2014-12-01

Understanding the Diversification of Central American Freshwater Fishes Using Comparative Phylogeography and Species Delimitation

Justin C. Bagley
Brigham Young University - Provo

Understanding the Diversification of Central American Freshwater Fishes

Using Comparative Phylogeography and Species Delimitation

Justin Colonial Bagley

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Jerald B. Johnson, Chair
Byron J. Adams
Mark C. Belk
Keith A. Crandall
Duke S. Rogers

Department of Biology

Brigham Young University

December 2014

ABSTRACT

Understanding the Diversification of Central American Freshwater Fishes
Using Comparative Phylogeography and Species Delimitation

Justin Colonial Bagley
Department of Biology, BYU
Doctor of Philosophy

Phylogeography and molecular phylogenetics have proven remarkably useful for understanding the patterns and processes influencing historical diversification of biotic lineages at and below the species level, as well as delimiting morphologically cryptic species. In this dissertation, I used an integrative approach coupling comparative phylogeography and coalescent-based species delimitation to improve our understanding of the biogeography and species limits of Central American freshwater fishes. In Chapter 1, I conducted a literature review of the contributions of phylogeography to understanding the origins and maintenance of lower Central American biodiversity, in light of the geological and ecological setting. I highlighted emerging phylogeographic patterns, along with the need for improving regional historical biogeographical inference and conservation efforts through statistical and comparative phylogeographic studies. In Chapter 2, I compared mitochondrial phylogeographic patterns among three species of livebearing fishes (Poeciliidae) codistributed in the lower Nicaraguan depression and proximate uplands. I found evidence for mixed spatial and temporal divergences, indicating phylogeographic "pseudocongruence" suggesting that multiple evolutionary responses to historical processes have shaped population structuring of regional freshwater biota, possibly linked to recent community assembly and/or the effects of ecological differences among species on their responses to late Cenozoic environmental events. In Chapter 3, I used coalescent-based species tree and species delimitation analyses of a multilocus dataset to delimit species and infer their evolutionary relationships in the *Poecilia sphenops* species complex (Poeciliidae), a widespread but morphologically conserved group of fishes. Results indicated that diversity is underestimated and overestimated in different clades by *c.* ±15% (including candidate species); that lineages diversified since the Miocene; and that some evidence exists for a more probable role of hybridization, rather than incomplete lineage sorting, in shaping observed gene tree discordances. Last, in Chapter 4, I used a comparative phylogeographical analysis of eight codistributed species/genera of freshwater fishes to test for shared evolutionary responses predicted by four drainage-based hypotheses of Neotropical fish diversification. Integrating phylogeographic analyses with paleodistribution modeling revealed incongruent genetic structuring among lineages despite overlapping ancestral Pleistocene distributions, suggesting multiple routes to community assembly. Hypotheses tests using the latest approximate Bayesian computation model averaging methods also supported one pulse of diversification in two lineages diverged in the San Carlos River, but multiple divergences of three lineages across the Sixaola River basin, Costa Rica, correlated to Neogene sea level events and continental shelf width. Results supported complex biogeographical patterns illustrating how species responses to historical drainage-controlling processes have influenced Neotropical fish diversification.

Keywords: approximate Bayesian computation, Central America, coalescent, comparative phylogeography, freshwater fishes, genetic breaks, Poeciliidae, species delimitation, species trees

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Jerald B. Johnson. It has been an honor to be his first Ph.D. student. I am very thankful for his investments of time, funding, equipment, and guidance with project design, planning and conducting fieldwork, and writing; all of his efforts have helped to make my dissertation and my broader experience at BYU highly successful. I will never forget that the Ph.D. is not just about where or what you study, it is about who you "become", and I thank Jerry for providing an encouraging example and tremendous academic freedom to me during this journey of becoming. I also extend special thanks to the secondary members of my Ph.D. committee, including Byron J. Adams (BYU), Mark C. Belk (BYU), Keith A. Crandall (George Washington University), and Duke S. Rogers (BYU), for their constructive criticism, career advice, and moral support.

Next, I would like to acknowledge several of my mentors and collaborators at BYU and at other institutions who were instrumental in helping me complete this thesis, and without whom this project would not have been possible. In particular, I am grateful for the friendship and help of former BYU postdoc Peter J. Unmack, who provided mentorship in the lab, including advice on DNA extraction, PCR primers, DNA sequence editing, alignment, and analysis and data submission to online repositories, as well as example datasets. I also thank him for many enjoyable times and lively conversations together on-campus and off, for his intensity, and for inspiring me with his enthusiasm for the evolution, biogeography, and conservation of freshwater fishes. I am also greatly indebted to M. Florencia Breitman (Centro Nacional Patagónico-Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina; BYU), who was an excellent collaborator with me on two of my dissertation chapters and also provided nurturing counsel throughout this process. Florencia assisted with several aspects of the project

through many valuable discussions and help with writing, voucher sample processing, DNA sequence editing and analysis, map-making, and proofreading manuscripts, alongside a thousand minor details. And I thank Mike Hickerson (City University of New York, CUNY) for detailed assistance with approximate Bayesian computation analyses in the msBayes pipeline, for providing workspace for us in his lab during the memorable days we spent in New York, and for so kindly opening his home in Harlem to us as well. Two other unsung champions in the development of this dissertation are Ana Laura Almendra (BYU) and Jessica A. Wooten (The University of Findlay), who I thank for providing GIS data and teaching me many practical aspects of ecological niche modeling, including techniques for manipulating data in DIVA-GIS and ArcMap and running analyses in MaxEnt. Ana also assisted me with helpful discussions of phylogeographic methods, the Unix command line, and Linux resource management scripts.

While writing this dissertation, I had the pleasure of working alongside a great cohort of graduate student peers who contributed to my personal and professional development while at BYU. Most notably, I thank my office mate Eduardo Castro Nallar, an excellent scientist that fascinated and distracted me with virus phylodynamics, for many fun conversations about life and molecular evolution, and his profound friendship. I also thank Cesar Aguilar, Ana Laura Almendra, Arley Camargo, Jared B. Lee, Rafael Leite, Luke J. Welton, Fernanda P. Werneck, and Perry Lee "JR" Wood Jr. for being so much fun to be around, for helpful discussions of methods in phylogeography, phylogenetics, and species delimitation, and for providing helpful comments on earlier versions of my manuscripts. In a similar spirit, I also appreciate enriching discussions, manuscript comments, and encouragement from BYU professor Jack W. Sites Jr.

Regarding collections material and data, I wish to thank numerous people whose efforts in field exploration of freshwater fishes laid the groundwork for this dissertation, and also those

who assisted with material or geographical coordinate data for this study. First, I'll acknowledge the dedicated ichthyological work of my forerunners, including the late Robert Rush Miller (University of Michigan) and George S. Meyers (Stanford University), William A. Bussing (University of Costa Rica), and Eldredge Bermingham (Smithsonian Tropical Research Institute [STRI], Panama; McGill University, Canada), whose years of studying Central American fishes provided a firmer distributional understanding of the regional freshwater fish fauna and basis upon which I could conduct my work, and to whom I am greatly indebted—they are my heroes. I also acknowledge members of the BYU Jerry Johnson Lab for their efforts collecting fishes from Central America for over a decade; you provided me with a convenient, albeit incomplete, starting point of preserved material and curatorial records from which to attempt multi-taxon comparative projects in a more informed fashion. While I collected many of the samples used during this dissertation, I could not have done this alone. I am grateful to Eduardo Castro Nallar, Joey T. Nelson, Ruth G. Reina, Aaron H. Smith, and Eric P. van den Berghe for valuable assistance conducting fieldwork for this project in Costa Rica and Nicaragua. Additionally, I thank Keith A. Crandall and Dennis K. Shiozawa (BYU), Javier Guevara Siquiera of the Sistema Nacional de Areas de Conservación and Ministerio de Ambiente y Energía (SINAC-MINAET, Costa Rica), Edilberto Duarte of the Ministerio del Ambiente y los Recursos Naturales (MARENA, Nicaragua), Eric P. van den Berghe (Zamorano University, Honduras), and Ruth G. Reina (STRI) for help obtaining and issuing collecting, import and export permits. Eldredge Bermingham (STRI) provided me with well-preserved fish tissue and DNA samples of *Poecilia* from throughout the Central American isthmus that comprised the majority of the samples of *Poecilia* for my 3rd and 4th Chapters. I also thank Andy Bentley (Kansas Biodiversity Institute & Natural History Museum), Prosanta Chakrabarty and Caleb D. McMahan (Louisiana State

University), Spencer J. Ingley (BYU), Michael "Michi" Tobler (Kansas State University), and Wilfredo A. Matamoros (University of Southern Mississippi) for graciously donating various ingroup and outgroup samples, most notably of *Astyanax*, *Poecilia* and *Limia* from Honduras, El Salvador, Mexico, and Panama.

Several people are thanked for additional assistance of technical nature during the course of this dissertation. In particular, I thank Simon Joly (Institut de Recherche en Biologie Végétale) and Frank Burbrink (CUNY) for helpful advice on running the JML analyses for my species delimitation chapter. I also thank several technical and support staff from our BYU Department of Biology for their services. In particular, I appreciate the work of Ed Wilcox and his staff at the BYU DNA Sequencing Center, who provided clean sequence data at the best rates I've ever seen. I also benefitted from administrative support from Gentri Glaittli, Christina George, and Jim Wooten, who it seems were always eager to answer questions, help me find and accurately present paperwork, and give advice for managing my professional finances.

I gratefully acknowledge various funding sources that made this dissertation possible. Research was supported by Brigham Young University, including a Mentoring Environment Grant and a Graduate Studies Graduate Research Fellowship; a U.S. National Science Foundation (NSF) Doctoral Dissertation Improvement Grant (DEB-1210883); and Idea Wild (http://www.ideawild.org/). I received stipend support from research assistantships and teaching assistantships provided by the BYU Department of Biology and College of Life Sciences, and from the NSF Partnerships for International Research and Education project OISE-PIRE-0530267. I was also aided by research and travel grants from the BYU Department of Biology, which provided financial assistance for supplies and dissemination of research.

I also acknowledge several people who mentored me towards a career in science, more

broadly, and in the ecology, evolution and biogeography of fishes, in particular. It was Stephen M. Secor, Martha J. Powell, and Guy and Kim Caldwell who ignited the initial flame of my passion for biological discovery while I was a fledgling Howard Hughes Medical Institute Undergraduate Researcher at The University of Alabama; they showed me how to read and write a scientific paper and taught me to conduct scientific research. Several key ichthyological mentors later "hooked" me on freshwater fishes and gave me the basic skills to conduct novel research on their ecology and evolution during my Master's degree at UA. These included Phillip M. Harris, Bernie Kuhajda, Brook Fluker, and Mike Sandel (UA); Patrick E. O'Neil and Maurice "Scott" F. Mettee (Geological Survey of Alabama); and Richard L. Mayden (Saint Louis University), and I offer each of them my sincerest thanks for leading me towards a labor of love. I also thank Richard Richards and Leslie J. Rissler of UA for opening my eyes to the philosophical and societal implications of evolution and conservation biology. Leslie is especially thanked for instilling in me the importance of integrative phylogeographical research, and for encouraging me to pursue my doctoral degree with Jerry at BYU; following her advice has served me very well.

Last, but not least, I thank my family (Flor, Mom, Dad, Bryant, and Aceituna) for all of their love, assistance, positivity, and encouragement during these last six years of my education. It was my parents who introduced me to organic nature, helped me develop my enjoyment of the outdoors from a very early age, and exhorted me to attain the highest level of education possible. However, I especially thank my faithful wife, Florencia, for selflessly and patiently allowing me to pursue my dreams, and whose help during this dissertation, like my love for nature itself, pales in comparison to my love for her. I dedicate this dissertation to my family for their constant support of my struggles to understand wild nature, especially the fascinating diversity of fishes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GENERAL INTRODUCTION

> "[G]eology and biogeography are both parts of natural history and, if they represent the independent and dependent variables respectively in a cause and effect relationship, … they can be reciprocally illuminating" (Rosen 1978, p. 776).

> "[C]omparative phylogeographic assessments within each of multiple codistributed species … offer perhaps the greatest hope for significant advances in understanding how … the demographic and natural histories of populations, can influence intraspecific phylogeographic patterns" (Avise 1998, pp. 376-377).

> "The species richness of Neotropical freshwater fishes … is unparalleled: with more than 5,600 species it represents a majority of the world's freshwater fishes and perhaps 10% of all known vertebrate species. … Any general understanding of vertebrate evolution must therefore address the spectacular evolutionary radiations of Neotropical fishes" (Albert & Reis 2011b, p. xi).

Phylogeography and molecular systematics have proven remarkably useful for understanding the processes influencing the historical diversification of biotic lineages at and below the species level, and also for delimiting morphologically 'cryptic' species (Avise 2000; Bickford *et al.* 2006; Pons *et al.* 2006). Phylogeography is a relatively young and integrative field of science that uses molecular data to infer the processes influencing geographical distributions of genetic lineages within and among species, especially at the intraspecific level (Avise *et al.* 1987; Avise 2000). The phylogeography literature base has grown remarkably fast over the nearly three decades since the birth of the field, yielding many insights into the histories of biotas in different ecosystems across the globe, including cryptic and often temporally 'deep' genealogical divergences within many species ranges (e.g. reviewed by Beheregaray 2008; Knowles 2009). By linking these patterns of population divergence with data on geographical barriers, earth

history events, distribution models, and speciation processes, phylogeography permits identifying and testing whether and over which temporal scales historical and recurrent processes have shaped intraspecific diversification in an area. Indeed, the reciprocal illumination that Rosen (1978; quoted above) envisioned for vicariance biogeography can also be achieved through phylogeography, as phylogeography provides means of inferring how species responded to geological processes without reliance on phylogenetic structure or areas of endemism required by other historical biogeography methods (Zink 2002); for example, historical processes from the unobservable past can still be inferred even when genetic breaks revealed by selectively neutral markers do not correspond to any present-day geographical barrier. Moreover, because population or species demographic histories influence the shapes of their gene genealogies, we can use coalescent theory (a stochastic, backwards-in-time theory of genealogical processes; Kingman 1982) to estimate their underlying genealogies (e.g. by simulating a distribution of them) as well as population parameters describing their histories (e.g. past changes in population sizes) while accounting for various population genetic processes (e.g. subdivision, speciation, mutation, recombination; reviewed by Kuhner 2009). An important paradigm shift in phylogeography has been the realization that phylogeographic inferences are improved by developing and discriminating among competing coalescent-based demographic models using statistical methods that account for the stochasticity of genetic processes (Knowles & Maddison 2002; Knowles 2009; Nielsen & Beaumont 2009). Indeed, coalescent models with one or multiple loci allow us to overcome problems associated with traditional 'pattern-matching' approaches to phylogeography (e.g. reviewed by Avise 2000), for example that gene tree point estimates are random, and that ancestral polymorphisms can substantially influence gene divergence dates or other parameter estimates and lead to erroneous biogeographical

interpretations (Edwards & Beerli 2000; Nielsen & Beaumont 2009).

In comparative phylogeography, genetic datasets are scaled up to test for congruent spatial and temporal population divergences across multiple codistributed species, which permits inferring the responses of whole communities, ecosystems, or species assemblages to geographic barriers and earth historical processes (e.g. historical contingencies such as the uplift of a mountain chain, rerouting of a river, or climate change) in a region (Bermingham & Moritz 1998; Arbogast & Kenagy 2001; Carstens & Richards 2007; Marske *et al.* 2012). Comparative phylogeography surfaced early in the history of phylogeography as improved restriction enzyme and DNA sequencing technologies aided the rapid buildup of mitochondrial DNA (mtDNA) surveys of natural populations facilitating multi-species analyses (Bermingham & Avise 1986; Avise 2000). Here, a key outcome was the realization that, by applying an approach similar to vicariance biogeography at the intraspecific level (discussed in Riddle *et al.* 2008), it became possible to test whether species responded in concerted fashion in space and time to historical processes in an area (e.g. Bermingham & Martin 1998; Sullivan *et al.* 2000; Arbogast & Kenagy 2001). Also, the resulting phylogeographical inferences built from patterns of population structuring replicated across multiple taxa reveal general, rather than lineage-specific, evolutionary patterns, and hold greater promise for discovering novel or unexpected patterns and inferring the geological history of a region than those of single-species studies (Zink 1996; Hickerson *et al.* 2010). Moreover, comparative phylogeography permits more rigorous tests of the predicted contributions of historical versus ecological processes towards generating species genetic variation and spatial-demographic histories (e.g. Bernatchez & Wilson 1998; Burridge *et al.* 2008). However, as with single-species analyses, comparative phylogeography is greatly improved through the use of statistical phylogeographical models of community divergence,

dispersal, and migration across geographical barriers, especially using recent methods based on approximate Bayesian computation and coalescent models (e.g. Hickerson *et al.* 2006, 2007, 2014; Huang *et al.* 2011). Although phylogeography can reinforce conclusions of traditional studies, among the greatest successes of phylogeography at the intraspecific and comparative levels has been the elucidation of cryptic speciation and deep community divergences challenging traditional denominations of species limits and biogeographical regions inferred from morphology-based taxonomy (e.g. Avise 2000; Riddle *et al.* 2000; Riddle & Hafner 2006).

An improved understanding of species limits and evolutionary relationships is also a key outcome of molecular systematics, which uses DNA-based reconstructions of phylogenies of species and higher taxa to improve our understanding of the histories of biotic lineages, communities, and whole biotas (Hillis *et al.* 1996; Felsenstein 2004). Especially when combined with data from ecology, phenotypic diversity, or natural history, phylogenies provide means of rigorously inferring the historical biogeographical processes (e.g. dispersal patterns, speciation and extinction rates) and speciation processes (e.g. geographical mode and tempo of speciation) that have given rise to modern-day diversity within clades or across communities (e.g. Webb *et al.* 2002; Losos & Glor 2003; Rabosky & Lovette 2008; Graham *et al.* 2009). However, phylogenetic methods rely on a variety of assumptions (e.g. DNA substitution models, character utility, appropriate model complexity), among the most important of which is that nominal taxa represent evolutionary species (Barraclough & Nee 2001). This is problematic because predefined taxonomy often provides an inexact fit to molecular data, and morphological taxonomy in particular is apt to underestimate species diversity due to the presence of cryptic species, among other issues (e.g. Pons *et al.* 2006; Satler *et al.* 2013). Moreover, it has until now been difficult to arrive at objective species delimitations, and morphological taxonomists have

often relied on subjective determinations of the distinctiveness of evolutionary lineages. That said, molecular data have been of increasing interest for species delimitation over the last decade, and workers have particularly focused on using single-locus analyses (e.g. of "DNA barcodes") as the basis for taxonomy (e.g. Hebert *et al.* 2004). Yet, this approach can also be confounded because single loci reflect realizations of random genealogical processes and variance in reproductive success, and thus may not accurately capture species limits (e.g. Nielsen & Beaumont 2009; Fujita *et al.* 2012). Fortunately, the proliferation of molecular data and the recent surge in importance to phylogenetics and species delimitation of coalescent models from statistical population genetics has ushered in new "coalescent-based species delimitation" methods that allow objectively delimiting species using multilocus genetic datasets and methods that overcome these shortcomings (reviewed by Fujita *et al.* 2012). This approach falls within integrative taxonomy, a synthetic field of study combining elements of molecular and morphological systematics (phylogenetics and taxonomy) with species delimitation, with the goal being to identify species limits and processes of lineage diversification based on multiple lines of evidence (Dayrat 2005; Padial *et al.* 2010; Fujita *et al.* 2012). Specifically, coalescent-based species delimitation provides a means of estimating support for speciation events on phylogenies using statistically rigorous methods, and with little if any investigator bias; these inferences can then be integrated with evidence from morphology, ecology, behavior and other fields to further test the distinctiveness of genetic lineages in an integrative taxonomy framework (Fujita *et al.* 2012). Overall, coalescent-based species delimitation methods allow testing hypotheses of species distinctiveness while inferring evolutionary processes that have led to the observed patterns by using coalescent models. Given species are the fundamental unit of biology, accurate species delimitation is thus of vital importance for accurately gauging species

diversity and devising effective conservation strategies (Sites & Marshall 2003; Mace 2004).

Perhaps nowhere on earth are the insights of phylogeography and species delimitation more significant or urgently needed than in the Neotropics. Ecologists and evolutionary biologists have been struggling to determine the mechanisms that explain the historical assembly and rich biological diversity of tropical North and South America ever since 19th Century naturalists such as Alfred Russel Wallace and his contemporaries Henry Walter Bates, Thomas Belt, Charles Darwin, Albert Günther, and Philip Sclater first seriously contemplated the biogeography of New World landscapes (e.g. Darwin 1859; Belt 1874; Wallace 1876). The traditions of biodiversity research in each of the two major areas of what Sclater (1858; using bird distributions) and Wallace (1876; studying distributions of multiple animal taxa) identified as constituting the "Neotropical" biogeographical realm (termed the "Neotropical ecozone" or "Neotropics" today), including southern Mexico and Central America as well as South America, are both longstanding and intermingled. Research on the Neotropics has revealed and emphasized among the "oldest" patterns in global ecology, for example the latitudinal gradient in species richness (declining species diversity towards the poles; Hawkins 2001; Wiens & Donoghue 2003). However, despite more than 100 years of study, we continue to remain fascinated with the diverse landscapes and biotas of these regions today (e.g. Bermingham *et al.* 2005; Hoorn & Wesselingh 2010; Albert & Reis 2011a). Indeed, the Neotropics remain as relevant as ever as the "preferred target" for biodiversity research (Rull 2008), and this is rightly so given they are among the most species-rich areas of the world, with multiple biodiversity "hotspots" of endemic species that are also threatened with imminent habitat loss due to human activities (Myers *et al.* 2000). The outstanding diversity of Neotropical biotas is continually threatened by habitat destruction (e.g. land conversion for agriculture), pollution, and human

population expansion, as well as the expansion of fungal diseases (reviewed by Leonard 1987; Robinson & Redford 1991; Stotz *et al.* 1996; Berger *et al.* 1998; Laurance *et al.* 2001; Olson *et al.* 2001; Klink & Machado 2005; Bagley & Johnson 2014a). As a consequence, the Neotropics comprise a region that is of particularly great interest for studying the historical and ecological processes promoting population divergence and speciation (e.g. Bermingham & Martin 1998; Rull 2008; Antonelli *et al.* 2009), and for objectively determining species limits to improve conservation prioritization (e.g. Fouquet *et al.* 2007, 2014).

Within the Neotropics, the most diverse group of vertebrates, and indeed a substantial proportion of global vertebrate diversity, is the Neotropical freshwater fish assemblage (Albert & Reis 2011b). Neotropical North and South America harbor the greatest diversity of freshwater fishes worldwide, with an estimated total of ~7000 described and undescribed species (around half of global freshwater fish species richness), including >525 species within the relatively modest areal extent of Central America, >1000 species in the Orinoco River basin, and 2173 species in the Amazon River superbasin (Reis *et al.* 2003; Albert & Reis 2011b; Matamoros *et al.* 2014). Comparing the 7000 species estimate of Albert & Reis (2011b) above to recent IUCN data summaries on vertebrate species diversity quoting a total of 62,305 species of vertebrates worldwide, this value if accurate would represent ~11.2% of vertebrate species richness (The World Conservation Union 2010). As a consequence, as noted by Albert & Reis (2011b, quoted above), it is axiomatic that any student of vertebrate diversity must confront the fascinating diversity of Neotropical freshwater fishes.

Central America forms a long (>1500 km), narrow isthmus extending from the Maya Highlands of southeastern Mexico and Motagua Fault Zone of Guatemala, southeast to Panama's Darien isthmus connection with Colombia. Along its length, the Central American Isthmus

encompasses about 0.4% of the earth's area (533,726 km^2) and connects the North and South American continents while isolating the Pacific Ocean from the Caribbean Sea; thus Central America makes up the only transoceanic, transcontinental isthmus worldwide. Central America is also a region of high ecological and geological complexity that has witnessed a variety of upheavals in landscape features and climate during its Miocene to recent geological evolution, especially due to its position at the intersection of five tectonic plates—the Cocos, North American, Caribbean, South American, and Nazca plates (Coates & Obando 1996; Coates *et al.* 2004). The landforms of the Central American Isthmus have been generated at a fascinating subduction factory where the Cocos and Nazca plates dip beneath the western Caribbean Plate at the Middle American Trench (along the Pacific coast of the isthmus), causing tectonic uplift that has formed a series of orogenic belts (mountain-building zones). When taken together, the five major volcanic segments uplifted at these belts comprise among the most tectonically active areas of the eastern Pacific "Ring of Fire". These segments are collectively referred to as the Central American Volcanic arc (CAVA; Mann *et al.* 2007) and have generated the northwest-trending volcanic cordilleras that largely define physical terrain and environments in the region. In turn, the volcanic cordilleras of the region play a major role in shaping regional atmospheric and oceanic circulation patterns, as well as the configurations and connectivity of freshwater drainage basins and terrestrial habitats on either side (e.g. Savage 2002; Hulsey & López-Fernández 2011). By dividing the region into two distinct coasts bisected by cordillera ranges and intermittent valleys and basins, the central cordilleras create a stepping-stone-like organization of terrestrial and aquatic populations but especially of the populations of obligate freshwater organisms, which are confined to discrete hydrological networks that interact with the landscape (e.g. Martin & Bermingham 2000; Hulsey & López-Fernández 2011). Thus, it should

be possible to use phylogeography to recover the history of interactions between drainage basins and the evolving Central American landscape, including historical connections among drainages, and patterns of dispersal and vicariance within and between coasts (e.g. Bermingham & Martin 1998).

In this rich landscape of mostly short coastal plains cut off by dramatically uplifting volcanic cordilleras and plateaus (up to >5700 m above sea level), relatively wider and more ancient continental areas of Nuclear Central America (NCA; Guatemala, Belize, El Salvador, Honduras and Nicaragua) are juxtaposed against more recent lands of lower Central America (LCA; Costa Rica and Panama) that are oceanic in origin and taper to a very narrow isthmus (Mann *et al.* 2007; Marshall 2007). However, despite being made up of geological blocks and terranes of different histories and geographical origins, these regions have been geologically linked since at least the Paleogene (~49 million years ago, Ma) to Miocene (23 Ma) and have experienced tectonic uplift, volcanism, and basin and erosional processes since the early-mid Miocene (~19–0 Ma) that caused the evolution of major landscape features and relief witnessed in Central America today (Coates & Obando 1996; Coates *et al.* 2004; Marshall *et al.* 2003; Mann *et al.* 2007; Rogers *et al.* 2007). Indeed, the prominent northwest-trending cordilleras of the region experienced their main phases of uplift during this timeframe, and have since effectively isolated terrestrial and aquatic species on either side of the continental divide (e.g. Abratis & Warner 2001; Rogers *et al.* 2002). In the late Pliocene–Pleistocene, the sequence of Central American Isthmus emergence ultimately culminated in the uplift of the Isthmus of Panama ~3.1–1.8 Ma (e.g. Coates & Obando 1996). This major earth history event facilitated a series of bidirectional dispersals of plants and animals between the nascent Central American isthmus and North and South America that is known as the 'Great American Biotic Interchange'

(Simpson 1950; Marshall *et al.* 1979; Stehli & Webb 1985), which led to important distributional shifts and *in situ* diversification of biotic lineages, all of which make Central America a region of central importance in biogeography (Stehli & Webb 1985; Webb 1995; Bermingham & Martin 1998; Savage 2002; Bagley & Johnson 2014a). Given many lineages of freshwater fishes are thought to have dispersed between Central America and outlying areas of northwestern South America before, during, and after the GABI, the fish assemblages of these two regions are linked and considered to form a single ichthyofauna (e.g. Bussing 1976, 1985; Albert & Reis 2011b; Chakrabarty & Albert 2011).

In this dissertation, I used an integrative approach coupling comparative phylogeography and coalescent-based species delimitation to improve our understanding of the biogeography and species limits of the Central American freshwater fish assemblage. I have chosen Central America as the locus of this project because the relatively recent and dynamic geological and ecological history of the Central American Isthmus has created a fascinating natural laboratory for biogeography that presents excellent opportunities for exploring the interplay between earth historical processes and the diversity and distributions of species (Bermingham & Martin 1998; Martin & Bermingham 2000; Bagley & Johnson 2014a). As discussed briefly in Chapter 4, there are also reasons that Central America presents a more approachable and suitable study area for evaluating the effects of historical processes on the genetic and distributional divergences of Neotropical freshwater fishes, with several benefits over larger Neotropical areas. Moreover, the presence of many freshwater fish species with overlapping ranges in Central America (e.g. Bussing 1998; Matamoros *et al.* 2014) lends itself to comparative studies providing improved, multi-taxon inferences of evolutionary history. The study taxa span three families of freshwater fishes from different orders of teleosts, including family Poeciliidae (Order Cyprinodontiformes),

family Characidae (Order Characiformes), and family Cichlidae (Order Perciformes). Based on family-level biogeographic evidence, all the focal lineages are classified as 'primary' (lacking salinity tolerance; Characidae) or 'secondary' freshwater fishes (possibly with physiological adaptations to brackish or salt water; Poeciliidae, Cichlidae; Myers 1938) that are obligate inhabitants of freshwater habitats. I specifically have excluded 'peripheral' fishes (Myers 1938) that despite inhabiting fresh waters are primarily marine forms that can disperse between rivers through saltwater and are unlikely affected by rising sea levels, making it difficult to distinguish among hypotheses using their biogeographical patterns. Prior research on the phylogeography of Central American freshwater fishes suggests that similar forces have acted to shape primary and secondary fish evolution and community composition at broad spatial scales in the region, regardless of such designations reflecting potential ecological differences in dispersal potential (Bermingham & Martin 1998; McCafferty *et al.* 2012). However, our recent analyses (from Chapter 2) of poeciliids support finer-scale patterns of phylogeographic structuring suggesting at least some role for ecological differences in shaping different responses among lineages to regional historical events among species from the same secondary freshwater fish family (Bagley & Johnson 2014b).

In Chapter 1, and in preparation for subsequent chapters, I conducted a literature review of the contributions of phylogeography studies to understanding the origins and maintenance of LCA biodiversity, in light of the regional geological and ecological setting (Bagley & Johnson 2014a). I highlighted emerging phylogeographic patterns in LCA, along with the need for improving regional historical biogeographical inference and conservation efforts through statistical and comparative phylogeographic studies. In Chapter 2, I compared mitochondrial phylogeographic patterns among three species of livebearing fishes (Poeciliidae) codistributed in

the lower Nicaraguan depression and proximal uplands (Bagley & Johnson 2014b). Our results revealed evidence for mixed spatial and temporal divergences among species, indicating phylogeographic "pseudocongruence" (spatially congruent yet temporally incongruent genetic breaks; Cunningham & Collins 1994; Donoghue & Moore 2003) suggesting that multiple evolutionary responses to historical processes have shaped population structuring of regional freshwater biota, possibly linked to recent community assembly and/or the effects of ecological differences among species on their responses to late Cenozoic environmental events. In Chapter 3, I used coalescent-based species tree and species delimitation analyses of a multilocus dataset to delimit species and infer their evolutionary relationships in the *Poecilia sphenops* species complex (Poeciliidae), a widespread but morphologically conserved group of fishes (e.g. Miller *et al.* 2005). Results indicated that diversity is underestimated and overestimated in different clades by *c.* ±15% (including two candidate species supported from collections from Nicaragua and Panama); that lineages diversified since the Miocene; and that some evidence exists for a more probable role of hybridization, rather than incomplete lineage sorting, in shaping observed gene tree discordances. The multilocus dataset generated in this chapter allowed us to better delimit species and develop coalescent-based models accounting for stochastic mutational and coalescent processes among populations and loci. Last, in Chapter 4, I used a comparative phylogeographical analysis of eight codistributed species/genera of freshwater fishes to test for shared evolutionary responses predicted by four drainage-based hypotheses that we outlined for Neotropical fish diversification in Central America—the 'tectonic vicariance hypothesis', 'marine vicariance hypothesis', 'continental shelf width hypothesis' (*sensu* Unmack *et al.* 2013, but tested for the first time using Central American fishes), and 'cross-cordillera exchange hypothesis'. By integrating mtDNA phylogeographic analyses with paleodistribution modeling

based on ecological niche models using paleoclimatic data layers (e.g. Waltari *et al.* 2007), we showed that incongruent genetic structuring has arisen among these lineages despite overlapping ancestral Pleistocene distributions, and this suggested that the biogeographical model for the regional fish assemblage (especially in the Nicaraguan depression and LCA) has most likely involved multiple routes to community assembly. Hypotheses tests using the latest hierarchical approximate Bayesian computation (ABC) model averaging methods in the msBayes bioinformatics pipeline (Hickerson *et al.* 2006, 2007, 2014; Huang *et al.* 2011) also supported a single pulse of diversification in two lineages diverged in the upper San Carlos River drainage, but multiple divergences of three lineages across the Sixaola River basin, Costa Rica. Moreover, the temporal pattern of diversification at these shared spatial breaks correlated best with Neogene sea level events, and the spatial continental shelf width. Seven focal lineages also displayed spatially congruent evidence for past drainage connections across the continental divide at the Guanacaste Cordillera. Overall, my Chapter 4 results supported complex biogeographical patterns of dispersal and vicariance illustrating how concerted responses and multiple responses across taxa to historical drainage-controlling processes have influenced Neotropical fish diversification.

REFERENCES

Abratis M, Wörner G (2001) Ridge collision, slab-window formation, and the flux of Pacific asthenosphere into the Caribbean realm. *Geology*, **29**, 127-130.

Albert JS, Reis RE (2011a) Historical Biogeography of Neotropical Freshwater Fishes. University of California Press, Berkeley, California.

Albert JS, Reis RE (2011b) Introduction to Neotropical freshwaters. In: Historical Biogeography of Neotropical Freshwater Fishes (eds Albert JS, Reis RE), pp. 3-20. University of California Press, Berkeley, California.

Antonelli A, Nylander JAA, Persson C, Sanmartín I (2009) Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 9749-9754.

Arbogast BS, Kenagy GJ (2001) Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, **28**, 819-825.

Avise JC (1998) The history and purview of phylogeography: a personal reflection. *Molecular Ecology*, **7**, 371-379.

Avise JC (2000) Phylogeography: the History and Formation of Species. Harvard University Press, Cambridge, Massachusetts.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.

Bagley JC, Johnson (2014a) Phylogeography and biogeography of the lower Central American Neotropics: diversification between two continents and between two seas. *Biological Reviews*, **89**, 767-790.

Bagley JC, Johnson JB (2014b) Testing for shared biogeographic history in the lower Central American freshwater fish assemblage using comparative phylogeography: concerted, independent, or multiple evolutionary responses? *Ecology and Evolution*, **4**, 1686-1705.

Barraclough T, Nee S (2001) Phylogenetics and speciation. *Trends in Ecology & Evolution*, **16**, 391-399.

Beheregaray LB (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, **17**, 3754-3774.

Belt T (1874) The Naturalist in Nicaragua. E. Bumpas, London.

Berger L, Speare R, Daszak P, Green DE, Cunningham AA, Goggin CL, Slocombe R, Ragan MA, Hyatt AD, McDonald KR, Hines HB, Lips KR, Marantelli G, Parkes H (1998) Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 9031-9036.

Bermingham E, Avise JC (1986) Molecular zoogeography of fresh-water fishes in the southeastern United States. *Genetics*, **113**, 939-965.

Bermingham E, Dick C, Moritz C (2005) Tropical Rainforests: Past, Present and Future. University of Chicago Press, Chicago, Illinois.

Bermingham E, Martin AP (1998) Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology*, **7**, 499-517.

Bermingham E, Moritz C (1998) Comparative phylogeography: concepts and applications. *Molecular Ecology*, **7**, 367-369.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, **7**, 431-452.
Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2006) Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, **22**, 148-155.
Burridge CP, Craw D, Jack DC, King TM, Waters JM (2008) Does fish ecology predict dispersal across a river drainage divide? *Evolution*, **62**(6), 1484-1499.
Bussing WA (1976) Geographic distribution of the San Juan ichthyofauna of Central America with remarks on its origin and ecology. In: Investigations of the Ichthyofauna of Nicaraguan Lakes (ed. Thorson TB), pp. 157-175. University of Nebraska, Lincoln, Nebraska.
Bussing WA (1985) Patterns of the distribution of the Central American ichthyofauna. In: The Great American Biotic Interchange (eds Stehli FG, Webb SD), pp. 453-473. Plenum Press, New York, New York.
Bussing WA (1998) Freshwater Fishes of Costa Rica, 2nd Edition. Editorial de la Universidad de Costa Rica, San José, Costa Rica.
Carstens BC, Richards CL (2007) Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution*, **61**(6), 1439-1454.
Chakrabarty P, Albert JS (2011) Not so fast: a new take on the Great American Biotic Interchange. In: Historical Biogeography of Neotropical Freshwater Fishes (eds Albert JS, Reis RE), pp. 293-305. University of California Press, Berkeley, California.
Coates AG, Collins LS, Aubry MP, Berggren WA (2004) The geology of the Darien, Panama, and the late Miocene-Pliocene collision of the Panama arc with northwestern South America. *Geological Society of America Bulletin*, **116**, 1327-1344.
Coates AG, Obando JA (1996) The geologic evolution of the Central American isthmus. In: Evolution and Environment in Tropical America (eds Jackson JBC, Budd AF, Coates AG), pp. 21-56. University of Chicago Press, Chicago, Illinois.
Cunningham CW, Collins T (1994) Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. In: *Molecular Ecology and Evolution: Approaches and Applications* (eds Schierwater B, Streit B, Wagner GP, DeSalle R), pp. 405-433. Birkhauser Verlag, Basel, Switzerland.
Darwin CR (1859) On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. John Murray, London.
Dayrat B (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society*, **85**, 407-415.
Donoghue MJ, Moore BR (2003) Toward an integrative historical biogeography. *Integrative and Comparative Biology*, **43**, 261-270.
Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839-1854.
Felsenstein J (2004) Inferring Phylogenies, 2nd Edition. Sinauer Associates, Sunderland, MA.
Fouquet A, Cassini CS, Haddad CFB, Pech N, Rodrigues MT (2014) Species delimitation, patterns of diversification and historical biogeography of the Neotropical frog genus *Adenomera* (Anura, Leptodactylidae). *Journal of Biogeography*, **41**, 855-870.
Fouquet A, Vences M, Salducci M-D, Meyer A, Marty C, Blanc M, Gilles A (2007) Revealing cryptic diversity using molecular phylogenetics and phylogeography in frogs of the *Scinax ruber* and *Rhinella margaritifera* species groups. *Molecular Phylogenetics and*

Evolution, **43**, 567-582.
Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, **27**, 480-488.
Graham CH, Parra JL, Rahbek C, McGuire JA (2009) Phylogenetic structure in tropical hummingbird communities. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19673-19678.
Hawkins BA (2001) Ecology's oldest pattern? *Trends in Ecology & Evolution*, **16**, 470.
Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 14812-14817.
Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, **54**, 291-301.
Hickerson MJ, Stahl EA, Lessios HA (2006) Test for simultaneous divergence using approximate Bayesian computation. *Evolution*, **60**(12), 2435-2453.
Hickerson MJ, Stahl E, Takebayashi N (2007) msBayes: pipeline for testing comparative phylogeographic histories using historical approximate Bayesian computation. *BMC Bioinformatics*, **8**, 268.
Hickerson MJ, Stone GN, Lohse K, Demos TC, Xie X, Landerer C, Takebayashi N (2014) Recommendations for using msBayes to incorporate uncertainty in selecting an ABC model prior: a response to Oaks et al. *Evolution*, **68**, 284-294.
Hillis DM, Moritz C, Mable BK (1996) Molecular Systematics, 2nd Edition. Sinauer Associates, Sunderland, Massachusetts.
Hoorn C, Wesselingh F (2010) Amazonia: Landscape and Species Evolution: a Look into the Past. Wiley-Blackwell, West Sussex, United Kingdom.
Huang W, Takebayashi N, Qi Y, Hickerson MJ (2011) MTML-msBayes: Approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. *BMC Bioinformatics*, **12**, 1.
Hulsey CD, López-Fernández H (2011) Nuclear Central America. In: Historical Biogeography of Neotropical Freshwater Fishes (eds Albert JS, Reis RE), pp. 279-291. University of California Press, Berkeley, California.
Kingman J (1982) The coalescent. *Stochastic Processes and their Applications*, **13**(3), 235-248.
Klink CA, Machado RB (2005) Conservation of the Brazilian Cerrado. *Conservation Biology*, **19**, 707-713.
Knowles LL (2009) Statistical phylogeography. *Annual Review in Ecology, Evolution, and Systematics*, **40**, 593-612.
Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623-2635.
Kuhner MK (2009) Coalescent genealogy samplers: windows into population history. *Trends in Ecology & Evolution*, **24**, 86-93.
Laurance WF, Cochrane MA, Bergen S, Fearnside PM, Dela-Monica P, Barber C, D'Angelo S, Fernandes T (2001) The future of the Brazilian Amazon. *Science*, **291**, 438-439.
Leonard HJ (1987) Natural Resources and Economic Development in Central America: A Regional Environmental Profile. Transaction Books, New Brunswick, New Jersey.
Losos JB, Glor RE (2003) Phylogenetic comparative methods and the geography of speciation.

Trends in Ecology & Evolution, **18**, 220-227.
Mace GM (2004) The role of taxonomy in species conservation. *Philosophical Transactions of the Royal Society B*, **359**, 711-719.
McCafferty SS, Martin A, Bermingham E (2012) Phylogeographic diversity of the lower Central American cichlid *Andinoacara coeruleopunctatus* (Cichlidae). *International Journal of Evolutionary Biology*, **2012**.
Mann P, Rogers RD, Gahagan L (2007) Overview of plate tectonic history and its unresolved tectonic problems. In Central America: Geology, Resources and Hazards (eds Bundschuh J, Alvarado GE), pp. 205-241. Taylor & Francis, Philadelphia, Pennsylvania.
Marshall JS (2007) The geomorphology and physiographic provinces of Central America. In: Central America: Geology, Resources and Hazards (eds Bundschuh J, Alvarado GE), pp. 1-51. Taylor & Francis, Philadelphia, Pennsylvania.
Marshall JS, Idleman BD, Gardner TW, Fisher DM (2003) Landscape evolution within a retreating volcanic arc, Costa Rica, Central America. *Geology*, **31**, 419-422.
Marshall LG, Butler RF, Drake RE, Curtis GH, Tedford RH (1979) Calibration of the Great American Interchange. *Science*, **204**, 272-279.
Marske KA, Leschen RAB, Buckley TR (2012) Concerted versus independent evolution and the search for multiple refugia: comparative phylogeography of four forest beetles. *Evolution*, **66**(6), 1862-1877.
Martin AP, Bermingham E (2000) Regional endemism and cryptic species revealed by molecular and morphological analysis of a widespread species of Neotropical catfish. *Proceedings of the Royal Society B*, **267**, 1135-1141.
Matamoros WA, McMahan CD, Chakrabarty P, Albert JS, Schaefer JF (2014) Derivation of the freshwater fish fauna of Central America revisited: Myers' hypothesis in the twenty-first century. *Cladistics*, **2014**, 1-12.
Miller RR, Minckley WL, Norris SM (2005) *Freshwater Fishes of México*. The University of Chicago Press, Chicago, Illinois.
Myers GS (1938) Fresh-water fishes and West Indian zoogeography. *Annual Report Smithsonian Institution*, **1937**, 339-364.
Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853-858.
Nielsen R, Beaumont MA (2009) Statistical inferences in phylogeography. *Molecular Ecology*, **18**, 1034-1047.
Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR (2001) Terrestrial ecoregions of the world: a new map of life on earth. *BioScience*, **51**(11), 933-938.
Padial JM, Miralles A, De la Riva I, Vences M (2010) The integrative future of taxonomy. *Frontiers in Zoology*, **7**, 16.
Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, **55**, 595-609.
Rabosky DL, Lovette IJ (2008) Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution*, **62**, 1866-1875.
Reis RE, Kullander SO, Ferraris CJ (2003) Checklist of the Freshwater Fishes of South and Central America. EDIPUCRS, Porto Alegre, Rio Grande do Sul, Brazil.

Riddle BR, Dawson MN, Hadly EA, Hafner DJ, Hickerson MJ, Mantooth SJ, Yoder AD (2008) The role of molecular genetics in sculpting the future of integrative biogeography. *Progress in Physical Geography*, **32**, 173-202.
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.
Riddle BR, Hafner DJ (2006) A step-wise approach to integrating phylogeographic and phylogenetic biogeographic perspectives on the history of a core North American warm deserts biota. *Journal of Arid Environments*, **66**, 435-461.
Robinson JG, Redford KH (1991) Neotropical Wildlife Use and Conservation. University of Chicago Press, Chicago, Illinois.
Rogers RD, Karason H, van der Hilst R (2002) Epirogenic uplift above a detached slab in northern Central America. *Geology*, **30**, 1031-1034.
Rogers RD, Mann P, Emmet PA (2007) Tectonic terranes of the Chortis block based on integration of regional aeromagnetic and geologic data. In: Geologic and Tectonic Development of the Caribbean Plate in Northern Central America (ed. Mann P), pp. 65-88. Geological Society of America Special Paper 428.
Rosen DE (1978) Vicariant patterns and historical explanation in biogeography. *Systematic Zoology*, **27**, 159-188.
Rull V (2008) Speciation timing and neotropical biodiversity: the Tertiary–Quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology*, **17**, 2722-2729.
Satler JD, Carstens BC, Hedin M (2013) Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*). *Systematic Biology*, **62**(6), 805-823.
Savage JM (2002) The Amphibians and Reptiles of Costa Rica: a Herpetofauna Between Two Continents, Between Two Seas. University of Chicago Press, Chicago, Illinois.
Sclater PL (1858) On the general geographical distribution of the members of the Class Aves. *Journal of the Proceedings of the Linnean Society, Zoology*, **2**, 130-145.
Simpson GG (1950) History of the fauna of Latin America. *American Scientist*, **38**, 361-389.
Sites JW Jr., Marshall JC (2003) Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology & Evolution*, **18**, 462-470.
Stehli FG, Webb SD (1985) The Great American Biotic Interchange. Plenum Press, New York, New York.
Stotz DF, Fitzpatrick JW, Parker TAI, Moskovits DK (1996) Neotropical Birds: Ecology and Conservation. University of Chicago Press, Chicago, Illinois.
Sullivan J, Arellano E, Rogers DS (2000) Comparative phylogeography of Mesoamerican highland rodents: concerted versus independent response to past climatic fluctuations. *American Naturalist*, **155**, 755-768.
The World Conservation Union (2010) IUCN Red List of Threatened Species. Summary Statistics for Globally Threatened Species. Table 1: Numbers of threatened species by major groups of organisms (1996-2010).
Unmack PJ, Hammer MP, Adams M, Johnson JB, Dowling TE (2013) The role of continental shelf width in determining freshwater phylogeographic patterns in south-eastern Australian pygmy perches (Teleostei: Percichthyidae). *Molecular Ecology*, **22**(6), 1683-1699.
Wallace AR (1876) The Geographical Distribution of Animals with a Study of the Relations of

Living and Extinct Faunas as Elucidating the Past Changes of the Earth's Surface. Macmillan, London.
Waltari E, Hijmans RJ, Peterson AT, Nyári AS, Perkins SL, Guralnick RP (2007) Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS One*, **2**(7), e563.
Webb CO, Ackerly DD, McPeek MA, Donoghue MJ (2002) Phylogenetics and community ecology. *Annual Reviews in Ecology and Systematics*, **33**, 475-505.
Webb SD (1995) Biological implications of the middle Miocene Amazon seaway. *Science*, **269**, 361-362.
Wiens JJ, Donoghue MJ (2004) Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, **19**, 639-644.
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.

Chapter 1: Phylogeography and biogeography of the lower Central American Neotropics: diversification between two continents and between two seas

BIOLOGICAL REVIEWS
Cambridge Philosophical Society
Biol. Rev. (2014), **89**, pp. 767–790.
doi: 10.1111/brv.12076

Phylogeography and biogeography of the lower Central American Neotropics: diversification between two continents and between two seas

Justin C. Bagley[1,*] and Jerald B. Johnson[1,2]

[1] *Evolutionary Ecology Laboratories, Department of Biology, Brigham Young University, 401 WIDB (Widtsoe Building), Provo, UT 84602, U.S.A.*
[2] *Monte L. Bean Life Science Museum, Brigham Young University, 645 E 1430 N, Provo, UT 84602, U.S.A.*

ABSTRACT

Lower Central America (LCA) provides a geologically complex and dynamic, richly biodiverse model for studying the recent assembly and diversification of a Neotropical biota. Here, we review the growing literature of LCA phylogeography studies and their contribution to understanding the origins, assembly, and diversification of the LCA biota against the backdrop of regional geologic and climatic history, and previous biogeographical inquiry. Studies to date reveal that phylogeographical signal within taxa of differing distributions reflects a diversity of patterns and processes rivalling the complexities of LCA landscapes themselves. Even so, phylogeography is providing novel insights into regional diversification (e.g. cryptic lineage divergences), and general evolutionary patterns are emerging. Congruent multi-taxon phylogeographic breaks are found across the Nicaraguan depression, Chorotega volcanic front, western and central Panama, and the Darién isthmus, indicating that a potentially shared history of responses to regional-scale (e.g. geological) processes has shaped the genetic diversity of LCA communities. By contrast, other species show unique demographic histories in response to overriding historical events, including no phylogeographic structure at all. These low-structure or incongruent patterns provide some evidence for a role of local, ecological factors (e.g. long-distance dispersal and gene flow in plants and bats) in shaping LCA communities. Temporally, comparative phylogeographical structuring reflects Pliocene–Pleistocene dispersal and vicariance events consistent with the timeline of emergence of the LCA isthmus and its major physiographic features, e.g. cordilleras. We emphasise the need to improve biogeographic inferences in LCA through in-depth comparative phylogeography projects capitalising on the latest statistical phylogeographical methods. While meeting the challenges of reconstructing the biogeographical history of this complex region, phylogeographers should also take up the critical service to society of applying their work to the conservation of its fascinating biodiversity.

Key words: Central American Isthmus, conservation, Costa Rica, environmental change, geology, historical biogeography, Panama, palaeogeography, phylogeography.

CONTENTS

* Author for correspondence (Tel: +1 801 422 2203; Fax: +1 801 422 0090; E-mail: justin.bagley@byu.edu).

I. INTRODUCTION

Large-scale geographical patterns of biodiversity (e.g. hotspots, coldspots, and latitudinal gradients in species richness) are increasingly well documented (Gaston, 2000; Myers *et al.*, 2000). However, understanding the mechanisms underlying global patterns of species richness and community composition remains one of the great challenges of ecology and biogeography (Gaston, 2000; Wiens & Donoghue, 2004; Lomolino *et al.*, 2010). Correlations between ecological factors (e.g. Kreft & Jetz, 2007), and ecological-drift models (Hubbell, 2001), have been shown to predict species richness and abundance accurately. Admirably, the latter approach even links local, deterministic processes (e.g. ecological interactions) and regional, historical processes (*sensu* Ricklefs, 1987). Such models are inadequate, however, to infer the historical origins and assembly of species-rich biotas, or the relative contributions of local- *versus* regional-scale processes towards shaping their diversification (e.g. Pennington, Cronk & Richardson, 2004; Ricklefs, 2006; Simon *et al.*, 2009). Here, historical biogeography (Arbogast & Kenagy, 2001; Posadas, Crisci & Katinas, 2006) is essential because biogeographical processes of dispersal, speciation, and extinction alter regional species pools *and* local community diversity through time (Ricklefs, 1987, 2006; Ricklefs & Schluter, 1993; Schneider, Cunningham & Moritz, 1998; Moritz *et al.*, 2000; Smith & Bermingham, 2005). Unfortunately, elucidating mechanisms underpinning the assembly and diversification of continental biotas has remained elusive because past attempts were limited to distributional data, which are often problematic due to inadequate taxonomic resolution, lack of fossil data, or historical range dynamics (e.g. Losos & Glor, 2003). Also, continental-scale insights into historical community fluctuations from molecular phylogeography have only recently become available for many areas (Bermingham & Avise, 1986; Bermingham & Martin, 1998; Avise, 2000; Beheregaray, 2008). Thus our understanding of the histories by which most biotas assembled and diversified remains limited.

Phylogeography is among the most integrative and fastest growing fields in biology today and is critical to understanding evolutionary diversification (e.g. Riddle *et al.*, 2008; Knowles, 2009; Hickerson *et al.*, 2010). Through illuminating geographical histories of genetic lineages within and among species, phylogeography provides tremendous insight into processes of lineage divergence (speciation) and spread and, therefore, historical biogeographical scenarios (Avise *et al.*, 1987; Avise, 2000; Kidd & Ritchie, 2006). Phylogeography offers an array of methods that, constantly debated and refined (e.g. Bloomquist, Lemey & Suchard, 2010), present exciting alternatives to traditional distribution-based biogeographical analyses. Phylogeography has proven very successful in historical biogeography due to its capacity for uncovering cryptic biodiversity, thus challenging traditional taxonomy (e.g. Avise, 2000; Riddle & Hafner, 2006); deciphering past movements and population dynamics of organisms (e.g. Hewitt, 2000); and integrating statistical frameworks and previously disjunct fields (e.g. Knowles & Maddison, 2002; Hickerson, Dolman & Moritz, 2006*a*; Hickerson, Stahl & Takebayashi, 2007; Kozak, Graham & Wiens, 2008).

Yet phylogeographic knowledge is markedly uneven with respect to geography. Of interest to the present review, while the Neotropical zone boasts seven of the world's 25 biodiversity hotspots (Mesoamerica, Caribbean, Chocó/Darién/western Ecuador, Tropical Andes, Brazil's Cerrado and Atlantic Forest, and Central Chile; Myers *et al.*, 2000), a recent worldwide survey of 2434 phylogeography publications found that Neotropical studies formed only ∼3% (Central America) to 6.3% (South America) of studies (Beheregaray, 2008). Remarkably, this means that the top two areas of vertebrate species richness, endemism and threat—the Tropical Andes and Mesoamerica (Myers *et al.*, 2000)—are largely underrepresented. Such general lack of phylogeographical information on Neotropical biotas limits our ability to gauge biodiversity levels and infer processes of diversification including the relative contributions of local *versus* regional processes (Wiens & Donoghue, 2004; Ricklefs,

2006; Simon *et al.*, 2009). More phylogeographic studies are clearly needed to understand Neotropical diversification.

Despite phylogeography's crucial role in understanding mechanisms of diversification, it is difficult to determine whether intraspecific phylogeographies represent patterns broadly imprinted across regional biodiversity (Avise, 2000; Castoe *et al.*, 2009). By comparing phylogeographical patterns across multiple lineages codistributed in a region, 'comparative phylogeography' provides a means of testing whether such general evolutionary patterns exist (Bermingham & Avise, 1986; Bermingham & Martin, 1998; Bermingham & Moritz, 1998; Sullivan, Arellano & Rogers, 2000; Arbogast & Kenagy, 2001; Hickerson, Stahl & Lessios, 2006*b*; Hickerson *et al.*, 2007). Spatially and temporally congruent patterns across multiple, independent lineages indicate a shared history of responses to the same overriding events, e.g. vicariance due to geological processes (Rosen, 1978; Nelson & Platnick, 1981; Ronquist, 1997). Phylogeography also enables inference of the environmental histories of landscapes (e.g. habitats; Crawford, Bermingham & Polania, 2007; Wang, Crawford & Bermingham, 2008), as 'ecological niche conservatism' tends to hold over evolutionary timescales across taxa (Peterson, Soberón & Sánchez-Cordero, 1999; Wiens & Graham, 2005); this can provide additional information on histories of species within a biogeographical region, independent of geological processes. Phylogeography, especially comparative phylogeography, of Neotropical biotas will therefore be most illuminating when applied in geographically and geologically complex areas lacking historical consensus (Arbogast & Kenagy, 2001; Castoe *et al.*, 2009; Daza, Castoe & Parkinson, 2010).

