of global environmental change*…" – Cracraft (2002, p. 127)

One of the most fundamental questions in systematic biology is also seemingly one of the simplest yet most vexing: what is a species? The utility of this question extends beyond theoretical arguments about the reality of units of nature above the organismal level (Cracraft 2002). Although other measures have been growing in importance, such as phylogenetic (Faith 1992) and genetic diversity (Moritz 2002), for better or for worse, the species remains the single most important unit for applied conservation efforts. Consider, for example, binding treaties such as the Convention on International Trade in Endangered Species (CITES), the International Union for the Conservation of Nature (IUCN) which maintains a Red List of Threatened Species, and individual countries that have enacted legislation to protect imperiled species, including the United States (ESA, 1973), Mexico (NOM-059-SEMARNAT, 2010), Australia (ESPA, 1992), and

Brazil (Act No. 37-N, IBAMA, 1992). Even prioritization schemes that emphasize the conservation of geographically defined high-biodiversity hotspots or wilderness areas over individual species themselves typically rely on species richness or endemism as metrics to quantify regional biodiversity (Myers et al. 2000, Mittermeier et al. 2003, Brooks et al. 2006). In either case, the flow of limited monetary resources towards conserving biodiversity is ultimately linked to species concepts and delimitation criteria, and the ultimate aim is to conserve the unique evolutionary history and ecological function that each species represents (Shaffer 2013). Despite this reliance on species as a measure of biodiversity, we are far from completing an inventory of the total number of species on Earth, so any technical or conceptual advances in our ability to objectively and efficiently delimit species numbers and boundaries should eventually have profound conservation implications (Dimmick et al. 1999, Adams et al. 2014). I stress objectivity because alpha taxonomy has traditionally been based on subjective morphological descriptions (Fujita et al. 2012) and efficiency because time is of the essence. Thus, although conservation biology largely grew out of an ecological tradition with evolutionary biology making a late arrival (Frankel and Soulé 1981, Shaffer 2013), modern systematists still have much to offer this burgeoning field (Funk et al. 2002, Cracraft 2002, Ogden et al. 2013).

The biological species concept (BSC; Mayr 1942, 1963) was widely adopted by most biologists during the second half of the twentieth century. However, many practicing systematists (those describing and delimiting species based on museum specimens) and conservation biologists (those breeding endangered

species in zoos and herbaria) had difficulty applying the BSC due to its ambiguous treatment of discrete morphological or genetic variation, whereby "diagnosable populations might be ranked either as a species or subspecies, or subspecific rank itself might be applied to diagnosably distinct forms as well as to arbitrary subdivisions of clinal variation" (Cracraft 2002, p. 131). The general lineage species concept (GLSC) of de Queiroz (1998, 2007, 2011), essentially extending the evolutionary species concept (ESC; Wiley 1978), attempted to circumvent this and other issues by challenging researchers to distinguish between the theoretical concept of species and the operational criteria that are used to put the concept into practice, setting the stage for new coalescent model-based approaches to test alternative hypotheses regarding the numbers and boundaries of species (Hey 2006, Yang and Rannala 2010, 2014, O'Meara 2010, Fujita et al. 2012, Lim et al. 2012, Carstens et al. 2013, Satler et al. 2013, Ruane et al. 2014). One promising development uses Bayes factors to rank species delimitation models with either sequence data (BFD; Grummer et al. 2014) or genome-wide SNP data (BFD*; Leaché et al. 2014) in a coalescent species tree framework, but few empirical studies have utilized these methods (but see Rittmeyer and Austin 2015). These advances provide ample opportunities to improve the objectivity and efficiency of systematic species delimitation studies.

Here, I further the field of systematic biology by utilizing nuclear DNA sequence and genome-wide SNP data to investigate the systematics of fringe-toed lizards (*Uma notata* complex) in the deserts of southwestern North America. *Uma* are renowned for specialized adaptations to life in desert sand dunes

(Stebbins 1944), including their namesake toe fringes (Carothers 1986) that have also emerged convergently in sand-dwelling lizards on other continents (Luke 1986). Norris (1958, p. 253) recognized that the complete restriction of *Uma* to the extreme conditions of wind-blown (aoelian) sand, "…when correlated with the geologic history of the present habitat, makes possible a zoogeographic study of considerable precision." Aeolian deposits occupied by *Uma* are insular and mobile, are able to migrate tens of meters downwind per year, and are tightly linked to the development of the lower Colorado River, its delta, and other local sand sources (Blount and Lancaster 1990, Muhs et al. 2003). Within *Uma*, two species (*U. exsul* and *U. paraphygas*, not included in this study) are restricted to the Chihuahuan Desert of Mexico, and are sister to the western *Uma* species of the Mojave, Sonoran and Colorado Deserts (Wilgenbusch and de Queiroz 2000). The western *Uma*, in turn are divided into two sister clades: *U. scoparia* of the Mojave Desert and the *U. notata* complex of the Colorado and Sonoran Deserts. The *U. notata* complex, itself arguably composed of several species or subspecies, represents an ideal example of the limitations that the BSC presents to diagnosable allopatric populations: "the systematic problem presented by the genus *Uma* is of the most difficult sort… the fact that all forms are completely isolated eliminates the possibility of obtaining direct criteria for name giving. The systematics of such a group must be based on the opinions of the taxonomist" (Norris 1958, p. 279). The situation was made unnecessarily more confusing by poorly labeled or lost type specimens with erroneous localities (Heifetz 1941). Thus, it is not a surprise that the taxonomy of the *U. notata*

complex has been unstable over the years, with anywhere from one to four species recognized.

Baird (1858, p. 253) described *Uma notata* from the "Mohave Desert", which was later revealed to be a locality error and was localized to the "Colorado Desert" by Heifetz (1941). Baird's description of the type (USNM 4124) was deemed inadequate by Cope (1866, p. 310) who described a second specimen and its toe fringes. However, this specimen (USNM 6063) later became the type specimen of *U. scoparia* (Cope 1894, p. 435). The type locality (Fort Buchanan, Arizona) was another error – no *Uma* are known to occur there, and due to morphological characters (crescents on throat), it must have originated from the Mojave Desert, California (Heifetz 1941). Cope (1895) described *U. inornata* from a single specimen from "the Colorado Desert, San Diego County, California", which lacks ventrolateral blotches. Again, this was doubtless in error and almost certainly from the Coachella Valley of Riverside County, California (Heifetz 1941). The type specimen (USNM 16500) was lost and later redescribed (Cope 1898, p. 281). These errors likely arose from the unfortunate practice of labeling specimens from where they were shipped rather than collected (Heifetz 1941). Cope (1895) also described *U. rufopunctata* (cotypes USNM 21846-52, Yuma Desert, Arizona) named for the reddish tint to its camouflage, matching the color of sands eroded from the Colorado Plateau. Mislead by convergent evolution of fringed toes in sand-dwelling *Callisaurus draconoides crinitus* and *C. d. rhodostictus*, Cope (1896, p. 1049) briefly synonymized *Uma* with *Callisaurus*, but revived *Uma* in 1898 (p. 276-284) after examining more morphological characters. Camp (1916,

p. 516), and later Van Denburgh (1922, p. 132) only recognized one species (*U. notata*), which was unsurprising given the data on hand. Heifetz (1941) reviewed the *Uma* literature, clarified the errors, and described a new subspecies (*U. n. cowlesi*) from Tepoca Bay, Sonora, Mexico (CAS 53370). This subspecies, identified by smaller ventrolateral blotches, is distributed on the northwest Sonoran coast from Puerto Peñasco to Tepoca Bay. Also, he synonymized *U. rufopunctata* with *U. notata*. Norris (1958) recognized two species: *U. scoparia* and *U. notata*, and treated *U. n. inornata*, *U. notata notata*, and *U. n. rufopunctata* as subspecies. He synonymized *U. n. cowlesi* with *U. n. rufopunctata*, although he recognized that "it is possible that it will be revived when more is known of the Sonoran coastal *Uma* from which [Heifetz] described it" (p. 281). Norris was the first to propose detailed biogeographic hypotheses for the origin of *Uma*, proposing that *U. scoparia* speciated from *U. notata* in the early Pleistocene or late Pliocene, dispersing northwest from Pleistocene refugia along the Colorado River.

The onset of molecular data collection did little to stabilize the taxonomy of *Uma*, despite making progress in resolving its phylogeny. Based on allozyme data that showed little variation, Adest (1977) only recognized one species, lumping all western forms into *U. notata*, essentially following Camp (1916) and Van Denburgh (1922). Utilizing morphological characters and allozymes, de Queiroz (1989, 1992) found that *U. scoparia* was sister to the *U. notata* complex (including *U. inornata*), a relationship later confirmed by sequencing mitochondrial cytochrome *b* and 16S regions (Wilgenbusch and de Queiroz 2000,

Schulte and de Queiroz 2008). De Queiroz (1989, 1992) recognized *U. scoparia*, *U. notata notata*, *U. n. rufopunctata*, and *U. inornata*, while acknowledging that "most of the variation among different taxonomies of the sand lizards is accounted for by disagreements concerning the ranks of taxa rather than about the existence of the entities that have been given names" (p. 6). Trepanier and Murphy (2001), using expanded phylogeographic sampling and mitochondrial DNA (mtDNA), recognized five species: *U. scoparia*, *U. inornata*, *U. notata* (west of the Colorado River), *U. rufopunctata* (east of Colorado River) and an undescribed species from Mohawk Dunes, Arizona that was sister to the rest of the *U. notata* complex. Furthermore, the mtDNA gene tree showed a nested pattern, in which the southeastern populations are basal and the northwestern most derived, so that *U. rufopunctata* was non-exclusive (paraphyletic), a pattern that would have been impossible to detect without denser population-level sampling. Despite this, Stebbins (2003) recognized only three species: *U. scoparia*, *U. inornata*, and *U. notata*. Jones and Lovich (2009) recognized *U. scoparia*, *U. inornata*, *U. notata*, and *U. rufopunctata*, with this taxonomy also being followed by Crother (2012) and the IUCN Red List (www.iucnredlist.org, accessed March 4, 2015). Gottscho et al. (2014) analyzed fourteen nuclear loci with coalescent models, and estimated Pleistocene speciation of *U. scoparia* from the *U. notata* complex. Within *U. notata*, the authors found a genetic diversity hotspot in the Algodones Dunes, supporting the hypothesis of Norris (1958) that *U. scoparia* may have originated by dispersing northwards along the Colorado River from the Gran Desierto de Altar, a 5,700 km^2 sand sea, the largest in North America (Brusca 2014). However,

this study, being focused on *U. scoparia*, lacked thorough sampling of the *U. notata* complex, including only one sample each of *U. inornata* and *U. rufopunctata* and entirely lacking samples from Mexico.

In a pragmatic sense, this messy taxonomy and biogeography ultimately has important conservation implications. In 1980, *U. inornata* was listed as 'threatened' under the Federal Endangered Species Act and 'endangered' under the California Endangered Species Act, yet since listing, its distribution has been reduced by ~60%, so that this species has lost 90-95% of its historical habitat in total (USFWS 1980, 2010). The remaining 15,000-20,000 acres of habitat faces degradation from a variety of sources (USFWS 2010), including off-highway vehicles (OHVs), invasive plants such as Saharan mustard (*Brassica tournefortii*), obstructions to blowsand transport systems, and changes in hydrology (lower water tables leading to loss of mesquite hummocks). Furthermore, small effective population sizes result from habitat fragmentation, resulting in low genetic diversity (Hedtke et al. 2007) and a higher risk of inbreeding depression (Soulé and Mills 1998, Frankham et al. 2002). Climate change is predicted to result in a 1-3º C increase in average temperature and a 12-35% decrease in precipitation in the Coachella valley by the year 2050 (Cayan et al. 2009), potentially with dire consequences for ectothermic lizards with poor dispersal abilities (McLaughlin et al. 2002, Sinervo et al. 2010). Chen et al. (2006) developed a model to predict the time to extinction based on habitat patch size and estimate the propensity of extinction of *U. inornata* in isolated habitat patches; the model predicted that the population on the Thousand Palms Conservation Area would go extinct in 78

years, which is troubling because this population is the largest and most robust of the species. Other members of the *U. notata* complex face the same suite of anthropogenic threats, albeit to a lesser degree. *Uma notata* is listed as a California Species of Special Concern, and faces intensive off-road vehicle activity at nearly every locality where this species is found in the state, including Borrego Springs, Ocotillo Wells, Truckhaven, Superstition Mountain, and Algodones Dunes (Figure 2.1). Impacts to *U. rufopunctata* are less severe in Mexico – the Gran Desierto de Altar, the largest dune complex in North America, is mostly protected as a biosphere reserve. *Uma notata* in California also faces potential threats from renewable energy development (Desert Renewable Energy Conservation Plan, www.drecp.org) and the international border fence, completed in 2008, which now likely prevents gene flow between conspecifics in the U.S. and Mexico. Thus, any changes in the taxonomy or insights into the processes shaping geographic patterns of genetic diversity might have important management implications.

For this chapter, I used a combination of ten nuclear loci (collected using Sanger sequencing) and genome-wide sequence and SNP data (collected using Illumina sequencing) to address four central questions regarding the *U. notata* complex, echoing four of the seven grand questions of systematics (Cracraft 2002):

1. Diversity: How do patterns of genetic diversity within the *U. notata* species complex inform us about the processes that generated them?

2. Phylogeny: What are the evolutionary relationships among the members of this complex?
3. Classification/species delimitation: How many species are represented in this complex, and where are their boundaries?
4. Biogeography: How have the combination of Pliocene and Pleistocene geological and climatic events, including the development of the San Andreas Fault system, Colorado River, and glacial cycles, affected diversification within this complex?

II. MATERIALS AND METHODS

(1) Field sampling and genomic DNA

Specimens and tissue samples were collected in the United States (California, Arizona) and Mexico (Baja California, Sonora) from 2008 to 2012. Lizards were either briefly captured to remove a piece of tail tissue before releasing the lizard otherwise unharmed, as was the case for all *U. inornata*, *U. scoparia*, and some *U. notata*, including all samples from Gottscho et al. (2014), or lizards were sacrificed as voucher specimens with MS222 injections (Conroy *et al*. 2009), including all Mexican samples and some *U. notata* in California. Liver and tail tip tissues were preserved in 96-100% ethanol and stored at -20° C. All vouchers are deposited at the San Diego Natural History Museum, and tissues are stored at -80 °C at San Diego State University. Following the most recent taxonomies of Jones and Lovich (2009) and Crothers (2012), I included 26

samples of *U. inornata*, 16 *U. notata*, 23 *U. rufopunctata* (including five samples from *U.* sp. nov. from Mohawk Dunes and 12 representing the sunk taxon *U. n. cowlesi* from the Sonoran coastal plain) and 31 *U. scoparia* as outgroups (96 total; Table 2.1, Figure 2.1). The DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) with RNAse A was used to extract genomic DNA (gDNA) following the manufacturer's instructions.

(2) Sanger data acquisition and processing

I included published data from ten nuclear loci, including four exons (BDNF, RAG-1, PNN, R35) and six anonymous loci (sun07, sun08, sun10, sun12, sun18, sun28) collected by Gottscho et al. (2014), representing 31 *U. scoparia,* nine *U. notata*, one *U. inornata* and one *U. rufopunctata* from Mohawk Dunes. This dataset was expanded using the polymerase chain reaction (PCR) primers and laboratory protocol of Gottscho et al. (2014) to amplify additional individuals for these ten loci. Cleaned PCR products were sequenced on a 96-capillary Sanger sequencer (ABI 3730XL, Applied Biosystems). In total, I analyzed Sanger data for 88 individuals, including 18 *U. inornata,* 16 *U. notata*, 6 *U. rufopunctata*, 12 *U. cowlesi*, five *U.* sp. nov., and 31 *U. scoparia* (Table 2.1). Sequencher v4.7 (Gene Codes Corp., Ann Arbor, MI) was used to analyze data quality, trim primer sequences, produce alignments and call heterozygous sites. PHASE v2.1 (Stephens et al. 2001) and seqPHASE (Flot 2010) were used to determine haplotypes using default settings. DNAsp v5.10 (Librado and Rozas 2009) was used to calculate summary statistics for the Sanger data, including number of

segregating sites per locus, nucleotide diversity for each species at each locus, fixed differences diagnosing species, and Tajima's D. Because this program cannot accommodate sequences with ambiguity codes, I chose the most probable set of alleles and base calls for all sites as estimated by PHASE to estimate summary statistics.

(3) RAD data acquisition and processing

I collected genomic SNP data using the double-digest Restriction-Associated-DNA sequencing (ddRADseq) protocol (Peterson et al. 2012), following the enzymes and size selection window of Leaché et al. (2014, 2015), including 13 *U. inornata*, 12 *U. notata*, six *U. rufopunctata*, 10 *U. cowlesi*, five *U.* sp. nov., and 18 *U. scoparia* (64 total, Table 2.1). DNA concentrations were measured using a Qubit 2.0 Fluorometer (Life Technologies, Grand Island, NY) and assessed samples for high molecular weight by running them out on 1% agarose gels. The high-fidelity restriction enzymes SbfI and MspI (New England Biolabs, Ipswich, MA) were used to digest 200-500 ng of gDNA per lizard. Digestions were then purified with Agencourt AMPure beads (Beckman Coulter, Danvers, MA) before attaching uniquely bar-coded adapters to each library with T4 Ligase (New England Biolabs, Ipswich, MA). To reduce the odds of incorrectly assigning samples due to sequencing error, each barcode differed by a minimum of two bases (assuming a Phred score of 20, the minimum quality threshold, 0.01% of reads would be incorrectly assigned). After a second AMPure cleanup, I generated eight pools of eight uniquely bar-coded libraries at equimolar

concentrations before size selecting fragments 415-515 bp long using a Pippen Prep (Sage Science, Beverly, MA) with 2% gel cassettes. Phusion *Taq* Polymerase (New England Biolabs) was used for the Polymerase Chain Reaction (PCR). I used a two-step cycle (98° C for 10 seconds, 72° C for 20 seconds) 12 times, followed by a final extension step of 72° C for 10 minutes. In this way, all 64 lizards could be identified using the hierarchical bar-coding scheme. An Agilent Bioanalyzer 2100 was used to ensure that the libraries were at the appropriate concentration and size distribution for sequencing. I sequenced the final libraries (100 bp single end reads) on half a flow-cell lane of a HiSeq 2500 (Illumina, San Diego, CA) at the Institute of Integrative Genome Biology (University of California, Riverside).

The Python pipeline pyRAD v2.1.2 was used for RAD data quality control, alignment, and genotype calling (Eaton 2014). The benefit of pyRAD is that it accommodates indel variation through its use of USEARCH (Edgar 2010) and MUSCLE (Edgar 2004), improving homolog identification across phylogenetically divergent samples (Eaton and Ree 2013, Leaché et al. 2015). The pipeline followed seven steps: 1) raw reads were de-multiplexed by barcode; 2) restriction cut sites and adapter sequences were trimmed, bases with Phred scores less than 20 were replaced with 'N' and reads with more than 10 Ns were removed; 3) reads were clustered within individuals into putative loci with USEARCH v7.0.1090 (Edgar 2010) using a clustering threshold of 0.85, retaining only loci with at least 10x coverage and with 10 or fewer Ns; 4) error rate and heterozygosity were jointly estimated; 5) consensus base calling was performed

using the mean error and heterozygosity rates of step four, while consensus sequences with more than five heterozygous sites were excluded (first paralog filter); 6) consensus sequences were clustered across individuals using MUSCLE v3.8.31 (Edgar 2004) with a 0.85 clustering threshold; 7) loci in which more than 50% of samples had shared heterozygosity at a site were removed (second paralog filter), and alignments of both full loci (89 bp) and variable sites (individual SNPs) were generated in a variety of formats (Nexus, Phylip, Structure, etc.) for further analysis.

There is still no consensus how to treat missing data in RADseq analyses – excessive levels of missing data may lead to spurious results from principal components analysis (PCA) and other clustering algorithms (Rheindt et al. 2014), wherein individuals with high levels of missing data may erroneously group out as separate populations, but low tolerance for missing data may bias the mutational rate spectrum of sampled loci, so that slowly evolving loci are disproportionately represented in the final dataset (Huang and Knowles 2014). Also, concatenated phylogenetic analyses have been shown to benefit from longer alignments even with a higher percentage of missing data (Wagner et al. 2013). Given this uncertainty, I generated two datasets with differing levels of completeness: a larger, less complete matrix with loci represented by at least 40 of 64 individuals, and a smaller matrix retaining only loci represented by at least 60 of 64 individuals.

(4) Concatenated phylogenetic inference

I constructed concatenated phylogenies using two different methods in order to discover putative lineages to be later validated using species delimitation models. First, a ML phylogeny of the concatenated data was generated with RAxML v8.1.1 (Stamatakis 2014). This rapid approach does not require *a priori* assignment to populations or species and thus is well suited to inform subsequent species tree analyses (de Queiroz and Gatesy 2007, Wagner et al. 2013). However, combining unlinked loci into a supermatrix and forcing a bifurcating hierarchy at the population level may result in erroneous groupings, so shallow-scale topologies and branch lengths should be interpreted cautiously. I concatenated 1,614 loci that were represented in a minimum of 40 of 64 individuals, resulting in a 149,784 base-pair (bp) alignment with 7,752 SNPs and 15.86% missing data, implemented the GTR+GAMMA substitution model, performed 500 rapid bootstrap replicates, and rooted the tree with the midpoint method, confirming that *U. scoparia* was the outgroup.

To estimate a time-calibrated phylogeny, I used BEAST v2.1.2 (Bouckaert et al. 2014) to perform an un-partitioned Bayesian analysis of the same 149,784 bp RADseq alignment with a uniform Yule model tree prior, a relaxed lognormal clock, and the GTR+GAMMA substitution model with empirical base frequencies. Calibrating the molecular clock is problematic because all loci are anonymous, no useful fossil calibrations are currently available for *Uma* (de Queiroz 1989), and comparative genomics of squamate reptiles is a field in its infancy. Therefore, I assumed an average mutation rate of 2.2×10^{-9}

subs/site/year based on the average genome-wide mammalian mutation rate (Kumar and Subramanian 2002), and compared my estimated timeline with that of Gottscho et al. (2014). I ran the Markov chain for 50 million generations, sampling trees every 1,000 steps. Convergence was assessed using Tracer v1.6 (Rambaut et al. 2014). After burn-in, I retained 40,000 trees and used TreeAnnotator v1.8.1 (Bouckaert et al. 2014) to produce a maximum clade credibility tree from the posterior distribution with mean node heights.

(5) Population structure

In order to guide subsequent species delimitation analyses, I used both parametric and non-parametric approaches to detect population structure without any *a priori* hypotheses about individual assignment. To infer population structure with the Sanger data, I utilized the R package Geneland v4.0.3 (Guillot et al. 2005, 2008). I encoded individual genotypes using the numbered alleles output by PHASE v2.1 (that is, entire sequences were coded as numerical alleles, rather than the individual SNPs within the sequences). Alternatively, I could have selected a single SNP per locus, but this would reduce the information content of the loci (one SNP can only have four possible alleles, but there is no such limit if unique sequences are coded as alleles). The advantage of Geneland lies in its use of a Bayesian spatial correlated allele frequencies model that incorporates geo-referenced coordinates for each individual and produces maps with posterior probabilities of population assignment. The purpose of the correlated allele frequencies model is to base inference of population structure on

a more informative prior than the default model of uncorrelated frequencies (e.g. Pritchard et al. 2000) and hence to get more accurate results, especially in cases of low genetic differentiation (Falush et al. 2003, Marchini et al. 2004, Guillot et al. 2008). While this correlated model may be more biologically realistic, it has been shown to artificially inflate K, the inferred number of populations (Guillot et al. 2008). Thus, I ran the model five times with K_{max} values ranging from 2 to 6, all of which were used to inform species delimitation models downstream. I ran the Markov Chain Monte Carlo (MCMC) for 20 million generations, sampling every 1,000 generations, discarding the first 10% as burn-in. Maps of population assignment were produced with a custom R script (Dryad).

For the RADseq data, first I conducted a PCA using smartPCA v1.02.10 (Patterson et al. 2006, Reich et al. 2008), analyzing 581 unlinked biallelic SNPs represented in at least 60/64 individuals. The main benefit of PCA is its ability to detect genetic structure without the computational burden of Bayesian clustering algorithms and the absence of assumptions about the underlying population genetic model, but it does not provide a group assessment. Therefore I also used Discriminant Analysis of Principal Components (DAPC) implemented in the R package adegenet v1.4-2 to assign individuals to populations (Jombart et al. 2010). DAPC uses PCA as a prior step to Discriminant Analysis (DA) to transform the data so that variables submitted to DA are perfectly uncorrelated and fewer than the number of sampled individuals, allowing DA to be applied to genome-wide SNP data. This approach follows a two-step process. First, I conducted the PCA, retained the first 20 PCs which retained >80% of the

cumulative variance, and used the Bayesian Information Criterion (BIC) to select the optimal number of clusters (in this case, 6). Second, K-means clustering was implemented by retaining the first 20 PCs and the first 10 discriminant functions to assign individuals to those six clusters.

Finally, I used Admixture v1.23 (Alexander et al. 2009) to detect population structure among identified clusters, analyzing the same data matrix as in the smartPCA analysis (581 unlinked SNPs present in 60 or more individuals). This maximum likelihood (ML) program implements the same underlying population genetic model as the widely used Bayesian program Structure (Pritchard et al. 2000). While both programs assign individuals into clusters in an attempt to minimize linkage disequilibrium and maximize Hardy-Weinberg equilibrium, Admixture has the added benefit of a fast numerical optimization algorithm to decrease computational time while avoiding problems with MCMC convergence. I used the cross-validation procedure to select the optimal K value (Alexander et al. 2009), testing values ranging from 2 to 10, and to estimate pair-wise F_{ST} values between inferred clusters. I used a customized R script, available on Dryad, to plot the results.

(6) Latitude vs. heterozygosity

In order to detect the signal of population bottlenecks or range expansions in northern populations related to Pleistocene glacial cycles, a linear regression of latitude vs. percent heterozygosity (output by pyRAD) was performed for all 64 individuals using the lm function in R. In order to correct for the phylogeny, I

used the phylogenetic generalized least squares (PGLS) method to fit the clock tree produced with BEAST under a maximum likelihood Brownian motion model (gls function, nlme package in R).

(7) Species tree inference

I constructed a coalescent species tree using the Sanger data (ten nuclear loci) with *BEAST (Heled and Drummond 2010) implemented in BEAST v1.8.1 (Drummond et al. 2012). Because this method requires *a priori* species groupings, I assigned individuals based on the results of the concatenated phylogenies and population structure analyses, using the most complex model that treats the six OTUs as distinct species (model H, Figures 2.1, 2.2): 1) *U. scoparia* (Mojave Desert, California and Arizona), 2) *U. inornata* (Coachella Valley, California), 3) *U. notata* (west of the Colorado River, California and Baja California), 4) *U. rufopunctata* (east of the Colorado River and west of El Pinacate, Sonora), 5) *U. cowlesi* (east of El Pinacate, Sonora), and 6) *U. rufopunctata* (Mohawk Dunes, Arizona). I assumed a HKY+GAMMA substitution model for all loci because earlier attempts to use more complex models resulted in a failure of the Markov chain to converge on a stable posterior distribution. I initially ran the analysis with a lognormal relaxed clock, but because the 95% HPD for the coefficient of variation for the substitution rate parameter included zero for each locus, I could not reject a strict molecular clock, thus for the final analysis a strict clock was assumed for all loci. I used a Yule process species tree prior and ran the MCMC for 300,000 generations, sampling the tree space every 10,000 generations. Convergence was

assessed with Tracer v1.6 (Rambaut et al. 2013). After discarding the first 10% of trees as burn-in, I produced a maximum clade consensus species tree with TreeAnnotator v.1.8.1 and visualized the entire post-burn-in posterior distribution with DensiTree v2.2.

For the RADseq data, I used SNAPP v1.1.1 (Bryant et al. 2012) implemented in BEAST v2.1.2 (Bouckaert et al. 2014) to infer a species tree using biallelic SNP data while bypassing gene tree inference in a fully coalescent model. SNAPP does not accommodate missing data at the species level, so I used a custom R script (Dryad) to produce a dataset of 468 unlinked biallelic SNPs present in a minimum of one individual for each of the six putative species listed above. A gamma rate prior distribution was used, and all prior values were left at their default settings, most of which have been shown not to have a large impact on the results (Rittmeyer and Austin 2015). The MCMC was run for 500,000 generations, sampling every 1,000 generations (xml file available on Dryad). After convergence was assessed with Tracer v1.6, I produced a maximum clade credibility tree with TreeAnnotator v1.8.1 after discarding the first 10% of trees as burn-in. The same posterior distribution of trees was also visualized with DensiTree v2.2 (Bouckaert 2010) after discarding burn-in.

(8) Species delimitation with Bayes factors

I tested eight different species delimitation models ranging from two species (*U. scoparia* and *U. notata*) to six species (*U. scoparia*, *U. inornata*, *U. notata*, *U. rufopunctata*, *U. cowlesi,* and *U.* sp. nov.). Intermediate models B-G lumped

these six species in various combinations, including all previously published taxonomies summarized in the introduction as well as novel groupings suggested by the concatenated phylogenies and clustering algorithms. I first used BFD (Grummer et al. 2014) to test alternative species delimitation models with the phased Sanger data. For each model, the species tree was first estimated with *BEAST (Heled and Drummond 2010) implemented in BEAST v1.8.1 (Drummond et al. 2012), using a HKY+GAMMA substitution model with empirical base frequencies for all loci, a strict molecular clock, a Yule process species tree prior, a MCMC chain length of 300,000 generations, sampling the tree space every 10,000 generations, and assessing convergence with Tracer v1.6. Marginal likelihoods for each model were estimated using both the stepping-stone and path sampling methods, with 100 path steps of 100,000 generations each, logging parameters every 1,000 generations (Baele et al. 2012). For the RADseq data, I used BFD* (Leaché et al. 2014) to test the same eight alternative species delimitation models as described above with the same 468 SNP dataset that I used to estimate the species tree with SNAPP. I conducted path sampling with 20 steps and a MCMC chain length of 20,000 steps (10% pre-burn-in) to estimate the marginal likelihood. For both BFD and BFD*, convergence was assessed by running the program multiple times and taking the mean of estimated marginal likelihoods, and effective sample size (ESS) estimates were also checked to ensure that the majority were >200. After averaging marginal likelihoods across three independent runs, I assessed the strength of support of alternative species delimitation models following the scale of Kass and Raftery

(1995). A positive BF statistic ($2\times\log_e$) favors model 1, whereas a negative BF value favors model 2. The scale is as follows: $0<2\times\ln(BF)<2$ is not worth mentioning, $2<2\times\ln(BF)<6$ is weak but positive evidence, $6<2\times\ln(BF)<10$ is strong support, and $2\times \ln(BF)>10$ is decisive support.

III. RESULTS

(1) Raw data characteristics

Statistics obtained from the Sanger sequence data are presented in Table 2.2. The combined length of all loci was 5,214 bp (mean 521.4 bp) containing 172 segregating sites (mean of 17.2 per locus) and 220 total haplotypes (mean 22 per locus). *Uma cowlesi*, *U. rufopunctata*, and *U. notata* had comparably high mean nucleotide diversity (0.47%, 0.46% and 0.45%, respectively) while *U. scoparia* and *U.* sp. nov. had the lowest (0.09% and 0.12%, respectively). *Uma inornata*, despite its small range and endangered status, had intermediate diversity (0.26%). Summary statistics for the RAD data obtained from the pyRAD output files are presented in Table 2.3. I obtained 56.2 million 100-bp reads for an average of 0.88 million reads per lizard after demultiplexing by the Illumina primer index and barcoded adapters. 51.8 million of these reads passed the trimming and quality filters and were aligned with a clustering threshold of 0.85 into a mean of 29,055 clusters per lizard with an average coverage (depth) of 26x. 3,961 clusters per lizard on average passed the minimum coverage threshold of 10x, and those clusters that passed had a mean depth of 161x. A mean of 3,472 loci (88%) per

lizard passed the paralog filter after estimating the heterozygosity and error rates and clustering loci across individuals with 0.85 similarity. The average number of polymorphic (heterozygous) sites per lizard was 821 (0.28% of total sites). As with the Sanger data, *U. cowlesi*, *U. rufopunctata*, and *U. notata* had a comparably high percentage of heterozygous sites (0.41%, 0.43%, and 0.35%, respectively), while *U. scoparia* had the lowest (0.15%), and *U.* sp. nov. (0.22%) and *U. inornata* (0.25%) were intermediate (Table 2.3, Figure 2.3). I obtained a final alignment with 1,614 loci present in a minimum of 40 of 64 individuals, resulting in a 149,784 base-pair (bp) matrix containing 7,752 SNPs, and a second alignment with 616 loci and 2,805 SNPs represented by at least 60 of 64 individuals, which I further pruned to 581 unlinked biallelic SNPs.

(2) Latitude vs. heterozygosity

There was a significant inverse relationship between latitude and the proportion of polymorphic sites (adjusted R-squared = 0.73, $p < 2.2e-16$, Figure 2.3), and this was still significant after correcting for phylogeny under the Brownian motion model using PGLS ($t = -3.642401$, $p = 6e-04$, 62 residual degrees of freedom). Thus the null hypothesis (that there is no effect of latitude on heterozygosity) can be rejected.

(3) Concatenated phylogenetic analyses

The ML estimate of the phylogeny of the *U. notata* complex and *U. scoparia* based on RADseq data is shown in Figure 2.4. This tree and associated map

clearly show that the six geographically distinct OTUs as shown in Figure 2.1 are supported as exclusive with >70% bootstrap support, with the exception of *U. rufopunctata*, which is non-exclusive with respect to *U. notata* and *U. inornata*. Bootstrap support values within these six OTUs are mostly low, likely indicating high gene flow at the population level, but strongly supported clades were revealed within *U. scoparia* (individuals from the northwestern Mojave Desert) and within *U. notata* (strong support for a clade in Baja California at Laguna Salada). The ML results are mostly corroborated by the Bayesian time tree analyses (Figure 2.5), which also found strong support for the exclusivity of five of six OTUs (>0.95 posterior probability), with the exception of *U. rufopunctat*a, which was again rendered as a non-exclusive assemblage. The topology of the maximum clade credibility tree from BEAST only differs from ML tree at the shallowest (most poorly supported) nodes within OTUs. The *U. notata* complex was estimated to have diverged from *U. scoparia* near the Pliocene/Pleistocene boundary after the formation of the modern Salton Trough and Colorado River delta, and before the onset of Pleistocene glacial cycles, although the 95% interval extends into the early Pleistocene and late Pliocene (1.83-3.45 mya), while divergence (speciation) within the *U. notata* complex was thus estimated to occur primarily during the mid to late Pleistocene, with the exception of the early divergence of *U.* sp. nov. from the remainder of the complex, which may have occurred as early as 2.97 mya (Figure 2.5).

(4) Population structure

The results of the Admixture analysis of 581 unlinked SNPs pruned from RAD loci are shown in Figure 2.5, and the cross validation error (CVE) values used to determine the optimal K value are shown in Table 2.4. K=6 was the preferred model (CVE=0.264). Under K=6, *U. scoparia* was split into two clusters that correspond with the northwestern and southeastern populations in the concatenated phylogenies, and *U. inornata*, *U. cowlesi*, and *U.* sp. nov., all monophyletic groups in the concatenated trees, were also recognized as discrete clusters. All individuals of *U. notata* and *U. cowlesi* were assigned to their respective populations with high confidence, while *U. rufopunctata* was shown to be composed of admixed individuals from *U. notata* and *U. cowlesi*. Pair-wise F_{ST} values between clusters ranged from 0.21 between northwestern and southeastern populations of *U. scoparia* to 0.78 between *U. scoparia* and *U*. sp. nov. (Table 2.5).

The results of the Geneland analysis of the Sanger data under K_{max}=5 are presented in Figure 2.5. The first cluster 1) consisted of *U. scoparia* (not shown), while the remaining four clusters consisted of 2) *U. inornata*, 3) *U. notata* + *U. rufopunctata*, 4) *U. cowlesi*, and 5) *U.* sp. nov. Results for other values of K_{max} are not shown as figures, but are described here as follows: under K_{max}=4, *U. cowlesi* clustered with *U. notata* + *U. rufopunctata*, while under K_{max}=3, *U. inornata* was lumped in as well (Figure 2.2). Only under K_{max}=2 was *U.* sp. nov. finally assigned to the rest of the *U. notata* complex. Under K_{max}=6, the program was no longer able to assign individuals to populations with high confidence (PP > 90).

The results of the smartPCA analysis are shown in Figure 2.7. The first principal component axis (PC1) separates *U. scoparia* from the *U. notata* complex, while the second PC corresponds with a northwest/southeast axis, separating *U. inornata* and *U.* sp. nov. as distinct groups, but *U. rufopunctata* and *U. cowlesi* are not clearly distinguishable in this plot. However, PC3 and PC4 were able to distinguish these groups. The results of the DAPC are shown in Figure 2.8, and the BIC scores used to rank alternative K values are shown in Table 2.4. Increasing K values higher than 6 provided negligible improvements in BIC, so the results under K=6 are presented. Under this model all individuals are assigned to their respective populations under the most complex model shown in Figures 2.1 and 2.2, except one individual (ADG010), geographically in proximity to other individuals of *U. cowlesi*, but which grouped with *U. rufopunctata*.

(5) Species tree analyses

The estimated species trees (maximum clade credibility consensus trees) under the most complex model H (Figure 2.2) are shown in Figure 2.9. Both the *BEAST analyses of Sanger data and the SNAPP analysis of the genome-wide SNP dataset found strong support for the monophyly of the *U. notata* complex with respect to *U. scoparia*, corroborating the concatenated analyses presented here (Figures 2.4, 2.5) and previously published hypotheses (Norris 1958, de Queiroz 1989). However, the *BEAST analyses of Sanger data had much weaker power to resolve the relationships, as is evident in Figure 2.9. By contrast, the

SNAPP analysis of RAD data estimated relationships with a high degree of confidence, as shown by the 1.0 PP values on all nodes. The topology of the SNAPP tree was largely consistent with the concatenated trees, except that the species tree has no way of demonstrating the non-exclusivity of *U. rufopunctata* with respect to *U. notata* and *U. inornata*. There is an interesting biogeographic pattern reflected in the SNAPP tree that mirrors the concatenated trees: southern and eastern populations are more basal, and northwestern more derived, such that *U.* sp. nov. split off first, and *U. inornata* is most nested.

(6) Species delimitation with Bayes factors

The results of the Bayes factor species delimitation analyses (BFD and BFD*) are shown in Table 2.6. In all cases, the most complex Model H (6 sp.) was decisively supported over the next best model (BF > 10), and in all cases, the simplest model A (2 sp.) was ranked the lowest. However, the methods differed in their relative rankings of the intermediate models, although in all cases models with more species were ranked higher. The results of this test demonstrate that the *Uma* sampled in this study are best modeled as six evolutionary distinct lineages or species, including *U. scoparia* and five species within the *U. notata* complex.