One such area, the lower Central American (LCA) isthmus, presents an exceptional natural laboratory for studying the recent historical assembly and diversification of a Neotropical biota using comparative phylogeography (Fig. 1). Here, we review and critically evaluate LCA phylogeography studies against a backdrop of the geologic and climatic setting, and previous biogeography studies, to provide a framework for subsequent work. We close our review by emphasising the need to improve inferences through in-depth comparative phylogeography analyses using the latest statistical phylogeographical approaches. We also discuss ways that future research can apply phylogeography to jointly refine our understanding of LCA biodiversity and regional conservation.

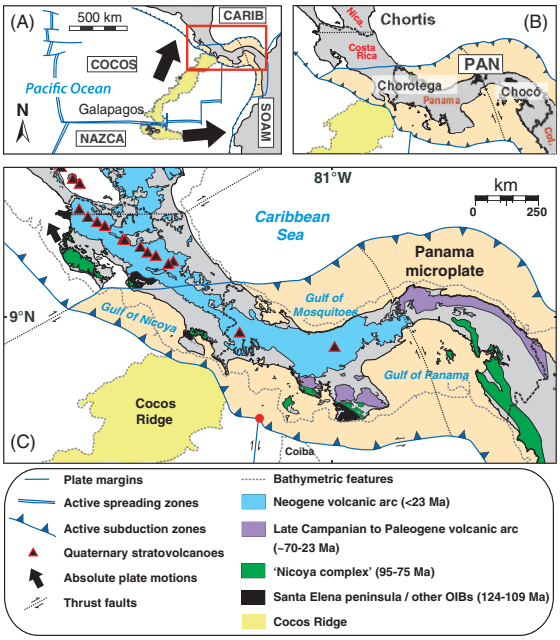

Fig. 1. Maps summarising the present-day tectonic setting and geology of lower Central America (LCA) and the Panama microplate (PAN; light orange shading). (A) Plate tectonic overview showing plate boundaries and absolute motions. Plate names are as follows: COCOS, Cocos; NAZCA, Nazca; CARIB, Caribbean; SOAM, South American. The Cocos plate and its prominent Cocos Ridge (yellow shading) subduct beneath CARIB, whereas NAZCA subducts beneath both CARIB and SOAM. Subaerial land is shaded grey. The red box delineates the study area, detailed in (B, C). (B) Basement blocks of the western Caribbean plate: Chortis, Chorotega, and Chocó. The Panama microplate comprises Chorotega (in part) and Chocó blocks. Country names are given in red (Col., Colombia; Nica., Nicaragua). The full extent of Chortis is shown in Fig. 4A. (C) Geological map of LCA showing major rock formations, quaternary stratovolcanoes of the Chorotega volcanic front (CVF), the Cocos–Nazca–Caribbean 'triple junction' (red dot), and major bathymetric features (fine-dotted grey lines) including the 200 m contour and Cocos and Coiba ridges (compiled after Coates & Obando, 1996; Carr *et al.*, 2007; Gazel *et al.*, 2008; Funk *et al.*, 2009; Buchs *et al.*, 2011).

II. LOWER CENTRAL AMERICA

We define lower Central America (7°11′–11°13′N, 77°10′–85°57′W) as the area spanning Costa Rica and Panama, plus nearby islands (e.g. Quepos, Bocas del Toro, Coiba, Las Perlas, Cocos). Part of the Caribbean plate (CARIB; Fig. 1A), LCA is underlain by Mesozoic oceanic basement formed by submarine volcanism > 80 million years ago (Ma; Mann, Rogers & Gahagan, 2007). LCA sits at the intersection of four tectonic plates, but it mostly forms the fault-bounded Panama microplate spanning the Chocó (in part) and Chorotega blocks (Fig. 1B). Along LCA's Pacific margin, the Cocos plate converges beneath CARIB at geological lightning speed (~85 mm/year), hindered by flat subduction of the aseismic Cocos Ridge (e.g. Funk *et al.*, 2009). Present-day active plate-boundary tectonics creates a high frequency of volcanic, earthquake, and mudslide hazards (Rose *et al.*, 2006; Sherrod *et al.*, 2007), which have likely contributed to localised population extinction and genetic isolation. LCA also carries risk of environmental damage from hurricanes (Atlantic), tsunamis (Pacific), and catastrophic flooding during wet season rains.

Despite covering only ~0.09% (127050 km^2) of earth's land area, only slightly larger than the state of Mississippi, LCA is one of the most physically and biologically complex areas on the planet (Fig. 1). Mainland LCA forms a long (~1170 km), narrow isthmus tapering from ~240 km across Costa Rica to merely 65 km at the Panama Canal basin. Physiography is largely defined by NW-trending, volcanic cordilleras intermittently bisected by fertile valley complexes (4 in Fig. 2A). The Chorotega volcanic front provides the most obvious regional geographical barrier including LCA's highest peak 3820 m above present sea level (a.s.l.), Cerro Chirripó Grande (Talamanca Cordillera). This and other Talamanca peaks create sky-islands of isolated montane habitat. Mountains of the Fila Costeña (mean ~1200 m a.s.l.; Fig. 2A), 'Nicoya complex' (> 600–900 m; Fig. 1C), Limón headland, and Darién (e.g. Cerro Sapo, 1145 m; Serranía del Darién, 1875 m; Fig. 2A) also add notable relief. In the Darién, these produce basin-and-range (e.g. Chucunaque basin–San Blas) topography. Elevations drop below 200 m in central Panama and along coastlines, except where steep-faced mountain ranges rise close to the ocean, constricting coastal plains to narrow corridors restricting movement of lowland species at Herradura headland, Bocas del Toro, Gulf of Mosquitoes, Soná peninsula, and Cerro Sapo. While much of LCA (36% land area) is tropical forest biome, it encompasses diverse vegetation zones from jungle-shrouded lowland wet and dry forests to mangrove estuaries, rolling savannas and grasslands, and once-pristine montane habitats (Fig. 2B; Marshall, 2007). Sharp climatic-vegetation transitions occur across headlands and the continental divide, which creates a Pacific-coastal rain-shadow effect. The resulting alternating pattern of wet forest, dry forest and savanna habitats along the Pacific versant has long been hypothesised to present climatic filter barriers limiting dispersal (e.g. Savage, 1966). Bocas del Toro, Perlas and Coiba islands are mostly forested and the closest islands (~35 km distance) of any real size to the LCA mainland. In comparison, Costa Rica's Cocos Island lies 550 km away. The hydrological network reveals many short incisive rivers; in NE Panama, essentially all rivers are < 15 km long (Bermingham & Martin, 1998). Major watersheds, e.g. Tuira and Chagres rivers, are spaced throughout and two of the largest tropical lakes worldwide, Lakes Managua (1042 km^2) and Nicaragua (8624 km^2), connect to the Caribbean through the Rio San Juan superbasin (Fig. 2B).

Despite its small size, LCA has among the highest levels of biodiversity per km^2 worldwide (Reid & Miller, 1989). Approximately 4–10% (~500000 species) of global biodiversity resides in Costa Rica alone, depending on the taxonomic group considered (Obando, 2002), and Panama may be more diverse. Beyond more than 300000 insect species, LCA harbours as many or more species of birds (> 970 species) and vascular plants (> 19500 species, 6.3–14.5% endemic) per 10000 km^2 as anywhere worldwide (Hurlbert & Villalobos-Figueroa, 1982; Stotz *et al.*, 1996; Davis *et al.*, 1997; Obando, 2002; Mutke & Barthlott, 2005). Its freshwater fishes (~170 species, 58% endemic), reptiles and amphibians (~830 species, 10–15% endemic), and mammals (~212 species in Costa Rica alone, < 5% endemic) are also highly species-rich or endemic (Savage, 1982, 2002; Obando, 2002; Smith & Bermingham, 2005; Abell *et al.*, 2008; Bolaños, Savage & Chaves, 2011; Fishbase, http://www.fishbase.org/). The Atlantic and Pacific coasts are often distinct biotic assemblages, e.g. for insects (Fig. 3A) and freshwater fishes (Fig. 3B). However, reptiles and amphibians are highly endemic in the Talamanca mountains (Fig. 3C) and herpetofauna, insect and plant areas of endemism overlap both central Panama coasts (Gentry, 1982; Savage, 1982; Morrone, 2006). These biodiversity patterns suggest that factors promoting *in situ* geographical isolation have played an important role in shaping LCA biotas. However, LCA's rich biodiversity is likely attributable to multiple factors including its position within tropical

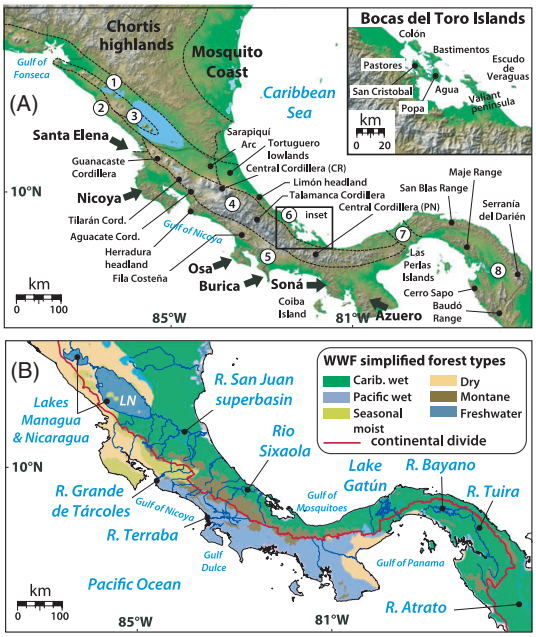

Fig. 2. Map of present-day physiography and vegetation cover of lower Central America (LCA). (A) Physiographic province boundaries (dashed lines) enclosing distinct LCA landform assemblages, drawn over digital elevation model derived from NASA SRTM image PIA03364 (after Marshall, 2007). Provinces: 1, Nicaraguan depression (ND); 2, Sandino fore arc; 3, Nicaraguan volcanic front (NVF); 4, Chorotega volcanic front (CVF); 5, Chorotega fore arc; 6, Chorotega back arc; 7, Panama Canal Zone (PCZ) lowlands; 8, Darién isthmus. Major peninsulas and headlands (arrows) and mountain ranges mentioned in the text are indicated. Names of major island chains are also given. The inset map describes the Bocas del Toro (BDT) archipelago and mainland. (B) World Wildlife Fund forest ecoregions (modified from Crawford *et al.*, 2007) and major freshwater drainages (LN, Lake Nicaragua) shown in reference to the continental divide (red line).

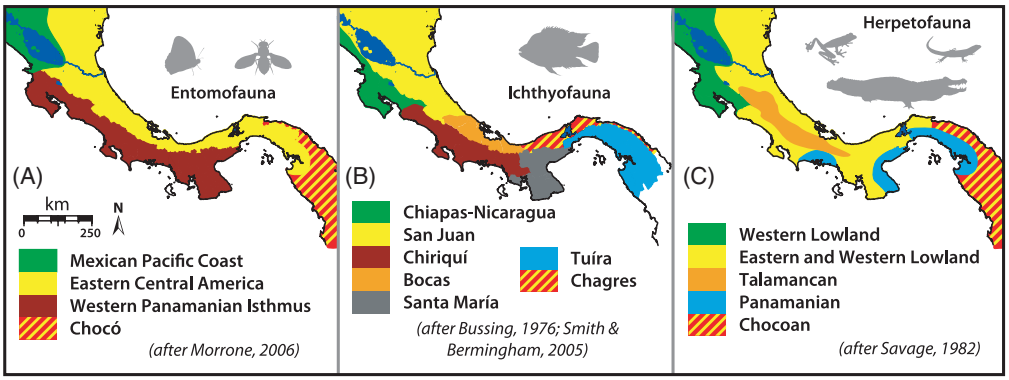

Fig. 3. Examples of biogeographical province boundaries in lower Central America (LCA) representing areas of high species turnover, shown for (A) insects, (B) freshwater fishes and (C) herpetofauna. Provinces (colours) reflect areas of endemism with distinct biotic assemblages. Below each panel, a legend of the province names is provided. Province names are given from their original sources (referenced within each panel).

latitudes, its role as a transition zone between North and South American biotas, its varied physiography and geomorphology, and its rich geologic history (Whitmore & Prance, 1987; Jackson, Budd & Coates, 1996).

III. GEOLOGICAL HISTORY

Geodynamic evolution in LCA dates back over 100 Ma, from the Early Cretaceous onset of Santa Elena peninsula formations (124 to 109 Ma; Hauff *et al.*, 2000) and Costa Rica arc volcanism (Gazel *et al.*, 2008), to Holocene isolation of Bocas del Toro islands ∼10 to 1 thousand years ago (ka) due to sea-level rise and continental submergence (Anderson & Handley, 2002). This interval witnessed dramatic geographic changes critical to the assembly of LCA landforms and biotas, altering probabilities of dispersal, vicariance, and extinction through time.

In the Late Cretaceous, LCA was incorporated along the western CARIB after fusing to the Chortis block, and the Chocó block (e.g. Baudó Range) was developing *via* submarine oceanic plateau volcanism. Early LCA is an enigma; however, as the dinosaurs were going extinct 65 Ma, Soná peninsula (and possibly Osa peninsula) was subaerial (Hauff *et al.*, 2000; Hoernle *et al.*, 2002). By Palaeocene times LCA was a Pacific island archipelago and dispersal into the region ∼60 to 50 Ma must have occurred over an up to ∼400 km ocean gap to the north and ∼400–1500 km ocean gap(s) to the south, based on plate reconstructions (Fig. 4A; Hauff *et al.*, 2000; Mann *et al.*, 2007; Scotese, 2008). The Baudó terrane remained submerged, but emergent Eocene lands included Nicoya complex terranes (Fig. 1; 95 to 75 Ma; Seyfried *et al.*, 1991; Hauff *et al.*, 2000) and active Azuero peninsula and San Blas arcs separated by ocean connections (Montes *et al.*, 2012). By mid-Eocene ∼50 Ma, volcanism increased in western Costa Rica and was peaking in the San Blas Range and part of the Atrato basin, where it slowed 38 to 15 Ma (Seyfried *et al.*, 1991; Montes *et al.*, 2012). The earliest remnants of *in situ* volcanism surface in Costa Rica–Nicaragua stratigraphic records in the Late Eocene–Miocene, beginning with alkaline Sarapiquí arc (Fig. 2A) eruptions (Abratis & Wörner, 2001; Gazel *et al.*, 2008).

LCA's major morphotectonic features formed largely since the Miocene. The Cocos plate formed ∼23 Ma *via* Farallón plate rifting, subducted beneath CARIB, and has since uplifted LCA substantially (Mann *et al.*, 2007; Marshall, 2007). The Baudó terrane surfaced in the Miocene. Yet LCA's overall Miocene configuration remains disputed. One 'peninsula model' (Fig. 4B) posits that a long, narrow peninsula jutting from Chortis ∼25 to 16 Ma progressively narrowed the Atrato seaway gap with Colombia. This is supported by land-mammal fossils (Whitmore & Stewart, 1965; Kirby & MacFadden, 2005; Retallack & Kirby, 2007) and stratigraphic dating analyses (Kirby, Jones & MacFadden, 2008). Upgraded Panama geological data and maps implicate that mountain ranges east of the Panama Canal were emergent around Late Eocene and helped form a contiguous peninsula since the Miocene (Montes *et al.*, 2012). The alternative 'island model' based on palaeobathymetric and sedimentary records posits that a Mid-Miocene–Pliocene volcanic archipelago spanning western Costa Rica to Colombia was disconnected from Nicaragua and South America, leaving marine connections open across the nascent isthmus (Fig. 4C–E; Coates & Obando, 1996; Coates *et al.*, 2004). The former model requires over-water dispersal by colonising propagules, followed by movement along contiguous land; the latter would require multiple over-water dispersal bouts between segments. The Early-Mid Miocene saw the accretion of Nicoya complex terranes and other basalts to Panama ∼20 to 15 Ma (Hauff *et al.*, 2000; Hoernle *et al.*, 2002). By Mid-Miocene, substantial land was emerging in central-SE Costa Rica: the Talamanca Cordillera began forming

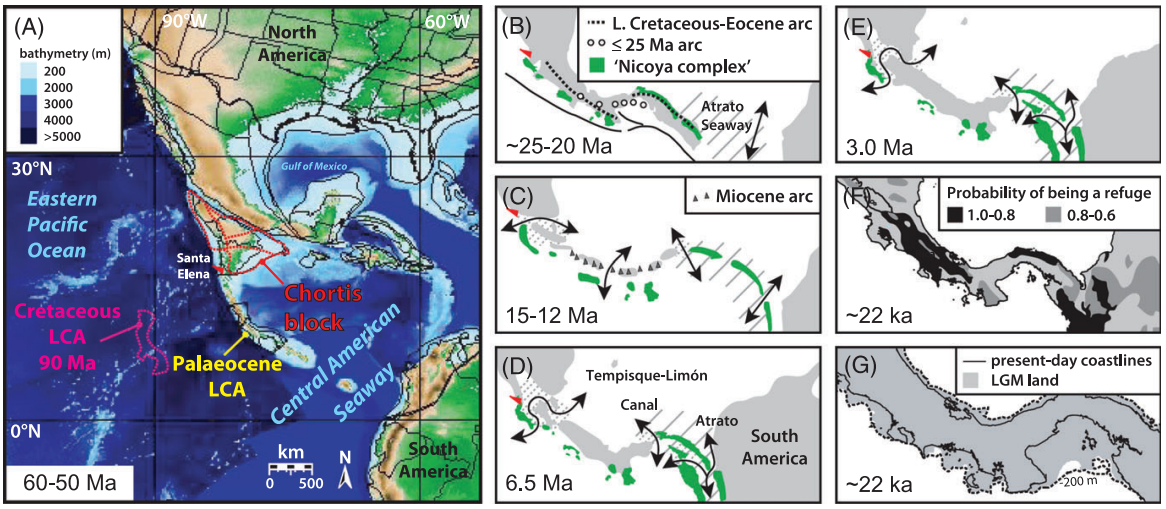

Fig. 4. Palaeogeography of lower Central America (LCA). (A) Palaeocene–Eocene plate reconstruction showing the position of LCA, the Chortis block (red-dashed lines), and the Santa Elena peninsula (red triangle) (modified from Scotese, 2008). Note that LCA was an ancient island archipelago at this time, isolated from nearby mainland areas to the north and southeast (shaded green to brown with increasing elevation above sea level) by ocean gaps (blue; see bathymetric legend) including the Central American Seaway. The position of LCA ~90 Ma in the Cretaceous (Mann *et al.*, 2007) is shown in magenta. (B–E) Miocene–Pliocene reconstructions of LCA: light grey, emergent land; grey diagonal lines, abyssal to bathyal (> 2000 m) depths; grey dotting, neritic depths; green shading, exotic oceanic terranes known as the 'Nicoya complex'; arrows, marine corridors; red triangles, subaerial Santa Elena peninsula. The geography of LCA ~25 to 20 Ma based on the 'peninsula model' (B; redrawn after Montes *et al.*, 2012), is contrasted against (C) middle Miocene (15 to 12 Ma), (D) Late Miocene (6.5 Ma, pre-Cocos Ridge), and (E) Pliocene (3 Ma, initial isthmus closure) 'island model' reconstructions (after Coates & Obando, 1996; Coates *et al.*, 2004, 2005). Names of marine corridors are given in (B, D). (F, G) LCA environments during the Last Glacial Maximum (LGM) ~22–19 ka. (F) Proposed remnant Pleistocene forests shaded according to their probabilities of being refugia, after Whitmore & Prance (1987). (G) Palaeobathymetry model showing LCA land (grey shading) extending over the continental shelf during the LGM (−110 m sea levels; modified from Smith & Bermingham, 2005) in the context of present-day coastlines and the 200 m contour.

17.5 Ma, and then the Aguacate Cordillera started 11.4 Ma and went extinct ~4.0 Ma (reviewed by Gazel *et al.*, 2008). LCA reached its modern position after a major Mid-Miocene collision with South America 12.8 to 7.1 Ma that created active left-lateral strike-slip faults along the Darién (Coates *et al.*, 2004). Around the same time, extensional forces at LCA's northern boundary formed the Nicaraguan depression (Fig. 2A), a long, fault-bounded rift valley spanning El Salvador's Median Trough to the Tortuguero lowlands basin. This depression opened SE–NW ~10 to 0 Ma, especially following Cocos slab break-off 10 to 4 Ma (Mann *et al.*, 2007; Funk *et al.*, 2009).

The most stunning changes occurred over Late Neogene–Quaternary, when gradual emergence of the LCA isthmus cut off the Central American Seaway and permanently linked the Americas for the first time in the Late Pliocene. Key events included: (*i*) collision and subduction of the Cocos Ridge 5.5 to 3.5 Ma beneath Costa Rica, which rapidly uplifted the Chorotega volcanic front and sparked increased volcanism, forming the Fila Costeña and Talamanca Cordillera (Abratis & Wörner, 2001; Mann *et al.*, 2007); and (*ii*) deposition of the Limón, Canal Zone, Chucunaque and Darién basins 7 to 0 Ma by crustal erosion associated with Chorotega and Andean uplift (reviewed by Coates & Obando, 1996). By ~4.6 Ma, ocean currents and ecosystems became reorganised (Keigwin, 1982; Haug & Tiedemann, 1998). The isthmus then became fully closed by at least 3.5 to 3.1 Ma before a permanent Isthmian Link with South America formed 3.1 to 1.8 Ma (Keller, Zenker & Stone, 1989; Duque-Caro, 1990; Coates *et al.*, 1992; Coates & Obando, 1996; Ibaraki, 2002). Combined with the simultaneous and rapid uplift of the Colombian Andes (Gregory-Wodzicki, 2000), development of the LCA isthmus played a role in Miocene–present global climate change by altering patterns of regional oceanic and atmospheric circulation, resulting in more intense Atlantic thermohaline circulation, more high-latitude Northern Hemisphere precipitation, and larger ice sheets (Keigwin, 1982; Schmidt, 2007; Lunt *et al.*, 2008). The last 2.2 to 0 Ma of the Quaternary were marked by activity of Chorotega volcanic front stratovolcanoes, which laid several sizeable debris fans that likely destroyed everything in their paths. The ~1.7 to 1.1 million year (Myr)-old Orotina debris fan (Avalancha formation) overlying the Rio Grande de Tárcoles (Marshall *et al.*, 2003) and Late Pleistocene Barú volcano debris fan (Sherrod *et al.*, 2007) provide good

examples of the latter. Overall, most of the LCA landscape formed since the Neogene. LCA thus provides a remarkable biogeographic experiment where, unlike continents, large subaerial areas are relatively young.

IV. CLIMATE AND SEA-LEVEL FLUCTUATIONS

Radical geologic evolution in LCA has been accompanied by climate and sea-level fluctuations altering the spatial and taxonomic habitat composition. Global climate progressively moistened and cooled through the Late Cenozoic, dropping to near present-day temperatures by ~4 to 2 Ma (Fig. 5A). Since at least that time (> 4 Ma), Pacific dry forest habitats of today have essentially been intact (Graham & Dilcher, 1995), although they probably oscillated between forest patches and savannas during the vicissitudes of the Late Pleistocene (Piperno & Pearsall, 1998). Around 39.4 to 28.1 ka during the Late Pleistocene, cold/humid conditions with relatively high seas and lower precipitation prevailed in LCA (González, Urrego & Martínez, 2006). Subsequently, LCA climate became much cooler and sea levels reached their lowest levels 28 to 14.5 ka (González *et al.*, 2006), overlapping the Last Glacial Maximum (LGM) 22 to 19 ka. LGM pollen records and other data show that mean 5–8°C cooling throughout LCA shifted montane forests down in elevation, creating a highland Costa Rica-Panama páramo corridor (Bush *et al.*, 1992; Colinvaux, 1996; Colinvaux *et al.*, 1997; Islebe & Hooghiemstra, 1997). The highest Talamancan peaks were simultaneously covered by small glaciers (< 50 km^2) that deglaciated ~10 ka (Lachniet, 2004). Whether corridors of savanna habitat existed over wide swathes of Neotropical lowlands during the LGM (or other Pleistocene periods) is intensely debated (e.g. Colinvaux, De Oliveira & Bush, 2000). The 'Pleistocene refugia hypothesis' (Haffer, 1969), which provided the impetus to spark this debate, explains terrestrial areas of endemism by predicting that glacial aridity fragmented LCA/Amazonian forests, isolating lowland taxa in persistent upland refugia separated by savannas (Fig. 4F; Haffer, 1969, 1997; Whitmore & Prance, 1987). The alternative 'disturbance-vicariance' hypothesis (Colinvaux, 1993, 1996; Colinvaux *et al.*, 1997) posits that LCA/Amazonia never had upland forest refugia fragmented by aridity, or wide savanna corridors (*cf*. Webb & Rancy, 1996). Rather, LCA was a mosaic of Late Pleistocene forest patches caused by glacial cooling cycles, and uplands (> 500 m a.s.l.) carried diverse montane biota without invading lowland species (Bush *et al.*, 1992; Colinvaux, 1993, 1996). Following the LGM, LCA experienced cooling and near present-day warming cycles 14.4 to 11.1 ka, while the ensuing Holocene was characterised by drying, human disturbance, and relative climatic stability (Bush *et al.*, 1992; Bush & Colinvaux, 1994; Leigh, O'Dea & Vermeij, 2014). Modern LCA climate is tropical (daily highs throughout the region range between 23.9 and 32.2°C, year round), and tropical moist forests that typically receive > 2000 mm total annual precipitation dominate land cover (Fig. 2B). However, high-elevation zones (e.g. Talamanca Cordillera) experience lower temperatures and have shrub- and grass-dominated páramo habitat; and Pacific environments of Santa Elena, Nicoya and Azuero peninsulas possess dry forests characterised by < 2000 mm total annual precipitation (Fig. 2B).

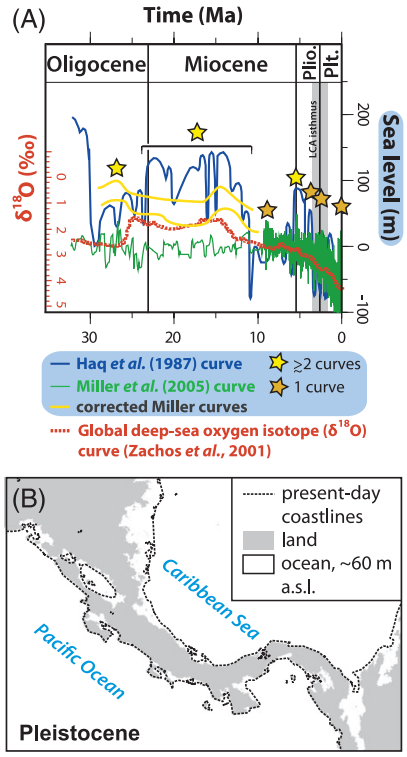

Fig. 5. Links between global changes in Late Cenozoic climate and sea levels, and their potential impacts on lower Central America (LCA). (A) Mean deep-sea oxygen isotope ($\delta^{18}O$, in parts per thousand) record (dark red dashed line), a temperature proxy positively correlated with global cooling, from Zachos *et al.* (2001). Several eustatic sea-level curves are also given. Dark green trend lines are from Miller *et al.* (2005); yellow trends are Miller *et al.*'s (2005) curves corrected and smoothed by Kominz *et al.* (2008); and the dark blue trendline is from Haq *et al.* (1987). Plio., Pliocene; Plt., Pleistocene. Stars indicate sea-level spikes greater than or equal to ~ 25 m a.s.l.; yellow stars indicate support from two or more curves; orange stars reflect support from a single curve. These sea-level highstands may have substantially inundated emergent LCA lowlands. (B) Marine inundation of LCA during the Pleistocene, modelled as a hypothetical high-sea stand ~ 60 m a.s.l. based on present-day digital elevation data (NASA SRTM, 90 m). This model presents a conservative estimate illustrating potential effects of extensive Pleistocene sea-level spikes, e.g. hypothesised by Nores (1999, 2004; see text).

Long-term eustatic sea-level estimates indicate that the seas have dropped rather continuously since LCA land began to emerge, especially since the Eocene–Oligocene

transition (Fig. 5A). However, multiple high-sea stands have affected LCA biogeography. Sedimentary records show that a marine corridor inundated the Nicaraguan depression until at least Late Pliocene (Coates & Obando, 1996), and this undoubtedly limited LCA–nuclear Central America (Guatemala to Nicaragua) (NCA) dispersals (e.g. of freshwater fishes; Bussing, 1976). Multiple eustatic curves converge on similar Miocene spikes ∼25 to 50 m a.s.l. around 20, 14, and 12 Ma (Kominz *et al.*, 2008; Müller *et al.*, 2008), and these likely created or maintained an LCA archipelago configuration, at least temporarily. Late Miocene–present spikes ≥ 25 m a.s.l. were inferred ∼9 Ma and 5 to 4.5 Ma (Fig. 5A) and sea levels breached the Panama Canal ∼7 to 6 Ma (Coates *et al.*, 2004). Reliably determined Pleistocene eustatic spikes ∼20 m a.s.l. occurred 2.4 to 1.8, 1.3 and 0.45 to 0.1 Ma and potentially connected the Atlantic and Pacific oceans (Keller *et al.*, 1989; Hearty *et al.*, 1999; Miller *et al.*, 2005). However, substantial Neotropical diversity is thought to have resulted from ∼100 m Miocene–Quaternary marine incursions supported by coastline and Amazonian studies (Webb, 1995; Nores, 1999, 2004). While the extent and timing of such incursions outside Amazonia remains controversial because evidence stems from tectonically uplifting areas, other reviews list large, potentially +85 m a.s.l. peaks around the 0.63 Ma interglacial (Mediterranean basin; Emig & Geistdoerfer, 2004). We modelled a slightly lower incursion, +60 m a.s.l., over modern elevations because LCA land approximated modern landmasses around this time [unpublished data, based on a 90 m-resolution NASA Shuttle Radar Topography Mission digital elevation model (http://www2.jpl.nasa.gov/srtm/)]. Results indicated that had major sea level highstands occurred during Pleistocene interglacials, these could have widely inundated lowland LCA habitats, altering lowland species distributions and causing genetic isolation (Fig. 5B).

V. CONSERVATION SIGNIFICANCE

As part of the Mesoamerica biodiversity hotspot, LCA contains exceptional biodiversity and endemism (Section II); unfortunately, hotspot membership is also predicated upon widespread and active human threats and therefore reflects growing recognition that LCA's fascinating biodiversity is in peril (Myers *et al.*, 2000). Land use is a leading proximate cause of global biodiversity loss and human-induced environmental change (Sala *et al.*, 2000), and a major threat to LCA biotas. Habitat destruction is spreading in LCA due to widespread land-clearing for agriculture and cattle ranching, combined with rampant human population growth: less than 40% of virgin forests remained in Central America by the late 1980s (Leonard, 1987), and today only 20% of original primary vegetation extent remains (Myers *et al.*, 2000). Some evidence suggests active restoration efforts in Costa Rica are helping to increase overall forest area back towards pre-1970s levels; however, problems such as hillside deforestation, forest conversion, and firewood acquisition remain widespread and ongoing. Invasive species, e.g. rainbow trout (*Oncorhynchus mykiss*) introduced to Costa Rica in 1925 (Hildebrand, 1938), and Mozambique tilapia (*Oreochromis mossambicus*) introduced in LCA since 1950 for aquaculture (Welcomme, 1988), threaten the fragile balance of LCA ecosystems. Environmental pollution is an enormous problem seen everywhere. Freshwater ecosystems are widely threatened by introduction of agrochemicals that poison local communities (e.g. localised fish-kills in Costa Rica due to pesticide poisonings), watershed destruction, and increased flooding and sediment loads due to logging and land cultivation (Leonard, 1987; Bussing, 1998). Commercial species of lobsters, shrimps, anchovies, and turtles have long been overexploited (Leonard, 1987). What is more, disease also poses a major threat; particularly alarming is the case of contagious fungal diseases. Well-known cases of amphibian declines in Costa Rica and Panama, which occur suddenly and sometimes result in species extinction, affect possibly up to half of extant amphibian species in the region (e.g. Lips, Reeve & Witters, 2003; Young *et al.*, 2004).

Despite many threats, LCA also is an international model of conservation efforts. This includes preservation of natural areas. Fully 28% of Costa Rica is legally protected land (93 total protected areas), and another 264228 ha are private nature reserves, while 34.4% of Panama (89 protected areas) is legally protected and 40000 ha are private reserves (Evans, 1999; Chacón, 2005, 2008; ANAM, 2010). LCA is also part of the United Nations Man and Biosphere Programme and the location of four UNESCO Biosphere Reserves, including two unique Costa Rican reserves (Cordillera Vocánica Central, Agua y Paz), one unique Panama reserve (Darién), and La Amistad reserve, a 'peace park' spanning the Costa Rica–Panama border (UNESCO, 2012*a*). This program integrates science, education and social programs to promote sustainable development, e.g. involving local communities in ecosystem management. An area of incredible cultural diversity due to its 'frontier' anthropological history, LCA is also multicultural, boasting > 80 different dialects; thus cultural and linguistic diversity will need to be maintained along with nature. Due to consideration as having outstanding universal value, seven LCA areas are designated UNESCO World Heritage Sites, including La Amistad Reserve, Cocos Island National Park, and the Area de Conservación Guanacaste in Costa Rica, as well as four unique Panamanian sites —Fortifications on the Caribbean Side of Panama: Portobelo-San Lorenzo (endangered due to lack of management and urban development), Darien National Park, La Amistad, the Archaeological Site of Panamá Viejo and Historic District of Panamá, and Coiba National Park and protected marine zone (UNESCO, 2012*b*). LCA countries also participate in a World Bank-funded regional partnership, the Mesoamerican Biological Corridor project (http://www.biomeso.net/), aimed at conserving ecological connectivity and promoting environmentally sustainable development through linking > 321000 km^2 of protected

areas from southern Mexico to Panama (IEG, 2011). However, despite multi-million dollar investments, conservation resources allocated to fresh waters lag well behind those committed to terrestrial and marine ecosystems.

VI. A BRIEF HISTORY OF LOWER CENTRAL AMERICAN BIOGEOGRAPHY

(1) A rare land bridge

LCA is earth's sole interoceanic and intercontinental landmass and most prominent land bridge. Land bridges are important in historical biogeography because they cut off marine connectivity, isolating communities on either side, and facilitate convergence of continental biotic components through inland dispersal (Lomolino *et al.*, 2010). Not surprisingly, the importance of LCA in shaping New World biogeography has long been recognised. Charles Darwin and other 19th Century naturalists thought that LCA served as a refuge where temperate North American and tropical American vegetation mixed, surviving the glacial stages (Darwin, 1859, pp. 338–340). Both Darwin and Alfred Russel Wallace cited work by Günther (1861) on marine fish communities, which, owing to considerable similarity in community composition on either side of the isthmus, indicated previous linkage(s) between the Pacific and Atlantic oceans (Darwin, 1859, p. 317; Wallace, 1876, p. 40). Based on the fish data and his familiarity with fossilised marine gastropods, Wallace (1876) explicitly hypothesised that the area where modern LCA resides became inundated during the Miocene. It was also Wallace (1876) who first articulated the importance of this 'small and insignificant' sliver of land as a driver of biotic convergence and interchange between North and South American biotas. Wallace considered LCA more effective in facilitating inland dispersals than other isthmuses, such as the desert Isthmus of Suez between Africa and Asia:

> 'The Isthmus of Panama is a more effectual line of union [biotic convergence], since it is hilly, well-watered, and covered with luxuriant vegetation; and we accordingly find that the main features of South American zoology are continued into Central America and Mexico' (Wallace, 1876, p. 38; our clarification in brackets).

However, Darwin and Wallace could not have known how right they were on these latter points. Like other 19th to mid-20th Century biogeographers, their approach to historical biogeography was limited to a vague understanding of phylogeny derived from morphology-based taxonomic lists and Charles Lyell's geologic model of continental and oceanic 'stasis', or permanence.

(2) The Great American Biotic Interchange and beyond

When the Isthmian Link emerged ∼3 Ma in the Late Pliocene, a 'Great American Biotic Interchange' (GABI) of terrestrial and freshwater species ensued overland, yielding increased species turnover and filling of open niches, and range expansions, speciation, and extinctions across the Americas (Stehli & Webb, 1985). By elucidating the sequence of this incredible natural experiment, 20th Century biogeographers made LCA famous worldwide as an example of the influence of continental convergence and land-bridge formation in shaping biotas. Wallace (1876, p. 131) had hypothesised that South American mammals invaded North America before the ice ages. However, it was not until classic studies by George Gaylord Simpson that detailed GABI histories of many lineages became fully known. Simpson (1940, 1950) recognised three 'strata' of South American land-mammalian fossils and derived an ecological and biogeographical explanation for their movements between North and South America before and after LCA isthmus emergence (reviewed by Stehli & Webb, 1985; Riddle & Hafner, 2010). Simpson showed that lineages moved predominantly southward across LCA to invade South America over Late Miocene–recent, more than had done so at previous times; however, interamerican dispersals had started before then and were bidirectional. For example, South American 'herald taxa' suddenly appeared in the Mid-Miocene mammal record of North America *via* (over-water) waif dispersals northward, or 'island hopping' (Simpson, 1950; Webb, 2006). A host of large-scale, distribution-based biogeography studies followed Simpson, including studies of Neotropical plants (Raven & Axelrod, 1974; Gentry, 1982; Gómez, 1986), insects (Halffter, 1987; McCafferty, 1998), freshwater fishes (Miller, 1966; Myers, 1966; Rosen, 1975; Bussing, 1976, 1985; Smith & Bermingham, 2005) and tetrapod amphibians and reptiles (Savage, 1966, 1982, 2002; Campbell, 1999), birds (Karr, 1990), and mammals (Marshall, 1979; Marshall *et al.*, 1979; Kirby & MacFadden, 2005). Drawing on improved 1960s–1970s field museum collections, plate tectonics theory, and geological mapping, these studies inferred historical scenarios, including GABI sequences, and heavily influenced LCA historical biogeography. Work through the early 1980s culminated in a synthesis of GABI histories of many organismal groups, and geological models, led by S. David Webb, Larry Marshall, and colleagues (Stehli & Webb, 1985). Over the 1990s to present, this synthesis has been updated by large-scale analyses of fossil and extant species records (e.g. Vermeij, 1991; Cadle & Greene, 1993; Webb & Rancy, 1996; Webb, 1997, 2006; Burnham & Graham, 1999; Leigh *et al.*, 2014). In contrast to land-mammalian patterns, the above studies demonstrated that rainforest plants, mayflies, freshwater fishes, herpetofauna, and rainforest birds dispersed predominantly northward from South America (despite bidirectional GABI movements) to become established in Central America, and beyond, since the Neogene.

The broad-scale studies mentioned above proposed numerous testable hypotheses of biogeographic provinces (e.g. based on areas of endemism) and elements, plus dates of Mesozoic-recent dispersals into LCA from outlying areas and *in situ* diversification or extinction of clades.

In particular, Savage (1966, 1982) proposed four Central American herpetofaunal elements — 'South American', 'Middle American', 'Old Northern' and 'Young Northern'. Savage hypothesised that the older extant lineages of these elements dispersed into LCA and NCA during the Late Cretaceous, went extinct in LCA over Eocene–Miocene, then reinvaded the isthmus during Miocene–Pliocene times (Savage, 1966, 1982). Bussing (1976) proposed three biogeographic elements of freshwater fishes and concluded, similarly to Savage, that South American fish lineages colonised NCA during the Cretaceous–Palaeogene *via* a temporary interamerican land bridge (that later disappeared), diversified in NCA and South America, and then invaded LCA in a second Miocene–Pliocene wave. Although LCA was never a Cretaceous land bridge, these models reconcile with current tectonic models: the proto-Greater Antilles arc passed through the ocean gap between North and South America over the Cretaceous–Paleogene (e.g. Hauff *et al.*, 2000; Hoernle *et al.*, 2002; Mann *et al.*, 2007). In a later work, Savage (2002) hypothesised that LCA highland herpetofaunal diversity originated as a result of Plio–Pleistocene climatic fluctuations (similar to refuge theory) as montane habitats shifted down in elevation during glacials, then up (being isolated again) during interglacials. Raven & Axelrod (1974) identified many northern-continent plant families thought to have migrated from South America during the Cenozoic. Likewise, Gentry (1982) hypothesised the presence of two floristic elements in LCA, 'Gondwanan' and 'Laurasian', and inferred that their biogeographic history was dominated by asymmetrical northward dispersals of South American lineages (lianas, canopy trees) into lowland plant communities following LCA isthmus emergence. While many biogeographical hypotheses from the above studies can be tested using phylogeography, their claims remain seldom tested by phylogeography studies today (but see, for example, Dick, Abdul-Salim & Bermingham, 2003; Castoe *et al.*, 2009; Streicher, Crawford & Edwards, 2009).

(3) Biogeographical paradigms

The 20th Century witnessed confrontations between several major biogeographical paradigms. In the 1960s, plate tectonics became accepted then superseded 'land-bridge biogeography' and 'oceanic dispersal' as the dominant theory explaining intercontinental biogeography (Raven & Axelrod, 1974; Lomolino *et al.*, 2010). MacArthur & Wilson's (1967) Equilibrium Theory of Island Biogeography (ETIB) revolutionised 'static' biogeographical thinking by modelling species diversity as the outcome of a dynamic balance between migration and extinction. Much subsequent debate in historical biogeography sought to establish the primacy of dispersal *versus* vicariance in explaining large-scale biogeographical patterns (Lomolino *et al.*, 2010). While vicariance (or 'cladistic') biogeography (e.g. Nelson & Platnick, 1981), which aligned plate tectonics and phylogenetics, came to dominate explanations of continental biogeography, island biogeography remained best explained by dispersal and related processes (Lomolino *et al.*, 2010). Land-bridge/oceanic-dispersal theories were generally abandoned in favour of vicariance; in LCA, however, available evidence supported land-bridge and oceanic-dispersal scenarios well before such debates emerged. Even as vicariance biogeography bloomed, biogeographers easily maintained the classic view that dispersal along an evolving land bridge, combined with vicariance and extinction events, explained LCA biogeography, e.g. species emplacement and diversification (Simpson, 1940, 1950; Savage, 1966, 1982; Raven & Axelrod, 1974; Marshall *et al.*, 1979; Gentry, 1982; Bussing, 1985). Geological evidence firmly supports a Pacific LCA origin, meaning no truly vicariant divergences (taxa) could even exist between LCA and outlying continental lineages, *sensu* Gondwanan vicariance; and fossils irrefutably show that taxa colonised LCA *via* oceanic dispersal. Thus the question in LCA biogeography is not whether vicariance and dispersal (oceanic or inland) occur, but what has been the sequence, effects, and relative importance of these events? Today, island biogeography is undergoing a paradigm shift setting aside the ETIB and vicariance biogeography (Heaney, 2007). Biogeographical data increasingly show that these models are inadequate to explain island life, and a new island biogeography paradigm is emerging combining elements of both ecological and evolutionary dynamics of relevant processes (reviewed by Heaney, 2007; Lomolino *et al.*, 2010). Species diversity patterns analysed in light of modern geography typically upheld ETIB predictions, whereas historical perspectives from phylogeography have repeatedly challenged this model [e.g. demonstrating very ancient island lineages (falsifying predicted high turnover rates), island–island migration, and intra-island speciation; Brown & Lomolino, 2000]. The eco-evolutionary shift also appears directed at remedying the well-known poor fit between the ETIB and systems and processes operating over geological timescales (i.e. that the model *sensu stricto* was limited to ecological timescales, assuming equivalent dispersal abilities/probabilities among species and no speciation; Heaney, 2007). Given its numerous land-bridge islands (Figs 1 and 2), LCA phylogeography is primed to contribute to the present period of testing and reshaping island biogeography theory, although few studies exist so far (but see below).

(4) The advent of molecular biogeography

Molecular data and analytical tools have assumed enormous importance in biogeography. Since the 1980s–1990s, advances in DNA sequencing technologies (e.g. polymerase chain reaction, automated sequencing, next generation sequencing) and molecular phylogenetic methods, and the exploding phylogeographic literature and toolkit, have fuelled an era of rejuvenated interest and growth in historical biogeography (Riddle *et al.*, 2008; Knowles, 2009; Hickerson *et al.*, 2010). Compared with traditional area-based inference (reviewed by Posadas *et al.*, 2006; Ebach & Tangney, 2007), DNA-based biogeography provides critical improvements such as molecular estimates of lineage divergence dates. Aside from dating species origins, molecular dating

permits empirically testing hypotheses (e.g. vicariance dates), thereby elucidating the timing *and* mechanisms underlying biogeographical patterns. Among their many advantages, molecular analyses provide billions more DNA characters than morphology/distribution-based approaches; explicitly model nucleotide substitution and other processes; accommodate evolutionary rate heterogeneity (e.g. relaxed clocks); permit splitting DNA matrices into separately modelled data partitions (e.g. by gene; no comparably sophisticated models exist for morphological evolution); and estimating lineage divergence times (T), with or without fossil information [best if rates or calibration points are well established, e.g. by taxonomic group or geological event (Lomolino *et al.*, 2010)]. With appropriate outgroup sampling, phylogeographical methods permit testing a variety of hypotheses, e.g. population demographic models and topological models, to infer the historical sequence of dispersal, vicariance, extinction and recolonisation events by which lineages arrived and diversified in LCA (Bermingham & Martin, 1998; Crawford *et al.*, 2007). Also, while area-based inference relies heavily on endemism (Nelson & Platnick, 1981), phylogeography infers historical events or processes (e.g. population expansion) even when spatial-genetic endemism is absent (Zink, 2002; Garrick, Caccone & Sunnucks, 2010). As noted above, phylogeography is also highly synthetic: by the 1990s, improved palaeogeographic models for LCA (Coates & Obando, 1996) and eustatic sea-level curves (Haq, Hardenbol & Vail, 1987) became available, and early studies showed that these could be used in conjunction with inferred phylogeographic relationships among populations to derive historical scenarios and test geological models (Bermingham & Martin, 1998). The advent of molecular biogeography has also provided impetus for refining GABI sequences, mainly using higher-level phylogenies (e.g. Weir, Bermingham & Schluter, 2009; Cody *et al.*, 2010), and sparked trans-isthmian marine studies developing LCA as a classical model of allopatric speciation in the oceans (e.g. Bermingham, McCafferty & Martin, 1997; Lessios, 2008).

VII. PHYLOGEOGRAPHICAL PATTERNS EMERGING FROM LOWER CENTRAL AMERICA

Here, we summarise (with select examples) major patterns emerging from phylogeographic studies of LCA taxa. Multi-taxon phylogeographical breaks and evidence for general patterns of dispersal, vicariance and other processes impacting the assembly and diversification of LCA biotas are emphasised. Our review draws on a database of 58 phylogeography studies, including 57 studies consistent with our goals published between 1996 and 2012, in addition to one of our own unpublished studies of three freshwater fish species (see online supporting information, Appendix S1). Studies to date represent ~94 nominal taxa sampled from multiple sites throughout their ranges, including LCA and surrounding areas. Mapping sampling localities from 66% of studies reveals that phylogeographers have sampled LCA widely, producing extensive geographical coverage (Fig. 6A). LCA's complex earth history and ecological heterogeneity predicts that generally complex patterns of phylogeographic congruence and incongruence are likely to be recovered within and among lineages. Consistent with this prediction, LCA taxa showcase a diversity of phylogeographical patterns rivaling the complexity of LCA landscapes (Fig. 6; Tables S1–S3). Apart from landscape diversity and history, this probably also reflects the diverse distributions of species sampled to date (Table S1).

(1) Phylogeographic structuring is common within LCA taxa and reveals cryptic biodiversity

Most (63.4%) LCA lineages show genetic structuring in the form of phylogeographic breaks (phylogenetic splits between mostly distinct geographical lineages). In total, LCA lineages support 31 major phylogeographic breaks, shown in Fig. 6B (with further details in Tables S1–S3), most of which have been recovered from mitochondrial DNA markers (see online Appendix S2). Apparently, long-term mechanisms of genetic isolation (physical, reproductive, etc.) have been at play in many species. Although phylogeographic studies have recovered clues to some species GABI histories (e.g. freshwater fishes, Bermingham & Martin, 1998; Reeves & Bermingham, 2006), these results suggest that isthmian

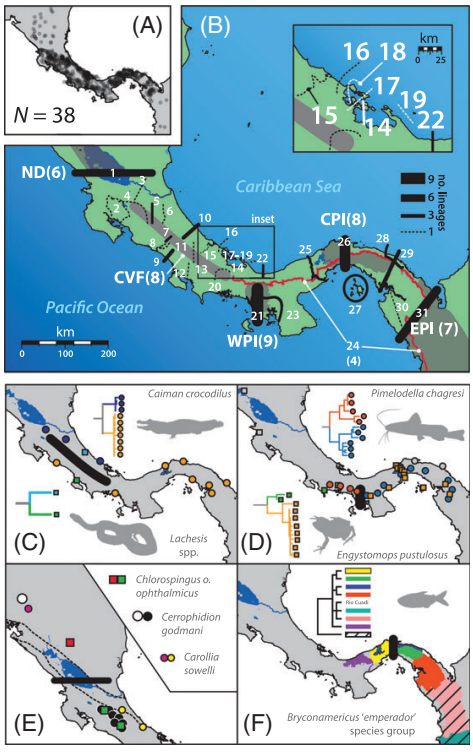

Fig. 6. Legend on next page.

environments, not just dispersal into or out of the region, have contributed significantly to within-LCA diversification. Results also underscore the point that LCA is more than a mere biogeographic crossroads between continents, but harbours unique genetic endemism (Wang *et al.*, 2008). Data from a synthesis of phylogeographical patterns from eastern North America show that ∼78% of organisms investigated exhibited clear phylogeographical structuring based mainly on organellar DNA markers (Soltis *et al.*, 2006; excluding taxa not analogous to those considered herein, e.g. marine organisms). Thus, LCA organisms exhibit just slightly less prevalent phylogeographical structuring than that of a much older and larger, truly continental area.

In a similar vein, LCA provides a classic, subcontinental-scale showcase of phylogeography's ability to make discoveries that might otherwise go unnoticed, principally cryptic lineage divergences. Overall, approximately 197 genetically distinct evolutionary lineages are recovered within 94 nominal taxa sampled to date, amounting to, on average, 2.1 lineages per taxon (Table S1). These patterns vary widely among taxa and have various biogeographical and taxonomic implications; however, amphibians and freshwater fishes harbour particularly exceptional cryptic diversity that appears informative for testing geological hypotheses. For example, a study of four nominal freshwater fish species uncovered ∼12–22 novel lineages (∼3–5 cryptic lineages/taxon; Bermingham & Martin, 1998). Their comparative phylogeographical inferences led Bermingham & Martin (1998) to propose a new model (B/M model) of landscape evolution for the LCA region. Also, one poison-dart frog species, *Oophaga pumilio*, apparently contains from several to up to 18–19 unique genetic lineages, depending how you count them (Wang & Shaffer, 2008). Venegas-Anaya *et al.* (2008) upheld genetic and geographical distinctiveness of *Caiman* crocodile subspecies but also discovered a novel cryptic lineage representing a new taxon. Several other studies have also identified cryptic lineages representing putative new species, or operational taxonomic unit-level biodiversity (Table S1; e.g. Martin & Bermingham, 2000; Reeves & Bermingham, 2006; Jones & Johnson, 2009; Vázquez-Miranda, Navarro-Sigüenza & Omland, 2009).

Fig. 6. Sampling and phylogeographical breaks emerging from lower Central American (LCA) phylogeography studies. (A) Map summarising geographical coverage of sampling localities, which were available from most (66%) studies in this review (see online Appendix S1). (B) Map of phylogeographical breaks (different coloured lines) discussed in the text, with abbreviations given for the five major breaks. Breaks (numbers of nominal taxa/lineages split across each break): 1, ND = Nicaraguan depression ($N = 6$); 2, BEB = Rio Bebedero ($N = 1$); 3, SJ1 = San Juan break 1 ($N = 1$); 4, SJ2 = San Juan break 2 ($N = 1$); 5, CC = Central Cordillera ($N = 2$); 6, SJ3 = San Juan break 3 ($N = 1$); 7, TCMF1 Talamanca Cordillera montane forest break 1 ($N = 2$); 8, SAV = Rio Savegre break ($N = 1$); 9, FILA = Fila Costeña ($N = 3$); 10, L1 = Limón ($N = 3$); 11, CVF = Chorotega volcanic front ($N = 8$); 12, PB = Piedras Blancas ($N = 1$); 13, TCMF2 = Talamanca Cordillera montane forest break 2 ($N = 1$); 14, POPA = Popa Island–mainland ($N = 2$); 15, SIXA = Sixaola–Changuinola ($N = 1$); 16, BDT1 = Bocas del Toro break 1 ($N = 1$); 17, BDT2 = Bocas del Toro break 2 ($N = 1$); 18, BDT3 = Bocas del Toro break 3 ($N = 1$); 19, ESCU = Escudo de Veraguas Island–mainland ($N = 1$); 20, BARU = Barú volcano ($N = 1$); 21, WPI = western Panama isthmus ($N = 9$); 22, MOSQ = Mosquito Gulf ($N = 2$); 23, AZUE = Azuero peninsula ($N = 2$); 24, PNSA = Panama–northern South America continental divide ($N = 4$); 25, VALLE = El Valle volcano ($N = 2$); 26, CPI = central Panama isthmus ($N = 8$); 27, PERL = Las Perlas Islands ($N = 2$); 28, CHIC = Rio Playón Chico basin ($N = 2$); 29, BT = Bayano–Tuira ($N = 3$); 30, SAPO = Sapo range ($N = 1$); 31, EPI = eastern Panama isthmus ($N = 7$). (C–F) Typical spatial-genetic splits recovered within different species, each contributing to major multi-taxon phylogeographic breaks shown in (B); examples show the (C) CVF break in *Caiman crocodilus* crocodiles (Venegas-Anaya *et al.*, 2008) and *Lachesis* spp. bushmasters (Zamudio & Greene, 1997); (D) WPI break in *Pimelodella chagresi* catfishes (e.g. Bermingham & Martin, 1998) and *Engystomops pustulosus* frogs (Weigt *et al.*, 2005); (E) ND break in *Chlorospingus ophthalmicus* birds (e.g. Weir *et al.*, 2008), *Cerrophidion godmani* pit-vipers (e.g. Castoe *et al.*, 2009), and *Carollia sowelli* bats (Hoffmann & Baker, 2003); and (F) CPI break in the *Bryconamericus* '*emperador*' species group (Reeves & Bermingham, 2006).

(2) Multi-taxon spatial structuring suggests general evolutionary patterns and highlights importance of regional processes shaping diversification

'To do science is to search for general patterns, not simply to accumulate facts . . . ' (MacArthur, 1972, p. 1).