IV. DISCUSSION

In this chapter, I addressed four major questions (listed at the end of the introduction) related to the phylogenetic systematics of the *U. notata* complex. Using a combination of Sanger and RAD data and thorough geographic sampling, 1) I described patterns of genetic diversity within this genus, 2) assessed the phylogenetic relationships among lineages using concatenated and coalescent species tree methods, 3) delimited the numbers and boundaries of species with Bayes factors, and 4) inferred how Pliocene and Pleistocene geological and climatic events influenced the historical biogeography of this complex. Here, I discuss each of these four themes individually in detail before I discuss the likely conservation implications of this work.

(1) Patterns of genetic diversity

Earlier work by Gottscho et al. (2014) found that *U. scoparia* had low genetic diversity (nucleotide diversity) relative to the *U. notata* complex. However, this study had relatively poor sampling of the *U. notata* complex as a whole, lacking samples from Mexico, and only including one sample each of *U. inornata* and *U.* sp. nov. In this study, I expanded the Sanger dataset of Gottscho et al. (2014) and supplemented it with a RAD dataset to provide increased resolution to examine patterns of genetic diversity. Both measures of nucleotide diversity with the Sanger data (Table 2.2), and percent heterozygosity with the RAD data (Table 2.3, Figure 2.3) show that *U. scoparia*, the northernmost species,

is the least diverse, while *U. cowlesi*, *U. rufopunctata*, and *U. notata* are most diverse, and *U. inornata* and *U.* sp. nov. are intermediate. Gottscho et al. (2014) hypothesized that founder effects and/or genetic bottlenecks related to Pleistocene glacial events and subsequent northwestern range expansions, as originally hypothesized by Norris (1958), could be responsible for the low genetic diversity observed in *U. scoparia* (Excoffier et al. 2009). Within the *U. notata* complex, it is important to note that the species with the highest diversity, *U. rufopunctata*, not only occupies the largest sand dune complex in North America, but also shows evidence of admixture between *U. notata* and *U. cowlesi* (Figure 2.5). *U. inornata* and *U.* sp. nov., which consist of small, isolated populations on the periphery of the range of the complex, have lower diversity levels.

(2) Phylogeny

Both ML and Bayesian analyses of concatenated RAD data resolved the following features in common with each other and the published mtDNA tree of Trepanier and Murphy (2001): 1) *U. scoparia* is monophyletic and sister to the *U. notata* complex, which is also strongly supported as monophyletic, 2) individuals from Mohawk Dunes (*U.* sp. nov.) are supported as monophyletic and are sister to the remainder of the complex, 3) individuals from the Gran Desierto/Yuma Dunes (*U. rufopunctata*) are paraphyletic with respect to *U. notata* and *U. inornata* west of the Colorado River, which are each monophyletic sister lineages. However, in the mtDNA tree presented by Trepanier and Murphy (2001), *U.*

cowlesi is not revealed as a monophyletic lineage, whereas it is strongly supported as monophyletic by my concatenated RAD trees. The topology of the species trees inferred using different data and methods of analysis (ten nuclear loci and *BEAST vs. 468 unlinked biallelic SNPS and SNAPP) were not congruent with each other (Figure 2.9). The SNAPP tree was strongly supported and consistent with the mtDNA and concatenated trees, showing a hierarchal nested pattern wherein the southeastern species are most basal, starting with *U.* sp. nov., while the northwestern species (*U. inornata*) is most derived. By contrast, the species tree inferred using ten nuclear loci with *BEAST did not support this hierarchal relationship (Figure 2.9), but because the relationships in this tree were not strongly supported, I favor the SNAPP tree as a more reliable result. These results demonstrate the power of next-generation sequencing approaches to obtain enough data to resolve phylogenetic relationships among closely related species with high confidence, a task which previously proved daunting with Sanger data alone.

(3) Species delimitation

The results of this study, using three alternative methods of Bayes factor delimitation, provided decisive support for the six-species model, including *U. scoparia* and five species within the *U. notata* complex (Table 2.6). *Uma inornata*, *U. rufopunctata*, and *U. notata* are currently recognized as valid taxa (Crothers 2012), while the other two are not. Heifetz (1941) provided an adequate description and type series of *U. n. cowlesi*, and I simply propose to elevate this

sunken subspecies to full species status – no new description is needed. While recognizing *U. cowlesi* would leave *U. rufopunctata* as a paraphyletic group as depicted in a concatenated analysis of individuals, there is no way to avoid this outcome without lumping the entire *U. notata* complex into one or two species, but these models were the most poorly supported by BFD and BFD* (Table 2.6). Thus, the recognition of a paraphyletic assemblage as a distinct species appears warranted in this case in order to fully recognize the lineage diversity within this group. While the description of the new *Uma* species from the Mohawk Dunes is beyond the scope of this study, from a biodiversity conservation standpoint it is critically important that it be described. Thus, I intend to prepare an additional manuscript later this year to describe this new species. This manuscript will integrate morphological data from museum specimens, establish a type series with quality tissue samples, and establish a new morphological identification key for the genus.

(4) Historical biogeography and phylogeography

The deserts of North America, despite their apparent desolation, represent one of only five global high-biodiversity wilderness areas (Mittermeier et al. 2003) and thus represent a top priority for global biodiversity conservation efforts. In order to protect this rich diversity, an understanding of the evolutionary mechanisms that generated it is essential. A major debate in the historical biogeography of the American southwest revolves around how complex Miocene and Pliocene geological events, including the evolution of the

San Andreas Fault system, the Baja California Peninsula, the Gulf of California, and the delta of the Colorado River, influenced allopatric speciation of desert flora and fauna (Mulcahy et al. 2006, Wood et al. 2013, Gottscho et al. 2014, Gottscho 2014). Many co-distributed taxa show spatially concordant phylogeographic breaks across features associated with these landscape features, implying a common underlying process that shaped this pattern (Wood et al. 2013, Gottscho 2014). These large-scale landscape changes may have driven speciation directly by fragmenting species distributions that were formerly continuous, also known as vicariance; alternatively, dispersal across a matrix previously shaped by large-scale landscape events, especially during Pleistocene glacial cycles, may have been the major driver of speciation and differentiation in the southwestern aridlands (Wood et al. 2013). The results of this study provide some evidence supporting both of these roles of landscape change in promoting speciation — as we shall see, the influences of plate tectonics and climate change are not so easily disentangled.

Prior to the discovery of plate tectonics, Norris (1958) hypothesized that *U. scoparia* diverged from the *U. notata* complex during the late Pliocene or early Pleistocene when a fraction of the ancestral species dispersed northward from the Gran Desierto de Altar along the Colorado River gorge into the southeastern Mojave Desert. After speciation, Norris hypothesized that glacial maxima confined *U. scoparia* to the lower, warmer southeastern portion of its range near the Colorado River before it colonized northwestern localities in the Mojave and Amargosa River drainages after the last glacial maximum. Gottscho et al. (2014),

using fourteen nuclear loci in an isolation-with-migration model, estimated the divergence date between *U. scoparia* and the *U. notata* complex to be in the mid-Pleistocene (95% confidence interval of 0.60 – 1.34 mya), whereas the concatenated analyses presented here suggest this divergence date to be older (95% confidence interval of 1.83 –3.45 mya), directly supporting the hypothesis of Norris (1958) over that of Gottscho et al. (2014). Although these studies used very similar estimates of mutation rate, it is important to note that concatenated methods may overestimate divergence dates relative to coalescent methods (Heled and Drummond 2010), so caution should be used while comparing these results directly; all else being equal, we should expect the concatenated method to produce older estimates. Norris (1958) further hypothesized that *U. scoparia* in the Mojave and Amargosa Rivers, and *U. inornata* in the Coachella Valley were both the results of northwestern range expansions in the late Pleistocene from ancestral stock to the southeast along the Colorado River. The topology and timing of the clock tree (Figure 2.5) strongly supports both of these hypotheses. Within *U. scoparia*, it is interesting to note that individuals from Pinto Basin (Joshua Tree National Park) and Dale Lake are nested deep within the Mojave River clade, sister to an individual from Barstow along the Mojave River, indicating that these populations likely descended from individuals that disperse back southeast from the Mojave River, following a sand transport corridor through Twenty-nine Palms.

Norris (1958) proposed his biogeographic reconstruction of the genus prior to the discovery of plate tectonics. We now know that the Colorado Desert

region is bisected by an active boundary between the Pacific and North American plates, with most of the relevant geological features having been formed in the Pliocene through the Oligocene (Stock and Hodges 1989). The San Andreas Fault system developed in southern California as early as 20 mya (Gottscho 2014), but it wasn't until 8-12 mya that a basin-and-range style proto-Gulf province began to form, and the Baja California Peninsula and Gulf of California were formed ~6 mya as strike-slip faulting commenced in the Gulf of California, including the Salton Trough (Elders et al. 1972, Stock and Hodges 1989). During the late Miocene and Pliocene, marine incursions extended the Gulf of California through much of the range currently occupied by the *U. notata* complex as far north as Whitewater at the northwestern end of Coachella Valley, as revealed by the Imperial formation (Smith 1991, McDougall et al. 1999, Oskin and Stock 2003). As early as 5.6 mya, the Colorado River began filling the north end of the Salton Trough with sediment eroded from the Grand Canyon and Colorado Plateau, and it reached the Yuma area by ~4 mya (Brusca 2014).

During the Pleistocene, lower sea levels and sediment deposited by the Colorado River conspired to expose the northern Salton Trough above sea level. During glacial maxima, ample precipitation produced perennial lakes and streams that eroded sediment from the mountains. During dry, warm interglacial periods, these freshwater sources evaporated, exposing sediments to wind and creating aeolian deposits, including the Gran Desierto de Altar (Blount and Lancaster 19990, Muhs et al. 2003). An additional complication is that the Colorado River itself was not stable throughout the Pleistocene, frequently

shifting its course among multiple drainage beds and occasionally draining into the prehistoric Lake Cahuilla in the center of the Salton Trough, currently occupied by the Salton Sea (Waters 1983). In addition to its role shaping the topography of the basin through sea level changes and dune formation, the direct influence of climate change itself likely influenced the distribution of *Uma* and other arid-adapted flora and fauna. Although the North American deserts were not directly impacted by ice sheets during the Pleistocene, pollen extracted from fossilized packrat (*Neotoma*) middens show that during the last glacial maximum (LGM) 18 kya, the Mojave Desert was dominated by a semi-arid coniferous woodland, confining arid-adapted flora to refugia at lower elevations and latitudes, primarily in the Sonoran and Colorado Deserts (Cole 1986). The modern desert flora later expanded northward to reach their present distributions in the Mojave Desert as recently as 6 kya in the Holocene epoch (Thompson and Anderson 2000). Presumably, similar fluctuations occurred throughout the glacial/interglacial cycles of the Pleistocene. However, a recent analysis of a Colorado Desert packrat midden series from the base of Guadalupe Canyon in Baja California (Holmgren et al. 2014) challenged this interpretation, finding that a community of chaparral, woodland and select desert flora inhabited the site during the last glacial maximum, refuting the hypothesis that the lower Colorado River basin served as a late Pleistocene refuge for arid-adapted flora. The rapid arrival of most missing desert species by the early Holocene suggests that these taxa probably survived the last glacial maximum as smaller, disparate populations in arid microhabitats within chaparral and

pinyon–juniper–oak woodlands. Either way, it is apparent that the climatic cycles of the Pleistocene exerted strong influences of the distributions of flora and fauna in the North American deserts, as is apparent in the strong latitudinal gradient of genetic diversity (Figure 2.3) and the Pleistocene divergence dates observed in the clock tree presented here (Figure 2.5).

The origin of the San Andreas Fault system in this region in the early Miocene (Gottscho 2014) and the development of the Colorado River delta in the late Miocene (Brusca 2014) are both older than the late Pliocene/Pleistocene divergences observed in this study. Thus, it appears that the primary role of geology/plate tectonics was to create the landscape matrix, and Pleistocene glacial cycles likely promoted repeated waves of dispersal and range contraction across this landscape, producing the young divergence dates and strong latitudinal gradient in genetic diversity that we observe today. However, one important exception concerns the role of the Colorado River in shaping diversity within this complex during the late Pleistocene. During most of the Pleistocene, the Colorado River delta was positioned further east of its eastern location, discharging into the Gulf of California in the vicinity of modern Bahia Adair (Figure 2.1). The fluvial sediments of this ancient paleodelta, partially cloaked with a mantle of aeolian sand, are still readily visible on satellite imagery (www.google.com/earth). The combination of the river delta positioned at the base of the El Pinacate volcanic crater, which began to form as early as 1.7 mya and erupted as recently as 8 kya (Lynch et al. 1993, Brusca 2014), was likely a sufficient barrier to allow for the divergence of *U. cowlesi* from lineages west of

the river. However, about 120 kya, as the Salton Trough continued to expand and deepen due to rifting along the San Andreas Fault system, the delta shifted westward to its current location west of the Gran Desierto (Blount and Lancaster 1990). This would mean that populations of *U. rufopunctata* would have no longer been isolated directly to the east from *U. cowlesi*, but would then be cut off from the west, parsimoniously explaining why *U. rufopunctata* is more closely related to *U. notata* than to *U. cowlesi* in all of the phylogenetic analyses presented here, why *U. rufopunctata* groups with *U. notata* in the Geneland analysis, as well as why this species appears to be composed of admixed populations of *U. notata* and *U. cowlesi* in the Admixture analysis under K=6 (Figure 2.5). Likely, *U. rufopunctata* had been exchanging genes first with *U. notata* prior to 120 kya, then possibly with *U. cowlesi* subsequent to the shift of the river. Thus, in this case it appears that landscape change associated with the continued evolution of the San Andreas Fault system and Colorado River directly contributed to the species richness that we observe in the region today, explaining why we see a break between two species in the vicinity of the paleodelta. Future studies could directly test this and other complex demographic hypotheses by modeling the allele frequency spectrum and/or utilizing approximate Bayesian computation (e.g. Rittmeyer and Austin 2015).

(5) Conservation implications

Earlier, I argued that systematic biology has much to offer applied biodiversity conservation efforts, particularly through the development of

objective species delimitation models. As outlined above, the results of this study support the need to elevate *U. cowlesi* to species status and to formally describe the new species from Mohawk Dunes. Because intraspecific levels of genetic diversity are important for future adaptation to changing environmental conditions (Soulé and Mills 1998, Frankham et al. 2002), these should be considered in conservation schemes as well (Vandergast et al. 2013). Here, I summarize what I view as the most important conservation implications to stem from this study, examining each of the six species separately, as they should be managed. Most of these implications somehow relate to the fact that three species are endemic to specific countries and/or states, while one species crosses state borders and two others cross international borders, so coordination among a multitude of government agencies will be crucial to manage the total species diversity within this clade. This summary is not meant to be exhaustive, but rather meant to outline areas where further research and collaboration is needed.

Gottscho et al. (2014) designated *Uma scoparia* as a single Evolutionarily Significant Unit (ESU) for conservation purposes. In this chapter, utilizing the increased resolving power of hundreds of unlinked SNPS, I detected novel population structure between southeastern populations of *U. scoparia* near the Colorado River (with high genetic diversity) and northwestern populations in the Mojave and Amargosa river drainages, including the populations by Dale Lake and Pinto Basin in Joshua Tree National Park (NP). This latter population has the lowest levels of genetic diversity measured in this study, even lower than the small, isolated populations of *U. inornata* and *U.* sp. nov., which may make

this species vulnerable to changing environment conditions. A large portion of this species range lies on BLM lands within the DRECP plan area, but also large chunks of habitat are protected in Death Valley NP, Mojave NP, Joshua Tree NP, and various BLM wilderness areas. Due to its low genetic diversity and potential impacts from expanded renewable energy development and climate change, I recommend that *U. scoparia* retain its status in California as a Species of Special Concern. Because the population in Arizona appears have especially high diversity, I recommend that land managers in Arizona monitor this population to ensure its persistence within the state.

Uma inornata has long been the flagship species for conservation of this genus. The results of this study strongly support the distinctiveness of this species, which was strongly supported as monophyletic in my concatenated analyses, identified as a discrete cluster by all clustering algorithms used in this study, and decisively supported as an independent lineage by BFD and BFD*. Thus, I recommend it continue to be managed as a threatened/endangered species. *Uma inornata* has moderate levels of genetic diversity, which is unexpected given its small, fragmented range and prior documentation of low genetic diversity (Hedtke et al. 2007). However, there is some evidence that this genetic diversity may need to be protected from attrition due to habitat fragmentation — population genetic (microsatellite) data taken at different time series indicate that loss of diversity has occurred since this species was listed (Wood et al., in prep). One action agencies could take to reduce the gradual loss of genetic diversity over the upcoming decades, beyond the top priority of

preventing further development and degradation of the remaining habitat, is to increase gene flow levels within the species by trans-locating individuals among adjacent dunes.

Uma notata as defined here is endemic to the Colorado Desert of California and Baja California west of the Colorado River. Potential conservation issues include the recently-completed international border fence (which necessitates separate management of the Laguna Salada population in Baja California), loss of ancestral habitat due to agricultural and urban development in the Imperial Valley, heavy off-road vehicle recreation impacts in California (including sites at Borrego Springs, Ocotillo Wells, Truckhaven, Superstition Mountain, and Algodones), and habitat degradation and fragmentation (pers. obs.). Fortunately, the relatively high genetic diversity of *U. notata* should make it more adaptable in the face of climate change, which is important because there is little elevational or latitudinal variation within the range of this species for it to escape warmer and drier conditions. Due to the heavy anthropogenic impacts to its range, I recommend that *U. notata* retain its status as a California Species of Special Concern, and should be monitored closely to ensure that it does not become threatened with extinction due to prolonged degradation of its habitat. The recent opening of the North Algodones Dunes Wilderness to off-road vehicle recreation is an especially grave concern as this area consists of the largest pristine section of dunes remaining for this species in California.

Uma rufopunctata has the highest genetic diversity as measured by RAD data of any species examined in this chapter. Its habitat, the largest chunk of

sand dunes in North America, is largely intact, protected in a biosphere reserve with very little development. Thus, this species is likely of the least conservation concern out of the six. However, it is worth noting that the new border fence may have cut off gene flow from Arizona populations in the vicinity of Yuma, and the damming of the Colorado River almost certainly has cut off the influx of new aeolian sand to the Yuma dunes. To maintain the high genetic diversity of these lizards in Arizona, allowing gene flow across the Arizona/Sonora border may be important, and quality of sand dune habit in Arizona should be monitored.

Uma cowlesi, validated by my analyses as a distinct species from *U. rufopunctata*, is endemic to the state of Sonora, Mexico (Heifetz 1941), and should be managed by the relevant local, state, and federal authorities as such. From examination of satellite maps, the majority of the range of this species appears to be relatively intact, although urban and agricultural encroachment is definitely a potential future issue. Unlike *U. rufopunctata*, most of the range of this species lies outside of the boundary of the local biosphere reserves to the northwest and thus could be vulnerable to expanded agricultural and urban development on the coastal plain.

Uma sp. nov., the most phylogenetically distinct species within the *U. notata* complex, is almost entirely restricted to the Barry M. Goldwater Air Force Range in the Sonoran Desert of Arizona, although the northernmost portion of its range extends into BLM lands near Interstate 8 (where the samples in this study were collected). The small range of this species and its location on a bombing and gunnery range make it vulnerable to military training activity. However, the U.S.

Air Force already has extensive experience dealing with other sensitive species such as pronghorn antelope on this range (Krausman et al. 2005), so continued responsible range management is essential to ensure that this species remains merely vulnerable and not threatened or endangered with extinction. I recommend that range managers should re-evaluate their land use priorities in light of having a new *Uma* species endemic to the Mohawk Dunes, one that is even more distinct phylogenetically than *U. inornata*, the current focus of conservation efforts.

(6) Summary and conclusions

In this chapter I utilized recent advances in next generation sequencing and coalescent-based species delimitation to study fringe-toed lizards of the *Uma notata* species complex, a group of special conservation concern with a long and unstable taxonomic history. The strong latitudinal gradient in genetic diversity observed across species is likely related to climatic cycles during the Pleistocene, whereby northwestern populations of *U. scoparia* and *U. inornata* are likely recent (late Pleistocene) range expansions. The Colorado River played an all-important role in the diversification of this group, providing much of the sand for its habitat, subdividing populations, and providing dispersal corridors across mountains paralleling the San Andreas Fault. I found that a model consisting of six species within this group (including the outgroup, *U. scoparia*) was decisively supported over all alternative models, including all previously published hypotheses and the current taxonomy. This study should have important

conservation implications because *U. cowlesi* should be elevated to full species status and *U*. sp. nov. from the Mohawk Dunes should be described as soon as possible; both of these new species are endemic to their respective states/countries and are vulnerable to anthropogenic impacts. This study highlights the continued importance of basic systematic biology for applied biodiversity conservation efforts.

V. REFERENCES

Adams M, Raadik TA, Burridge CP, Georges A. 2014. Global biodiversity assessment and hyper-cryptic species complexes: more than one species of elephant in the room? Systematic Biology, 63:518–533.

Adest GA. 1977. Genetic relationships in genus *Uma* (Iguanidae). Copeia:47–52.

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 19:1655–1664.

Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko AV. 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Molecular Biology and Evolution, 29:2157–2167.

Baird SF. 1858. Description of new genera and species of North American lizards in the museum of the Smithsonian Institution. Proceedings of the Academy of Natural Sciences of Philadelphia, 1858:253–256.

Blount G, Lancaster N. 1990. Development of the Gran Desierto sand sea, northwestern Mexico. Geology, 18:724–728.

Bouckaert RR. 2010. DensiTree: making sense of sets of phylogenetic trees. Bioinformatics, 26:1372–1373.

Bouckaert RR, Heled J, Kuehnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. PLoS Computational Biology, 10:e1003537.

Brooks TM, Mittermeier RA, da Fonseca GAB, Gerlach J, Hoffmann M, Lamoreux JF, Mittermeier CG, Pilgrim JD, Rodrigues ASL. 2006. Global biodiversity conservation priorities. Science, 313:58–61.

Brusca RC. 2014. A brief historical geology of northwestern Mexico. Unpublished manuscript, v29, p. 1–67.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene

trees in a full coalescent analysis. Molecular Biology and Evolution, 29:1917–1932.

Buising AV. 1990. The Bouse formation and bracketing units, southeastern California and western Arizona – implications for the evolution of the proto-gulf of California and the Colorado River. Journal of Geophysical Research-Solid Earth and Planets, 95:20111–20132.

Camp CL. 1916. Notes on the local distribution and habits of the amphibians and reptiles of southeastern California in the vicinity of the Turtle Mountains. University of California Publications in Zoology, 12:503–544.

Carothers JH. 1986. An experimental confirmation of morphological adaptation – toe fringes in the sand-dwelling lizard *Uma scoparia*. Evolution, 40:871–874.

Carstens BC, Pelletier TA, Reid NM, Satler JD. 2013. How to fail at species delimitation. Molecular Ecology, 22:4369–4383.

Cayan D, Tyree M, Dettinger M, Hidalgo H, Das T, Maurer E, Bromirski P, Graham N, Flick R. 2009. Climate change scenarios and sea level rise estimates for the California 2008 climate change scenarios assessment. California Climate Change Center, Sacramento, California, p. 1–62.

Chen X, Barrows CW, Li B-L. 2006. Is the Coachella Valley Fringe-toed Lizard (*Uma inornata*) on the edge of extinction at Thousand Palms Preserve in California? Southwestern Naturalist, 51:28–34.

Cole KL. 1986. The lower Colorado River Valley: a Pleistocene desert. Quaternary Research, 25:392–400.

Conroy CJ, Papenfuss T, Parker J, Hahn NE. 2009. Use of tricaine methanesulfonate (MS222) for euthanasia of reptiles. Journal of the American Association for Laboratory Animal Science, 48:28.

Cope ED. 1866. On Reptilia and Batrachia of the Sonoran Province of the Nearctic Region. Proceedings of the Academy of Natural Sciences of Philadelphia, 18:300–314.

Cope ED. 1894. On the Iguanian genus *Uma*. American Naturalist 28:434–435.

Cope ED. 1895. On the species of *Uma* and *Xantusia*. American Naturalist 29: 938–939.

Cope ED. 1896. On the genus *Callisaurus*. Ibid, 30:1049–1050.

Cope ED. 1898 (1900). The crocodilians, lizards, and snakes of North America. Annual Report of the United States National Museum,1898:155–1294.

Cracraft J. 2002. The seven great questions of systematic biology: An essential foundation for conservation the sustainable use of biodiversity. Annals of the Missouri Botanical Garden, 89:127–144.

Crother BI. 2012. Standard Common and Current Scientific Names for North American Amphibians, Turtles, Reptiles, and Crocodilians, Seventh Edition. Herpetological Circular 39:1–92.

de Queiroz A, Gatesy J. 2007. The supermatrix approach to systematics. Trends in Ecology and Evolution, 22:34–41.

de Queiroz K. 1989. Morphological and biochemical evolution in the sand lizards. PhD Dissertation, University of California, Berkeley.

de Queiroz K. 1992. Phylogenetic relationships and rates of allozyme evolution among the lineages of sceloporine sand lizards. Biological Journal of the Linnean Society, 45:333–362.

de Queiroz K. 1998. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. In: Howard DJ, Berlocher SH editors. Endless Forms: Species and Speciation. Oxford University Press, Oxford, p. 57–75.

de Queiroz K. 2007. Species concepts and species delimitation. Systematic Biology, 56:879–886.

de Queiroz K. 2011. Branches in the lines of descent: Charles Darwin and the evolution of the species concept. Biological Journal of the Linnean Society, 103:19–35.

Dimmick WW, Ghedotti MJ, Grose MJ, Maglia AM, Meinhardt DJ, Pennock DS. 1999. The importance of systematic biology in defining units of conservation. Conservation Biology, 13:653–660.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution, 29:1969–1973.

Eaton DAR. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics, 30:1844–1849.

Eaton DAR, Ree RH. 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). Systematic Biology, 62:689–706.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32:1792–1797.

Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics, 26:2460–2461.

Elders WA, Biehler S, Rex RW, Robinson PT, Meidav T. 1972. Crustal spreading in southern California. Science, 178:15–24.

Excoffier L, Foll M, Petit RJ. 2009. Genetic consequences of range expansions. Annual Review of Ecology Evolution and Systematics, 40:481–501.

Faith DP. 1992. Conservation evaluation and phylogenetic diversity. Biological Conservation, 61:1–10.

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics, 164:1567–1587.

Flot JF. 2010. SEQPHASE: a web tool for interconverting phase input/output files and fasta sequence alignments. Molecular Ecology Resources, 10:162–166.

Frankel OH, Soulé ME. 1981. Conservation and Evolution. Cambridge University Press, Cambridge.

Frankham R, Ballou JD, Briscoe DA. 2002. Introduction to Conservation Genetics. Cambridge, pp. 227–253, 336–358.

Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C. 2012. Coalescent-based species delimitation in an integrative taxonomy. Trends in Ecology and Evolution, 27:480–488.

Funk VA, Sakai AK, Richardson K. 2002. Biodiversity: The interface between systematics and conservation. Systematic Biology, 51:235–237.

Gottscho AD, Marks SB, Jennings WB. 2014. Speciation, population structure, and demographic history of the Mojave Fringe-toed Lizard (*Uma scoparia*), a species of conservation concern. Ecology and Evolution, 4:2546–2562.

Gottscho AD. 2014. Zoogeography of the San Andreas Fault system: Great Pacific Fracture Zones correspond with spatially concordant phylogeographic boundaries in western North America. Biological Reviews, doi: 10.1111/brv.12167.

Grummer JA, Bryson RW, Jr., Reeder TW. 2014. Species delimitation using Bayes Factors: simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). Systematic Biology, 63:119–133.

Guillot G. 2008. Inference of structure in subdivided populations at low levels of genetic differentiation-the correlated allele frequencies model revisited. Bioinformatics, 24:2222–2228.

Guillot G, Mortier F, Estoup A. 2005. GENELAND: a computer package for landscape genetics. Molecular Ecology Notes, 5:712–715.

Guillot G, Santos F, Estoup A. 2008. Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. Bioinformatics, 24:1406–1407.

Hedtke SM, Zamudio KR, Phillips CA, Losos J, Brylski P. 2007. Conservation genetics of the endangered Coachella Valley Fringe-toed Lizard (*Uma inornata*). Herpetologica, 63:411–420.

Heifetz W. 1941. A review of the lizards of the genus *Uma*. Copeia, 1941:99–111.

Heled J, Drummond AJ. 2008. Bayesian inference of population size history from multiple loci. BMC Evolutionary Biology, 8:289.

Hey J. 2006. On the failure of modern species concepts. Trends in Ecology and Evolution, 21:447–450.

Holmgren CA, Betancourt JL, Cristina Penalba M, Delgadillo J, Zuravnsky K, Hunter KL, Rylander KA, Weiss JL. 2014. Evidence against a Pleistocene desert refugium in the Lower Colorado River Basin. Journal of Biogeography, 41:1769–1780.

Huang H, Knowles LL. 2014. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. Systematic Biology, doi:10.1093/sysbio/syu046.

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics, 11:94.

Jones LL, Lovich RE. 2009. Lizards of the American Southwest. Rio Nuevo Publishers, Tucson, p. 1–568.

Kass RE, Raftery AE. 1995. Bayes factors. Journal of the American Statistical Association, 90:773–795.

Krausman PR, Harris LK, Haas SK, Koenen KKG, Devers P, Bunting D, Barb M. 2005. Sonoran pronghorn habitat use on landscapes disturbed by military activities. Wildlife Society Bulletin, 33:16–23.

Kumar S, Subramanian S. 2002. Mutation rates in mammalian genomes. Proceedings of the National Academy of Sciences of the United States of America, 99:803–808.

Leaché AD, Fujita MK, Minin VN, Bouckaert RR. 2014. Species delimitation using genome-wide SNP data. Systematic Biology, 63:534–542.

Leaché AD, Chavez AS, Jones LN, Grummer JA, Gottscho AD, Linkem CW. 2015. Phylogenomics of phrynosomatid lizards: conflicting signals from sequence capture versus restriction site associated DNA sequencing. Genome Biology and Evolution, doi:10.1093/gbe/evv026.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25:1451–1452.

Lim GS, Balke M, Meier R. 2012. Determining species boundaries in a world full of rarity: singletons, species delimitation methods. Systematic Biology, 61:165–169.

Luke C. 1986. Convergent evolution of lizard toe fringes. Biological Journal of the Linnean Society, 27:1–16.

Lynch DJ, Musselman TE, Gutman JT, Patchett PJ. 1993. Isotopic evidence for the origin of Cenozoic volcanic rocks in the Pinacate volcanic field, northwestern Mexico. Lithos 29:295–302.

Marchini J, Cardon LR, Phillips MS, Donnelly P. 2004. The effects of human population structure on large genetic association studies. Nature Genetics, 36:512–517.

Mayr E. 1942. Systematics and the Origin of Species. Columbia University Press, New York.

Mayr E. 1963. Animal Species and Evolution. Harvard University Press, Cambridge, Massachusetts.

McDougall K, Poore RZ, Matti J. 1999. Age and paleoenvironment of the Imperial formation near San Gorgonio Pass, southern California. Journal of Foraminiferal Research, 29:4–25.

McLaughlin JF, Hellmann JJ, Boggs CL, Ehrlich PR. 2002. Climate change hastens population extinctions. Proceedings of the National Academy of Sciences of the United States of America, 99:6070–6074.

Mittermeier RA, Mittermeier CG, Brooks TM, Pilgrim JD, Konstant WR, da Fonseca GAB, Kormos C. 2003. Wilderness and biodiversity conservation. Proceedings of the National Academy of Sciences of the United States of America, 100:10309–10313.

Moritz C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology, 51:238–254.

Muhs DR, Reynolds RL, Been J, Skipp G. 2003. Eolian sand transport pathways in the southwestern United States: importance of the Colorado River and local sources. Quaternary International, 104:3–18.

Mulcahy DG, Spaulding AW, Mendelson JR, Brodie ED. 2006. Phylogeography of the flat-tailed horned lizard (*Phrynosoma mcallii*) and systematics of the *P. mcallii-platyrhinos* mtDNA complex. Molecular Ecology, 15:1807–1826.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. Nature, 403:853–858.

Norris KS. 1958. The evolution and systematics of the iguanid genus *Uma* and its relation to the evolution of other North American desert reptiles. Bulletin of the American Museum of Natural History, 114:251–326.

Ogden R, Gharbi K, Mugue N, Martinsohn J, Senn H, Davey JW, Pourkazemi M, McEwing R, Eland C, Vidotto M, et al. 2013. Sturgeon conservation genomics: SNP discovery and validation using RAD sequencing. Molecular Ecology, 22:3112–3123.

O'Meara BC. 2010. New heuristic methods for joint species delimitation and species tree inference. Systematic Biology, 59:59–73.

Oskin M, Stock J. 2003. Marine incursion synchronous with plate-boundary localization in the Gulf of California. Geology, 31:23–26.

Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. PLoS Genetics, 2:2074–2093.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One, 7:e37135.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics, 155:945–959.

Rambaut A, Suchard M, Drummond A. 2013. Tracer v1.6. Available from: http://tree.bio.ed.ac.uk/software/tracer/

Reich D, Price AL, Patterson N. 2008. Principal component analysis of genetic data. Nature Genetics, 40:491–492.

Rheindt FE, Fujita MK, Wilton PR, Edwards SV. 2014. Introgression and phenotypic assimilation in *Zimmerius* flycatchers (Tyrannidae): population genetic and phylogenetic inferences from genome-wide SNPs. Systematic Biology, 63:134–152.

Rittmeyer EN, Austin CC. 2015. Combined next-generation sequencing and morphology reveal fine-scale speciation in Crocodile Skinks (Squamata: Scincidae: *Tribolonotus*). Molecular Ecology 24, 466–483.

Ruane S, Bryson RW, Jr., Pyron RA, Burbrink FT. 2014. Coalescent species delimitation in milksnakes (genus *Lampropeltis*) and impacts on phylogenetic comparative analyses. Systematic Biology, 63:231–250.

Satler JD, Carstens BC, Hedin M. 2013. Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*). Systematic Biology, 62:805–823.

Schulte JA, II, de Queiroz K. 2008. Phylogenetic relationships and heterogeneous evolutionary processes among phrynosomatine sand lizards (Squamata, Iguanidae) revisited. Molecular Phylogenetics and Evolution, 47:700–716.

Shaffer HB. 2013. Evolution and Conservation. In: Losos J editor. The Princeton Guide to Evolution, Princeton University Press, p. 766–773.

Sinervo B, Mendez-de-la-Cruz F, Miles DB, Heulin B, Bastiaans E, Villagran-Santa Cruz M, Lara-Resendiz R, Martinez-Mendez N, Lucia Calderon-Espinosa M, Nelsi Meza-Lazaro R, et al. 2010. Erosion of Lizard Diversity by Climate Change and Altered Thermal Niches. Science, 328:894–899.

Smith JT. 1991. Cenozoic marine mollusks and paleogeography of the Gulf of California. In: Dauphin JT, Simoneit RT editors. The Gulf and Peninsular Province of the Californias, Association of Petroleum Geologists, Memoir 47, p. 447–480.

Soulé ME, Mills LS. 1998. Population genetics. Science, 282:1658–1659.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30:1312–1313.

Stebbins RC. 1944. Some aspects of the ecology of the iguanid genus *Uma*. Ecological Monographs, 14:311–322.

Stebbins RC. 2003. A field guide to western reptiles and amphibians: third edition. Houghton Mifflin Company, Boston, New York.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics, 68:978–989.

Stock JM, Hodges KV. 1989. Pre-Pliocene extension around the Gulf of California and the transfer of Baja California to the Pacific plate. Tectonics, 8:99–115.

Thompson RS, Anderson KH. 2000. Biomes of western North America at 18,000, 6000 and 0 C-14 yr BP reconstructed from pollen and packrat midden data. Journal of Biogeography, 27:555–584.

Trepanier TL, Murphy RW. 2001. The Coachella Valley Fringe-toed Lizard (*Uma inornata*): Genetic diversity and phylogenetic relationships of an endangered species. Molecular Phylogenetics and Evolution, 18:327–334.

[USFWS] U.S. Fish and Wildlife Service. 1980. Endangered and threatened wildlife and plants; listing as threatened with critical habitat for the Coachella Valley Fringe-toed Lizard. Federal Register, 50:63812–63820.

[USFWS] U.S. Fish and Wildlife Service. 2010. Coachella Valley Fringe-toed Lizard (*Uma inornata*): 5-Year review: summary and evaluation. Carlsbad Fish and Wildlife Office, Carlsbad, California, p. 1–53.

Van Denburgh J. 1922. The reptiles of western North America, Volume I: Lizards. Occasional Papers of the California Academy of Sciences, 10:1–612.

Vandergast AG, Inman RD, Barr KR, Nussear K, Esque T, Hathaway SA, Wood DA, Medica PA, Breinholt JW, Stephen CL, Gottscho AD, Marks SB, Jennings WB, Fisher RN. 2013. Evolutionary hotspots in the Mojave Desert. Diversity, 5:293–319.

Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, Seehausen O. 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Molecular Ecology, 22:787–798.

Waters MR. 1983. Late Holocene lacustrine chronology and archaeology of ancient Lake Cahuilla, California. Quaternary Research 19:373–387.

Wiley EO. 1978. Evolutionary species concept reconsidered. Systematic Zoology, 27:17–26.

Wilgenbusch J, de Queiroz K. 2000. Phylogenetic relationships among the phrynosomatid sand lizards inferred from mitochondrial DNA sequences generated by heterogeneous evolutionary processes. Systematic Biology, 49:592–612.

Wood DA, Vandergast AG, Barr KR, Inman RD, Esque TC, Nussear KE, Fisher RN. 2013. Comparative phylogeography reveals deep lineages and regional evolutionary hotspots in the Mojave and Sonoran Deserts. Diversity and Distributions, 19:722–737.

Yang Z, Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. Proceedings of the National Academy of Sciences of the United States of America, 107:9264–9269.

Yang Z, Rannala B. 2014. Unguided species delimitation using DNA sequence data from multiple loci. Molecular Biology and Evolution, 31:3125–3135.

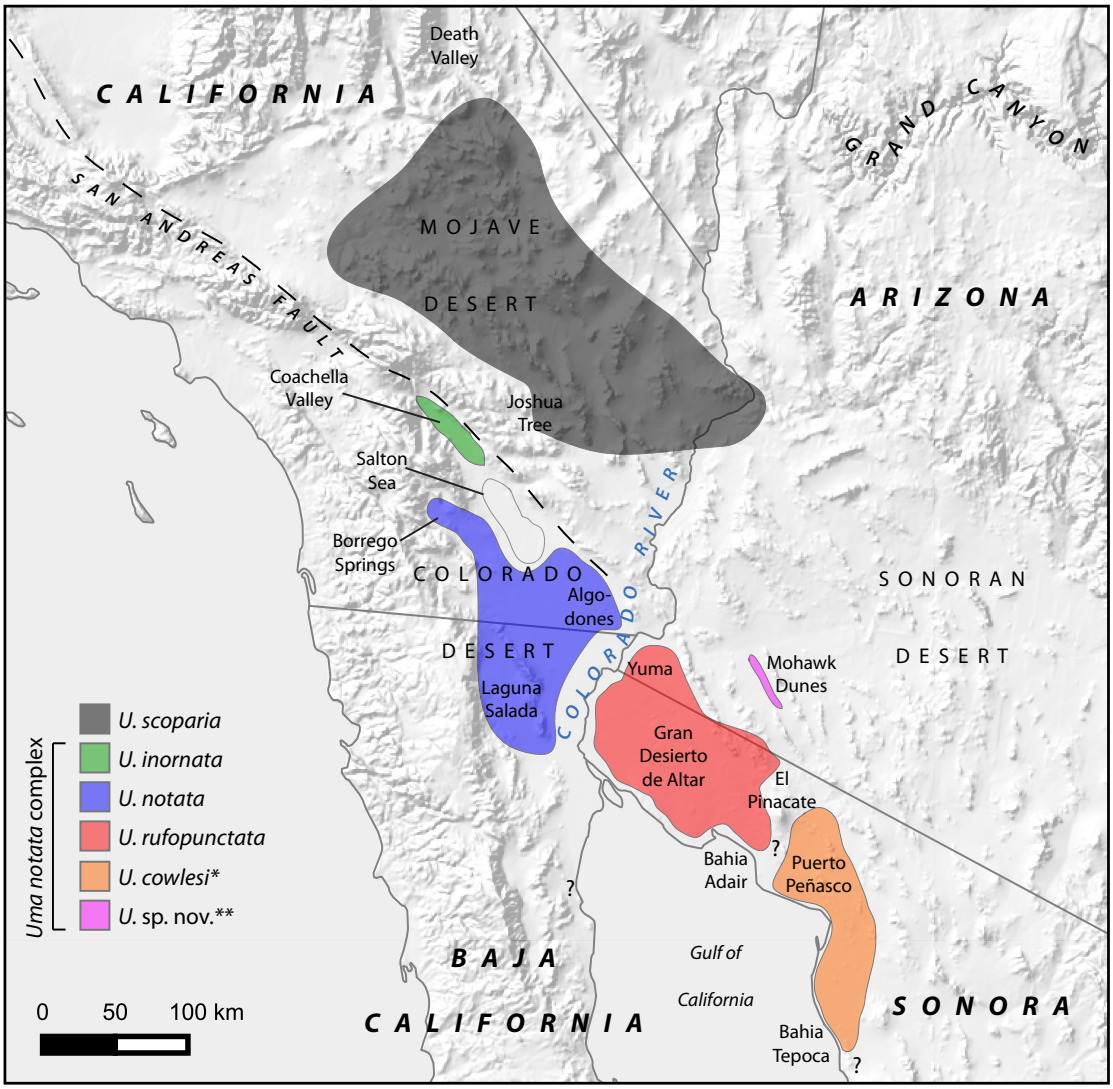

Figure 2.1. Approximate historical geographic distributions of the six operational taxonomic units (OTUs) within the genus *Uma* examined in this chapter. Four of the OTUs (*U. scoparia*, *U. inornata*, *U. notata*, and *U. rufopunctata*) have been described as species and are currently recognized as valid (Crothers 2012), while the other two are not. **Uma cowlesi* (Heifetz 1941; Tepoca Bay, Sonora, Mexico) was described as a subspecies of *U. notata* with an associated type series but later sunk within *U. n. rufopunctata* by Norris (1958), yet my analyses suggest that this taxon may be valid, so is notated as a species. **The proposed novel species from the Mohawk Dunes, Arizona (Trepanier and Murphy 2001) has yet to be formally named, described, and recognized, so is also notated as a species.

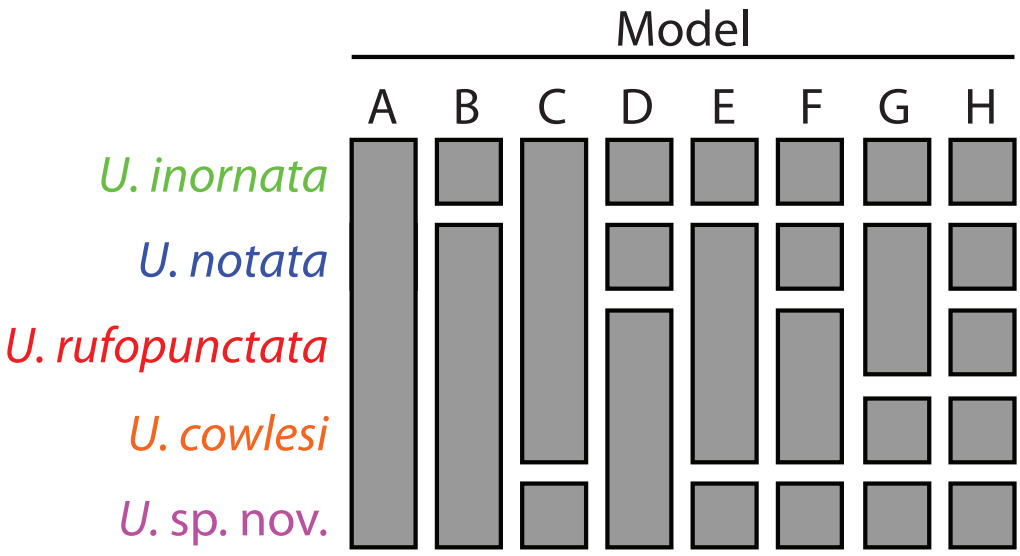

Figure 2.2. Eight alternative species delimitation hypotheses tested for *Uma* in this chapter. Rows represent the six operational taxonomic units shown in Figure 2.1, while columns represent the species delimitation models A–H. Model A represents the simplest model, treating the *U. notata* complex as a single species sister to *U. scoparia* (the outgroup), while the most complex model (H) treats each OTU as a separate species. All other tested models are based on previously published taxonomies or results of the Geneland analyses under different values for the maximum number of populations (K_{max}) as follows: Model A (2 sp., Norris 1958), Model B (3 sp., Stebbins 2003), Model C (3 sp., Geneland, K_{max}=3), Model D (4 sp., Jones and Lovich 2009), Model E (4 sp., Geneland, K_{max}=4), Model F (5 sp., Trepanier and Murphy 2001), Model G (5 sp., Geneland, K_{max}=5), and Model H (6 sp.).

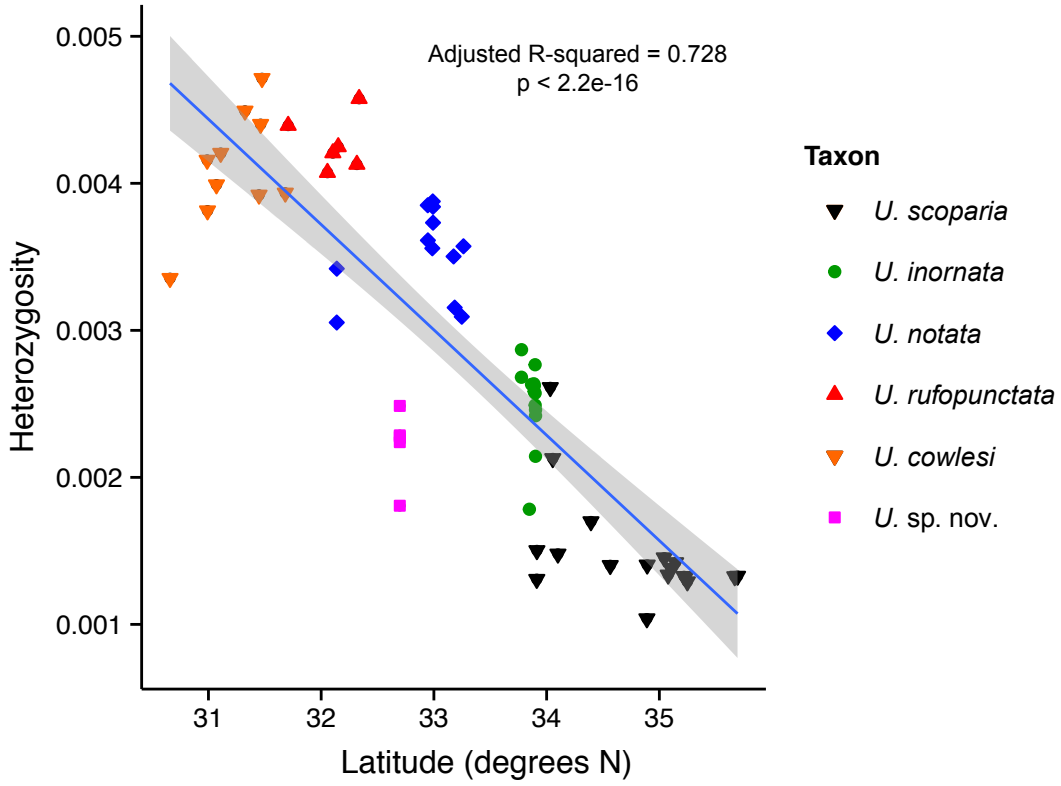

Figure 2.3. Linear regression of latitude (degrees North) vs. heterozygosity (percent polymorphism in the RADseq data) across all *Uma* species.

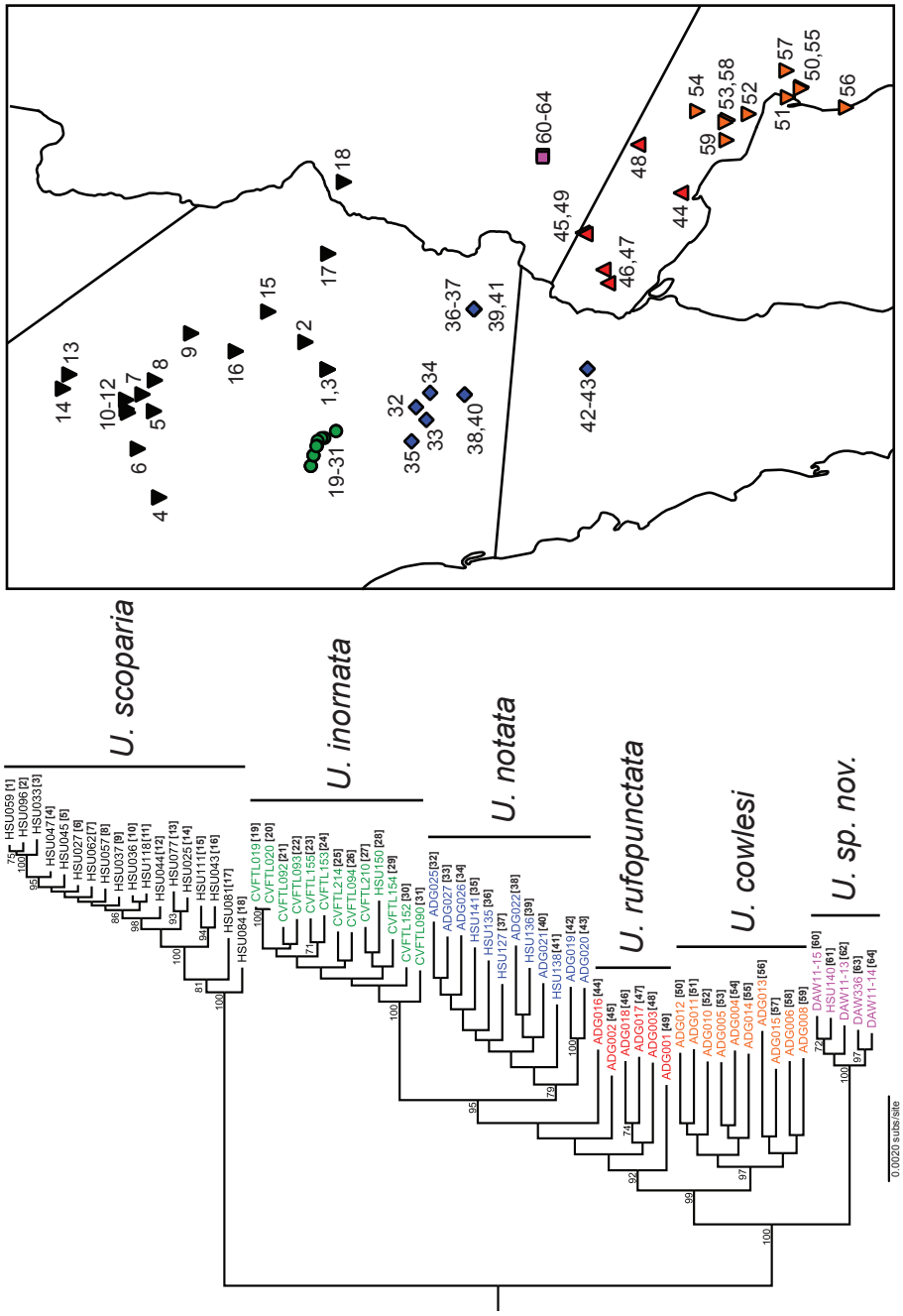

Figure 2.4. Left: *Uma* maximum likelihood (ML) tree estimated with RAxML v8.1.1 of 1,614 concatenated RAD loci (149,784 bp matrix, 7,752 variable sites) representing at least 40 of 64 individuals. Right: Map of the 64 individuals with RAD sequence data. Colored names at tips correspond to individual field numbers (Table 2.1), and the numbers in brackets correspond to the numbered points on the map. Numbers above nodes represent bootstrap support values.

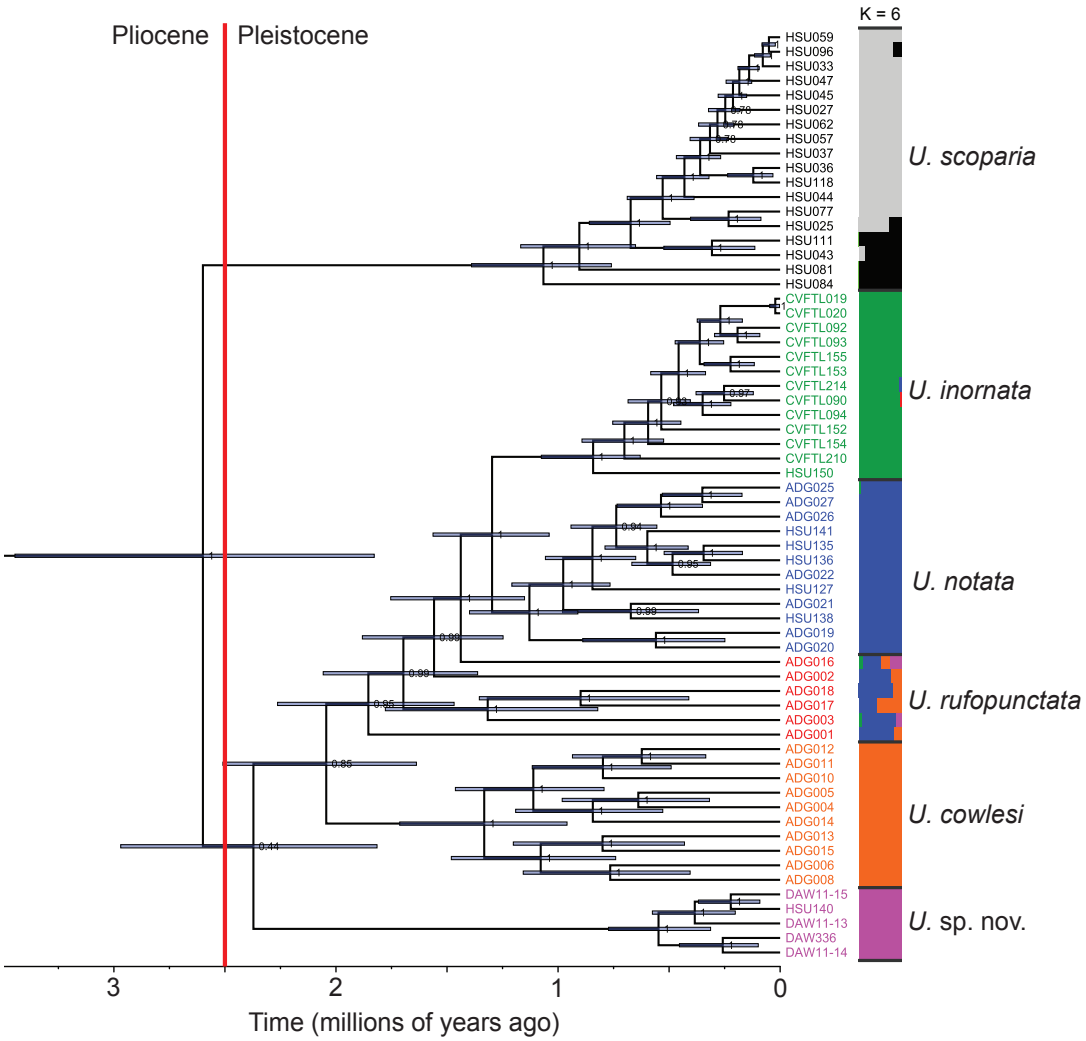

Figure 2.5. *Uma* Bayesian clock tree of 1,614 concatenated loci (149,784 bp matrix, 7,752 variable sites) present in 40+ of 64 individuals (left) aligned with the results of ADMIXTURE at K=6 (best-fit model) and K=7 using the 468 unlinked SNPs present in 60+ of 64 individuals (right). The shaded bars on nodes represent 95% Highest Posterior Density (HPD) confidence intervals.

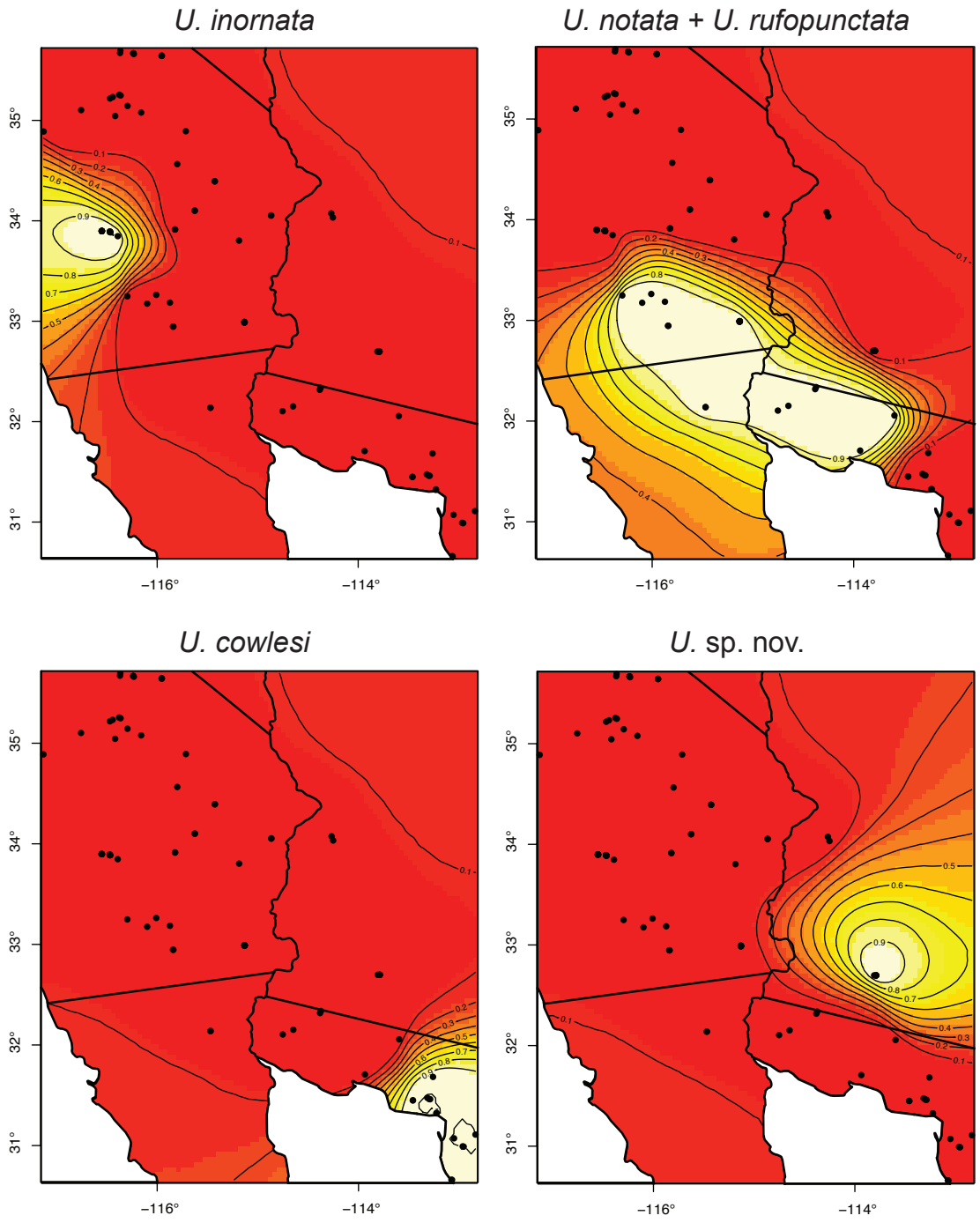

Figure 2.6. Results of the Geneland analysis of *Uma* populations under K_{max}=5 showing maps of the posterior probability (PP) of population assignment (the fifth cluster, *U. scoparia*, is not shown). All individuals were assigned to their respective populations with PP > 0.9.

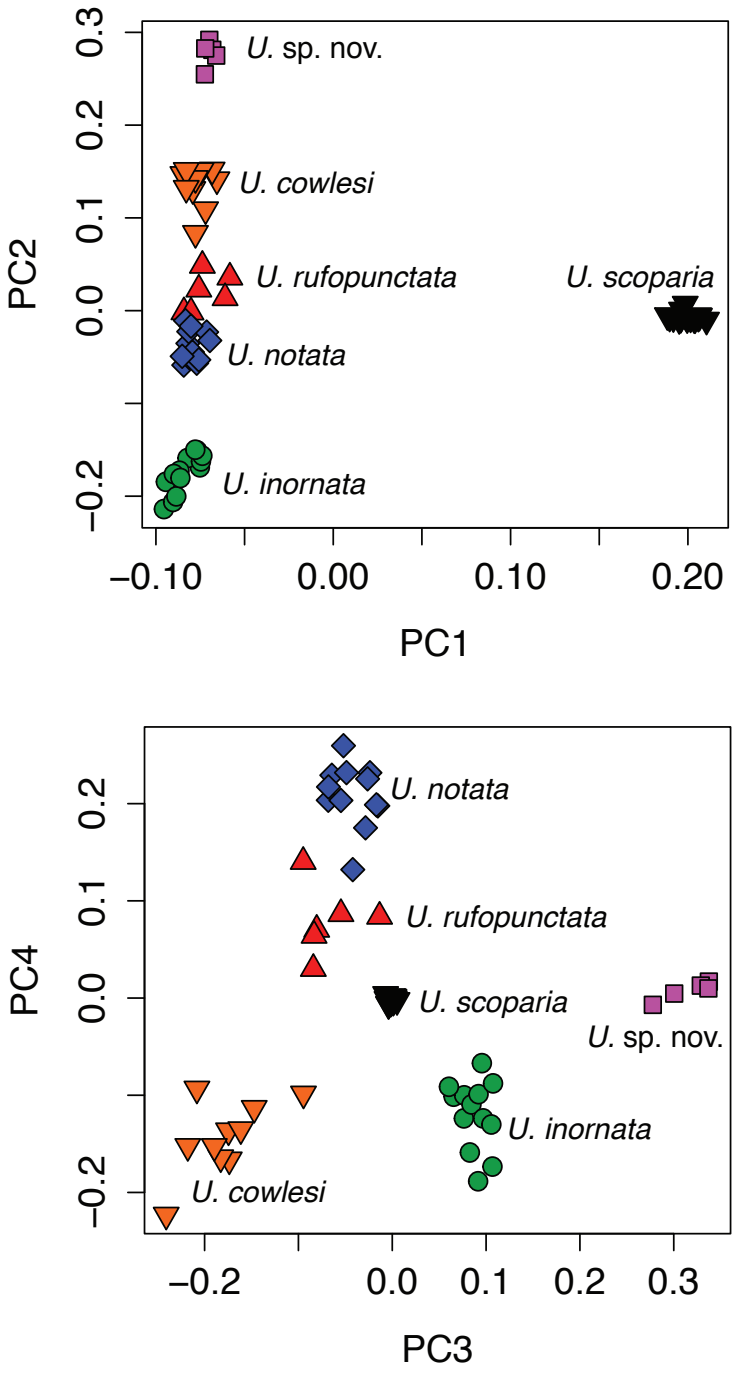

Figure 2.7. Principal Components Analysis (PCA) plots of 468 unlinked biallelic SNPs present in at least 60 of 64 *Uma* individuals. PC1, PC2, PC3, and PC4 represent 21.5%, 7.1%, 5.4%, and 4.8% of the variance in the dataset, respectively.

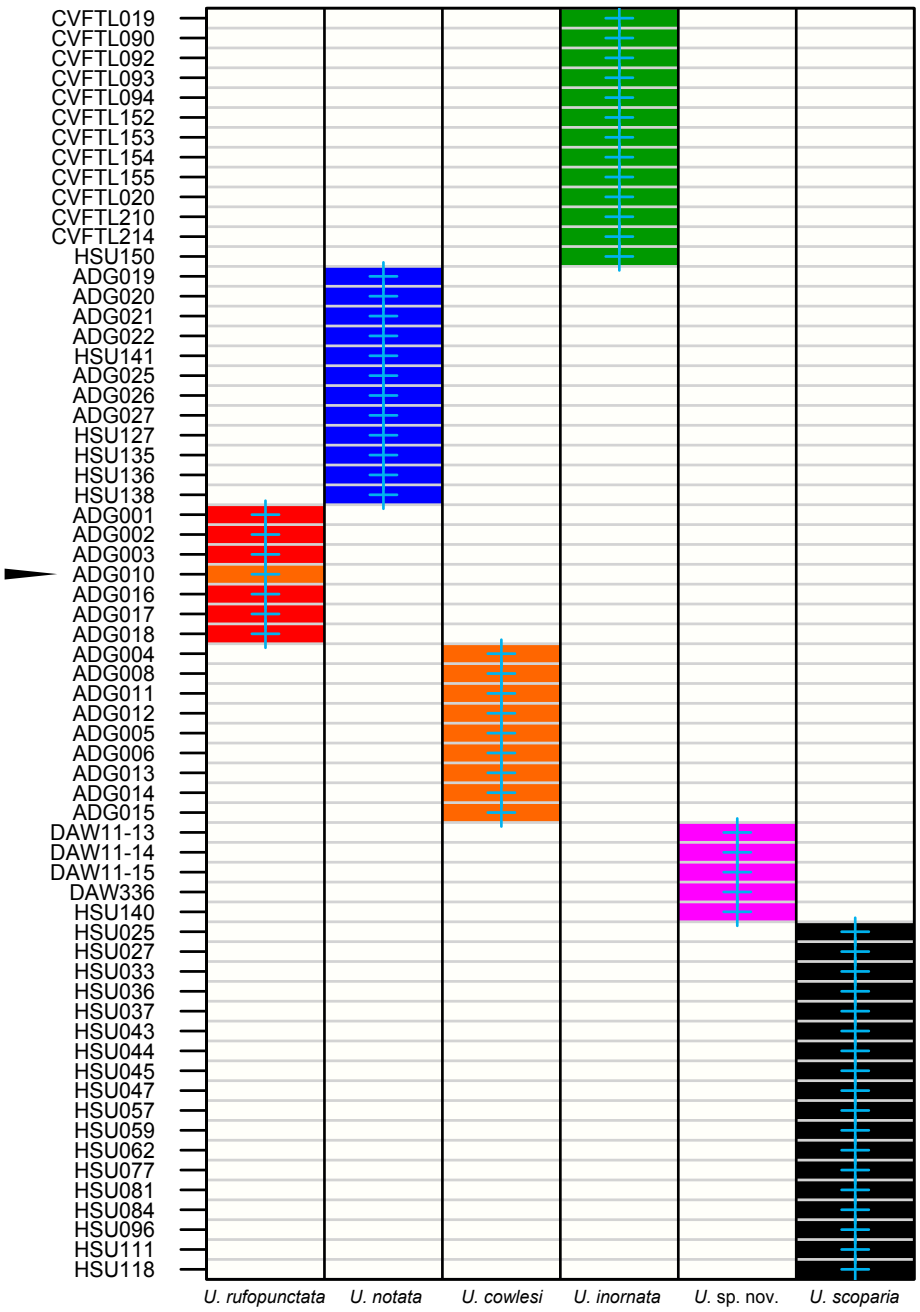

Figure 2.8. *Uma* population assignment as determined by Discriminant Analysis of Principal Components (DAPC) under K=6. Note that the assignment of individuals matches the six OTUs identified in Figures 2.1, 2.4, and 2.5, with the exception of ADG010, which was assigned to *U. rufopunctata* in this analysis but was grouped with *U. cowlesi* in all other analyses.

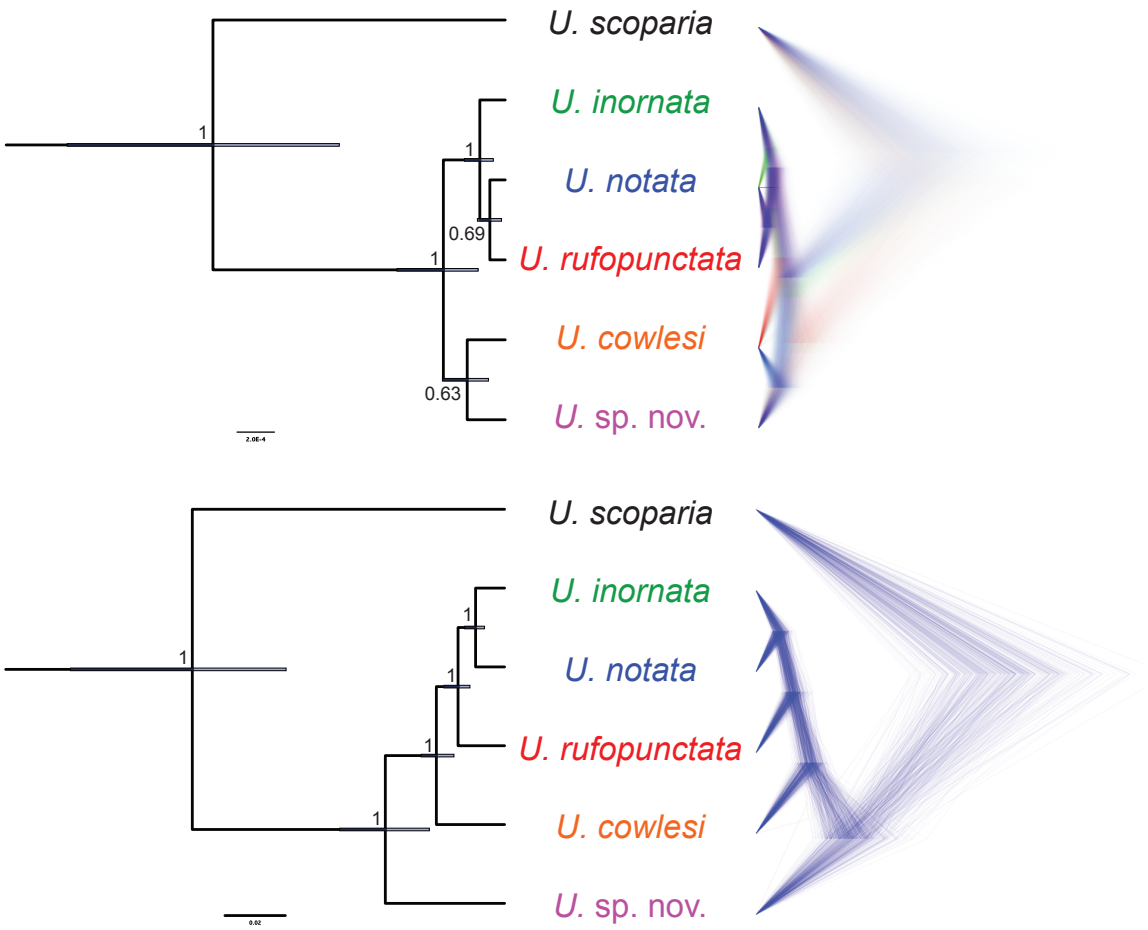

Figure 2.9. *Uma* species trees for the six OTUs under the most complex delimitation model (H). On the left are the maximum clade credibility trees, and on the right are the posterior distributions after burn-in visualized with DensiTree. On the top is the species tree estimated using *BEAST and 10 nuclear loci collected with Sanger sequencing, while on the bottom is the species tree estimated with SNAPP and 468 unlinked SNPs.

Table 2.1. Basic data for *Uma* specimens included in this chapter.

Field #	Voucher #	Taxon	Latitude	Longitude	Sanger	RADseq
ADG001	SDSNH76142	*U. rufopunctata*	32.3365500	-114.3740000	Y	Y
ADG002	SDSNH76143	*U. rufopunctata*	32.3166333	-114.3820000	Y	Y
ADG003	SDSNH76144	*U. rufopunctata*	32.0555833	-113.5961667	Y	Y
ADG004	SDSNH76145	*U. cowlesi*	31.6819500	-113.2592833	Y	Y
ADG005	SDSNH76146	*U. cowlesi*	31.4752833	-113.3171833	Y	Y
ADG006	SDSNH76147	*U. cowlesi*	31.4625833	-113.2981833	Y	Y
ADG007	SDSNH76148	*U. cowlesi*	31.4601667	-113.2882667	Y	-
ADG008	SDSNH76149	*U. cowlesi*	31.4486667	-113.4605333	Y	Y
ADG009	n/a	*U. cowlesi*	31.4490833	-113.4603000	Y	-
ADG010	SDSNH76150	*U. cowlesi*	31.3241667	-113.2261167	Y	Y
ADG011	SDSNH76151	*U. cowlesi*	31.0711667	-113.0507500	Y	Y
ADG012	SDSNH76152	*U. cowlesi*	30.9917667	-112.9654500	Y	Y
ADG013	SDSNH76153	*U. cowlesi*	30.6593000	-113.0705667	Y	Y
ADG014	SDSNH76154	*U. cowlesi*	30.9886667	-112.9550833	Y	Y
ADG015	SDSNH76155	*U. cowlesi*	31.1088833	-112.8384167	Y	Y
ADG016	SDSNH76156	*U. rufopunctata*	31.7072500	-113.9366167	Y	Y
ADG017	SDSNH76157	*U. rufopunctata*	32.1036333	-114.7519000	Y	Y
ADG018	SDSNH76158	*U. rufopunctata*	32.1513333	-114.6482833	Y	Y
ADG019	SDSNH76159	*U. notata*	32.1383667	-115.4697833	Y	Y
ADG020	SDSNH76160	*U. notata*	32.1381667	-115.4697833	Y	Y
ADG021	SDSNH76161	*U. notata*	32.9455500	-115.8394333	Y	Y
ADG022	SDSNH76162	*U. notata*	32.9468000	-115.8425833	Y	Y
ADG025	SDSNH76164	*U. notata*	33.2622667	-116.0097000	Y	Y
ADG026	SDSNH76165	*U. notata*	33.1843500	-115.8731333	Y	Y
ADG027	SDSNH76166	*U. notata*	33.1746500	-116.0992667	Y	Y
CVFTL019	n/a	*U. inornata*	33.7766200	-116.3167200	-	Y
CVFTL020	n/a	*U. inornata*	33.7772100	-116.3186100	-	Y
CVFTL086	n/a	*U. inornata*	33.8916651	-116.4710778	Y	-
CVFTL087	n/a	*U. inornata*	33.8912552	-116.4701395	Y	-
CVFTL088	n/a	*U. inornata*	33.8914641	-116.4704951	Y	-
CVFTL089	n/a	*U. inornata*	33.8918908	-116.4711413	Y	-
CVFTL090	n/a	*U. inornata*	33.8911293	-116.4702268	Y	Y
CVFTL091	n/a	*U. inornata*	33.8977795	-116.5504586	Y	-
CVFTL092	n/a	*U. inornata*	33.8977431	-116.5503939	Y	Y
CVFTL093	n/a	*U. inornata*	33.8982112	-116.5501318	Y	Y
CVFTL094	n/a	*U. inornata*	33.8979362	-116.5513987	Y	Y
CVFTL095	n/a	*U. inornata*	33.8983638	-116.5524131	Y	-
CVFTL096	n/a	*U. inornata*	33.8980533	-116.5513548	Y	-
CVFTL097	n/a	*U. inornata*	33.8979162	-116.5508688	Y	-
CVFTL098	n/a	*U. inornata*	33.8974724	-116.5503520	Y	-
CVFTL099	n/a	*U. inornata*	33.8982842	-116.5503802	Y	-
CVFTL152	n/a	*U. inornata*	33.8866700	-116.4665700	-	Y
CVFTL153	n/a	*U. inornata*	33.9014400	-116.6414500	-	Y
CVFTL154	n/a	*U. inornata*	33.9016400	-116.6419100	-	Y
CVFTL155	n/a	*U. inornata*	33.9016100	-116.6421000	-	Y
CVFTL210	n/a	*U. inornata*	33.8662700	-116.3971200	-	Y
CVFTL214	n/a	*U. inornata*	33.8895100	-116.4256100	-	Y
DAW11-13	n/a	*U.* sp. nov.	32.6974900	-113.7890400	Y	Y

Field #	Voucher #	Taxon	Latitude	Longitude	Sanger	RADseq
DAW11-14	n/a	*U.* sp. nov.	32.6974900	-113.7890400	Y	Y
DAW11-15	n/a	*U.* sp. nov.	32.6974900	-113.7890400	Y	Y
DAW336	n/a	*U.* sp. nov.	32.6974900	-113.7890400	Y	Y
DH051	n/a	*U. inornata*	33.8875373	-116.4675563	Y	-
DH052	n/a	*U. inornata*	33.8870658	-116.4669968	Y	-
DH053	n/a	*U. inornata*	33.8871933	-116.4672664	Y	-
HSU005	n/a	*U. scoparia*	35.6441667	-115.9543833	Y	-
HSU014	n/a	*U. scoparia*	35.2547333	-116.3778000	Y	-
HSU023	n/a	*U. scoparia*	35.6441667	-115.9544000	Y	-
HSU024	n/a	*U. scoparia*	35.6425000	-115.9545333	Y	-
HSU025	n/a	*U. scoparia*	35.6934500	-116.3659667	Y	Y
HSU027	n/a	*U. scoparia*	35.1022667	-116.7570500	Y	Y
HSU029	n/a	*U. scoparia*	35.6718333	-116.3695667	Y	-
HSU033	n/a	*U. scoparia*	33.9126333	-115.8213000	Y	Y
HSU036	n/a	*U. scoparia*	35.2323167	-116.4418167	Y	Y
HSU037	n/a	*U. scoparia*	34.8920833	-115.7141667	Y	Y
HSU043	n/a	*U. scoparia*	34.5642500	-115.8003167	Y	Y
HSU044	n/a	*U. scoparia*	35.2483167	-116.3643333	Y	Y
HSU045	n/a	*U. scoparia*	35.0434667	-116.4180833	Y	Y
HSU047	n/a	*U. scoparia*	34.8893667	-117.1299333	Y	Y
HSU057	n/a	*U. scoparia*	35.0783167	-116.1585500	Y	Y
HSU059	n/a	*U. scoparia*	33.9136833	-115.8231667	Y	Y
HSU062	n/a	*U. scoparia*	35.1441333	-116.2946333	Y	Y
HSU075	n/a	*U. scoparia*	35.6677667	-116.2381500	Y	-
HSU076	n/a	*U. scoparia*	35.6663333	-116.2382333	Y	-
HSU077	n/a	*U. scoparia*	35.6687167	-116.2363833	Y	Y
HSU078	n/a	*U. scoparia*	35.6615833	-116.2321000	Y	-
HSU080	n/a	*U. scoparia*	34.0511500	-114.8666000	Y	-
HSU081	n/a	*U. scoparia*	34.0543833	-114.8651833	Y	Y
HSU084	n/a	*U. scoparia*	34.0325500	-114.2516167	Y	Y
HSU085	n/a	*U. scoparia*	34.0742333	-114.2673167	Y	-
HSU096	n/a	*U. scoparia*	34.1002000	-115.6266833	Y	Y
HSU097	n/a	*U. scoparia*	34.1011833	-115.6252667	Y	-
HSU111	n/a	*U. scoparia*	34.3932000	-115.4260167	Y	Y
HSU112	n/a	*U. scoparia*	34.3934833	-115.4281167	Y	-
HSU118	n/a	*U. scoparia*	35.2187000	-116.4691667	Y	Y
HSU124	n/a	*U. scoparia*	33.8023167	-115.1849000	Y	-
HSU127	n/a	*U. notata*	32.9859000	-115.1324833	Y	Y
HSU128	n/a	*U. notata*	32.9875667	-115.1314000	Y	-
HSU134	n/a	*U. notata*	32.9922833	-115.1328500	Y	-
HSU135	n/a	*U. notata*	32.9921333	-115.1322333	Y	Y
HSU136	n/a	*U. notata*	32.9917667	-115.1318167	Y	Y
HSU137	n/a	*U. notata*	32.9894833	-115.1333333	Y	-
HSU138	n/a	*U. notata*	32.9900167	-115.1307000	Y	Y
HSU139	n/a	*U. notata*	32.9885000	-115.1297167	Y	-
HSU140	n/a	*U. sp. nov*	32.6973833	-113.8095167	Y	Y
HSU141	n/a	*U. notata*	33.2474333	-116.2957333	Y	Y
HSU150	1743CAP	*U. inornata*	33.8470000	-116.3919000	Y	Y

Table 2.1 (continued).