Many ($N = 17$; 54.8%) of the phylogeographical breaks recovered in LCA to date are spatially congruent across multiple taxa (Fig. 6B; Table S3), over small to regional scales. This supports the existence of generalised evolutionary patterns in LCA. Furthermore, congruence among taxonomically and ecologically divergent but codistributed lineages indicates historical associations of genotypes possibly due to shared biogeographic history in the same local communities (Arbogast & Kenagy, 2001; Zink, 2002). Notably, multi-taxon breaks are recovered across the Chorotega volcanic front (CVF, Fig. 6B, C); western Panamanian Isthmus (WPI, Fig. 6B, D); Nicaraguan depression (ND, Fig. 6B, E; representing LCA–NCA divergences); central Panama at or east of the Panama Canal Zone (CPI, Fig. 6B, F); and eastern Panama (EPI, Fig. 6B; representing LCA–South America divergences directly within or in the vicinity of the Darién isthmus). The striking correlation between these breaks and physiography, particularly major geographical barriers (Section III), suggests that regional processes (e.g. orogeny, oceanic terrane accretion and uplift, and other geological processes) have played a major role in shaping intraspecific diversification across LCA biodiversity, by promoting and maintaining long-term zoogeographical barriers.

(3) Low phylogeographic structuring and incongruent patterns highlight roles of chance, ecological differences, and local processes shaping LCA communities

In contrast to the above patterns, a considerable proportion of LCA lineages ($N = 34$, or 36.6%; representing 22 studies) exhibit zero to limited phylogeographic structure (Table S1). Around a third of these studies conducted inadequate spatial or numerical sampling (e.g. coarse sampling grain or density) of widely distributed taxa (e.g. *Uroderma bilobatum* bats; Hoffmann, Owen & Baker, 2003). In such cases, determining the degree to which the observed lack of structuring is attributable to actual evolutionary genetic patterns (e.g. 'phylogeographic category V'; Avise *et al.*, 1987) *versus* sampling artifacts (e.g. inadequate phylogenetic signal or poor marker selection leading to unresolved phylogenetic topologies) is problematic. However, species may exhibit low intraspecific structuring due to high gene flow (e.g. migration-drift non-equilibrium), hybridisation, large historical effective population sizes (N_e; e.g. making equilibrium and complete lineage sorting hard to attain), or recent colonisation (e.g. founder events) combined with low mutation rates (Avise, 2000; Wakeley, 2002). Processes contributing to lack of phylogeographical structuring within these species are therefore likely to vary and may reflect ecological differences.

At a comparative level, phylogeographic incongruence can arise from historical differences among species at the same parameters surrounding the evolutionary circumstances of zero–low intraspecific phylogeographic structuring discussed above. However, in comparative phylogeography, spatial incongruence indicates potentially independent responses of species to the series of geologic and palaeoclimatic changes that have occurred within an area (Avise, 2000; Arbogast & Kenagy, 2001; Zink, 2002). With that said, cases of low phylogeographic structuring in LCA (Table S1) are incongruent relative to the multi-taxon patterns described above, suggesting that species may have experienced different responses to historical events within shared distributions. Phylogeographic incongruence is commonly inferred from comparative LCA studies. Phylogeographical comparisons of codistributed LCA bat (Hoffmann & Baker, 2003; Martins *et al.*, 2009), frog (Crawford *et al.*, 2007), snake (Castoe *et al.*, 2009) and freshwater fish species (Bermingham & Martin, 1998; Reeves & Bermingham, 2006; J. C. Bagley & J. B. Johnson, unpublished data)—in many cases, focal taxa that combined range throughout much or all of LCA or Central America—reveal idiosyncratic patterns of area relationships and gene flow patterns, up to regional scales. This also supports a potential lack of shared biogeographic history. Common processes may not have influenced diversification of some ecological communities at broader spatial scales, leading to different historical responses by habitat, within and among taxonomic groups. However, chance, including stochastic differences in the timing of LCA colonisation among lineages, might partly account for this. Differential dispersals into LCA could reflect the influence of extrinsic ecological factors (e.g. presence of available suitable habitat in the target area) or intrinsic ecological differences of lineages (e.g. dispersal abilities) in shaping biodiversity distributions (Bermingham & Martin, 1998; Arbogast & Kenagy, 2001; Zink, 2002).

(4) Deep phylogenetic subdivisions within plants appear rare in LCA: truth or illusion?

Taxonomic sampling biases have favoured animals over plant taxa at a 9:1 ratio, precluding robust comparisons between plant and animal phylogeographies (see online Appendix S2). Strikingly, however, available data reveal that no plant species possess deep phylogeographic breaks or contribute to multi-taxon breaks within LCA. Yet is this a representative portrait of the evolutionary history of LCA plant species, or an illusion? It would be tempting to conclude from these data that LCA plant species share a congruent *lack* of phylogeographic structure, suggesting that they have been largely unaffected by historical barriers and processes shaping genetic isolation in animal taxa. However, the observed lack of phylogeographical structure in plants more likely reflects a combination of (*i*) low genetic marker resolution and (*ii*) higher relative dispersal potential of plant species studied to date, facilitated by intrinsic and extrinsic ecological factors promoting dispersal to and establishment in new areas. For example, most plant studies have relied on chloroplast DNA (cpDNA), which may evolve 10–100 times more slowly than animal mtDNA, limiting the ability of this marker to detect phylogeographic structure (see Avise, 2000, and references therein). Moreover, regarding dispersal, most LCA plants studied to date have been large tree species that by their nature are more dispersive than other plant types, a situation which lends itself to less genetic structuring in these species (Petit & Hampe, 2006). Despite such potential biases, previous studies have concluded that Neotropical plant species are more dispersal-prone than animals based on fossil pollen and molecular phylogenetic data showing that multiple plant lineages reached LCA before many vertebrate GABI participants (Raven & Axelrod, 1974; Cody *et al.*, 2010). Indeed, over-water dispersal apparently has played a more important role in shaping LCA plant distributions than anticipated (Cody *et al.*, 2010). Taking one species as an example, despite water-intolerant seeds, phylogeographical analyses demonstrate that *Symphonia globulifera* trees reached LCA from South America before Late Pliocene isthmus completion, *via* long-distance oceanic dispersal (Dick *et al.*, 2003; Dick & Heuertz, 2008). However, while *S. globulifera* demonstrates that over-water dispersal is a mechanism that has operated during the assembly of the LCA flora, it is important to note that populations are differentiated based on genetic data from DNA sequences and microsatellites (Dick & Heuertz, 2008); therefore, despite containing no deep phylogeographical structuring, this species apparently experiences dispersal limitation after it colonises new areas. This example illustrates the importance of factoring in the peculiarities of plant species genetics and ecologies when conducting phylogeographic analyses and highlights how

incorporating more rapidly evolving markers (e.g. nDNA or cpDNA microsatellites) could provide better avenues to geographical inference in future studies of LCA plants.

(5) Other dispersal-demography connections

The previous sections highlight a strength of phylogeography—its ability to link ecology and demography to broader macroevolutionary and biogeographical patterns (Avise *et al.*, 1987). Limited phylogeographical structuring is expected in superior-dispersing and -colonising species in the absence of strong physical barriers to dispersal/gene flow, whereas progressively monophyletic gene tree lineages are expected across strong environmental gene flow barriers through time (Avise *et al.*, 1987; Avise, 2000). Phylogeographic structure should therefore correlate inversely with behavioural preference and physiological capacity for dispersal (e.g. high rates, over large distances). Aside from some of the plant patterns (but see caveats and discussion in Section VII.4), patterns from other taxa also support this prediction. For example, widely distributed *Atta* leafcutter ants (Solomon *et al.*, 2008) and bees capable of long-distance dispersal (Dick *et al.*, 2004) show limited genetic structuring. Conversely, dispersal-limited montane salamanders and frogs show substantial Tertiary–Quaternary diversification (García-París *et al.*, 2000; Streicher *et al.*, 2009). Congruent with expectations, livebearing 'secondary' freshwater fishes with presumed salt tolerance (Myers, 1938) lack isolation in Atlantic-coast drainage basins (J.C. Bagley & J.B. Johnson, unpublished data). Unexpectedly, however, salt-intolerant 'primary' freshwater fishes, considered to have relatively lower dispersal potential (Myers, 1938), display evidence for rapid deployment across the landscape, recent clades, and sometimes no phylogeographical signal at all (e.g. *Cyphocharax magdalenae*, Reeves & Bermingham, 2006).

(6) Filter barriers: biogeographic province boundaries and other features are permeable barriers to dispersal

Whereas Wallace (1876) thought that the LCA isthmus allowed relatively unimpeded dispersal through the region (Section VI.1), Simpson (1950) viewed the LCA isthmus as a historical 'filter barrier' reducing, without eliminating, inland movements of species. Simpson's view has since become widely accepted, and it is supported by phylogeographic patterns, particularly across biogeographic province boundaries (e.g. Fig. 3). Province boundaries are thought to explain species turnover and reflect localised vicariant barriers historically limiting gene flow and species distributions (Avise *et al.*, 1987; Ronquist, 1997). Therefore, historical processes should have promoted LCA lineage divergence at province boundaries (Lee & Johnson, 2009); however, evidence for this prediction is mixed. On the one hand, multi-taxon breaks span province boundaries (Fig. 3), e.g. the Chorotega volcanic front (Fig. 6B, C; Table S3). On the other hand, gene flow, or a mixed phylogeographical structuring, has been inferred across this and other province boundaries. Freshwater fish communities have mixed genetically across the Chorotega front (Jones & Johnson, 2009; Lee & Johnson, 2009) and the eastern-central Panamanian Isthmus (Reeves & Bermingham, 2006), aided by headwater river capture events. Gene flow has probably also occurred across the West Panama portion of the CVF in pseudoscorpions (Zeh, Zeh & Bonilla, 2003b*b*). Additionally, bird data indicate that *Chlorospingus ophthalmicus* bush-tanagers possibly exchanged genes across montane areas of bird endemism (Guatemalan *versus* Talamancan; Weir, 2009) separated by the Nicaraguan depression (Bonaccorso *et al.*, 2008), while *Glyphorynchus spirurus* woodcreepers apparently experienced long-distance dispersal/gene flow across the Colombian Andes (Marks, Hackett & Capparella, 2002), or EPI break area (Fig. 6B). Assuming these cases represent actual gene-flow events (not incomplete lineage sorting), then (*i*) patterns of LCA vicariance, dispersal and gene flow vary not only across spatial scales (Smith & Bermingham, 2005), but also according to physiographic barrier (province boundary) considered; (*ii*) vicariant barriers have had mixed impacts on community formation and species distributions; and (*iii*) the above examples support the interpretation that the corresponding province boundaries represent filter barriers. Building on the third of these points, a growing list of studies reveals phylogeographic breaks broadly correlated to Bocas del Toro, Panama (BDT; Fig. 2A) environments, and this indicates that this area presented a historically important filter barrier in LCA. East of BDT, phylogenetic splits across the Caribbean Gulf of Mosquitoes break (MOSQ; 22 in Fig. 6B; Tables S1, S3) are supported by mtDNA lineages of *Caiman crocodilus* crocodiles (Venegas-Anaya *et al.*, 2008) and *Pristimantis ridens* frogs (Wang *et al.*, 2008). Phylogeographic breaks correlated with the BDT region are also recovered in mtDNA variation in catfishes (Perdices *et al.*, 2002) and frogs (Crawford *et al.*, 2007; Robertson, Duryea & Zamudio, 2009) and differentiated nuclear ribosomal spacer sequences in trees (Dick & Heuertz, 2008). Interestingly, these patterns have arisen despite a lack of obvious geographical barriers (e.g. contiguous Caribbean wet and mangrove forests dominate the coastline, Fig. 2B). This illustrates the ability of phylogeography to derive and test new biogeographical explanations as required when phylogeographic breaks fit no known historical events (Riddle, 1996; Gascon *et al.*, 2000). This area is low in elevation and forms a young, contiguous Limón–Bocas del Toro coast exposing Neogene sediment and rock formations (Marshall, 2007). The BDT embayment and mainland were also partly inundated after a nearby 1991 earthquake (Marshall *et al.*, 2003), and our sea level model (Fig. 5B) suggests that this area could have been extensively affected by Pleistocene seas. Thus, isolation caused by marine incursions presents an alternative to tectonic uplift (Venegas-Anaya *et al.*, 2008) and restricted coastal dispersal corridors (Crawford *et al.*, 2007) as a potential explanation for this filter barrier.

(7) Phylogeography and island biogeography theory in LCA

Traditional island biogeography theory (Section VI.3) makes several predictions germane to LCA, including that (*i*) islands have likely been colonised from nearby mainland source-pools; (*ii*) due to rapid turnover, island species should be young; and (*iii*) larger, more distant islands should harbour older, more genetically divergent lineages due to lower extinction (remnants of former radiations persist). From coastal geology (Section III; Coates & Obando, 1996), bathymetry (e.g. Fig. 4G), and sea level dynamics (Figs 4G and 5A; Fleming *et al.*, 1998), we also predict (*iv*) land-bridge islands located over the LCA continental shelf originally had species similar to mainland communities (due to colonisation during low seas, when exposed shelf habitat created mainland connections; e.g. as shown in Fig. 4G for the Bocas del Toro archipelago) whose populations recently became isolated (during high seas) from mainland populations. Phylogeographic datasets from several LCA taxa support the above predictions. For example, poison-dart frog phylogeography does not fit a vicariance model (sequential mainland–island isolation events) at Bocas del Toro; these frogs apparently originated from nearby Costa Rican mainland frogs and achieved their present distributions through multiple mainland–island, island–island, and island–mainland dispersal events (Wang & Shaffer, 2008). These poison-dart frog lineages are also predictably shallow, or recently isolated (see below). However, island biogeography prediction *ii* above is rejected: extinction has apparently not been strong enough to remove many novel insular lineages, e.g. the poison-dart frog radiation persists and various other frogs, birds, and freshwater fishes exhibit mainland– or island–island phylogeographical breaks, e.g. BDT1–3, ESCU, PERL (Fig. 6B; Tables S1–S3). The data suggest that different taxa/lineages have colonised LCA islands at vastly different times and persisted, forming endemic lineages through genetic drift. The most striking example is the ancient Las Perlas Islands lineage of salt-tolerant *Synbranchus marmoratus* swamp-eels, which possibly colonised the islands > 50 Ma *via* oceanic dispersal (Perdices, Doadrio & Bermingham, 2005; PERL break, Fig. 6B). These data corroborate the prediction from 'new' island biogeography theory that island biotas are typified by persistence, rather than extinction (Heaney, 2007).

(8) Temporal patterns suggest a mainly Plio–Pleistocene timeframe for biotic diversification in LCA

The largely Late Neogene–recent timing of landscape evolution (Section III) and interamerican biotic exchange (e.g. GABI; Section VI.2) in LCA predicts that biotic assembly and diversification in LCA, and between LCA clades and sister clades in outlying continental areas, should coincide with this interval. A meta-analysis (see online Appendix S1) of divergence times inferred from

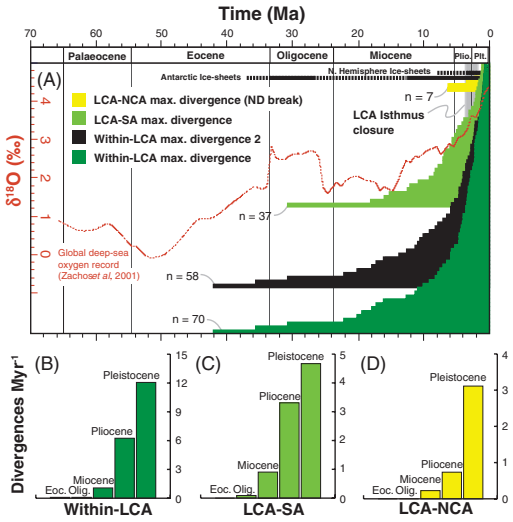

Fig. 7. (A) Time ranges of initial within-lower Central America (LCA), LCA–Nuclear Central America (NCA), and LCA–South America (SA) lineage divergence events estimated for LCA lineages studied to date. Thin horizontal bars span the time in millions of years ago (Ma) since the initial speciation or divergence event for each clade to the present and are plotted in increasing chronological order by geographical class (colours). Within-LCA diversification time ranges are shown for clades confined to the study area (within-LCA max. divergence), and for splits including clades containing samples from outlying areas (within-LCA max. divergence 2); see text, and Appendix S1 and Table S2 for the raw data. Time-range data are presented over major earth history parameters/events: dashed red line, mean deep-sea oxygen isotope ($\delta^{18}O$) curve, a temperature proxy mostly controlled by changes in continental ice-sheet volume (thick black horizontal bars, permanent ice sheets; thick dashes, times with partial or melting ice sheets; modified from Zachos *et al.*, 2001); vertical grey bar, timeframe of LCA isthmus closure (Section III). (B–D) Exponential increases in speciation/lineage divergence rates, or possibly declining extinction rates, within LCA (by geological epoch) over Eocene–Pleistocene, inferred based on divergence time distributions in (A) (with corresponding colours).

time-calibrated molecular phylogenetic divergences from studies reviewed herein supports this prediction. Divergence dates for stem and crown nodes recovered in LCA phylogeography studies (summarised in Table S2) exhibit an over 40 Myr range, from 42.1 Ma (max. crown age, *Craugastor podiciferus* frogs; Streicher *et al.*, 2009) to a mere 235 ka (LCA–SA divergence within *Anopheles albimanus* mosquitoes; Loaiza *et al.*, 2010*a*). However, most lineage divergence can be constrained to less than 20 Ma, and a broadly exponential pattern of lineage diversification since ~14 Ma is evident in the time ranges of estimated initial lineage diversification events (Fig. 7A). Most lineage divergences, including 69% of within-LCA divergences and 57% of LCA–SA divergences, are constrained to Pliocene–recent, with maximum divergence dates ranging no later than Early

Pliocene (5.3 Ma). These results imply potentially higher speciation rates since the Pliocene, a suspicion corroborated by rate calculations. Crudely estimating lineage divergence rates, which presumably reflect speciation rates, shows a likewise exponential pattern of increase across estimates over Oligocene–Pleistocene (Fig. 7B–D). This might reflect marginally higher divergence rates due to redundant samples in our time-range plots but is unlikely to be due to variation in sampling, given the large number of lineages we sampled and that we observed the highest speciation rates more recently (whereas unsampled lineages are expected to cause declines in recent speciation rates). Another potential explanation for the pattern of higher speciation rates since the Pliocene is the likely extinction of older lineages. Nonetheless, the observed period of exponentially increasing diversification rates is synchronous with or just follows Early Pliocene high-sea stands that, combined with East Panama microplate vertical positioning, maintained a partially drowned Panama Isthmus ~7 to 3.7 Ma (Duque-Caro, 1990; Coates *et al.*, 2004). Subsequent divergences in our compilation overlap global cooling, LCA land bridge emergence, GABI exchanges, and Quaternary intensification of glacio-eustatic cycles. Mirroring similar meta-analyses based on recent molecular phylogenetic evidence from Neotropical taxa, these results suggest that Neogene–recent geological and palaeogeographic events *and* Quaternary glacio-eustatic cycles are likely to have been important drivers of biotic diversification in LCA and surrounding Neotropical areas (Cody *et al.*, 2010; Rull, 2011, and references therein).

VIII. CHALLENGES: IMPROVING METHODS AND INFERENCES IN LCA PHYLOGEOGRAPHY

The field of LCA phylogeography must face several challenges to ensure continued progress and improved historical biogeographical inferences. First, more phylogeography studies using better data are needed to elucidate further the historical origins, assembly, and diversification of LCA biotas. Conducting phylogeographical analyses of LCA organisms can be difficult due to landscape complexity and logistical issues. Still, meeting this challenge through amassing more single-species datasets will enhance our knowledge of the processes underlying genetic variation within LCA species. In turn, increasing the number of codistributed species datasets will permit expanded comparative analyses needed to test further the generality of the emerging phylogeographical patterns herein. However, future studies should proactively work to improve finer-scale sampling and counteract existing geographical and taxonomic sampling biases. In particular, more studies sampling plants and taxa with premontane to montane distributions are needed (Section VII.4; online Appendix S2), e.g. in the under-sampled Talamanca Cordillera (Fig. 6A). Indeed, while few comparative analyses of highland taxa have been conducted (e.g. Castoe *et al.*, 2009), filling this gap will likely continue unveiling distinct lineages that are new to science along with insights into how highland diversity is maintained (e.g. García-París *et al.*, 2000; Streicher *et al.*, 2009). On the related issue of gene sampling, workers have relied principally on single-locus analyses of mtDNA and cpDNA (see online Appendix S2); however, the maturation of the field will require development of nuclear phylogeography perspectives. Although mtDNA is highly informative and a robust indicator of population history and species limits (Avise, 2000; Zink & Barrowclough, 2008), single gene trees have both historical and random components and can be discordant topologically with one another as well as 'true' species/population trees, e.g. due to incomplete lineage sorting (Maddison, 1997). Thus, sequence data from multiple unlinked loci are needed to overcome noisy historical gene tree signals and correctly infer phylogenetic relationships in a species tree framework (e.g. Liu & Pearl, 2007; Kubatko *et al.*, 2009), and multi-locus data also provide a robust framework for accurately estimating divergence times and demographic parameters (e.g. migration rates) to infer phylogeography (Edwards & Beerli, 2000; Hey & Machado, 2003). To accelerate nDNA marker development, we recommend new methods identifying exon-primed intron-crossing markers and anonymous nuclear loci based on genome-enabled approaches (reviewed by Thomson, Wang & Johnson, 2010). However, we note that comparative phylogeography (Arbogast & Kenagy, 2001; Riddle *et al.*, 2008) accounts for the gene tree variance problem by testing for replicated population divergences across taxa, indicating common historical events in a region (though inferences are most reliable when congruence is demonstrated across many taxa).

Developing a more hypothesis-driven and statistically rigorous research program capitalising on novel advances in statistical population genetics and geospatial analysis presents a second, arguably more formidable challenge. Building 'just-so' stories or *ad hoc* explanations from observed genetic patterns has been commonplace in LCA studies, and this is not all bad: exploratory analyses cover 'scenario space' and often yield unexpected discoveries (Garrick *et al.*, 2010) such as the cryptic divergences discussed above. And, aside from our review, this is supported by previous syntheses of phylogeographical data from Amazonia (e.g. Patton, Da Silva & Malcolm, 1994) and the southeastern US Coastal Plain (Avise *et al.*, 1987; Avise, 2000) that also highlight cryptic genetic breaks useful for interpreting regional historical biogeography. However, reliance on pattern discovery and matching is nonstatistical and embodies the major criticisms of phylogeography (references in Edwards & Beerli, 2000; Arbogast & Kenagy, 2001; Knowles & Maddison, 2002; Posadas *et al.*, 2006; Hickerson *et al.*, 2010). By contrast, 'statistical phylogeography' (Knowles & Maddison, 2002; Knowles, 2009; Garrick *et al.*, 2010) provides more objective and statistically rigorous methods to infer demographic history while taking geography and stochastic population genetic processes into account. Drawing on coalescent theory (Wakeley, 2002) and probabilistic simulations, these sophisticated model-based methods can estimate population parameters (e.g. Kuhner, 2009) and statistically discriminate

among *a priori* demographic and biogeographic models while accounting for coalescent variance (e.g. Knowles & Carstens, 2007). Statistical phylogeographical approaches naturally lend themselves to developing and testing *a priori* hypotheses (e.g. modeling explicit historical scenarios, with varying population sizes, divergence times, and migration rates), thus their application will be essential for developing a more hypothesis-driven focus. And this will also benefit comparative analyses: as larger comparative datasets are assembled, approximate Bayesian computation methods will permit tests of co-dispersal, co-vicariance, or other patterns of shared demographic histories (*versus* multiple divergences) using highly parameterised statistical phylogeographical models (Hickerson *et al.*, 2006*a*,*b*, 2007, 2010; Knowles, 2009; Huang *et al.*, 2011). Perspectives from other underutilised tools such as geospatial modelling (e.g. ecological niche models; Kozak *et al.*, 2008) and phyloclimatic modelling (Yesson & Culham, 2006) should also be developed and used to enhance phylogeographical inferences. For example, geospatial-modelling applications can be used to predict the impacts of historical environments on species palaeodistributions (e.g. range dynamics), and this information can subsequently aid the generation and testing of *a priori* models/hypotheses using statistical phylogeography (e.g. Kidd & Ritchie, 2006; Richards, Carstens & Knowles, 2007; Chan, Brown & Yoder, 2011).

A third, more general future challenge for LCA phylogeography will be integrating insights into organismal evolutionary history with theory and methods from other disciplines including ecology, other historical biogeographical techniques (Posadas *et al.*, 2006), comparative phylogenetics, palaeontology and the geosciences. The general trend of fragmentation between historical biogeography and ecology, and among the various sub-fields of historical biogeography, is widely recognised (Wiens & Donoghue, 2004; Ebach & Tangney, 2007; Sanmartín, 2010). However, methodological integration with other fields will be increasingly critical as LCA workers address more interdisciplinary questions, including those surrounding spatial-genetic analyses of selection, adaptation, functional trait evolution, community assembly and more. Virtually no studies link LCA phylogeography and adaptation; yet which came first, environmental adaptations or phylogeographical lineage divergences? Also, which lineages more likely diversified during Pleistocene climatic fluctuations—physiologically plastic ones or those with higher genotypic diversity (limited plastic adaptations)? In addition, contributions of adaptive and phenotypic differences to phylogeographical patterning are poorly explored; however, if addressed from a genomic perspective, such questions may also yield insight into the origins of adaptive and functional diversity.

IX. PHYLOGEOGRAPHY AND LCA CONSERVATION

Phylogeography stands to make several critical contributions to conservation in lower Central America, in light of the region's threatened status. First, as illustrated in Section VII.1, phylogeography is poised to provide the unique service of uncovering cryptic lineage divergence and speciation patterns in LCA. This can aid conservation efforts in several ways. Knowledge of cryptic speciation patterns aids identifying and refining species limits and therefore regional species diversity patterns, which may highlight the need to alter conservation plans (e.g. if today's most species-rich areas are not the hottest regional hotspots tomorrow). The challenges for, and impacts on, systematics and taxonomy are obvious: it will be critical to formally describe novel biodiversity as it is revealed, and to determine its place in the tree of life. Also, conservation strategies typically prioritise species with various extinction-risk correlates, including small geographic ranges, low abundances, and specialised life histories or feeding phenotypes; however, phylogeographers often discover cryptic lineages by studying common, widespread species. Detecting natural species with smaller ranges contained within widespread, non-target 'species' may identify 'new' species at greater risk of extinction, thus warranting conservation resources. Second, where comparative phylogeography uncovers replicated cryptic divergences, areas of diversity and endemism can be better identified—areas where processes generating biodiversity have acted and presumably still are acting across multiple taxa to maintain biodiversity (Moritz, 2002). Sub-areas spanning environmental gradients within those areas can then be selected to preserve adaptive differences (Moritz, 2002), and highly threatened sub-areas can be pinpointed. If multi-taxon breaks are found to occur over distinct geographical barriers that limit gene flow between areas, and one area or the other is most susceptible to climate change, this may determine primacy of areas due to increased likelihood of future extirpation or extinctions. Thirdly, expounding on this latter theme, it will be important for LCA conservation that phylogeographical inferences are integrated with predictive niche-based models of the past, present, and future distributions of lineages. As illustrated by work on Brazil's Atlantic Forest by Ana Carnaval, Craig Moritz, and colleagues identifying areas of high genetic diversity and environmental stability, integrative phylogeographical approaches can highlight areas that have contributed to local endemism and that are most likely to withstand climate change over the coming centuries (Carnaval & Moritz, 2008; Carnaval *et al.*, 2009). A similar approach combining geographic information systems (GIS)-based modelling and statistical phylogeography seems to hold promise for predicting, and explaining, patterns of LCA biodiversity as well (Chan *et al.*, 2011). Undoubtedly, the unique physiography and complex earth history of lower Central America will continue to provide a fascinating backdrop for addressing integrative questions relevant

to Neotropical biogeography and diversification, through phylogeography. However, the above steps may help us not only understand, but also conserve, the patterns and processes of diversification in this unique Neotropical region.

X. CONCLUSIONS

(1) Despite covering only ∼0.09% of earth's land area, lower Central America is among the most physically and biologically complex areas worldwide. LCA boasts distinct landform (Fig. 2A) and biotic assemblages (Fig. 3), each reflecting unique histories of landscape *versus* organic evolution. Indeed, levels of physiographic variation in LCA are usually only attained at whole-continental scales (Marshall *et al.*, 2003; Marshall, 2007).

(2) Throughout the past 100 Myr, LCA has experienced diverse geographical changes (particularly since the Miocene) that altered probabilities of biogeographical processes (e.g. dispersal) and earth surface processes (orogeny and sea-level and climatic fluctuations, plus major earth history events e.g. formation of the Isthmus of Panama). These undoubtedly figured prominently in structuring modern Neotropical biogeography patterns. Combined with the observation that LCA landscapes are mostly geologically young (many major physiographic features formed Neogene–recent), which should limit the complexity of biogeographical patterns, this makes LCA particularly attractive for phylogeography.

(3) Whereas the extent of Plio-Pleistocene sea-level rise and fall remains debated, elevation data and consensus from sea-level curves suggest that sea-level change may have significantly impacted LCA biodiversity over Neogene–recent. However, phylogeographical data have rarely been used explicitly to address sea-level events as drivers of LCA diversification (but see Jones & Johnson, 2009). Whether and to what extent marine incursions have influenced broad-to-fine-scale diversification of LCA biota should be tested further, and geologically correlated.

(4) As part of the Mesoamerica biodiversity hotspot, LCA is of great conservation significance; however, despite landmark-scale conservation efforts, its biodiversity remains highly threatened by anthropogenic factors. More conservation efforts are essential to ensure the persistence of LCA biota; certain habitats including fresh waters warrant more recognition and conservation resources.

(5) LCA presents a rare intercontinental and interoceanic land bridge that facilitated massive interamerican exchanges of species since ∼3 Ma, the 'Great American Biotic Interchange' (Stehli & Webb, 1985). GABI histories of taxa have been a primary focus of historical biogeographical research in LCA; however, these and other key biogeographical hypotheses (scenarios) proposed for Central America, e.g. Savage's (2002) Pleistocene model, warrant more testing. Overemphasising species GABI histories has left finer-scale patterns of colonisation and post-colonisation diversification lesser known, and more phylogeographical inquiry is essential to overcome this knowledge gap.

(6) Most LCA taxa show phylogeographical breaks indicating that LCA environments apparently have contributed to within-LCA diversification. Therefore, LCA is more than a mere biogeographic crossroads between continents, but harbours unique genetic endemism.

(7) LCA provides a classic showcase of phylogeography's ability to make otherwise unnoticed discoveries, especially cryptic lineage divergences. Amphibians and freshwater fishes harbour particularly exceptional cryptic diversity and appear informative for testing geological hypotheses. Given the prevalence of cryptic divergences, LCA would be ideal for studying population divergence/speciation by comparing multiple diverged lineage pairs; this could provide robust tests for 'suture zones' (Remington, 1968), which have not been rigorously evaluated.

(8) Comparative phylogeographical patterns reveal at least 17 multi-taxon phylogeographical breaks in LCA, highlighting the importance of regional processes in shaping genetic diversity and composition of LCA biotic communities. Yet more work is needed (*i*) to test the evolutionary generality of these breaks across biodiversity (e.g. through comparisons with new datasets from other species), and (*ii*) to evaluate temporal congruence (e.g. simultaneous divergence) *versus* other patterns, e.g. pseudocongruence (identical area relationships caused by different underlying events; Cunningham & Collins, 1994), and to test more rigorously for underlying causal factors.

(9) A surprisingly large proportion of LCA taxa exhibit little or no phylogeographical structuring, particularly plants. Although such patterns could reflect a diversity of alternative mechanisms, they indicate a role for chance, ecological differences (e.g. dispersal potential) and local processes (e.g. ecological interactions) in shaping regional patterns of biotic assembly and diversification. However, using more rapidly evolving genetic markers will likely recover greater levels of genetic structure in LCA plant species.

(10) Phylogeographic data agree with the long-held view that LCA, and areas of the subcontinent (e.g. coastal headlands, Bocas del Toro), presents biological filter barriers (e.g. Simpson, 1950; Savage, 1966; Crawford et al., 2007). More study will be instructive, however, in determining (*i*) which filter barriers have figured most prominently in shaping LCA biogeographical patterns; (*ii*) why LCA environments filter dispersal/gene flow in some taxa, but not close relatives; and (*iii*) whether filter barriers are also areas of secondary contact (creating suture zones, as per above).

(11) A new metamorphosis of theory is brewing in the field of island biogeography, resulting in the setting aside of classic theories (ETIB, vicariance biogeography) in explaining patterns of island life (e.g. Heaney, 2007;

Lomolino *et al.*, 2010). LCA contains land-bridge islands, and inland habitat 'islands' (e.g. Talamancan sky-islands), that can be used to test island biogeography predictions. Phylogeographical data from LCA taxa support some predictions of the traditional models, but suggest that inter- and intra-island patterns of dispersal, diversification (e.g. radiation), and persistence deserve more attention (e.g. Wang & Shaffer, 2008).

(12) Phylogeographical studies recover historical patterns of dispersal and population divergence coincident with formation of LCA's Isthmus and other major physiographic features, especially since the Pliocene. Quantitatively summarising the timeline of diversification inferred from studies to date suggests that both Neogene and Quaternary events have driven LCA diversification, with an exponential Pliocene–recent increase in rates of diversification.

(13) Although LCA phylogeography studies are uncovering many novel insights into the assembly and diversification of this recent Neotropical biota, more phylogeography studies using (*i*) better data, including larger numbers of unlinked molecular markers; (*ii*) geographically and taxonomically expanded, comparative sampling strategies; (*iii*) more hypothesis-driven approaches; and (*iv*) the latest statistical phylogeographical methods accounting for coalescent stochasticity and other potentially confounding processes, are needed. Adopting these approaches should greatly improve biogeographical inferences in the region; however, given the long-recognised state of fragmentation in historical biogeography, future work should also (*v*) integrate phylogeographical inferences with data and methods from disjunct fields of biogeography (e.g. ecological biogeography) and other disciplines, e.g. ecology and geospatial modelling.

(14) Phylogeography is poised to make critical contributions to conservation biology in LCA, including the way we view and prioritize areas and species for conservation resources. We encourage phylogeographers to apply their work to conservation; here, inferences into cryptic intraspecific diversification, genetic endemism, and predicting past-to-future species persistence and environmental stability across multiple taxa and areas seem particularly promising.

XI. ACKNOWLEDGEMENTS

We thank M. F. Breitman, B. J. Adams, and two anonymous reviewers for very helpful comments on earlier drafts of this manuscript. J.C.B. was funded by research and teaching assistantships from the Brigham Young University (BYU) Department of Biology, a Graduate Research Fellowship from BYU Graduate Studies, and a BYU Mentoring Environment Grant (to J.B.J). J.C.B.'s PhD research received additional funding from an NSF Doctoral Dissertation Improvement Grant (DEB-1210883), and Idea Wild (http://www.ideawild.org/).

XII. REFERENCES

**References marked with asterisk have been cited within the supporting information.*

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., Coad, B., Mandrak, N., Contreras-Balderas, S., Bussing, W., Stiassny, M. L. J., Skelton, P., Allen, G. R., Unmack, P., Naseka, A., Ng, R., Sindorf, N., Robertson, J., Armijo, E., Higgins, J. V., Heibel, T. J., Wikramanayake, E., Olson, D., López, H. L., Reis, R., Lundberg, J. G., Sabaj Pérez, M. H. & Petry, P. (2008). Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. *Bioscience* **58**, 403–414.

Abratis, M. & Wörner, G. (2001). Ridge collision, slab-window formation, and the flux fo Pacific asthenosphere into the Caribbean realm. *Geology* **29**, 127–130.

ANAM (2010). *Cuarto Informe Nacional de Panamá Ante el Convenio Sobre la Diversidad Biológica*. ANAM, Panamá. Available at http://www.anam.gob.pa/images/file/CUARTO_INFORME_NACIONAL.pdf. Accessed 09.11.2012

Anderson, R. P. & Handley, C. O. Jr. (2002). Dwarfism in insular sloths: biogeography, selection, and evolutionary rate. *Evolution* **56**, 1045–1058.

*Arbeláez-Cortés, E., Nyári, A. S. & Navarro-Sigüenza, A. G. (2010). The differential effect of lowlands on the phylogeographic pattern of a Mesoamerican montane species (*Lepidocolaptes affinis*, Aves: Furnariidae). *Molecular Phylogenetics and Evolution* **57**, 658–668.

Arbogast, B. S. & Kenagy, G. J. (2001). Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography* **28**, 819–825.

*Avise, J. C. (1998). The history and purview of phylogeography: a personal reflection. *Molecular Ecology* **7**, 371–379.

Avise, J. C. (2000). *Phylogeography: the History and Formation of Species*. Harvard University Press, Cambridge.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. & Saunders, N. C. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**, 489–522.

Beheregaray, L. B. (2008). Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology* **17**, 3754–3774.

*Bergemann, S. E., Smith, M. A., Parrent, J. L., Gilbert, G. S. & Garbelotto, M. (2009). Genetic population structure and distribution of a fungal polypore, *Datronia caperata* (Polyporaceae), in mangrove forests of Central America. *Journal of Biogeography* **36**, 266–279.

Bermingham, E. & Avise, J. C. (1986). Molecular zoogeography of fresh-water fishes in the southeastern United States. *Genetics* **113**, 939–965.

Bermingham, E. & Martin, A. P. (1998). Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology* **7**, 499–517.

Bermingham, E. & Moritz, C. (1998). Comparative phylogeography: concepts and applications. *Molecular Ecology* **7**, 367–369.

Bermingham, E., McCafferty, S. & Martin, A. (1997). Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. In *Molecular Systematics of Fishes* (eds T. Kocher and C. Stepien), pp. 113-128. Academic Press, San Diego.

Bloomquist, E. W., Lemey, P. & Suchard, M. A. (2010). Three roads diverged? Routes to phylogeographic inference. *Trends in Ecology & Evolution* **25**, 626–632.

Bolaños, F., Savage, J. M. & Chaves, G. (2011). *Amphibians and Reptiles of Costa Rica*. Listas Zoológicas Actualizadas, Museo de Zoología de la Universidad de Costa Rica, San Pedro. Available at http://museo.biologia.ucr.ac.cr/Listas/LZAPublicaciones.htm.

Bonaccorso, E., Navarro-Siguenza, A. G., Sanchez-Gonzalez, L. A., Peterson, A. T. & Garcia-Moreno, J. (2008). Genetic differentiation of the *Chlorospingus ophthalmicus* complex in Mexico and Central America. *Journal of Avian Biology* **39**, 311–321.

*Brower, A. V. Z. (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 6491–6495.

Brown, J. H. & Lomolino, M. V. (2000). Concluding remarks: historical perspective and the future of island biogeography theory. *Global Ecology and Biogeography* **9**, 87–92.

*Brown, W. M., George, M. Jr. & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 1967–1971.

*Brown, J. L., Maan, M. E., Cummings, M. E. & Summers, K. (2010). Evidence for selection on coloration in a Panamanian poison frog: a coalescent-based approach. *Journal of Biogeography* **37**, 891–901.

Buchs, D. M., Arculus, R. J., Baumgartner, P. O. & Ulianov, A. (2011). Oceanic intraplate volcanoes exposed: example from seamounts accreted in Panama. *Geology* **39**, 335–338.

Burnham, R. J. & Graham, A. (1999). The history of neotropical vegetation: new developments and status. *Annals of the Missouri Botanical Garden* **86**, 546–589.

Bush, M. B. & Colinvaux, P. A. (1994). Tropical forest disturbance: paleoecological records from Darien, Panama. *Ecology* **75**, 1761–1768.

Bush, M. B., Piperno, D. R., Colinvaux, P. A., Deoliveira, P. E., Krissek, L. A., Miller, M. C. & Rowe, W. E. (1992). A 14,300-yr paleoecological profile of a lowland tropical lake in Panama. *Ecological Monographs* **62**, 251–275.

Bussing, W. A. (1976). Geographic distribution of the San Juan ichthyofauna of Central America with remarks on its origin and ecology. In *Investigations of the Ichthyofauna of Nicaraguan Lakes* (ed. T. B. Thorson), pp. 157–175. University of Nebraska, Lincoln.

Bussing, W. A. (1985). Patterns of the distribution of the Central American ichthyofauna. In *The Great American Biotic Interchange* (eds F. G. Stehli and S. D. Webb), pp. 453–473. Plenum Press, New York.

Bussing, W. A. (1998). *Freshwater Fishes of Costa Rica*. Revista de Biologica Tropical, Universidad de Costa Rica, San José.

Cadle, J. E. & Greene, H. W. (1993). Phylogenetic patterns, biogeography, and the ecological structure of Neotropical snake assemblages. In *Species Diversity in Ecological Communities: Historical and Geographical Perspectives* (eds R. E. Ricklefs and D. Schluter), pp. 281–293. University of Chicago Press, Chicago.

Campbell, J. A. (1999). Distribution patterns of amphibians in Middle America. In *Patterns of Distribution of Amphibians: A Global Perspective* (ed. W. E. Duellman), pp. 111–210. Johns Hopkins University Press, Baltimore.

Carnaval, A. C. & Moritz, C. (2008). Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. *Journal of Biogeography* **35**, 1187–1201.

Carnaval, A. C., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T. & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. *Science* **323**, 785–789.

Carr, M. J., Saginor, I., Alvarado, G. E., Bolge, L. L., Lindsay, F. N., Milidakis, K., Turrin, B. D., Feigenson, M. D. & Swisher, C. C. III (2007). Element fluxes from the volcanic front of Nicaragua and Costa Rica. *Geochemistry, Geophysics, Geosystems* **8**, 1–22.

Castoe, T. A., Daza, J. M., Smith, E. N., Sasa, M. M., Kuch, U., Campbell, J. A., Chippindale, P. T. & Parkinson, C. L. (2009). Comparative phylogeography of pitvipers suggests a consensus of ancient Middle American highland biogeography. *Journal of Biogeography* **36**, 88–103.

*Cavers, S., Navarro, C. & Lowe, A. J. (2003). Chloroplast DNA phylogeography reveals colonization history of a Neotropical tree, *Cedrela odorata* L., in Mesoamerica. *Molecular Ecology* **12**, 1451–1460.

Chacón, C. (2005). Fostering conservation of key priority sites and rural development in Central America: the role of private protected areas. *Parks* **15**, 39–47.

Chacón, C. (2008). Private lands conservation in Mesoamerica. In *IUCN World Conservation Congress 2008*, Barcelona, Spain.

*Chacón, M. I., Pickersgill, B., Debouck, D. G. & Arias, J. S. (2007). Phylogeographic analysis of the chloroplast DNA variation in wild common bean (*Phaseolus vulgaris* L.) in the Americas. *Plant Systematics and Evolution* **266**, 175–195.

Chan, L. M., Brown, J. L. & Yoder, A. D. (2011). Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Molecular Phylogenetics and Evolution* **59**, 523–537.

*Cheviron, Z. A., Hackett, S. J. & Capparella, A. P. (2005). Complex evolutionary history of a Neotropical lowland forest bird (*Lepidothrix coronata*) and its implications for historical hypotheses of the origin of Neotropical avian diversity. *Molecular Phylogenetics and Evolution* **36**, 338–357.

*Clement, M., Posada, D. & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657–1659.

Coates, A. G. & Obando, J. A. (1996). The geologic evolution of the Central American Isthmus. In *Evolution and Environment in Tropical America* (eds J. B. C. Jackson, A. F. Budd and A. G. Coates), pp. 21–56. University of Chicago Press, Chicago.

Coates, A. G., Jackson, J. B., Collins, L. S., Cronin, T. M., Dowsett, H. J., Bybell, L. M., Jung, P. & Obando, J. A. (1992). Closure of the Isthmus of Panama: the near-shore marine record of Costa Rica and Panama. *Geological Society of America Bulletin* **104**, 814–828.

Coates, A. G., Collins, L. S., Aubry, M. P. & Berggren, W. A. (2004). The geology of the Darien, Panama, and the late Miocene-Pliocene collision of the Panama arc with northwestern South America. *Geological Society of America Bulletin* **116**, 1327–1344.

Coates, A. G., McNeill, D. F., Aubry, M. P., Berggren, W. A. & Collins, L. S. (2005). An introduction to the geology of the Bocas del Toro archipelago, Panama. *Caribbean Journal of Science* **41**, 374–391.

Cody, S., Richardson, J. E., Rull, V., Ellis, C. & Pennington, R. T. (2010). The Great American Biotic Interchange revisited. *Ecography* **33**, 326–332.

Colinvaux, P. (1993). Pleistocene biogeography and diversity in tropical forests of South America. In *Biological Relationships between Africa and South America* (ed. P. Goldblatt), pp. 473–499. Yale University Press, New Haven.

Colinvaux, P. A. (1996). Quaternary environmental history and forest diversity in the Neotropics. In *Evolution and Environment in Tropical America* (eds J. B. C. Jackson, A. F. Budd and A. G. Coates), pp. 359–405. University of Chicago Press, Chicago.

Colinvaux, P. A., Bush, M. B., SteinitzKannan, M. & Miller, M. C. (1997). Glacial and postglacial pollen records from the Ecuadorian Andes and Amazon. *Quaternary Research* **48**, 69–78.

Colinvaux, P. A., De Oliveira, P. E. & Bush, M. B. (2000). Amazonian and neotropical plant communities on glacial time-scales: the failure of the aridity and refuge hypotheses. *Quaternary Science Reviews* **19**, 141–169.

*Crawford, A. J. (2003). Huge populations and old species of Costa Rican and Panamanian dirt frogs inferred from mitochondrial and nuclear gene sequences. *Molecular Ecology* **12**, 2525–2540.

Crawford, A. J., Bermingham, E. & Polania, C. (2007). The role of tropical dry forest as a long-term barrier to dispersal: a comparative phylogeographical analysis of dry forest tolerant and intolerant frogs. *Molecular Ecology* **16**, 4789–4807.

Cunningham, C. W. & Collins, T. (1994). Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. In *Molecular Ecology and Evolution: Approaches and Applications* (eds B. Schierwater, B. Streit, G. P. Wagner and R. DeSalle), pp. 406–433. Birkhauser Verlag, Basel.

Darwin, C. (1859). *On the Origin of Species*. John Murray, London.

Davis, S. D., Heywood, V. H., Herrera-MacBryde, O., Villa-Lobos, J. & Hamilton, A. (1997). *Centres of Plant Diversity: a Guide and Strategy for Their Conservation, Volume 3: The Americas*. IUCN Publications Unit, Cambridge. Available at http://botany.si.edu/projects/cpd. Accessed 11.09.2012

*Daza, J. M., Smith, E. N., Paez, V. P. & Parkinson, C. L. (2009). Complex evolution in the Neotropics: the origin and diversification of the widespread genus *Leptodeira* (Serpentes: Colubridae). *Molecular Phylogenetics and Evolution* **53**, 653–667.

Daza, J. M., Castoe, T. A. & Parkinson, C. L. (2010). Using regional comparative phylogeographic data from snake lineages to infer historical processes in Middle America. *Ecography* **33**, 343–354.

*Demastes, J. W., Hafner, M. S. & Hafner, D. J. (1996). Phylogeographic variation in two Central American pocket gophers (*Orthogeomys*). *Journal of Mammalogy* **77**, 917–927.

Dick, C. W. & Heuertz, M. (2008). The complex biogeographic history of a widespread tropical tree species. *Evolution* **62**, 2760–2774.

Dick, C. W., Abdul-Salim, K. & Bermingham, E. (2003). Molecular systematic analysis reveals cryptic Tertiary diversification of a widespread tropical rain forest tree. *American Naturalist* **162**, 691–703.

Dick, C. W., Roubik, D. W., Gruber, K. F. & Bermingham, E. (2004). Long-distance gene flow and cross-Andean dispersal of lowland rainforest bees (Apidae: Euglossini) revealed by comparative mitochondrial DNA phylogeography. *Molecular Ecology* **13**, 3775–3785.

*Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**, e88.

Duque-Caro, H. (1990). Neogene stratigraphy, paleoceanography and paleobiogeography in northwest South America and the evolution of the Panama seaway. *Palaeogeography Palaeoclimatology Palaeoecology* **77**, 203–234.

Ebach, M. C. & Tangney, R. S. (2007). *Biogeography in a Changing World*. CRC Press, Boca Raton.

Edwards, S. V. & Beerli, P. (2000). Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* **54**, 1839–1854.

*Eizirik, E., Kim, J.-H., Menotti-Raymond, M., Crawshaw, P. G. Jr., O'Brein, S. J. & Johnson, W. E. (2001). Phylogeography, population history and conservation genetics of jaguars (*Panthera onca*, Mammalia, Felidae). *Molecular Ecology* **10**, 65–79.

Emig, C. C. & Geistdoerfer, P. (2004). The Mediterranean deep-sea fauna: historical evolution, bathymetric variations and geographical changes. *Carnets de Géologie* **2004**, 1–10.

Evans, S. (1999). *The Green Republic: A Conservation History of Costa Rica*. University of Texas Press, Austin.

*Felsenstein, J. (2004). *Inferring Phylogenies*. Sinauer Associates, Sunderland.

Fleming, K., Johnston, P., Zwartz, D., Yokoyama, Y., Lambeck, K. & Chappell, J. (1998). Refining the eustatic sea-level curve since the Last Glacial Maximum using far- and intermediate-field sites. *Earth and Planetary Science Letters* **163**, 327–342.

Funk, J., Mann, P., McIntosh, K. & Stephens, J. (2009). Cenozoic tectonics of the Nicaraguan depression, Nicaragua, and Median Trough, El Salvador, based on seismic-reflection profiling and remote-sensing data. *Geological Society of America Bulletin* **121**, 1491–1521.

García-París, M., Good, D. A., Parra-Olea, G. & Wake, D. B. (2000). Biodiversity of Costa Rican salamanders: implications of high levels of genetic differentiation and phylogeographic structure for species formation. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 1640–1647.

Garrick, R. C., Caccone, A. & Sunnucks, P. (2010). Inference of population history by coupling exploratory and model-driven phylogeographic analyses. *International Journal of Molecular Sciences* **11**, 1190–1227.

Gascon, C., Malcolm, J. R., Patton, J. L., da Silva, M. N. F., Bogart, J. P., Lougheed, S. C., Peres, C. A., Neckel, S. & Boag, P. T. (2000). Riverine barriers and the geographic distribution of Amazonian species. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 13672–13677.

Gaston, K. J. (2000). Global patterns in biodiversity. *Nature* **405**, 220–227.