Table 2.2. Summary statistics for phased Sanger data. L = length of locus, S = number of segregating (polymorphic) sites, N = number of alleles (number of individuals x 2), h = number of haplotypes, D = Tajima's D statistic (none of the values were found to be significant, P > 10), and π = nucleotide diversity.

Locus	L	N	S	h	D	$\pi_{accipiria}$	$\pi_{cinereata}$	π_{notata}	$\pi_{trifasciata}$	$\pi_{cavalesi}$	$\pi_{sp. nov.}$
BDNF	613	166	8	10	-0.22	0.0000	0.0020	0.0015	0.0019	0.0016	0.0000
PNN	600	130	18	20	-1.18	0.0012	0.0026	0.0038	0.0042	0.0032	0.0000
R35	450	154	18	26	0.33	0.0019	0.0022	0.0063	0.0065	0.0076	0.0000
RAG-1	660	152	15	14	-1.09	0.0010	0.0003	0.0010	0.0007	0.0010	0.0008
sun07	530	142	16	19	-1.09	0.0001	0.0006	0.0012	0.0028	0.0032	0.0004
sun08	555	84	12	13	-0.60	0.0008	0.0014	0.0036	0.0025	0.0040	0.0000
sun10	599	130	32	37	-1.21	0.0012	0.0034	0.0066	0.0120	0.0090	0.0039
sun12	359	152	17	24	-1.44	0.0005	0.0002	0.0063	0.0052	0.0023	0.0014
sun18	379	74	12	14	1.15	0.0006	0.0095	0.0085	0.0018	0.0069	0.0000
sun28	469	152	24	43	-0.26	0.0018	0.0039	0.0058	0.0090	0.0085	0.0056
MEAN	521	134	17.2	22	-	0.0009	0.0026	0.0045	0.0046	0.0047	0.0012
TOTAL	5,214	1,336	172	220	-	0.0092	0.0259	0.0445	0.0465	0.0474	0.0121

Table 2.3. Summary statistics for RADseq data. N_{reads} = total number of reads after demultiplexing, pass.total = total number of reads that passed quality filter, D_{total} = average depth of the total clusters, C_{pass} = number of clusters that passed the coverage threshold (10x), D_{pass} = average depth of coverage of clusters that passed the threshold, nloci = total number of loci per lizard after aligning across individuals, floci = number of loci with >10x depth and passed paralog filter, nsites = number of sites across all loci, npoly = number of polymorphic sites, %poly = percentage of sites that are polymorphic.

Sample	nreads	pass.total	D_{total}	C_{pass}	D_{pass}	floci	nsites	npoly	%poly
ADG001	1,284,910	1,177,214	27	4,336	218	3,750	317,962	1,455	0.458%
ADG002	779,794	709,362	27	3,690	153	3,241	275,163	1,136	0.413%
ADG003	1,371,402	1,250,520	43	4,288	238	3,692	313,021	1,275	0.407%
ADG004	873,311	798,605	30	4,013	159	3,481	295,353	1,162	0.393%
ADG005	1,180,477	1,077,981	18	4,972	170	4,316	365,945	1,725	0.471%
ADG006	833,580	776,822	22	4,174	145	3,669	311,469	1,371	0.440%
ADG008	966,476	886,017	26	3,545	198	3,068	260,415	1,021	0.392%
ADG010	846,809	773,762	31	4,262	142	3,725	316,114	1,420	0.449%
ADG011	893,114	813,136	30	3,948	164	3,425	290,683	1,160	0.399%
ADG012	806,560	742,676	38	3,253	192	2,832	240,458	917	0.381%
ADG013	736,111	676,307	24	3,885	142	3,387	287,447	964	0.335%
ADG014	1,467,692	1,369,699	31	4,607	239	3,978	337,193	1,402	0.416%
ADG015	847,823	782,698	32	4,529	135	3,965	336,279	1,414	0.420%
ADG016	1,122,689	1,040,489	32	3,758	233	3,245	275,412	1,210	0.439%
ADG017	704,450	647,833	24	3,894	131	3,380	286,922	1,207	0.421%
ADG018	944,433	875,328	29	3,432	215	3,010	255,645	1,086	0.425%
ADG019	932,669	856,819	27	4,754	142	4,145	351,230	1,201	0.342%
ADG020	759,934	711,455	21	3,272	185	2,889	245,316	749	0.305%
ADG021	710,943	661,913	29	4,172	133	3,672	311,621	1,200	0.385%
ADG022	1,108,250	1,020,587	40	3,785	227	3,262	276,850	1,000	0.361%
ADG025	983,329	896,755	36	4,345	165	3,769	319,460	1,141	0.357%
ADG026	1,142,653	1,043,669	29	3,587	243	3,093	262,475	828	0.315%
ADG027	806,846	742,280	29	3,652	160	3,213	272,656	955	0.350%
CVFTL019	172,957	157,632	7	2,535	45	2,247	190,597	511	0.268%
CVFTL020	410,353	382,265	14	3,153	94	2,826	239,866	688	0.287%
CVFTL090	474,278	442,835	18	2,864	121	2,582	219,146	575	0.262%
CVFTL092	660,781	605,566	18	4,143	108	3,596	304,760	843	0.277%
CVFTL093	120,672	109,643	12	2,039	42	1,877	159,345	397	0.249%
CVFTL094	528,737	487,556	11	3,018	124	2,669	226,610	583	0.257%
CVFTL152	390,269	356,094	11	2,876	96	2,538	215,450	568	0.264%
CVFTL153	513,743	471,670	24	2,929	136	2,605	221,146	474	0.214%
CVFTL154	880,300	823,086	15	4,649	139	4,028	341,436	826	0.242%
CVFTL155	955,557	872,750	27	3,523	203	3,089	262,238	645	0.246%
CVFTL210	950,755	875,150	28	5,245	127	4,566	386,745	1,018	0.263%
CVFTL214	800,461	738,785	29	4,507	135	3,973	336,977	871	0.258%
DAW11-13	875,560	809,799	26	4,827	128	4,235	358,821	892	0.249%
DAW11-14	488,066	453,095	26	3,379	112	3,023	256,616	584	0.228%
DAW11-15	714,705	657,154	21	3,743	142	3,240	274,726	628	0.229%
DAW336	310,857	284,172	24	2,581	94	2,336	198,256	444	0.224%
HSU025	1,172,133	1,083,963	32	4,311	214	3,815	323,758	430	0.133%
HSU027	768,100	712,839	27	4,253	136	3,740	317,003	438	0.138%

Sample	nreads	pass.total	D_{total}	C_{pass}	D_{pass}	floci	nsites	npoly	%poly
HSU033	1,049,491	976,276	15	4,937	157	4,280	362,682	474	0.131%
HSU036	633,096	583,526	20	3,819	120	3,375	286,143	379	0.132%
HSU037	1,721,110	1,604,027	37	5,791	236	5,115	433,646	609	0.140%
HSU043	927,181	853,297	18	3,796	185	3,280	278,304	390	0.140%
HSU044	1,517,162	1,392,543	11	5,348	193	4,452	377,462	487	0.129%
HSU045	909,521	840,129	31	3,451	208	3,040	258,023	375	0.145%
HSU047	1,515,434	1,405,750	26	5,183	227	4,529	384,044	399	0.104%
HSU057	731,099	671,711	10	3,368	154	2,927	248,436	332	0.134%
HSU059	1,654,180	1,532,194	48	5,298	248	4,718	399,977	601	0.150%
HSU062	994,259	918,867	32	3,647	212	3,195	271,039	385	0.142%
HSU077	710,432	658,070	33	4,107	134	3,666	310,897	412	0.133%
HSU081	713,945	660,991	30	3,603	152	3,206	272,053	579	0.213%
HSU084	1,764,750	1,621,851	64	5,564	247	4,946	419,281	1,095	0.261%
HSU096	685,232	635,498	18	3,984	124	3,530	299,548	443	0.148%
HSU111	921,183	850,584	28	4,060	174	3,633	308,388	524	0.170%
HSU118	1,450,267	1,359,135	14	5,107	212	4,364	370,141	491	0.133%
HSU127	790,404	725,963	26	3,727	165	3,281	278,522	991	0.356%
HSU135	884,436	810,715	36	3,902	168	3,412	289,640	1,081	0.373%
HSU136	823,345	764,650	38	3,641	180	3,218	273,142	1,049	0.384%
HSU138	730,140	676,558	33	4,459	126	3,916	332,220	1,288	0.388%
HSU140	848,113	782,459	38	3,234	197	2,875	244,054	441	0.181%
HSU141	644,175	589,672	16	3,833	118	3,369	285,854	884	0.309%
HSU150	298,304	278,419	15	2,943	73	2,663	226,012	403	0.178%
MEAN	878,278	810,139	26	3,961	161	3,472	294,501	821	0.280%
TOTAL	56,209,798	51,848,866	-	253,530	-	222,182	18,848,093	52,526	-

Table 2.3 (continued).

Table 2.4. Cross-validation error (CVE) and Bayesian information criterion (BIC) values for different K values estimated from Admixture and Discriminant Analysis of Principal Components (DAPC), respectively.

K	BIC	CVE
2	172.44	0.303
3	166.13	0.274
4	160.70	0.272
5	154.17	0.272
6	**152.91**	**0.264**
7	152.79	0.285
8	153.51	0.303
9	152.75	0.314
10	152.18	0.334

Table 2.5. Pair-wise F_{ST} values calculated among *Uma* populations inferred with Admixture under K=6.

	U. inornata	***U. notata***	***U. cowlesi***	***U. scoparia* (NW)**	***U. sp. nov.***
U. notata	0.30				
U. cowlesi	0.33	0.22			
***U. scoparia* (NW)**	0.74	0.63	0.60		
***U.* sp. nov.**	0.57	0.41	0.36	0.78	
***U. scoparia* (SE)**	0.70	0.59	0.57	0.21	0.73

Table 2.6. Results of the Bayes Factor Delimitation (BFD) analyses using Sanger data (Grummer et al. 2014) with path sampling (PS) and stepping-stone (SS) of marginal likelihoods, and BFD* analyses using RAD data (Leaché et al. 2014) using PS. Bayes factor (BF) values are used to rank species models relative to the best species model (the one with the lowest marginal likelihood). Model H was decisively supported as the best-fit model (BF 0) under all three methods.

	BFD (PS)			**BFD (SS)**			**BFD* (PS)**		
Model	**ML**	**Rank**	**BF**	**ML**	**Rank**	**BF**	**ML**	**Rank**	**BF**
A	-10252.1	8	560	-10277.0	8	564	-9114.2	8	3160
B	-10105.5	6	267	-10129.7	6	270	-8367.2	7	1667
C	-10180.3	7	417	-10204.6	7	419	-8432.9	6	1798
D	-10019.8	4	96	-10043.8	4	98	-8076.2	5	1085
E	-10034.6	5	125	-10058.6	5	127	-7769.1	4	470
F	-9980.5	2	17	-10003.7	2	18	-7573.6	3	79
G	-9988.4	3	33	-10011.7	3	34	-7561.1	2	54
H	**-9971.9**	**1**	**0**	**-9994.9**	**1**	**0**	**-7534.0**	**1**	**0**

Chapter 3: Comparative Phylogeography of Phrynosomatid Lizards of the Baja California Peninsula: Testing the Simultaneous Vicariance and Range Expansion Hypotheses

One of the primary goals of comparative phylogeography is to understand population histories and the relative influences of vicariance and dispersal in generating patterns such as the geographic distributions of multiple species. The Baja California Peninsula and Gulf of California of western North America are among Earth's youngest and most striking geological features, offering a unique opportunity to study the mechanisms of speciation using a comparative phylogeographic approach. In this chapter I studied peninsular populations of several co-distributed lineages of lizards (Phrynosomatidae), including the genera *Callisaurus*, *Petrosaurus*, *Urosaurus*, and *Sceloporus*. Each of these lineages occupies numerous islands in the gulf and has deep genetic breaks in one or more locations along the peninsula. I collected genome-wide sequence and single-nucleotide polymorphism data from 228 individual lizards from the peninsula and eight islands using restriction-associated DNA sequencing (RADseq). I used concatenated phylogenetic methods, and parametric and non-parametric clustering algorithms to detect spatially and temporally concordant phylogeographic breaks among co-distributed taxa. The estimated divergence dates across co-distributed clades revealed that for each of the five regions where comparisons are possible, at least two or three episodes of divergence are required to explain the observed patterns. I explored the effects of latitude on heterozygosity using linear regression and phylogenetic generalized least

squares (PGLS) to test for range expansions or bottlenecks associated with Pleistocene glacial cycles or continent-to-peninsula invasions. These results are consistent with the hypothesis that plate tectonic events established a complex landscape matrix with numerous barriers to dispersal that ultimately facilitated divergence, generating the idiosyncratic patterns observed among co-distributed lineages.

I. INTRODUCTION

"*The peaceful union of such contrasting environmental elements as desert and sea embellishes [Baja California] with a mystical beauty... Its often violent and complex geological evolution has in turn contributed to the evolution of forms of life found nowhere else on earth. This uniqueness, coupled with the rugged and unforgiving terrain, leaves Baja California a region steeped in mystery and intrigue, one of the last biological frontiers in North America*." – Grismer (2002, p. *xi*)

One of the primary goals of historical biogeography is to understand the processes that generate patterns such as the geographic distributions of species (Riddle and Hafner 2010). Large-scale landscape reconfigurations (e.g. the tectonic rifting of continents, the uplift of mountain ranges, and the formation of rivers and seaways) may drive speciation directly by fragmenting formerly continuous species distributions through vicariance, generating spatially and temporally congruent patterns of genetic divergence in co-distributed lineages,

especially those with similar ecological traits (Nelson and Platnick 1981). For example, the fragmentation of rainforests by semi-arid savannahs during cool/dry periods of the Pleistocene and the uplift of the Andes during the Tertiary in South America have been hypothesized to have driven speciation of Neotropical birds (Haffer 1969, Hoorn et al. 2010), and the gradual closure of the Isthmus of Panama (Keigwin 1978) has been hypothesized to have caused speciation in marine sister species in the Caribbean and Pacific (Marko 2002). An alternative model is that allopatric speciation is largely driven by dispersal events, often climate-related, over a matrix previously shaped by large-scale landscape change, so that older lineages would have a higher probability of dispersing across geographical barriers and diversifying (Smith et al. 2014a,b). Under this model, organism-specific abilities to persist and disperse across a structured landscape are the primary contributors to speciation, while plate tectonics play a more indirect role by establishing the barriers to dispersal, such as mountain ranges and ocean basins, and their associated climatic ecotones. For example, molecular clock analyses have revealed trans-Atlantic dispersal in primates (Schrago and Russo 2003), crocodilians (Meredith et al. 2011), and amphisbaenian reptiles (Longrich et al. 2015), dispelling long-standing vicariance hypotheses that relied on the break-up of Gondwana to explain the distributions of these taxa. However, when examining an entire biota for a given region, these alternative hypotheses are not necessarily mutually exclusive, frustrating efforts to disentangle the forces generating Earth's biodiversity. This is critically

important because we cannot protect biodiversity, much of it cryptic, unless we know what is out there and how it evolved.

The burgeoning field of phylogeography (Avise 2000) provides an analytical toolkit that can be used to test these hypotheses in an evolutionary context. The ultimate goal of phylogeography is to understand the historical processes that underlie the geographic distribution of genetic variation within species and species complexes (Avise 2000, Knowles 2009). Phylogeography provides detailed species-specific data regarding how geologic events, climate change, and other factors have interacted with aspects of a species ecological niche to influence its evolution (Knowles 2009). In a comparative context, these causal links can reveal ways in which assemblages or communities are structured by shared biogeographic history or idiosyncratic species-specific events (Knowles 2009), thus 'comparative phylogeography' extends this approach to multiple co-distributed taxa that span a biogeographic barrier or ecotone of interest (Zink 1996, 2002, Arbogast and Kenagy 2001). It had long been assumed that the most parsimonious explanation for multiple taxa that demonstrate spatially congruent phylogeographic breaks is that they share a common history (=simultaneous divergence) (Arbogast and Kenagy 2001). However, in the past decade, the field of 'statistical phylogeography' (Knowles and Maddison 2002, Knowles 2009) has matured to the point where this hypothesis can be tested explicitly using maximum-likelihood and/or Bayesian models that incorporate uncertainty in the nucleotide substitution and/or coalescent processes (e.g. Lee and Edwards 2008, Carstens et al. 2009). An important part of this development

has been the growing recognition that multi-locus data is important to model stochastic demographic processes (particularly incomplete lineage sorting), estimate divergence times among individuals or species, and to capture genome-wide patterns of variation and subdivision (Knowles 2009), something that could not be accomplished by the earlier wave of phylogeographic literature, which relied chiefly on single-locus mitochondrial DNA (mtDNA). In the past decade, several phylogeographic studies have adopted both comparative and statistical approaches to test alternative hypotheses (e.g. Carstens and Richards 2007, Hickerson et al. 2007, Leaché et al. 2007, Smith et al. 2014a, 2014b), prompting continued interest in the development of new methods in this field (Oaks 2014).

Most of the published comparative and statistical phylogeography literature has relied on either single or multi-locus data collected using Sanger sequencing methods, but in recent years, the widespread use of next-generation sequencing (NGS), also known as massively parallel or high-throughput sequencing, has revolutionized phylogeography, increasing the sizes of datasets (number of loci) by two or more orders of magnitude over that of most Sanger studies (Emerson et al. 2010, Carstens et al. 2012, McCormack et al. 2013). Because obtaining whole genome sequences for the tens or hundreds of individuals required for a high-resolution phylogeographic analysis is still not financially or technically feasible for most researchers, reduced representation libraries (RRLs) have become a popular method of choice (McCormack et al. 2013). One approach is to use a target capture method to target ultra-conserved elements (UCEs) obtained by aligning genomes of divergent tetrapod lineages,

using the standing variation in the outer flanking regions for phylogeographic analysis (Faircloth et al. 2012). This method is primarily of interest for deeper level phylogenetic problems, such as the phylogenetic placement of turtles relative to other reptiles (Crawford et al. 2012). The use of UCEs in phylogeography has been more limited, although it has been successfully used in a statistical comparative phylogeographic analysis of Neotropical birds (Smith et al. 2014a,b). However, this method is relatively expensive and it is likely that these loci are influenced by strong stabilizing selection, accounting for their ease of alignment across deeply divergent lineages, but violating the assumption of neutrality (Faircloth et al. 2012). A popular alternative (utilized in this study) is restriction-associated DNA sequencing (RADseq), a cost-effective method to collect short-read sequence data for tens to hundreds of individuals. This method works by digesting the genome with restriction enzymes, sequencing the regions flanking restriction sites, and assembling the raw data *de novo*. Because no reference genome is needed for assembly, it is well suited for non-model organisms (Peterson et al. 2012). However, this method suffers from non-random 'locus drop-out' due to mutations in restriction sites as increasingly distantly related taxa are included (Leaché et al. 2015), so is best suited for addressing population-level questions. Although RADseq has been demonstrated to be useful for recent phylogeographic analyses (e.g. Emerson et al. 2010, Streicher et al. 2014), to the best of my knowledge, to date there have been no published comparative phylogeographic studies that have utilized RADseq data to test for simultaneous divergence (but see the unpublished manuscript of Harvey et al.).

Thus, the use of RADseq data in comparative phylogeography remains a modern frontier of evolutionary biology.

Study System. The Baja California Peninsula (BCP), one of the most striking geological features on Earth (Figure 3.1), is an ideal region of high endemism to apply a novel comparative phylogeographic approach. In a geological and biogeographic sense, the BCP spans the USA/México border, extending ~1,400 km from San Gorgonio Pass in southern California to Cabo San Lucas in Baja California Sur. Rugged and remote, with high topographic and climatic complexity, it comprises part of North America's last remaining high-biodiversity wilderness area (the only such area lying outside of the tropics), and thus represents a top priority for global proactive conservation efforts (Mittermeier et al. 2003). The diverse archipelago in the Gulf of California (GoC) is renowned as an UNESCO World Heritage Site and is protected by the Mexican government as a Biosphere Reserve (Ezcurra et al. 2002). The BCP is geologically young, having been gradually transferred from the North American plate to the Pacific plate from 6 to 3.5 mya as oblique extensional faulting and later seafloor spreading along the San Andreas Fault system formed the modern GoC, separating the BCP from the continent by nearly 500 km to the northwest (Moore and Buffington 1968, Lonsdale 1989, Stock and Hodges 1989, Umhoefer et al. 2002). Glacial cycles of the Pleistocene epoch caused cyclical fluctuations in sea level, temperature, and precipitation, further impacting the endemic biota through repeated episodes of range contraction and expansion (Garrick et al. 2009, Holmgren et al. 2014).

Due to its complex geological history and diverse array of habitats, the BCP and its offshore islands offers a plethora of unique natural experiments to test hypotheses regarding lineage diversification and historical biogeography, and has long captured the interest of ecologists and evolutionary biologists (Savage 1960, Soulé and Sloan 1966, Soulé and Yang 1973, Case 1975). Prior to the discovery of plate tectonics and the advent of modern phylogeographic methods, Savage (1960) hypothesized that the herpetofauna of the southern BCP must have originated by dispersing southwards from the North American continent. Later, after the advent of plate tectonics, Murphy (1983a,b) hypothesized that vicariant events associated with the tectonic rifting of the peninsula and "mid-peninsular seaways" were important in structuring the biogeographic patterns observed today. Proponents of the mid-peninsular seaway hypothesis (Aguirre et al. 1999, Murphy and Aguirre-Leon 2002) have thus described the ancient BCP as a "peninsular archipelago" (not to be confused with the modern islands in the GoC). Supporters of this hypothesis generally cite the spatially concordant single locus-based phylogeographic boundaries observed across a variety of marine and terrestrial taxa (Zink 2002, Riddle et al. 2000, Riginos 2005, Lindell et al. 2006) as supporting evidence, since direct geological evidence for these past seaways is lacking (Crews and Hedin 2006). These breaks frequently occur in the Vizcaíno Desert or its adjacent mountains (Riddle et al. 2000, Gottscho 2014), the Loreto basin and the Sierra de la Giganta (Lindell et al. 2005, 2008), and the low-lying Isthmus of La Paz which isolates the dry tropical cape region including the Sierra de la Laguna and Cabo San Lucas (Grismer 2002, Lindell et al. 2005, 2008).

Regarding the mid-peninsular break specifically, Leaché et al. (2007) pointed out that "despite the effort expended on debating the existence of a putative seaway, previous studies did not consider that the appreciable differences in genetic divergence observed across different lineage pairs separated at the mid-peninsula are an indication that diversification may have occurred in multiple waves." Leaché et al. (2007) used a Bayesian analysis of a hierarchical model to test for simultaneous vicariance across 12 co-distributed sister lineages (six mammals and six reptiles) sharing a single locus –based genealogical break at the mid-peninsula, and found that at least two waves of divergence were required to explain the observed mitochondrial data. A third hypothesis is that climatic cycles related to Pleistocene glacial cycles were important in generating phylogeographic patterns, specifically, that arid-adapted taxa were confined to southern or low-elevation refugia during glacial maxima, and expanded to northern and high-elevation localities since the last glacial maximum (Hewitt 1996), as thought to be the case for columnar cacti in the BCP (Nason et al. 2002; Clark-Tapia and Molina-Freaner 2003). All of these studies have used one or a few loci; none have utilized genome-wide sequence or SNP data.

These alternative hypotheses all have predictions that can be tested using a comparative phylogeographic approach. For a given lineage distributed throughout the peninsula, if the lineage originated in the north and dispersed south (Savage 1960), I predict that individual heterozygosity will show a significant positive correlation with latitude after correcting for phylogeny (Excoffier et al. 2009). For example, this logic has been used to infer the

expansion of modern humans from Eastern Africa using the observed steady decrease of heterozygosity with distance from Ethiopia (Prugnolle et al. 2005, Liu et al. 2006). Under this scenario one might also expect to find that southern populations are derived with respect to northern ones. Alternatively, if Pleistocene glacial maxima affected northern populations through extinctions, range expansions, and/or bottlenecks, we would expect to observe the opposite pattern, that northern individuals will be derived and nested within a clade of southern individuals, and that latitude will be significantly and inversely correlated with heterozygosity (Hewitt 1996, Excoffier et al. 2009). Finally, the simultaneous vicariance hypothesis can be tested in a comparative statistical framework by comparing divergence dates estimated between various clades. Following (Smith et al. 2014a,b), if the 95% highest posterior densities (HPDs) of the estimated divergence times for a given pair of co-distributed sister lineages do not overlap, the simultaneous divergence hypothesis can be rejected for that break.

Here, I use genome-wide DNA sequence and SNP data collected using RADseq to test these alternative hypotheses using lizards (family Phrynosomatidae) in the BCP. Phrynosomatids are ideal organisms to study the historical biogeography of the BCP for three reasons: 1) they are abundant, ecologically diverse, and widespread throughout the BCP and its offshore islands, 2) they are relatively well-studied, as detailed morphological monographs and mitochondrial data have been published for most species, and 3) they often show spatially concordant phylogeographic or morphological

breaks in the same regions. I focus on peninsular lineages of four genera (Figure 3.2): zebra-tailed lizards (*Callisaurus*), brush lizards (*Urosaurus*), banded rock lizards (*Petrosaurus*), and spiny lizards (*Sceloporus*). Within each genus numerous lineages have been described, although their status as species or subspecies is often under contention. My working list of operational taxonomic units (OTUs) based on the published literature is shown in Table 3.1. Unlike the previous chapter, it is outside the scope of this chapter to delimit species boundaries or propose formal taxonomic revisions (although I do advise where future work may be needed in this area). Rather, I use a comparative phylogeographic approach to address four questions:

1. Are the phylogeographic patterns revealed by analysis of genome-wide sequence and SNP data (using both population genetic and phylogenetic methods) congruent with the patterns previously discovered using morphology, allozymes, and mtDNA?
2. For each co-distributed lineage, do genome-wide patterns of heterozygosity reject the hypotheses that genetic diversity is reduced in insular populations (presumably due to founder effects and/or bottlenecks), in northern populations (due to recent post-glacial range expansions from southern refugia), or in southern populations (due to dispersal from a continental population)?
3. Did the various lineages of phrynosomatid lizards included in this study diverge across the same areas at the same times (simultaneous

divergence)? Or, alternatively, did the divergences occur in multiple temporal waves?

4. Can divergence times among the various lineages be linked to specific geological and/or climatic events?

II. MATERIALS AND METHODS

(1) Field sampling

Specimens and tissue samples were collected in the United States (southern California) and Mexico (Baja California, Baja California Sur, Sonora) between 1996 and 2014. Liver and tail tip tissues were preserved in 96-100% ethanol and stored at -20° C at San Diego State University (SDSU). All voucher specimens associated with this study have been deposited and accessioned at the SDNHM. I included 61 *Callisaurus*, 40 *Petrosaurus*, 47 *Urosaurus*, and 80 *Sceloporus* (228 total) spanning the BCP and eight offshore islands.

(2) RAD library preparation

I collected genome-wide DNA sequence/SNP data using the double-digest Restriction-Associated-DNA sequencing (ddRADseq) protocol (Peterson et al. 2012) using the enzymes and size selection window of Leaché et al. (2014, 2015). Following the manufacturer's instructions, I used the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) with RNAse A to extract genomic DNA (gDNA). DNA concentrations were measured using a Qubit 2.0 Fluorometer

(Life Technologies, Grand Island, NY) and assessed for high molecular weight using 1% agarose gel electrophoresis. The high-fidelity restriction enzymes *SbfI* and *MspI* (New England Biolabs, Ipswich, MA) were used to digest 200-500 ng of gDNA per sample and the digestions were purified using Agencourt AMPure beads (Beckman Coulter, Danvers, MA) before I attached uniquely bar-coded adapters to each library with T4 Ligase (New England Biolabs, Ipswich, MA). After a second AMPure cleanup, samples were combined to create pools of eight uniquely bar-coded libraries at equimolar concentrations. I then selected fragments 415-515 bp long using a Pippen Prep (Sage Science, Beverly, MA) with 2% gel cassettes. I used Phusion *Taq* Polymerase (New England Biolabs) for the Polymerase Chain Reaction (PCR) with a two-step cycle (98° C for 10 seconds, 72° C for 20 seconds) 12 times, followed by a final extension step of 72° C for 10 minutes. An Agilent Bioanalyzer 2100 was used to ensure that the libraries were at the appropriate concentration and size distribution for sequencing. The final libraries (100 bp single end reads) were sequenced on five different flow-cell lanes of a HiSeq 2500 (Illumina, San Diego, CA) at the Institute of Integrative Genome Biology (University of California, Riverside).

(3) Bioinformatics

I used the python pipeline pyRAD v2.1.2 for quality control, alignment, and genotype calling (Eaton and Ree 2013, Eaton 2014, Leaché et al. 2015). The pipeline followed seven steps: 1) raw reads were de-multiplexed by barcode; 2) restriction cut sites and adapter sequences were trimmed, bases with Phred

scores less than 20 were replaced with 'N', and reads with >10 Ns were removed; 3) reads were clustered into putative loci within individuals using USEARCH v7.0.1090 (Edgar 2010) using a clustering threshold of 0.85, retaining only loci with at least 10x coverage and with 10 or fewer Ns; 4) heterozygosity and error rates were jointly estimated for each individual; 5) consensus base calling was performed using the mean error and heterozygosity rates estimated during step four, and consensus sequences with >5 heterozygous sites were excluded as potential paralogs; 6) consensus sequences were clustered across individuals using MUSCLE v3.8.31 (Edgar 2004) following a 0.85 clustering threshold; 7) loci in which > 50% of samples had shared heterozygosity at a site were filtered (another paralog filter), and finally, alignments of both full loci (89 bp) and variable sites (individual SNPs) were generated for further analysis.

Treatment of missing data in RADseq analyses is still a problematic endeavor, as there is an inverse relationship between the final number of loci recovered in the alignment and the minimum number of individuals required for a locus to be included in the dataset. Low tolerance for missing data may bias the mutational rate spectrum of sampled loci, so that slowly evolving loci are disproportionately represented in the final matrix (Huang and Knowles 2014). Excessive levels of missing data in larger matrices sometimes leads to spurious results from clustering algorithms (Rheindt et al. 2014), wherein individuals with high levels of missing data may erroneously cluster as separate populations. However, concatenated phylogenetic analyses have also been shown to benefit from longer alignments even with a higher percentage of missing data (e.g.,

Rokas et al. 2003, Wagner et al. 2013). Given this uncertainty, I arbitrarily chose a cut-off of 1,000 or more loci as my minimum matrix size. Thus, four independent datasets were created, one for each genus: *Callisaurus* (1,027 loci, with all loci present in a minimum of 47/61 individuals), *Petrosaurus* (1,018 loci, minimum 26/40), *Urosaurus* (1,062 loci, minimum 22/47) and *Sceloporus* (1,024 loci, minimum 43/80).

(4) Concatenated phylogenetic inference

I used BEAST v2.1.2 (Bouckaert et al. 2014) to estimate a time-calibrated phylogeny of the concatenated data for each genus with a uniform Yule model tree prior, a relaxed lognormal clock, and the GTR+GAMMA substitution model with empirical base frequencies using the supermatrix approach (de Queiroz and Gatesy 2007), as it can readily handle the missing data problem inherent to RAD sequencing. By concatenating unlinked loci into a single matrix and analyzing it as a single "supergene," it is assumed the divergence times obtained are reflective of the actual divergence times between species. In recent years species tree methods have been developed to address this issue using SNP data (Bryant et al. 2012), however, *BEAST and SNAPP are not able to handle missing data at the species level, and pruning my datasets resulted in very few SNPs conserved across early Pliocene or late Miocene divergences, particularly for *Petrosaurus* (47 SNPs) and *Sceloporus* (7 SNPs), so these analyses are not included here. Calibrating the molecular clock is also problematic because all loci are anonymous and no useful fossil calibrations are currently available for

peninsular phrynosomatids. Therefore, I assumed an average mutation rate of 2.2×10^{-9} subs/site/year for all taxa based on the average genome-wide mammalian mutation rate (Kumar and Subramanian 2002) and compared the divergence dates to several geological calibrations post-hoc. This rate is close to that published for avian introns (2.56×10^{-9} subs/site/year; Lee and Edwards 2008). Fortunately, several geological calibrations are available, including the age of the BCP and of its islands that would provide upper bounds for divergence dates of species or populations associated with these geographic features (but not lower bounds, as dispersal across these boundaries cannot be ruled out). Another implicit assumption is that the mutation rate is the same across genera, despite possible differences in generation time and other important life history parameters. This might seem like a reasonable assumption given that all the lizards examined here belong to the same family (Phrynosomatidae), but this hypothesis cannot be tested directly with the data at hand due to locus dropout. The Markov chains were run for 50 million generations each, sampling trees every 1,000 steps. Convergence and the length of burn-in were assessed using Tracer v1.6 (Rambaut et al. 2014). I discarded the first 50% of trees for *Callisaurus* and *Petrosaurus*, the first 20% for *Urosaurus*, and the first 10% for *Sceloporus*. Effective sample size (ESS) estimates were then re-assessed with Tracer to ensure these were greater than 200 for the posterior and likelihood distributions before using TreeAnnotator v1.8.1 (Bouckaert et al. 2014) to produce a maximum clade credibility tree from the posterior distribution with mean node heights for each genus. A maximum-likelihood (ML) phylogeny of the concatenated data was

also constructed separately for each genus with RAxML v8.1.1 (Stamatakis 2014) HPC BlackBox on the CIPRES Science Gateway (Miller et al. 2015). The GTR+GAMMA substitution model using empirical base frequencies was implemented, rapid bootstrap replicates were performed (letting RAxML halt bootstrapping automatically as recommended), and the trees were rooted based on the root locations inferred using BEAST.

(5) Population genetic structure

In order to test whether the RADseq data support the existence of my ML inferred phylogeographic units, as well as previously defined units, and to validate the units of comparison in my comparative divergence dating analyses, two different clustering algorithms were used. Because these methods require multiple unlinked loci, I used a custom R script to select the most informative biallelic SNP for each locus (defined as the SNP with the least amount of missing data). If there was a tie for the most informative SNP, the SNP was chosen at random. I performed a principal components analysis (PCA) using smartPCA v1.02.10 (Patterson et al. 2006, Reich et al. 2008) for each dataset separately. The main benefits of PCA are its ability to detect genetic structure without assumptions about the underlying population genetic model and its computational speed (Jombart et al. 2010), but it does not directly infer clusters. Therefore, Admixture v1.23 (Alexander et al. 2009) was also used to detect population structure among identified clusters, analyzing the same data matrices as in the smartPCA analysis. This ML program implements the same underlying

population genetic model as the widely used Bayesian program Structure (Pritchard et al. 2000). Both methods assign individuals to clusters in an attempt to minimize linkage disequilibrium and maximize Hardy-Weinberg equilibrium within clusters. The primary advantage of Admixture is its fast numerical optimization algorithm to decrease computational time while avoiding problems with MCMC convergence, which is especially useful for large SNP datasets. I used the cross-validation procedure to select the optimal K value ranging from 2 to 9 (any lower would not be useful for detecting structure, any higher would be difficult to visualize in a single plot) and to estimate pair-wise F_{ST} values between clusters and used a custom R script to plot the results (Alexander et al. 2009).

(6) Latitude vs. heterozygosity

In order to test for possible population bottlenecks or geographic range expansions, I used the proportion of heterozygous sites across all reads for a given individual (as output by pyRAD) as a metric to explore the impact of latitude. I interpret a significant positive correlation of latitude and heterozygosity to reject the hypothesis of a recent northward range expansion or a bottleneck (Nason et al. 2002, Clark-Tapia and Molina-Freaner 2003, Prugnolle et al. 2005, Liu et al. 2006), and conversely, I interpret a negative significant correlation to reject the southward dispersal hypothesis (Savage 1960). I interpret a lack of statistical significance to mean that neither hypothesis can be rejected. In order to detect possible range expansions related to Pleistocene climatic cycles, I used the lm function in R to perform linear regressions of latitude vs. the

proportion of heterozygous sites within each genus and across genera, including and excluding insular samples. Because heterozygosity may be correlated with phylogeny, I also fit a linear model to the data for each genus using phylogenetic generalized least squares (PGLS) under the Brownian motion model of evolution (Pinheiro and Bates 2000) using the ultrametric tree produced with BEAST for each genus to produce the phylogenetic correlation matrix.

III. RESULTS

(1) RAD summary statistics

Summary statistics for the RADseq data are presented in Table 3.2. In total, I analyzed >255 million raw reads obtained from the Illumina HiSeq 2500 (average of 1.1 million reads per individual) after demultiplexing. Of these, a total of 241 million (average 1.1 million per lizard) passed the initial trimming and quality filter, and were aligned using a clustering threshold of 0.85 into an average of 31,455 clusters per lizard at an average depth of 35x using USEARCH. An average of 5,456 clusters passed the minimum coverage threshold (10x), and the remaining clusters had a mean coverage of 198x. After aligning clusters across individuals using MUSCLE, a mean of 4,977 loci per lizard remained. As described in the methods, concatenated alignments containing the minimum number of individuals necessary to obtain at least 1,000 loci for *Callisaurus* (95,568 bp matrix containing 7,218 variable sites), *Petrosaurus* (94,615 bp matrix containing 5,930 variable sites), *Urosaurus* (99,155 bp matrix containing 9,488

variable sites), and *Sceloporus* (95,447 bp matrix containing 11,014 variable sites) were constructed. For the Admixture and PCA analyses, selecting one SNP per locus and filtering for biallelic SNPs resulted in 994 unlinked SNPs for *Callisaurus*, 932 for *Petrosaurus*, 1036 for *Urosaurus*, and 1015 for *Sceloporus*.