Gazel, E., Carr, M. J., Hoernle, K., Feigenson, M. D., Szymanski, D., Hauff, F. & van den Bogaard, P. (2008). Galapagos-OIB signature in southern Central

America: mantle refertilization by arc-hot spot interaction. *Geochemistry, Geophysics, Geosystems* **10**, 1–32.
Gentry, A. H. (1982). Neotropical floristic diversity: phytogeographical connections between Central and South America, Pleistocene climatic fluctuations, or an accident of Andean orogeny? *Annals of the Missouri Botanical Garden* **69**, 557–593.
*Gibbard, P. L., Head, M. J., Walker, M. J. C. & the Subcommission on Quaternary Stratigraphy (2009). Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. *Journal of Quaternary Science* **25**, 96–102.
Gómez, L. D. (1986). *Vegetación de Costa Rica, apuntes para una biogeografía Costarricense. Vegetación y clima de Costa Rica*. Editorial Universidad Estatal a Distancia, San José.
*González, M. A., Eberhard, J. R., Lovette, I. J., Olson, S. L. & Bermingham, E. (2003). Mitochondrial DNA phylogeography of the bay wren (Troglodytidae: *Thryothorus nigricapillus*) complex. *Condor* **105**, 228–238.
González, C., Urrego, L. E. & Martínez, J. I. (2006). Late quaternary vegetation and climate change in the Panama basin: palynological evidence from marine cores ODP 677 and TR 163-38. *Palaeogeography, Palaeoclimatology, Palaeoecology* **234**, 62–80.
Graham, A. & Dilcher, D. (1995). The Cenozoic record of tropical dry forest in northern Latin America and the southern United States. In *Seasonally Dry Tropical Forests* (eds S. H. Bullock, H. A. Mooney and E. Medina), pp. 124–145. Cambridge University Press, Cambridge.
Gregory-Wodzicki, K. M. (2000). Uplift history of the central and northern Andes: a review. *Geological Society of America Bulletin* **112**, 1091–1105.
Günther, A. (1861). On a collection of fishes sent by Capt Dow from the Pacific coast of Central America. *Proceedings of the Zoological Society of London* **1861**, 370–376.
Haffer, J. (1969). Speciation in Amazonian forest birds. *Science* **165**, 131–137.
Haffer, J. (1997). Alternative models of vertebrate speciation in Amazonia: an overview. *Biodiversity and Conservation* **6**, 451–476.
*Hagemann, S. & Pröhl, H. (2007). Mitochondrial paraphyly in a polymorphic poison frog species (Dendrobatidae; *D. pumilio*). *Molecular Phylogenetics and Evolution* **45**, 740–747.
Halffter, G. (1987). Biogeography of the montane entomofauna of Mexico and Central America. *Annual Review of Entomology* **32**, 95–114.
Haq, B. U., Hardenbol, J. & Vail, P. R. (1987). Chronology of fluctuating sea levels since the Triassic. *Science* **235**, 1156–1167.
*Hasbún, C. R., Gomez, A., Kohler, G. & Lunt, D. H. (2005). Mitochondrial DNA phylogeography of the Mesoamerican spiny-tailed lizards (*Ctenosaura quinquecarinata* complex): historical biogeography, species status and conservation. *Molecular Ecology* **14**, 3095–3107.
*Hasegawa, M., Kishino, H. & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**, 160–174.
Hauff, F., Hoernle, K., van den Bogaard, P., Alvarado, G. & Garbe-Schönberg, D. (2000). Age and geochemistry of basaltic complexes in western Costa Rica: contributions to the geotectonic evolution of Central America. *Geochemistry, Geophysics, Geosystems* **1**, 1–41.
Haug, G. H. & Tiedemann, R. (1998). Effect of the formation of the Isthmus of Panama on Atlantic Ocean thermohaline circulation. *Nature* **393**, 673–676.
*Hauswaldt, J. S., Ludewig, A. K., Vences, M. & Prohl, H. (2011). Widespread co-occurrence of divergent mitochondrial haplotype lineages in a Central American species of poison frog (*Oophaga pumilio*). *Journal of Biogeography* **38**, 711–726.
Heaney, L. R. (2007). Is a new paradigm emerging for oceanic island biogeography? *Journal of Biogeography* **34**, 753–757.
Hearty, P. J., Kindler, P., Cheng, H. & Edwards, R. L. (1999). A +20 m middle Pleistocene sea-level highstand (Bermuda and the Bahamas) due to partial collapse of Antarctic ice. *Geology* **27**, 375–378.
Hewitt, G. M. (2000). The genetic legacy of the ice ages. *Nature* **405**, 907–913.
Hey, J. & Machado, C. A. (2003). The study of structured populations—new hope for a difficult and divided science. *Nature Reviews Genetics* **4**, 535–543.
Hickerson, M. J., Dolman, G. & Moritz, C. (2006*a*). Comparative phylogeographic summary statistics for testing simultaneous vicariance. *Molecular Ecology* **15**, 209–223.
Hickerson, M. J., Stahl, E. A. & Lessios, H. A. (2006*b*). Test for simultaneous divergence using approximate Bayesian computation. *Evolution* **60**, 2435–2453.
Hickerson, M. J., Stahl, E. & Takebayashi, N. (2007). msBayes: pipeline for testing comparative phylogeographic histories using hierarchical approximate Bayesian computation. *BMC Bioinformatics* **8**, 268.
Hickerson, M. J., Carstens, B. C., Cavender-Bares, J., Crandall, K. A., Graham, C. H., Johnson, J. B., Rissler, L., Victoriano, P. F. & Yoder, A. D. (2010). Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* **54**, 291–301.
Hildebrand, S. F. (1938). A new catalogue of the fresh-water fishes of Panama. *Field Museum of Natural History Zoological Series* **22**, 219–359.
*Hillis, D. M. & Bull, J. J. (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* **42**, 182–192.
Hoernle, K., van den Bogaard, P., Werner, R., Lissinna, B., Hauff, F., Alvarado, G. & Garbe-Schönberg, D. (2002). Missing history (16–71 Ma) of the Galápagos hotspot: implications for the tectonic and biological evolution of the Americas. *Geology* **30**, 795–798.
Hoffmann, F. G. & Baker, R. J. (2003). Comparative phylogeography of short-tailed bats (*Carollia*: Phyllostomidae). *Molecular Ecology* **12**, 3403–3414.
Hoffmann, F. G., Owen, J. G. & Baker, R. J. (2003). mtDNA perspective of chromosomal diversification and hybridization in Peters' tent-making bat (*Uroderma bilobatum*: Phyllostomidae). *Molecular Ecology* **12**, 2981–2993.
*Hoorn, C., Wesselingh, F. P., ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., Sanmartín, I., Sanchez-Meseguer, A., Anderson, C. L., Figueiredo, J. P., Jaramillo, C., Riff, D., Negri, F. R., Hooghiemstra, H., Lundberg, J., Stadler, T., Särkinen, T. & Antonelli, A. (2010). Amazonia through time: Andean uplift, climate change, landscape evolution and biodiversity. *Science* **330**, 927–931.
Huang, W., Takebayashi, N., Qi, Y. & Hickerson, M. J. (2011). MTML-msBayes: approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. *BMC Bioinformatics* **12**, 1.
Hubbell, S. P. (2001). *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton.
Hurlbert, S. H. & Villalobos-Figueroa, A. (1982). *Aquatic Biota of México, Central America, and the West Indies*. San Diego State University, San Diego.
*Hynková, I., Starostova, Z. & Frynta, D. (2009). Mitochondrial DNA variation reveals recent evolutionary history of main *Boa constrictor* clades. *Zoological Science* **26**, 623–631.
Ibaraki, M. (2002). Responses of planktonic foraminifera to the emergence of the Isthmus of Panama. *Revista Mexicana de Ciencias Geológicas* **19**, 152–160.
IEG (Independent Evaluation Group) (2011). *The Mesoamerican Biological Corridor*. Regional Program Review 5. The World Bank Group, Washington.
Islebe, G. A. & Hooghiemstra, H. (1997). Vegetation and climate history of montane Costa Rica since the last glacial. *Quaternary Science Reviews* **16**, 589–604.
Jackson, J. C. B., Budd, A. F. & Coates, A. G. (eds) (1996). *Evolution and Environment in Tropical America*. University of Chicago Press, Chicago.
Jones, C. P. & Johnson, J. B. (2009). Phylogeography of the livebearer *Xenophallus umbratilis* (Teleostei: Poeciliidae): glacial cycles and sea level change predict diversification of a freshwater tropical fish. *Molecular Ecology* **18**, 1640–1653.
Karr, J. R. (1990). Birds of tropical rainforest: comparative biogeography and ecology. In *Biogeography and Ecology of Forest Bird Communities* (ed. A. Keast), pp. 215–228. SPB Academic Publishing, The Hague.
Keigwin, L. (1982). Isotopic paleoceanography of the Caribbean and East Pacific: role of Panama uplift in Late Neogene time. *Science* **217**, 350–353.
Keller, G., Zenker, C. E. & Stone, S. M. (1989). Late Neogene history of the Pacific-Caribbean gateway. *Journal of South American Earth Sciences* **2**, 73–108.
Kidd, D. M. & Ritchie, M. G. (2006). Phylogeographic information systems: putting the geography into phylogeography. *Journal of Biogeography* **33**, 1851–1865.
*Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111–120.
Kirby, M. X. & MacFadden, B. (2005). Was southern Central America an archipelago or a peninsula in the middle Miocene? A test using land-mammal body size. *Palaeogeography, Palaeoclimatology, Palaeoecology* **228**, 193–202.
Kirby, M. X., Jones, D. S. & MacFadden, B. J. (2008). Lower Miocene stratigraphy along the Panama Canal and its bearing on the Central American Peninsula. *PLoS One* **3**, e2791.
Knowles, L. L. (2009). Statistical phylogeography. *Annual Review of Ecology, Evolution, and Systematics* **40**, 593–612.
Knowles, L. L. & Carstens, B. C. (2007). Estimating a geographically explicit model of population divergence. *Evolution* **61**, 477–493.
Knowles, L. L. & Maddison, W. P. (2002). Statistical phylogeography. *Molecular Ecology* **11**, 2623–2635.
Kominz, M. A., Browning, J. V., Miller, K. G., Sugarman, P. J., Mizintseva, S. & Scotese, C. R. (2008). Late Cretaceous to Miocene sea-level estimates from the New Jersey and Delaware coastal plain coreholes: an error analysis. *Basin Research* **20**, 211–226.
Kozak, K. H., Graham, C. H. & Wiens, J. J. (2008). Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology & Evolution* **23**, 141–148.
Kreft, H. & Jetz, W. (2007). Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 5925–5930.
*Kubatko, L. S., Carstens, B. C. & Knowles, L. L. (2009). STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics* **25**, 971–973.
Kuhner, M. K. (2009). Coalescent genealogy samplers: windows into population history. *Trends in Ecology & Evolution* **24**, 86–93.
Lachniet, M. S. (2004). Late Quaternary glaciation of Costa Rica and Guatemala, Central America. In *Quaternary Glaciations-Extent and Chronology, Part III: South America, Asia, Africa, Australia, Antarctica, Developments in Quaternary Science* (eds J. Ehlers and P. Gibbard), pp. 135–138. Elsevier, Amsterdam.
*Larget, B. & Simon, D. L. (1999). Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Molecular Biology and Evolution* **16**, 750–759.

Lee, J. B. & Johnson, J. B. (2009). Biogeography of the livebearing fish *Poecilia gillii* in Costa Rica: are phylogeographical breaks congruent with fish community boundaries? *Molecular Ecology* **18**, 4088–4101.
Leigh, E. G., O'Dea, A. & Vermeij, G. J. (2014). Historical biogeography of the Isthmus of Panama. *Biological Reviews* **89**, 148–172.
Leonard, H. J. (1987). *Natural Resources and Economic Development in Central America: A Regional Environmental Profile*. Transaction Books, New Brunswick.
Lessios, H. A. (2008). The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. *Annual Review of Ecology, Evolution, and Systematics* **39**, 63–91.
Lips, K. R., Reeve, J. D. & Witters, L. R. (2003). Ecological traits predicting amphibian population declines in Central America. *Conservation Biology* **17**, 1078–1088.
Liu, L. & Pearl, D. K. (2007). Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Systematic Biology* **56**, 504–514.
Loaiza, J. R., Scott, M. E., Bermingham, E., Rovira, J. & Conn, J. E. (2010*a*). Evidence for Pleistocene population divergence and expansion of *Anopheles albimanus* in southern Central America. *The American Journal of Tropical Medicine and Hygiene* **82**, 156–164.
*Loaiza, J. R., Scott, M. E., Bermingham, E., Sanjur, O. I., Wilkerson, R., Rovira, J., Gutiérrez, L. A., Correa, M. M., Grijalva, M. J., Birnberg, L., Bickersmith, S. & Conn, J. E. (2010*b*). Late Pleistocene environmental changes lead to unstable demography and population divergence of *Anopheles albimanus* in the northern Neotropics. *Molecular Phylogenetics and Evolution* **57**, 1341–1346.
Lomolino, M. V., Riddle, B. R., Whittaker, R. J. & Brown, J. H. (2010). *Biogeography*. Fourth Edition. Sinauer Associates, Sunderland.
Losos, J. B. & Glor, R. E. (2003). Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology & Evolution* **18**, 220–227.
Lunt, D. J., Valdes, P. J., Haywood, A. & Rutt, I. C. (2008). Closure of the Panama Seaway during the Pliocene: implications for climate and Northern Hemisphere glaciation. *Climate Dynamics* **30**, 1–18.
MacArthur, R. H. (1972). *Geographical Ecology: Patterns in the Distribution of Species*. Harper and Row, New York.
MacArthur, R. H. & Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton University Press, Princeton.
*Macey, J. R., Schulte, J. A. II, Larson, A., Fang, A., Wang, Y., Tuniyev, B. S. & Papenfuss, T. J. (1998). Phylogenetic relationships of toads in the *Bufo bufo* species group from the eastern escarpment of the Tibetan Plateau: a case of vicariance and dispersal. *Molecular Phylogenetics and Evolution* **9**, 80–87.
*Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology* **46**, 523–536.
Mann, P., Rogers, R. D. & Gahagan, L. (2007). Overview of plate tectonic history and its unresolved tectonic problems. In *Central America: Geology, Resources and Hazards* (eds J. Bundschuh and G. E. Alvarado), pp. 205–241. Taylor & Francis, Philadelphia.
Marks, B. D., Hackett, S. J. & Capparella, A. P. (2002). Historical relationships among Neotropical lowland forest areas of endemism as determined by mitochondrial DNA sequence variation within the Wedge-billed Woodcreeper (Aves: Dendrocolaptidae: *Glyphorynchus spirurus*). *Molecular Phylogenetics and Evolution* **24**, 153–167.
Marshall, L. G. (1979). A model for paleobiogeography of South American cricetine rodents. *Paleobiology* **5**, 126–132.
Marshall, J. S. (2007). The geomorphology and physiographic provinces of Central America. In *Central America: Geology, Resources and Hazards* (eds J. Bundschuh and G. E. Alvarado), pp. 1–51. Taylor & Francis, Philadelphia.
Marshall, L., Butler, R. F., Drake, R. E., Curtis, G. A. & Tedforth, R. H. (1979). Calibration of the Great American Interchange. *Science* **204**, 272–279.
Marshall, J. S., Idleman, B. D., Gardner, T. W. & Fisher, D. M. (2003). Landscape evolution within a retreating volcanic arc, Costa Rica, Central America. *Geology* **31**, 419–422.
Martin, A. P. & Bermingham, E. (2000). Regional endemism and cryptic species revealed by molecular and morphological analysis of a widespread species of Neotropical catfish. *Proceedings of the Royal Society B* **267**, 1135–1141.
Martins, F. M., Templeton, A. R., Pavan, A. C. O., Kohlbach, B. C. & Morgante, J. S. (2009). Phylogeography of the common vampire bat (*Desmodus rotundus*): marked population structure, Neotropical Pleistocene vicariance and incongruence between nuclear and mtDNA markers. *BMC Evolutionary Biology* **9**, 294.
McCafferty, W. P. (1998). Ephemeroptera and the great American interchange. *Journal of the North American Benthological Society* **17**, 1–20.
*McCafferty, S. S., Martin, A. & Bermingham, E. (2012). Phylogeographic diversity of the lower Central American cichlid *Andinoacara coeruleopunctatus* (Cichlidae). *International Journal of Evolutionary Biology* **2012**, 1–12 (doi:10.1155/2012/780169).
Miller, R. R. (1966). Geographical distribution of Central American freshwater fishes. *Copeia* **1966**, 773–802.
Miller, K. G., Kominz, M. A., Browning, J. V., Wright, J. D., Mountain, G. S., Katz, M. E., Sugarman, P. J., Cramer, B. S., Christie-Blick, N. & Pekar, S. F. (2005). The phanerozoic record of global sea-level change. *Science* **310**, 1293–1298.
*Miller, M. J., Bermingham, E., Klicka, J., Escalante, P., Do Amaral, F. S. R., Weir, J. T. & Winker, K. (2008). Out of Amazonia again and again: episodic crossing of the Andes promotes diversification in a lowland forest flycatcher. *Proceedings of the Royal Society B* **275**, 1133–1142.
Montes, C., Cardona, A., McFadden, R., Morón, S. E., Silva, C. A., Restrepo-Moreno, S., Ramírez, D. A., Hoyos, N., Wilson, J., Farris, D., Bayona, G. A., Jaramillo, C. A., Valencia, V., Bryan, J. & Flores, J. A. (2012). Evidence for middle Eocene and younger land emergence in central Panama: implications for Isthmus closure. *Geological Society of America Bulletin* **124**, 780–799.
Moritz, C. (2002). Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology* **51**, 238–254.
Moritz, C., Patton, J. L., Schneider, C. J. & Smith, T. B. (2000). Diversification of rainforest faunas: an integrated molecular approach. *Annual Review of Ecology and Systematics* **31**, 533–563.
Morrone, J. J. (2006). Biogeographic areas and transition zones of Latin America and the Caribbean islands based on panbiogeographic and cladistic analyses of the entomofauna. *Annual Review of Entomology* **51**, 467–494.
*Morse, G. E. & Farrell, B. D. (2005). Interspecific phylogeography of the *Stator limbatus* species complex: the geographic context of speciation and specialization. *Molecular Phylogenetics and Evolution* **36**, 201–213.
Müller, R. D., Sdrolias, M., Gaina, C., Steinberger, B. & Heine, C. (2008). Long-term sea-level fluctuations driven by ocean basin dynamics. *Science* **319**, 1357–1362.
Mutke, J. & Barthlott, W. (2005). Patterns of vascular plant diversity at continental to global scales. *Biologiske Skrifter* **55**, 521–531.
Myers, G. S. (1938). Fresh-water fishes and West Indian zoogeography. *Annual Report of the Smithsonian Institution* **1937**, 339–364.
Myers, G. S. (1966). Derivation of the freshwater fish fauna of Central America. *Copeia* **4**, 766–773.
Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858.
*Navarro-Sigüenza, A. G., Peterson, A. T., Nyári, A., Garcia-Deras, G. M. & Garcia-Moreno, J. (2008). Phylogeography of the *Buarremon* brush-finch complex (Aves, Emberizidae) in Mesoamerica. *Molecular Phylogenetics and Evolution* **47**, 21–35.
*Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York.
*Nei, M. & Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.
Nelson, G. & Platnick, N. (1981). *Systematics and Biogeography: Cladistics and Vicariance*. Columbia University Press, New York.
Nores, M. (1999). An alternative hypothesis for the origin of Amazonian bird diversity. *Journal of Biogeography* **26**, 475–485.
Nores, M. (2004). The implications of Tertiary and Quaternary sea level rise events for avian distribution patterns in the lowlands of northern South America. *Global Ecology and Biogeography* **13**, 149–161.
*Novick, R. R., Dick, C. W., Lemes, M. R., Navarro, C., Caccone, A. & Bermingham, E. (2003). Genetic structure of Mesoamerican populations of Big-leaf mahogany (*Swietenia macrophylla*) inferred from microsatellite analysis. *Molecular Ecology* **12**, 2885–2893.
*Nyári, A. S. (2007). Phylogeographic patterns, molecular and vocal differentiation, and species limits in *Schiffornis turdina* (Aves). *Molecular Phylogenetics and Evolution* **44**, 154–164.
Obando, V. (2002). *Biodiversidad en Costa Rica: Estado del Conocimiento y Gestión*. InBio, Heredia.
*Ornelas-García, C. P., Dominguez-Dominguez, O. & Doadrio, I. (2008). Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evolutionary Biology* **8**, 340.
Patton, J. L., Da Silva, M. N. F. & Malcolm, J. R. (1994). Gene genealogy and differentiation among arboreal spiny rats (Rodentia: Echimyidae) of the Amazon Basin: a test of the riverine barrier hypothesis. *Evolution* **48**, 1314–1323.
Pennington, R. T., Cronk, Q. C. B. & Richardson, J. A. (2004). Introduction and synthesis: plant phylogeny and the origin of major biomes. *Proceedings of the Royal Society B* **359**, 1455–1464.
Perdices, A., Bermingham, E., Montilla, A. & Doadrio, I. (2002). Evolutionary history of the genus *Rhamdia* (Teleostei: Pimelodidae) in Central America. *Molecular Phylogenetics and Evolution* **25**, 172–189.
Perdices, A., Doadrio, I. & Bermingham, E. (2005). Evolutionary history of the synbranchid eels (Teleostei: Synbranchidae) in Central America and the Caribbean islands inferred from their molecular phylogeny. *Molecular Phylogenetics and Evolution* **37**, 460–473.
Peterson, A. T., Soberón, J. & Sánchez-Cordero, V. (1999). Conservatism of ecological niches in evolutionary time. *Science* **285**, 1265–1267.
Petit, R. J. & Hampe, A. (2006). Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics* **37**, 187–214.
Piperno, D. R. & Pearsall, D. M. (1998). *The Origins of Agriculture in the Lowland Neotropics*. Academic Press, San Diego.

Posadas, P., Crisci, J. V. & Katinas, L. (2006). Historical biogeography: a review of its basic concepts and critical issues. *Journal of Arid Environments* **66**, 389–403.
*Rabosky, D. L. & Lovette, I. J. (2008). Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution* **62**, 1866–1875.
Raven, P. H. & Axelrod, D. I. (1974). Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden* **61**, 539–673.
Reeves, R. G. & Bermingham, E. (2006). Colonization, population expansion, and lineage turnover: phylogeography of Mesoamerican characiform fish. *Biological Journal of the Linnean Society* **88**, 235–255.
Reid, W. V. & Miller, K. R. (1989). *Keeping Options Alive: the Scientific Basis for Conserving Biodiversity*. World Resources Institute, Washington.
Remington, C. L. (1968). Suture-zones of hybrid interaction between recently joined biotas. In *Evolutionary Biology* (eds T. Dobzhansky, M. K. Hecht and W. C. Steere), pp. 321–428. Plenum Press, New York.
Retallack, G. J. & Kirby, M. X. (2007). Middle Miocene global change and paleogeography of Panama. *Palaios* **22**, 667–679.
Richards, C. L., Carstens, B. C. & Knowles, L. L. (2007). Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *Journal of Biogeography* **34**, 1833–1845.
Ricklefs, R. E. (1987). Community diversity: relative roles of local and regional processes. *Science* **235**, 167–171.
Ricklefs, R. E. (2006). Evolutionary diversification and the origin of the diversity-environment relationship. *Ecology* **87**, S3–S13.
Ricklefs, R. E. & Schluter, D. (1993). *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. University of Chicago Press, Chicago.
Riddle, B. R. (1996). The molecular phylogeographic bridge between deep and shallow history in continental biotas. *Trends in Ecology & Evolution* **11**, 207–211.
Riddle, B. R. & Hafner, D. J. (2006). A step-wise approach to integrating phylogeographic and phylogenetic biogeographic perspectives on the history of a core North American warm deserts biota. *Journal of Arid Environments* **66**, 435–461.
Riddle, B. R. & Hafner, D. J. (2010). Integrating pattern with process at biogeographic boundaries: the legacy of Wallace. *Ecography* **33**, 321–325.
Riddle, B. R., Dawson, M. N., Hadly, E. A., Hafner, D. J., Hickerson, M. J., Mantooth, S. J. & Yoder, A. D. (2008). The role of molecular genetics in sculpting the future of integrative biogeography. *Progress in Physical Geography* **32**, 173–202.
*Robertson, J. M. & Zamudio, K. R. (2009). Genetic diversification, vicariance, and selection in a polytypic frog. *Journal of Heredity* **100**, 715–731.
Robertson, J. M., Duryea, M. C. & Zamudio, K. R. (2009). Discordant patterns of evolutionary differentiation in two Neotropical treefrogs. *Molecular Ecology* **18**, 1375–1395.
Ronquist, F. (1997). Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology* **46**, 195–203.
Rose, W. I., Bluth, G. J. S., Carr, M. J., Ewert, J. W., Patino, L. C. & Vallance, J. W. (eds) (2006). *Volcanic Hazards in Central America*. Geological Society of America Special Paper 412. Geological Society of America, Boulder.
Rosen, D. E. (1975). A vicariance model of Caribbean biogeography. *Systematic Zoology* **24**, 431–464.
Rosen, D. E. (1978). Vicariant patterns and historical explanation in biogeography. *Systematic Zoology* **27**, 159–188.
*Rull, V. (2007). On the origin of present Neotropical biodiversity: a preliminary meta-analysis about speciation timing using molecular phylogenies. *Orsis* **22**, 105–119.
*Rull, V. (2008). Speciation timing and neotropical biodiversity: the Tertiary–Quaternary debate in the light of molecular phylogenetic evidence. *Molecular Ecology* **17**, 2722–2729.
Rull, V. (2011). Neotropical biodiversity: timing and potential drivers. *Trends in Ecology & Evolution* **26**, 508–513.
*Rutschmann, F. (2006). Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times. *Diversity and Distributions* **12**, 35–48.
Sala, O. E., Chapin, F. S. III, Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L. F., Jackson, R. B., Kinzig, A., Leemans, R., Lodge, D. M., Mooney, H. A., Oesterheld, M., Poff, N. L., Sykes, M. T., Walker, B. H., Walker, M. & Wall, D. H. (2000). Global biodiversity scenarios for the year 2100. *Science* **287**, 1770–1774.
*Sanderson, M. J. (1997). A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* **14**, 1218–1231.
Sanmartín, I. (2010). Evolutionary biogeography: an integrative approach. *Systematic Biology* **59**, 486–488.
Savage, J. M. (1966). The origins and history of the Central American herpetofauna. *Copeia* **1966**, 719–766.
Savage, J. M. (1982). The enigma of the Central American herpetofauna: dispersals or vicariance? *Annals of the Missouri Botanical Garden* **69**, 464–547.
Savage, J. M. (2002). *The Amphibians and Reptiles of Costa Rica: a Herpetofauna Between Two Continents, Between Two Seas*. University of Chicago Press, Chicago.
Schmidt, D. N. (2007). The closure history of the Panama Isthmus: evidence from isotopes and fossils to models and molecules. In *Deep Time Perspectives on Climate Change--Marrying the Signal from Computer Models and Biological Proxies* (eds M. Williams, A. M. Haywood, J. F. Gregory and D. N. Schmidt), pp. 427–442. Geological Society of London, London.
Schneider, C. J., Cunningham, M. & Moritz, C. (1998). Comparative phylogeography and the history of endemic vertebrates in the Wet Tropics rainforests of Australia. *Molecular Ecology* **7**, 487–498.
Scotese, C. (2008). Plate tectonic and paleogeographic mapping: state of the art. Search and Discovery Article 40312, In *American Association of Petroleum Geologists, Annual Convention*, San Antonio, Texas. Available at http://www.searchanddiscovery.net/documents/2008/08029scotese/index.htm. Accessed 11.10.2012
*Serrano-Serrano, M. L., Hernandez-Torres, J., Castillo-Villamizar, G., Debouck, D. G. & Sanchez, M. I. (2010). Gene pools in wild Lima bean (*Phaseolus lunatus* L.) from the Americas: evidences for an Andean origin and past migrations. *Molecular Phylogenetics and Evolution* **54**, 76–87.
Seyfried, H., Astorga, A., Amann, H., Calvo, C., Kolb, W., Schmidt, H. & Winsemann, J. (1991). Anatomy of an evolving island arc: tectonic and eustatic control in the south Central American fore-arc area. *Special Publications of the International Association of Sedimentologists* **12**, 217–240.
Sherrod, D. R., Vallance, J. W., Espinosa, T. & McGeehin, A. (2007). Volcán Barú: eruptive history and volcanic-hazards assessment. United States Geological Survey Open-File Report 2007-1401, 1-33.
Simon, M. F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T. & Hughes, C. E. (2009). Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 20359–20364.
Simpson, G. G. (1940). Mammals and land bridges. *Journal of the Washington Academy of Sciences* **30**, 137–163.
Simpson, G. G. (1950). History of the fauna of Latin America. *American Scientist* **38**, 361–389.
*Slade, R. W. & Moritz, C. (1998). Phylogeography of *Bufo marinus* from its natural and introduced ranges. *Proceedings of the Royal Society B* **265**, 769–777.
Smith, S. A. & Bermingham, E. (2005). The biogeography of lower Mesoamerican freshwater fishes. *Journal of Biogeography* **32**, 1835–1854.
Solomon, S. E., Bacci, M., Martins, J., Vinha, G. G. & Mueller, U. G. (2008). Paleodistributions and comparative molecular phylogeography of leafcutter ants (*Atta* spp.) provide new insight into the origins of Amazonian diversity. *PLoS One* **3**, e2738.
Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S. & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology* **15**, 4261–4293.
Stehli, F. G. & Webb, S. D. (eds) (1985). *The Great American Biotic Interchange*. Plenum Press, New York.
Stotz, D. F., Fitzpatrick, J. W., Parker, T. A. I. & Moskovits, D. K. (1996). *Neotropical Birds: Ecology and Conservation*. University of Chicago Press, Chicago.
Streicher, J. W., Crawford, A. J. & Edwards, C. W. (2009). Multilocus molecular phylogenetic analysis of the montane *Craugastor podiciferus* species complex (Anura: Craugastoridae) in Isthmian Central America. *Molecular Phylogenetics and Evolution* **53**, 620–630.
Sullivan, J., Arellano, E. & Rogers, D. (2000). Comparative phylogeography of Mesoamerican highland rodents: concerted versus independent response to past climatic fluctuations. *American Naturalist* **155**, 755–768.
Thomson, R. C., Wang, I. J. & Johnson, J. R. (2010). Genome-enabled development of DNA markers for ecology, evolution and conservation. *Molecular Ecology* **19**, 2184–2195.
UNESCO (2012a). *Biosphere Reserves*. UNESCO. Available at http://www.unesco.org/new/en/natural-sciences/environment/ecological-sciences/biosphere-reserves. Accessed 09.11.2012
UNESCO (2012b). *World Heritage List*. UNESCO. Available at http://whc.unesco.org/en/list. Accessed 09.11.2012
Vázquez-Miranda, H., Navarro-Sigüenza, A. G. & Omland, K. E. (2009). Phylogeography of the rufous-naped wren (*Campylorhynchus rufinucha*): speciation and hybridization in Mesoamerica. *The Auk* **126**, 765–778.
Venegas-Anaya, M., Crawford, A. J., Galvan, A. H. E., Sanjur, O. I., Densmore, L. D. & Bermingham, E. (2008). Mitochondrial DNA phylogeography of *Caiman crocodilus* in Mesoamerica and South America. *Journal of Experimental Zoology Part A* **309A**, 614–627.
Vermeij, G. J. (1991). When biotas meet: understanding biotic interchange. *Science* **253**, 1099–1104.
Wakeley, J. (2002). Inferences about the structure and history of populations: coalescents and intraspecific phylogeography. In *The Evolution of Population Biology* (eds R. S. Singh and M. K. Uyenoyama), pp. 193–215. Cambridge University Press, Cambridge.
*Walker, J. D. & Geissman, J. W. (2009). *Geologic Time Scale*. Geological Society of America. Available at http://geosociety.org/science/timescale. Accessed 05.06.2010
Wallace, A. R. (1876). *The Geographical Distribution of Animals*. Macmillan, London.

Wang, I. J. & Shaffer, H. B. (2008). Rapid color evolution in an aposematic species: a phylogenetic analysis of color variation in the strikingly polymorphic strawberry poison-dart frog. *Evolution* **62**, 2742–2759.
Wang, I. J., Crawford, A. J. & Bermingham, E. (2008). Phylogeography of the pygmy rain frog (*Pristimantis ridens*) across the lowland wet forests of isthmian Central America. *Molecular Phylogenetics and Evolution* **47**, 992–1004.
Webb, S. D. (1995). Biological implications of the middle Miocene Amazon seaway. *Science* **269**, 361–362.
Webb, S. D. (1997). The great American faunal interchange. In *Central America: A Natural and Cultural History* (ed. A. G. Coates), pp. 97–122. Yale University Press, New Haven.
Webb, S. D. (2006). The Great American Biotic Interchange: patterns and processes. *Annals of the Missouri Botanical Garden* **93**, 245–257.
Webb, S. D. & Rancy, A. (1996). Late Cenozoic evolution of the Neotropical mammal fauna. In *Evolution and Environment in Tropical America* (eds J. B. C. Jackson, A. F. Budd and A. G. Coates), pp. 335–358. The University of Chicago Press, Chicago.
Weigt, L. A., Crawford, A. J., Rand, A. S. & Ryan, M. J. (2005). Biogeography of the túngara frog, *Physalaemus pustulosus*: a molecular perspective. *Molecular Ecology* **14**, 3857–3876.
Weir, J. T. (2009). Implications of genetic differentiation in Neotropical montane forest birds. *Annals of the Missouri Botanical Garden* **96**, 410–433.
Weir, J. T., Bermingham, E., Miller, M. J., Klicka, J. & Gonzalez, M. A. (2008). Phylogeography of a morphologically diverse Neotropical montane species, the Common Bush-Tanager (*Chlorospingus ophthalmicus*). *Molecular Phylogenetics and Evolution* **47**, 650–664.
Weir, J. T., Bermingham, E. & Schluter, D. (2009). The Great American Biotic Interchange in birds. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 21737–21742.
Welcomme, R. L. (1988). International introductions of inland aquatic species. *FAO Fisheries Technical Paper* **294**, 1-318.
Whitmore, T. C. & Prance, G. T. (1987). *Biogeography and Quaternary History in Tropical America*. Oxford University Press, Oxford.
Whitmore, F. C. & Stewart, R. H. (1965). Miocene mammals and Central American seaways. *Science* **148**, 180–185.
Wiens, J. J. & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution* **19**, 639–644.
Wiens, J. J. & Graham, C. H. (2005). Niche conservatism: integrating evolution, ecology and conservation biology. *Annual Review of Ecology, Evolution, and Systematics* **36**, 519–539.
*Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U. B., Helm, K. M., Bychowski, R., Higuchi, R. G., Palumbi, S. R., Prager, E. M., Sage, R. D. & Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society* **26**, 385–400.
Yesson, C. & Culham, A. (2006). Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. *Systematic Biology* **55**, 785–802.
Young, B. E., Lips, K. R., Reaser, J. K., Ibáñez, R., Salas, A. W., Cedeño, J. R., Coloma, L. A., Ron, S., La Marca, E., Meyer, J. R., Muñoz, A., Bolaños, F., Chaves, G. & Romo, D. (2004). Population declines and priorities for amphibian conservation in Latin America. *Conservation Biology* **15**, 1213–1223.
Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. (2001). Trends, rhythyms, and aberrations in global climate 65 Ma to present. *Science* **292**, 686–693.
Zamudio, K. R. & Greene, H. W. (1997). Phylogeography of the bushmaster (*Lachesis muta*: Viperidae): implications for neotropical biogeography, systematics, and conservation. *Biological Journal of the Linnean Society* **62**, 421–442.
*Zeh, D. W., Zeh, J. A. & Bonilla, M. M. (2003a). Phylogeography of the giant harlequin beetle (*Acrocinus longimanus*). *Journal of Biogeography* **30**, 747–753.
Zeh, J. A., Zeh, D. W. & Bonilla, M. M. (2003b). Phylogeography of the harlequin beetle-riding pseudoscorpion and the rise of the Isthmus of Panama. *Molecular Ecology* **12**, 2759–2769.
Zink, R. M. (2002). Methods in comparative phylogeography, and their application to studying evolution in the North American aridlands. *Integrative and Comparative Biology* **42**, 953–959.
Zink, R. M. & Barrowclough, G. F. (2008). Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology* **17**, 2107–2121.

XIII. SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.

Table S1. Summary of patterns emerging from lower Central American (LCA) phylogeography studies ($N = 58$), including phylogeographical breaks.

Table S2. Summary of molecular dating methods and divergence dates for major lineages recovered in lower Central American phylogeography studies.

Table S3. Summary of species contributions to phylogeographic breaks recovered in LCA.

Appendix S1. Materials and methods.

Appendix S2. Literature search results.

(*Received 23 December 2012; revised 29 November 2013; accepted 3 December 2013; published online 3 February 2014*)

Chapter 1 – Supplementary Material

Table S1. Summary of patterns emerging from lower Central American (LCA) phylogeography studies ($N = 58$), including phylogeographical breaks shown in Fig. 6B–F of the main text. In parentheses next to taxon names, the number of nominal LCA taxonomic lineages sampled phylogeographically (left of slash) is contrasted against the approximate number of genetically distinct evolutionary lineages recovered within LCA (right side of slash). Abbreviations: Allo., allozymes; BDT, Bocas del Toro; Cp, chloroplast; cpDNA chloroplast DNA sequences; IBD, isolation by distance; microsat., microsatellite DNA; Mt, mtDNA sequences; N, nDNA sequences; NCA, Nuclear Central America (Guatemala to Honduras); RFLP, mtDNA (animals) or cpDNA (plants) restriction fragment data; SA, South America. §Studies without explicitly stated *a priori* geographical or ecological hypotheses.

No. Reference	Taxon	Molecular markers	Patterns
Demastes *et al.* (1996)	***Orthogeomys* mammals (2/5)**	Mt, RFLP	Within *Orthogeomys cherriei*, two clades exhibit E–W Central Cordillera split (CC break), no IBD; in *O. underwoodi*, linear/pectinate phylogeographic structure and IBD, support for Atlantic-coast Costa Rica breaks at Rio Savegre (SAV) and Fila Costeña (FILA).
Zamudio & Greene (1997)§	***Lachesis* snakes (1/2)**	Mt	Sampling *Lachesis* phylogeographically yielded an E–W Talamanca Cordillera split (CVF break), with one *Lachesis* lineage isolated on either side of Chorotega volcanic front (*L. melanocephala–L. stenophrys*; but disjunct distribution).
Bermingham & Martin (1998)	***Roeboides, Brachyhypopomus, Pimelodella* fishes (4/~12–22)**	Mt, RFLP	Multiple (~6–9+) clades per genus connected by short internodes; drainage basin isolation; El Valle volcano break (VALLE break) supported within *Roeboides guatemalensis*; western Panama isthmus (WPI) break supported within *R. guatemalensis–R. occidentalis*, *Brachyhypopomus occidentalis*, and *Pimelodella chagresi*; LCA–northern SA split (EPI break) supported within three lineages—*Roeboides*, *B. occidentalis*, and *P. chagresi*; and Panama–northern SA continental divide break (PNSA) supported in *Roeboides* and *P. chagresi*. Rapid dispersals from SA? Overall, four nominal taxa contained ~12–22 distinct genetic lineages.
Slade & Moritz (1998)	***Bufo marinus* toads (1/1)**	Mt	No LCA structure, although LCA samples fell into same trans-Andes subclade (along with northwestern Venezuela; but inadequate sampling); cis–trans Andes divergence. Pre-isthmus LCA dispersal(s)?
García-París *et al.* (2000)§	***Bolitoglossa pesrubra* salamanders (2/4)**	Mt, Allo.	Significant structure across LCA mountain tops within Talamanca Cordillera, Costa Rica, resulting in three unique *B. pesrubra* lineages; "Salsipuedes" clade *versus* "Villa Mills" clade split supported the Talamanca Cordillera break 1 (TCMF1). Overall, four parapatric Talamanca Cordillera lineages were recovered within two nominal taxa (*B. pesrubra*, *B.* sp. B).

Martin & Bermingham (2000)	***Pimelodella chagresi* fish (2/~6–9)**	Mt, RFLP	Two main mtDNA lineages—first one larger, sister to Magdalena basin, second smaller and sister to Atrato basin; mtDNA clades with distinct endonuclease haplogroups (e.g. "AAA" unique to larger lineage; only second smaller clade had "ECC" and "GDD"), plus additional samples; split between Osa and Burica peninsulas at Piedras Blancas (PB break); WPI, Panama–northern SA continental divide (PNSA), CPI, CHIC, and EPI breaks also supported; drainage basin isolation, interdrainage gene flow. LCA colonisation from SA by two different lineages at two different times? Overall, two nominal *Pimelodella* species sampled phylogeographically contained about 6–9 genetic lineages (clades/subclades).
Eizirik *et al.* (2001)	***Panthera onca* jaguars (1/1)**	Mt, Microsat.	Mixed result: no LCA structure, mtDNA genome (no well-supported, distinct sister clades in phylogeny; *P. onca* control region sequences fell into one minimum-spanning network); however, Nicaraguan depression split (ND break) supported by phylogeny of 29 microsatellites (but inadequate sampling, no internal phylogenetic support values); some geographical differentiation suggested reduced gene flow across Amazon river and eastern Panama isthmus (Darién), but, overall, evidence for high historical gene flow levels.
Marks *et al*. (2002)	***Glyphorynchus spirurus* birds (1/2)**	Mt	Panama–northern SA continental divide break supported (PNSA; although inadequate sampling) by split between clade containing Atlantic LCA plus Imerí (SA, east of Andes) *versus* western Panama + Chocó (SA). Colonization from SA following isthmus closure? Overall, *G. spirurus* contained about eight unique genetic lineages.
Perdices *et al.* (2002)§	***Rhamdia* catfishes (3/~4–10)**	Mt	Approximately four distinct major LCA (trans-) clades (1–2/lineage; two substructured); drainage basin isolation within Rio Bebedero (BEB break); and cis–trans Andes structure; surprisingly, no EPI break strongly supported. Rapid migration into/through LCA from SA? Overall, three nominal *Rhamdia* spp. contained up to ~4–10 unique genetic lineages (up to 23, including clades plus subclades).
Cavers *et al*. (2003)§	***Cedrela odorata* plants (1/1)**	Cp	No LCA structure; limited genetic diversity recovered in three cpDNA network haplogroups (separated by only ≤6 mutations), but all recovered in one minimum spanning network; wet- *versus* dry-adapted "types". Pre-isthmus dispersal of a northern haplotype from NCA, and two post-isthmus dispersals from SA: one after isthmus formation, the other very recent (post-Pleistocene, ~13 ka)?
Crawford (2003)	***Eleutherodactylus* frogs (3/5)**	Mt, N	Three nominal taxa sampled in LCA phylogeographically contained ~5 major lineages. Within *E. stejnegerianus*, three clades fell into a basal polytomy; but the split between *E. stejnegerianus* and *E. persimilis*

			supported the CVF break; also the interspecific *E. bransfordii–E. polyptychus* split supported the L1 break. In addition, a clade of three populations from Atlantic lowlands/island (La Selva + Fila Carbón + Isla Colón) were sister to a clade containing the Atlantic montane area plus Pacific versant (*E. persimilis* + *E. stejnegerianus*). However, no LCA structure within *E. polyptychus*. Overall, five taxa sampled contained seven unique genetic lineages.
Dick *et al*. (2003)	***Symphonia globulifera* plants (1/1)**	N	No LCA structure; genetic differentiation within (1–9 mutational differences between haplotypes); cis–trans Andes structure. Pre-isthmus LCA dispersal?
González *et al.* (2003)	***Thryothorus nigricapillus* birds (1/3)**	Mt	Two main LCA clades with E–W Panama isthmus structure (BT break); island–mainland substructure at Bocas del Toro; LCA–SA break between eastern Darién (Panama) and Ecuador within "*nigricapillus* group". LCA dispersal around isthmus closure? Overall, three unique genetic lineages were supported within *T. nigricapillus*.
Hoffman & Baker (2003)§	***Carollia* bats (2/4)**	Mt	Within *C. sowelli*, a split between one LCA clade with some substructure sister to a NCA clade from Honduras supported the ND break (although sampling was inadequate or disjunct); a cis–trans Andes split plus two separate LCA clades strongly supported the AZUE break at the west-central Azuero peninsula, Panama within *C. castanea*; the EPI split is also supported within *C. castanea*. Post-isthmus LCA dispersal(s)? Overall, two LCA *Carollia* species sampled phylogeographically contained four unique genetic lineages.
Hoffman *et al*. (2003)	***Uroderma bilobatum* bats (1/1)**	Mt	No LCA structure (probably unrelated to inadequate spatial sampling employed), but found evidence for three "chromosomal races". Post-isthmus dispersal and diversification?
Novick *et al.* (2003)	***Swietenia macrophylla* plants (1/1)**	Microsat.	Guanacaste Cordillera differentiation but no phylogeographic breaks, no LCA structure (otherwise inadequate sampling); structure within putative Pleistocene refuge; IBD supported. Pre-isthmus LCA dispersal?
Zeh *et al*. (2003*a*)§	***Acrocinus longimanus* beetles (1/1)**	Mt	No LCA structure (most variation within populations); among two Panama clades, only one was strongly supported, and internal branching relationships likewise were not strongly supported.
Zeh *et al*. (2003*b*)§	***Cordylochernes scorpioides* pseudoscorpions (1/~2–3)**	Mt	Two to three distinct but geographically overlapping clades recovered in LCA, and no clear spatial breaks. Two LCA dispersals from SA? Bermingham/Martin model (Bermingham & Martin, 1998) supported? Overall, about three to four unique *Cordylochernes* lineages recovered in the Neotropics.

Dick *et al.* (2004)§	**Euglossine bees (12/~13)**	Mt	No LCA structure; long-range dispersal (cross-Andes), gene flow, rapid expansion. Overall, 12 nominal euglossine species sampled within LCA contained approximately 13 unique genetic lineages.
Cheviron *et al*. (2005)	***Lepidothrix coronata* birds (1/1)**	Mt	No LCA structure; stable demography; cis–trans Andes split. Riverine barriers? Overall, despite containing only a single LCA lineage, *L. coronata* contained about 4–6 unique genetic lineages throughout the Neotropics.
Hasbún *et al.* (2005)	***Ctenosaura quinquecarinata* lizards (1/1)**	Mt	No LCA structure, but LCA–NCA split (Nicaragua + Costa Rica clade was sister to Mexico clade). Overall, despite containing only a single LCA lineage, *C. quinquecarinata* contained three unique genetic lineages throughout the study area.
Morse & Farrell (2005)§	***Stator limbatus* beetles (1/2)**	Mt	North *versus* South American clades split in LCA, but inadequate sampling; no LCA structure defined because geographical position of split indeterminate (phylogeographical break occurs somewhere between Guanacaste in southwest Costa Rica and Azuero peninsula, Panama). Northward dispersal from SA? Overall, despite containing only two LCA lineages, *S. limbatus* contained 3–4 unique genetic lineages throughout southern North America, the Caribbean and the Neotropics.
Perdices *et al*. (2005)	***Synbranchus marmoratus* swamp eels (fish) (1/3)**	Mt, N	This one species sampled phylogeographically throughout LCA yielded three unique genetic lineages based on several complex patterns (e.g. spatially multi-dimensional phylogeographical breaks) that we do not lump but consider to provide support for individual CVF and WPI breaks (split between "SyMCA" *versus* "SyLCA" clades), as well as CPI and BT breaks ("SyLCA Bayano" clade *versus* other mtDNA *Synbranchus*). Pre-isthmus LCA dispersal(s)? Overall, whereas one species sampled in LCA (*S. marmoratus*) contained three genetic lineages, two nominal eel taxa (*S. marmoratus* and *Ophisternon aenigmaticum* eels) sampled contained ~12 unique genetic lineages throughout Central America and the Caribbean.
Weigt *et al.* (2005)	***Engystomops pustulosus* frogs (1/4)**	Mt, Allo.	Four clades were recovered in LCA and data strongly supported the western Panama isthmus (WPI), Azuero peninsula (AZUE), and Las Perlas (PERL) breaks (but inadequate sampling). One split between Costa Rica and Panama (northern *versus* southern clades) could not be geographically determined but corresponded to the basal split within the species. Pre-isthmus dispersal(s)? Overall, *E. pustulosus* contained six unique genetic lineages throughout the Neotropics.
Reeves & Bermingham (2006)§	***Brycon*, *Bryconamericus*, *Cyphocharax*, and *Eretmobrycon***	Mt, RFLP	Multiple clades (approximately six per "major lineage") within 17 nominal taxa sampled phylogeographically in LCA (but minimum evolution trees presented with no phylogenetic support indices limit confidence in relationships); some drainage basin isolation; SIXA break supported within

	characid fishes (17/29–31)		"*Bryconamericus scleroparius* species group"; CPI break supported within western *Brycon* major lineage; eastern *Bryconamericus*—the "*emperador* species group"—supported the CPI and Bayano–Tuira (BT) breaks; and the eastern *Brycon* major-lineage supported the El Valle (VALLE) and Playón Chico (CHIC) breaks. Temporally staggered, post-isthmus LCA dispersals? As noted by McCafferty *et al*. (2012), this study supports cross-cordillera exchange/gene flow in the Cocle del Sur/Norte and Chagres/Tuira drainage pairs, in two lineages—eastern *Bryconamericus* and *Brycon argenteus*. No phylogeographic structuring was observed within *Cyphocharax magdalenae* or *Eretmobrycon bayano*, except for differentiation between the Sixaola and Changuinola rivers in the latter species. Overall, the authors sampled four genera and ~17–20 species, and discovered 29–31 unique lineages.
Chacón *et al.* (2007)§	***Phaseolus vulgaris* plants (1/1)**	Cp, RFLP	No LCA structure despite three closely related one-step cpDNA clades (but inadequate sampling). Post-isthmus divergence?
Crawford *et al*. (2007)	***Craugastor* frogs (three spp., main focus of study) (3/~9)**	Mt	Three main focal LCA *Craugastor* species contained ~9–13 unique genetic lineages (with 9 being a more conservative estimate); WPI break supported within *C. crassidigitus*; *C. talamancae* "Caribbean" clade structure supported POPA break; Sapo range split (SAPO break), supported within *C. raniformis*; E–W Atlantic coast break (either L1 or BDT1) inferred between Limón headland and BDT in *C. fitzingeri* ("western PA"–"CR" + "HN", all within "Caribbean" clade) but geographical position of split indeterminate (inadequate sampling). Pre-isthmus dispersal?
Hagemann & Pröhl (2007)§	***Oophaga pumilio* frogs (1/3)**	Mt	Three main mtDNA clades recovered in LCA; within one clade, *O. pumilio* sister to *Oophaga arboreus* supported the CVF break; the POPA, L1, BDT1, BDT2, and ESCU breaks were all supported, mostly within *O. pumilio*; overlapping sister clades recovered along the Tortuguero lowlands in Costa Rica, including one clade from Caño Negro to Pueblo Nuevo lying atop a second clade from Upala to Tortuguero; *in situ* LCA diversification? Two colonisation events?
Nyári (2007)§	***Schiffornis turdina* birds (1/2)**	Mt	Two non-sister LCA clades were recovered within this species. One break between clades "1" *versus* "3" occurred similar to but did not coincide with the PNSA break—this break corresponds roughly to the eastern Panama isthmus (EPI) break (despite inadequate sampling), and possibly reflects replacement or competitive exclusion of lineage "1" by lineage "2" in eastern Panama. Overall, seven unique genetic lineages from throughout the Neotropics were recovered within *S. turdina*.

Bonaccorso *et al.* (2008)	***Chlorospingus ophthalmicus* birds (2/2–3)**	Mt	Among two subspecies taxa sampled in LCA, results supported the Nicaraguan depression break (ND break) and a shallow mid-southern Talamanca Cordillera split (TCMF2 break). LCA dispersal around or following isthmus closure? Insufficient phylogenetic signal? While 2–3 major lineages were recovered within the two LCA subspecies of *C. ophthalmicus*, ~10 subspecies of *C. ophthalmicus* were sampled overall and contained around one unique genetic lineage each.
Dick & Heuertz (2008)	***Symphonia globulifera* plants (1/1)**	Cp, N, microsat.	No LCA structure; cis–trans Andes structure; long-distance gene flow. Overall, *S. globulifera* contained approximately three unique genetic lineages throughout the Neotropics, but only a single LCA clade.
Miller *et al.* (2008)	***Mionectes oleagineus* birds (1/4)**	Mt, N	Cis–trans Andes structure, plus four trans-Andes clades, but no clear phylogeographical sister clade relationships, no clear phylogeographic structure fully within LCA; however, the PNSA break is supported, when samples from western Ecuador are also taken into account, by the northern LCA–western Ecuador split (despite a large disjunction). Post-isthmus, three cross-Andes dispersals from SA? Overall, *M. oleagineus* contained six unique genetic lineages distributed throughout the Neotropics, but only four LCA clades.
Navarro-Sigüenza *et al.* (2008)§	***Buarremon brunneinucha* birds (1/1)**	Mt	No LCA structure, but widely disjunct El Salvador–SA break recovered; population expansion inferred. Post-isthmus dispersal(s) into LCA and SA? Overall, despite containing only a single LCA lineage, *B. brunneinucha* contained eight unique genetic lineages throughout the Neotropics.
Ornelas-García *et al.* (2008)§	***Astyanax* fishes (4/8, over all LCA taxa)**	Mt, N	ND break weakly supported, due to sharing of haplotypes across ND clades; multiple areas of drainage basin isolation. Pre-isthmus LCA dispersal through Central America? B/M model support? Four nominal species sampled phylogeographically in LCA contained eight lineages. Overall, five nominal LCA taxa sampled in LCA—*A. aeneus*, *A. fasciatus*, *A. nasutus*, *A. nicaraguensis*, and *A. orthodus*—contained around nine distinct lineages (~1 clade/lineage; but inadequate sampling in some lineages).
Solomon *et al.* (2008)	***Atta* ants (1/1)**	Mt	*A. cephalotes*: no LCA structure; Pleistocene LCA population diversification; population expansion and riverine model rejected (Amazon only). Insufficient data? Pleistocene refugia (Amazon only)? Post-isthmus LCA dispersal from SA? Overall, *A. cephalotes*, *A. sexdens*, and *A. laevigata* contained eight distinct genetic lineages.

Venegas-Anaya *et al.* (2008)§	***Caiman crocidilus* crocodiles (1/2)**	Mt	Two main LCA clades, with Talamanca Cordillera (CVF break) and MOSQ breaks supported by a complex phylogeographical pattern; beyond LCA, cis–trans Andes split also supported by basal divergence of NCA + LCA lineage *versus* east-of-Andes SA lineage. Post-isthmus LCA dispersal? Overall, the three *Caiman* subspecies contained five distinct genetic lineages throughout the Neotropics.
Wang *et al*. (2008)	***Pristimantis ridens* frogs (1/3)**	Mt	Three main lineages recovered in LCA within this species; MOSQ break supported along western Pacific, Panama isthmus supported; CPI break also supported; Costa Rican haplotypes had Golfo Dulce origin; Tilarán subclade, population expansion. Pre-isthmus LCA dispersal?
Wang & Shaffer (2008)	***Oophaga pumilio* frogs (1/~7–19)**	Mt, N	Four divergent colour morph clades were recovered within LCA, and these corresponded to phylogeographic breaks within and around Bocas del Toro supported, including POPA, BDT1, and BDT3 breaks, as well as the ESCU break supported by island–mainland divergence between Escudo de Veraguas Island and mainland populations. Overall, at least seven unique genetic lineages supported in LCA within one nominal taxon.
Weir *et al.* (2008)	***Chlorospingus ophthalmicus* birds (2/3)**	Mt	The phylogenetic split between *regionalis* and *punctulatus* subspecies along mid-southern Talamanca Cordillera supported the TCMF2 break; ND break supported by N–S split within *C. o. regionalis*; phylogeographic structure unexpectedly strongest in Central America, not the Andes. Overall, 11 *C. ophthalmicus* subspecies sampled contained 12 lineages (~1.1 distinct genetic lineages per subspecies).
Castoe *et al.* (2009)	**Middle American pitvipers (LCA *Atropoides* and *Cerrophidion*: 3/6)**	Mt	Three species sampled phylogeographically in LCA (*Atropoides picadoi*, *A. mexicanus*, *Cerrophidion godmani*) contained six lineages (three lineages endemic to LCA); ND break supported within *C. godmani* (although disjunct distribution/sampling); *A. picadoi* and *A. mexicanus*, no significant phylogeographic structure in either species. Pre-isthmus LCA dispersal southward from NCA? One or multiple dispersal-vicariance events? Overall, across three main genera examined, 17 nominal taxa contained 22 distinct genetic lineages, and the six species sampled phylogeographically throughout Central America yielded 12 distinct genetic lineages.
Daza *et al.* (2009)	***Leptodeira septentrionalis* snakes (1/2)**	Mt, N	Primarily sampled *L. s. ornata* phylogeographically in LCA; recovered two major *L. s. ornata* LCA clades supporting CVF break due to Atlantic–Pacific coast split between Limón (Costa Rica) + BDT (Panama) *versus* Puntarenas (Costa Rica) clades; cis–trans Andes LCA–SA break supported. Northward and southward dispersals into/through LCA, SA? Overall, the 12 main *Leptodeira* species/subspecies sampled contained ~15 unique genetic lineages.