(2) Heterozygosity analyses

Using a standard linear regression, a significant positive correlation between latitude and the proportion of heterozygous sites was found for *Callisaurus* (Figure 3.3), while significant negative correlations were found for *Petrosaurus* (Figure 3.4) and *Urosaurus* (Figure 3.5). However, no significant correlations were found for *Sceloporus* (Figure 3.6). Excluding insular samples improved the R-squared and *p* values for *Callisaurus*, *Petrosaurus*, and *Urosaurus*, but not for *Sceloporus*. I also analyzed the two main sister clades within *Sceloporus* separately (*S. magister* + *S. zosteromus* [magister complex] vs. *S. orcutti* + *S. licki* + *S. hunsakeri* [orcutti complex]) with and without insular samples, but I still did not detect any significant correlations (results not shown). After correcting for phylogeny using PGLS under the Brownian motion model, a significant inverse relationship between latitude and heterozygosity was found in *Petrosaurus* and *Urosaurus*, a significant negative relationship was found in *Sceloporus*, and no significant relationship was found in *Callisaurus* (Table 3.3).

(3) *Callisaurus* phylogeography and population structure

The RAxML analysis produced a phylogram with seven well-supported clades (bootstrap values ≥0.70), six of which neatly correspond with previously described subspecies in Table 3.1 (Figure 3.7): *C. d. rhodostictus* from San Gorgonio Pass to Bahía de los Ángeles, *C. d. splendidus* from Isla Ángel de la Guarda, *C. d. crinitus* of the Vizcaíno Desert, *C. d. carmenensis* from south of the Sierra San Francisco to Isla Carmen, *C. d. plasticus* from Isla Danzante to the Isthmus of La Paz (including Isla San Francisco), and *C. d. draconoides* from the cape region (including the Sierra de la Laguna, Sierra de la Trinidad and Isla Espiritu Santo). The seventh clade consists of the "Central" population east of the Sierra de la Libertad/Vizcaíno Desert and north of the Sierra San Francisco, hypothesized by Grismer (2002) to be an intergrade between *C. d. rhodostictus*, *C. d. crinitus*, and *C. d. carmenensis* based on morphological intermediacy. These seven clades were also strongly supported (posterior probability ≥0.95) in the maximum clade credibility tree produced by BEAST. The results of the Admixture analysis found that K=7 was the best supported model (Table 3.4), and these seven clusters nearly perfectly match the seven clades, with the exception of one individual (ADG126) from just west of Bahía de los Ángeles, which was shown to belong to *C. d. rhodostictus* in the phylogenies, but demonstrates possible admixture between *C. d. rhodostictus* and the Central population (Figure 3.8). The fact that this individual also has the highest proportion of heterozygous sites of any *Callisaurus* (Figure 3.3) and was collected from a narrow pass at the northern end of the eastern escarpment of the Sierra de

la Libertad (Figures 3.1, 3.7) further suggest that this individual maybe a hybrid. F_{ST} values among the seven clusters ranged from 0.41 between the Central population and *C. d. rhodostictus* to 0.78 between *C. d. draconoides* and *C. d. splendidus,* indicating a substantial amount of population subdivision along the BCP (Table 3.5). The initial divergence between the northern and southern clades of *Callisaurus,* which currently contact in the vicinity of Loreto (F_{ST} = 0.58 between *C. d. plasticus* and *C. d. carmenensis*), was estimated to have occurred in the late Pliocene. Mean divergence time estimates indicate that most diversification within *Callisaurus* in the BCP occurred in the late Pliocene or early Pleistocene, including the split between *C. d. splendidus* on Isla Ángel de la Guarda and *C. d. rhodostictus* (F_{ST} = 0.47). The populations on Isla Espiritu Santo, Isla San Francisco, Isla Danzante, and Isla Carmen all originated during the mid-late Pleistocene, but were not recognized as distinct clusters by the Admixture analysis. The PCA plots (Figure 3.9) were clearly able to differentiate the seven clades using the first four principal components (PCs) – the plot of PC1 vs. PC2 revealed six clades laid out in latitudinal order, with the seventh (*C. d. draconoides*) positioned off the axis, although the distributions of *C. d. carmenensis* and *C. d. crinitus* overlapped, but the plot of PC3 vs. PC4 was able to discriminate between these clades.

(4) *Petrosaurus* phylogeography and population structure

The RAxML tree (Figure 3.10) and the BEAST tree (Figure 3.11) for *Petrosaurus* revealed five strongly supported clades that were also identified by

Admixture: 1) a northern clade of *P. mearnsi* that ranges from San Gorgonio Pass to the east slope of the Sierra San Pedro Martír, 2) a southern clade of *P. mearnsi* ranging from Cataviña to Bahía de los Ángeles (including *P. slevini* from Isla Ángel de la Guarda), 3) a northern clade of *P. repens* found northwest of Bahía de los Ángeles, 4) a central clade of *P. repens* found in the Sierra San Francisco, Sierra Guadalupe, and northern Sierra la Giganta, and 5) a southern clade of *P. repens* from the southern Sierra la Giganta and the Isthmus of La Paz. *Petrosaurus thalassinus* from the cape region (Isla Partida to Cabo San Lucas) was also strongly supported as a clade in the BEAST and RAxML trees, but was split into two non-geographic clusters by Admixture under the preferred model (K=7; Table 3.4), although these clusters do not match any phylogenetic pattern, suggesting this model may be over-splitting. F_{ST} values ranged from 0.41 between northern *P. mearnsi* and northern *P. repens* to 0.78 between southern *P. mearnsi* and southern *P. repens* (Table 3.6). Mean divergence time estimates indicate that the northern and southern clades of *P. mearnsi/slevini* split from *P. repens/thalassinus* at the Miocene/Pliocene boundary. The split between *P. repens* and *P. thalassinus* across the Isthmus of La Paz in the early Pleistocene was coincident with the split between northern and southern clades of *P. mearnsi.* In the late Pleistocene, northern, central, and southern clades of *P. repens* diverged from each other, *P. slevini* on Isla Ángel de la Guarda diverged from the southern clade of *P. mearnsi*, and the Isla Partida population split from the rest of the cape region. The PCA plots were able to distinguish *P. mearnsi*, *P. repens*, and *P. thalassinus* along the first two PC axes, while the third PC axis separated the

northern and southern clades of *P. mearnsi*, and the fourth PC axis accounts for variation within these latter clades (Figure 3.12).

(5) *Urosaurus* phylogeography and population structure

Urosaurus lahtelai, restricted to the vicinity of Cataviña, and *Urosaurus nigricaudus,* ranging from the Magdalena Plain to Cabo San Lucas, were the only clades identified in both of the phylogenetic analyses and the Admixture analysis under the preferred model (K=4) (Figures 3.13, 3.14). The northern clade of *U. nigricaudus* from the Magdalena Plain and Isla Santa Margarita was recovered in both phylogenetic analyses but not by Admixture, while the southern clade from the Isthmus of La Paz, Isla Espiritu Santo, Isla Partida, and the cape region was strongly supported by the BEAST analysis only. Within *U. microscutatus*, which was found to be monophyletic in both phylogenetic analyses, Admixture detected a northern cluster ranging from the Sierra de la Libertad to southern California and a southern cluster found south of the Sierra San Francisco to the Sierra la Giganta, with three geographically intermediate individuals from the Sierra San Francisco showing limited admixture between the two clusters (Figure 3.14). In both phylogenies, the northern cluster was found to be monophyletic but the southern cluster was found to be paraphyletic with respect to the northern cluster. F_{ST} values ranged from 0.41 between northern and southern clusters of *U. microscutatus* to 0.88 between *U. lahtelai* and northern U. *microscutatus* (Table 3.7). Mean divergence time estimates between *U. lahtelai*, *U. nigricaudus*, and *U. microscutatus* lie in the late Miocene, and the divergence of

northern from southern clades of *U. nigricaudus* most likely occurred in the mid-Pliocene. Meanwhile, the origins of northern *U. microscutatus* and insular populations of *U. nigricaudus* on Isla Espiritu Santo, Isla Partida, and Isla Santa Margarita did not occur until the mid-Pleistocene. The first PC axis separated *U. lahtelai* from the remaining species, the second PC axis separated *U. nigricaudus* from *U. microscutatus*, the third PC axis separated northern and southern clusters of *U. microscutatus*, and the fourth PC axis divided northern and southern clusters of *U. nigricaudus* (Figure 3.15).

(6) *Sceloporus* phylogeography and population structure

For ease of viewing, the RAxML tree for *Sceloporus* was divided into two figures for each of the two co-distributed sister lineages: *S. orcutti*, *S. licki*, and *S. hunsakeri* (Figure 3.16), and *S. magister* and *S. zosteromus* (Figure 3.17). This tree revealed seven strongly supported clades that were also identified by BEAST and Admixture (Figure 3.18): 1) a northern clade of *S. orcutti* that ranges from San Gorgonio Pass to Cataviña, 2) a southern clade of *S. orcutti* that ranges from the Sierra San Francisco to the Sierra la Giganta, 3) *S. licki* from the Sierra de la Laguna and Sierra de la Trinidad, 4) *S. hunsakeri* from the cape region (Isla Partida to Cabo San Lucas), 5) *S. magister* east of San Gorgonio Pass, Sierra Juarez, Paso San Matías, and the Sierra San Pedro Martír and north of the GoC, 6) a northern clade of *S. zosteromus* including *S. z. rufidorsum* from the Sierra Juarez to the Sierra de la Libertad and a putative intergrade population from the Vizcaíno Desert (Grismer 2002), and 7) a southern clade of *S. zosteromus*

extending from south of Laguna San Ignacio to Cabo San Lucas, including *S. z. monserratensis*, *S. z. zosteromus*, and a hypothesized intergrade zone (based on morphological intermediacy of specimens) from the Isthmus of La Paz (Grismer 2002). Pairwise F_{ST} values between clusters ranged from 0.52 between northern and southern clusters of *S. zosteromus* to 0.83 between northern *S. zosteromus* and *S. hunsakeri* (Table 3.8). Mean divergence time estimates indicate that the initial divergence of the *S. zosteromus/magister* clade from the *S. orcutti/hunsakeri/licki* clade most likely occurred in the late Miocene. *Sceloporus zosteromus* and *S. licki* originated in the early Pliocene, and *S. hunsakeri* originated in the mid-Pliocene. Northern and southern clades of *S. zosteromus* diverged in the late Pliocene (F_{ST} =0.52), and northern and southern clades of *S. orcutti* split in the early Pleistocene (F_{ST} = 0.53). Finally, insular populations of *S. orcutti* (Isla Tortuga) and *S. hunsakeri* (Isla Partida) originated in the mid and late Pleistocene, respectively. The first PC axis separated *S. hunsakeri* from the remaining species, the second axis distinguished *S. zosteromus* and *S. magister*, from *S. orcutti* and *S. licki*, the third axis separated *S. licki* from *S. orcutti*, and the fourth axis distinguished *S. magister* from all remaining groups (Figure 3.19). However, northern and southern clades of *S. zosteromus* and *S. orcutti* were not distinguishable using the first four PC axes.

IV. DISCUSSION

(1) Comparison with previously published studies

In this study, I used genome-wide RADseq data to examine the comparative phylogeography of phrynosomatid lizards of the BCP. The first major question I addressed was: are the phylogeographic patterns revealed by analysis of genome-wide sequence and SNP data (using both population genetic and phylogenetic methods) congruent with the patterns previously discovered using morphology, allozymes, and mtDNA? Adest (1987) used allozymes to study geographic variation within *C. draconoides* and to assess divergence among the various subspecies. He found that *Callisaurus* exhibits little genetic differentiation over a vast geographic area – mainland samples were essentially undifferentiated, except for samples from Loreto, Baja California Sur, and Mazatlan, Sinaloa, and concluded that this genus is best treated as monotypic. Grismer (2002) also considered *C. draconoides* to be a single species based on qualitative morphological descriptions, although he recognized the five subspecies within the BCP as diagnosable "pattern classes" connected by two intergrade zones. The divergent seven clades/clusters revealed by this study (Figures 3.8, 3.9) are a near perfect fit to the map in Grismer (2002, p. 137). The central clade recovered in my analyses corresponds to an intergrade zone in this map, and *C. d. plasticus* from my analyses nearly matches his southern intergrade between *C. d. draconoides* and *C. d. carmenensis*, except that the range of *C. d. plasticus* extends farther north (to Isla Danzante) than that of his intergrade zone,

confined to the Isthmus of La Paz. Grismer (2002) did not recognize *C. d. plasticus* because the type locality of *C. d. plasticus* (Agua Verde Bay) is included in the range of *C. d. carmenensis* in his map (p. 137). Most recently, Lindell et al. (2005) used mtDNA to examine the phylogeography of *C. draconoides* across its range. These authors also concluded that *Callisaurus* is a single species, highlighting three major genetic breaks within the BCP: 1) across the Vizcaíno Desert (the "mid-peninsular" break), inferred to be late Miocene/early Pliocene in age, 2) a late Pliocene break in the vicinity of Loreto, and 3) a late Pliocene/early Pleistocene break across the Isthmus of La Paz. Although Lindell et al. (2005) concluded that "specimens from the Vizcaíno Desert are not genetically isolated" a closer examination of their tree reveals that the two samples of *C. d. crinitus* from the Vizcaino Desert are strongly supported as monophyletic and sister to a monophyletic *C. d. carmenensis*. Finally, Lindell et al. (2005) revealed that the one individual from Isla Ángel de la Guarda was sister to *C. d. rhodostictus* from the northern BCP and the Mojave/Sonoran deserts, and assumed that *C. d. splendidus* originated through vicariance 2-3 mya. All three of these breaks are geographically congruent with the findings of my study, and the timing of all but the mid-peninsular break (where I found evidence for an early Pleistocene break) matches my study as well. However, Lindell et al. (2005) did not detect the central lineage due to sparse sampling.

The most recent molecular genetic analysis of *Petrosaurus* used allozymes (Aguilars-S. et al. 1988), and there has been no phylogeographic studies using mtDNA or nDNA. Aguilars-S. et al. (1988) confirmed the distinctiveness of *P.*

mearnsi (including P. *slevini*) from *P. thalassinus* (*including P. repens*), which corresponds to the deepest split (late Miocene) in my *Petrosaurus* trees (Figures 3.11, 3.12), but they found little support for any structure beyond that, concluding that *P. thalassinus* demonstrates a "high gene-flow pattern characteristic of more two-dimensional geographic ranges." This latter conclusion was not supported by my data: the Admixture analysis found three discrete clusters within *P. repens* that correspond to northern, central, and southern clades (Figure 3.12). As demonstrated with *Callisaurus*, I suspect that this reflects the much greater power of DNA sequence data, especially multi-locus RADseq data, to resolve phylogeographic structure compared to allozyme data. Based on qualitative morphological diagnoses, Grismer (2002) recognized *P. slevini* as a distinct species from *P. mearnsi*. My phylogenetic analyses found that *P. slevini* is nested within *P. mearnsi,* and the clustering analyses reflected this – the northern clade/cluster ranges from San Gorgonio Pass to the Sierra San Pedro Martír, including the type locality of *P. mearnsi* (Mountain Springs), while the southern clade/cluster includes *P. mearnsi* from Cataviña and Bahía de los Ángeles and *P. slevini*. Future work could explicitly test this alternative species delimitation hypothesis in a coalescent framework, but this is beyond the scope of this chapter. Grismer (2002) also split *P. repens* from *P. thalassinus*; my results are consistent with the map of Grismer (2002) showing *P. repens* and *P. thalassinus* divided at the Isthmus of La Paz.

The phylogeography of peninsular *Urosaurus* was first assessed by Aguirre-León et al. (1999) who used allozyme data to demonstrate support for a

mid-peninsular break east of the Vizcaino Desert (mid-Pleistocene in age) and an older Pliocene break across the Isthmus of La Paz. Both of these breaks, and their inferred timing, are consistent with my results. Aguirre-León et al. (1999) tentatively recognized *U. lahtelai* pending further investigation but recommended that *U. microscutatus* be synonymized with *U. nigricaudus* because patterns of allelic distribution indicate ongoing gene flow across the Vizcaíno Desert. Grismer (2002) also followed this taxonomy and did not recognize any morphological pattern classes within *U. nigricaudus*. Although my Admixture analysis found support for some introgression from *U. microscutatus* in the northernmost individual of *U. nigricaudus,* overall these two lineages were found to be reciprocally monophyletic and deeply divergent (late Miocene) in my phylogenetic analyses (Figures 3.14, 3.15). Lindell et al. (2008) analyzed mtDNA to re-assess phylogeographic structure within peninsular *Urosaurus*, and although they did not find that either *U. microscutatus* or *U. nigricaudus* were monophyletic, unlike the current study, they did recover four phylogeographic breaks: 1) in the Vizcaíno Desert (late Miocene), 2) in the Loreto region (Pliocene), 3) across the Isthmus of La Paz (late Miocene), and 4) north of Cabo San Lucas (late Miocene/early Pliocene). My study found support for the first three breaks, but not the fourth. Feldman et al. (2011) analyzed mtDNA and three nuclear loci and found that *U. nigricaudus* renders *U. microscutatus* paraphyletic, not a result supported by my study, but their study only included one sample of *U. nigricaudus* and two samples of *U. microscutatus*, and their locality descriptions are vague, making it difficult to compare results directly.

However, their range map showing the distribution of *U. microscutatus* contacting *U. nigricaudus* in a diagonal line along the western slope of the Sierra la Giganta perfectly matches the phylogeographic pattern recovered here (Figure 3.14). In summary, my results confirm earlier hypotheses that phylogeographic breaks exist in the mid-peninsula, Loreto, and the Isthmus of La Paz, although my inferred divergence times for these events are more consistent with Aguirre-León et al. (1999) than with Lindell et al. (2008). Finally, the reciprocal monophyly of *U. microscutatus* and *U. nigricaudus* is a novel result, so a future coalescent-based species delimitation study is needed to explicitly test and evaluate the species limits within the *U. nigricaudus* complex.

Peninsular *Sceloporus* have had the most comprehensive systematic work completed of any studied in this chapter, and numerous previously published hypotheses regarding species limits, taxonomy, phylogeny, and divergence are now available to test. Following Leaché and Mulcahy (2007), two sister complexes are co-distributed along the length of the BCP: the *orcutti* complex consisting of *S. orcutti*, *S. licki*, and *S. hunsakeri* (Figure 3.17) and the *magister* complex consisting of the sister clades *S. zosteromus* and *S. magister* (Figure 3.18). Murphy (1983a) hypothesized that the *magister* complex split from the *orcutti* complex 5–10 mya. Within the BCP, he recognized three endemic species of the *magister* complex (*S. rufidorsum*, *S. monserratensis*, and *S. zosteromus*) and hypothesized that these three taxa diverged from *S. magister* in the early Pliocene (5 mya). Within the *orcutti* complex, he hypothesized that *S. licki* and *S. hunsakeri* were sister lineages that diverged from *S. orcutti* during the late Miocene (5.2–6.6

mya). Welsh (1988) and Schulte et al. (2006) hypothesized that *S. zosteromus* and *S. magister* diverged during the Miocene (6–14 and 9.8 mya, respectively). Grismer (1994) hypothesized much younger divergence dates: 1–3 mya for the *S. zosteromus/magister* divergence and 1.6–3 mya for the divergence of *S. orcutti* from *S. licki/hunsakeri*. Grismer and McGuire (1996) analyzed karyotypes, allozymes, color pattern, and distribution and found that the three endemic peninsular taxa form a single, continuously distributed, interbreeding population (*S. zosteromus*) sister to *S. magister* which ranges into the Colorado Desert region (the two taxa are parapatric at Paso San Matías), although Grismer (2002) did recognize these taxa as morphological pattern classes. Leaché and Mulcahy (2007) completed the most detailed analysis of this group to date, analyzing mtDNA and four nuclear loci. Within the *magister* complex, they found evidence for a mid-peninsular break at the Vizcaíno Desert roughly corresponding to the southern limit of *S. z. rufidorsum*, and within the *orcutti* complex, they found strong support for a sister relationship of *S. orcutti* and *S. hunsakeri* with respect to *S. licki*. Both of these novel relationships are also supported by my study (Figures 3.17-3.19). Utilizing the geological age of Isla Monserrat as a calibration point, Leaché and Mulcahy (2007) estimated that the *orcutti* and *magister* complexes diverged 6.48 mya, *S. magister* split from *S. zosteromus* 5.57 mya, northern and southern clades of *S. zosteromus* split 2.97 mya, *S. licki* split from *S. orcutti/hunsakeri* 4.23 mya, and *S. hunsakeri* diverged from *S. orcutti* 2.65 mya. In all but the last case, where I found evidence for an early Pliocene divergence, these mean estimates fall within the 95% HPDs of this study (Figure 3.19). My 95% HPDs do reject the older Miocene

(> 6 mya) divergence hypothesis (Welsh 1988, Schulte et al. 2006) and the younger Pleistocene/Pliocene (1-3 mya) divergence hypothesis for the divergence of *S. magister* from *S. zosteromus* (Grismer 1994). The other key distinction of my study from that of Leaché and Mulcahy (2007) is that they did not recover support for two distinct northern and southern clades of *S. orcutti*; this appears to be a novel finding to my study, although unpublished mtDNA data collected by Reeder et al. has suggested this relationship before.

(2) Heterozygosity as a function of latitude

The second question I asked was: do genome-wide patterns of heterozygosity support the prediction that genetic diversity should be reduced in northern populations due to range expansions or bottlenecks (Hewitt 1996) or in southern populations due to dispersal from an ancestral continental population (Savage 1960)? Standard linear regressions revealed significant inverse correlations between heterozygosity and latitude in *Petrosaurus* and *Urosaurus*, and excluding insular samples improved the R-squared values. However, no pattern at all was detected *in Sceloporus*, and the opposite pattern was detected in *Callisaurus* (latitude was positively correlated with heterozygosity). After correcting for phylogeny using PGLS, the patterns for *Urosaurus* and *Petrosaurus* were still significant, while the pattern for *Callisaurus* was not. I interpret these results to reject the Savage (1960) hypothesis for *Urosaurus* and *Petrosaurus*. The significant inverse correlations shown by these two genera are instead consistent with recent northward range expansions or population bottlenecks (Hewitt 1996,

Excoffier et al. 2009), especially for *Urosaurus*, considering that the deeply nested northern clade of *U. microscutatus* is of Pleistocene age and renders southern *U. microscutatus* paraphyletic. This pattern has also been documented in co-distributed cacti based on linear regressions of heterozygosity with latitude, and the authors likewise interpreted their results to support northward range expansions related to postglacial cycles (Nason et al. 2002; Clark-Tapia and Molina-Freaner 2003). I do not have an explanation for why *Sceloporus* does not show a significant correlation between latitude and heterozygosity in the standard regression. However, there is a significant different in the PGLS test, but it appears that the variance within species at any given latitude is quite large.

(3) Comparative divergence dating

The third question I addressed was, did the various lineages of phrynosomatid lizards included in this study diverge across the same areas at the same times (Lindell et al. 2006)? Or, alternatively, did divergence occur in multiple waves for each area (Leaché et al. 2007)? Bearing in mind the assumption of equal mutation rate across genera, I reject the null hypothesis of simultaneous divergence if the 95% highest posterior density (HPD) of divergence times for any given set of taxa do not overlap. In Table 3.9, I report selected divergence times obtained from my BEAST trees for each co-distributed lineage across five different geographic regions where two or more taxa showed spatially concordant breaks.

Starting with the Vizcaino Desert, at least three episodes of lineage divergence are required to explain the results – the youngest divergence was found in *P. repens* in the late Pleistocene (0.55-0.87 mya) which did not overlap with the 95% HPD for the next youngest divergence in *U. microscutatus* (1.33-1.99 mya) which in turn did not overlap with that of the oldest divergence seen in *S. zosteromus* (2.10-3.69 mya). The divergence of northern and southern *S. orcutti* (1.48-2.42 mya) overlapped with that of *S. zosteromus,* so I cannot reject the null hypothesis that they diverged simultaneously sometime between 2.10-2.42 mya. Likewise, because the 95% HPD for *Callisaurus* (1.62-2.71 mya) overlaps with that of *Urosaurus*, I cannot reject simultaneous divergence of the two genera from 1.62-1.99 mya, nor can I reject simultaneous divergence of *Callisaurus* and *S. orcutti* from 1.62-2.42 mya, nor of *Callisaurus* and *S. zosteromus* from 2.10-2.71 mya. There are two periods of time in which I cannot reject simultaneous divergence across the Vizcaíno Desert for three or more lineages. First, I cannot reject simultaneous divergence of *Callisaurus* (central vs. *C. d. carmenensis*), *S. orcutti* (north vs. south), and *S. zosteromus* (north vs. south) from 2.1-2.42 mya. Likewise, I cannot reject simultaneous divergence of *Callisaurus*, *U. microscutatus* (north vs. south), and *S. orcutti* between 1.62-1.99 mya.

In the Loreto/Sierra de la Giganta region, I can definitively reject the hypothesis of simultaneous divergence among all of the three taxa that show breaks in that region. None of these 95% HPDs overlap, requiring three separate episodes of divergence. *Petrosaurus repens* shows the youngest break in the late Pleistocene (0.66-1.05 mya) while *C. d. carmenensis/plasticus* diverged in the late

Pliocene (2.50-3.54 mya) and *U. microscutatus/nigricaudus* diverged in the early Pliocene to late Miocene (4.50-7.41 mya). No *Sceloporus* showed a break here so they were not included for comparison.

Across the Isthmus of La Paz, the youngest divergence was again seen between *P. repens* and *P. thalassinus* in the early-mid Pleistocene (1.26-2.22 mya) while the oldest was observed between *S. licki* and *S. hunsakeri/orcutti* (3.38-5.12 mya). Because these two intervals do not overlap, at least two episodes of divergence are required, one in the Pleistocene (1.52-2.22 mya) that could account for divergence of both *C. draconoides/plasticus* and *P. repens/thalassinus* and another in the Pliocene (3.38-4.44 mya) that could account for divergence between northern and southern *U. nigricaudus* as well as among members of the *S. orcutti* complex. *Sceloporus zosteromus* did not show a break across the Isthmus, so was not included for comparison.

Only two taxa in this study were present on Isla Ángel de la Guarda. *Callisaurus d. splendidus* diverged from *C. d. rhodostictus* in the late Pliocene/early Pleistocene (1.65-2.63 mya), which did not overlap for the divergence of *P. slevini* from southern *P. mearnsi* in the mid-late Pleistocene (0.68-1.35 mya). Thus, I reject simultaneous divergence of these two taxa. Isla Espiritu Santa and Isla Partida are only separated by a tidal channel a few meters wide, which would have receded with sea levels even a few meters lower than present, I treat them as a single island for comparison. All four genera were collected on either or both of these islands as well as the cape region. A minimum of two episodes of divergence are required – one for the insular *U. nigricaudus* in the early

Pleistocene/late Pliocene (1.65-2.78 mya) and another narrow window in the late Pleistocene (0.66 and 0.74 mya) that could account for divergence of insular *C. d. draconoides*, *S. hunsakeri*, and *P. thalassinus*.

Taking these results together, I conclude that divergence of the various lineages defined in this study did not occur simultaneously, supporting the hypothesis of Leaché et al. (2007). For any given phylogeographic break, at least two or three episodes of divergence must have occurred, ranging in timing from the late Pleistocene to the early Pliocene. However, there are also numerous instances of overlap between 95% HPDs where I cannot reject the hypothesis of simultaneous divergence of two or three taxa at the same break.

(4) Influence of geology and climate on divergence

The fourth and final question I asked was: can divergence times among the various lineages be linked to specific geological and/or climatic events? As mentioned earlier, the following interpretation is based on the assumption that the mammalian substitution rate used here is accurate. Fortunately, several geological calibrations are available that can be used as upper bounds on various divergences, and indicate that this rate is not likely to be much slower (although it could be faster). First, the origin of the modern BCP began with transform faulting in the GoC ~6 mya and the transfer of plate motions into the basin was completed by ~3.5 mya (Lonsdale 1989; Umhoefer et al. 1994, 2002); therefore, we should not expect to find divergences of intrapeninsular clades to pre-date 6 mya (as measured by the lower bound of the 95% HPD). Likewise, Isla Ángel de la

Guarda separated from the BCP during the Pliocene/Pleistocene ~2-3 mya due to transform faulting along the San Andreas Fault system (Stock 2000, Carreño and Helenes 2002), so populations of *C. d. splendidus* and *P. slevini* should not be older than this. I apply the same logic to populations on the remaining islands – Isla San Francisco formed in the Pliocene, and Islas Tortuga, Carmen, Danzante, Espiritu Santo, Partida are of Pleistocene origin (Carreño and Helenes 2002). With the mutation rate I assumed, none of the 95% HPDs for the divergence observed in any trees pre-date these boundaries. The only mean divergence that pre-dates the origin of the BCP is that of *U. lahtelai* from *U. nigricaudus/microscutatus* (mean estimate 6.69 mya), but the lower bound of the HPD (4.96 mya) falls well within this geological constraint.

The islands included in this study can be divided into three types. Continental (land-bridge) islands are separated by shallow seaways that would have been exposed as dry land during the last glacial maximum when sea levels were up to 100 m lower than present (Carreño and Helenes 2002), although the actual age of the island itself may predate this. Islas San Francisco, Carmen, Danzante, Espiritu Santo, Partida, and Santa Margarita belong to this category. For continental islands, the most parsimonious explanation for the colonization of the island is that of vicariance; lizards were able to disperse into the island when it was connected to the peninsula, subsequently isolated by the rising sea level. Conversely, oceanic islands are volcanoes that emerge from the sea floor and were never connected to a continent. It can thus be inferred that any lizards occupying these islands must have dispersed there naturally over water or been

introduced by humans. The sole example of this type of island in my study is Isla Tortuga and its population of *S. orcutti*. My mid-Pleistocene divergence estimate is consistent with the known age of this island (Carreño and Helenes 2002) and indicates that this population is naturally occurring and must have dispersed over water. This is significant because it cannot be assumed that insular populations of lizards in the GoC are all naturally occurring – Seri Indians in particular have been documented to transplant populations of chuckwallas (*Sauromalus sp.*) for food (Felger and Moser 1985). Although the Seri are not known to have eaten *S. orcutti*, it is still satisfying to be able to reject this hypothesis directly. Finally, Isla Ángel de la Guarda, the third type, does not neatly fit into the categories of continental or oceanic islands. As mentioned above, this island fragmented from the trailing edge of the BCP as it migrated northwest (Stock 2000). Thus, the structural rocks of the island are part of the BCP itself, but since the early Pleistocene a deep basin developed in the Canal de Ballenas that separates the island from the peninsula. With the elevation of the island and depths of the channel each exceeding 1000 meters, if the water were to be removed from this basin the scene might resemble the Grand Canyon. Thus, glacial maxima of the late Pleistocene would not have connected this island to the peninsula (Carreño and Helenes 2002). This presents an interesting opportunity to test hypotheses of vicariance vs. dispersal. The mean divergence date estimate of 2.11 mya for the split of *C. d. splendidus* from *C. d. rhodostictus* is more consistent with the vicariance hypothesis – lizards were likely already on the landmass that would become the island, and were simply transported along

for the ride. In contrast, my mean divergence date estimate of 0.99 mya for the divergence of *P. slevini* from southern populations of *P. mearnsi* supports the over-water dispersal hypothesis. *Petrosaurus mearnsi* is a common inhabitant of the steep canyons that drain from the eastern escarpment of the peninsular ranges into the GoC, and these canyons are infrequently subjected to powerful flash floods associated with the occasional tropical storm that wanders up the gulf. It is not difficult to imagine a *P. mearnsi*, with its superb ability to cling to the undersides of boulders, getting caught up in such a torrent yet successfully clinging to a raft of debris. If the rafting lizard survived long enough to reach the gulf, powerful tidal currents could easily transport it across the channel in a matter of days, if not hours. One gravid female might be all it takes to start a new population.

Another area of contention is the mid-peninsular break of the Vizcaíno Desert. The varieties of the vicariance hypothesis all suggest that a seaway (hypothesized to be anywhere from late Miocene to mid-Pleistocene in age, depending on the study) severed the BCP across the Vizcaíno Desert around the latitude of San Ignacio, creating spatially concordant phylogeographic breaks across a variety of species (Murphy et al. 1983a, 1983b, Upton and Murphy 1997, Aguirre-León et al. 1999, Riddle et al. 2000, Lindell et al. 2005, 2006, 2008). However, this model has been disputed, as no direct geological evidence for such a seaway exists, and mtDNA from 12 species of small mammals and squamate reptiles instead support two events differentially structuring genetic diversity in this region (Leaché et al. 2007). Gottscho (2014) reviewed the geological and

zoogeographic literature regarding the mid-peninsular break, which is important for not only phrynosomatid lizards but for a variety of other terrestrial and marine animals, and hypothesized that this break was directly related to the alignment of the Vizcaíno Desert and Punta Eugenia with the Molokai and Shirley Fracture Zones, which represent fossilized microplate boundaries (see Chapter 1). However, Gottscho (2014) could not distinguish between the two major alternative hypotheses of vicariance or dispersal across the Vizcaíno Desert. In thus study, I found that a minimum of three waves of divergence are necessary to explain the phylogeographic breaks observed, but the Pliocene/Pleistocene timing of these divergences postdates the formation of the fracture zones and the Vizcaíno Peninsula. *Urosaurus microscutatus* probably shows the strongest effect of a northern range expansion in the Pleistocene across the mid-peninsular break, based on the divergence time estimated using BEAST, the nested south-to-north tree structures, as well as the inverse relationship between heterozygosity and latitude. The western Vizcaino Desert, including its Pacific beaches, also contains extensive deposits of aeolian sand dunes, home to *C. d. crinitus*, which has evolved elongated toe fringes to improve traction in this substrate, convergent with sand-dwelling lizards elsewhere in North America and the world (Norris 1958, Luke 1986). *Callisaurus d. crinitus* avoids such hard-packed substrates even where they occur in isolated pockets within the western part of its range (Grismer 2002). The result that *C. d. crinitus* is a monophyletic unit endemic to the sandy substrates of the Vizcaino Desert suggests that natural

selection may be playing a strong role in promoting speciation, although more work is needed to test this hypothesis.

It is unclear what geological and climatic forces may have influenced divergence across the Loreto/Sierra de la Giganta region. The dramatic eastern escarpment of this range exposes the volcanic Comondú formation, which was deposited in the early Miocene and predates the origin of the BCP, and there is no evidence that a seaway ever severed the steep escarpment of this range (Umhoefer et al. 2001). The Loreto basin itself was flooded by higher sea levels ~3.4 mya (Umhoefer et al. 1994, 2002), and it is possible that the combination of inundation of coastal areas and the escarpment of the Sierra la Giganta may have acted as a barrier to dispersal for *Callisaurus*, as the 95% HPD for the divergence across this barrier (between *C. d. plasticus* and *C. d. carmenensis*) extends from 2.50 to 3.54 mya. However, the 95% HPDs for the divergences of *U. microscutatus* and *U. nigricaudus* and of *P. repens* (central vs. southern) do not include this time period. The contact zone of *U. nigricaudus* and *U. microscutatus* lies along the gentler Pacific slope of the Sierra la Giganta, in contrast to *C. d. carmenensis/plasticus* which both contact in the vicinity of Loreto.

The Isthmus of La Paz is another controversial biogeographic barrier, and again at least two waves of divergence are required to explain my results. Unlike the Vizcaíno Desert and Loreto breaks, which lie along the main axis of the peninsular ranges and were unlikely to have been submerged by higher sea levels, the Isthmus of La Paz is a low-lying area along the La Paz fault zone (Hausback 1984), marine inundations are likely to have occurred. The 95% HPDs

for the divergences of *C. d. draconoides/plasticus*, northern vs. southern *U. nigricaudus*, and *S. orcutti/hunsakeri* include an especially warm period of the Pliocene 2.7 – 3.2 mya when global sea levels were ~22 m higher than present (Miller et al. 2012), which would be sufficient to inundate most of the isthmus – indeed, Pliocene marine fossils are known from low-lying areas of the neighboring cape region (Miller 1980). However, the divergence of *P. thalassinus* from *P. repens* occurred more recently, and the split of *S. licki* from *S. orcutti/hunsakeri* was older. The idea that the cape region existed as an island during the Pliocene is widespread in the literature: mitochondrial data from a variety of taxa suggest a late Pliocene/early Pleistocene seaway across the isthmus along the western side of the fault zone (Riddle et al. 2000, Lindell et al. 2005).

(5) Future directions

This research raises a series of new questions and hypotheses for future studies to address. First, it is clear that taxonomic revisions for all four genera may be warranted, as my study identified several clades/clusters that either correspond with subspecies or represent undescribed biodiversity. These include the six subspecies and central clade of *Callisaurus*, *P. mearnsi* and *P. slevini*, *U. microscutatus*, and northern and southern clades of *S. orcutti* and *S. zosteromus*. The Bayes factor delimitation I used in Chapter 2 (Leaché et al. 2014) would be well applied here, although when I attempted to do so I ran into the problem of locus dropout for *Sceloporus* and *Petrosaurus*. For this reason, a top research

priority of mine in the summer of 2015 is to collect such a dataset that does not suffer from locus dropout using a target capture method (e.g. Faircloth et al. 2012). Not only will this method allow me to align SNPs across more deeply divergent lineages for species tree inference, but such a dataset will also me to test for substitution rate heterogeneity among genera, and a fossil-calibrated ultrametric tree that includes multiple samples from each phylogeographic unit in a single matrix without excessive missing data is sorely needed to better test the hypothesis of simultaneous divergence and to link speciation with specific geological and/or climatic events. Another area where future work is needed is to estimate gene flow rates in a coalescent framework. Two promising methods are the D-statistic (Eaton and Ree 2013) and utilizing the allele frequency spectrum (Gutenkunst et al. 2009). Denser sampling where hybrid zones are thought to occur would also be helpful, especially for *Callisaurus* in the Vizcaíno Desert and the Isthmus of La Paz. Estimating gene flow directly not only should aid in species delimitation, but should also yield insights into the biological mechanisms underlying speciation. Finally, hierarchal approximate Bayesian computation (hABC) could be used to test simultaneous divergence (Hickerson et al. 2007, Leaché et al. 2007).

(6) Summary and conclusions

Utilizing a large RADseq dataset, I studied the comparative phylogeography of phrynosomatid lizards of the BCP to better understand the relationship between landscape evolution and speciation. I found that

heterozygosity was positively correlated with latitude in *Callisaurus*, negatively correlated in *Urosaurus* and *Petrosaurus*, and not correlated at all in *Sceloporus*, suggesting that these taxa did not all respond in the same way to changing climatic conditions. Overall, insular lizards had significantly lower heterozygosity than peninsular lizards, and this pattern held true for *Callisaurus*, *Petrosaurus*, and *Sceloporus*, but not *Urosaurus*. The Admixture analyses revealed 25 clusters among the four genera, 22 of which correspond with strongly supported clades in my concatenated phylogenetic analyses. I identified five regions of the BCP where two or more taxa demonstrate spatially concordant phylogeographic boundaries, and I infer multiple episodes of divergence must have occurred across each of these boundaries. Thus, for each phylogeographic break, I reject the hypothesis that all taxa diverged simultaneously due to single vicariance event (e.g. a trans-peninsular seaway). Consistent with the findings of Smith et al. (2014), who applied the same basic questions to a different study system (Neotropical rainforest birds), I instead interpret these results to support the hypothesis that plate tectonic events established a complex landscape matrix with numerous barriers to dispersal that ultimately facilitated speciation indirectly, generating the idiosyncratic patterns observed among co-distributed lineages.

V. REFERENCES

Adest GA. 1987. Genetic differentiation among populations of the zebratail lizard, *Callisaurus draconoides* (Sauria, Iguanidae). Copeia:854–859.

Aguilars-S MA, Sites JW, Murphy RW. 1988. Genetic variability and population structure in the genus *Petrosaurus* (Iguanidae). Journal of Herpetology, 22:135–145.

Aguirre-León G, Morafka DJ, Murphy RW. 1999. The peninsular archipelago of Baja California: a thousand kilometers of tree lizard genetics. Herpetologica, 55:369–381.

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 19:1655–1664.

Arbogast BS, Kenagy GJ. 2001. Comparative phylogeography as an integrative approach to historical biogeography. Journal of Biogeography, 28:819–825.

Avise JC. 2000. Phylogeography: the history and formation of species. Harvard University Press, Cambridge, MA.

Blainville HMD. 1835. Description de quelques espèces de reptiles de la Californie précédée de l'analyse d'un système général d'herpétologie et d'amphibologie. Nouvelles annales du Muséum d'histoire naturelle Paris, 4:232–296.

Bouckaert R, Heled J, Kuehnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Computational Biology, 10:e1003537.

Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Molecular Biology and Evolution, 29:1917–1932.

Carreño AL, Helenes J. 2002. Geology and ages of the islands. In: Case TJ, Cody ML, Ezcurra E editors. A new island biogeography of the Sea of Cortés, Oxford University Press, p. 14–40.

Carstens BC, Lemmon AR, Lemmon EM. 2012. The promises and pitfalls of next-generation sequencing data in phylogeography. Systematic Biology, 61:713–715.

Carstens BC, Richards CL. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. Evolution, 61:1439–1454.

Carstens BC, Stoute HN, Reid NM. 2009. An information-theoretical approach to phylogeography. Molecular Ecology, 18:4270–4282.

Case TJ. 1975. Species numbers, density compensation, and colonizing ability of lizards on islands in the Gulf of California. Ecology, 56:3–18.

Clark-Tapia R, Molina-Freaner F. 2003. The genetic structure of a columnar cactus with a disjunct distribution: *Stenocereus gummosus* in the Sonoran desert. Heredity, 90:443–450.

Conroy CJ, Papenfuss T, Parker J, Hahn NE. 2009. Use of tricaine methanesulfonate (MS222) for euthanasia of reptiles. Journal of the American Association for Laboratory Animal Science, 48:28.

Cope ED. 1863. Descriptions of new American Squamata in the museum of the Smithsonian Institution, Washington. Proceedings of the Academy of Natural Sciences of Philadelphia, 15:100–106.

Cope ED. 1864. Contributions to the herpetology of tropical America. Proceedings of the Academy of Natural Sciences of Philadelphia, 16:166–181.

Cope ED. 1896. On the genus *Callisaurus*. American Naturalist, 30:1049–1050.

Crawford NG, Faircloth BC, McCormack JE, Brumfield RT, Winker K, Glenn TC. 2012. More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs. Biology Letters, 8:783–786.

Dickerson MC. 1919. Diagnoses of twenty-three new species and a new genus of lizards from Lower California. Bulletin of the American Museum of Natural History, 61:461–477.

Eaton DAR. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics, 30:1844–1849.

Eaton DAR, Ree RH. 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). Systematic Biology, 62:689–706.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32:1792–1797.

Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics, 26:2460–2461.

Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE, Holzapfel CM. 2010. Resolving postglacial phylogeography using high-throughput sequencing. Proceedings of the National Academy of Sciences of the United States of America, 107:16196–16200.

Excoffier L, Foll M, Petit RJ. 2009. Genetic consequences of range expansions. Annual Review of Ecology Evolution and Systematics, 40:481–501.

Ezcurra E, Bourillón L, Cantú A, Martínez ME, Robles A. 2002. Ecological conservation. In: Case TJ, Cody ML, Ezcurra E editors. A new island biogeography of the Sea of Cortés, Oxford University Press, p. 417–444.

Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology, 61:717–726.

Feldman CR, Flores-Villela O, Papenfuss TJ. 2011. Phylogeny, biogeography, and display evolution in the tree and brush lizard genus *Urosaurus* (Squamata: Phrynosomatidae). Molecular Phylogenetics and Evolution, 61:714–725.

Felger RS, Moser MB. 1985. People of the desert and sea. University of Arizona Press, Tucson.

Garrick RC, Nason JD, Meadows CA, Dyer RJ. 2009. Not just vicariance: phylogeography of a Sonoran Desert euphorb indicates a major role of range expansion along the Baja peninsula. Molecular Ecology, 18:1916–1931.

Gottscho AD. 2014. Zoogeography of the San Andreas Fault system: Great Pacific Fracture Zones correspond with spatially concordant phylogeographic boundaries in western North America. Biological Reviews, in press (doi:10.1111/brv.12167).

Grismer LL. 1994. The origin and evolution of the peninsular herpetofauna of Baja California, Mexico. Herpetological Natural History, 2:51–106.

Grismer LL. 2002. Amphibians and reptiles of Baja California, including its Pacific islands and the islands in the Sea of Cortés. University of California Press, Berkeley.

Grismer LL, McGuire JA. 1996. Taxonomy and biogeography of the *Sceloporus magister* complex (Squamata: Phrynosomatidae) in Baja California, Mexico. Herpetologica, 52:416–427.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genetics, e1000695.

Haffer J. 1969. Speciation in Amazonian forest birds. Science, 165:131-137.

Hall WP, Smith HM. 1979. Lizards of the *Sceloporus orcutti* complex of the Cape Region of Baja California. Breviora, 452:1–26.

Hallowell E. 1854. Descriptions of new reptiles from California. Proceedings of the Academy of Natural Sciences of Philadelphia, 7:91–97.

Harvey MG, Smith BT, Glenn TC, Faircloth BC, Brumfield RT. Sequence capture versus restriction site associated DNA sequencing for phylogeography. Unpublished manuscript, doi:arxiv.org/pdf/1312.6439

Hausback BP. 1984. Cenozoic volcanic and tectonic evolution of Baja California Sur, Mexico. In: Frizzell VA editor. Geology of the Baja California peninsula, Pacific Section, Society of Economic Paleontologists and Mineralogists, Los Angeles, p. 219–236.

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

Hickerson MJ, Stahl E, Takebayashi N. 2007. msBayes: pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation. BMC Bioinformatics, 8:268.

Holmgren CA, Betancourt JL, Cristina Penalba M, Delgadillo J, Zuravnsky K, Hunter KL, Rylander KA, Weiss JL. 2014. Evidence against a Pleistocene

desert refugium in the Lower Colorado River Basin. Journal of Biogeography, 41:1769–1780.

Hoorn C, Wesselingh FP, ter Steege H, Bermudez MA, Mora A, Sevink J, Sanmartin I, Sanchez-Meseguer A, Anderson CL, Figueiredo JP, et al. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science, 330:927–931.

Huang H, Knowles LL. 2014. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. Systematic Biology, in press (doi:10.1093/sysbio/syu046).

Keigwin LD. 1978. Pliocene closing of Isthmus of Panama, based on biostratigraphic evidence from nearby Pacific Ocean and Caribbean Sea cores. Geology, 6:630–634.

Knowles LL. 2009. Statistical phylogeography. Annual Review of Ecology Evolution and Systematics, 40:593–612.

Knowles LL, Maddison WP. 2002. Statistical phylogeography. Molecular Ecology, 11:2623–2635.

Kumar S, Subramanian S. 2002. Mutation rates in mammalian genomes. Proceedings of the National Academy of Sciences of the United States of America, 99:803–808.

Leaché AD, Chavez AS, Jones LN, Grummer JA, Gottscho AD, Linkem CW. 2015. Phylogenomics of phrynosomatid lizards: conflicting signals from sequence capture versus restriction site associated DNA sequencing. Genome Biology and Evolution, 7:706–719.

Leaché AD, Crews SC, Hickerson MJ. 2007. Two waves of diversification in mammals and reptiles of Baja California revealed by hierarchical Bayesian analysis. Biology Letters, 3:646–650.

Leaché AD, Fujita MK, Minin VN, Bouckaert RR. 2014. Species delimitation using genome-wide SNP data. Systematic Biology, 63:534–542.

Leaché AD, Harris RB, Maliska ME, Linkem CW. 2013. Comparative species divergence across eight triplets of spiny lizards (*Sceloporus*) using genomic sequence data. Genome Biology and Evolution, 5:2410–2419.

Leaché AD, Mulcahy DG. 2007. Phylogeny, divergence times and species limits of spiny lizards (*Sceloporus magister* species group) in western North American deserts and Baja California. Molecular Ecology, 16:5216–5233.

Lee JY, Edwards SV. 2008. Divergence across Australia's Carpentarian Barrier: statistical phylogeography of the red-backed fairy wren (*Malurus melanocephalus*). Evolution, 62:3117–3134.

Lindell J, Mendez-de la Cruz FR, Murphy RW. 2005. Deep genealogical history without population differentiation: discordance between mtDNA and allozyme divergence in the zebra-tailed lizard (*Callisaurus draconoides*). Molecular Phylogenetics and Evolution, 36:682–694.

Lindell J, Mendez-De La Cruz FR, Murphy RW. 2008. Deep biogeographical history and cytonuclear discordance in the black-tailed brush lizard (*Urosaurus nigricaudus*) of Baja California. Biological Journal of the Linnean Society, 94:89–104.

Lindell J, Ngo A, Murphy RW. 2006. Deep genealogies and the mid-peninsular seaway of Baja California. Journal of Biogeography, 33:1327–1331.

Liu H, Prugnolle F, Manica A, Balloux F. 2006. A geographically explicit genetic model of worldwide human-settlement history. American Journal of Human Genetics, 79:230–237.

Longrich NR, Vinther J, Pyron RA, Pisani D, Gauthier JA. 2015. Biogeography of worm lizards (Amphisbaenia) driven by end-Cretaceous mass extinction. Proceedings of the Royal Society B, 282.

Lonsdale P. 1989. Geology and tectonic history of the Gulf of California. In: Winterer EL, Hussong DM, Decker RW editors. The geology of North America, Vol. N: the eastern Pacific Ocean and Hawaii, The Geological Society of America, Boulder, Colorado, p. 499–521.

Luke C. 1986. Convergent evolution of lizard toe fringes. Biological Journal of the Linnean Society, 27:1–16.

Marko PB. 2002. Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. Molecular Biology and Evolution, 19:2005–2021.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution, 66:526–538.

Meredith RW, Hekkala ER, Amato G, Gatesy J. 2011. A phylogenetic hypothesis for *Crocodylus* (Crocodylia) based on mitochondrial DNA: evidence for a trans-Atlantic voyage from Africa to the New World. Molecular Phylogenetics and Evolution, 60:183–191.

Miller MA, Schwartz T, Pickett BE, He S, Klem EB, Scheuermann RH, Passarotti M, Kaufman S, O'Leary MA. 2015. A RESTful API for access to phylogenetic tools via the CIPRES science gateway. Evolutionary Bioinformatics, 11:43–48.

Miller KG, Wright JD, Browning JV, Kulpecz A, Kominz M, Naish TR, Cramer BS, Rosenthal Y, Peltier WR, Sosdian S. 2012. High tide of the warm Pliocene: implications of global sea level for Antarctic deglaciation. Geology, 40:407–410.

Miller WE. 1980. The late Pliocene Las Tunas local fauna from southernmost Baja California, Mexico. Journal of Paleontology, 54:762–805.

Mittermeier RA, Mittermeier CG, Brooks TM, Pilgrim JD, Konstant WR, da Fonseca GAB, Kormos C. 2003. Wilderness and biodiversity conservation. Proceedings of the National Academy of Sciences of the United States of America, 100:10309–10313.

Moore DG, Buffington. 1968. Transform faulting and growth of the Gulf of California since the late Pliocene. Science, 161:1238–1241.

Murphy RW. 1983a. Paleobiogeography and genetic differentiation of the Baja California herpetofauna. Occasional Papers of the California Academy of Sciences, 137:1–48.

Murphy RW. 1983b. The reptiles: origin and evolution. In: Case TJ, Cody ML editors. Island biogeography in the Sea of Cortéz, University of California Press, Berkeley, p. 130–158.

Murphy RW, Aguirre-Leon G. 2002. The nonavian reptiles: origins and evolution. In: Case TJ, Cody ML, Ezcurra E editors. A new island biogeography of the Sea of Cortés, Oxford University Press, p. 181–220.

Nason JD, Hamrick JL, Fleming TH. 2002. Historical vicariance and postglacial colonization effects on the evolution of genetic structure in *Lophocereus*, a Sonoran Desert columnar cactus. Evolution, 56:2214–2226.

Nelson G, Platnick N. 1981. Systematics and Biogeography: Cladistics and Vicariance. Columbia University Press, New York.

Norris KS. 1958. The evolution and systematics of the iguanid genus *Uma* and its relation to the evolution of other North American desert reptiles. Bulletin of the American Museum of Natural History, 114:251–326.

Oaks JR. 2014. An improved approximate-Bayesian model-choice method for estimating shared evolutionary history. BMC Evolutionary Biology, 14:150.

Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. PLoS Genetics, 2:2074–2093.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One, 7:e37135.

Pinheiro JC, Bates DM. 2000. Mixed-effects models in S and S-PLUS. Springer.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics, 155:945–959.

Prugnolle F, Manica A, Balloux F. 2005. Geography predicts neutral genetic diversity of human populations. Current Biology, 15:159–160.

Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer

Rau CS, Loomis RB. 1977. A new species of *Urosaurus* (Reptilia, Lacertilia, Iguanidae) from Baja California, Mexico. Journal of Herpetology, 11:25–29.

Reich D, Price AL, Patterson N. 2008. Principal component analysis of genetic data. Nature Genetics, 40:491–492.

Rheindt FE, Fujita MK, Wilton PR, Edwards SV. 2014. Introgression and phenotypic assimilation in *Zimmerius* flycatchers (Tyrannidae): population genetic and phylogenetic inferences from genome-wide SNPs. Systematic Biology, 63:134–152.

Riddle BR, Hafner DJ. 2010. Integrating pattern with process at biogeographic boundaries: the legacy of Wallace. Ecography, 33:321–325.

Riddle BR, Hafner DJ, Alexander LF, Jaeger JR. 2000. Cryptic vicariance in the historical assembly of a Baja California peninsular desert biota. Proceedings of the National Academy of Sciences of the United States of America, 97:14438–14443.

Riginos C. 2005. Cryptic vicariance in Gulf of California fishes parallels vicariant patterns found in Baja California mammals and reptiles. Evolution, 59:2678–2690.

Rokas A, Williams BL, King N, Carroll SB. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature, 425:798–804.

Savage JM. 1960. Evolution of a peninsular herpetofauna. Systematic Zoology, 9:184–212.

Schrago CG, Russo CAM. 2003. Timing the origin of New World monkeys. Molecular Biology and Evolution, 20:1620–1625.

Schulte JA, Macey JR, Papenfuss TJ. 2006. A genetic perspective on the geographic association of taxa among and North American lizards of the *Sceloporus magister* complex (Squamata: Iguanidae: Phrynosomatinae). Molecular Phylogenetics and Evolution, 39:873–880.

Smith BT, Harvey MG, Faircloth BC, Glenn TC, Brumfield RT. 2014b. Target capture and massively parallel sequencing of ultraconserved elements for comparative studies at shallow evolutionary time scales. Systematic Biology, 63:83–95.

Smith BT, McCormack JE, Cuervo AM, Hickerson MJ, Aleixo A, Cadena CD, Perez-Eman J, Burney CW, Xie X, Harvey MG, et al. 2014a. The drivers of tropical speciation. Nature, 515:406–409.

Soulé M, Sloan AJ. 1966. Biogeography and distribution of the reptiles and amphibians on islands in the Gulf of California, Mexico. Transactions of the San Diego Society of Natural History, 14:137–156.

Soulé M, Yang SY. 1973. Genetic variation in side-blotched lizards on islands in the Gulf of California. Evolution, 27:593–600.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30:1312–1313.

Stejneger LH. 1893. Annotated list of the reptiles and batrachians collected by the Death Valley Expedition in 1891, with descriptions of new species. North American Fauna, 7:159–228.

Stejneger L. 1894. Description of *Uta mearnsi*, a new lizard from California. Proceedings of the United States National Museum, 17:589–591.

Stock JM, Hodges KV. 1989. Pre-Pliocene extension around the Gulf of California and the transfer of Baja California to the Pacific plate. Tectonics, 8:99–115.

Stock JM. 2000. Relation of the Puertecitos volcanic province, Baja California, Mexico, to development of the plate boundary in the Gulf of California. In: Delgado-Granados H, Aguirre-Diaz G, Stock JM editors. Cenozoic volcanism and tectonics of Mexico, Geological Society of America Special Paper, 334:143–156.

Streicher JW, Devitt TJ, Goldberg CS, Malone JH, Blackmon H, Fujita MK. 2014. Diversification and asymmetrical gene flow across time and space: lineage sorting and hybridization in polytypic barking frogs. Molecular Ecology, 23:3273–3291.

Umhoefer PJ, Dorsey RJ, Renne P. 1994. Tectonics of the Pliocene Loreto Basin, Baja California Sur, Mexico, and the Evolution of the Gulf of California. Geology, 22:649–652.

Umhoefer PJ, Dorsey RJ, Willsey S, Mayer L, Renne P. 2001. Stratigraphy and geochronology of the Comondú group near Loreto, Baja California Sur, Mexico. Sedimentary Geology, 144:125–147.

Umhoefer PJ, Mayer L, Dorsey RJ. 2002. Evolution of the margin of the Gulf of California near Loreto, Baja California Peninsula, Mexico. Geological Society of America Bulletin, 114:849–868.

Upton DE, Murphy RW. 1997. Phylogeny of the side-blotched lizards (Phrynosomatidae: *Uta*) based on mtDNA sequences: support for a midpeninsular seaway in Baja California. Molecular Phylogenetics and Evolution, 8:104–113.

Van Denburgh J. 1894. Descriptions of three new lizards from California and lower California, with a note on *Phrynosoma blainvillii*. Proceedings of the California Academy of Sciences, 4:296–301.

Van Denburgh J. 1895. Review of the herpetology of lower California. Proceedings of the California Academy of Sciences, 2:77–163.

Van Denburgh J. 1922. The reptiles of western North America, Vol 1: Lizards. Proceedings of the California Academy of Sciences, 1:1–611.

Van Denburgh J, Slevin JR. 1921. Preliminary diagnoses of new species of reptiles from islands in the Gulf of California, Mexico. Proceedings of the California Academy of Sciences, 11: 95–98.

Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, Seehausen O. 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Molecular Ecology, 22:787–798.

Welsh HH. 1988. An ecogeographic analysis of the herpetofauna of the Sierra San Pedro Mártir Region, Baja California, with a contribution to the biogeography of the Baja California herpetofauna. Proceedings of the California Academy of Sciences, 46:1–72.

Yarrow HC. 1882. Description of new species of reptiles and amphibians in the U.S. National Museum. Proceedings of the United States National Museum, 5:438–443.

Zink RM. 1996. Comparative phylogeography in North American birds. Evolution, 50:308–317.

Zink RM. 2002. Methods in comparative phylogeography, and their application to studying evolution in the North American aridlands. Integrative and Comparative Biology, 42:953–959.

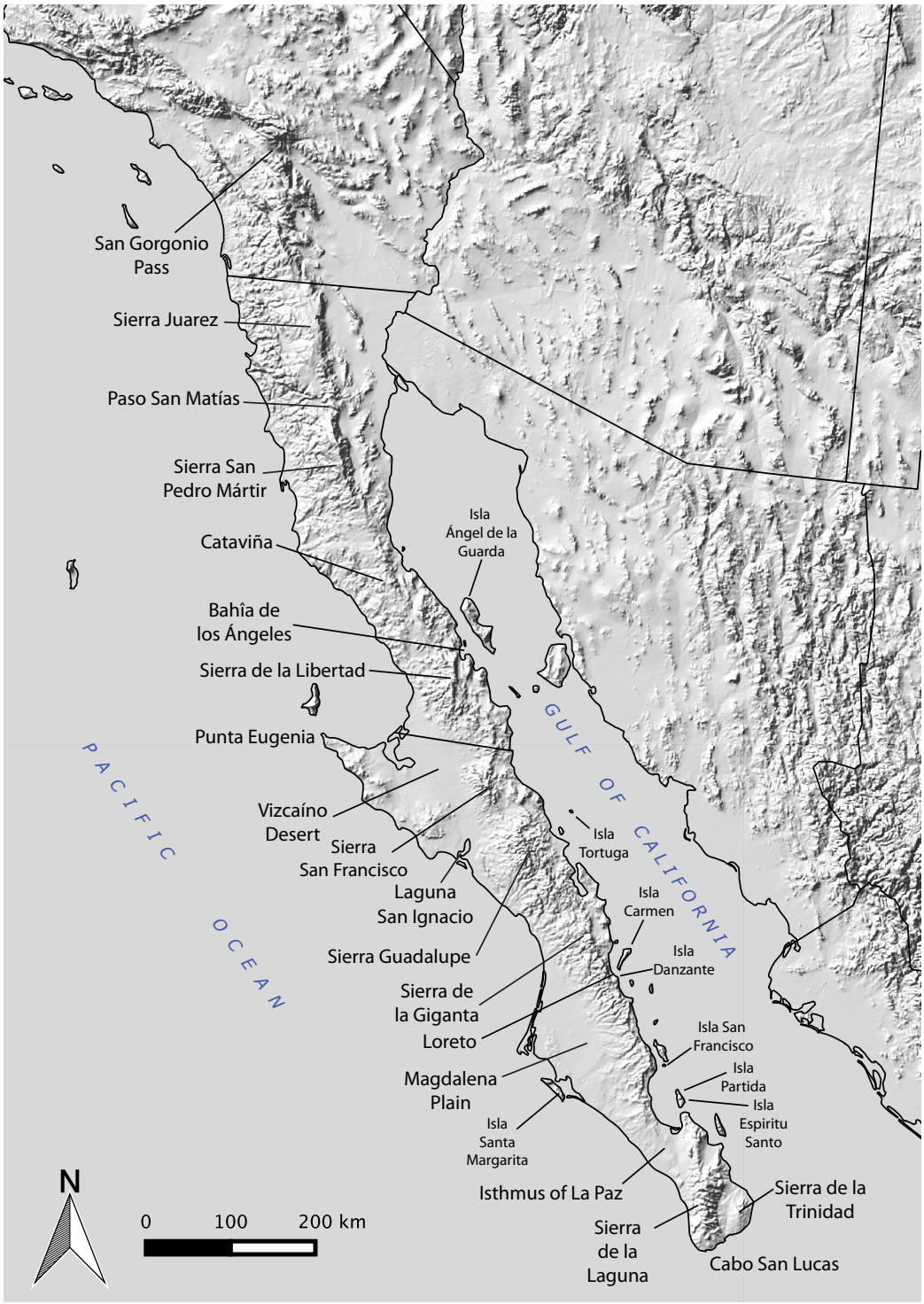

Figure 3.1. Map of the Baja California Peninsula and Gulf of California with relevant geographic features.

Figure 3.2. Examples of the four genera of phrynosomatid lizards featured in this chapter. Clockwise from top left: zebra-tailed lizard (*Callisaurus draconoides draconoides*), east of La Paz, Baja California Sur; banded rock lizard (*Petrosaurus thalassinus*), Sierra de la Laguna, Baja California Sur; black-tailed brush lizard (*Urosaurus microscutatus*), Bahía Concepción, Baja California Sur; granite spiny lizard (*Sceloporus orcutti*), Riverside, California.

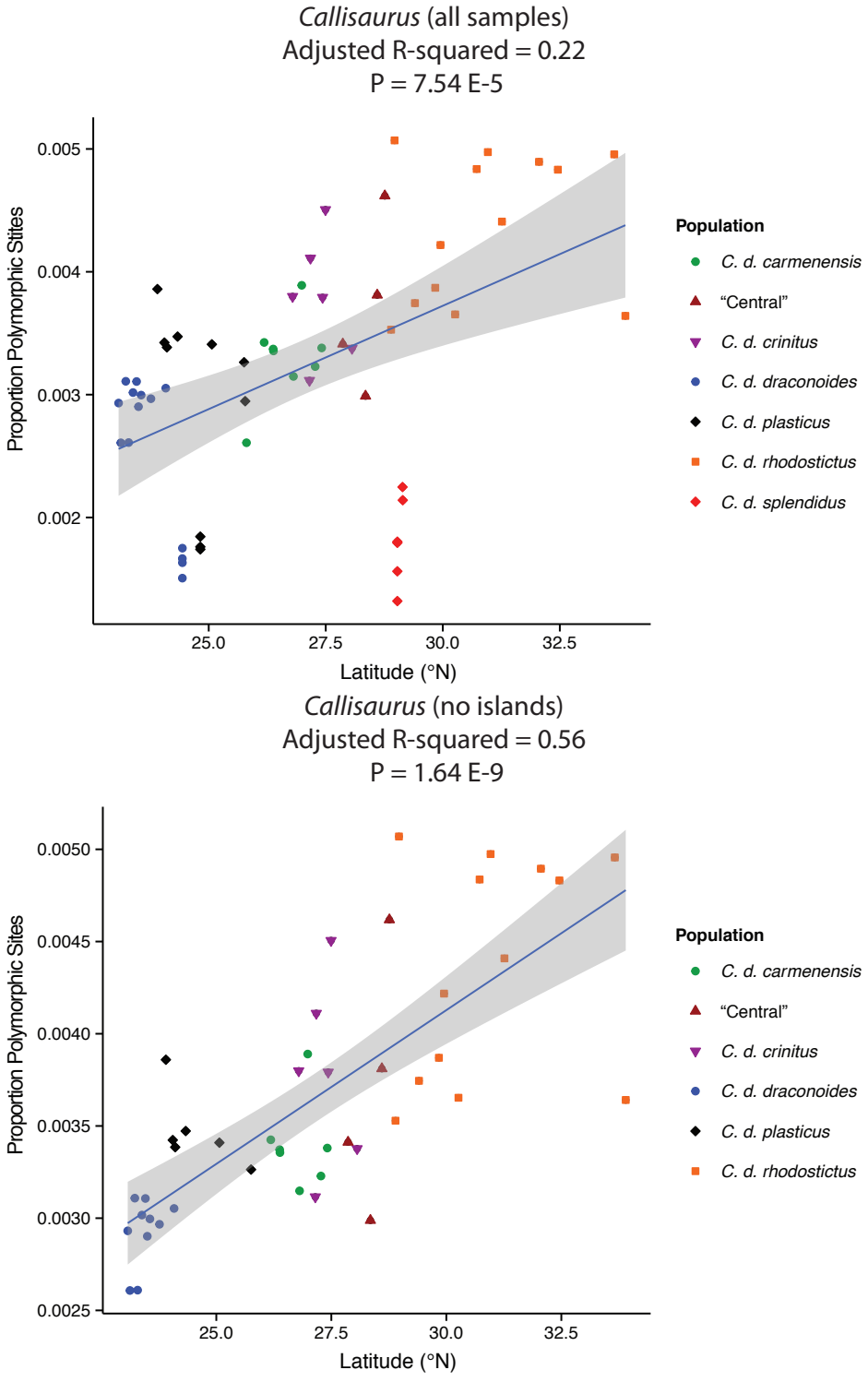

Figure 3.3. Linear regression of the proportion of heterozygous sites vs. latitude for *Callisaurus*, including (top) and excluding (bottom) insular samples.

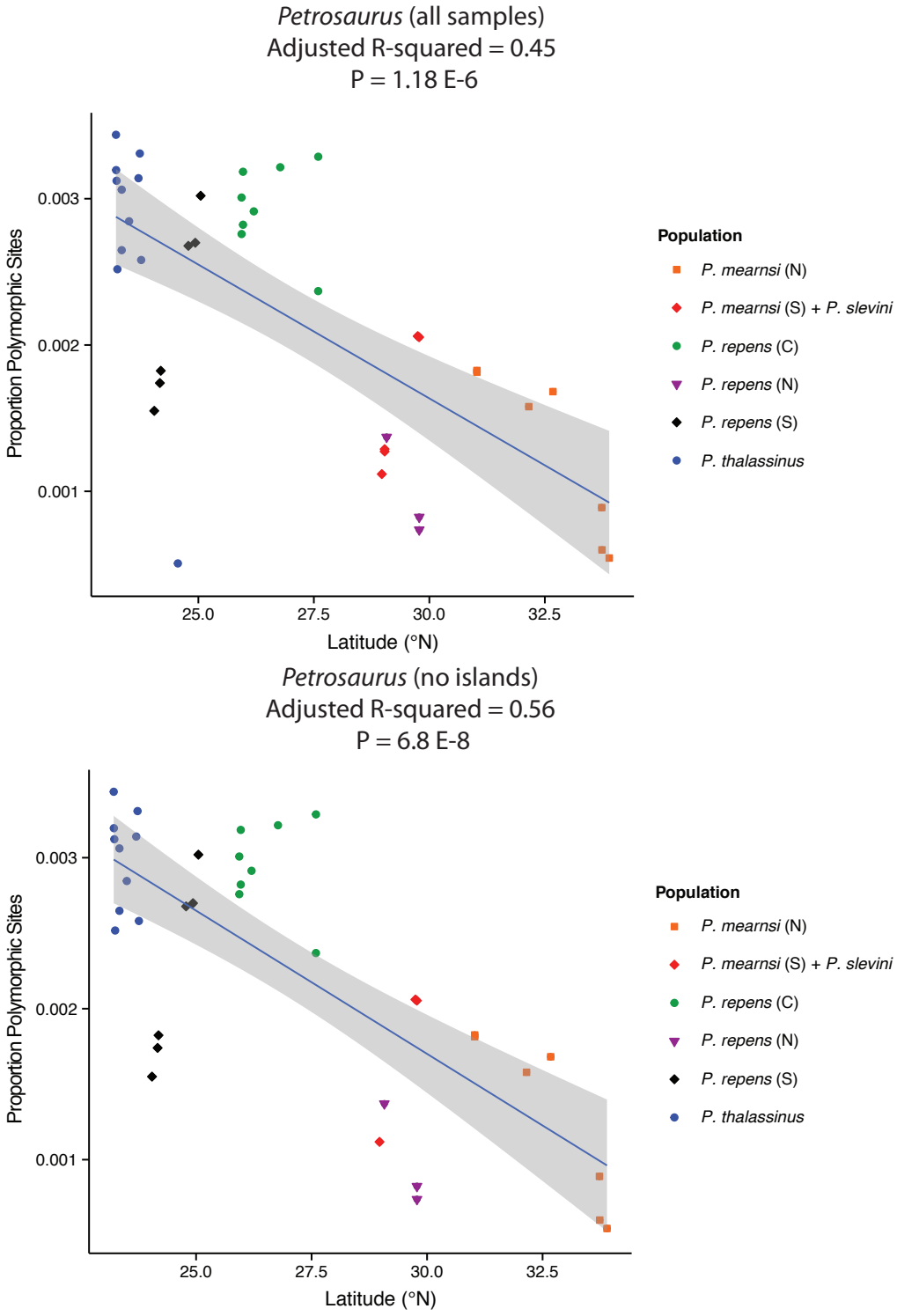

Figure 3.4. Linear regression of the proportion of heterozygous sites vs. latitude for *Petrosaurus*, including (top) and excluding (bottom) insular samples.

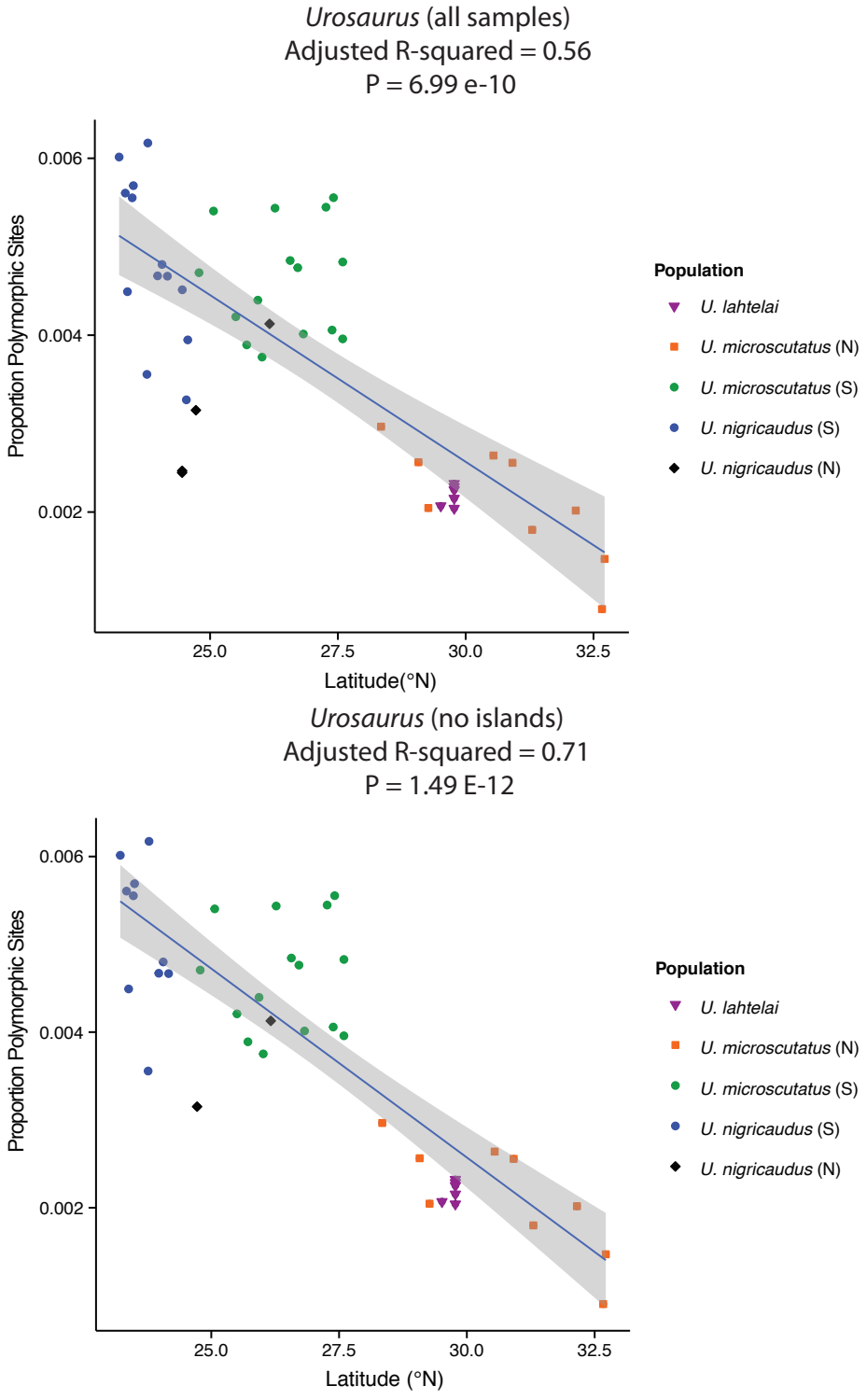

Figure 3.5. Linear regression of the proportion of heterozygous sites vs. latitude for *Urosaurus*, including (top) and excluding (bottom) insular samples.

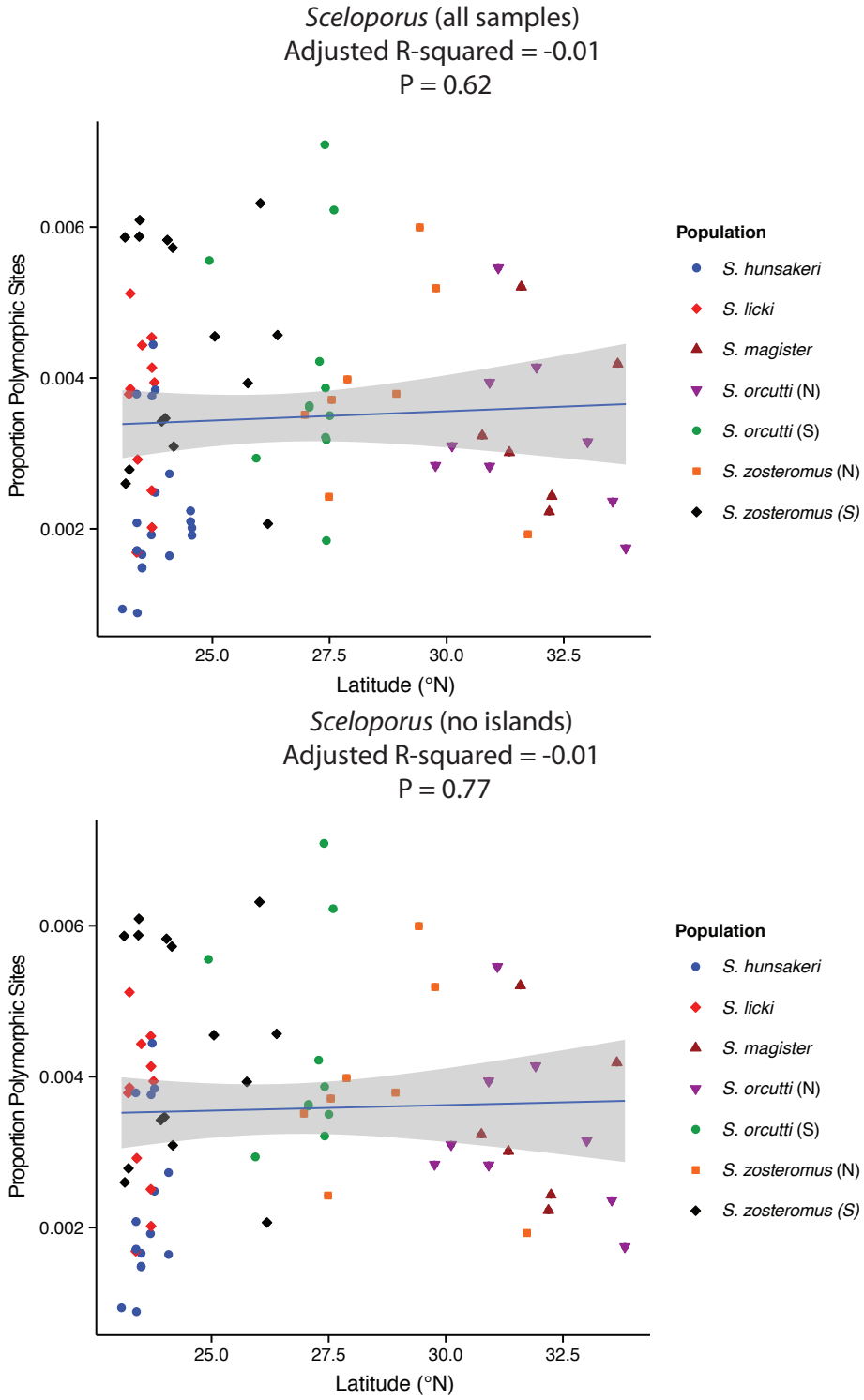

Figure 3.6. Linear regression of the proportion of heterozygous sites vs. latitude for *Sceloporus*, including (top) and excluding (bottom) insular samples.