Hynková *et al*. (2009)§	***Boa constrictor* snakes (1/2)**	Mt	Two main clades supported—one per subspecies (*B. c. constrictor*, *B. c. imperator*; plus around two differentiated subclades per subspecies)—with mostly CA–SA basal divergence, but only weak support for CA–SA split due to shared haplotypes between areas (e.g. Nicaragua sample "NIC2") and uncertainty about the geographical origin of samples; EPI break well-supported by split within *imperator* clade "3". LCA dispersal from South America approximately at time of isthmus formation?
Jones & Johnson (2009)	***Xenophallus umbratilis* fish (1/4)**	Mt	Two major lineages within Costa Rica, whose divergence supported SJ2 break; overall, four unique clade/subclade lineages from the San Juan superbasin, correlated with historical eustatic sea levels. Historical vicariance/habitat fragmentation due to marine incursion(s)?
Lee & Johnson (2009)	***Poecilia gillii* fish (1/1)**	Mt, N	The nine shallow subclades within Costa Rica are here considered part of one major genetic lineage—thus no significant LCA phylogeographic structure exists; some significant population genetic partitioning across biogeographic province boundaries; but gene flow across biogeographic province boundaries also supported, including movements across the Herradura headland region (e.g. Rio Grande de Tárcoles); evidence for historical population stasis preceding recent population declines within provinces.
Martins *et al.* (2009)	***Desmodus rotundus* bats (1/1)**	Mt, N	No LCA structure (but inadequate sampling); cis–trans Andes split supported between LCA ("CA") and Pantanal ("PAN") clades diverged in the Pleistocene. Overall, however, *D. rotundus* contained five unique genetic lineages.
Robertson *et al.* (2009)	***Agalychnis callidryas* and *Dendropsophus ebraccatus* frogs (2/9)**	Mt, microsat.	Four *versus* five major lineages recovered within *Dendropsophus* and *Agalychnis*, respectively; both species, phylogeographic breaks supported across the Fila Costeña (FILA) and western Panama isthmus (WPI), though FILA only weakly supported in *A. callidryas*; CVF break supported within *A. callidryas*; Limón headland (L1) break strongly supported in *D. ebraccatus*; IBD supported in both species.
Robertson & Zamudio (2009)	***Agalychnis callidryas* frogs (1/5)**	Mt	Five clades, with phylogeographical patterns identical to Robertson *et al*. (2009) and CVF, FILA and WPI breaks supported.
Streicher *et al.* (2009)§	***Craugastor podiciferus* frogs (2/7)**	Mt, N	Six clades recovered and CC and TCMF1 breaks supported within *C. podiciferus*. Within *C.* sp. A (clade "G"), the BARU break was supported. Overall, between two taxa recognised *a priori* as species-level that were sampled phylogeographically, seven unique LCA genetic lineages were recovered.
Vázquez-Miranda *et al.* (2009)§	***Campylorhynchus rufinucha* birds (1/1)**	Mt	LCA phylogeographic breaks could not be rigorously defined for this species, but evidence for multiple distinct clades and IBD was recovered, and

			they inferred NCA population expansions not shared in LCA. Although the authors recommended that three species be recognised, *C. rufinucha* appeared to contain two, not three, major lineages (one of which occurs in LCA) that are well supported by multiple phylogenetic analyses, and around three to five genetically differentiated groups (e.g. including major subclades).
Arbeláez-Cortés *et al.* (2010)	***Lepidocolaptes affinis* birds (1/2)**	Mt	Allopatric split supported across ND break (but disjunct distribution; sparse sampling in LCA); significant IBD supported; surprisingly no phylogeographic structure across Isthmus of Tehuantepec. Long-distance dispersal? Overall, while eight "haplotype groups" were recovered, *L. affinis* contained three well-supported major phylogeographic lineages (two of which occur in LCA).
Brown *et al.* (2010)	***Oophaga pumilio* frogs (1/1)**	Mt	No LCA structure; mtDNA D-loop tree characterised by a widespread lack of support for inferred relationships; apparently inadequate phylogenetic signal; used coalescent simulations to test for a significant role of natural selection over drift in shaping phenotypic divergence; found significant lineage sorting in colour not attributable to drift; significant incomplete lineage sorting. Hypothesised *in situ* Quaternary divergence consistent with geology?
Daza *et al*. (2010)§	**Middle American snakes (3/6)**	Mt	CVF break supported across Talamanca Cordillera in *Leptodeira septentrionalis* and *Bothrops asper*, with single, ~3.9 Ma diversification event inferred from simulations; ND break supported by *Cerrophidion godmani* (but inadequate sampling); inferred temporally staggered diversification events across the ND in multiple lineages; LCA–NCA split within *L. nigrofasciata*; cis–trans Andes split within *Leptodeira*; two pulses of LCA–SA diversification independent of isthmus closure. Results favour peninsula model? The three lineages sampled phylogeographically in LCA (*Lachesis*, two sister species; *L. septentrionalis*; *B. asper*) contained six unique genetic lineages. Overall, six of the sampled lineages yielded relevant breaks in and around LCA (*C. godmani*, *L. septentrionalis* + *L. annulata*, *L. nigrofasciata*, *B. asper*, *Bothriechis schlegelii*, and *Lachesis*), and these breaks corresponded to approximately 12 unique genetic lineages.
Loaiza *et al.* (2010*a*)	***Anopheles albimanus* mosquitoes (1/1)**	Mt	No significant LCA structure (all haplotypes fell into a single network), although groups of populations were genetically differentiated east-to-west within LCA; shallow divergence and population expansion (~22 ka) inferred in late Pleistocene; no IBD. Post-isthmus dispersals and secondary contact?

Loaiza *et al.* (2010*b*)	***Anopheles albimanus* mosquitoes (1/2)**	Mt, N	Weak support for EPI break due to LCA–SA split (this break was rather well supported by mtDNA but not supported by nDNA data); 3–4 presumably differentiated populations ("haplogroups"), as well as evidence for population expansions dated within the Pleistocene within each population; IBD supported within each haplogroup by Mantel tests.
Serrano-Serrano *et al.* (2010)	***Phaseolus lunatus* plants (1/1)**	Cp, N	No LCA structure (all LCA samples fell into a single widely distributed clade, "MII") and no well-supported breaks in the Neotropics; demographic population expansion within clade containing LCA samples supported by multiple metrics. Post-isthmus dispersal or diversification? Overall, *P. lunatus* contained ~3–5 unique "gene pools".
Hauswaldt *et al.* (2011)	***Oophaga pumilio* frogs (1/2)**	Mt, N, microsat.	Two geographically overlapping mtDNA lineages, with no clear geographical break; population expansions inferred. Coexistence of divergent haplotype lineages within a single species, rather than cryptic/incipient speciation?
McCafferty *et al.* (2012)	***Andinoacara coeruleopunctatus* fish (1/6)**	Mt	Recovered six major mtDNA lineages (focal taxon paraphyletic); WPI break supported; rapid colonisation and diversification of lineages in Panama; cross-cordillera exchange (sharing of mtDNA haplotypes) in Rio Cocle del Sur–Rio Cocle del Norte and Rio Chagres–Rio Tuira drainage pairs. Dispersal into LCA around time of LCA isthmus in the Pliocene (although pre-Pliocene dispersal not ruled out)?
J.C. Bagley & J.B. Johnson (unpublished data)	**Poeciliid fishes (three spp.; 3/6)**	Mt	Within *Alfaro cultratus*, three mtDNA lineages recovered, with SJ1 break strongly supported and SJ3 break weakly supported; ND break approximately supported between *A. cultratus* and sister lineage (*A. huberi*) in Honduras (but inadequate sampling across this break). Within *Poecilia gillii*, inferred southeastern (N–S) colonisation history despite incomplete lineage sorting from coalescent simulations; essentially no LCA structure. *Xenophallus* patterns essentially identical to Jones & Johnson (2009). Patterns influenced by incomplete lineage sorting? Overall, similar timelines of diversification but spatial incongruence (biogeographic incongruence) supported potentially different histories within region, across species; drainage basins not major barrier to gene flow across taxa (e.g. some evidence for gene flow or stream capture in all three species); some evidence for Quaternary population expansions.

Table S2. Summary of molecular dating methods and divergence dates for major lineages recovered in lower Central American phylogeography studies. Lineage divergence dates correspond to Fig. 7. Reference order corresponds to Table S1. Under Methods, "strict" indicates that a DNA sequence-based rate calibration was estimated or taken from a previous work, assumed to be a global rate, and used to convert genetic distances to absolute time; clock rates are presented in parentheses as pairwise per cent per million years (Myr^{-1}) or substitutions per site per unit time (e.g. subs. s^{-1} y^{-1}). By contrast, "relaxed" indicates that clock methods were used that account for or incorporate among-lineage rate variation in DNA sequence evolution, or 'rate heterogeneity' (rate het.). Columns F and P indicate that fossil data or palaeogeographic data (e.g. geological events, formations), respectively, were used to calibrate molecular clocks. N indicates that the calibrations were used to constrain one or more nodes of the tree during phylogeny estimation (allows rate calculation), whereas R indicates that calibrations were used to derive rates of evolution taken as global rates. In the two right-hand columns, asterisks (*) indicate crude divergence time estimates that we calculated based on pairwise mtDNA sequence divergence (seq div) conversions using the conventional 2% Myr^{-1} vertebrate rate, unless stated otherwise; corresponding values of approximate sequence divergences are given as percentages in parentheses. Other abbreviations are as in Table S1, except the following: BEAST, BEAST software program; HKY, Hasegawa-Kishino-Yano model of DNA sequence evolution (Hasegawa, Kishino & Yano, 1985); IM, isolation with migration software program; K2P, Kimura 2-parameter distance (Kimura, 1980); K_s, the number of synonymous substitutions; MDIV, MDIV software program; MRCA, most recent common ancestor; msBayes, the msBayes software pipeline (Hickerson *et al*., 2007); MULTI, MULTIDIVTIME software program; NPRS, non-parametric rate smoothing; PL, penalized likelihood; Q, Quaternary period (2.58 Ma to present); RAG1, nuclear recombination activating gene 1; T, Tertiary period (2.58–65.5 Ma). See Fig. 6 for definitions of breaks.

Reference	Taxon	Methods (rates)	F	P	Within-LCA divergence dates	LCA–SA (and other) divergence dates
Demastes *et al.* (1996)	***Orthogeomys* mammals**	Strict (2% Myr^{-1})	—	—	**T–Q:** 0.5 Ma basal *O. cherriei* divergence (CC break) and 3.6 Ma interspecific stem. *We calculated divergence dates at 1.73 Ma for the *O. underwoodi* FILA split, 580 ka for the *O. underwoodi* SAV split, and 750 ka for basal divergence within *O. cherriei* (CC break) based on seq div; we used these values in Fig. 7.	—
Zamudio & Greene (1997)	***Lachesis* snakes**	Strict (0.47–1.32% Myr^{-1})	R	R	**T:** 4–11 Ma crown (CVF break)	**T:** 6.4–17.9 Ma stem divergence

Bermingham & Martin (1998)	***Roeboides, Brachyhypopomus, Pimelodella* fishes**	Strict (1.3% Myr^{-1})	—	R	**T–Q:** ~1.0–3.0 Ma to recent divergences across VALLE, WPI, EPI breaks *Using approximate HKY distances (%) from this paper and the 1.3% Myr^{-1} rate at left, we estimated basal divergences of all LCA samples within lineages at 5.38 Ma (~7%, *Roeboides*), 6.15 Ma (~8%, *B. occidentalis*), and 4.0 Ma (~5.2%, *Pimelodella*). A 1.85 Ma (2.4%) divergence date was similarly estimated for the VALLE break; WPI break estimates were 4.54 Ma (~5.9%) in *Roeboides*, 1.92 Ma (~2.5%) in *B. occidentalis*, and 3.38 Ma (~4.4%) in *P. chagresi*); and 3.15 Ma (~4.1%, *Roeboides*) and 1.31 Ma (~1.7%, *P. chagresi*) dates were estimated for the PNSA break using the same substitution rate. We used these data in Fig. 7, rather than estimates from this paper.	**T:** ~4–7 Ma stems *Using methods at left, we calculated EPI divergences at 5.62 Ma (~7.3%) in *Roeboides*, 1.92 Ma (~2.5%) in *B. occidentalis*, and 1.77 Ma (2.3%) in *P. chagresi*.
Slade & Moritz (1998)	***Bufo marinus* toads**	Strict (2% Myr^{-1})	—	—	***Q:** we calculated a 235 ka crown for the clade containing LCA samples (0.47% max. within-Costa Rica seq div, LCA MRCA)	**T:** 2.7 Ma stem (LCA + SA–SA, cis–trans Andes) ***LCA–NCA (Q):** 740 ka (1.48% CR–southeastern Mexico seq div, LCA clade stem)
García-París *et al.* (2000)	***Bolitoglossa pesrubra* salamanders**	—	—	—	***Q:** we calculated a 1.95 Ma crown divergence at the TCMF1 break (max. 3.9% seq div) and 2.0 Ma intraspecific crown (max. 4.1% seq div)	—
Martin & Bermingham (2000)	***Pimelodella* fishes**	Strict (1.3% Myr^{-1})	—	R	**(T–Q):** †1.5 Ma (~2%, PB break), 2.8 Ma (~3.6%, WPI break), 1.7 Ma and 3.1 Ma (~2.3% and 4%, PNSA breaks), †923 ka (~1.2%, CPI break), and †1.0 Ma (~1.3%, CHIC break) divergences	**(T–Q):** †7.2 Ma (9.3%, EPI break, LCA–Magdalena) and †2.1 Ma (2.7%, EPI break, LCA–Rio Atrato)

Eizirik *et al.* (2001)	***Panthera onca* jaguars**	Strict (1.2% Myr^{-1}, coalescent)	R	—	***Q:** we calculated a ~625 ka crown for Central America based on max. K2P values.	**Overall (Q):** 137–830 ka crown, coalescence of mtDNA haplotypes (provides constraint on origin and expansion).
Marks *et al.* (2002)	***Glyphorynchus spirurus* birds**	—	—	—	***Q:** 1.0 Ma crown (2% LCA–Imerí seq div) and 1.3 Ma stem (2.6% seq div, PNSA break) divergences *Note*: PNSA split overlapped LCA and northern SA.	—
Perdices *et al.* (2002)	***Rhamdia* catfishes**	Strict (1.3–1.5% Myr^{-1})	—	R	**T–Q:** 2.95 Ma stem (*R. guatemalensis*, BEB break), 1.7–2.0 Ma *R. guatemalensis* and 2.5–2.9 Ma and *R. laticauda* "expansion" crowns, 5.6–6.5 trans-Andean clade crown *Using rates at left, we estimated BEB divergence as 33.3–38.5 ka.	**T:** 8.4 (7.7–8.8) Ma stem (cis–trans Andes)
Cavers *et al.* (2003)	***Cedrela odorata* plants**	—	—	—	—	—
Crawford (2003)	***Eleutherodactylus* frogs**	Strict (*ND2*: 0.96%; c-*myc*: 1.38–2.01 × 10^{-9} subs. s^{-1} y^{-1}, per lineage)	—	R	**T:** 10.0 Ma (7.63–12.3 Ma) *E. bransfordii–E. polyptychus* stem (L1 break), 11.8 Ma (9.32–13.7 Ma) *E. stejnegerianus–E. persimilis* stem (CVF break), 8.09 Ma (6.25–10.3 Ma) *E. stejnegerianus* crown	—
Dick *et al.* (2003)	***Symphonia globulifera* plants**	Strict (≥7 × 10^{-10} subs. s^{-1} y^{-1}) and Relaxed (NPRS, PL)	N R	R	**T–Q:** 7.18 Ma trans-Andean crown date, and approximately Quaternary within-LCA crown (their Fig. 2)	**T:** 15 Ma stem (cis–trans Andes)
González *et al.* (2003)	***Thryothorus nigricapillus* birds**	—	—	—	***Q:** 2.0–2.8 Ma stem (4–5.6% seq div) for BT break), 600 ka *T. castaneus* crown group (max. 1.2% seq div), 5.7 Ma *T. nigricapillus* complex–*T. semibadius* (11.4% seq div) divergences	***Q:** 1.6 Ma "*T. nigricapillus* group" crown (max. 3.2% seq div)

Hoffman & Baker (2003)	***Carollia* bats**	Strict (2.3–5% Myr^{-1})	—	—	***Q:** 1.2–2.61 Ma (~6% seq div) *C. castanea* AZUE break, 0.53–1.16 Ma EPI break (2.66% seq div, LCA–west Ecuador split within *C. castanea*), 0.8–1.8 Ma ND break (max. 4.1% seq div within *C. sowelli*). We calculated these values using the rates at left.	***T–Q:** 1.7–3.6 Ma *C. castanea* crown divergence (max. 8.3% seq div within *C. castanea*, cis–trans Andes). We calculated these values using pairwise rates at left (discussed in Hoffman *et al.*, 2003). **Overall (T–Q):** 1.0–4.5 Ma *Carollia* crown
Hoffman *et al.* (2003)	***Uroderma bilobatum* bats**	Strict (2.3–5% Myr^{-1})	—	—	***Q:** 260–565 ka stem (mean 1.3%, 1–1.6% net seq div) and 160–348 ka crown (mean 0.8%, 0.6–1% net seq div) divergences for the LCA chromosomal race. We calculated these values using rates at left.	***Q:** Despite no defined phylogeographic break between sister lineages across these regions, we estimated a hypothetical 696 ka divergence between LCA *versus* SA "races". **Overall (Q)**: 0.2–0.9 Ma stems, across "races"
Novick *et al.* (2003)	***Swietenia macrophylla* plants**	—	—	—	—	—
Zeh *et al.* (2003*a*)	***Acrocinus longimanus* beetles**	Strict (1.5% Myr^{-1})	—	—	**Q:** 540 ka crown divergence (mean 0.81% uncorrected seq div, Panama clades "A" and "B")	**Q:** 707 ka divergence (mean 1.06% uncorrected seq div, Panama *versus* Trinidad samples)
Zeh *et al.* (2003*b*)	***Cordylochernes scorpioides* pseudoscorpions**	Strict (‡2.3% Myr^{-1})	—	—	**T–Q:** 1.7–3.2 Ma crown divergence (6.2–7.8% K2P, Panama clades "A" *versus* "B")	**T:** 5.5–6.3 Ma stem (11.7–15.4% K2P)
Dick *et al.* (2004)	**Euglossine bees**	Strict (1.2–1.5% Myr^{-1})	—	R	**(Q):** 0.0–1.42 Ma crown divergences (mean 0–1.9% haplotype seq div) across 11 bee species (*E. cognata/mixta* omitted)	—
Cheviron *et al.* (2005)	***Lepidothrix coronata* birds**	Strict (1.6% Myr^{-1})	—	R	***Q:** 1.19 Ma crown (0–1.91% mean seq div, trans-Andean clade)	**T–Q:** 1.3–3.3 Ma divergence (2.3 ± 1.0 Ma, mean 3.04–5.53% uncorrected cis–trans Andes seq div) *Mean K2P seq div (3.61–7.32%) was equivalent to 2.26–4.58 Ma cis–trans Andes divergence
Hasbún *et al.* (2005)	***Ctenosaura quinquecarinata* lizards**	Strict (1.36–1.44% Myr^{-1})	—	—	***Q:** Using the 1.36% rate at left and numbers of changes across the tree (their Fig. 1), we estimated a ~1.17 Ma crown age for the LCA "Southern lineage" (Nicaragua and Costa Rica	**LCA–NCA (Q):** 1.47 Ma divergence (mean 2% seq div)

					samples).	
Morse & Farrell (2005)	***Stator limbatus* beetles**	Strict (‡2.3% Myr^{-1})	—	—	***T:** we calculated a 1.09 Ma crown for LCA clade ("Mesoamerican phylogroup") using corrected intra-lineage seq div in their manuscript and rate at left.	**T–Q:** 4.4 Ma divergence between North American and South American lineages (corrected pairwise seq div)
Perdices *et al.* (2005)	***Synbranchus marmoratus* swamp eels (fish)**	Strict (1.05–1.3% Myr^{-1}) and Relaxed (NPRS)	—	R	**T:** mean 8.5–10.7 Ma (11.2 ± 0.72% mean ± S.D. seq div) and range 7.7–12.4 Ma (10.1–13%) CVF/WPI breaks divergence; mean 22.0–27.4 Ma (28.8 ± 1.8%) and range 18.9–30.7 Ma (24.8–32.2%) crown, CPI/BT breaks; mean 58.5–73 Ma (76.7 ± 2.8%) and range 54.7–78.5 Ma (71.6–82.4%) stem, PERL break	**T:** ~8 Ma, constrained to approximate origin of the Rio Orinoco in northern SA ("SyNSA"–"SyMCA" + "SyLCA" seq div, cis–trans Andes)
Weigt *et al.* (2005)	***Engystomops pustulosus* frogs**	Strict (0.86–1.38% Myr^{-1}) and Relaxed (MULTI)	—	N	**T–Q:** mean ~9 Ma and range 4.1–15.9 Ma species crown (northern *versus* southern clades) ¶Other estimates: 2.1–9.5 Ma northern lineage crown, 2.3–10.8 Ma southern lineage crown, mean 4.4 and range 1.5–8.8 Ma stem (WPI break), 0.3–5.8 Ma "Western Panama" crown, mean 2.6 Ma and range 0.6–6.3 Ma "Central Panama" crown (also basal divergence for AZUE break), mean 1.1 and range ‡0.0–3.5 Ma (PERL break)	**T:** ¶8.0–15.8 Ma (*E. pustulosus versus* Amazonia species, thus cis–trans Andes)
Reeves & Bermingham (2006)	***Brycon*, *Bryconamericus*, *Cyphocharax*, and *Eretmobrycon* characid fishes**	Strict (3.6% Myr^{-1})	—	R	**Q:** 90–442 ka intra-lineage crowns (within-LCA, including Colombia) *Major-lineage crown estimates calculated using the rate at left and max. pairwise within-lineage K_S (HKY distances): western *Brycon*, ~458 ka (3.3%); eastern *Brycon*, 1.85 Ma	—

					(13.3%); *Bryconamericus* '*scleroparius'* species group, ~1.79 Ma (12.9%); *Bryconamericus* '*emperador*' species group, ~1.07 Ma (7.7%); *Cyphocharax magdalenae*, ~194 ka (1.4%). Divergence estimates for phylogeographic breaks within characid fishes calculated using the same method, except based on % seq div (pairwise, unless stated otherwise) roughly estimated by summing K_s values in their Fig. 2: 472 ka (~1.7% lineage^{-1}, SIXA), 417 ka and 458 ka (~1.5% lineage^{-1} and ~3.3% for *Bryconamericus* '*emperador*' species group and western *Brycon* CPI breaks, respectively), 1.17 Ma (~8.4%, BT), 444 ka (~1.6% lineage^{-1}, VALLE), 2.42 Ma (~8.7% lineage^{-1}, CHIC).	
Chacón *et al.* (2007)	***Phaseolus vulgaris* plants**	Relaxed (PL)	N	—	—	**Overall (Q):** mean 600 ka *P. vulgaris* crown (range 300–900 ka), mean 1.3 Ma stem (range 0.6–1.6 Ma, *P. vulgaris*–sister *Phaseolus* lineage split; NCA + LCA + Andes–Ecuador + Peru)
Crawford *et al.* (2007)	***Craugastor* frogs**	Strict (1.91% Myr^{-1}) and Relaxed (NPRS)	—	R	**T–Q:** 8.0–20.0 Ma *C. crassidigitus* crown, 5.5–17 Ma *C. talamancae* crown, 2.6–9.8 Ma *C. fitzingeri* crown, and 12–36 Ma *C. talamancae* + *C. crassidigitus* + *C. fitzingeri* stem. Data for most phylogeographic breaks not available, except 4.2–9.3 Ma WPI divergence was inferred. *Note*: WPI split not well supported within *C. fitzingeri*, but they inferred 0.9–4.7 Ma divergence across this area in this species.	—
Hagemann & Pröhl	***Oophaga pumilio***	—	—	—	***§Q:** 1.2–1.3 Ma (max. 2.4% seq div,	—

(2007)	**frogs**				CVF break), 50–56 ka (max. 0.1% seq div, POPA break), 450–500 ka (max. 0.9% seq div, L1 break), 2.0–2.2 Ma (max. 4% seq div, BDT1 break), 450–500 ka (max. 0.9% seq div, BDT2 break), 1.5–1.7 Ma (max. 3% seq div, ESCU break), 800–890 ka (max. 1.6% seq div, "N" clade), 2.2–2.4 Ma (max. 4.4% seq div, "S" clade) crowns. He also calculated an 800–890 ka mean (1.6% seq div) for the CVF break and a 2.25–2.5 Ma LCA stem for *Oophaga* sampled in this study (4.5% seq div, *O. pumilio* + *O. arboreus* + *O. speciosus*). Estimates above were based on data presented in the text or their Appendix.	
Nyári (2007)	***Schiffornis turdina* birds**	—	—	—	***Q:** we calculated a 400 ka stem (mean 0.8% uncorrected *p*-distance between clades 1 *versus* 3; LCA–Chocó SA, EPI break), and a 4.2 Ma LCA crown (8.4% uncorrected *p*-distance between phylogroups 1 *versus* 2, MRCA of all LCA samples).	***T–Q:** 2.6–2.7 Ma (5.1–5.4% eastern Panama–Amazon [cis–trans Andes] uncorrected *p*-distance between clades 2 *versus* 5 and 2 *versus* 6, respectively), 1.6 Ma (3.1–3.2% clades 1 *versus* 4 and 3 *versus* 4 [cis–trans Andes] uncorrected *p*-distance) ***Overall (T):** 4.8 Ma species crown (9.6% max. uncorrected *p*-distance), 7.5 Ma stem (15% *S. turdina–S. major p*-distance)
Bonaccorso *et al.* (2008)	***Chlorospingus ophthalmicus* birds**	—	—	—	***Q:** we calculated a mean 2.55 Ma and range 2.15–3.1 Ma divergence time for the Nicaraguan depression break (mean 5.1%, range 4.3–6.2% seq div); and 850 ka or 500 ka (mean 1% or max. 1.7% within-LCA seq div) provide maximum constraints for the TCMF2 break.	***T:** we estimated a 3.0–3.75 Ma (6–7.5% seq div) LCA–northern SA divergence date.
Dick & Heuertz (2008)	***Symphonia globulifera* plants**	—	—	—	See Dick *et al.* (2003).	See Dick *et al.* (2003).
Miller *et al.* (2008)	***Mionectes***	Strict (2%	N	N	—	**Q:** 1.6 Ma crown (PNSA break)

	***oleagineus* birds**	Myr^{-1}) and Relaxed (NPRS)	?	?		**T–Q:** 1.0 Ma and 200 ka cis–trans Andes divergences
Navarro-Sigüenza *et al.* (2008)	***Buarremon brunneinucha* birds**	—	—	—	***Q:** 250–400 ka crown (0.5–0.8% max. likelihood-corrected seq div)	***Q:** 1.7 Ma (3.4% max. likelihood-corrected seq div) ***NCA–SA (Q)**: 1.5 Ma (3% max. likelihood-corrected seq div)
Ornelas-García *et al.* (2008)	***Astyanax* fishes**	Strict (0.8–1.1% Myr^{-1}) and Relaxed (NPRS, PL)	N R	N R	***Q:** For *A. nicaraguensis*, we calculated a 0.7 Ma crown (1.4% seq div, ND break).	**T:** 6.4–12.2 Ma (LCA + NCA–SA).
Solomon *et al.* (2008)	***Atta* ants**	Strict (9.5 subs. s^{-1} Myr^{-1}; coalescent, IM)	N R ?	N R ?	— *Note*: shallow mtDNA gene tree structure suggested a most likely Quaternary MRCA for LCA samples	**T–Q:** 1.42 Ma (95% CIs: 0.82–4.893 Ma) basal divergence within *A. cephalotes* (including populations from Mexico, Belize, LCA, and SA)
Venegas-Anaya *et al.* (2008)	***Caiman crocodilus* crocodiles**	Strict (0.69–0.8% Myr^{-1})	R	—	**T–Q:** 1.6–1.8 Ma (this split supported both CVF and MOSQ breaks, and it also represents the basal divergence of all LCA samples)	**T:** 2.9–6.7 Ma **LCA–NCA (T–Q):** 2.5–2.9 Ma **Overall (T):** 5.7–6.7 Ma species crown (this node represents a cis–trans Andes split between a clade including NCA + LCA *versus* a clade of east-of-Andes SA populations)
Wang *et al.* (2008)	***Pristimantis ridens* frogs**	Strict (1.91% Myr^{-1}) and Relaxed (NPRS)	—	R	**T–Q:** 10–22 Ma species crown (MRCA, NCA–LCA), 0.66–5.3 Ma Costa Rica crown (all Costa Rica samples), 0.08–3.4 Ma Tilarán–Tortuguero lowlands divergence *Using the rate at left, we calculated ~2.77 Ma divergence across the MOSQ break, and ~733 ka divergence across the CPI break.	—
Wang & Shaffer (2008)	***Oophaga pumilio* frogs**	—	—	—	***§T–Q:** ~100–111 ka (0.2%, POPA break stem), 1.65–1.8 Ma (3.3%, Bocas del Toro–Puerto Viejo divergence, BDT1 break stem), 250–278 ka (0.5%,	—

					BDT3 break stem), 4.94 Ma (8.9%, ESCU break stem), and 4.95–5.5 Ma (9.9%, species crown) (all based on max. net corrected seq div)	
Weir *et al.* (2008)	***Chlorospingus ophthalmicus* birds**	Strict (2% Myr^{-1})	—	—	**Q:** mean ~0.9 Ma (~95% CIs: 0.7–1.0 Ma, TCMF2 break) and ~500 ka (ND break, their Fig. 4) divergences	**T–Q:** mean 4.7 Ma (95% CIs: 4.2–5.3, NCA + LCA *versus* SA) and 3.2 Ma (*C. o. honduratius–C. o. regionalis*, NCA *versus* LCA) divergences
Castoe *et al.* (2009)	**Middle American pitvipers (two *Atropoides* species, and *Cerrophidion godmani*)**	Relaxed (MULTI, PL)	N	N	**T:** mean 4.39 Ma (95% CIs: 3.06–6.03) *C. godmani* split across the ND break	**Overall (T–Q):** mean 8.56 Ma (95% CIs: 6.77–10.61) *A. picadoi* stem, mean ~1.0 Ma (95% CIs: 0.5–1.5) *A. mexicanus* crown, mean 3.05 Ma (95% CIs: 2.18–4.15, LCA–NCA divergence) *A. mexicanus* stem (to MRCA with other Mexican *Atropoides*), mean 5.73 Ma (95% CIs: 4.31–7.37) *C. godmani* crown, mean ~6.5 Ma (95% CIs: 5.1–8.1) *C. godmani* stem
Daza *et al.* (2009)	***Leptodeira septentrionalis* snakes**	Relaxed (BEAST, MULTI)	N	—	**Q:** mean 1.89 Ma (95% CIs: 1.05–2.82) CVF break divergence (*L. s. ornata* Atlantic *versus* Pacific Costa Rica population divergence, LCA lineage crown)	**T–Q:** mean 3.26 Ma (95% CIs: 2.12–4.6) *L. s. ornata* stem (*L. s. ornata–L. a. annulata* divergence, cis–trans Andes) **LCA–NCA (T):** mean 6.37 Ma (95% CIs: 3.86–9.22) *L. nigrofasciata* species crown (Mexico *versus* Costa Rica divergence)
Hynková *et al.* (2009)	***Boa constrictor* snakes**	Strict (2% Myr^{-1})	—	—	—	**T–Q:** ~0.95–3.05 Ma divergence within *imperator* clade "3" consistent with EPI break, 2.5–3.5 Ma basal divergence (5–7% uncorrected seq div) major clades mostly distributed in CA and SA
Jones & Johnson (2009)	***Xenophallus umbratilis* fish**	Strict (1–2% Myr^{-1}; coalescent, BEAST)	—	—	**T–Q:** 4.4 Ma (95% CIs: 2.3–7.1 Ma) crown, SJ2 break	—
Lee & Johnson (2009)	***Poecilia gillii* fish**	—	—	—	***Q:** ~2.4 Ma (4.8% max. uncorrected seq div)	—
Martins *et al.* (2009)	***Desmodus rotundus* bats**	Strict (2.6–5% Myr^{-1},	R	—	***Q:** We calculated a ~337–648 ka (1.68% seq div) divergence date for the	**Q:** 0.69–1.6 Ma stem ("CA"–"PAN" split, cis–trans Andes)

		Mt; 0.19% Myr^{-1}, based on the nDNA RAG1 gene; coalescent, MDIV)			LCA ("CA") clade crown using the rates at left.	
Robertson *et al.* (2009)	***Agalychnis callidryas* and *Dendropsophus ebraccatus* frogs**	—	—	—	***§T:** ~3.33–3.69 Ma (6.65%, *A. callidryas* FILA break) and ~3.82–4.24 Ma (7.63%, *D. ebraccatus* WPI break) crown divergences; ~245–272 ka (0.49%, *D. ebraccatus*) and ~2.03–2.25 Ma (4.05%, *A. callidryas*) "Northwestern CR"–"Northeastern CR" divergences; ~1.73–1.92 Ma (3.46%, *D. ebraccatus* FILA break) and ~3.07–3.41 Ma (6.14%, *A. callidryas* CVF break) "Northwestern CR"–"Southwestern CR" divergences; ~3.13–3.47 Ma (6.25%, *A. callidryas* WPI break) clade divergences against "Central Panama" samples; and ~1.11–1.23 Ma (2.21%) *D. ebraccatus* L1 break divergence	—
Robertson & Zamudio (2009)	***Agalychnis callidryas* frogs**	—	—	—	—	—
Streicher *et al.* (2009)	***Craugastor podiciferus* frogs**	Strict (0.22–0.75% Myr^{-1}; coalescent, BEAST) and Relaxed (BEAST)	N	N	**T:** overall ~14.4 Ma mean (range 4.70–42.11) *C. podiciferus* crown and overall ~12 Ma mean (range 3.61–35.63) crown for "Group I" clade west of Barú volcano (west of BARU break); ~3.7 Ma (range 0.6–17.69) *C.* sp. A (clade "G") crown divergence; and ~19.7 Ma mean (range 6.64–51.48) *C. podiciferus*–*C.* sp. A divergence	—
Vázquez-Miranda *et al.* (2009)	***Campylorhynchus rufinucha* birds**	Strict (1.6–2.7%	—	—	**T–Q:** 0.8–2.4 Ma crown (mean 4.1% uncorrected seq div "between the large	—

		Myr^{-1})			and small or medium forms") *Note:* Given the lack of resolution of the internal branching order of major lineages in this study, this estimate also applies as a maximum constraint on LCA–NCA divergence.	
Arbeláez-Cortés *et al.* (2010)	***Lepidocolaptes affinis* birds**	2.1% Myr^{-1} rate cited but not applied	—	—	***Q:** ~857 ka *L. affinis* crown (max. 1.8% seq div from uncorrected *p*-distances, ND break)	***Overall (T):** ~3.52 Ma *Lepidocolaptes* crown (7.4% max. seq div)
Brown *et al.* (2010)	***Oophaga pumilio* frogs**	—	—	—	—	—
Daza *et al.* (2010)	**Middle American snakes**	Strict (coalescent, msBayes) and Relaxed (BEAST)	N	N	**T–Q:** mean ~4.1 Ma (95% CIs: 2.4–5.9) *C. godmani* ND break, mean ~2.5 Ma (95% CIs: 1.4–3.6) *L. septentrionalis* CVF break, mean ~2.8 Ma (95% CIs: 1.7–3.9) *B. asper* CVF break, mean ~2.8 Ma (95% CIs: 1.3–3.2) *Lachesis* CVF break	**T–Q:** mean ~5.4 Ma (95% CIs: 3.4–7.9) *Lachesis* crown (*L. stenophrys–L. melanocephala* stem), mean ~4.0 Ma (95% CIs: 2.4–5.3) *L. s. ornata-L. a. annulata* split (cis–trans Andes, this is the stem of Costa Rican *L. septentrionalis*), overall 0.8–22.8 Ma (LCA–SA, seven splits) **LCA–NCA (T):** mean 6.8 Ma (max. 9.9 Ma) LCA–NCA split within *L. nigrofasciata* **Other (T):** max. 11.9 Ma (upper 95% CI) interspecific ND break within *Atropoides* (genus crown), mean ~4.2 Ma (95% CIs: 3.4–6.0) *C. godmani* crown, max. 5.4 Ma (upper 95% CI) interspecific break across CVF in *Porthidium*
Loaiza *et al.* (2010*a*)	***Anopheles albimanus* mosquitoes**	Strict (1.0×10^{-8} subs. s^{-1} y^{-1})	—	—	**Q:** ~200 ka (95% CIs: 165–235 ka) crown (0.4% net seq div)	—
Loaiza *et al.* (2010*b*)	***Anopheles albimanus* mosquitoes**	Strict (1.2×10^{-8} subs. s^{-1} y^{-1})	—	—	**Q:** 250 ka (95% CIs: 215–285) crown ("NCRWP" *versus* "CEPCO") divergence	**Q:** 827 ka (95% CIs: 702–952) EPI break divergence ("CEPCO" *versus* "PCOLE")

Serrano-Serrano *et al.* (2010)	***Phaseolus lunatus* plants**	Relaxed (PL)	N	—	**Q:** 413 ka (95% CIs: 350–530) stem for "MII" clade (which contained LCA samples); thus an at most middle Pleistocene basal divergence of LCA genes	—
Hauswaldt *et al.* (2011)	***Oophaga pumilio* frogs**	—	—	—	***§(T):** 3.85–4.28 Ma species crown (7.7% max. uncorrected intraspecific *p*-distance)	—
McCafferty *et al.* (2012)	***Andinoacara coeruleopunctatus* fish**	Relaxed (BEAST, 1.3% Myr^{-1})	—	R	**T–Q:** 3.4 Ma crown (95% CIs: 1.5–5.5) ***T–Q:** ~2.02 Ma WPI break divergence (max. ~2.63% seq div, "G"–"H" split; estimated from their Fig. 2 using rate at left)	**T:** 5.9 Ma (2.7–10.5 Ma), crown also containing Rio Atrato samples of *A. coeruleopunctatus* as well as *A. pulcher*
J.C. Bagley & J.B. Johnson (unpublished data)	**Poeciliid fishes**	Strict (0.68–1.52% Myr^{-1}; coalescent, BEAST)	—	—	**T–Q:** mean 3.66 Ma (95% CIs: 1.59–6.74) *Alfaro cultratus* crown, mean 4.1 Ma (95% CIs: 1.84–7.51) *Poecilia gillii* crown, mean 4.81 Ma (95% CIs: 2.31–8.81) *Xenophallus umbratilis* crown	**LCA–NCA (T):** ~6.0 Ma *A. cultratus–A. huberi* stem (ND break)

†Dates we estimated from sums of pairwise HKY distances presented in Martin & Bermingham's (2000) Fig. 2; other dates were estimated from their Fig. 3.

‡'Arthropod clock', global substitution rate estimated by regression using arthropod rate data from multiple sources (e.g. some biogeographic or fossil data; Brower, 1994).

¶Approximate maximum divergence times estimated by eye from the upper 95% CIs presented in Fig. 4 of Weigt *et al.* (2005).

§Dates estimated using the conventional 2% Myr^{-1} rate and a pairwise 1.8% Myr^{-1} 'frog clock' rate derived from Mongolian toads (Macey *et al.*, 1998).

?Rate calibration or topological constraint used, but basis for choice unclear or lacking, other than citing previous studies.

Table S3. Summary of species contributions to phylogeographic breaks recovered in LCA.

Break	Species (no. lineages containing break)	References	Total no.
1. ND, Nicaraguan depression break	†*Panthera onca* jaguars (1) *Carollia sowelli* bats (1) †*Chlorospingus ophthalmicus* birds (1) †*Astyanax nicaraguensis* fish (1) *Chlorospingus ophthalmicus* birds *Cerrophidion godmani* pitvipers (1) *Lepidocolaptes affinis* birds (1) *Cerrophidion godmani* snakes	Eizirik *et al.* (2001) Hoffman & Baker (2003) Bonaccorso *et al.* (2008) Ornelas-García *et al.* (2008) Weir *et al.* (2008) Castoe *et al.* (2009) Arbeláez-Cortés *et al.* (2010) Daza *et al.* (2010)	6
2. BEB, Rio Bebedero break	*Rhamdia guatemalensis* catfish (1)	Perdices *et al.* (2002)	1
3. SJ1, San Juan break 1	*Alfaro cultratus* fish (1)	J.C. Bagley & J.B. Johnson (unpublished data)	1
4. SJ2, San Juan break 2	*Xenophallus umbratilis* fish (1)	Jones & Johnson (2009)	1
5. CC, Central Cordillera break	*Orthogeomys cherriei* pocket gophers (1) *Craugastor podiciferus* frogs (1)	Demastes *et al.* (1996) Streicher *et al.* (2009)	2
6. SJ3, San Juan break 3	†*Alfaro cultratus* fishes (1)	J.C. Bagley & J.B. Johnson (unpublished data)	1
7. TCMF1, Talamanca Cordillera montane forest break 1	*Bolitoglossa pesrubra* salamanders (1) *Craugastor podiciferus* frogs (1)	García-París *et al.* (2000) Streicher *et al.* (2009)	2
8. SAV, Rio Savegre break	*Orthogeomys underwoodi* pocket gophers (1)	Demastes *et al.* (1996)	1
9. FILA, Fila Costeña break	*Orthogeomys underwoodi* pocket gophers (1) *Dendropsophus ebraccatus* frogs (1); †*Agalychnis callidryas* frogs (1) †*Agalychnis callidryas* frogs	Demastes *et al.* (1996) Robertson *et al.* (2009) Robertson & Zamudio (2009)	3
10. L1, Limón break	*Eleutherodactylus* dirt frogs (1) *Oophaga pumilio* frogs (1) *Dendropsophus ebraccatus* (1)	Crawford (2003) Hagemann & Pröhl (2007) Robertson *et al.* (2009)	3
11. CVF, Chorotega volcanic front break	*Lachesis* snakes (1) *Eleutherodactylus* dirt frogs (1) *Synbranchus* swamp eels (fishes; 1) *Oophaga* frogs (1) *Caiman crocidilus* crocodiles (1) *Leptodeira septentrionalis ornata* snakes (1)	Zamudio & Greene (1997) Crawford (2003) Perdices *et al.* (2005) Hagemann & Pröhl (2007) Venegas-Anaya *et al.* (2008) Daza *et al.* (2009)	8

	†*Agalychnis callidryas* frogs (1) †*Agalychnis callidryas* frogs *Bothrops asper* snakes (1); *Leptodeira s. ornata* snakes; *Lachesis* snakes	Robertson *et al.* (2009) Robertson & Zamudio (2009) Daza *et al.* (2010)	
12. PB, Piedras Blancas break	*Pimelodella chagresi* fishes (1)	Martin & Bermingham (2000)	1
13. TCMF2, Talamanca Cordillera montane forest break 2	*Chlorospingus ophthalmicus* birds (1)	Bonaccorso *et al.* (2008) Weir *et al.* (2008)	1
14. POPA, Popa Island–mainland break	*Craugastor talamancae* frogs (1) †*Oophaga pumilio* frogs (1)	Crawford *et al.* (2007) Wang & Shaffer (2008)	2
15. SIXA, Sixaola–Changuinola break	*Bryconamericus* tetras (fish; 1)	Reeves & Bermingham (2006)	1
16. BDT1, Bocas del Toro break 1	*Oophaga pumilio* frogs (1) *Oophaga pumilio* frogs	Hagemann & Pröhl (2007) Wang & Shaffer (2008)	1
17. BDT2, Bocas del Toro break 2	*Oophaga pumilio* frogs (1)	Hagemann & Pröhl (2007)	1
18. BDT3, Bocas del Toro break 3	*Oophaga pumilio* frogs (1)	Wang & Shaffer (2008)	1
19. ESCU, Escudo de Veraguas Island–mainland break	*Oophaga pumilio* frogs (1) *Oophaga pumilio* frogs	Hagemann & Pröhl (2007) Wang & Shaffer (2008)	1
20. BARU, Barú volcano break	*Craugastor* frogs (1)	Streicher *et al.* (2009)	1
21. WPI, Western Panama isthmus break	*Roeboides* tetras (fishes; 1); *Brachyhypopomus (Hypopomus) occidentalis* knifefishes (1); *Pimelodella chagresi* catfish (1) *Pimelodella chagresi* catfish *Synbranchus* swamp eels (fishes; 1) *Engystomops pustulosus* frogs (1) *Craugastor crassidigitus* frogs (1) *Dendropsophus ebraccatus* frogs (1); *Agalychnis callidryas* frogs (1) *Andinoacara coeruleopunctatus* fish (1)	Bermingham & Martin (1998) Martin & Bermingham (2000) Perdices *et al.* (2005) Weigt *et al.* (2005) Crawford *et al.* (2007) Robertson *et al.* (2009) McCafferty *et al.* (2012)	9
22. MOSQ, Mosquito Gulf break	*Caiman crocidilus* crocodiles (1) *Pristimantis ridens* frogs (1)	Venegas-Anaya *et al.* (2008) Wang *et al.* (2008)	2
23. AZUE, Azuero peninsula break	*Carollia castanea* bats (1) *Engystomops pustulosus* frogs (1)	Hoffman & Baker (2003) Weigt *et al.* (2005)	2
24. PNSA, Panama–northern South America continental divide break	*Roeboides occidentalis–R. guatemalensis* tetra fishes (1) *Pimelodella chagresi* catfish (1) *Pimelodella chagresi* catfish	Bermingham & Martin (1998) Martin & Bermingham (2000)	4

	Glyphorynchus spirurus birds (1) *Mionectes oleagineus* birds (1)	Marks *et al.* (2002) Miller *et al.* (2008)	
25. VALLE, El Valle volcano break	*Roeboides guatemalensis* tetras (fish; 1) *Brycon obscurus–B. petrosus* tetra fishes (1)	Bermingham & Martin (1998) Reeves & Bermingham (2006)	2
26. CPI, Central Panama isthmus break	*Pimelodella chagresi* catfish (1) *Pimelodella chagresi* catfish *Synbranchus* swamp eels (fishes; 1) *Engystomops pustulosus* frogs (1) *Brycon* (western) tetras (fishes; 1) *Bryconamericus* (eastern) tetras (fishes; 1) *Pristimantis ridens* frogs (1) *Leptodeira* cat-eyed snakes (1) and *Bothrops* pitvipers (1)	Bermingham & Martin (1998) Martin & Bermingham (2000) Perdices *et al.* (2005) Weigt *et al.* (2005) Reeves & Bermingham (2006) Wang *et al.* (2008) Daza *et al.* (2010)	8
27. PERL, Las Perlas Islands break	*Synbranchus* swamp eels (fishes; 1) *Engystomops pustulosus* frogs (1)	Perdices *et al.* (2005) Weigt *et al.* (2005)	2
28. CHIC, Rio Playón Chico basin break	*Pimelodella eutaenia* catfishes (1) *Brycon* (eastern) tetras (fishes; 1)	Martin & Bermingham (2000) Reeves & Bermingham (2006)	2
29. BT, Bayano–Tuira break	*Thryothorus nigricapillus* birds (1) *Synbranchus* swamp eels (fishes; 1) *Brycon* (eastern) tetras (fishes; 1)	González *et al.* (2003) Perdices *et al.* (2005) Reeves & Bermingham (2006)	3
30. SAPO, Sapo range break	*Craugastor raniformis* frogs (1)	Crawford *et al.* (2007)	1
31. EPI, eastern Panama isthmus	*Roeboides* tetras (fishes; 1); *Brachyhypopomus* (*Hypopomus*) knifefishes (1); *Pimelodella chagresi* catfishes (1) *Pimelodella chagresi* catfishes *Carollia castanea* bats (1) *Schiffornis turdina* birds (1) *Boa constrictor* snakes (1) †*Anopheles albimanus* mosquitoes (1)	Bermingham & Martin (1998) Martin & Bermingham (2000) Hoffman & Baker (2003) Nyári (2007) Hynková *et al.* (2009) Loaiza *et al.* (2010b)	7

†Weakly supported due to either (*i*) sharing of haplotypes between sister clades supporting the break [e.g. possibly reflecting gene flow in birds (Bonaccorso *et al.*, 2008); reflects some para-/polyphyly of areas], or because (*ii*) the break was supported by one dataset and not another within the same study (e.g. jaguars; Eizirik *et al.*, 2001). The number of lineages (no. lineages) containing a break refers to the number of parent lineages, i.e. nominal taxa or stem groups, containing the split.

Supporting Information for Bagley, J. C. & Johnson, J. B. (2014). Phylogeography and biogeography of the lower Central American Neotropics: diversification between two continents and between two seas. *Biological Reviews.*

APPENDIX S1: MATERIALS AND METHODS

(1) Literature survey

From 2011 to 2012, we compiled a database for this review based on *ISI Web of Knowledge*. All databases were searched (>8700 research journals; last accessed September 19, 2012). We looked for broad phylogeographical patterns; therefore, we considered studies with experimental designs including phylogenetic analyses and multiple (at least two) geographical sampling sites throughout a single lower Central American (LCA) species range to constitute phylogeographic analyses. J.C.B. read research article titles and abstracts retrieved from multiple searches (e.g. using terms "phylogeograph*" and "Central America") and pruned papers with insufficient sampling (zero or one LCA samples) or topics beyond the scope of this review. Topics considered beyond the scope of our review of the LCA phylogeography literature included other reviews, as well as studies focused primarily on interspecific phylogenetic relationships, population genetics, sympatric and ecological speciation, landscape genetics (primarily microsatellite-based studies of landscape effects), marine ecology and evolution, and virus or disease biology. We then obtained recently published articles from premier phylogeography journals *Molecular Ecology* and *Molecular Phylogenetics and Evolution*, which publish the most studies (Beheregaray, 2008). To this database, we added results from one of our unpublished comparative phylogeography studies of three LCA freshwater fish species (J.C. Bagley & J.B. Johnson, unpublished data).

Studies in this refined database were scored for taxonomic focus (binning taxa into seven groups discussed below), sampling design, and methods used. We recorded, or georeferenced, geographical coordinates of sampling localities from 38 studies with accessible sampling information. When recording details on genome sampling, we grouped ribosomal DNA with nuclear DNA (nDNA). For genetic marker sampling, we recorded the numbers and names of genes and loci sampled. Where necessary, we converted geographical coordinate data printed in the articles from other formats into decimal degrees, using the UTM & Lat/Lon Conversions tool of the Montana State University Research Coordination Network (http://www.rcn.montana.edu/resources/tools/coordinates.aspx/). However, for a select number of studies with legible maps and few sampling localities that did not report geographical

sampling coordinates (e.g. Demastes, Hafner & Hafner, 1996), J.C.B. georeferenced sites using the software program ArcMap 10 (Environmental Systems Research Institute, Redlands, CA) or DIVA-GIS (http://www.diva-gis.org/) based on locality details and map figures. We summarised sampling coverage by mapping available coordinate data from these studies in ArcMap 10.

(2) Evaluating phylogeographic patterns

We recorded the major qualitative and quantitative phylogeographic findings of each study. These included phylogenetic and network patterns. We also recorded information on the presence-absence, number, taxonomic composition, and geographical patterning of genetic lineages. Genetic lineages are synonymous with 'clades': reciprocally monophyletic groups that include an ancestor and all of its descendants. The first step in a comparative phylogeographical analysis is to identify the major lineages present in a set of molecular datasets from different species codistributed in an area (Zink, 2002). We identified major lineages in two principal ways. First, major lineages were identified from well-supported bifurcations or 'splits' in phylogenies (see Nei & Kumar, 2000, and Felsenstein, 2004, for reviews of phylogenetic methods). Phylogenetic splits with nonparametric bootstrap proportions ≥50 (usually, ≥70; Hillis & Bull, 1993) or Bayesian marginal posterior probabilities (BPP) ≥95 (Larget & Simon, 1999) were considered to provide superior internal support; however, nodes with Bayesian posterior probabilities 89–94.9 were considered positive support. An attempt was made to give priority to BPP values when available, because their interpretation is straightforward: Bayesian trees, thus BPPs, reflect posterior probabilities of clade support given the DNA data and model of DNA sequence evolution (Larget & Simon, 1999). Second, we also identified major lineages where genetic samples (e.g. haplotypes) split into two or more distinct networks. For example, clades are separated at 95% parsimony probability (connection limits) during statistical parsimony analysis in the commonly used software program TCS (Clement, Posada & Crandall, 2000), which infers relationships among a set of genetic samples reflecting phylogenetic and population genetic processes. Although the number of the mutational steps representing the maximum number of connection steps at 95% will vary among studies and datasets, clades are often distinct at ≥95% parsimony probability when distinguished by many unsampled mutations (e.g. >10) in parsimony networks of large datasets [e.g. with >100 samples and ≥600–1000 base pairs (bp) mitochondrial DNA (mtDNA) matrix length]. For example, in a recent TCS analysis a dataset of 601 bp of mtDNA from 355 individuals, including 55 haplotypes, of the Central American livebearing fish *Alfaro cultratus* had 10 maximum mutational connection steps at 95% parsimony probability (J.C. Bagley & J.B. Johnson, unpublished data). This means that a

connection of 10 mutational steps was the maximum justified between pairs of sequences by the parsimony criterion. If, however, all of the DNA sequence samples fell into a single clade during a network analysis (in parsimony or median-joining networks, with few mutations between them), then we concluded that no structure was present in the dataset (e.g. *Anopheles albimanus* mosquitoes; Loaiza *et al*., 2010*a*).

The second step in a comparative phylogeographical analysis is to determine whether the major lineages recovered are sister clades, and their geographical patterning (i.e. whether sister clades exhibit geographical overlap *versus* structuring in the form of phylogeographic breaks). In turn, phylogeographic breaks are most commonly interpreted as at least some mechanism of physical/genetic isolation having been at play within species (Zink, 2002). It is particularly important to determine whether sister clades in multiple species exhibit geographically congruent breaks i.e. comparative 'phylogeographical congruence' (Avise, 2000; Zink, 2002). This step is critical, because comparative phylogeographical congruence indicates a potentially shared history of responses to physiographic and palaeoclimatic changes that have occurred within an area (Bermingham & Martin, 1998; Avise, 2000; Arbogast & Kenagy, 2001; Zink, 2002). We identified phylogeographic breaks as phylogenetic breaks separating mostly distinct geographical lineages. Here, we allowed some para-/polyphyly of areas: phylogeographic breaks were still deemed present when up to ~5 samples/populations created an interdigitating pattern of areas between sister clades. We mapped the position of the (mostly unambiguous) phylogeographic breaks recovered within lineages. We evaluated correspondence of LCA phylogeographic breaks to geographical barriers marked by major physiographic features (Fig. 2) and biogeographical province boundaries (Fig. 3). We also tested for comparative phylogeographic congruence—whether congruent breaks across multiple taxa corresponded to predicted features/boundaries. Congruence across multiple codistributed taxa would provide evidence that general patterns of evolutionary diversification exist in LCA. Moreover, we examined phylogeographic congruence among plant *versus* animal species.

(3) Molecular clocks and divergence dating

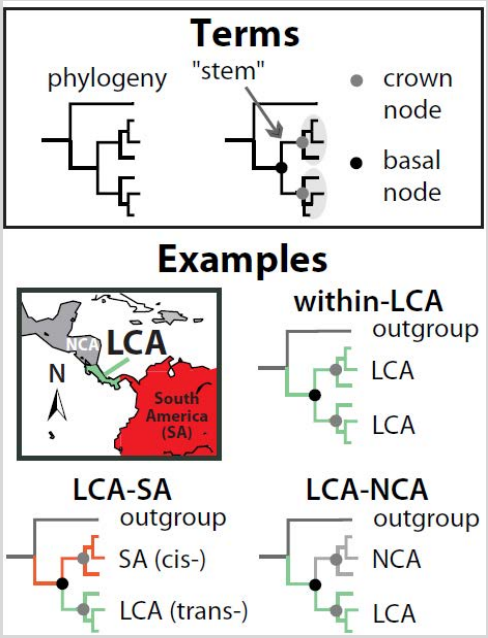

Fig. S1. Crown nodes and basal nodes, and examples of phylogeographical nodes of interest in the present study. The box on top illustrates that two sister clades share a common ancestor, represented by a 'basal' node. Basal, or ancestral, node divergences are synonymous with "stem node" divergences and we use these terms interchangeably. Furthermore, a 'crown' group refers to a collection of phylogenetic tip taxa (distal tips of phylogenetic trees, which in most phylogeography studies stem from individuals representing populations of living organisms) together with their ancestors back to their most recent common ancestor (MRCA) and all of that ancestor's descendants. Below the box, we provide several examples. The hypothetical within-LCA phylogeographical split shows the nodes of interest for which divergence dates might be recorded; where possible, we would have recorded or estimated crown and basal node dates. In the LCA–SA example, obtaining crown divergence date estimates as well as the stem age for the two cis–trans Andes clades would be desirable. The LCA–Nuclear Central America (NCA) example is similar to the LCA–SA split, but shows lineages diverged between Nicaragua and LCA, e.g. across the Nicaraguan depression (Fig. 2A).

To summarise temporal patterns of LCA assembly and diversification, including correlations between the timeline of phylogeographic diversification and earth history, we collated divergence time estimates from each study that inferred unambiguous phylogenetic splits. Phylogenetic splits with strong internal support values (see above) were emphasised, especially estimates for species 'crown' nodes and 'basal' or 'stem' nodes (Fig. S1), and splits corresponding to phylogeographic breaks. Three at times overlapping categories of divergence time estimates were collated: (*i*) splits within and among LCA lineages (Fig. S1; Within-LCA estimates); (*ii*) splits between LCA and South American lineages (Fig. S1; LCA–SA estimates; including splits representing the EPI break, Fig. 6B), interpreted as representing potential dispersal events followed by genetic drift in LCA *versus* South American areas; and (*iii*) splits between LCA and Nuclear Central American (NCA) lineages across the Nicaraguan depression or ND break (LCA–NCA estimates), interpreted as vicariance of montane taxa due to ND graben formation, and interpreted as dispersal followed by isolation in other taxa. Where available, we recorded divergence time estimates and their ranges or 95% confidence intervals (CIs) directly from the literature. If multiple calibrations were used, we recorded the ranges of resulting estimates to capture the full interval of the ages and obtain the oldest estimated dates.

Quantitative divergence time estimates were not always reported and, when they were, they were not always stated explicitly. Where clear divergence time (*T*) estimates were unavailable for nodes of interest (species crowns and stems, phylogenetic splits representing phylogeographic breaks), we applied a molecular clock to pairwise DNA sequence divergence

(*d*) data gleaned from the studies to roughly estimate *T* values. The conventional 2% Myr^{-1} rate of vertebrate mtDNA evolution (Brown, George & Wilson, 1979; Wilson *et al.*, 1985; see references in Miller *et al.*, 2008) was used unless other more suitable rates were available e.g. referenced in the studies. The 2% rate is equivalent to 1×10^{-8} substitutions $site^{-1}$ $year^{-1}$ (subs. s^{-1} y^{-1}) or 1% change Myr^{-1} $lineage^{-1}$, where site(s) refers to a single nucleotide position in a DNA sequence matrix (Brown *et al.*, 1979; Wilson *et al.*, 1985; also see references in Miller *et al.*, 2008). Assuming that genetic divergences are pairwise, the simplest calculation for the time (*T*) of divergence in Ma for a given lineage is calculated by taking d_{xy}, the mean number of nucleotide substitutions per site between two clades/populations of sequences, dividing by the nucleotide-substitution rate (μ; global rate of evolution), and dividing by 2, so that $T = (d_{xy}/\mu)/2$ (Nei, 1987). Given a per-lineage estimate of sequence divergence (*d*), divergence time is simply calculated as $T = d/\mu$. Where necessary, we estimated *T* for unambiguous phylogenetic splits of interest using these equations. We used these crude methods (*i*) for comparative purposes, and (*ii*) because a full re-analysis of the many datasets unearthed in our survey was beyond the scope of our review.

T estimates in the resulting database were interpreted as representing maximum or potential maximum divergence dates, thus they represented maximum initial constraints on the corresponding lineage divergence events. This is because, in all cases, divergence date estimates were derived from gene trees or gene divergences and not species trees (e.g. Maddison, 1997; Kubatko, Carstens & Knowles, 2009), and a well-known property of divergence date estimates calculated from gene trees/divergences is that they are likely to overestimate the actual timing of speciation or population divergence events (reviewed by Edwards & Beerli, 2000). Thus, regardless of the actual estimate obtained (mean, max., *etc*.), estimated divergence dates herein reflect potential maximum ages, and thus they allow us to place approximate upper bounds on the ages of the splits considered.

We ranked each *T* estimate as Tertiary (2.58–65.5 Ma) or Quaternary (0–2.58 Ma) and matched to its corresponding geological epoch (Gibbard *et al.*, 2009). We also matched each estimate to its corresponding phylogeographic break(s). We summarised the temporal and taxonomic structures of the divergence times, thus the inferred origins and diversification of modern LCA lineages, by creating time ranges for each split using the maximum *T* available and a 0 ka lower bound; for example, we set the upper bound of the range based on upper 95% CIs or estimate ranges where available, instead of means. We plotted these time ranges chronologically in order of ascending maximum *T* (e.g. Rull, 2007, 2008). Each time range graphically depicted

the estimated initial divergence events within each of the lineages sampled to date in LCA phylogeography studies (for which data were available). We did this separately for three geographical classes of *T* estimates: (*i*) within-LCA, (*ii*) LCA–SA, and (*iii*) LCA–NCA divergences (Fig. 7). However, some clades with LCA samples contained samples from other areas, e.g. South America; therefore, we summarised within-LCA time ranges a second time, factoring in this added level of complexity (Within-LCA max. divergence 2; Fig. 7). We estimated rates of lineage divergence, for comparative purposes, by dividing the number of lineage divergence time ranges whose maximum recorded values fell within each epoch by the length of that epoch in millions of years, for the Oligocene–Pleistocene (Gibbard *et al.*, 2009; Walker & Geissman, 2009), assuming that *T* estimates registered in the LCA literature reflected random lineage sampling. These rate calculations are likely to reflect errors in the structure of the divergence time estimates (which were derived from various analyses of different DNA regions, with different mutational tendencies, using various methods), and errors due to the finite sampling of lineages reflected in the literature database, thus the presence of unsampled lineages. Unsampled lineages produce spurious declines in speciation rates nearer to the present (documented in related analyses; see references and discussion in Rabosky & Lovette, 2008); however, our results show little sign of this trend, as highest speciation rates inferred were those of more recent epochs (see main text). Nonetheless, rate calculations allowed us to capture what the present LCA phylogeography literature says about relative rates and magnitude of differences in diversification in different geological epochs, across areas.

APPENDIX S2: LITERATURE SEARCH RESULTS

(1) Sampling and hypotheses testing in LCA phylogeography studies

In our database, the first phylogeography study published for LCA found contrasting patterns of phylogeographical structure (supporting two different breaks) and genetic structure (e.g. isolation by distance) in two species of *Orthogeomys* mammals with non-overlapping distributions in Costa Rica based on mtDNA restriction fragment data (Demastes *et al.*, 1996; Table S1). Subsequent studies revealed a growing and maturing field, yet the growth of the literature has not been exponential (Fig. S2) and disparities are evident in taxonomic focus, sampling and study design.

Taxonomically speaking, studies have largely (90%) focused on animals (Fig. S2), and herpetofauna have been the most popular before birds and fishes, plants and insects, mammals, and non-insect invertebrates. There have been no phylogeographic studies of inland fungi. These taxonomic biases mirror trends in unglaciated eastern North America (Avise, 1998; Soltis *et al.*, 2006) and partly are due to practical limitations. For example, plant chloroplast DNA (cpDNA) may evolve 10–100 times more slowly than animal mtDNA (see main text), whereas plant mtDNA evolve even slower and are often confounded by intramolecular recombination. Moreover, fungi studies are limited by low-rate evolving nDNA (e.g. nuclear ribosomal internal transcribed spacer, or 'ITS') and confounding factors such as mobile mtDNA genetic elements (Bergemann *et al.*, 2009). Bias against inland invertebrates might also reflect preference for marine invertebrates, e.g. living and fossil foraminifera and calcareous nannofossils (e.g. Coates *et al.*, 2004) and echinoids (e.g. Hickerson *et al.*, 2006*a*). Lower Central American phylogeography studies have also been biased towards single species (70.7%, $N = 41$ studies). While 29.3% of studies analyzed multiple (mean ≈ 4) ingroup lineages, only around 47.1–58.8% of these (13.8–17.2% of all studies) would be considered comparative phylogeography (*sensu* Arbogast & Kenagy, 2001; Avise, 2000; Bermingham & Martin, 1998; Bermingham & Moritz, 1998), as several studies analysed species complexes or did not evaluate phylogeographic congruence.

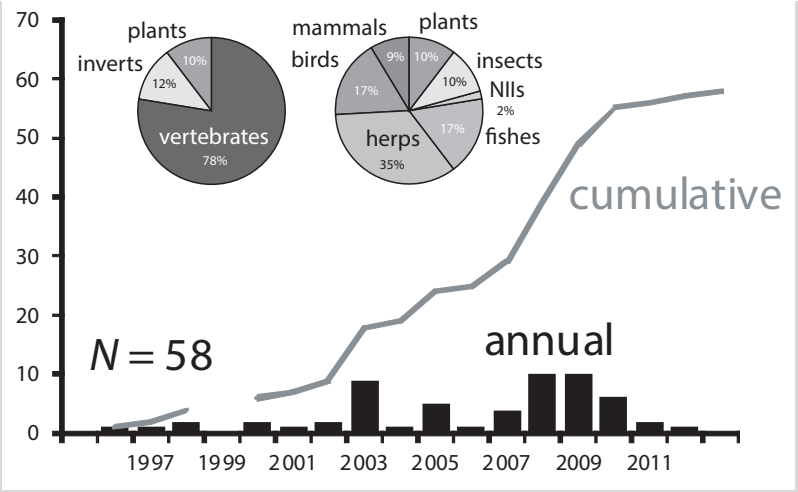

Fig. S2. Patterns of taxonomic sampling as well as annual *versus* cumulative growth of the literature base of LCA phylogeography studies to date. Pie charts show the percentage of the total number of studies filled by studies of each of three major taxonomic groups (left) or for a more detailed scheme including all seven taxonomic groups considered in this study (right). Growth has not been exponential.

Spatial sampling varied widely among studies. Broad-scale spatial sampling strategies were prevalent; for example, 12 studies represented in Fig. 6A sampled very few LCA populations (only 1–7 populations). Sampling has been spatially biased towards west-central and coastal Panama lowlands e.g. Bocas del Toro, reflecting access and focus on colonisation patterns of species from putative South American source populations. As a result, highland areas have been relatively poorly sampled. Variable sampling strategies reflect practical trade-offs as well as the diverse set of goals and objectives addressed by LCA phylogeographers. Yet previous sampling strategies have rendered post-colonisation diversification lesser known (Jones & Johnson, 2009). Within-LCA diversification is increasingly addressed at finer spatial scales (Robertson, Duryea & Zamudio, 2009; Robertson & Zamudio, 2009; Jones & Johnson, 2009; Streicher, Crawford & Edwards, 2009; J.C. Bagley & J.B. Johnson, unpublished data); yet more work clearly remains to be done to improve sampling. Of course, sampling grain and extent will continue to vary according to species distributions and ecology, as well as the different questions addressed in phylogeography studies; however, sampling greater percentages of species ranges at higher densities probably increases the likelihood of inferring phylogeographical breaks and is likely to uncover many additional insights over and beyond that attainable with only broad spatial sampling strategies.

In terms of character sampling, the kinds of genetic markers used in each study varied widely, and molecular marker choices are summarised for each study in Table S1. Fully 75.9% of studies used mtDNA or cpDNA alone, and a minority of studies analysed multi-locus organellar and nuclear DNA markers

(20.7%) or nDNA alone (3.4%). About two genes were amplified, on average, and mtDNA cytochrome *b* (cyt*b*; 44.8% of studies) and cytochrome oxidase I (COI; 31%) and nDNA recombination activating gene 1 (RAG1; 8.6%) were the most common sequences analysed.

(2) Divergence dating methods in LCA phylogeography

Molecular dating methods (reviewed by Rutschmann, 2006) have been common in LCA phylogeography, being used in 65.5% of studies (Table S2). Strict and relaxed molecular clocks were frequently employed together (~18.4% of the time); however, strict clocks were used in 51.7% of studies, and most (71.1%) studies with clocks used strict clocks alone. Molecular clocks increased in popularity, with historical precedence of strict clocks followed by an influx of alternatives, e.g. nonparametric rate smoothing (Sanderson, 1997) and relaxed clock methods (e.g. Drummond *et al.*, 2006, and references therein). Around 24% of LCA phylogeography studies calibrated their molecular clocks with fossils. Other calibrations relied on assumptions about absolute evolutionary rates, or palaeogeographic calibrations such as a Pliocene date for final LCA isthmus closure and trans-isthmian divergence [e.g. Martin & Bermingham (2000) used a rate calculated from geminate marine fishes that they published previously (Bermingham, McCafferty & Martin, 1997)] or the age of strata modern populations occur on. Although the best methods for estimating divergence times remain a topic of research and controversy (reviewed by Rutschmann, 2006), these estimates provide a key source of information for comparing the timing of lineage multiplication events and historical (e.g. geological) events to evaluate the validity of biogeographical hypotheses (e.g. Hoorn *et al.*, 2010).

Chapter 2: Testing for shared biogeographic history in the lower Central American freshwater fish assemblage using comparative phylogeography: concerted, independent, or multiple evolutionary responses?

Ecology and Evolution

Open Access

Testing for shared biogeographic history in the lower Central American freshwater fish assemblage using comparative phylogeography: concerted, independent, or multiple evolutionary responses?

Justin C. Bagley[1] & Jerald B. Johnson[1,2]

[1]Evolutionary Ecology Laboratories, Department of Biology, Brigham Young University, Provo, Utah 84602
[2]Monte L. Bean Life Science Museum, Brigham Young University, Provo, Utah 84602

Keywords
Comparative phylogeography, freshwater fishes, hierarchical approximate Bayesian computation, Neotropics, Nicaraguan depression, Poeciliidae.

Correspondence
Justin C. Bagley, Department of Biology, 401 WIDB (Widtsoe Building), Provo, UT 84602. Tel: +1 801 422 2203; Fax: +1 801 422 0090; E-mail: justin.bagley@byu.edu

Funding Information
Funded by Brigham Young University, including a Mentoring Environment Grant and a Graduate Studies Graduate Research Fellowship, and through stipend support from US National Science Foundation PIRE project OISE-PIRE 0530267.

Received: 23 January 2014; Revised: 13 March 2014; Accepted: 14 March 2014

Ecology and Evolution **2014; 4(9): 1686–1705**

doi: 10.1002/ece3.1058

Abstract

A central goal of comparative phylogeography is determining whether codistributed species experienced (1) concerted evolutionary responses to past geological and climatic events, indicated by congruent spatial and temporal patterns ("concerted-response hypothesis"); (2) independent responses, indicated by spatial incongruence ("independent-response hypothesis"); or (3) multiple responses ("multiple-response hypothesis"), indicated by spatial congruence but temporal incongruence ("pseudocongruence") or spatial and temporal incongruence ("pseudoincongruence"). We tested these competing hypotheses using DNA sequence data from three livebearing fish species codistributed in the Nicaraguan depression of Central America (*Alfaro cultratus*, *Poecilia gillii*, and *Xenophallus umbratilis*) that we predicted might display congruent responses due to co-occurrence in identical freshwater drainages. Spatial analyses recovered different subdivisions of genetic structure for each species, despite shared finer-scale breaks in northwestern Costa Rica (also supported by phylogenetic results). Isolation-with-migration models estimated incongruent timelines of among-region divergences, with *A. cultratus* and *Xenophallus* populations diverging over Miocene–mid-Pleistocene while *P. gillii* populations diverged over mid-late Pleistocene. Approximate Bayesian computation also lent substantial support to multiple discrete divergences over a model of simultaneous divergence across shared spatial breaks (e.g., Bayes factor $[B_{10}] = 4.303$ for Ψ [no. of divergences] > 1 vs. $\Psi = 1$). Thus, the data support phylogeographic pseudoincongruence consistent with the multiple-response hypothesis. Model comparisons also indicated incongruence in historical demography, for example, support for intraspecific late Pleistocene population growth was unique to *P. gillii*, despite evidence for finer-scale population expansions in the other taxa. Empirical tests for phylogeographic congruence indicate that multiple evolutionary responses to historical events have shaped the population structure of freshwater species codistributed within the complex landscapes in/around the Nicaraguan depression. Recent community assembly through different routes (i.e., different past distributions or colonization routes), and intrinsic ecological differences among species, has likely contributed to the unique phylogeographical patterns displayed by these Neotropical fishes.

Introduction

Comparative phylogeographic studies provide an important means of elucidating the relative influence of shared earth history events on contemporary biodiversity. By comparing spatial-genetic divergences, divergence times, gene flow, and population dynamics (e.g., N_e, effective population size) across multiple codistributed species,

comparative studies provide critical assessments of phylogeographical congruence, forming a basis for historical inferences (Bermingham and Martin 1998; Avise 2000; Arbogast and Kenagy 2001; Hickerson et al. 2010). Any of several outcomes may result, embodied by at least three general competing hypotheses with unique biogeographical implications. The first hypothesis, the "concerted-response hypothesis" predicts that codistributed species responded in lockstep fashion to geological and palaeoclimatic events, and correlated habitat shifts, within their overlapping distributions (Sullivan et al. 2000). Concerted responses are supported by congruent genetic breaks across taxa in space and time (Donoghue and Moore 2003), which should be common among ecologically and phylogenetically similar taxa due to codependence on similar habitats (Bermingham and Martin 1998; Feldman and Spicer 2006). Congruent responses point to causal factors underlying diversification and are consistent with long coassociations in local communities resulting in similar evolutionary trajectories (Avise 2000). Comparative phylogeographic congruence also predicts similar patterns in codistributed yet un-sampled taxa (Avise 2000; Sullivan et al. 2000).

One alternative to the concerted-response hypothesis, the "independent-response hypothesis", predicts codistributed species will bear genetic signatures of independent evolutionary responses to regional historical processes (Sullivan et al. 2000). This hypothesis is supported by phylogeographical incongruence in space, not time. This is because "incongruence" is identified when different spatial-genetic breaks derive from synchronous diversification (Cunningham and Collins 1994; Donoghue and Moore 2003). Incongruence thus occurs because species show different responses to the same earth history events or to the same deterministic biological factors (e.g., predation environment), and such independent but synchronous evolutionary trajectories are thought most likely to arise due to intrinsic differences in biological attributes among the species sampled (Cunningham and Collins 1994; Bermingham and Martin 1998; Avise 2000; Arbogast and Kenagy 2001; Donoghue and Moore 2003).

In contrast with the two competing hypotheses above, which propose temporal congruence, "pseudocongruence" and "pseudoincongruence" arise when spatial-genetic divergences are respectively congruent or incongruent but asynchronous, reflecting different responses correlated to different events (Cunningham and Collins 1994; Donoghue and Moore 2003). We refer to these scenarios defined by temporal incongruence as variations of a "multiple-response hypothesis". Complex pseudocongruent or pseudoincongruent patterns indicate little or no history of community coevolution; rather, different past distributions and recent community assembly, or stochastic dispersal or lineage sorting events, best explain such phylogeographic patterns (e.g., Donoghue and Moore 2003; Nielsen and Beaumont 2009). Biological factors also influence multiple-response scenarios, for example, to the extent that ecological differences determine species propensities or rates of dispersing into and becoming established in novel areas.

The lower Central American (LCA) isthmus is famous worldwide as an example of a land-bridge formation that has shaped continental Neotropical biotas by facilitating widespread dispersals, speciation, and extinctions in North and South America (Marshall et al. 1979; Stehli and Webb 1985). However, the LCA subcontinent is also increasingly appreciated as a system of highly endemic assemblages with interwoven histories of community assembly and species diversification (Bermingham and Martin 1998; Reeves and Bermingham 2006; Bagley and Johnson 2014). Much of our understanding of this history comes from studies of LCA species phylogeographic histories conducted in recent years (reviewed by Bagley and Johnson 2014). The LCA freshwater fish assemblage presents a particularly interesting model for phylogeography. This group is composed of >170 species from ecologically and morphologically diverse clades, including a wide representation of the teleost families Cichlidae (33 species) and Poeciliidae, that is, "livebearing fishes" (~28 species) (Bussing 1998; Smith and Bermingham 2005). Many species in this assemblage occupy identical river systems, making them well suited to test for concerted evolutionary responses, as they were potentially affected by the same past environmental changes (Bermingham and Martin 1998; Smith and Bermingham 2005). Previous studies indicate that freshwater fishes colonized LCA in multiple waves, mostly during the Pliocene–Pleistocene, but were more restricted than terrestrial taxa in tracking habitat disturbances as orogeny and drainage boundary formation progressed (Bermingham and Martin 1998; Streicher et al. 2009; Loaiza et al. 2010). This led to cryptic dispersals, vicariance, drainage isolation, and speciation within the assemblage (e.g., Martin and Bermingham 2000). For example, mtDNA phylogeography studies have recovered complex, but spatially correlated, genetic breaks across multiple fish species in Panama and between LCA and northwestern South America (e.g., Bermingham and Martin 1998; Martin and Bermingham 2000; Perdices et al. 2005; Reeves and Bermingham 2006), suggesting potential commonalities of evolutionary history (reviewed by Bagley and Johnson 2014). However, a limited subset of species phylogeographies have been inferred to date, compared with the total species diversity of the LCA fish assemblage. Thus, the question of whether LCA freshwater fish communities experienced common spatial, temporal, and demographic responses to past environmental changes, or not, remains an open one, particularly for northern LCA (Costa

Rica), where few phylogeography studies to date have sampled freshwater fishes (Bagley and Johnson 2014).

Here, we conduct a comparative analysis of three species from the LCA freshwater fish assemblage, in order to empirically evaluate predictions of the concerted-, independent-, and multiple-response hypotheses. Within LCA, we focus on the Nicaraguan depression (ND; Fig. 1A) and surrounding uplands of the San Juan biogeographical province (Bussing 1985, 1998; Smith and Bermingham 2005), where a unique history of factors likely influenced species evolution (discussed below) and several livebearing fish species ranges overlap (Bussing 1998). Phylogeography studies have been conducted on two livebearers from this area: *X. umbratilis* (Meek 1912) (monotypic, hereafter "*Xenophallus*"), and molly, *P. gillii* (Kner 1863). Within *Xenophallus*, Jones and Johnson (2009) discovered two deeply-diverged mitochondrial (mt) DNA lineages in the upland San Carlos basin, and the ND lowlands (Fig. 1) presumably correlated with sea-level dynamics; genetic diversification since the Pliocene (~4.5 Ma); and significant genetic partitioning by drainages. In contrast, Lee and Johnson (2009) found evidence for shallow haplotype divergences, limited among-region differentiation, and complex gene flow between populations of *P. gillii* from the same area. However, Bayesian demographic models published by Jones and Johnson (2009) and Lee and Johnson (2009) seemingly indicate overlapping Pleistocene-recent population bottleneck events in these taxa, assuming similar mtDNA substitution rates. These studies denote progress toward comparative perspectives on the evolutionary history of northern LCA freshwater communities. However, these species phylogeographies have never been rigorously compared using identical analyses. Thus, their degree of spatial and temporal congruence, and whether their genetic patterns represent general evolutionary patterns, remains unclear. By combining phylogeographical analyses of new DNA sequences from knife-edged livebearer, *A. cultratus* (Regan 1908), with analyses of existing data from *P. gillii* and *Xenophallus* populations codistributed in the ND, we test two predictions of the concerted-response hypothesis, against the independent- and multiple-response hypotheses. These predictions include (1) that these fishes should exhibit congruent spatial-genetic structuring and (2) that spatial subdivisions should be temporally congruent, due to co-occurrence in drainages correlated with LCA geomorphology (Marshall et al. 2003; Smith and Bermingham 2005).

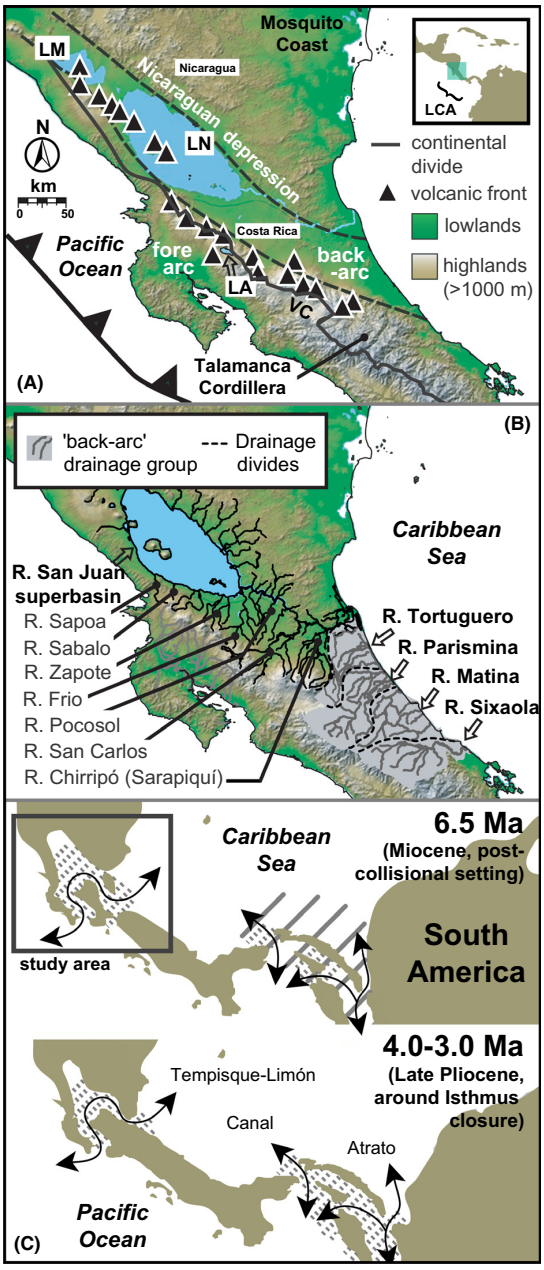

Figure 1. Study area. Major physiographic elements include: (A) the Nicaraguan depression (ND), Lakes Managua (LM), and Nicaragua (LN), the Rio San Juan superbasin (SJ), Lake Arenal (LA), the Valle Central (VC), and Quaternary stratovolcanoes (black triangles with white trim); inset map: position relative to greater Central American/Caribbean realm. Drainages subdivide into two a priori drainage groups (B): tributaries connected through freshwater in the San Juan superbasin (dark lines), and Caribbean "back-arc" drainages (light gray shading, dotted black borders). Palaeogeographic reconstructions (C; after Coates and Obando 1996; Coates et al. 2004) indicate that the study area was partly inundated by marine corridors (arrows) over Miocene–Pliocene (brown, land; diagonal lines, abyssal depths; stippling, neritic depths). Digital elevation layers were derived from NASA Shuttle Radar Topography Mission image PIA03364.

Additionally, coalescent theory predicts that genetic variation changes with historical N_e, for example, during population growth, decline, or range shifts (Wakeley 2000, 2003). Thus, to enhance our understanding of potential connections between historical biogeography and demography in these species, we also evaluate intraspecific DNA polymorphism and neutrality, and then compare historical-demographic responses among species.

Materials and Methods

Study area and sampling

The study area encompasses ~18,000 km^2 in and around the ND in southern Nicaragua and northern Costa Rica (Fig. 1). Here, our focal species co-occur in five major drainages at elevations ranging from 35 to 346 m. Based on geomorphology, these drainages subdivide into two a priori groups shown in Figure 1B: (1) the Rio San Juan superbasin, including lakes Managua and Nicaragua and tributaries to the southeast associated with Pliocene–Holocene formations of the Chorotega volcanic front (Fig. 1A); and (2) four Caribbean drainages along the LCA "back-arc" isolated from each other by saltwater, whose headwaters are associated with the Miocene–recent Talamanca Cordillera (Marshall 2007). The ND is a long, fault-bounded rift valley spanning El Salvador's Median Trough to the Tortuguero lowlands basin, Costa Rica. The ND formed by extensional forces at LCA's northern boundary, resulting in southeast–northwestward opening since 10 Ma (Mann et al. 2007; Funk et al. 2009). Sedimentary records show that, during Miocene–Pliocene high seas ~50–100 m above present sea level and even moderate late-Pliocene seas (Haq et al. 1987; Miller et al. 2005), a marine corridor inundated the ND until at least late Pliocene (Fig. 1C; Coates and Obando 1996; Coates et al. 2004). This corridor limited dispersal of many organisms between Nicaragua and Costa Rica, including freshwater fishes (Bussing 1976). Subsequently, the ND study area was above water (had surface freshwaters) by ~3.0–2.1 Ma (Coates and Obando 1996; Marshall et al. 2003). Proto-Chirripó drainage headwaters were redirected to the Pacific during creation of the Valle Central ~0.8–0.3 Ma (Fig. 1A; Marshall et al. 2003). The nearby Central Cordillera formed by Late Pleistocene, leaving active volcanoes amid drainage headwaters (Fig. 1A; Marshall et al. 2003), a source of periodic local extinctions. Because the steep Caribbean continental shelf restricts river anastomosis, over 100 m drops in sea level during Pleistocene glacial maxima (Lambeck et al. 2002) probably altered coastal freshwater connectivity minimally in this area (Smith and Bermingham 2005). Thus drainages modulated in length and elevation during Pleistocene

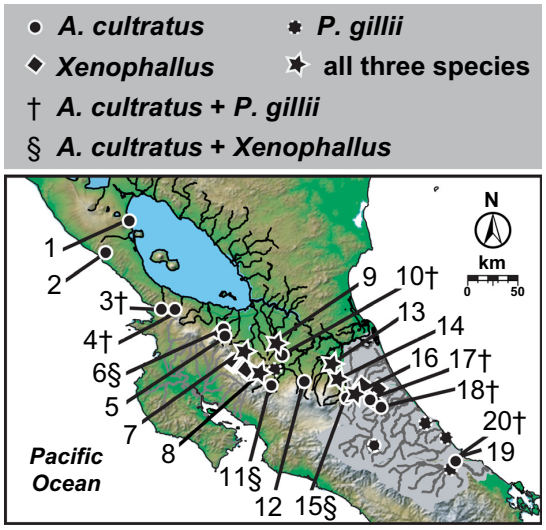

Figure 2. Geographical sampling localities. Sites where *Alfaro cultratus* (circles, 1–20), *Poecilia gillii* (dotted circles), and *Xenophallus umbratilis* (diamonds) were sampled are shown, including sites where we sampled all three species (stars) and combinations of two species (indicated by † and § symbols next to site numbers). Sites correspond to exact localities and sample sizes listed in Table S1.

sea-level cycles associated with glacial stages. Whereas Pleistocene–recent patterns of sea-level fall are widely agreed upon, the extent of eustatic sea-level highstands of the Quaternary remains debated among geologists; however, available data indicate large correlated spikes ≥20–30 m above present sea level ~2.4–1.8 Ma and 1.3 Ma (Miller et al. 2005) and ~550 ka (Hearty et al. 1999), that probably inundated LCA lowlands. Each of these Pliocene–Holocene environmental disturbances might have importantly shaped ND species phylogeographies, producing range fragmentation, upland isolation, or extinction-recolonization dynamics.

We sampled *A. cultratus* from 18 localities (sites 3–20 in Fig. 2) in four of the five study area drainages (Fig. 1B). We obtained samples of *Alfaro huberi* (Fowler 1923), the allopatric sister species to *A. cultratus*, from five sites in Honduras. Specimens were preserved in 95% ethanol in the field. We augmented data from these samples with published mtDNA cytochrome *b* (cyt*b*) sequences from *A. cultratus* [N = 3, sites 1–2; from Hrbek et al. (2007), Doadrio et al. (2009)] and codistributed *P. gillii* and *Xenophallus* populations (Jones and Johnson 2009; Lee and Johnson 2009), including 19 *P. gillii* localities and 23 *Xenophallus* localities (Fig. 2). There were six sites in the San Juan and Tortuguero drainages where we sampled all three species. Table S1 provides detailed sampling data and GenBank accession numbers. Outgroups used in the analyses below are described in Appendix S1.

Laboratory methods

We collected DNA sequence data from *A. cultratus* and *A. huberi* samples. After isolating DNA using the Qiagen DNeasy96 tissue protocol (Qiagen Sciences, Germantown, MD, USA), we amplified cyt*b* fragments for each sample by PCR using forward primer GLU31 (Unmack et al. 2009) and reverse primer HD (15680; Schmidt et al. 1998). Amplification and sequencing reactions, clean up, and sequence visualization followed Lee and Johnson (2009). We aligned mtDNA sequences manually in SEQUENCHER 4.8 (Gene Codes Corporation, Ann Arbor, MI, USA) and checked amino acid coding for errors (stop codons) while viewing electropherograms. We collapsed identical cyt*b* sequences into unique haplotypes using DnaSP 5.10 (Librado and Rozas 2009). We obtained a total of 355 *A. cultratus* and seven *A. huberi* sequences of a cyt*b* fragment 601 bp in length. Cyt*b* data encompassed 46 *A. cultratus* haplotypes, 37 *P. gillii* haplotypes (from 143 sequences; 1140 bp), and 29 *Xenophallus* haplotypes (from 131 sequences; 1140 bp), plus additional outgroup sequences.

Analyzing multiple unlinked loci can improve phylogeographical inferences, including population divergence-time and summary-statistics estimates (Edwards and Beerli 2000; Wakeley 2003), and provide perspective on putative sex-based asymmetries in gene flow and population structure (e.g., Avise 2000; Zink and Barrowclough 2008). Thus, we additionally screened nuclear ribosomal protein *S7* (*RPS7*; $N = 72$) introns 1 and 2 from multiple *A. cultratus* populations. Unfortunately, these sequences were uninformative in pilot analyses (e.g., star phylogeny, ~0.8% overall pairwise divergence), so we excluded them from our analyses. One limitation of basing our phylogeographical inferences on the matrilineal signal of mitochondrial DNA is that our results may not necessarily be congruent with patterns of population history in nuclear genomes. Despite such concerns, we are confident that our mtDNA analyses are appropriate for the questions we have addressed; for example, mtDNA is a robust indicator of population history and species histories, especially across multiple codistributed taxa, and thus has been a workhorse of comparative phylogeography (e.g., due to high information content, faster coalescence, etc., Avise 2000; Zink and Barrowclough 2008). Moreover, our use of mitochondrial markers makes our results comparable to several other LCA studies (e.g., Sullivan et al. 2000; Jones and Johnson 2009; Lee and Johnson 2009).

Genetic diversity and neutrality

We compared intraspecific genetic diversity levels across taxa by calculating segregating sites (S), haplotype diversity ($Hd \pm$ SE [standard error]), nucleotide diversity (π), and Watterson's (1975) θ_w (per site) for each locality and species using DnaSP. We calculated the same summary statistics in DnaSP for each population group (see BARRIER Results, Fig. 3). We also computed summary-statistic averages across localities within drainages. Patterns captured by these statistics may reflect sampling differences, for example, denser within-locality sampling in *A. cultratus*; however, Hd and π are less sensitive to such sampling effects (Li 1997). We assessed selective neutrality of each cyt*b* dataset–an assumption of most of our analyses–using Hudson-Kreitman-Aguadé (HKA; Hudson et al. 1987) tests, testing significance using 10^4 coalescent simulations in DnaSP; these tests used outgroups identical to phylogenetic outgroups below.

Spatial patterns

To test for spatial-genetic congruence, as predicted under the concerted-response hypothesis, we evaluated genetic structuring and breaks across the study area while taking spatial sampling patterns into account, but without prior knowledge of population structure or genetic barriers. First, we used the simulated annealing algorithm implemented in SAMOVA 1.0 (Dupanloup et al. 2002) to define genetically homogeneous, maximally differentiated spatial population clusters (K). We modeled $K = 2$–10 groups, drawing from 100 initial conditions, and noted fixation index (Φ_{CT}) trends. Second, we identified genetic barriers among populations using BARRIER 2.2 (Manni et al. 2004a,b). In BARRIER, we laid Delaunay triangulation networks over sampling sites (based on Voroni tessellation). We then used Monmonier's (1973) algorithm to sequentially identify genetic "barriers" as locations of maximum pairwise Tamura and Nei (Tamura and Nei 1993; TrN) genetic distances between localities across each network, calculated in ARLEQUIN 3.5 (Excoffier and Lischer 2010; 1000 nonparametric permutations). We assessed relative support for barriers by calculating bootstrap proportions (BP) from 100 bootstrapped barriers, generated by supplying BARRIER with bootstrapped TrN distance matrices (resampling the original datasets within populations, using PopTools; Hood 2008); we considered it strong support when BP $\geq$ 50. We did not apply this procedure to *Xenophallus*, because low within-site genetic diversity rendered bootstrapping ineffective. We independently tested spatial configurations inferred in SAMOVA and BARRIER using analyses of molecular variance (AMOVA) performed in ARLEQUIN (1000 nonparametric permutations). When faced with isolation-by-distance, SAMOVA and Monmonier's algorithm are more likely to misidentify populations and genetic barriers between them (Dupanloup et al. 2002). Thus, we tested correspondence between linearized genetic distance $[F_{ST}/(1–F_{ST})]$ and natural log-transformed geographic

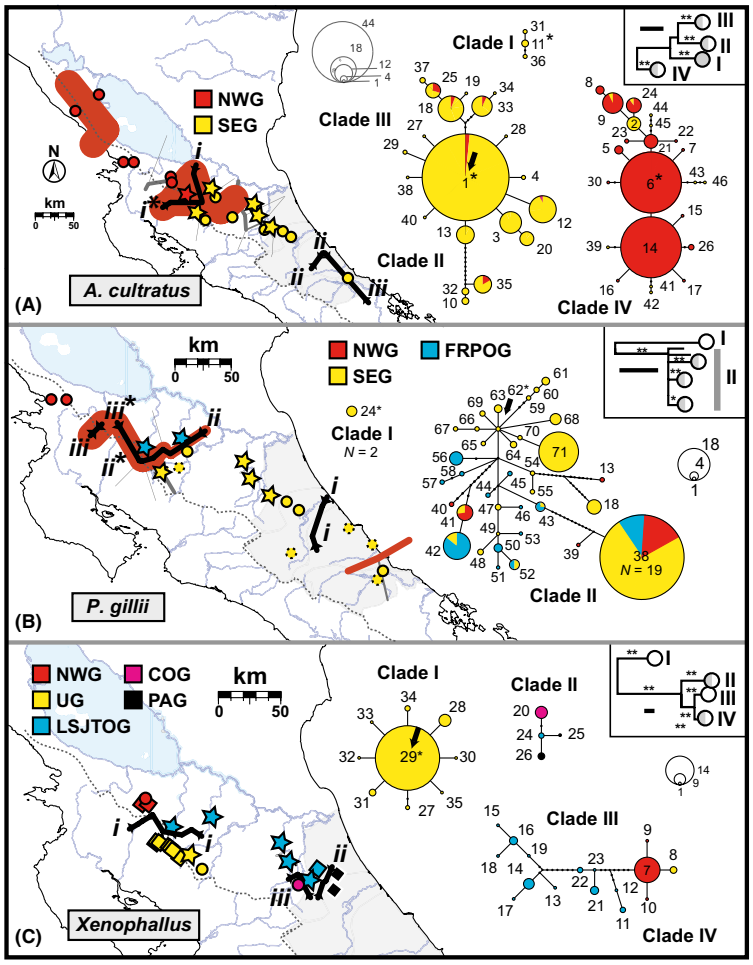

Figure 3. Incongruent spatial-genetic structuring among Nicaraguan depression livebearing fish species, based on mtDNA *cytb* variation. Solid black lines indicate genetic barriers (i–iii) delimiting distinct population groups (represented with different colors and abbreviations; population groups are described in the text) inferred using Monmonier's algorithm in BARRIER. Asterisks indicate significant barriers, based on bootstrapping; and maximum TrN genetic distances across each break are given as percentages. Table 1 presents summary-statistics and neutrality tests for these groups. Corresponding *cytb* parsimony networks are also shown presenting haplotypes as network circles, scaled according to their frequency and colored to show proportions of their distributions in each population group. Networks were separated based on a 95% parsimony criterion. Phylogenetic relationships are shown within each species map (inset boxes), with nodal support (BP: *50–70, **>70) and a scale bar (0.01 subs/site) for the simplified maximum-likelihood tree; tip circles summarize clade geography with respect to drainage groups in Figure 1B (dark gray, San Juan drainage; light gray, back-arc). Haplotype numbers correspond to labels used in (A) Table S1, (B) Lee and Johnson (2009), and (C) Jones and Johnson (2009).

distance between localities using standard regression, and Mantel tests (Mantel 1967) with significance tested using 10^4 permutations in PASSAGE 2 (Rosenberg and Anderson 2011). Details of SAMOVA and BARRIER analyses and interpretation are given in Appendix S2.

We also tested for congruent hierarchical genetic partitioning among San Juan basin tributaries and drainage groups using two a priori biogeographic AMOVAs. These AMOVAs were similar to those employed by Jones and Johnson (2009), except we grouped localities using drainage and drainage groups as defined in Figure 1B. We qualitatively tested for similar population groups (SAMOVA, BARRIER), genetic barriers between major drainages (BARRIER, AMOVAS), and among-drainage partitioning across taxa (AMOVAS) to identify shared effects of drainage boundaries as historical barriers to gene flow.

We compared phylogenetic relationships (haplotype gene trees) and nodal support among *cytb* haplotypes inferred for each focal species using maximum likelihood tree searches and bootstrap (500 pseudoreplicates) searches in GARLI 0.97 (Zwickl 2006). Likelihood analyses relied on substitution models (Table S2) selected using a decision theory algorithm, DT-ModSel (Minin et al. 2003), and partitioned data by codon position, ([1 + 2], 3). We independently inferred relationships among phylogenetic clades using statistical parsimony analyses in TCS 1.21 (Clement et al. 2000; 95% Connection Limit). We estimated sequence divergence over haplotype pairs among clades as pairwise maximum composite likelihood means in MEGA5 (Tamura et al. 2011).

Temporal patterns

We evaluated temporal congruence, the second prediction of the concerted-response hypothesis, by using the Bayesian coalescent dating approach implemented in IMa2

(Hey 2010) to estimate divergence times (t) among adjacent population groups from BARRIER. While we were mainly interested in estimating t, IMa2 also estimates population migration rates (m_1, m_2) and sizes of current (θ_1, θ_2) and ancestral populations (θ_A) using Hey and Nielsen's (2004) "isolation-with-migration" model. We conducted several pilot runs to estimate appropriate Markov chain Monte Carlo (MCMC) sampling chain lengths and priors. Subsequently, we ran three final runs per population pair starting from different random seeds, with 10 chains each. After logging 10^6 states discarded as "burn-in", we ensured chain mixing and convergence, judged by (1) ≥10% update rates for t, (2) appropriate chain-swapping rates, and (3) runs converging on similar parameter estimates. Fossil data and species-specific substitution rates were unavailable to us, thus we specified uniform mutation rate (μ) priors spanning lower and upper per-lineage mutation rates published for teleost fish protein-coding mtDNA, 1.7×10^{-9} and 1.4×10^{-8} substitutions/site/year (subs/site/year) [refs. in Waters and Burridge (1999), Burridge et al. (2008)]. See Appendix S3 for details of our IMa2 runs, for example, prior settings. Resulting t estimates were converted to absolute time (T_{div}) using the equation $T_{div} = t/\mu k$ (where k = sequence length), assuming species generation times equivalent to 1 year/generation (Winemiller 1993). To cover a range of possible mutation rates, we estimated T_{div} twice per population pair, setting μ equal to (1) the standard 2% rate (1.0×10^{-8} subs/site/year, per lineage) for vertebrate mtDNA (Brown et al. 1979; Wilson et al. 1985) and (2) a "slower" 0.9% pairwise rate (4.5×10^{-9} subs/site/year, per lineage) estimated for trout species mtDNA (Salmonidae; Martin and Palumbi 1993) that has previously been applied to higher-order teleosts (e.g., Waters and Burridge 1999).

We further tested the temporal congruence of shared genealogical breaks within northwestern Costa Rica (see Results) using hierarchical approximate Bayesian computation as implemented in the bioinformatics pipeline, MTML-msBayes (Huang et al. 2011). Using MTML-msBayes, we tested for simultaneous divergence among three population-group pairs diverged in this area under a finite sites coalescent model, allowing lineages to diverge, experience different migration patterns, and change population sizes (θ) independently while accounting for coalescent gene-tree stochasticity (Huang et al. 2011; refs. therein). After calculating a vector of observed summary statistics for each population pair, we used coalescent simulations to generate 5×10^6 simulated DNA datasets for three population pairs. Simulations assumed no migration or recombination, consistent with general IMa2 and neutrality test results (see Results). We generated hyper-posteriors for the mtDNA, representing 1000 random draws from the joint posterior distribution, by comparing the observed versus simulated summary statistics vectors using the pipeline's standard rejection/acceptance algorithm. Posterior estimates of the number of discrete co-divergences (Ψ) were obtained via polychotomous regression; posterior estimates of population divergence time ($E[\tau]$; units of μ/generation) and Ω (dispersion index representing the ratio of variance to the mean divergence times across Y taxon pairs) were obtained by local linear regression (Beaumont et al. 2002). To evaluate the "weight of evidence" in favor of simultaneous divergence, we calculated Bayes factors (B_{10}) to compare the level of posterior support for simultaneous versus nonsimultaneous divergence ($\Psi = 1$ vs. $\Psi > 1$; and $\Omega = 0.05$ vs. $\Omega > 0.05$), using Jeffreys' (1961) criteria for B_{10} "weight of evidence". We also used B_{10} values to evaluate support for continuous divergence ($\Psi = 3$ vs. $\Psi < 3$). Bayes factor calculations accounted for prior support for each hypothesis. To explore our data, we conducted multiple msBayes runs across a range of prior values (upper $\theta = 0.005$–0.05; upper ancestral $\theta = 0.25$–0.5; $Nm = 0$) to evaluate the effects of the prior on the models, and we conducted Bayes factor hypotheses testing using each model.

Historical demographic patterns

We qualitatively evaluated historical demographic congruence by comparing estimates of past population size fluctuations through time captured using the Bayesian skyline plot method (Drummond et al. 2005) implemented in BEAST 1.7.5 (Drummond et al. 2012). Pilot runs (MCMC=10^6) showed that marginal relaxed clock standard deviations ("ucld.stdev" parameter) clumped at zero, indicating highly clock-like data. Therefore, we conducted Bayesian skyline model runs using strict clocks (MCMC=2×10^8; burn-in=2×10^7; "Piecewise-constant" skyline model; "Coalescent: Constant Size" tree priors). We specified uniform priors spanning teleost mtDNA mutation rates (see IMa2 methods) and substitution models selected by DT-ModSel (Table S2). We partitioned sites by codon position ([1 + 2], 3), unlinking parameters across subsets. We calculated posterior distributions of $N_e\tau$ through time, and node ages (t_{MRCA}s), and their 95% highest posterior densities (HPD) using TRACER 1.5 (Rambaut and Drummond 2009). We then tested whether Bayesian skyline plots were more appropriate than constant-size (Hudson 1990), exponential growth, or logistic growth models run with equivalent priors. The best model had the highest smoothed marginal likelihood ($\ln L \pm$ SE, 1000 bootstrap pseudoreplicates) and was compared to alternatives using log Bayes factors ($\log_{10} B_{10}$) calculated in TRACER (Suchard et al. 2001), and established support criteria (Drummond et al.

2005). To complement our BEAST analyses, we estimated Ramos-Onsins and Rozas' (2002) R_2 and Tajima's (1989) D neutrality statistics and their 95% confidence intervals using coalescent simulations in DnaSP (10^4 replicates). To distinguish population expansions from purifying or positive natural selection, we tested for neutrality using similar simulations of Fay and Wu's (2000) H statistic. Agreement across methods (support for skylines showing expansions; positive, significant R_2; negative, significant D; and nonsignificant H estimates) and taxa would provide strong evidence for congruent past population dynamics.

Results

Genetic diversity and neutrality

Among localities, mtDNA genetic diversity was highly spatially variable: π ranged from 0 to 0.0244, *Hd* ranged 0 to 1, and θ_w ranged 0 to 25.92 (Table S1). When averaged over local subpopulations, intraspecific *Hd* (range = 0.384–0.539; cross-species mean≈0.473), π (range = 0.0004–0.0076), and θ_w (range = 0.733–5.586) varied from low to moderate, but were much higher in *A. cultratus* and *P. gillii* than *X. umbratilis* (Table S1). MtDNA genetic diversity also varied greatly among population groups (see below): π ranged from 0.541 to 7.505, *Hd* ranged 0 to 0.944, and θ_w ranged 0.542 to 12.074 (Table 1). However, diversity peaked in groups located in the southeast of the study area in all three species. Haplotype diversity and θ_w peaked in the southeastern *P. gillii* population group, the southeastern *A. cultratus* group displayed the highest π, and *Xenophallus* diversity peaked in the lower San Juan-Tortuguero group (Table 1; see group designations below). Likewise, genetic diversity was higher within back-arc drainages than the San Juan, although local *A. cultratus* subpopulations had slightly higher mean intradrainage diversity relative to the other taxa (Table S3). Cyt*b* variation met expectations of neutral evolution in all three species (HKA test: *A. cultratus* $\chi^2 = 0.004$, $P = 0.951$; *P. gillii* $\chi^2 = 0.256$, $P = 0.613$; *Xenophallus* $\chi^2 = 0.244$, $P = 0.622$).

Overall spatial incongruence

Overall, mtDNA analyses recovered incongruent spatial-genetic structuring among the three focal species in this study. The best SAMOVA grouping schemes partitioned the sampling area into two *A. cultratus* groups ($K = 2$, $\Phi_{CT} = 0.709$), six *P. gillii* groups ($K = 6$, $\Phi_{CT} = 0.543$), and seven *Xenophallus* groups ($K = 7$, $\Phi_{CT} = 0.973$), indicating differing numbers and positions of spatial subdivisions between homogeneous populations within species (Figs. S1, S2, and Appendix S2). Genetic barriers detected using Monmonier's algorithm were similar but not

Table 1. Summary statistics and neutrality test results for *Alfaro cultratus*, *Poecilia gillii*, and *Xenophallus umbratilis* and homogeneous populations inferred within species using BARRIER.

	A. cultratus			*P. gillii*				*Xenophallus*					
Parameter	NWG	SEG	All	NWG	FRPOG	SEG	All	NWG	UG	LSJTOG	COG	PAG	All
N	147	208	355	16	24	103	143	24	61	30	8	8	131
No. localities	8	12	20	2	2	15	19	2	13	6	1	2	24
S	44	52	58	6	11	93	98	4	9	20	0	2	96
h	20	33	46	5	8	29	37	4	9	12	1	3	29
Hd (±SE)	0.811 (±0.021)	0.880 (±0.016)	0.924 (±0.007)	0.717 (±0.095)	0.707 (±0.082)	0.939 (±0.010)	0.946 (±0.007)	0.424 (±0.113)	0.404 (±0.079)	0.892 (±0.031)	0.000 (±0.000)	0.679 (±0.122)	0.843 (±0.027)
π	0.0063	0.0125	0.0217	0.0018	0.0012	0.0101	0.0093	0.0005	0.0005	0.0061	NA	0.0007	0.0286
θ_w	0.0132	0.0146	0.0150	0.0016	0.0026	0.00157	0.0155	0.0009	0.0017	0.0044	NA	0.0007	0.0155
R_2	**0.086***	**0.081***	**0.075***	**0.155***	**0.139***	**0.089***	**0.087***	0.156 ns	**0.116***	0.119 ns	NA	0.241 ns	**0.088***
Tajima's D	−0.082 ns	−0.112 ns	−0.119 ns	−0.058 ns	−0.046 ns	−0.098 ns	−0.117 ns	−0.027 ns	−0.017 ns	−0.080 ns	NA	−0.018 ns	−0.110 ns
Fay and Wu's H	−0.032 ns	−0.034 ns	−0.027 ns	0.028 ns	−0.009 ns	−0.064 ns	−0.063 ns	0.002 ns	−0.001 ns	−0.076 ns	NA	0.018 ns	0.323 ns

NA, not available; ns, not significant.
Significant results are shown in bold (*$P < 0.01$).

Table 2. AMOVA tests of models reflecting the best grouping schemes inferred using SAMOVA and BARRIER, plus two a priori biogeographical hypotheses of hierarchical genetic structuring within/among drainages.

	Source of variation (percentage)			Φ-statistics		
Comparison (number of groups)	Among groups	Among subpopulations, within groups	Within subpopulations	Φ_{CT}	Φ_{SC}	Φ_{ST}
Alfaro cultratus						
SAMOVA / BARRIER model (K = 2)	70.9	9.0	20.1	**0.71****	0.31**	0.80**
1. Rio San Juan vs. back-arc drainages (2)	17.7	54.9	27.4	0.18 ns	0.67**	0.73**
2. Rio San Juan trib. by trib. (6)	15.8	58.7	25.5	0.16 ns	0.70**	0.75**
Poecilia gillii						
SAMOVA model (K = 6)	54.3	13.7	32.0	**0.54****	0.30**	0.68**
BARRIER model (K = 3)	17.8	47.4	34.8	**0.18***	0.58**	0.65**
1. Rio San Juan vs. back-arc drainages (2)	22.1	44.0	33.9	**0.22****	0.56**	0.66**
2. Rio San Juan trib. by trib. (6)	−24.1	79.0	45.0	−0.24 ns	0.55**	0.64**
Xenophallus						
SAMOVA model (K = 7)	97.3	1.4	1.3	**0.97****	0.52**	0.99**
BARRIER model (K = 5)	95.0	4.1	0.9	**0.95****	0.82**	0.99**
1. Rio San Juan vs. back-arc drainages (2)	22.7	76.2	1.2	0.23 ns	0.99*	0.99*
2. Rio San Juan trib. by trib. (5)	93.0	6.0	1.0	**0.93****	0.86**	0.99**

ns, not significant.

"Comparisons" are models (trib., tributary) and "number of groups" corresponds to population groups compared under each model. Sources of variation are percentages representing hierarchical partitioning of diversity across levels (negative percentages are interpreted as not significantly different from zero), and Φ-statistics range from 0, indicating no genetic structure, to 1, indicating complete isolation. Φ_{CT} is the correlation of random haplotypes within a group relative to the whole dataset (i.e., among groups), with significant results bolded (see Appendix S2 for further details).

*P < 0.05; **P < 0.01.

identical to SAMOVA results; for example, BARRIER yielded two *A. cultratus* groups that were identical to those from SAMOVA, but fewer *P. gillii* and *Xenophallus* groups (Fig. 3 and Table 2). Still, grouping schemes resulting from both methods yielded matching barriers within each species, and were supported by independent AMOVAs (Table 2). Together, these results suggested that a significant barrier to gene flow (mean BP = 88.2%, across five segments) divided *A. cultratus* range into northwestern (NWG) and southeastern groups (SEG) in the lowlands between Rio Frio and Rio Pocosol flanked by Tenorio and Arenal volcanoes (Fig. 3A). This barrier separated Lake Nicaragua tributaries (e.g., Rio Frio) from others, including the nearby San Carlos basin where we sampled all three taxa at site 8 (Fig. 2 and Table S1). In *P. gillii*, a similar well-supported break (mean BP = 79.3%, six segments) occurred just east of Rio Pocosol, dividing Frio and Pocosol rivers subpopulations (FRPOG) from a southeastern group (SEG) of lower San Juan and Rio Tortuguero samples (Fig. 3B), including the San Carlos site. A strongly supported *P. gillii* break (BP = 97%) between Sabalo and Frio rivers formed a northwestern group (NWG) of westernmost Lake Nicaragua-tributary subpopulations (Rio Sapoa, Rio Sabalo). In *Xenophallus* (Fig. 3C), three barriers delimited an upland group (UG) confined to upper San Carlos tributaries separated from a large group localized in low-elevation San Juan and Rio Tortuguero tributaries (LSJTOG), and two smaller groups confined to the Corinto (COG; Rio Chirripó drainage) and Parismina rivers (PAG). The UG–LSJTOG barrier separated higher-elevation *Xenophallus* sites from low-elevation ones. Combined, SAMOVA and BARRIER results supported a similar pattern delimiting a fifth, northwestern *Xenophallus* group (NWG) divided from the other groups just northwest of the active Arenal volcano within the Guanacaste Cordillera highlands (Fig. 3C).

Mantel tests supported isolation-by-distance in *Xenophallus* (normalized Mantel coefficient = 0.307, t = 3.512, P = 0.003) but not the other species (*A. cultratus*: normalized Mantel coefficient = 0.0394, t = 0.464, P = 0.320; *P. gillii*: normalized Mantel coefficient = 0.138, t = 1.428, P = 0.0752). Likewise, *Xenophallus* genetic and ln-geographic distances showed a positive regression relationship not recovered in *A. cultratus* or *P. gillii* (Fig. S3).