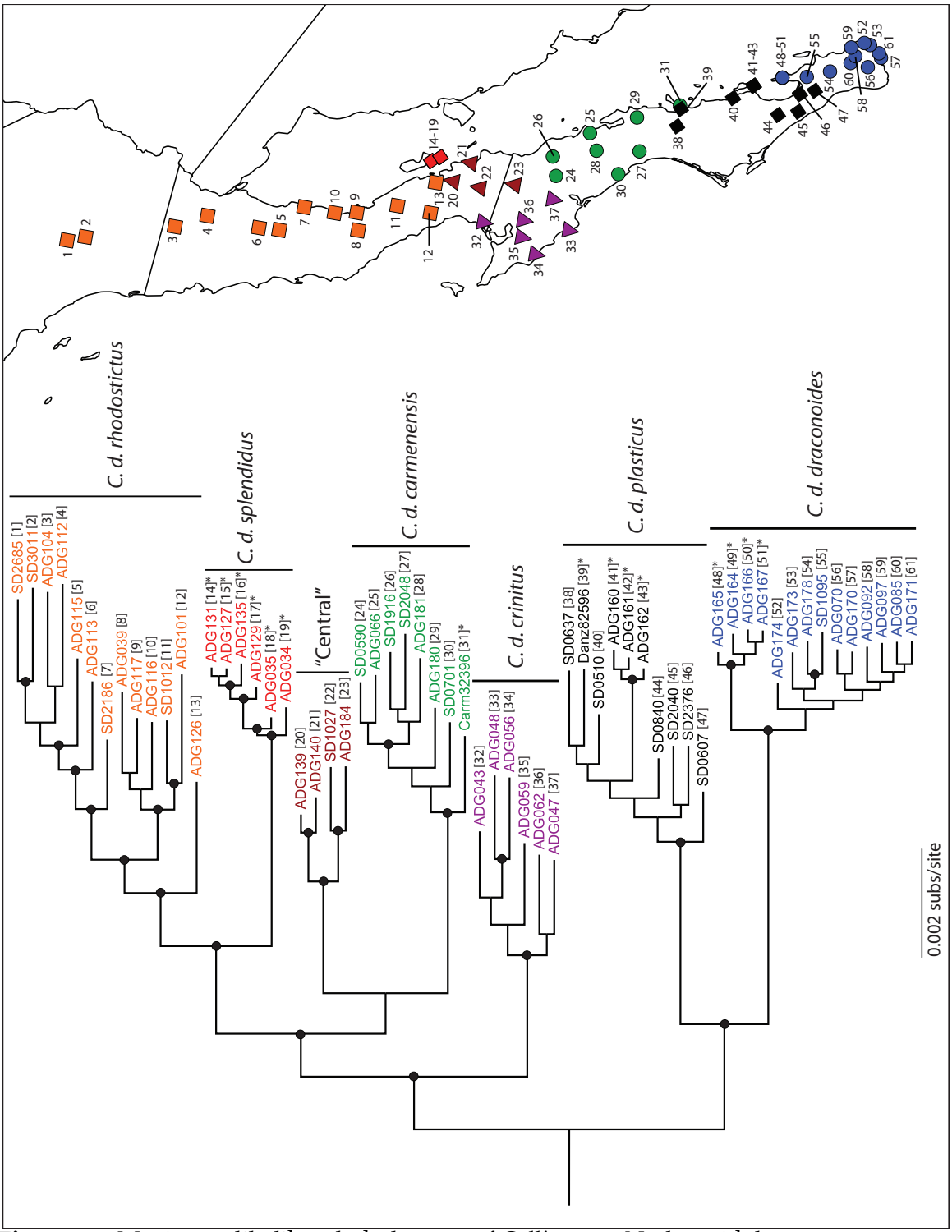

Figure 3.7. Maximum likelihood phylogram of *Callisaurus*. Nodes with bootstrap support > 70% are denoted with solid black circles on nodes, and numbers in brackets next to taxa labels (left) correspond to the numbers on the map (right).

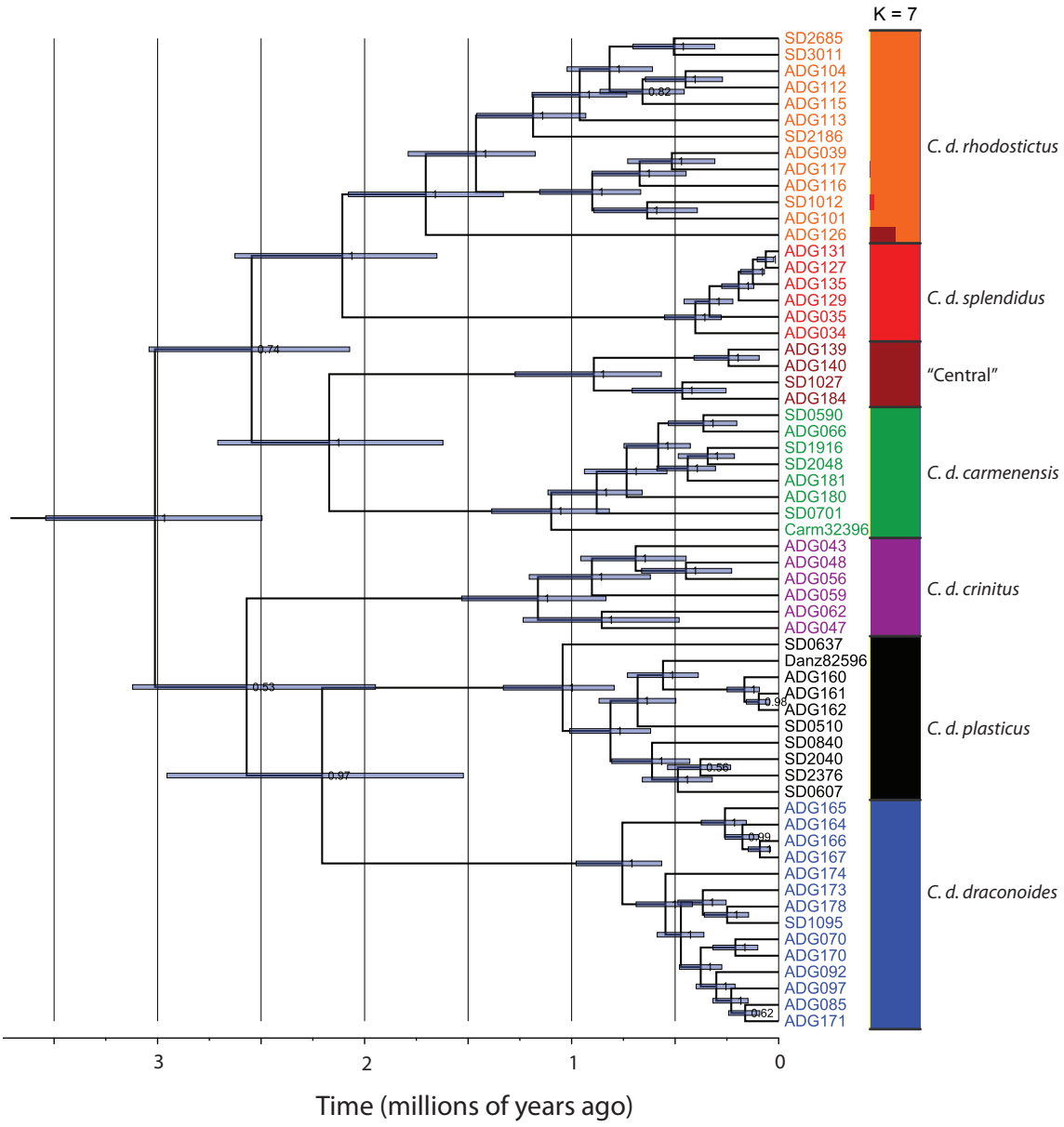

Figure 3.8. Clock tree for *Callisaurus* (left), with Admixture plot (right) under K=7 (994 unlinked SNPs).

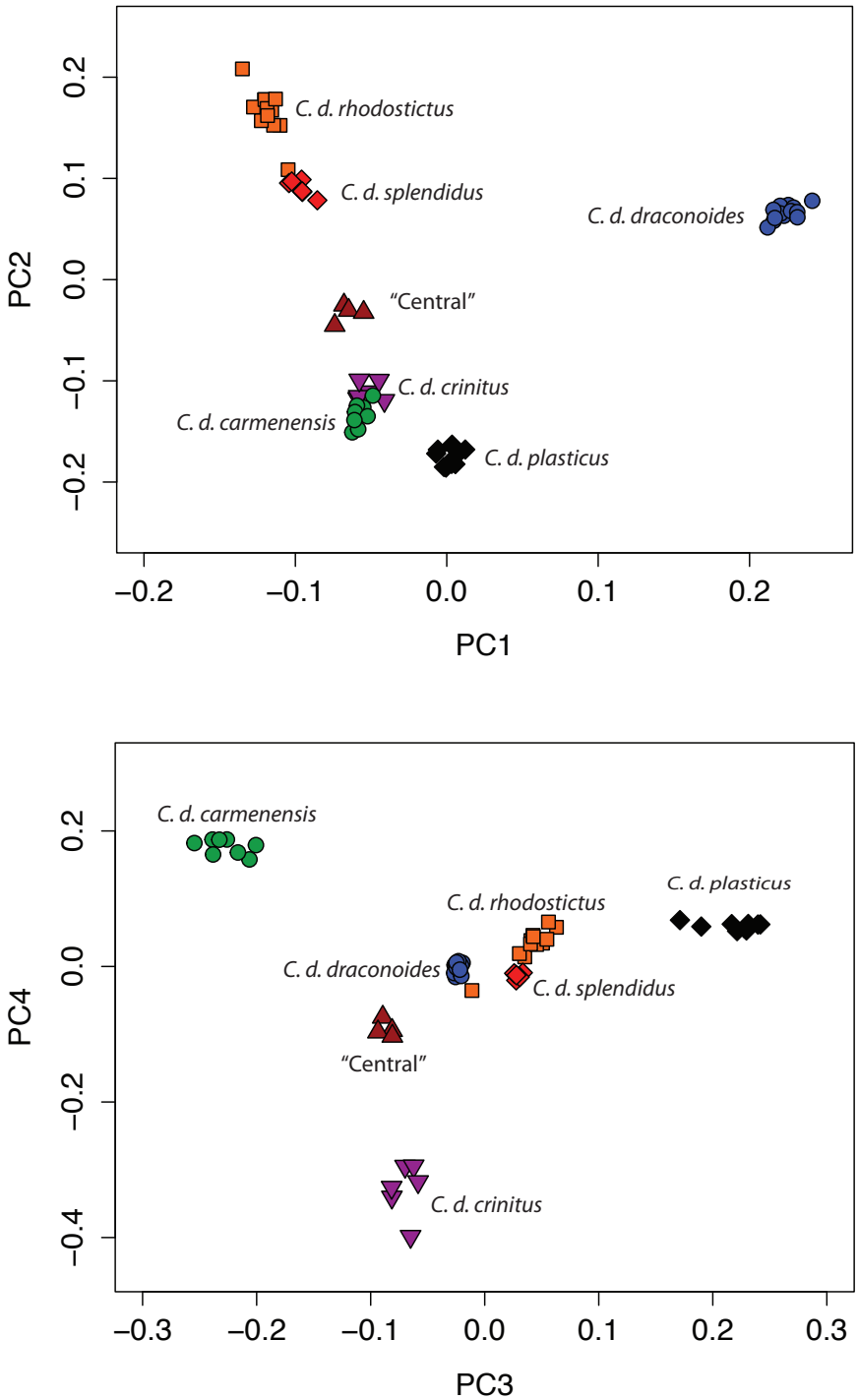

Figure 3.9. PCA analysis of *Callisaurus* using 994 unlinked SNPs. PC1, PC2, PC3, and PC4 account for 13.8%, 9.2%, 7.1%, and 5.7% of the variance, respectively.

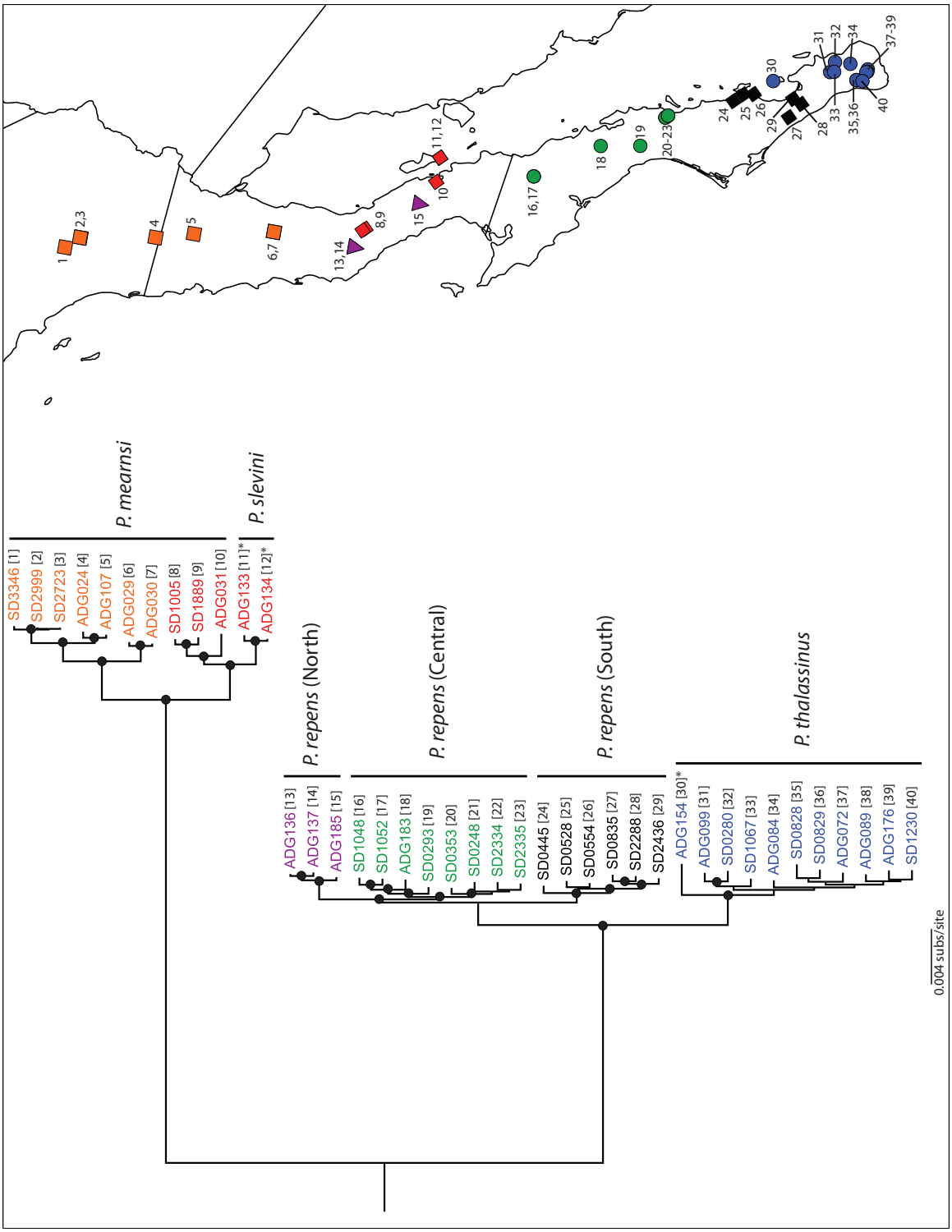

Figure 3.10. Maximum likelihood phylogram of *Petrosaurus*. Nodes with bootstrap support > 70% are denoted with solid black circles on nodes, and numbers in brackets next to taxa labels (left) correspond to the numbers on the map (right).

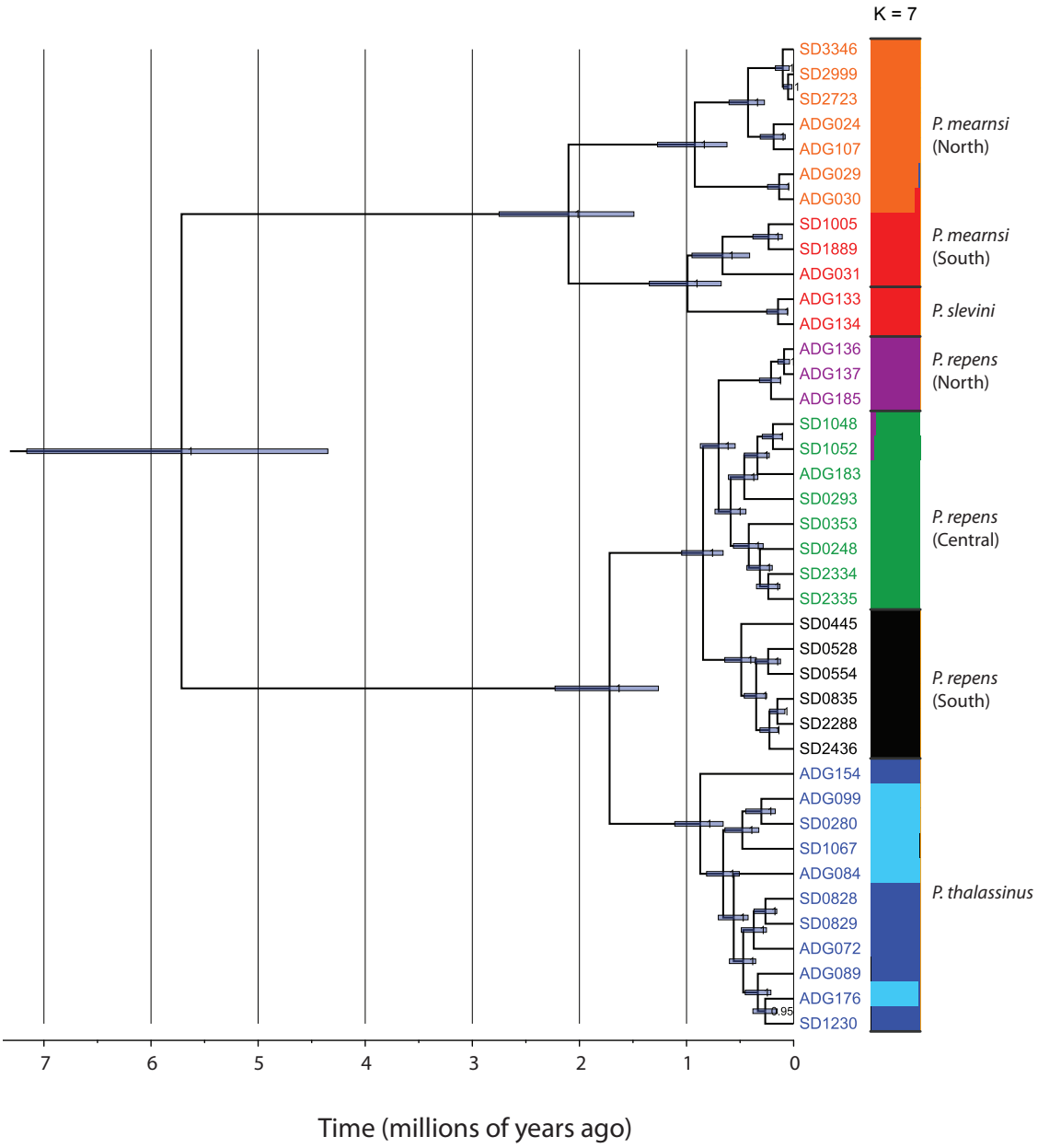

Figure 3.11. Clock tree for *Petrosaurus* (left), with Admixture plot (right) under K=7 (932 unlinked SNPs).

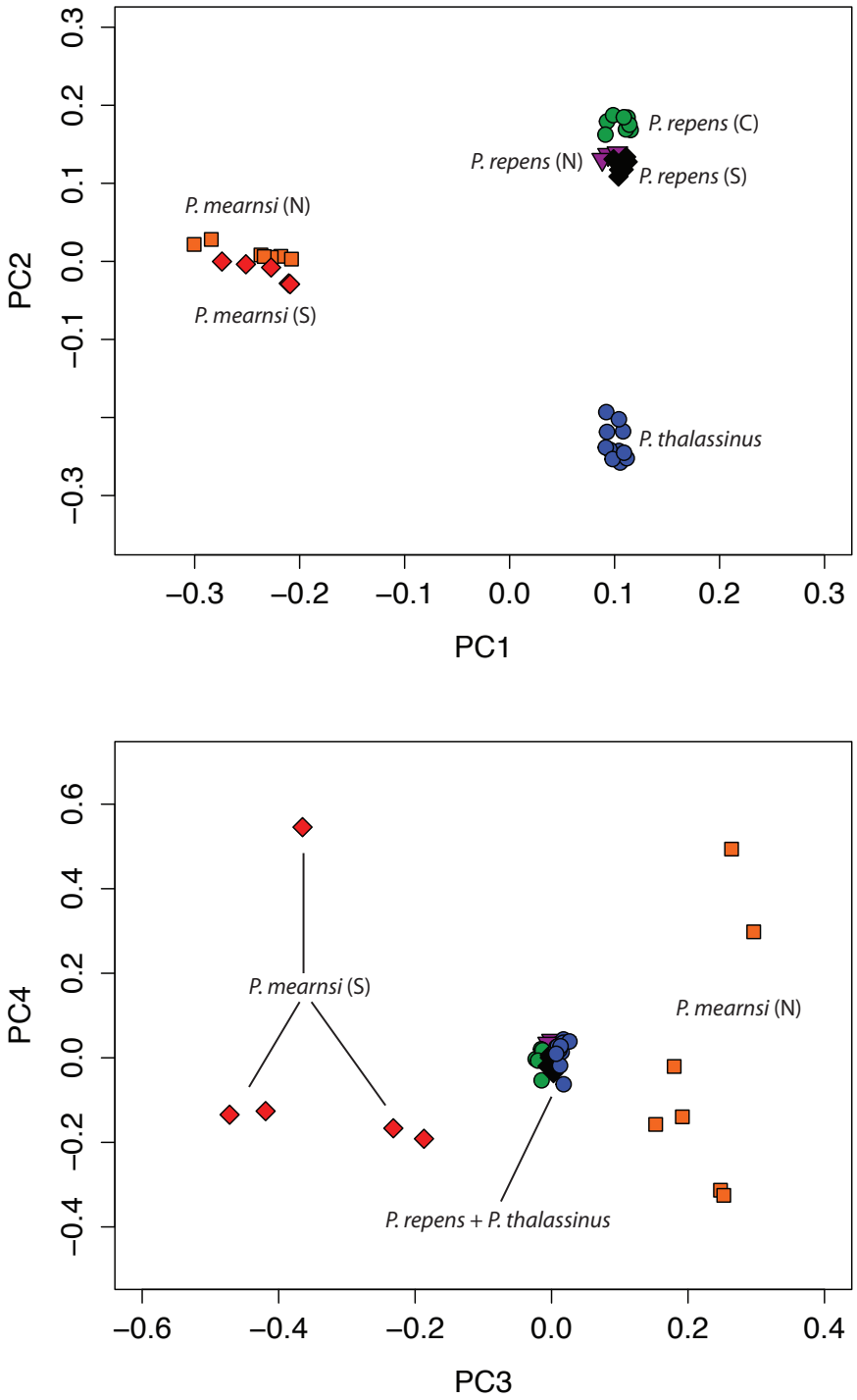

Figure 3.12. PCA analysis of *Petrosaurus* (932 unlinked SNPs). PC1, PC2, PC3, and PC4 account for 36.0%, 8.6%, 6.0%, and 3.7% of the variance, respectively.

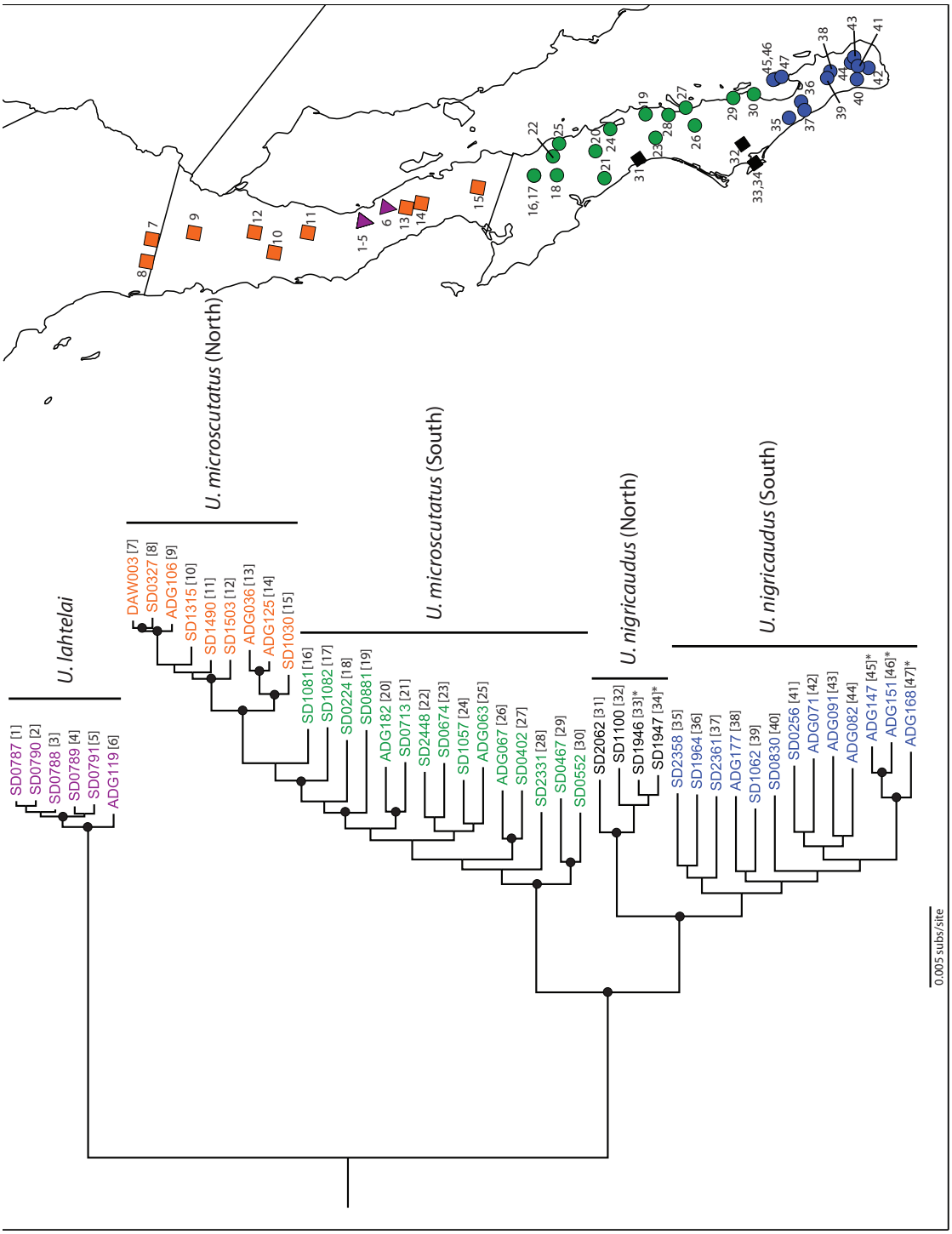

Figure 3.13. Maximum likelihood phylogram of *Urosaurus*. Nodes with bootstrap support > 70% are denoted with solid black circles on nodes, and numbers in brackets next to taxa labels (left) correspond to the numbers on the map (right).

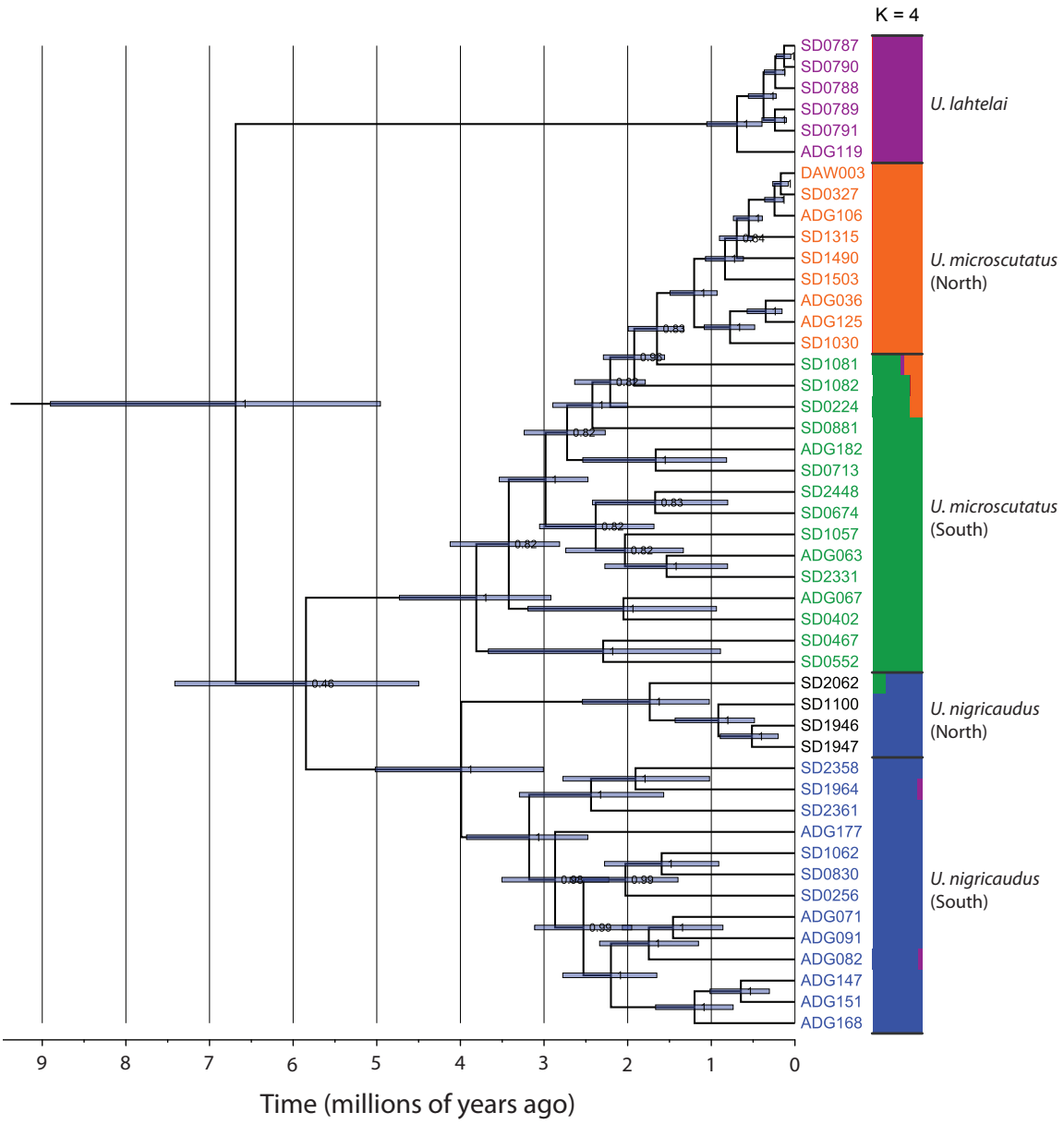

Figure 3.14. Clock tree for *Urosaurus* (left), with Admixture plot (right) under K=4 (1036 unlinked SNPs).

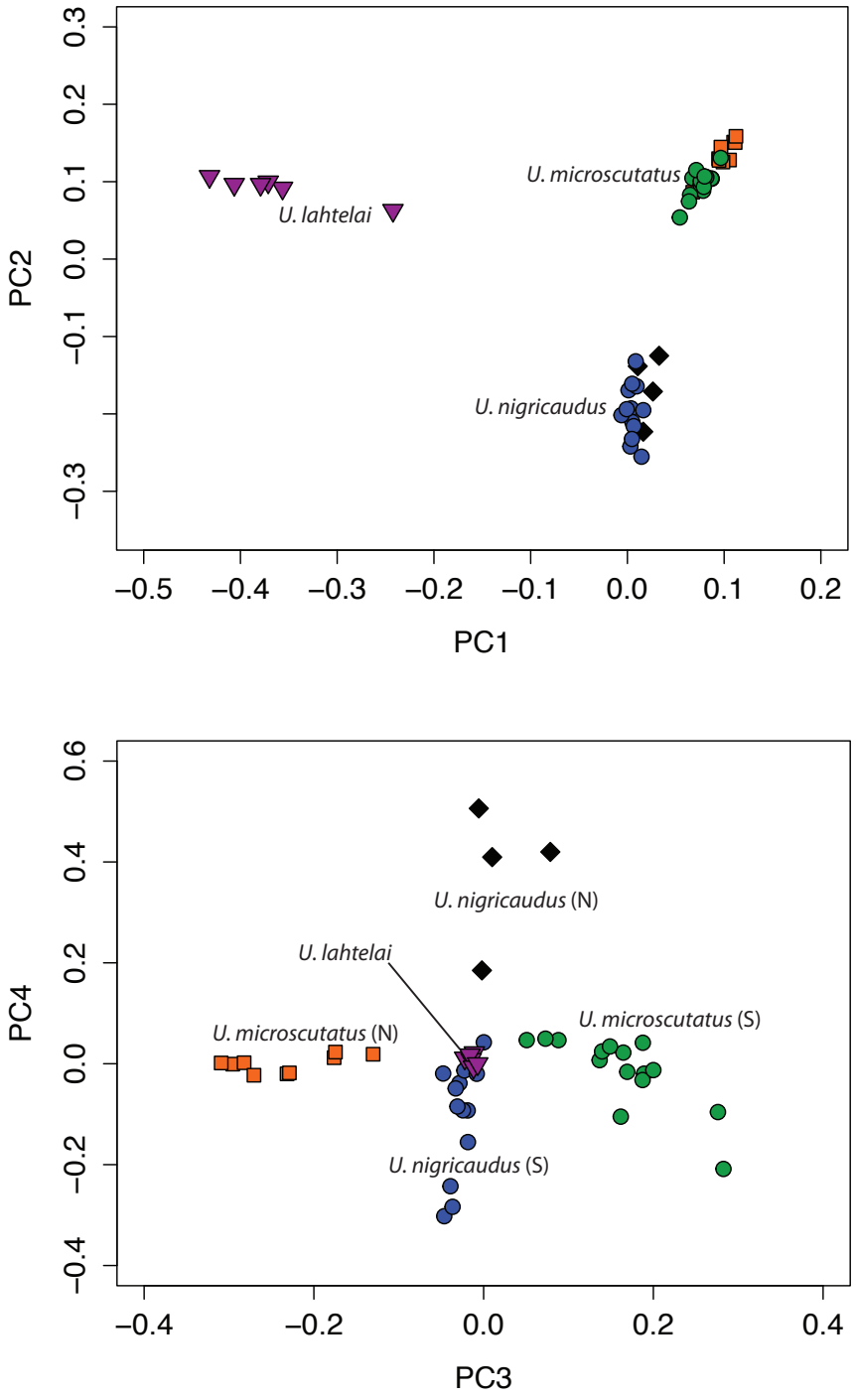

Figure 3.15. PCA analysis of *Urosaurus* (1036 unlinked SNPs). PC1, PC2, PC3, and PC4 account for 14.2%, 11.1%, 5.2%, and 3.6% of the variance, respectively.

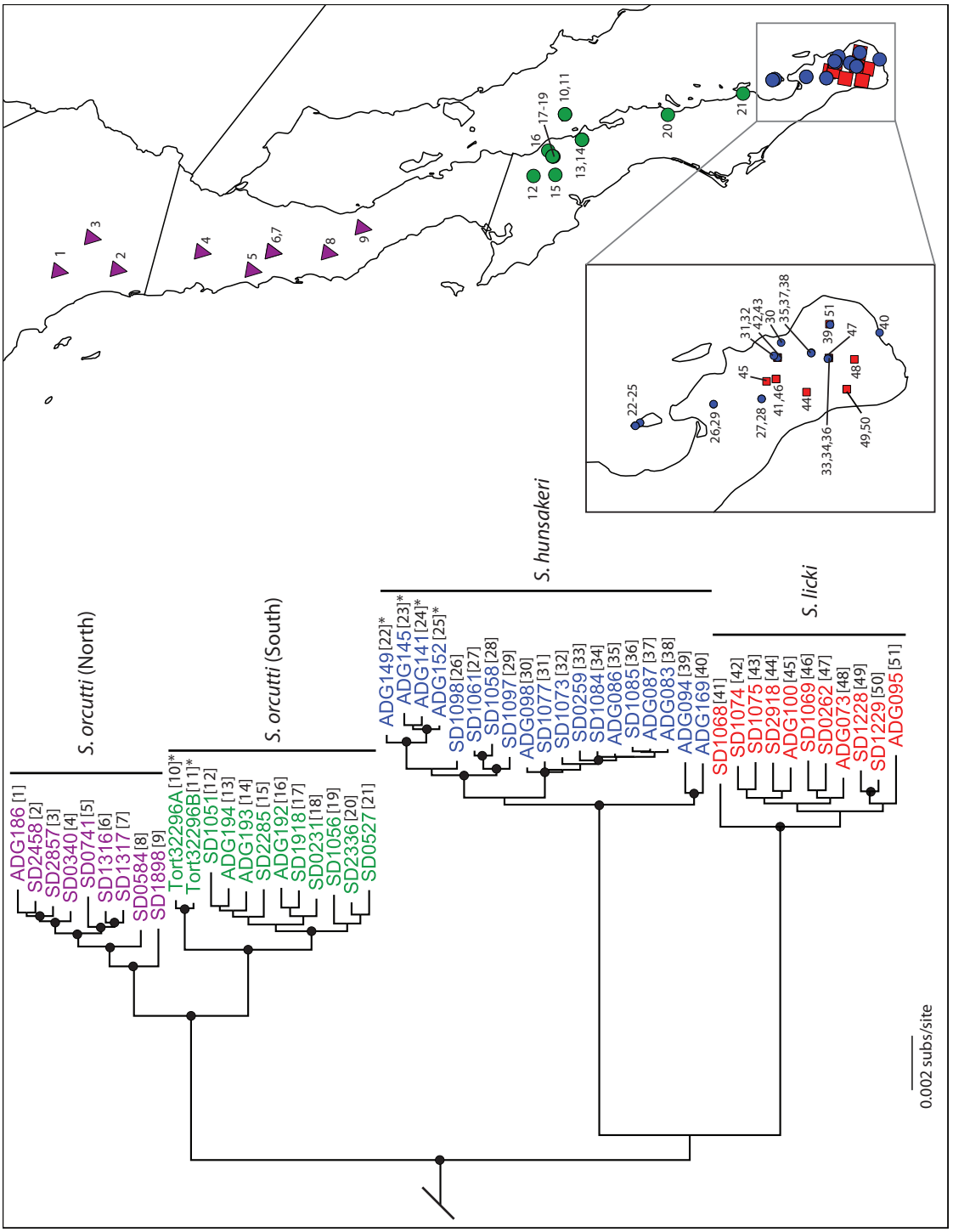

Figure 3.16. Maximum likelihood phylogram of the *Sceloporus orcutti* group. Nodes with bootstrap support > 70% are denoted with solid black circles on nodes, and numbers in brackets next to taxa labels (left) correspond to the numbers on the map (right). For ease of viewing, the other half of the tree is displayed in Figure 3.17.