Based on biogeographical AMOVA results, we rejected congruent genetic structuring across taxa within and among a priori drainage groups and San Juan tributary drainages (Table 2). Consistent with gene flow of alleles among demes in different drainages, *A. cultratus* showed nonsignificant genetic structuring among San Juan–back-arc drainage groups (AMOVA model 1) and *A. cultratus* and *P. gillii* had nonsignificant structuring among San Juan

tributaries (AMOVA model 2). However, *P. gillii* were significantly differentiated between San Juan and back-arc drainage groups. In contrast, *Xenophallus* AMOVA 2 indicated San Juan tributaries were distinct from one another, but that San Juan–back-arc drainage structuring was nonsignificant (Table 2).

Gene tree analyses further highlighted phylogeographic incongruence. *Alfaro cultratus* was monophyletic with three well-supported clades (Figs. 3A, S4) diverged on average 3.2% at cyt*b*. *Alfaro cultratus* clades were mostly overlapping mosaics of subpopulations that mapped poorly to drainages but closely matched population groups inferred using SAMOVA/BARRIER: with few exceptions, haplotypes comprising clades I–III were confined to the SEG group, while those of clade IV fell into the NWG. Located respectively west versus east of the NWG–SEG barrier, clades I and IV were maximally diverged (4.2%). Clades III and IV had star-like networks with ancestral Sapoa, San Carlos, and Sixaola drainage haplotypes possibly indicating recent population expansions in these regions, and we estimated a San Carlos origin for the network root (haplotype 1). *Poecilia gillii* (Figs. 3B, S4) displayed two geographically overlapping mtDNA clades, limited spatially isolated or genetically distinct variation, ~2–4% divergences at cyt*b* (max. divergence: 4.6%, clades I versus II), and small well-supported San Juan and back-arc drainage subclades (subclade II-a: haplotypes 38–39, mainly San Carlos; II-b: 40–42, mainly Sapoa, Sabalo and Frio; and II-c: 59–61, Matina). Haplotype 24 (San Carlos) was ancestral, sister to all other *P. gillii* haplotypes; however, haplotype 62 (Parismina) was the network root and showed a star-like pattern consistent with recent expansion. While the SEG harbored most *P. gillii* alleles, haplotypes 13 and 38–41 were confined to the NWG group. *Xenophallus* differed from the other species in having four well-supported, nonallopatric clades mostly isolated in drainage basins. The *Xenophallus* cyt*b* topology contrasted deep (6.0%) divergence of San Carlos (clade I) haplotypes (e.g., haplotype 29, the network root), from all others, against shallow intradrainage variation (Figs. 3C, S4). As in the other species, *Xenophallus* gene tree and network results also supported the genetic barriers inferred in BARRIER, with UG samples largely constituting clade I, COG and PAG samples largely constituting clade II, and clade III presenting a mixture of NWG, UG, and LSJTOG haplotypes. A star-like network pattern was only recovered among *Xenophallus* UG haplotypes, consistent with intradrainage expansion (Fig. 3C).

Finer-scale spatial congruence

Whereas the above results indicated overall spatial incongruence, congruent genetic structuring was supported over finer spatial scales in one area. SAMOVA, BARRIER, and phylogenetic results revealed genetic differentiation in the same subregion of northwestern Costa Rica in all three species, with common differentiation just west of the San Carlos basin or between lowland-to-upland Frio and San Carlos sites, but all along the western edge of the San Carlos (Figs. 3, S1, S4). The pertinent breaks split the *A. cultratus* NWG–SEG, *Xenophallus* UG–LSJTOG, and *P. gillii* FRPOG–SEG groups (Fig. 3). These "northwest Costa Rica breaks" corresponded to species main pairwise population divergences, including BP-supported barriers, or barriers with the highest TrN distances identified in BARRIER.

Temporal incongruence

Coalescent-based dating analyses in IMa2 yielded reliable estimates of BARRIER population group sizes (θ) and divergence times (t) in most runs, indicated by likelihood surface peaks (Fig. 3, Table 3, and Appendix S3). Obtaining confidence intervals for t was difficult, however, because some runs peaked at lower values before converging to positive values at larger t, representing infinite migration; thus, we accepted likelihood peaks as the best parameter estimates. So, although we found congruent northwestern Costa Rica spatial breaks, peak posterior t estimates revealed temporal incongruence for the three species overall and at the shared break (Fig. 3 and Table 3). Miocene–mid-Pleistocene divergences in northwest Costa Rica were much more likely in *A. cultratus* (NWG–SEG: T_{div} range = 3.583–1.612 Ma) and *Xenophallus* (UG–LSJTOG: T_{div} range = 13.731–2.334 Ma) than *P. gillii*. All divergences within *P. gillii* ranged over mid-late Pleistocene, including the FRPOG–SEG population pair (T_{div} range = 0.130–0.0221 Ma). A similar pattern of incongruence arose when comparing all T_{div} estimates together. Whereas we estimated nonzero migration rates in *A. cultratus*, posterior m distributions peaked at the lower limit of resolution or 95% HPDs included zero in *P. gillii* and *Xenophallus*, indicating no ongoing gene flow.

Akin to IMa2 results above, tests for simultaneous diversification at a finer-scale level within northwestern Costa Rica using approximate Bayesian computation models also revealed a striking pattern of temporal incongruence. MTML-msBayes results were nearly identical across four models with slightly different priors; therefore, we present results from one representative model (M2, upper $\theta = 0.01$, upper $E[\tau] = 2$, $Nm = 0$), though prior settings and results for all models can be found in Table S4. The Ψ (mean = 2.291) and Ω (mean = 0.269, 95% HPD range = 0.000–0.657; Bayesian posterior probability of one divergence event from polychotomous regression = 0.149) parameter estimates indicated that a model

Table 3. Coalescent divergence time analysis parameter estimates.

Species	Comparison (1–2)	θ_1	θ_2	$m_{1\rightarrow2}$	$m_{2\rightarrow1}$	t	T_{div}, 2% rate (Ma)	T_{div}, 0.9% rate (Ma)
Alfaro	NWG–SEG	14.920	39.560	0.348	0.166	9.690	1.612	3.583
cultratus	95% HPDs	7.720, 24.280	27.160, NA	0.108, 0.923	0.178, 0.438	3.013, NA	0.953, NA	2.119, NA
Poecilia	NWG–SEG	6.750	53.25	0.0005	0.528	1.278	0.112	0.659
gillii	95% HPDs	1.750, NA	29.250, 85.750	0.000, NA	0.257, NA	–	–	–
	NWG–FRPOG	7.350	5.850	0.00171	0.341	1.468	0.129	0.757
	95% HPDs	1.770, 25.410	1.410, 18.570	0.000, 2.359	0.000, 2.349	0.755, NA	0.0662, NA	0.390, NA
	FRPOG–SEG	4.750	45.250	0.000	0.000	0.253	0.0221	0.130
	95% HPDs	1.250, 14.250	22.250, 82.250	–	–	0.0975, 0.548	0.0009, 0.0480	0.050, 0.283
Xenophallus	UG–LSJTOG	9.750	29.750	0.000	0.000	26.610	2.334	13.731
	95% HPDs	3.750, 21.750	16.250, 49.750	–	–	5.850, 35.550	0.513, 3.118	3.019, 18.344
	LSJTOG-PAG	29.500	3.500	0.000	0.000	11.140	0.977	5.748
	95% HPDs	14.500, 50.500	0.000, NA	–	–	4.537, NA	0.398, NA	2.341, NA
	COG–UG	0.500	9.500	0.0004	0.0004	2.138	0.188	1.103
	95% HPDs	0.000, 11.500	3.500, 23.500	0.000, 0.6764	0.000, 0.301	0.613, NA	0.0537, NA	0.316, NA
	COG–LSJTOG	0.250	28.750	0.000	0.000	10.190	0.894	5.258
	95% HPDs	0.000, NA	16.250, 49.250	–	–	4.388, 16.860	0.385, 1.479	2.264, 8.700

Estimates of population sizes (θ_1, θ_2); migration rates (m); mutation-scaled population divergence times (t); and absolute divergence times (T_{div}) in millions of years ago, inferred in IMa2 are shown for pairwise comparisons of diverged population groups (regions) from BARRIER (see Results, Fig. 3). Estimates were similar across three final runs using different random seeds, so results from best runs are presented. In brackets, 95% highest posterior density intervals (HPDs) are given where complete posterior distributions appeared to be estimated; whereas bounds that could not be estimated are listed as not available (NA), and zeros (with no density intervals) are given for m estimates in models for which pilot runs recovered zero migration hence m priors were set to zero in final runs. We calculated T_{div} using different mutation rates (μ), including the standard 2% rate for vertebrate mtDNA genes (Brown et al. 1979; Wilson et al. 1985), and a more slowly evolving 0.9% salmonid mtDNA rate (Martin and Palumbi 1993).

of multiple discrete divergences, rather than simultaneous divergence, was supported by the data (Fig. 3 and Table S4). This was also supported by hypotheses testing: based on Bayes factors of 4.303 for $\Psi > 1$ versus $\Psi = 1$, the data provide substantial support for a model with multiple divergences. In contrast, Bayes factors of 0.930 and 0.947 indicated only marginal weight of evidence (Jeffreys 1961) for simultaneous divergence ($\Psi = 1$ vs. $\Psi > 1$) and continuous divergence ($\Psi = 3$ vs. $\Psi < 3$) respectively. Based on a Bayes factor of 1.535 for $\Omega > 0.01$ vs. $\Omega < 0.01$, dispersion index Ω also indicated evidence against a simultaneous divergence model; however, the weight of the evidence was marginal, suggesting a potentially weaker ability of Ω to reject simultaneous divergence for our data. Modal divergence time estimates across the three population pairs from MTML-msBayes were similar to T_{div} estimates from IMa2 falling mostly within the Pliocene–Pleistocene (Table S4).

Historical-demographic incongruence

The *P. gillii* data provided substantial support for broadly incongruent historical demography. The data supported the *P. gillii* Bayesian skyline plot over the other competing demographic models based on Bayes factors (Table 4), and plotting the skyline reconstruction of population dynamics through time revealed *P. gillii* late Pleistocene growth following slight population bottlenecking ~40 ka (Fig. S5). *Poecilia gillii* population expansion was also supported by significant ($P < 0.01$) and positive R_2 statistics (Table 1), as well as a star-like pattern of haplotypes radiating from the network root (Fig. 3B); however, the expansion signal was not recovered by Tajima's D, which was negative but nonsignificant. *Alfaro cultratus* results were intermediate to those of *P. gillii*: whereas Bayes factors strongly supported the constant model over the other competing models (Table 4), significant and positive R_2 statistics supported expansions overall and within *A. cultratus* population groups despite negative and nonsignificant Tajima's D values (Table 1). This was surprising, given parsimony networks showed evidence for finer-scale *A. cultratus* expansions within regions (see above, Fig. 3A). Contrasting patterns in the other taxa, essentially all *Xenophallus* results pointed to a constant population size over time. Bayes factors less than 0.5 indicated that *Xenophallus* models were indistinguishable (Jeffreys 1961); thus, by parsimony, the

Table 4. Bayes factor tests comparing Bayesian coalescent demographic models.

Species	Model	t_{MRCA} (Ma)	Smoothed ln likelihood (L) ± SE	Bayes factors ($\log_{10} B_{10}$)			
				BSP	Constant	Exponential	Logistic
Alfaro cultratus	BSP	1.358 [0.448, 4.193]	−1508.083 ± 0.188	–	−1.586	−1.793	−1.826
	Constant	1.398 [0.460, 4.272]	−1504.432 ± 0.185	**1.586***	–	−0.207	−0.240
	Exponential	1.086 [0.423, 2.863]	−1503.955 ± 0.186	**1.793***	0.207	–	−0.033
	Logistic	1.329 [0.449, 4.061]	−1503.879 ± 0.159	**1.826***	0.240	0.033	–
Poecilia gillii	**BSP**	1.937 [0.677, 5.859]	−2335.298 ± 0.119	–	**1.978***	**1.682***	**1.971***
	Constant	1.842 [0.656, 5.605]	−2339.853 ± 0.135	−1.978	–	−0.296	−0.006
	Exponential	1.472 [0.628, 3.827]	−2339.172 ± 0.137	−1.682	0.296	–	0.289
	Logistic	1.740 [0.633, 5.219]	−2339.838 ± 0.134	−1.971	0.006	−0.289	–
Xenophallus	BSP	1.937 [0.677, 5.859]	−2263.109 ± 0.094	–	0.008	−0.208	0.079
	Constant	1.842 [0.656, 5.605]	−2263.128 ± 0.099	−0.008	–	−0.216	0.071
	Exponential	1.597 [0.735, 4.217]	−2262.630 ± 0.093	0.208	0.216	–	0.287
	Logistic	1.823 [0.716, 5.507]	−2263.291 ± 0.092	−0.079	−0.071	−0.287	–

Geometric mean t_{MRCA} estimates based on sufficient MCMC-chain mixing in Beast (ESS > 400) are shown in millions of years ago with 95% HPDs in brackets, followed by ln-likelihood estimates from Tracer (±standard error [SE]). Bayes factors are presented as row-by-column comparisons. Best-fit models based on Jeffreys' (1961) "weight of evidence" criteria (*strong support) are presented in bold.

constant model (model with the fewest parameters) is the most likely best-fit *Xenophallus* model. Consistent with population stasis, most *Xenophallus* networks suggested stable population structuring, and most groups had nonsignificant R_2 and Tajima's D values (Fig. 3B and Table 1). However, the finer-scale *Xenophallus* UG population expansion in the San Carlos basin revealed by the network (above) was supported by a positive and significant R_2 (Table 1). Results of Fay and Wu's H tests supported a scarcity of high frequency variants suggesting that the historical demographic inferences above do not reflect purifying or positive natural selection ($P > 0.05$; Table 1). Demographic models in Beast yielded intraspecific t_{MRCA}s that were comparable to IMa2 and MTML-msBayes estimates, peaking ~1.9–1.4 Ma around early Pleistocene, with overlapping Miocene–mid-Pleistocene confidence intervals (Table 4).

Discussion

The paradigm view in historical biogeography holds that congruent spatial-genetic subdivisions among codistributed taxa are most parsimoniously explained by a shared biogeographic history, whereas spatially incongruent patterns reflect independent responses owing to biological differences (Arbogast and Kenagy 2001; Avise 2000; Bermingham and Martin 1998; Hickerson et al. 2010; Sullivan et al. 2000; Donoghue and Moore 2003; Feldman and Spicer 2006; Bagley and Johnson 2014; refs. therein). Moreover, temporally incongruent patterns are thought to reflect multiple divergences in response to different events (Cunningham and Collins 1994; Donoghue and Moore 2003). Thus, a key question in historical biogeography is whether testing for shared biogeographic history supports concerted, independent, or multiple evolutionary responses. From empirical tests for spatial and temporal phylogeographic congruence among three livebearing fish species from the Nicaraguan depression of Central America, we find considerable evidence that the evolution of these taxa has not been concerted. Instead, these fishes display strikingly incongruent spatial-genetic structuring (Figs. 3, S1–S4, Appendix S2, and Table 2) *and* temporal population divergences (Fig. 4, Tables 3, 4, S4)–an overall pattern of pseudoincongruence. We therefore reject the concerted- and independent-response hypotheses. Our results suggest that our focal species have neither responded *solely* in lockstep fashion nor *solely* individualistically to long-term effects of shared biogeographic history, but that multiple geological or climatic events within the complex Nicaraguan depression landscape have shaped their population structuring. Multiple responses during recent community assembly involving different geographical distributions or colonization routes appear to have played a role in shaping the phylogeographic and community composition of the northern lower Central American freshwater fish assemblage. While drawing more robust conclusions about the precise number and underlying causes of population divergences inferred herein using mtDNA will require additional data from multiple unlinked nuclear loci, our study represents an important first step toward unraveling the history of the fish communities in this region. Indeed, ours is the first comparative analysis establishing a geographical and temporal framework for understanding diversification of northern LCA freshwater biota. Our results also provide some evidence that multiple evolutionary responses across these species were overlaid by incongruent demographic

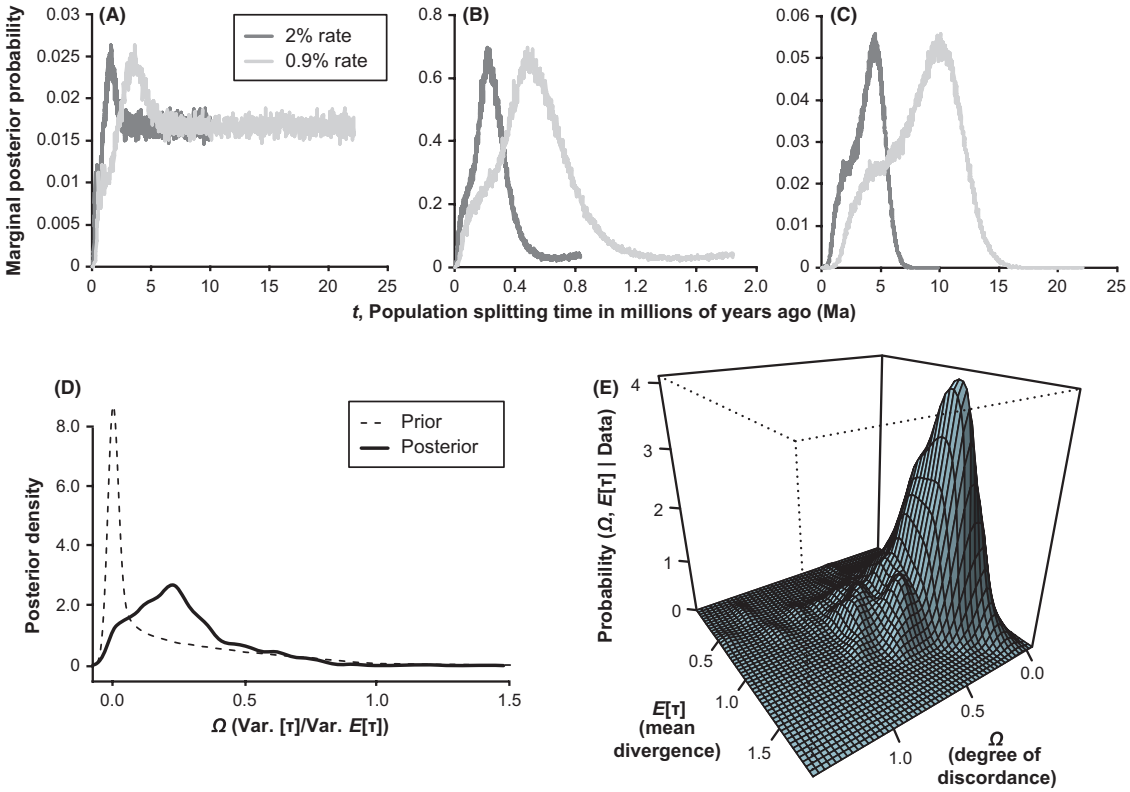

Figure 4. Posterior distributions and temporal incongruence of divergence times among northwest Costa Rican population pairs. The first three panels display marginal posterior probabilities of t parameters estimated by IMa2 for *Alfaro cultratus* (A), *Poecilia gillii* (B), and *Xenophallus umbratilis* (C), with probabilities estimated from two separate substitution rates. The two lower panels present results from MTML-msBayes analyses, including a comparison of the prior distribution versus the posterior densities of the number of divergence times across population pairs, Ψ (D), as well as a surface plot of the joint posterior probability densities of $E[\tau]$ (E).

histories (Tables 1, 4 and Fig. S5). Here, we explore each level of incongruence among our results–temporal, spatial, and demographic, as well as ecological factors that, in addition to a multiple-response scenario, potentially explain the patterns we observed.

Statistical phylogeography studies have repeatedly underscored the importance of testing for temporal congruence while accounting for potentially confounding factors influencing divergence time estimates, such as mutational and coalescent gene-tree stochasticity (e.g., Edwards and Beerli 2000; Wakeley 2003; Nielsen and Beaumont 2009). We empirically tested for temporal congruence across multiple codistributed species while using methods that explicitly model isolation processes, demographic events, and coalescent variance, for example, Bayesian simulations sampling many coalescent gene genealogies. At the broadest temporal scales relevant for our data (i.e., thousands to millions of years), we inferred Miocene–mid-Pleistocene divergences between regions in *A. cultratus* and *X. umbratilis*, but mid-late Pleistocene divergences across barriers in *P. gillii* (Fig. 3 and Table 3), using full-Bayesian IMa2 analyses. Their different timescales of diversification are in greatest accord with the interpretation that these ND livebearing fishes represent possibly ancient but asynchronously evolved lineages that did not disperse into the study area at the same time. In other words, a parsimonious explanation of these patterns is that these species had different past distributions, thus experienced different dispersal and vicariance events at different times. Our results preclude a scenario of ancient dispersal or vicariance in *P. gillii*, however, because regional divergence estimates agree with a possibly recent origin of this species in the study area. This is consistent with substantial evidence for a late Pleistocene bottleneck-expansion event in *P. gillii* (Fig. S5 and Tables 1, 4) that may have occurred during recent recolonization, or post-colonization expansion. Nevertheless, the well-known limitations of single-locus phylogeography studies warrant careful

consideration of the effects of potential sources of uncertainty while interpreting these results. In particular, it is difficult to estimate demographic parameters from mtDNA, from which we should expect wider confidence intervals reflecting (1) the inherent stochasticity of coalescent processes (Hudson 1990; Edwards and Beerli 2000) and (2) the influence of varying levels of N_e or m across ancestral populations (Wakeley 2003; Nielsen and Beaumont 2009). While our IMa2 runs converged for most parameter estimates and estimated error in the inferences, chance events have likely influenced our results. Estimating confidence in population divergence times was problematic and gave broad, overlapping confidence intervals (Table 3). Moreover, approximately flat likelihood surfaces yielded large ancestral population size estimates, possibly indicating retained polymorphisms or that gene flow occurred between the ancestral population and other populations not included in the simple 2-population IMa2 models we employed. However, independent estimates of divergence times converged on similar results supporting conclusions drawn from the IMa2 analyses: multi-population coalescent models (i.e., Bayesian skyline plots) inferred intraspecific t_{MRCA}s whose confidence intervals bracketed the majority of the IMa2 t-estimate probability densities, and we also estimated similar Pliocene–Pleistocene regional divergences in MTML-msBayes (Fig. 3 and Tables 4, S4). Hypotheses tests using Bayes factors to compare divergence models based on approximate Bayesian computation simulations also provided moderate to strong support for temporal incongruence (Table S4). Thus, despite potential issues with mtDNA time estimates, different methods support the inferred pattern of multiple evolutionary responses over a Miocene–Pleistocene timescale of diversification.

Qualitatively similar patterns of idiosyncratic temporal divergences have been reported in other comparative phylogeographic studies, including analyses of three codistributed freshwater fish lineages from southern LCA (Bermingham and Martin 1998; Reeves and Bermingham 2006), Mesoamerican rodents across the Isthmus of Tehuantepec (Sullivan et al. 2000), and California herpetofauna (Feldman and Spicer 2006). Our results therefore add to a growing body of evidence from different study systems worldwide supporting a commonality of temporally incongruent phylogeographic patterns in codistributed taxa. The divergence time estimates we report also closely approximate levels, thus the potential timing, of population divergences found in previous studies of lower Central American taxa. For example, a mtDNA-RFLP study of *Orthogeomys cherriei* pocket gophers found haplotypes were up to 1.5% diverged in the Costa Rican Central Cordillera (Demastes et al. 1996), which roughly correlates to mid-Pleistocene assuming the standard 2% pairwise vertebrate mtDNA rate (Bagley and Johnson 2014). A study of *Rhamdia guatemalensis* catfishes found that western Costa Rican populations isolated in the Rio Bebedero basin diverged from all other haplotypes just prior to the final closure of the LCA isthmus ~3 Ma (Perdices et al. 2002). And multiple studies along the Panamanian Isthmus in southern LCA show that various lineages of electric knifefishes (Hypopomidae), seven-spine catfishes (Heptapteridae), and tetras (Characidae) also display Pliocene-late Pleistocene divergences similar to our findings (Bermingham and Martin 1998; Martin and Bermingham 2000; Reeves and Bermingham 2006). Combined with our results, these examples show that the relatively recent (~7-0 Ma) geological history of emergent LCA isthmus lands (e.g., Fig. 1C) appears to have significantly constrained regional patterns and processes of evolutionary divergence, and this is consistent with a recent meta-analysis of divergence times reported in LCA phylogeography studies (Bagley and Johnson 2014).

Given we used the same locus (similar mutation rates) to compare the phylogeographies of closely related species with overlapping ranges, inadequate phylogenetic signal cannot account for the broad-scale pattern of spatial incongruence among *A. cultratus*, *P. gillii*, and *Xenophallus* (e.g., Figs. 3, S1). Instead, while our tests of temporal congruence show that the incongruent spatial-genetic subdivisions in these taxa arose during responses to different events at different times, other factors also potentially explain the observed spatial differences, including species-specific responses driven by different biological attributes (*cf.* Burney and Brumfield 2009; Fouquet et al. 2012). The livebearing fishes we studied share complex ecological adaptations for viviparity, benthopelagic habits, and non-superfetating reproduction (Winemiller 1993; Reznick and Miles 1989; Johnson and Bagley 2011; J. C. Bagley and J. B. Johnson, unpubl. data). Still, these species differ along key ecological axes indicating potentially superior dispersal propensity and wider physiological tolerances in *P. gillii* and *A. cultratus*, relative to *Xenophallus*. Most notably, *P. gillii* achieve larger maximum body size (105 mm; compared with 45–65 mm), a broader range of elevations (0–1220 m; compared with 0–590 m) and thermal environments (J. B. Johnson, pers. obs.), and a much larger geographic range (Guatemala through Panama, except southwestern Panama) than the other species (Bussing 1998; Smith and Bermingham 2005; but see Alda et al. 2013). *Alfaro cultratus* also display a much larger range (northern Nicaragua to western Panama) than *Xenophallus*, which is endemic to the study area (Bussing 1998) and is usually more abundant at the upper elevations of its range (J. C. Bagley and J. B. Johnson, pers. obs.). Furthermore, consistent with salinity tolerance and propensity for movement into peripheral habitats in other *Poecilia* (e.g., *P. mexicana*; Schlupp et al. 2002), *P. gillii*

occur in brackish water, while our other focal species do not (Bussing 1998). These differences at ecological traits correlated to dispersal propensity, population size, and competitive ability have likely influenced unique phylogeographic signals in these species. Indeed, our results, combined with available ecological data, are consistent with the general prediction that phylogeographic structure within species should correlate inversely with behavioral preference or potential for dispersal (e.g., dispersal rates, distances) (Avise 2000; Bagley and Johnson 2014). This is best illustrated in *Xenophallus*, which displays evidence for relatively lower dispersal propensity, yet a higher degree of phylogeographic structuring indicated by deep phylogenetic divergences, haplotypes/clades mostly isolated in drainages (Figs. 3C, S4), and zero ongoing gene flow (Table 3) consistent with limited inter-drainage and -population migration. Mantel tests and regression analyses also supported isolation-by-distance only in *Xenophallus* (Fig. S3). Thus, landscape barriers and geographical distances have influenced phylogeographical structuring to a greater degree in *Xenophallus* than the other taxa most likely due to lower dispersal propensity. That *Xenophallus* AMOVA model 1 supported no structuring between the San Juan and Tortuguero rivers conflicts with this view (Table 2). However, this may reflect a recent stream capture event unrelated to dispersal ecology, or an artifact of limited *Xenophallus* sampling ($N = 9$) and genetic diversity (e.g., $S = 2$) within Rio Tortuguero (Table S1).

Population-level processes may also have contributed to the spatially incongruent subdivisions among the ND livebearers in this study. Because gene flow among demes, incomplete lineage sorting, and demographic fluctuations can produce similar genetic imprints (e.g., converging due to chance, or regional extinctions), teasing these processes apart is difficult. However, nonequilibrium statistical phylogeography tools such as IMa2 that jointly estimate demographic parameters while modeling coalescent and mutational stochasticity, implicitly test whether alleles shared between populations reflect gene flow versus incomplete lineage sorting (assuming uniform priors with $m \neq 0$, as in all of our pilot IMa2 runs and several of our long runs). If marginal distributions of m parameters include 0, gene flow can be rejected in favor of incomplete lineage sorting; otherwise, there is sufficient information to resolve migration as influencing the data. In IMa2 we inferred discordant migration rates between regions among taxa, with nonzero migration between *A. cultratus* groups while other results suggested no ongoing gene flow (Table 3; Appendix S3). Thus shared alleles between diverged populations (Figs. 3, S4) are best explained by gene flow in *A. cultratus*, but appear more consistent with incomplete lineage sorting in the other taxa. While multiple unlinked loci are needed to obtain more accurate parameter estimates to test this initial mtDNA characterization, our migration estimates generally agree with the ecological context above (e.g., zero gene flow within presumably poorer dispersing *Xenophallus*). Moreover, we expect that samples from demes with higher m should be more polymorphic than those from demes with lower m (Wakeley and Aliacar 2001), and this is met by the polarized genetic variation displayed in *A. cultratus* (higher π and *Hd*) and *Xenophallus* (much lower π and *Hd*; Tables 1, S1). Zero-gene-flow inferences in *P. gillii* are exceptional to this (Table 3), as other data suggest this taxon may be a stronger disperser; however, low m estimates may indicate insufficient data for estimating migration in this species while fitting a six-parameter model. Alternatively, the small (<1) nonzero peak m estimates between *P. gillii* population groups (e.g., SEG and FRPOG into NWG) may simply indicate a trivial number of migrants relative to overall population size (Tables 1, 3).

Natural selection is another process with consequences for population genetic variation that can cause spatially incongruent phylogeographic breaks across codistributed species (Irwin 2002). However, while selection can play a role in shaping mtDNA genetic patterns (e.g., Machado and Hey 2003), it has unlikely influenced major patterns among our results. Coalescent simulations of neutrality test statistics demonstrate that our data conform to expectations of selectively neutral evolution (Table 1). Furthermore, we evaluated spatial phylogeographic congruence based on tests of whether genetic barriers were supported by bootstrapping (a randomization procedure) in BARRIER, which allows us to rule out a random pattern of barriers due to natural selection. Despite the utility of this approach, other studies drawing similar conclusions (e.g., Fouquet et al. 2012), have not evaluated this possibility; however, doing so seems more important in cases such as ours where broad-scale spatially incongruent patterns are recovered than in other cases.

In addition to temporally and spatially incongruent phylogeographic histories, we find evidence for incongruent patterns of historical-demographic fluctuations over recent timescales among ND livebearers (Figs. 3, S5 and Tables 1, 4). Our results support recent broad-scale expansion in *P. gillii* and overall stasis despite finer-scale expansions within regions in *A. cultratus* and *Xenophallus*. However, while Bayes factors strongly rejected the null model in *P. gillii* and strongly supported it in *A. cultratus*, they could not distinguish between skyline, constant, exponential and logistic demographic models in *Xenophallus*. Clearly, failing to reject the null model (size-constancy) provides a weaker basis for making inferences about past population dynamics than rejecting the null model would. However, size-constancy rather than

bottleneck-expansions or other growth trends still appears to be the most likely historical demographic scenario for *Xenophallus* based on a parsimonious interpretation of Bayes factors (Table 4). The gene tree and network patterns (Figs. 3, S4) and neutrality test results (Table 1) also support size-constancy in *Xenophallus*.

Notwithstanding the many points of incongruence among our focal taxa, we have established that these species share genetic breaks just north of present-day Lake Arenal in northwestern Costa Rica, in between two large Rio San Juan tributaries, the Frio and San Carlos rivers (Figs. 1, 3). Considering the ecological and geological heterogeneity of LCA landscapes in and around the Nicaraguan depression (e.g., Fig. 1; reviewed in Bagley and Johnson 2014; Funk et al. 2009; Mann et al. 2007; Coates and Obando 1996; Coates et al. 2004), as well as the incongruent temporal, spatial, ecological, and demographic patterns discussed above, the fact that we find evidence for such congruent spatial breaks at such a fine spatial scale (<~25 km) is rather astonishing. This spatial-genetic subdivision has also never been observed in other Costa Rican taxa aside from the livebearing fishes in this study (Bagley and Johnson 2014). Whereas paleoclimatic effects are often cited as the cause of phylogeographic breaks in terrestrial taxa (e.g., Avise 2000; Hewitt 2000), the geographical distributions of LCA freshwater fishes are principally controlled by drainage basin geomorphology and connectivity, (e.g., Bermingham and Martin 1998; Bussing 1998; Smith and Bermingham 2005). Thus, the observed Frio-San Carlos break most likely reflects a direct influence of different geological and sea level events on the drainage networks of the southern San Juan superbasin. As shown in Fig. 1C, LCA has experienced radical geological transitions and landscape changes (e.g., Coates and Obando 1996; Coates et al. 2004) and is physiographically defined by northwest-southeast-trending volcanic cordilleras of Quaternary age (e.g., Marshall et al. 2003; Marshall 2007). In the vicinity of the shared break, the upper reaches the Frio and San Carlos rivers and nearby drainages interact with the Guanacaste Cordillera; the Plio-Pleistocene activity of volcanoes within this part of the Central American volcanic arc seems most likely to have triggered vicariance or extinction-recolonization events responsible for the Frio-San Carlos break. Modern topography may also have contributed to a common pattern of genetic divergence between tributaries and drainages in this region, including the maintenance of isolation, for example as steep drainage gradients limited connectivity. Given evidence for finer-scale population expansions within the San Carlos basin in all three taxa, the present position of this barrier may reflect an ongoing process of secondary expansion following genetic drift in moderate to long-term isolation.

In summary, through multiple empirical tests for congruence, our study has demonstrated that spatially and temporally incongruent phylogeographic and demographic patterns are evident in three species of livebearing fishes that are codependent upon freshwater habitats within the Nicaraguan depression landscape. The majority of our results point to multiple evolutionary responses among these taxa, and we have statistically shown that these corresponded to multiple historical dispersal and vicariance events, possibly suggesting waves of dispersion through the area. Despite overall pseudoincongruence supporting a "multiple-response hypothesis", however, landscape history appears to have promoted commonalities of phylogeographical structuring, albeit over fine spatial scales. More nuclear loci and expanded spatial sampling covering the entire species ranges are necessary to better tease apart the exact histories responsible for the varying evolutionary trajectories in these taxa. However, a comparative perspective has afforded us a view of the lower Central American freshwater fish assemblage that has provided insights into historical as well as ecological influences on population structure, and which permits drawing several future predictions. First, additional studies of individual taxa similarly confined to these freshwater habitats of the Nicaraguan depression should show similar phylogeographic patterns, although it is likely that even further evidence for a multiple-response scenario will be uncovered. And, secondly, we predict that additional comparative studies will yield many new insights into the relative roles of concerted, independent, and multiple responses in shaping the assembly and diversification of species rich and endemic Central American ecosystems.

Acknowledgments

This research was funded by Brigham Young University, including a Mentoring Environment Grant and a Graduate Studies Graduate Research Fellowship, and through stipend support from US National Science Foundation PIRE project OISE-PIRE 0530267. Sampling was conducted under appropriate SINAC-MINAET permits, and we thank Javier Guevara Siquiera (San José, Costa Rica) for help issuing collecting permits and undergraduate Joey Nelson for valuable assistance in the field. We are indebted to Wilfredo Matamoros for generously providing *Alfaro huberi* samples, Mike Hickerson for assistance with MTML-msBayes analyses and interpretation, and Brian Barber, Laura Belk, M. Florencia Breitman, Jesse Brienholt, Keith Crandall, Jared Lee, Dan Mulcahy, Peter Unmack, and Fernanda Werneck for helpful discussions on earlier versions of this manuscript.

Conflict of Interest

None declared.

References

Alda, F., R. G. Reina, I. Doadrio, and E. Bermingham. 2013. Phylogeny and biogeography of the *Poecilia sphenops* species complex (Actinopterygii, Poeciliidae) in Central America. Mol. Phylogenet. Evol. 66:1011–1026.

Arbogast, B. S., and G. J. Kenagy. 2001. Comparative phylogeography as an integrative approach to historical biogeography. J. Biogeogr. 28:819–825.

Avise, J. C. 2000. Phylogeography: the history and formation of species. Harvard Univ. Press, Cambridge, MA. vii, 447p.

Bagley, J. C., and J. B. Johnson. 2014. Phylogeography and biogeography of the lower Central American Neotropics: diversification between two continents and between two seas. Biol. Rev. doi: 10.1111/brv.12076.

Beaumont, M. A., W. Zhang, and D. J. Balding. 2002. Approximate Bayesian computation in population genetics. Genetics 162:2025–2035.

Bermingham, E., and A. P. Martin. 1998. Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. Mol. Ecol. 7: 499–517.

Brown, W. M., M. George Jr, and A. C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl Acad. Sci. USA 76:1967–1971.

Burney, C. W., and R. T. Brumfield. 2009. Ecology predicts levels of genetic differentiation in Neotropical birds. Am. Nat. 174:358–368.

Burridge, C. P., D. Craw, D. Fletcher, and J. M. Waters. 2008. Geological dates and molecular rates: fish DNA sheds light on time dependency. Mol. Biol. Evol. 18:624–633.

Bussing, W. A. 1976. Geographic distribution of the San Juan ichthyofauna of Central America with remarks on its origin and ecology. Pp. 157–175 *in* T. B. Thorson, ed. Investigations of the ichthyofauna of Nicaraguan lakes. Univ. of Nebraska, Lincoln, NE.

Bussing, W. A. 1985. Patterns of the distribution of the Central American ichthyofauna. Pp. 453–473 *in* F. G. Stehli and S. D. Webb, eds. The Great American Biotic Interchange. Plenum Press, New York, NY.

Bussing, W. A. 1998. Freshwater fishes of Costa Rica, 2nd ed. Editorial de la Universidad de Costa Rica, San José, Costa Rica, 468p.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9:1657–1659.

Coates, A. G., and J. A. Obando. 1996. The geologic evolution of the Central American isthmus. Pp. 21–56 *in* J. B. C. Jackson, A. F. Budd and A. G. Coates, eds. Evolution and environment in tropical America. Univ. of Chicago Press, Chicago, IL.

Coates, A. G., L. S. Collins, M.-P. Aubry, and W. A. Berggren. 2004. The geology of the Darien, Panama, and the late Miocene-Pliocene collision of the Panama arc with northwestern South America. Geol. Soc. Am. Bull. 116:1327–1344.

Cunningham, C. W., and T. Collins. 1994. Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. Pp. 405–433 *in* R. DeSalle, B. Schierwater and G. Wagner, eds. Molecular ecology and evolution: approaches and applications. Birkhäuser-Verlag, Basel, Switzerland.

Demastes, J. W., M. S. Hafner, and D. J. Hafner. 1996. Phylogeographic variation in two Central American pocket gophers (*Orthogeomys*). J. Mammal. 77:917–927.

Doadrio, I., S. Perea, L. Alcaraz, and N. Hernandez. 2009. Molecular phylogeny and biogeography of the Cuban genus Girardinus Poey, 1854 and relationships within the tribe Girardinini (Actinopterygii, Poeciliidae). Mol. Phylogenet. Evol. 50:16–30.

Donoghue, M. J., and B. R. Moore. 2003. Toward an integrative historical biogeography. Integr. Comp. Biol. 43:261–270.

Drummond, A. J., A. Rambaut, B. Shapiro, and O. G. Pybus. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22:1185–1192.

Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29:1969–1973.

Dupanloup, I., S. Schneider, and L. Excoffier. 2002. A simulated annealing approach to define the genetic structure of populations. Mol. Ecol. 11:2571–2581.

Edwards, S. V., and P. Beerli. 2000. Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. Evolution 54:1839–1854.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10:564–567.

Fay, J. C., and C. I. Wu. 2000. Hitchhiking under positive Darwinian selection. Genetics 155:1405–1413.

Feldman, C. R., and G. S. Spicer. 2006. Comparative phylogeography of woodland reptiles in California: repeated patterns of cladogenesis and population expansion. Mol. Ecol. 15:2201–2222.

Fouquet, A., B. P. Noonan, M. T. Rodrigues, N. Pech, A. Gilles, and N. J. Gemmell. 2012. Multiple Quaternary refugia in the eastern Guiana Shield revealed by comparative phylogeography of 12 frog species. Syst. Biol. 61:461–489.

Funk, J., P. Mann, K. McIntosh, and J. Stephens. 2009. Cenozoic tectonics of the Nicaraguan depression, Nicaragua, and Median Trough, El Salvador, based on

seismic-reflection profiling and remote-sensing data. Geol. Soc. Am. Bull. 121:1491–1521.

Haq, B. U., J. Hardenbol, and P. R. Vail. 1987. Chronology of fluctuating sea levels since the Triassic. Science 235:1156–1167.

Hearty, P. J., P. Kindler, H. Cheng, and R. L. Edwards. 1999. A + 20 m middle Pleistocene sea-level highstand (Bermuda and the Bahamas) due to partial collapse of Antarctic ice. Geology 27:375–378.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907–913.

Hey, J. 2010. Isolation with migration models for more than two populations. Mol. Biol. Evol. 27:905–920.

Hey, J., and R. Nielsen. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics 167:747–760.

Hickerson, M. J., B. C. Carstens, J. Cavender-Bares, K. A. Crandall, C. H. Graham, J. B. Johnson, et al. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. Mol. Phylogenet. Evol. 54:291–301.

Hood, G. M. (2008) PopTools. Version 3.0.3. Available at http://www.cse.csiro.au/poptools (accessed 17 January 2013).

Hrbek, T., J. Seckinger, and A. Meyer. 2007. A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. Mol. Phylogenet. Evol. 43:986–998.

Huang, W., Y. Qi, N. Takebayashi, and M. J. Hickerson. 2011. MTML-msBayes: Approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. BMC Bioinformatics 12:1.

Hudson, R. R. 1990. Gene genealogies and the coalescent process. Oxford Surv. Evol. Biol. 7:1–44.

Hudson, R. R., M. Kreitman, and M. Aguadé. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics 116:153–159.

Irwin, D. E. 2002. Phylogeographic breaks without geographic barriers to gene flow. Evolution 56:2383–2394.

Jeffreys, H. 1961. Theory of probability, 3rd ed. Oxford Univ. Press, Oxford, U.K., 447p.

Johnson, J. B., and J. C. Bagley. 2011. Ecological drivers of life-history evolution. Pp. 38–49 *in* J. P. Evans, A. Pilastro and I. Schlupp, eds. Ecology and evolution of poeciliid fishes. Univ. of Chicago Press, Chicago, IL.

Jones, C. P., and J. B. Johnson. 2009. Phylogeography of the livebearer *Xenophallus umbratilis* (Teleostei: Poeciliidae): glacial cycles and sea level change predict diversification of a freshwater tropical fish. Mol. Ecol. 18:1640–1653.

Lambeck, K., T. M. Esat, and E. K. Potter. 2002. Links between climate and sea levels for the past three million years. Nature 419:199–206.

Lee, J. B., and J. B. Johnson. 2009. Biogeography of the livebearing fish *Poecilia gillii* in Costa Rica: are phylogeographical breaks congruent with fish community boundaries? Mol. Ecol. 18:4088–4101.

Li, W.-H. 1997. Molecular evolution. Sinauer Associates, Sunderland, MA, 432p.

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

Loaiza, J. R., M. E. Scott, E. Bermingham, J. Rovira, and J. E. Conn. 2010. Evidence for Pleistocene population divergence and expansion of *Anopheles albimanus* in southern Central America. Am. J. Trop. Med. Hyg. 82:156–164.

Machado, C. A., and J. Hey. 2003. The causes of phylogenetic conflict in a classic *Drosophila* species group. Proc. R. Soc. B. 270:1193–1202.

Mann, P., R. D. Rogers, and L. Gahagan. 2007. Overview of plate tectonic history and its unresolved tectonic problems. Pp. 205–241 *in* J. Bundschuh and G. E. Alvarado, eds. Central America: geology, resources and hazards. Taylor & Francis, Philadelphia, PA.

Manni, F. E., E. Guerard, and E. Heyer. 2004a. Geographical patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by "Monmonier's algorithm". Hum. Biol. 76:173–190.

Manni, F. E., E. Guerard, and E. Heyer (2004b) BARRIER 2.2. Museum of Mankind, Paris, France. Available at http://www.mnhn.fr/mnhn/ecoanthropologie/software/barrier.html (accessed 9 January 2012).

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27: 209–220.

Marshall, J. S. 2007. The geomorphology and physiographic provinces of Central America. Pp. 1–51 *in* J. Bundschuh and G. E. Alvarado, eds. Central America: geology, resources and hazards. Taylor & Francis, Philadelphia, PA.

Marshall, L., R. F. Butler, R. E. Drake, G. A. Curtis, and R. H. Tedforth. 1979. Calibration of the Great American Interchange. Science 204:272–279.

Marshall, J. S., B. D. Idleman, T. W. Gardner, and D. M. Fisher. 2003. Landscape evolution within a retreating volcanic arc, Costa Rica, Central America. Geology 31:419–422.

Martin, A. P., and E. Bermingham. 2000. Regional endemism and cryptic species revealed by molecular and morphological analysis of a widespread species of Neotropical catfish. Proc. R. Soc. B. 267:1135–1141.

Martin, A. P., and S. R. Palumbi. 1993. Body size, metabolic rate, generation time, and the molecular clock. Proc. Natl Acad. Sci. USA 90:4087–4091.

Miller, K. G., M. A. Kominz, J. V. Browning, J. D. Wright, G. S. Mountain, M. E. Katz, et al. 2005. The Phanerozoic record of global sea-level change. Science 310:1293–1298.

Minin, V., Z. Abdo, P. Joyce, and J. Sullivan. 2003. Performance-based selection of likelihood models for phylogeny estimation. Syst. Biol. 52:674–683.

Monmonier, M. S. 1973. Maximum-difference barriers: an alternative numerical regionalization method. Geogr. Anal. 3:245–261.

Nielsen, R., and M. A. Beaumont. 2009. Statistical inferences in phylogeography. Mol. Ecol. 18:1034–1047.

Perdices, A., E. Bermingham, A. Montilla, and I. Doadrio. 2002. Evolutionary history of the genus *Rhamdia* (Teleostei: Pimelodidae) in Central America. Mol. Phylogenet. Evol. 25:172–189.

Perdices, A., I. Doadrio, and E. Bermingham. 2005. Evolutionary history of the synbranchid eels (Teleostei: Synbranchidae) in Central America and the Caribbean islands inferred from their molecular phylogeny. Mol. Phylogenet. Evol. 37:460–473.

Rambaut, A., and A. J. Drummond (2009) Tracer. Version 1.5. Available at http://beast.bio.ed.ac.uk/Tracer (accessed March 15 2013).

Ramos-Onsins, S. E., and J. Rozas. 2002. Statistical properties of new neutrality tests against population growth. Mol. Biol. Evol. 19:2092–2100.

Reeves, R. G., and E. Bermingham. 2006. Colonization, population expansion, and lineage turnover: phylogeography of Mesoamerican characiform fish. Biol. J. Linn. Soc. 88:235–255.

Reznick, D. N., and D. B. Miles. 1989. A review of life history patterns in poeciliid fishes. Pp. 125–148 *in* G. K. Meffe and F. F. J. Snelson, eds. Ecology and evolution of livebearing fishes (Poeciliidae). Prentice-Hall, Englewood Cliffs, NJ.

Rosenberg, M. S., and C. D. Anderson. 2011. PASSaGE: Pattern Analysis, Spatial Statistics, and Geographic Exegesis. Method. Ecol. Evol. 2:229–232.

Schlupp, I., J. Parzefall, and M. Schartl. 2002. Biogeography of the Amazon molly, *Poecilia formosa*. J. Biogeogr. 29:1–6.

Schmidt, T. R., J. P. Bielawski, and J. R. Gold. 1998. Molecular phylogenetics and evolution of the cytochrome *b* gene in the cyprinid genus *Lythrurus* (Actinopterygii: Cypriniformes). Copeia 1998:14–22.

Smith, S. A., and E. Bermingham. 2005. The biogeography of lower Mesoamerican freshwater fishes. J. Biogeogr. 32:1835–1854.

Stehli, F. G., and S. D. Webb, eds. 1985. The Great American Biotic Interchange. Plenum Press, New York, NY. 550 p.

Streicher, J. W., A. J. Crawford, and C. W. Edwards. 2009. Multilocus molecular phylogenetic analysis of the montane *Craugastor podiciferus* species complex (Anura: Craugastoridae) in Isthmian Central America. Mol. Phylogenet. Evol. 53:620–630.

Suchard, M. A., R. E. Weiss, and J. S. Sinsheimer. 2001. Bayesian selection of continuous-time Markov chain evolutionary models. Mol. Biol. Evol. 18:1001–1013.

Sullivan, J., E. Arellano, and D. S. Rogers. 2000. Comparative phylogeography of mesoamerican highland rodents: concerted versus independent response to past climatic fluctuations. Am. Nat. 155:755–768.

Tajima, F. 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10:512–526.

Tamura, K., D. Peterson, N. Peterson, N. Peterson, G. Stecher, M. Nei, et al. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731–2739.

Unmack, P. J., A. P. Bennin, E. M. Habit, P. F. Victoriano, and J. B. Johnson. 2009. Impact of ocean barriers, topography, and glaciation on the phylogeography of the catfish *Trichomycterus areolatus* (Teleostei: Trichomycteridae) in Chile. Biol. J. Linn. Soc. 97:876–892.

Wakeley, J. 2000. The effects of subdivision on the genetic divergence of populations and species. Evolution 54:1092–1101.

Wakeley, J. 2003. Inferences about the structure and history of populations: coalescents and intraspecific phylogeography. Pp. 193–215 *in* R. Singh and M. Uyenoyama, eds. The evolution of population biology. Cambridge Univ. Press, Cambridge, U.K.

Wakeley, J., and N. Aliacar. 2001. Gene genealogies in a metapopulation. Genetics 159:893–905.

Waters, J. M., and C. P. Burridge. 1999. Extreme intraspecific mitochondrial DNA sequence divergence in *Galaxias maculatus* (Osteichthys: Galaxiidae). One of the world's most widespread freshwater fish. Mol. Phylogenet. Evol. 11:1–12.

Watterson, G. A. 1975. Number of segregating sites in genetic models without recombination. Theor. Popul. Biol. 7:256–276.

Wilson, A. C., R. L. Cann, S. M. Carr, M. George, U. B. Gyllensten, K. M. Helm, et al. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. 26:385–400.

Winemiller, K. O. 1993. Seasonality of reproduction by livebearing fishes in tropical rainforest streams. Oecologia 95:266–276.

Zink, R. M., and G. F. Barrowclough. 2008. Mitochondrial DNA under siege in avian phylogeography. Mol. Ecol. 17:2107–2121.

Zwickl, D. J. (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD diss. The University of Texas at Austin.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. SAMOVA maps with maximally genetically differentiated groups of samples separated by thick red

lines. Dotted (thin) lines are Voroni diagram edges used during calculations. All models produced statistically significant fixation indices confirmed by independent AMOVAs ($P < 0.001$; Table 2).

Figure S2. Fixation index scores (Φ_{CT}) from SAMOVA analyses plotted against the K value imposed during each run.

Figure S3. Regressions showing relationships between genetic distances and geographical distances [(ln) km] across all sampling sites (Fig. 2) for three species of LCA livebearing freshwater fishes.

Figure S4. "Best" maximum likelihood gene tree topologies with nodal support.

Figure S5. Bayesian skyline plots of effective population size changes. Historical skyline reconstructions of population size ($N_e\tau$, converted to N_e using generation time) through time for each species, correlated with the late Pleistocene eustatic sea level curve of (Lambeck et al. 2002; cited in the main text).

Table S1. Locality details, population group assignments, GenBank accession numbers, and DNA polymorphism levels across subpopulations.

Table S2. DNA substitution models selected using DT-ModSel.

Table S3. Cyt*b* DNA polymorphism levels within and among drainage basins.

Table S4. Model priors, estimated number and timing of divergence events, and Bayes factors from MTML-msBayes. Results are presented for four coalescent models (M1–M4) run in MTML-msBayes. Bayes factors were used to conduct hypotheses tests of posterior support for simultaneous divergence (e.g., $\Psi = 1$) versus other hypotheses.

Appendix S1. Sampling and outgroups details.

Appendix S2. SAMOVA and BARRIER methods and results.

Appendix S3. Coalescent divergence time estimation: IMa2 methods.