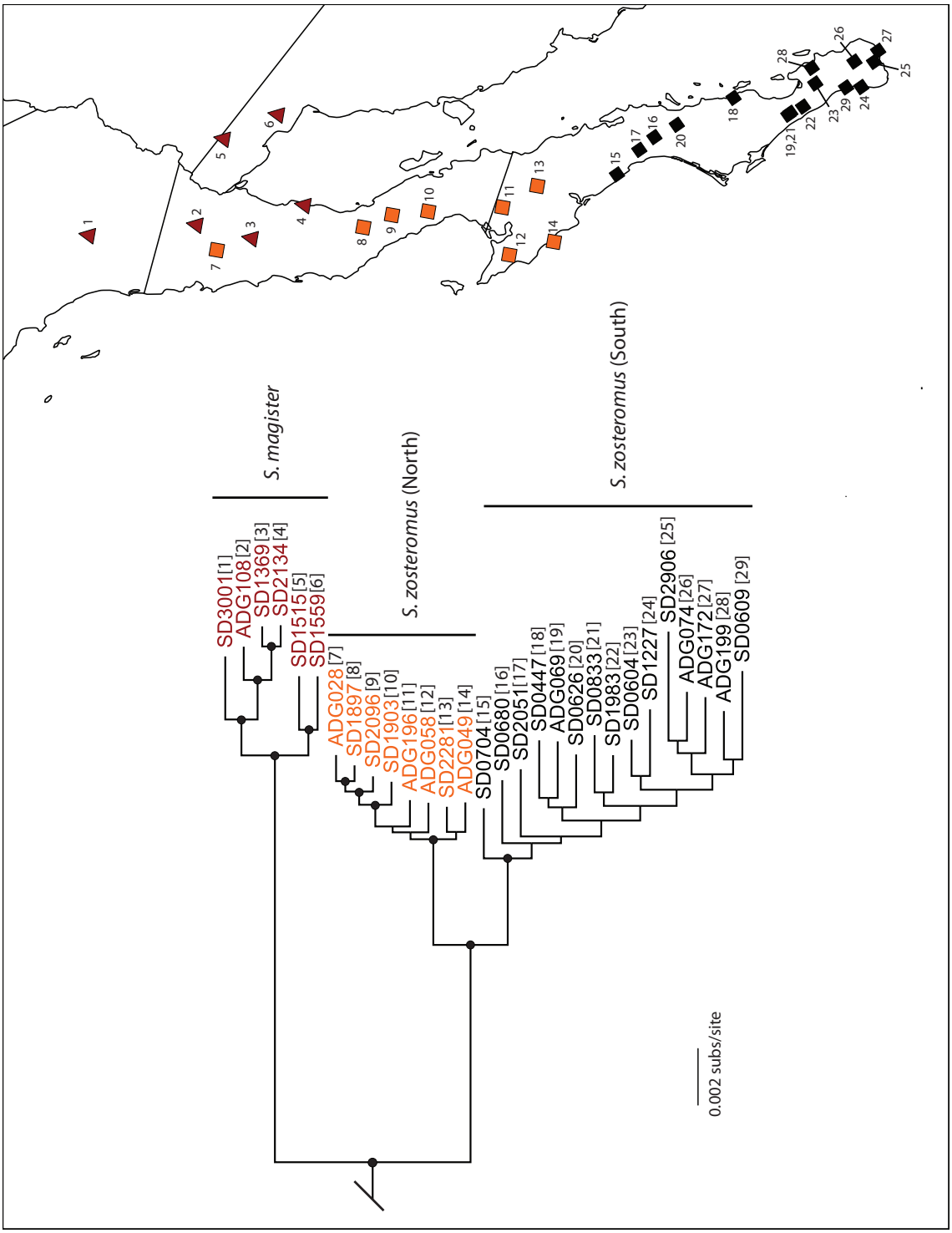

Figure 3.17. Maximum likelihood phylogram of the *Sceloporus zosteromus* group. Nodes with bootstrap support > 70% are denoted with solid black circles on nodes, and numbers in brackets next to taxa labels (left) correspond to the numbers on the map (right).

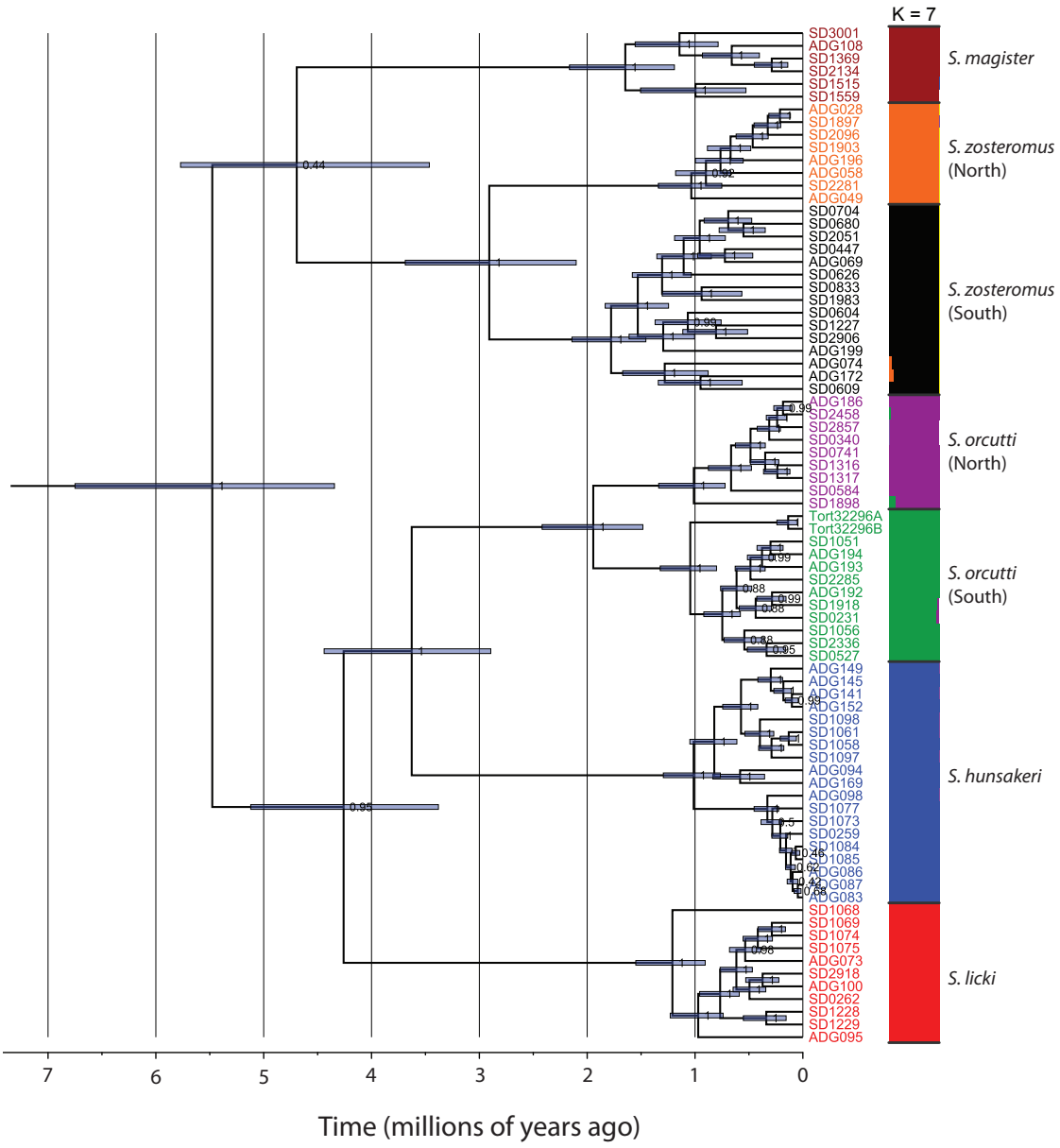

Figure 3.18. Clock tree for *Sceloporus* (left) with Admixture plot (right) under K=7 (1015 unlinked SNPs).

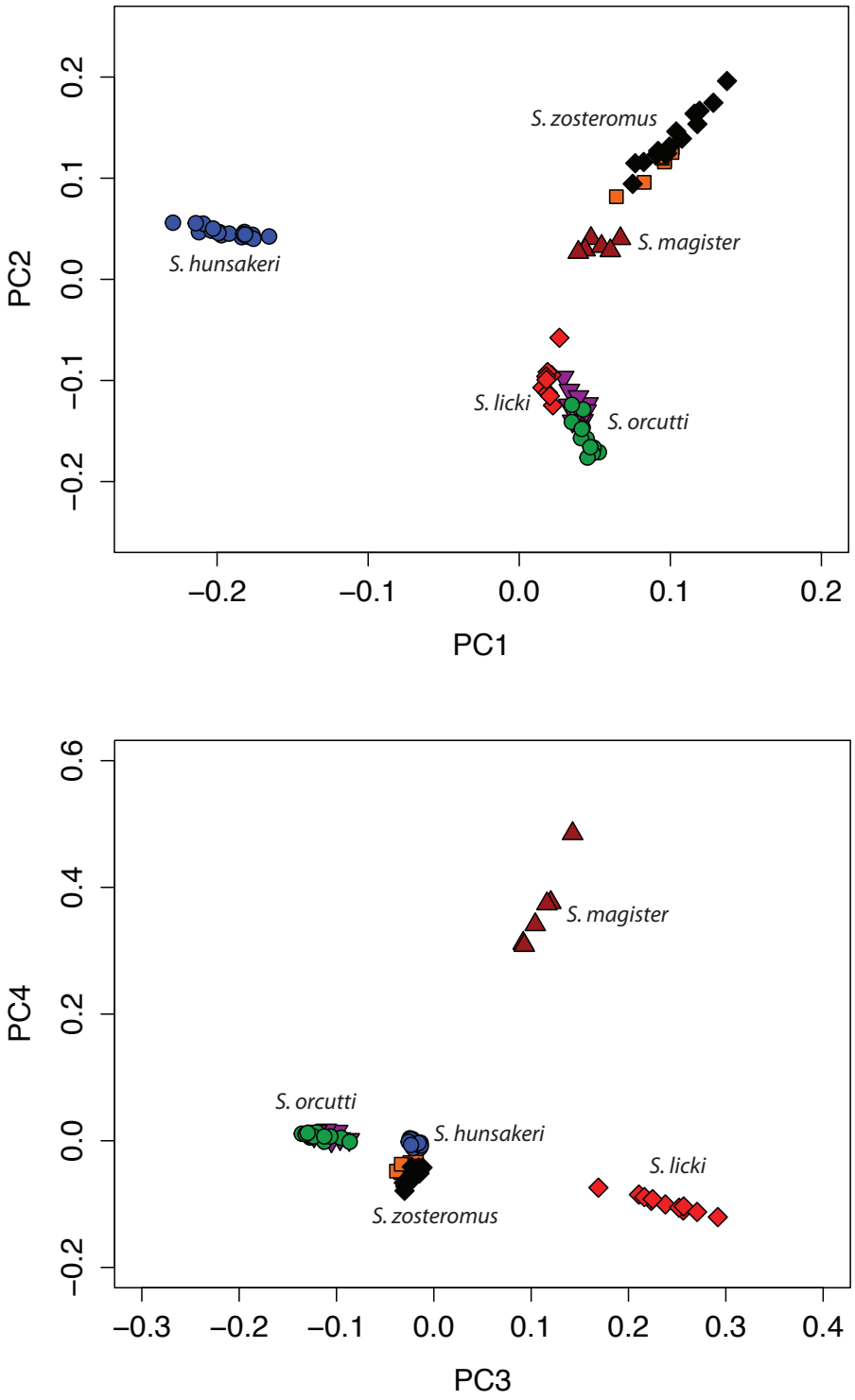

Figure 3.19. PCA *Sceloporus* (1015 unlinked SNPs). PC1, PC2, PC3, and PC4 account for 14.0%, 11.1%, 8.9%, and 8.4% of the variance, respectively.

Table 3.1. Operational Taxonomic Units (OTUs) included in this study. The taxonomy follows Grismer (2002), except that Grismer's intraspecific "pattern classes" are here treated as subspecies, and Grismer did not recognize those taxa marked with an asterisk (*).

OTU	Distribution	Habitats	Citation(s)
Callisaurus draconoides rhodostictus	Northern BCP from W of Bahía de los Angeles to San Gorgonio Pass and beyond, E of Sierra Juarez and Sierra San Pedro Mártir	Washes, bajadas, dunes, beaches in Colorado Desert scrub	Cope 1896
Callisaurus draconoides splendidus	Isla Ángel de la Guarda	Insular endemic; rocky shores, cobblestone beaches, gravel washes	Dickerson 1919
Callisaurus draconoides crinitus	Vizcaíno Desert	Sand dunes and sandy beaches with Pacific climatic influence	Cope 1896
Callisaurus draconoides carmenensis	S of Sierra San Francisco, E of Laguna San Ignacio, N of Loreto and Magdalena Plain, including Isla Carmen	Subtropical arroyos, washes, bajadas, beaches, islands	Dickerson 1919
**Callisaurus draconoides plasticus*	Isthmus of La Paz and S Magdalena Plain, including Isla San Francisco	Washes and arroyos in subtropical desert scrub, bajadas, beaches, islands	Dickerson 1919
Callisaurus draconoides draconoides	Cape region, SE of Isthmus of La Paz, including Isla Espiritu Santo, Sierra de la Laguna, and Cabo San Lucas	Dry tropical mountain arroyos, thorn scrub, dry forest, beaches, islands	Blainville 1835
Petrosaurus mearnsi	Northern BCP from W of Bahía de los Angeles to San Gorgonio Pass E of Sierra Juarez and Sierra San Pedro Mártir	Steep, rocky canyons and hillsides with large boulders and vertical cliffs	Stejneger 1894
Petrosaurus slevini	Isla Ángel de la Guarda and Isla Mejia	Steep, rocky canyons and hillsides with large boulders and vertical cliffs	Van Denburgh 1922
Petrosaurus repens	Central BCP mountains from S of Sierra San Pedro Martír to Isthmus of La Paz	Steep, rocky canyons and hillsides with large boulders and vertical cliffs	Van Denburgh 1895
Petrosaurus thalassinus	Cape region, SE of Isthmus of La Paz, including Isla Espiritu Santo, Sierra de la Laguna, and Cabo San Lucas	Steep, rocky canyons and hillsides with large boulders and vertical cliffs	Cope 1863

OTU	Distribution	Habitats	Citation(s)
Urosaurus nigricaudus	Southern BCP, including cape region (Isla Espiritu Santo, Sierra de la Laguna, and Cabo San Lucas), Isthmus of La Paz, and Magdalena Plain	Rocky and brushy hillsides, canyons, and washes	Cope 1864
**Urosaurus microscutatus*	Northern and central BCP from N of Sierra Juarez to Isthmus of La Paz, excluding Vizcaino Desert and Magdalena Plain	Rocky and brushy hillsides, canyons, and washes	Van Denburgh 1894
Urosaurus lahtelai	Cataviña region	Restricted to granite outcrops and palm oases	Rau and Loomis 1977
Sceloporus zosteromus rufidorsum	Northern BCP from W of Sierra San Pedro Martír to Vizcaino Desert	Coastal succulent scrub, chaparral, desert scrub, beaches, sand dunes, washes	Yarrow 1882
Sceloporus zosteromus monserratensis	Central BCP from Vizcaino Desert to Isthmus of La Paz	Desert plains, thorn scrub, washes	Van Denburgh and Slevin 1921
Sceloporus zosteromus zosteromus	Cape Region, including Sierra de la Laguna	Tropical thorn scrub, tropical dry forest, washes	Cope 1863
Sceloporus magister	Northeastern BCP, E of Sierra Juarez and Sierra San Pedro Martír	Open desert bajadas and washes	Hallowell 1854
Sceloporus orcutti	Entire BCP except Cape Region and Vizcaino Desert, from San Gorgonio Pass to Isthmus of La Paz	Rocky outcrops in coastal scrub, chaparral, and desert, especially near riparian zones	Stejneger 1893
Sceloporus licki	Cape Region mountains (Sierra de la Laguna, Sierra la Trinidad)	Rocky outcrops and wooded areas in tropical dry forest	Van Denburgh 1895
Sceloporus hunsakeri	Cape Region, including Sierra de la Laguna, Isla Espirtuo Santo and Partida, Cabo San Lucas)	Rocky outcrops in tropical thorn scrub and tropical dry forest	Hall and Smith 1979

Table 3.1 (continued).

Table 3.2. Summary statistics for RADseq data. Nreads = total number of reads after demultiplexing, Pass.total = total number of reads that passed quality filter, Cpass = number of clusters that passed the coverage threshold (10x), Dpass = average depth of coverage of clusters that passed the threshold, nloci = total number of loci per lizard after aligning across individuals, floci = number of loci with >10x depth and passed the paralog filter, nsites = number of sites across all loci, npoly = number of polymorphic sites.

	Nreads	**Pass.total**	**Cpass**	**Dpass**	**Floci**	**Nsites**	**Npoly**
Total (*Callisaurus*)	73,284,891	69,008,599	348,979	12,055	323,088	27,434,654	94,193
Mean (*Callisaurus*)	1,201,392	1,131,289	5,721	198	5,297	449,748	1,544
Total (*Petrosaurus*)	44,722,564	41,944,862	263,447	7,940	240,750	20,447,818	44,273
Mean (*Petrosaurus*)	1,118,064	1,048,622	6,586	198	6,019	511,195	1,107
Total (*Urosaurus*)	57,748,107	54,025,229	175,086	11,333	159,874	13,576,078	53,038
Mean (*Urosaurus*)	1,228,683	1,149,473	3,725	241	3,402	288,853	1,128
Total (*Sceloporus*)	79,825,923	75,824,766	456,483	13,897	411,141	34,863,624	127,257
Mean (*Sceloporus*)	997,824	947,810	5,706	174	5,139	435,795	1,591
Total	255,581,485	240,803,456	1,243,995	45,225	1,134,853	96,322,174	318,761
Mean	1,120,971	1,056,156	5,456	198	4,977	422,466	1,398

Table 3.3. Results of the phylogenetic general least squares analysis of latitude (independent variable) and heterozygosity (dependent variable). Significant correlations were found for *Petrosaurus*, *Urosaurus*, and *Sceloporus*, but not *Callisaurus* ($p < 0.05$). Std. Error = Standard Error, df = degrees of freedom.

Genus	Slope	Std. Error	*t*	*p*	df
Callisaurus	6.61E-05	9.37E-05	0.706	0.48	61
Petrosaurus	-2.55E-04	1.09E-04	-2.339	0.02	40
Urosaurus	-4.48E-04	1.05E-04	**-4.283**	1.0E-04	47
Sceloporus	-3.66E-04	1.77E-04	-2.069	0.04	80

Table 3.4. Cross-validation error (CVE) values estimated with Admixture. Bold-faced values represent selected K values for subsequent plots.

K	*Callisaurus*	*Petrosaurus*	*Urosaurus*	*Sceloporus*
2	0.405	0.440	0.605	0.448
3	0.341	0.357	0.557	0.340
4	0.318	0.348	**0.530**	0.272
5	0.308	0.282	0.567	0.236
6	0.312	0.330	0.573	0.256
7	**0.288**	**0.255**	0.598	**0.207**
8	0.296	0.288	0.568	0.213
9	0.300	0.379	0.636	0.219

Table 3.5: F_{ST} values estimated by Admixture for *Callisaurus* (K=7). CDCR = *C. d. crinitus*, CDSP = *C. d. splendidus*, CDCE= "Central" population, CDCA = *C. d. carmenensis*, CDDR = *C. d. draconoides*, CDPL = *C. d. plasticus*.

	CDCR	CDSP	CDCE	CDCA	CDRH	CDDR
CDCR						
CDSP	0.61					
CDCE	0.43	0.63				
CDCA	0.48	0.69	0.49			
CDRH	0.43	0.47	0.41	0.47		
CDDR	0.61	0.78	0.64	0.66	0.57	
CDPL	0.52	0.72	0.56	0.58	0.51	0.64

Table 3.6: F_{ST} values estimated by Admixture for *Petrosaurus* (K=7). PTH1 = *P. thalassinus* (first cluster), PRES = *P. repens* (north), PMEN = *P. mearnsi* (north), PTH2 = *P. thalassinus* (second cluster), PREN = *P. repens* (north), PMES = *P. mearnsi* (south), PREC = *P. repens* (central).

	PTH1	PRES	PMEN	PTH2	PREN	PMES
PTH1						
PRES	0.61					
PMEN	0.43	0.63				
PTH2	0.48	0.69	0.49			
PREN	0.43	0.47	0.41	0.47		
PMES	0.61	0.78	0.64	0.66	0.57	
PREC	0.52	0.72	0.56	0.58	0.51	0.64

Table 3.7. F_{ST} values estimated by Admixture for *Urosaurus* (K=4). UMIS = *U. microscutatus* (south), UNI = *U. nigricaudus*, ULA = *U. lahtelai*, UMIN = *U. microscutatus* (north).

	UMIS	UNI	ULA
UMIS			
UNI	0.45		
ULA	0.74	0.70	
UMIN	0.41	0.59	0.88

Table 3.8. F_{ST} values estimated by Admixture for *Sceloporus* (K=7). SZON = *S. zosteromus* (north), SZOS = *S. zosteromus* (south), SHU = *S. hunsakeri*, SORS = *S. orcutti* (south), SMA = *S. magister*, SORN = *S. orcutti* (north), SLI = *S. licki*.

	SZON	SZOS	SHU	SORS	SMA	SORN
SZON						
SZOS	0.52					
SHU	0.83	0.75				
SORS	0.78	0.70	0.79			
SMA	0.78	0.68	0.81	0.76		
SORN	0.79	0.71	0.80	0.53	0.77	
SLI	0.79	0.70	0.81	0.74	0.78	0.77

Table 3.9. Comparison of timing of spatially concordant phylogeographic breaks across taxa. Mean divergence date estimates are in millions of years ago (mya) while numbers in parentheses represent 95% highest posterior density (HPD) intervals.

Lineage	Vizcaíno Desert	Loreto / Sierra de la Giganta	Isthmus of La Paz	Isla Ángel de la Guarda	Islas Espiritu Santo / Partida
Callisaurus	2.17 (1.62-2.71)	3.01 (2.50-3.54)	2.21 (1.52-2.95)	2.11 (1.65-2.63)	0.75 (0.57-0.98)
Petrosaurus	0.70 (0.55-0.87)	0.85 (0.66-1.05)	1.72 (1.26-2.22)	0.99 (0.68-1.35)	0.87 (0.66-1.11)
Urosaurus	1.65 (1.33-1.99)	5.85 (4.50-7.41)	3.99 (3.01-5.02)	n/a	2.20 (1.65-2.78)
Sceloporus orcutti complex	1.94 (1.48-2.42)	n/a	3.63 (2.89-4.44), 4.26 (3.38-5.12)	n/a	0.57 (0.42-0.74)
Sceloporus magister complex	2.91 (2.10-3.69)	n/a	n/a	n/a	n/a

CONCLUSIONS

The overall theme of this dissertation revolves around the interplay of geological evolution and biological evolution in western North America along the San Andreas Fault (SAF) system. In the first chapter, I reviewed the zoogeographic literature and demonstrated that four GPFZs in the Pacific plate correspond with zoogeographic and phylogeographic boundaries along the SAF system in western North America. These four boundaries act as phylogeographic breaks or distributional limits for a wide variety of marine and terrestrial animals, mostly those with limited dispersal abilities. By reviewing published geological literature, I showed that the GPFZs are the sole remaining trace of ancient transform faults in the now-destroyed EPR before it met the Farallon Trench bordering North America in the late Oligocene. The GPFZs also constrained the positions and orientations of microplates and triple junctions during the evolution of the SAF system, and I hypothesize that this ultimately resulted in zoogeographic boundaries that reflect tectonic boundaries. However, a variety of proximate mechanisms could drive speciation across these boundaries, and more research is needed in this area. These boundaries may act directly as barriers to dispersal or through vicariance, or indirectly by establishing ecological or climatic gradients. This review is important because despite being readily visible with free internet software, discussion of the GPFZs is almost entirely absent from the phylogeography literature. Comparative phylogeographic analysis of co-distributed lineages with similar life-history

traits is likely the most promising approach for testing the hypotheses presented here. More collaboration among marine and terrestrial phylogeographers with geologists is needed to advance this field.

In the second chapter, I utilized advances in next generation sequencing and coalescent-models to delimit species of fringe-toed lizards of the *Uma notata* species complex. I observed a strong latitudinal gradient in genetic diversity across species, which is likely related to climatic cycles during the Pleistocene. Northwestern populations of *U. scoparia* and *U. inornata* likely represent recent range expansions. I found that a model consisting of six species within this group (including the outgroup, *U. scoparia*) was decisively supported over all alternative models, including all previously published hypotheses and the current taxonomy. This study may have important conservation implications if *U. cowlesi* is elevated to full species status and *U*. sp. nov. from the Mohawk Dunes would be described, as these lineages could be vulnerable to anthropogenic impacts due to their restricted ranges. This chapter highlights the continued importance of basic systematic biology for biodiversity conservation.

In the third chapter, I analyzed a large RADseq dataset to investigate the comparative phylogeography of phrynosomatid lizards of the Baja California Peninsula (BCP) to better understand the relationship between tectonic evolution and speciation. I found that heterozygosity was positively correlated with latitude in *Callisaurus*, negatively correlated in *Urosaurus* and *Petrosaurus*, and not correlated at all in *Sceloporus*, suggesting that these taxa responded in idiosyncratic ways to changing environmental conditions. I identified five

regions of the BCP where two or more taxa demonstrate spatially concordant phylogeographic boundaries, and I infer multiple episodes of divergence must have occurred across each of these boundaries. Thus, I reject the hypothesis that all taxa diverged simultaneously at each barrier due to a single vicariance event such as a trans-peninsular seaway. Instead, I interpret these results to support the hypothesis that plate tectonic events ultimately facilitated speciation indirectly by establishing a complex landscape matrix with numerous barriers to dispersal that generated the idiosyncratic patterns observed here.

Taken as a whole, why is this work important in a broader, pragmatic sense? Southwestern North America is a region containing rich biodiversity, much of it endemic. It is difficult to disentangle the complex landscape from the life forms that inhabit it. Unfortunately, this region is currently under a great deal of pressure from human activities. People are fragmenting the landscape with roads, cities, farms, and other developments that eliminate dispersal corridors for some species and open new corridors for other species. Climate change is already occurring and is predicted to intensify heat waves and drought, and is especially insidious because no part of the globe, no matter how pristine and remote, can escape its effects. Continued population growth and development can only amplify the pressure on biodiversity. I argue that from a conservation and management perspective, it would be ideal if we could *predict* how these human impacts might affect the continued diversification of life in southwestern North America, but constructing predictive models is impossible without first understanding how historical geological and climatic events

influenced previous bouts of speciation. The bottom line is that historical biogeography and species delimitation are important for conservation because we cannot effectively protect biodiversity unless we know exactly what is out there and how it evolved in the first place.

APPENDIX A: GLOSSARY

To minimize confusion, I have provided a brief glossary of geological terms and local landscape features frequently used in Chapter 1. All definitions are paraphrased based on references cited herein.

Geological Terms

Fracture Zone (FZ): a linear band of mountainous topography, consisting of long ridges, escarpments, and lines of volcanoes, which offsets and is roughly perpendicular to the crest of a mid-oceanic ridge. A fracture zone is an aseismic extension of a transform fault between plates of oceanic crust moving away from seafloor spreading centres (Menard and Atwater, 1969).

Great Pacific Fracture Zone (GPFZ): any of a series of especially large FZs that extends thousands of kilometers across the Pacific Ocean, running parallel to other GPFZs, including the Mendocino, Murray, Molokai, and Clarion FZs (Menard, 1967).

Lithosphere: the outermost layer of the solid Earth, consisting of the rigid crust and the elastic upper mantle, which is fragmented into a series of mobile and dynamic tectonic plates. Oceanic lithosphere is continuously generated at spreading centres and destroyed at subduction zones (McKenzie and Parker, 1967).

Magnetic anomaly: a local disturbance in the Earth's magnetic field arising from geomagnetism of the Earth's crust. Bands of magnetic anomalies of the same age (isochrons) typically run parallel to the axis of a spreading centre. The magnetic bands result from iron molecules orienting themselves with respect to the direction of the Earth's magnetic field at the time the new lithosphere solidified. Because the polarity of Earth's magnetic field periodically reverses, the isochrons can be dated to specific time intervals. Maps of magnetic anomaly isochrons are one of the most important sources of data for reconstructing past plate movements (McKenzie and Parker, 1967; Atwater, 1989).

Microplate: a tectonic plate small enough to be easily subducted beneath or accreted to larger tectonic plates, e.g. the Monterey microplate (Nicholson *et al*., 1994).

Palinspastic reconstruction: a map showing the previous location of geological features while correcting for any plate movements along fault lines, such as strike-slip displacement or microplate rotation (e.g. Hall, 2002*a*).

Propagating rift: an extensional plate boundary that progressively breaks through rigid lithosphere, which when in combination with sea-floor spreading, produces a V-shaped wedging of lithosphere formed at the propagating spreading centre (Hey *et al.*,1989). Propagating rifts associated with left-offset transform faults may have been responsible for both the rotation of the Transverse Ranges and opening of the Gulf of California in the Miocene (Hey, 1998). See also **spreading centre**.

Ridge: see **spreading centre.**

Rift: see **spreading centre.**

Rise: see **spreading centre.**

Spreading centre (a.k.a. **rift**, **rise**, or **ridge**): a divergent plate boundary, where sea-floor spreading (rifting) occurs, allowing magma to well up to the crustal surface and cool into new oceanic lithosphere**,** e.g. the **East Pacific Rise**. Sea-floor spreading accounts for why lithosphere typically grows older as one travels away from a ridge at a high angle along the "conveyor belt."

Strike-slip fault: see **transform fault.**

Subduction zone (a.k.a. **trench**): a convergent plate boundary where one plate (usually a dense oceanic plate) is melted into the magma beneath another plate (usually a buoyant continental plate), creating a volcanic arc in the process, e.g. the Cascadia subduction zone**.**

Tectonic plate: a major block of the lithosphere that behaves as a mostly rigid unit while interacting with other plates and drifting atop the Earth's molten core. There are at least twelve commonly recognized major plates covering the Earth, e.g. the Pacific and North American plates. See also: **microplate.**

Transform fault (a.k.a. **strike-slip fault**): a tectonic plate boundary in which the direction of plate movement is completely or mostly parallel to the orientation of the fault, so that the two plates slide past one another (Wilson, 1965). This frequently occurs in the portion of a fracture zone between offset spreading centres, and is responsible for the large offsets in magnetic anomalies observed across the Great Pacific Fracture Zones (Atwater, 1970).

The SAF is a unique transform fault that resulted when the Pacific and North American plates made contact, splintering the Farallon plate into a series of plates and microplates (Atwater, 1970, 1989). If plate motions are oblique with respect to the orientation of the transform fault, transpression or transtension may result in uplift, rotation and deformation of plate margins, as observed in the Transverse Ranges (Nicholson *et al*., 1994).

Transpression: a type of strike-slip deformation deviating from simple shear due to oblique shortening motion and compression, e.g. Pliocene uplift of the Transverse Ranges along the SAF following Miocene rotation (Nicholson *et al*., 1994).

Transtension: a type of strike-slip deformation deviating from simple shear due to oblique extension motion, often resulting in crustal stretching, thinning, or rifting, e.g. oblique rifting in the Gulf of California (Lonsdale, 1989).

Trench: see **subduction zone.**

Triple junction: a point of contact of three plate boundaries, e.g. the Mendocino triple junction in the north or the Rivera triple junction in the south (McKenzie and Morgan, 1969; Furlong and Schwartz, 2004).

Landscape Features

Baja California Peninsula: a 1,200 km peninsula, separated from mainland México by the Gulf of California, now attached to the Pacific plate south of California (Fig. 3).

Cabo Corrientes: the prominent headland located to the east of the Rivera triple junction (southern limit of the SAF system) at approximately 20º N latitude in Jalisco, México, marking the southernmost limit of the Gulf of California geomorphic province (Fig. 3).

Cape Mendocino: the prominent headland associated with the Mendocino triple junction (northern limit of the SAF system) in northwestern California at approximately 40º N latitude (Fig. 2).

Cascades: the volcanic arc extending from southern British Columbia to northern California, associated with the subduction of the Juan de Fuca plate (Fig. 2).

Cascadia subduction zone: the offshore trench lying north of Cape Mendocino and to the west of northern California, Oregon, Washington, and British Columbia, that forms a convergent boundary between the Juan de Fuca and North American plates and has generated the volcanic Cascades (Fig. 2).

Central Mexican Plateau: a large arid plateau in mainland México extending from the USA border in the north to the Sierra Transvolcanica in the south, and bounded by the Sierra Madre Occidental and Sierra Madre Oriental to the west and east, respectively.

Central Valley: the broad valley in California bounded in the north by the Cascades, in the east by the Sierra Nevada, in the south by the Transverse Ranges, and in the west by the Coast Ranges (Fig. 2).

Clarion Fracture Zone: a GPFZ extending to the west of the Rivera triple junction in mainland México (Fig. 1).

Coast Ranges: the belt of mountains south of Cape Mendocino, west of the Central Valley, and north of Point Conception and the Transverse Ranges (Fig. 2).

Cocos plate: A medium-sized plate, the southern remnant of the ancient Farallon plate, that lies to the south of the Mesoamerican subduction zone in México and Central America (Fig. 1).

Continental Borderland: the offshore waters of southern California and northern Baja California between Point Conception and Punta Eugenia (Fig. 2), characterized by topographically complex series of islands and basins, most likely formed by transtensional extension in the wake of the rotating Transverse Ranges (Shepard and Emery 1941; Nicholson *et al*., 1994).

East Pacific Rise (EPR): the spreading centre bordering the Pacific Plate on the east from the Cocos plate (in ancient times, the Farallon plate) and the Nazca plate. See also **spreading centre.**

Farallon plate: the ancient oceanic plate, bordered by the East Pacific Rise to the west and the North American Plate to the east, that was split into the Juan de Fuca and Cocos plates (and several microplates) when the East Pacific Rise contacted the Farallon subduction zone in the late Oligocene (Fig. 4).

Farallon subduction zone: see **Farallon Trench**.

Farallon Trench (a.k.a. **Farallon subduction zone**): the ancient convergent plate boundary that separated the Farallon plate from the North American plate. Destruction of the Farallon plate in this trench during the Mesozoic was responsible for the formation of massive granitic batholiths (now exposed by

uplift and erosion) in the Sierra Nevada and Baja California Peninsula. The Farallon subduction zone was split into the Cascadia and Mesoamerican subduction zones in the late Oligocene (Fig. 4).

Gorda Ridge: the spreading centre (the northern remnant of the EPR) that separates the Pacific plate from the southwest portion of the Juan de Fuca plate.

Guatemalan Trench (a.k.a. **Mesoamerican subduction zone**): the trench that separates the Mesoamerican highlands of southern México and Central America from the Cocos Plate, bounded on the northwest by the RTJ (Fig. 3).

Gulf of California: an extension of the Pacific Ocean 100–150 km wide situated between the Mexican mainland and the Baja California peninsula, which floods part of a structural depression that was formed by detachment and oblique separation of the peninsula from the continent along the Pacific-North America plate boundary (Fig. 3).

Juan de Fuca plate: a small plate, the northern remnant of the ancient Farallon plate, that lies to the west of northern California, Oregon, Washington, and British Columbia, to the north of the Mendocino triple junction (Fig. 1).

Klamath Mountains: an ancient mountain range sandwiched between the Cascades and the Coast Ranges in northwestern California and southwestern Oregon (Fig. 2).

Mendocino Fracture Zone: a GPFZ extending to the west of the Mendocino triple junction in northern California (Fig. 1).

Mendocino triple junction (MTJ): the meeting location of the North American, Pacific, and Juan de Fuca plates (transform-transform-trench) off northern California, forming the northern boundary of the SAF system (Fig. 4).

Mesoamerican subduction zone: see **Guatemalan trench**.

Molokai Fracture Zone: a GPFZ extending to the west of the Shirley FZ and Punta Eugenia in Baja California (Fig. 3).

Murray Fracture Zone: a GPFZ extending to the west of Point Conception in southern California (Fig. 2).

North American plate: a large continental plate that constitutes the entire North American continent except Baja California and southwestern California (west of the SAF *sensu stricto*), which are now attached to the Pacific plate (Fig. 1).

Pacific plate: the massive tectonic plate that comprises most of the Pacific Ocean. The Pacific plate borders the North American plate along the SAF system, and is separated from the Juan de Fuca and Cocos plates by the Gorda Ridge and East Pacific Rise, respectively (Fig. 1).

Peninsular Ranges: the axis of mountains forming the Baja California peninsula, bounded in the north by the SAF and Transverse Ranges, in the east by the Gulf of California, and in the west by the Pacific Ocean (Fig. 2).

Point Conception: the prominent headland at the west end of the Transverse Ranges in southern California located at approximately 34.5° N latitude (Fig. 2).

Punta Eugenia: the prominent headland extending west from the Vizcaíno Desert at approximately 28º N latitude in the Baja California peninsula (Fig. 3).

Rivera triple junction (RTJ): the meeting location of the North American, Pacific, and Rivera plates off southern México, forming the southern boundary of the SAF system (Fig. 4). It has two limbs that are trenches and one limb that is an oblique extensional zone (Stock and Lee, 1994).

San Andreas Fault (SAF) *sensu stricto*: the northernmost portion of the SAF system that consists of a right-lateral strike-slip fault between Cape Mendocino and the eastern Transverse Ranges. Bounded in the north by the MTJ, the SAF turns into a rift zone in the Gulf of California geomorphic province (Fig. 4).

San Andreas Fault (SAF) system: the entire Pacific-North America plate boundary between the Mendocino and Rivera triple junctions, including both the Gulf of California where transtensional rifting occurs and the SAF *sensu stricto* where the plate boundary is predominantly strike-slip (Fig. 4).

Shirley Fracture Zone: a FZ connecting the Molokai FZ to Punta Eugenia in Baja California (Fig. 3).

Sierra Madre Occidental: the Mexican mountain range bordered by the Gulf of California on the west and the Central Mexican Plateau on the east (Fig. 3).

Sierra Madre Oriental: the Mexican mountain range bordered by the Central Mexican Plateau on the west and the Gulf of México on the east (Fig. 3).

Sierra Nevada: the mostly igneous mountain range in California, east of the Central Valley (Fig. 2).

Sierra Transvolcanica: the east-west trending volcanic belt extending southeast from Cabo Corrientes at approximately 20º N latitude in mainland México, bounded by the Sierra Madre Occidental, Central Mexican Plateau, and Sierra Madre Oriental to the north (Fig. 3).

Transverse Ranges: the east-west trending mountain ranges in southern California, aligned with Point Conception and the Murray Fracture Zone. Includes the Santa Ynez, Sierra Madre, San Rafael, Sierra Pelona, Santa Monica, San Gabriel, San Bernardino and Little San Bernardino ranges (Fig. 2).

Vizcaíno Desert: the triangle-shaped desert plain to the east of Punta Eugenia at approximately 28º N latitude in the Baja California peninsula (Fig. 3).