Chapter 2 – Supplementary Material

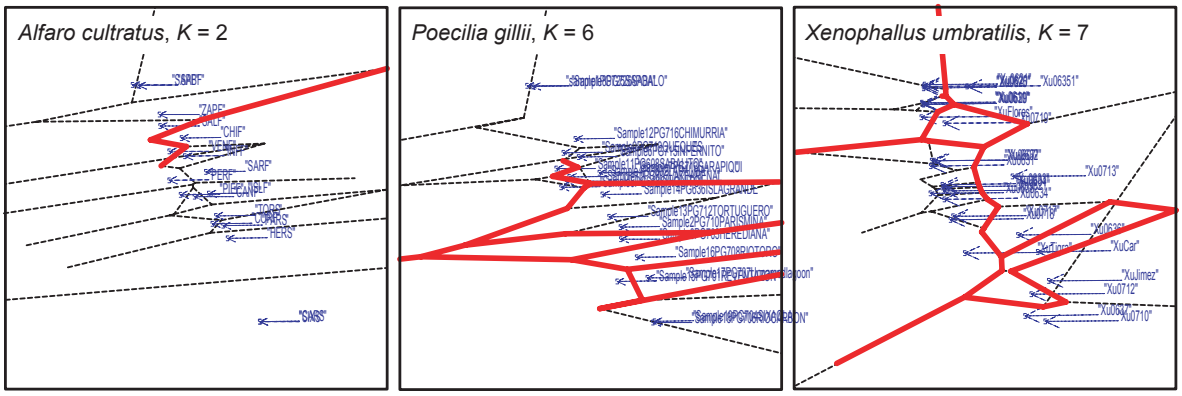
Alfaro cultratus, K = 2
Poecilia gillii, K = 6
Xenophallus umbratilis, K = 7
"Sample12PG716CHIMURRIA"
"Sample2PG710PARISMINA"
"Sample16PG708PIOTOPO"
"Xu06351"
"Xu0713"
"XuCar"
"XuJimez"
"Xu0712"
"Xu0710"

Figure S1

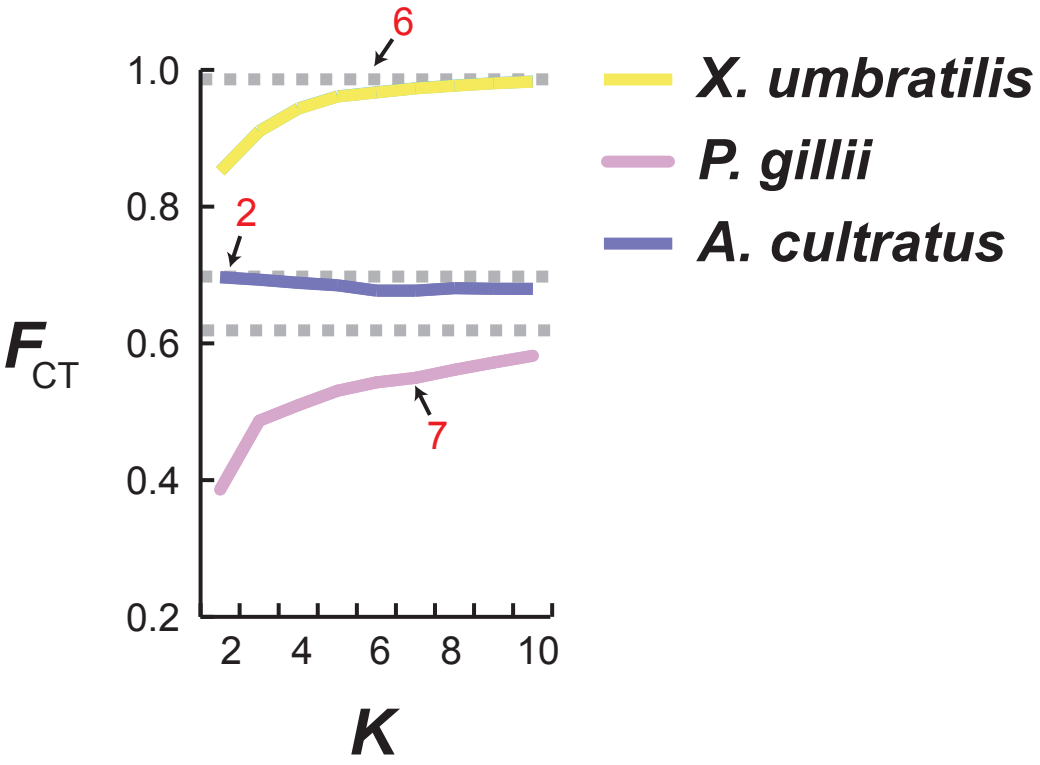
6
2
7
1.0
0.8
0.6
0.4
0.2
F_{CT}
2
4
6
8
10
K
X. umbratilis
P. gillii
A. cultratus

Figure S2

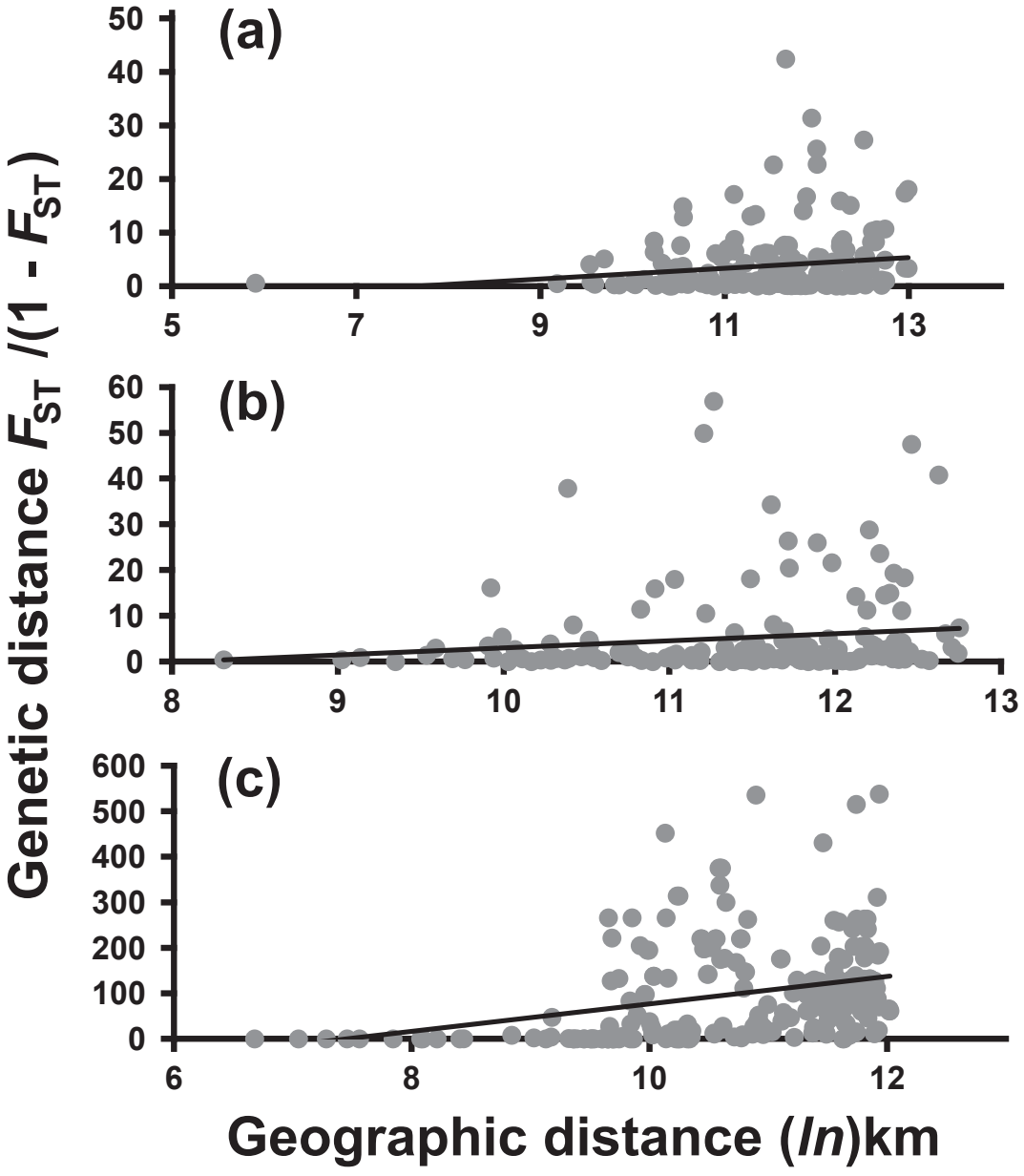

Figure S3

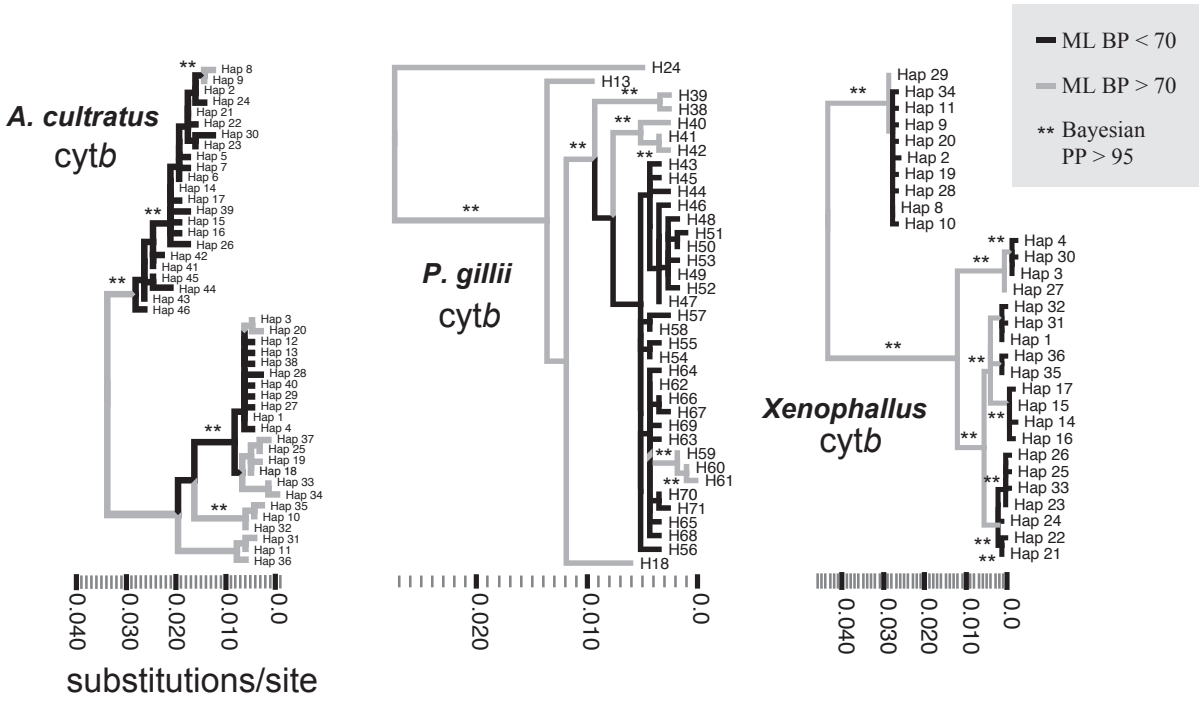
A. cultratus
cytb
P. gillii
cytb
Xenophallus
cytb
ML BP < 70
ML BP > 70
** Bayesian PP > 95
substitutions/site
0.040
0.030
0.020
0.010
0.0

Figure S4

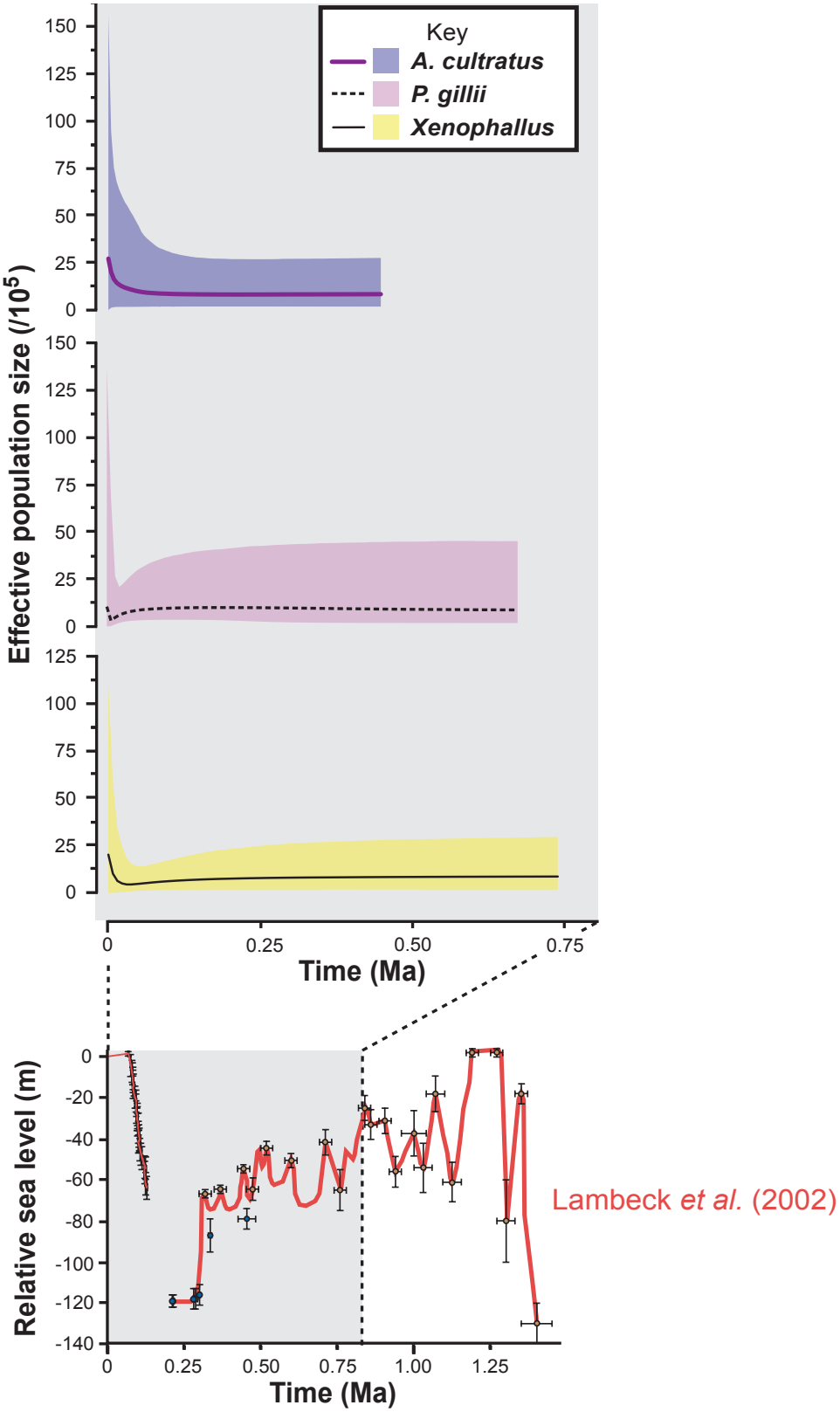
Key
A. cultratus
P. gillii
Xenophallus
Effective population size (/10^5)
150
125
100
75
50
25
0
0
0.25
0.50
0.75
Time (Ma)
Relative sea level (m)
0
-20
-40
-60
-80
-100
-120
-140
0
0.25
0.50
0.75
1.00
1.25
Time (Ma)
Lambeck et al. (2002)

Figure S5

Table S1 Locality details, population group assignments, GenBank accession numbers, and DNA polymorphism levels across sub-populations

NOTE TO EDITOR: GenBank accession numbers for new sequence data generated in this study are pending (denoted "XXXXXXXX").

Species [Ref.]	Locality	Drainage	ID	CODE	BARRIER population group	*N*	Lat. (°N)	Long. (°W)	Country	GenBank nos. cyt*b*	Summary statistics *S*	*h*	*Hd*	s.e. *Hd*	π	θ_w	*N*
Alfaro cultratus						**355**											
A. cultratus [1]	NA	NA	—	—		1	—	—	Costa Rica	EF017531	NA	NA	NA	NA	NA	NA	1
A. cultratus [2]	Lake Nicaragua (LN)	San Juan (1)	1	LN	NWG	2	11.9240	-85.9423	Nicaragua	FJ178773, FJ178772	NA	NA	NA	NA	NA	NA	2
A. cultratus [2]	Rio El Monje (Lake Managua, LM)	San Juan (1)	2	MONJE	NWG	1	11.6333	-86.3000	Nicaragua	FJ178774	NA	NA	NA	NA	NA	NA	1
A. cultratus (this study)	Rio Sapoa (Sapoa)	San Juan (1)	3	SAPF	NWG	49	11.0444	-85.6159	Costa Rica	XXXXXXXX	2	3	0.232	0.076	0.0004	0.453	47§
A. cultratus (this study)	Rio Sabalo (Sabalo)	San Juan (1)	4	SABF	NWG	54	11.0428	-85.4892	Costa Rica	XXXXXXXX	31	7	0.335	0.082	0.0056	6.803	54
A. cultratus (this study)	Rio Zapote (Zapote)	San Juan (1)	5	ZAPF	NWG	12	10.8665	-85.0339	Costa Rica	XXXXXXXX	2	3	0.439	0.158	0.0008	0.662	12
A. cultratus (this study)	Rio Salto (Zapote)	San Juan (1)	6	SALF	NWG	20	10.7982	-85.0233	Costa Rica	XXXXXXXX	25	3	0.511	0.091	0.0047	7.047	20
A. cultratus (this study)	Rio Venado (Frio)	San Juan (1)	7	VENF	NWG	9	10.6448	-84.8222	Costa Rica	XXXXXXXX	0	1	0.000	0.000	0.0000	0.000	9
A. cultratus (this study)	Quebrada Perez (San Carlos)	San Juan (1)	8	PERF	SEG	21	10.4735	-84.8223	Costa Rica	XXXXXXXX	5	3	0.567	0.056	0.0022	1.390	21
A. cultratus (this study)	Rio Chimurria (Pocosol)	San Juan (1)	9	CHIF	SEG	27	10.7274	-84.5582	Costa Rica	XXXXXXXX	29	6	0.708	0.066	0.0206	7.857	23§
A. cultratus (this study)	Quebrada Piecueca (San Carlos)	San Juan (1)	10	PIEF	SEG	7	10.3861	-84.5790	Costa Rica	XXXXXXXX	27	3	0.524	0.209	0.0132	11.020	7
A. cultratus (this study)	Rio Infernito (San Carlos)	San Juan (1)	11	INFF	SEG	19	10.6180	-84.4842	Costa Rica	XXXXXXXX	0	1	0.000	0.000	0.0000	0.000	19
A. cultratus (this study)	Rio Caño Negro (San Carlos)	San Juan (1)	12	CANF	SEG	4	10.3728	-84.2782	Costa Rica	XXXXXXXX	2	2	0.500	0.265	0.0017	1.091	4
A. cultratus (this study)	Rio Sarapiquí (Chirripó)	San Juan (1)	13	SARF	SEG	20	10.5245	-84.0313	Costa Rica	XXXXXXXX	24	7	0.753	0.079	0.0077	6.765	20
A. cultratus (this study)	Rio Isla Grande (Chirripó)	San Juan (1)	14	ISLF	SEG	19	10.3930	-83.9682	Costa Rica	XXXXXXXX	7	5	0.591	0.118	0.0037	2.003	19
A. cultratus (this study)	Rio Corinto (Chirripó)	San Juan (1)	15	CORF	SEG	21	10.2119	-83.8865	Costa Rica	XXXXXXXX	30	4	0.681	0.059	0.0096	8.339	21
A. cultratus (this study)	Upper Río Tortuguero	Tortuguero (2)	16	TORS	SEG	12	10.2594	-83.8122	Costa Rica	XXXXXXXX	27	6	0.818	0.084	0.0162	8.941	12
A. cultratus (this study)	Unnamed river	Parismina (2)	17	PARS	SEG	20	10.1977	-83.6521	Costa Rica	XXXXXXXX	7	4	0.575	0.115	0.0034	2.110	16§
A. cultratus (this study)	Rio Herediana	Parismina (2)	18	PARS	SEG	16	10.1242	-83.5562	Costa Rica	XXXXXXXX	8	4	0.617	0.096	0.0049	2.411	16
A. cultratus (this study)	Unnamed river	Sixaola (2)	19	SIXS	SEG	19	9.6209	-82.8577	Costa Rica	XXXXXXXX	33	7	0.854	0.043	0.0178	9.442	19
A. cultratus (this study)	Rio Carbon	Sixaola (2)	20	SIXS	SEG	3	9.6231	-82.8552	Costa Rica	XXXXXXXX	22	3	1.000	0.272	0.0244	14.667	3
Means:						**16.952**					**15.611**	**4**	**0.539**	**0.104**	**0.0076**	**5.056**	**16.476**
Alfaro huberi						**7**					—	—	—	—	—	—	—
A. huberi (this study)	—	Lis Lis	—	—	—	2	15.6659	-86.5802	Honduras	XXXXXXXX	—	—	—	—	—	—	—
A. huberi (this study)	—	Cangrejal	—	—	—	1	15.6530	-86.0600	Honduras	XXXXXXXX	—	—	—	—	—	—	—
A. huberi (this study)	—	Patuca	—	—	—	2	14.8475	-88.8749	Honduras	XXXXXXXX	—	—	—	—	—	—	—
A. huberi (this study)	—	Motagua	—	—	—	1	14.9054	-89.1618	Honduras	XXXXXXXX	—	—	—	—	—	—	—
A. huberi (this study)	Unnamed tributary	Unnamed tributary	—	—	—	1	14.9978	-89.1321	Honduras	XXXXXXXX	—	—	—	—	—	—	—
Poecilia gillii						**143**											
P. gillii [3]	Rio Sapoa (Sapoa)	San Juan (1)	3	PG725	NWG	8	11.0444	-85.6159	Costa Rica	—	5	4	0.768	0.113	0.0020	1.928	8
P. gillii [3]	Rio Sabalo (Sabalo)	San Juan (1)	4	PG726	NWG	8	11.0428	-85.4892	Costa Rica	—	0	1	0.000	0.000	0.0000	0.000	8
P. gillii [3]	Rio Venado (Frio)	San Juan (1)	7	PG719	FRPOG	16	10.6448	-84.8222	Costa Rica	—	7	5	0.450	0.151	0.0008	2.110	16
P. gillii [3]	Rio Sabalito (San Carlos)	San Juan (1)	—	PG608	SEG	5	10.5486	-84.9808	Costa Rica	—	1	2	0.600	0.175	0.0005	0.480	5
P. gillii [3]	Rio Chiquito (San Carlos)	San Juan (1)	—	PG612	SEG	5	10.4377	-84.8682	Costa Rica	—	22	4	0.900	0.161	0.0109	10.560	5
P. gillii [3]	Lake Arenal (San Carlos)	San Juan (1)	—	PG603	SEG	2	10.4721	-84.7693	Costa Rica	—	12	2	1.000	0.500	0.0105	12.000	2
P. gillii [3]	Rio La Palma (San Carlos)	San Juan (1)	—	PG602	SEG	5	10.4988	-84.6890	Costa Rica	—	0	1	0.000	0.000	0.0000	0.000	5
P. gillii [3]	Rio Infernito (San Carlos)	San Juan (1)	11	PG715	SEG	8	10.618	-84.4842	Costa Rica	—	3	2	0.250	0.180	0.0007	1.157	8
P. gillii [3]	Rio Chimurria (Pocosol)	San Juan (1)	9	PG716	FRPOG	8	10.7274	-84.5582	Costa Rica	—	6	3	0.464	0.200	0.0013	2.314	8
P. gillii [3]	Rio Sarapiquí (Chirripó)	San Juan (1)	13	PG713	SEG	13	10.5246	-84.0313	Costa Rica	—	21	6	0.872	0.054	0.0066	6.767	13
P. gillii [3]	Rio Isla Grande (Chirripó)	San Juan (1)	14	PG636	SEG	1	10.3930	-83.9682	Costa Rica	—	NA	NA	NA	NA	NA	NA	1
P. gillii [3]	Rio Tortuguero (Tortuguero)	Tortuguero (2)	16	PG712	SEG	15	10.2594	-83.8122	Costa Rica	—	18	5	0.705	0.088	0.0037	5.536	15
P. gillii [3]	Unnamed tributary	Parismina (2)	17	PG710	SEG	7	10.1977	-83.5687	Costa Rica	—	0	1	0.000	0.000	0.0000	0.000	7
P. gillii [3]	Rio Herediana	Parismina (2)	18	PG703	SEG	5	10.1242	-83.5562	Costa Rica	—	54	3	0.800	0.164	0.0214	25.920	5
P. gillii [3]	Rio Reventazon	Parismina (2)	—	PG701	SEG	5	9.8723	-83.6332	Costa Rica	—	18	2	0.400	0.237	0.0063	8.640	5
P. gillii [3]	Rio Toro	Matina (2)	—	PG708	SEG	8	10.0168	-83.2102	Costa Rica	—	46	3	0.679	0.122	0.0106	17.741	8
P. gillii [3]	Unnamed lagoon	Matina (2)	—	PG707	SEG	8	9.8926	-82.9723	Costa Rica	—	2	3	0.464	0.200	0.0004	0.771	8
P. gillii [3]	Rio Carbon	Sixaola (2)	20	PG706	SEG	8	9.6231	-82.8552	Costa Rica	—	12	3	0.607	0.164	0.0054	4.628	8
P. gillii [3]	Rio Sixaola	Sixaola (2)	—	PG704	SEG	8	9.6320	-82.8192	Costa Rica	—	0	1	0.000	0.000	0.0000	0.000	8
Means:						**7.526**					**12.611**	**3**	**0.498**	**0.139**	**0.0045**	**5.586**	**7.526**
Xenophallus umbratilis						**131**											
X. umbratilis [4]	Rio Zapote (Zapote)	San Juan (1)	—	Xu0625	NWG	8	10.7242	-85.0664	Costa Rica	—	1	2	0.429	0.169	0.00038	0.386	8
X. umbratilis [4]	Rio Bijagua (Zapote)	San Juan (1)	—	Xu0620	NWG	8	10.7277	-84.0313	Costa Rica	—	3	3	0.464	0.200	0.00066	1.157	8

X. umbratilis [4]	Unnamed tributary (Zapote)	San Juan (1)	—	Xu0621	NWG	8	10.7314	-85.0553	Costa Rica	—	1	2	0.429	0.169	0.00038	0.386	8
X. umbratilis [4]	Rio Venado (Frio)	San Juan (1)	7	Xu0719	LSJTORG	4	10.6448	-84.8222	Costa Rica	—	1	2	0.500	0.265	0.00044	0.545	4
X. umbratilis [4]	Rio Chimurria (Pocosol)	San Juan (1)	9	Xu0635	LSJTORG	8	10.7274	-84.5582	Costa Rica	—	1	2	0.250	0.180	0.00022	0.386	8
X. umbratilis [4]	Rio Sarapiquí (Chirripó)	San Juan (1)	13	Xu0713	LSJTORG	1	10.5245	-84.0313	Costa Rica	—	NA	NA	NA	NA	NA	NA	1
X. umbratilis [4]	Rio Isla Grande (Chirripó)	San Juan (1)	14	Xu0636	LSJTORG	8	10.3930	-83.9682	Costa Rica	—	5	4	0.643	0.184	0.00110	1.928	8
X. umbratilis [4]	Rio Corinto (Chirripó)	San Juan (1)	15	Xu0637	COG	8	10.2119	-83.8865	Costa Rica	—	0	1	0.000	0.000	0.00000	0.000	8
X. umbratilis [4]	Upper Rio Tortuguero	Tortuguero (2)	16	Xu0712	LSJTORG	8	10.2594	-83.8122	Costa Rica	—	2	2	0.536	0.123	0.00094	0.771	8
X. umbratilis [4]	Upper Rio Tortuguero	Tortuguero (2)	—	XuCar	LSJTORG	1	10.3553	-83.7375	Costa Rica	—	NA	NA	NA	NA	NA	NA	1
X. umbratilis [4]	Unknown tributary	Parismina (2)	—	XuJimez	PAG	2	10.2894	-83.6100	Costa Rica	—	1	2	1.000	0.500	0.00088	1.000	2
X. umbratilis [4]	Unnamed tributary	Parismina (2)	—	Xu0710	PAG	6	10.1978	-83.6519	Costa Rica	—	1	2	0.533	0.172	0.00047	0.438	6
X. umbratilis [4]	Quebrada Piecueca (San Carlos)	San Juan (1)	10	XuTigra	UG	2	10.3517	-84.5881	Costa Rica	—	0	1	0.000	0.000	0.00000	0.000	2
X. umbratilis [4]	Lake Arenal (San Carlos)	San Juan (1)	—	Xu0631	UG	8	10.5486	-84.9808	Costa Rica	—	1	2	0.250	0.180	0.00022	0.386	8
X. umbratilis [4]	Lake Arenal (San Carlos)	San Juan (1)	—	Xu0607	UG	8	10.5597	-84.9697	Costa Rica	—	3	4	0.750	0.139	0.00081	1.157	8
X. umbratilis [4]	Lake Arenal (San Carlos)	San Juan (1)	—	Xu0632	UG	8	10.5603	-84.9403	Costa Rica	—	1	2	0.536	0.123	0.00047	0.386	8
X. umbratilis [4]	Lake Arenal (San Carlos)	San Juan (1)	—	Xu0633	UG	4	10.5064	-84.8458	Costa Rica	—	0	1	0.000	0.000	0.00000	0.000	4
X. umbratilis [4]	Lake Arenal (San Carlos)	San Juan (1)	—	Xu0604	UG	3	10.5014	-84.8406	Costa Rica	—	0	1	0.000	0.000	0.00000	0.000	3
X. umbratilis [4]	Lake Arenal (San Carlos)	San Juan (1)	—	Xu9821	UG	8	10.4922	-84.8358	Costa Rica	—	2	3	0.464	0.200	0.00044	0.771	8
X. umbratilis [4]	Unknown tributary (San Carlos)	San Juan (1)	—	Xu9829	UG	8	NA	NA	Costa Rica	—	1	2	0.250	0.180	0.00022	0.386	8
X. umbratilis [4]	Lake Arenal (San Carlos)	San Juan (1)	—	Xu0634	UG	4	10.4736	-84.8222	Costa Rica	—	0	1	0.000	0.000	0.00000	0.000	4
X. umbratilis [4]	Rio Agua Caliente (San Carlos)	San Juan (1)	—	Xu0717	UG	4	10.4350	-84.7233	Costa Rica	—	1	2	0.500	0.265	0.00044	5.450	4
X. umbratilis [4]	La Vuelta del Borracho (San Carlos)	San Juan (1)	—	Xu0718	UG	4	10.4275	-84.7522	Costa Rica	—	0	1	0.000	0.000	0.00000	0.000	4
Means:						**5.696**					**1.190**	**2**	**0.359**	**0.145**	**0.00038**	**0.740**	**5.696**
Overall minimum value:											**0**	**1**	**0**	**0**	**0**	**0**	**1**
Overall maximum value:											**54**	**7**	**1**	**0.5**	**0.0244**	**25.92**	**54**

This table presents detailed sampling information for all three species in our study, including site names (with bodies of water that each sampled tributary flows into given in parentheses) where local sub-populations were sampled, site IDs corresponding to map numbers in **Fig. 2**, population group membership as inferred from BARRIER analyses (**Fig. 3**), number of samples (*N*), and latitude and longitude data in decimal degrees. For each site, we also list GenBank accession numbers (nos.) corresponding to sequences generated and/or analyzed for all individuals from that site. Numbers of segregating sites (*S*) determining the number of haplotypes (*h*), haplotype diversity (*Hd*) and its standard error (s.e.), nucleotide diversity (π), Watterson's theta (θ_w, an estimator of population mutation rate), and *N* used for summary statistics calculations are presented. At the bottom of each species list, intraspecific mean values are given in bold for *N* and summary statistics (see text for further details). The overall ranges (min. and max. across sites within species, across taxa) of each statistic are summarized at the bottom of the table in bold. Symbols and abbreviations: §, indicates some samples were removed prior to analyses for a given site (e.g., due to missing data) and this explains discrepancies between overall *N*s (column 7) and sample sizes summary statistics were calculated from; dr., drainage(s); NA, not available.

References

1. Hrbek T, Seckinger J, Meyer A (2007) A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. Mol Phylogenet Evol 43: 986-998.
2. Doadrio I, Perea S, Alcaraz L, Hernandez N (2009) Molecular phylogeny and biogeography of the Cuban genus *Girardinus* Poey, 1854 and relationships within the tribe Girardinini (Actinopterygii, Poeciliidae). Mol Phylogenet Evol 50: 16-30.
3. Lee JB, Johnson JB (2009) Biogeography of the livebearing fish *Poecilia gillii* in Costa Rica: are phylogeographical breaks congruent with community boundaries? Mol Ecol 18: 4088-4101.
4. Jones CP, Johnson JB (2009) Phylogeography of the livebearer *Xenophallus umbratilis* (Teleostei: Poeciliidae): glacial cycles and sea level change predict diversification of a freshwater tropical fish. Mol Ecol 18: 1640-1653.

Table S2 DNA substitution models selected using DT-MODSEL

Species	DNA dataset	*N*	bp	Best model	Analysis
Alfaro cultratus					
	Full cyt*b* database	355	601	HKY+*I*	BEAST (BSP), DNASP, TCS (network)
	1st codon pos.	355	201	K80	BEAST (BSP)
	2nd codon pos.	355	200	F81	BEAST (BSP)
	3rd codon pos.	355	200	TrN+Γ	BEAST (BSP)
	A. cultratus cyt*b* haplotypes (*N* = 46) + 1 *A. huberi* outgroup sequence	47	601	HKY+Γ	GARLI (ML)
	1st codon pos.	47	201	TrNef	GARLI (ML)
	2nd codon pos.	47	200	F81	GARLI (ML)
	3rd codon pos.	47	200	HKY+Γ	GARLI (ML)
Poecilia gillii					
	Full cyt*b* database	143	1140	TrN+*I*	BEAST (BSP), DNASP, TCS (network)
	1st codon pos.	143	379	K80	BEAST (BSP)
	2nd codon pos.	143	379	F81	BEAST (BSP)
	3rd codon pos.	143	379	TrN+Γ	BEAST (BSP)
	P. gillii cyt*b* haplotypes (*N* = 37) + 1 *P. mexicana* outgroup sequence	38	1140	TrN+*I*	GARLI (ML)
	1st codon pos.	38	379	K80	GARLI (ML)
	2nd codon pos.	38	379	F81	GARLI (ML)
	3rd codon pos.	38	379	TrN+*I*	GARLI (ML)
Xenophallus umbratilis					
	Full cyt*b* database	131	1140	TrN+Γ	BEAST (BSP), DNASP, TCS (network)
	1st codon pos.	131	379	K80	BEAST (BSP)
	2nd codon pos.	131	379	F81	BEAST (BSP)
	3rd codon pos.	131	379	TrN	BEAST (BSP)
	Xenophallus cyt*b* haplotypes (*N* = 36) + 1 *Priapichthys annectens* outgroup sequence	37	1140	HKY+Γ	GARLI (ML)
	1st codon pos.	37	379	SYM	GARLI (ML)
	2nd codon pos.	37	379	HKY	GARLI (ML)
	3rd codon pos.	37	379	TrN	GARLI (ML)

Model selection analyses using the decision theory algorithm in DT-MODSEL [1] supported different best-fit models of DNA evolution for different datasets across taxa, including datasets filtered by codon positions. This table lists model selection results for intraspecific cyt*b* datasets analyzed in this study, as well as the analyses that each dataset (thus molecular model) was used in. Symbols and abbreviations: *Γ*, gamma-distributed rate variation; bp, number of nucleotide base pairs; BSP, Bayesian skyline plot and associated demographic modeling and Bayes factor analyses; DNASP, DNA polymorphism, mismatch distribution, and neutrality statistics analyses conducted in the program by the same name; *I*, parameter

representing proportion of invariable sites; ML, phylogenetic maximum likelihood analyses estimating haplotype gene trees; *N*, sample size. Although different model selection algorithms, such as MODELTEST [2] and JMODELTEST [3], are available that have historically been more widely used than DT-MODSEL, we preferred to use this software for our substitution model selection analyses because DT-MODSEL has been shown to recover better models than these other programs [1].

References

1. Minin V, Abdo Z, Joyce P, Sullivan J (2003) Performance-based selection of likelihood models for phylogeny estimation. Syst Biol 52: 674-683.
2. Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817-818.
3. Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25: 1253-1256.

Table S3 Cyt*b* DNA polymorphism levels within and among drainage basins

Species	Drainage	*S*	*h*	*Hd*	s.e.	π	*k*	θ_w	Mean *N*
A. cultratus									
	San Juan	14.154	3.692	0.449	0.097	0.005	3.242	4.110	22.250
	Tortuguero	27.000	6.000	0.818	0.084	0.016	9.758	8.941	12.000
	Parismina	7.500	4.000	0.596	0.106	0.004	2.492	2.261	16.000
	Sixaola	27.500	5.000	0.927	0.158	0.021	12.676	12.055	11.000
	mean:	**19.038**	**4.673**	**0.698**	**0.111**	**0.012**	**7.042**	**6.842**	**15.313**
P. gillii									
	San Juan	7.700	3.000	0.530	0.153	0.0033	3.795	3.732	7.182
	Tortuguero	18.000	5.000	0.705	0.088	0.0037	4.210	5.536	15.000
	Parismina	24.000	2.000	0.400	0.134	0.0092	10.533	11.520	5.667
	Matina	24.000	3.000	0.572	0.161	0.0055	6.286	9.256	8.000
	Sixaola	6.000	2.000	0.304	0.082	0.0027	3.072	2.314	8.000
	mean:	**15.940**	**3.000**	**0.502**	**0.124**	**0.0049**	**5.579**	**6.472**	**8.770**
Xenophallus									
	San Juan	1.176	2.000	0.307	0.122	0.00033	0.372	0.761	5.889
	Tortuguero	2.000	2.000	0.536	0.123	0.00094	1.071	0.771	4.500
	Parismina	1.000	2.000	0.767	0.336	0.00068	0.767	0.719	4.000
	mean:	**1.392**	**2.000**	**0.536**	**0.194**	**0.00065**	**0.736**	**0.750**	**4.796**

Refer to text for description of DNA polymorphism statistics. *Xenophallus* samples from Rio Tempisque, while included in our other analyses for this species (e.g., of the full cyt*b* dataset; see Table S1), are excluded here because this drainage does not occur in the main study area (thus, data from this drainage are excluded from mean back-arc drainage group comparisons in the text).

Table S4. Model priors, estimated number and timing of divergence events, and Bayes factors from MTML-msBayes. Results are presented for four coalescent models (M1-M4) run in MTML-msBayes. Bayes factors were used to conduct hypotheses tests of posterior support for simultaneous divergence (e.g., $\Psi=1$) versus other hypotheses.

Model:	M1	M2	M3	M4	mean
Prior settings					
upper θ	0.0049	0.01	0.05	0.0049	–
lower θ	4×10^{-8}	4×10^{-8}	4×10^{-8}	4×10^{-8}	–
upper τ	2	2	2	2	–
no. τ classes (Ψ)	[1]0	0	0	0	–
Nm	0	0	0	0	–
up. ancestral θ	0.5	0.5	0.5	0.25	–
constrain=	0	0	0	0	–
Posterior estimates					
mean Ψ [2]	2.582	2.291	2.132	2.157	2.219
BPP Ψ=1	0.0859	0.149	0.218	0.203	0.211
mode $E[\tau]$	0.644	1.173	0.395	1.0497	0.815
$E[\tau]$ 95% Cis	[0.000,1.312]	[0.445,1.527]	[0.0208,0.585]	[0.270,1.652]	–
Div. time (Ma) [3]	0.644	1.173	0.395	1.0497	0.815
Div. time 95% CIs	[0.000,1.312]	[0.445,1.527]	[0.0208,0.585]	[0.270,1.652]	–
Div. time (Ma) [4]	1.431	2.607	0.878	2.333	1.812
Div. time 95% CIs	[0.000,2.916]	[0.989,3.393]	[0.0462,1.300]	[0.600,3.671]	–
mean Ω	0.626	0.269	0.312	0.481	0.367
Ω 95% HPDs	[0.252,1.218]	[0.000,0.657]	[0.000,0.734]	[0.104,1.0001]	–
Bayes factors (B_{10})					
Comparison					
Ψ=1 vs. Ψ>1	0.9955	0.9296	0.7634	1.0181	1.0175
Ψ>1 vs. Ψ=1	**4.018**	**4.3028**	**5.2391**	**3.9287**	**4.0771**
Ω>0.01 vs. Ω<0.01	1.535	1.525	1.539	**1.558**	1.539
Ψ=3 vs. Ψ<3	0.9643	0.9469	1.0833	1.009	0.960

The population mutation parameter θ is in units of per site per generation. In the mean divergence time hyper-parameter $E[\tau]$, τ is the mean divergence time of the population pairs (calculated from $\tau_1, \ldots, \tau_Y$ population pairs), in coalescent time units of $4N$ generations. In the Nm parameter representing the effective number of migrants per generation, m denotes the probability of symmetric post-divergence migration between sister lineages. Results given in this table are based on coalescent simulations (5×10^6 iterations) of Y=3 population pairs, following which an accept/reject algorithm with tolerance set to 0.0002 was used to create a distribution of 999 draws from the prior, to approximate the joint posterior distribution. The only exceptions are the Ω Bayes factor comparisons (B_{10}), which were based on 9,999 draws from the prior. Estimated 95%

confidence intervals are given in brackets. B_{10} Bayes factors >2 are presented in bold, and B_{10} values indicating 'substantial' support in favor of the alternative hypothesis (values >3.2; based on guidelines in Jeffreys [1]) are further underlined. Abbreviations: BPP, Bayesian posterior probability; CIs, confidence intervals; Div. time, divergence time; Ma, millions of years ago (assuming generation time = 1 yr/generation); no., number of; up., upper.

[1]Here, zero specifies that Ψ were drawn from a set of 1–3, or up to the total number of taxon pairs.

[2]Posterior probability estimated from polychotomous regression (i.e., local multinomial logit regression).

[3]Calculated as Div. time = $E[\tau] \times [(0.5 \times \text{upper } \theta)/\mu]$, assuming the standard pairwise 2% vertebrate mtDNA rate [2,3].

[4]Similar to [3], but calculated assuming the pairwise 0.9% salmonid mtDNA rate [4].

References

1. Jeffreys H (1961) Theory of probability. Third Edition. Oxford, UK: Oxford University Press. 447 p.
2. Brown WM, George M Jr, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 76: 1967-1971.
3. Wilson AC, Cann RL, Carr SM, George M, Gyllensten UB, Helm KM, Bychowski R, Higuchi RG, Palumbi SR, Prager EM, Sage RD, Stoneking M (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. Biol J Linn Soc 26: 385-400.
4. Martin AP, Palumbi SR (1993) Body size, metabolic rate, generation time, and the molecular clock. Proc Natl Acad Sci USA 90: 4087-4091.

Appendix S1: Sampling and outgroups details

In addition to sampling information provided in the main text and figures, we provide detailed information on collection localities for each of our focal species in **Table S1** of the Supporting Information, including site names and geographical coordinates. Geographical coordinates are given in decimal degrees format (these were used during SAMOVA and BARRIER analyses; see **Appendix S2**) and should be highly accurate as we took them from the ground in Costa Rica using hand-held GPS devices (Honduran samples were similarly derived on-site by W. Matamoros, who provide the samples). **Table S1** also lists GenBank accession numbers for new *Alfaro cultratus* sequences generated in this study, as well as those for sequences of *Poecilia gillii* (from Lee and Johnson [1]) and *Xenophallus umbratilis* (from Jones and Johnson [2]) used in this study.

As noted in the main text, we used one or more *Alfaro huberi* samples collected for this study as outgroups during our *A. cultratus* analyses. For maximum likelihood (ML) phylogenetic analysis in GARLI, we used one *A. huberi* sample (GenBank accession no.: XXXXXXX, haplotype 47, **Table S1**), while for the Hudson-Kreitman-Aguadé test (HKA; [3]) conducted in DNASP, we used as data for the outgroup species all seven *A. huberi* samples listed in **Table S1**, which collapsed into three distinct cyt*b* haplotypes (H47-H49). Outgroups added to the other species cyt*b* alignments during GARLI analyses included published cyt*b* sequences for two additional species of livebearing fishes (Poeciliidae) obtained from GenBank. Specifically, based on phylogenetic hypotheses of Hrbek *et al.* [4] and Ptacek and Breden [5], *Priapichthys annectens* (GenBank no: EF017542, genotype/isolate ID "Panne" from Hrbek *et al.* [4]) was the outgroup for *Xenophallus*, and *Poecilia mexicana* (GenBank no: FJ178776, genotype/isolate ID "3211MEX", locality "Col River, Veracruz" Mexico, from Doadrio *et al.* [6]) was the outgroup for *P. gillii*. These same sequences served as outgroups during HKA tests conducted on the full cyt*b* databases of *P. gillii* and *Xenophallus* used in this study.

References

1. Lee JB, Johnson JB (2009) Biogeography of the livebearing fish *Poecilia gillii* in Costa Rica: are phylogeographical breaks congruent with community boundaries? Mol Ecol 18: 4088-4101.

2. Jones CP, Johnson JB (2009) Phylogeography of the livebearer *Xenophallus umbratilis* (Teleostei: Poeciliidae): glacial cycles and sea level change predict diversification of a freshwater tropical fish. Mol Ecol 18: 1640–1653.
3. Hudson RR, Kreitman M, Aguadé M (1987) A test of neutral molecular evolution based on nucleotide data. Genetics 116: 153-159.
4. Hrbek T, Seckinger J, Meyer A (2007) A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. Mol Phylogenet Evol 43: 986-998.
5. Ptacek MB, Breden F (1998) Phylogenetic relationships among the mollies (Poeciliidae: *Poecilia: Mollienesia* group) based on mitochondrial DNA sequences. J Fish Biol 53: 64-81.
6. Doadrio I, Perea S, Alcaraz L, Hernandez N (2009) Molecular phylogeny and biogeography of the Cuban genus *Girardinus* Poey, 1854 and relationships within the tribe Girardinini (Actinopterygii, Poeciliidae). Mol Phylogenet Evol 50: 16-30.

Appendix S2: SAMOVA and BARRIER methods and results

The SAMOVA algorithm [1] and Monmonier's [2] algorithm, as implemented for studying phylogeographic data in BARRIER [3,4], comprise two recent and widely used methods for detecting the presence of genetic barriers and population structure. Several previous single species phylogeography studies have used BARRIER and SAMOVA to identify groups of populations, which were then used for further statistical population genetics analyses of the same datasets (e.g., to conform to the expectation of the methods/models that there is no underlying population structure in the data influencing the results), and to estimate areas where important landscape features or environmental changes may have historically isolated local populations or impeded gene flow (e.g., [5]). It is also clear from the literature that these methods are well suited for comparative analyses. In-line with our study, comparative phylogeographical analyses such as a well-known review and meta-analysis of eastern North American phylogeography by Soltis *et al.* [6], and a recent analysis by Poelchau and Hamrick [7] of three codistributed lower Central American tree species that today share overlapping distributions relative to our study taxa, have used Monmonier's algorithm to identify important genetic barriers within multiple codistributed taxa, in order to test for spatial phylogeographical congruence.

Genetic 'barriers' are areas of maximum rates of genetic change across a landscape (discussed in [1]), and while both SAMOVA and BARRIER use Voroni network-based methods for defining genetic barriers, these methods are different and therefore highly complementary. For example, because SAMOVA directly estimates population structure (positions of homogeneous, maximally genetically differentiated groups or "populations") while taking spatial sampling positions into account and indirectly defining genetic barriers as areas between the inferred populations, whereas Monmonier's algorithm directly reconstructs genetic barriers and thus indirectly identifies population grouping schemes [1]. In either case, both of these methods permit recovering an estimate of the spatial positions of the unknown number, *K*, of actual (presumably panmictic) homogeneous breeding populations within a species. Dupanloup *et al.* [1] showed, through population genetics simulations, that SAMOVA performs best out of the two methods at identifying maximally genetically diverged groups, whereas BARRIER is more proficient at finding the actual number of *K* population groupings.

We implemented both of these methods as a combined test of spatial-genetic congruence among our three focal freshwater fish taxa, to evaluate whether these species exhibited shared

patterns of genetic barriers reflecting potentially shared evolutionary history. We used each method because it is not firmly established which method is best for identifying comparative phylogeographical congruence. Thus we preferred to look for cross-validation across methods as evidence that our comparative inferences were 'strongly supported', i.e., repeatable and robust to different underlying assumptions of different methods. Here, it is important to note, as pointed out by Garrick *et al.* [8], that seeking cross-validation in this way is only valid when results are compared across methods that have similar underlying purposes, as in our study. In our study, comparative phylogeographical congruence would be strongly supported by similar geographical positions of inferred population groups and barriers across all three species. However, we assumed that rigid spatial congruence (of inferred barriers/populations) across taxa along all network edges was not a requirement for arriving at a basis for biologically meaningful interpretation of the data. Instead, we recognized that identifying partial spatial-genetic congruence in a limited part of the study area would still present an opportunity for making further inferences, if only over smaller spatial scales than the entire sampling extent.

Population genetic simulations suggest the largest mean F_{CT} value among a series of SAMOVA models with different initial settings may accurately recover the unknown number of groups (K), and that the point at which increasing F_{CT} values asymptote often represents a meaningful estimate of K [1]. Our rationale behind interpreting the 'best' number of groups determined from our SAMOVA model results stemmed explicitly from these findings. However, we used Φ_{CT}, the F_{CT} analog for DNA sequences analyzed under the analysis of variance framework [9], as the basis of our interpretations. Both of these "$_{CT}$"-subscripted statistics represent the amount of molecular genetic variance present in the overall sample that is explained by among-group variation. It is also noteworthy to point out that *Xenophallus* and *P. gillii* SAMOVA results conformed to the expectation that Φ_{CT} increase with K [1], with Φ-value plateaus respectively supporting $K = 9$ and $K = 6$ distinct groups (**Fig. S2**). However, this behavior was not observed in *A. cultratus*. In light of the inferred patterns of phylogenetic clades of *A. cultratus* and their relationships based on maximum-likelihood phylogenetic gene tree analyses and network analyses, which corresponded to the inferred $K = 2$ SAMOVA groups (and thus also to the barriers inferred by running Monmonier's algorithm on the *A. cultratus* data), it seemed highly appropriate to interpret this deviation in *A. cultratus* as a natural outcome of $K = 2$ being the best model. In other words, our interpretation in light of additional evidence is that *A.*

cultratus likely deviates from the expectation that Φ_{CT} increase with K, as a consequence of Φ_{CT} peaking at $K = 2$ groups.

To round out our discussion of Φ-statistics above and in the main text, we note here that in contrast to Φ_{CT}, Φ_{SC} is the correlation of the diversity of random haplotypes within sub-populations (localities) relative to random pairs from the same group of sub-populations (within regions); whereas Φ_{ST} is the correlation of random haplotypes within sub-populations relative to random pairs drawn from the entire dataset (analogous to F_{ST}). In addition to Φ_{CT}, we also report Φ_{SC} and Φ_{ST} from independent AMOVAs testing what we determined to be the best grouping schemes inferred from our SAMOVA/BARRIER models (see Table 2).

References

1. Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11: 2571-2581.
2. Monmonier MS (1973) Maximum-difference barriers: an alternative numerical regionalization method. Geogr Anal 3: 245-261.
3. Manni FE, Guerard E, Heyer E (2004a) Geographical patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by "Monmonier's algorithm". Hum Biol 76: 173-190.
4. Manni FE, Guerard E, Heyer E (2004b) BARRIER 2.2. Museum of Mankind, Paris, France. Available at: http://www.mnhn.fr/mnhn/ecoanthropologie/software/barrier.html.
5. Ribeiro RA, Lemos-Filho JP, Ramos ACS, Lovato MB (2011) Phylogeography of the endangered rosewood *Dalbergia nigra* (Fabaceae): insights into the evolutionary history and conservation of the Brazilian Atlantic Forest. Heredity 106: 46-57.
6. Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006). Comparative phylogeography of unglaciated eastern North America. Mol Ecol 15: 4261-4293.
7. Poelchau MF, Hamrick JL (2011) Comparative phylogeography of three common Neotropical tree species. J Biogeogr, doi:10.1111/j.1365-2699.2011.02599.x.
8. Garrick RC, Caccone A, Sunnucks P (2010) Inference of population history by coupling exploratory and model-driven phylogeographic analyses. Int J Mol Sci 11: 1190-1227.
9. Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479-491.

Appendix S3: Coalescent divergence time estimation: IMA2 methods

The 'isolation-with-migration' model implemented through the MCMC procedure available in IMA2 uses a procedure that samples many coalescent genealogies, uses them to capture what the data says about the parameters in the model (m_1, m_2, θ_1, θ_2, θ_A, t), and then uses the genealogies to estimate the posterior density of the parameters [1-3]. IMA2 also estimates the TMRCA from the many genealogies sampled. We ran IMA2 with simple two-population models, although the program accommodates >1 ancestral population thus >2 modern populations, because we analyzed a single locus, and these other more complicated models require much more data [4]. It is important to note that, despite providing reliable methods for modeling population history while accounting for potentially confounding processes (e.g., migration, mutational stochasticity), coalescent-genealogy sampling methods including IMA2 make several limiting assumptions and have their own peculiarities. Because space was not permitting in the main text we briefly discuss (i) the assumptions and limitations of this program here, and we also give more detailed information on our IMA2 (ii) analyses and (iii) results.

The first assumption that IMA2 makes is (1) that the data being analyzed are neutrally evolving DNA markers and not influenced by the effects of directional selection or purifying selection (e.g., selective sweeps) [2]. We tested this assumption and found that our mtDNA data met the expectation of neutrality, e.g., based on HKA tests (see "Genetic diversity and neutrality" section, Results). (2) IMA2 also assumes no recombination, and our data meet this criterion: mtDNA are not subject to detectable recombination events. (3) The model implemented in IMA2 also assumes that the populations are not exchanging migrants with any other populations than those modeled and that migration, and that gene flow occurs at a constant rate following population splitting events [2]. Several of the genetically meaningful population groups that we conducted IMA2 analyses on (see BARRIER results) are allopatrically distributed and bounded on their southwestern sides by the North American continental divide (Figs. 1-3); thus these groups seem to fit assumption 2 above, as the next proximal populations sampled seem sufficiently geographically close and isolated as to exclude the possibility of exchange with other (e.g., unsampled) populations. (4) IMA2 also (unlike its predecessor, IM) assumes constant population sizes following initial population splitting (assuming a two-population case or model). Given that our mismatch distribution and neutrality tests generally inferred a shared pattern of population size-constancy for each BARRIER-inferred population group that we

modeled in IMA2, our data are also well suited for IMA2 analysis because they fit this assumption (see "Historical demographic congruence" section, in Results). Moreover, even at the species level, there was only strong evidence for past population dynamics (e.g., in Bayesian skyline models) in *P. gillii*, not the other species (Table 4).

Another relevant point to note is that IMA2 and similar programs cannot identify the timing of migration (whether it occurred before during or after population splitting, or only at present, etc.), although coalescent-inferred migration events most likely occur (or are observed) near the present [5]. As a result, we did not attempt to infer, or test hypotheses based on, posterior-derived estimates of migration timing, although developing methods to address such questions would be a worthwhile endeavor for future research.

In terms of settings, our IMA2 runs employed Hasegawa-Kishino-Yano (HKY) substitution models [6,7]. We used this model because it is the most appropriate model implemented in IMA2 for DNA sequence data, allowing for multiple substitutions and different transition and transversion rates. In contrast, other models selected for the data by DT-MODSEL for our population groups (data not shown) are not implemented in the IMA2 program. Using burn-in periods of 10^6 steps followed by 3×10^6 post-burn-in steps yielded reliable estimates of most parameters in most cases, based on sufficient convergence (e.g., stable trendline plots) and swapping rates of chains (e.g., splitting times were updated at higher rates in higher numbered chains, suggesting acceptable update rates).

In terms of results, our finding that the posterior distributions of *t* values (and estimates of other parameters, but usually only when *m* was not set equal to zero) often peaked at relatively lower *t* values, dropped, and then converged to approximately constant non-zero values is unremarkable. This pattern in the posterior is a common result of single-locus analyses that, despite being non-optimal, still allows excluding the equilibrium migration hypothesis in many cases [8], including our study. In other words, this pattern in our data indicates the peak likelihood represents a model with diverged populations, and this model was more likely than infinite, equilibrium migration, allowing us to exclude this latter hypothesis. Space was also prohibiting in the main text to permit some discussion of other results. For example, whereas we estimated non-zero *m* in *A. cultratus*, peak posterior m values or HPD ranges indicated that on-going gene flow was effectively zero in *P. gillii* and *Xenophallus*. Here, a practical point of note is that, in such cases, uniform *m* priors (the default) are 'truly' non-informative in IMA2. Thus,

for 'zero-migration' population pairs found in *P. gillii* and *Xenophallus*, we conducted additional runs specifying $m = 0$ and these allowed us to achieve better convergence and θ and t parameter estimates, which we report. To account for this issue while permitting low levels of migration, JCB re-ran the IMA2 models for these zero-migration pairs under exponential m priors (-j7 option), modeling migration as a decreasing function with a peak at zero. Results of these exponential-migration runs did not substantially alter or depart from results inferred in the other runs (unpublished data).

References

1. Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics 167: 747-760.
2. Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc Natl Acad Sci USA 104: 2785-2790.
3. Hey J (2010) Isolation with migration models for more than two populations. Mol Biol Evol 27: 905-920.
4. Pinho C, Hey J (2010) Divergence with gene flow: models and data. Annu Rev Ecol Evol S 41: 215-230.
5. Sousa VC, Grelaud A, Hey J (2011) On the nonidentifiability of migration time estimates in isolation with migration models. Mol Ecol 20: 3956-3962.
6. Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22: 160-174.
7. Palsbøll PJ, Berube M, Aguilar A, Notarbartolo di Sciara G, Nielsen R (2004) Discerning between recurrent gene flow and recent divergence under a finite-site mutation model applied to North Atlantic and Mediterranean sea fin whale (*Balaenoptera physalus*) populations. Evolution 58: 670-675.
8. Nielsen R, Beaumont MA (2009) Statistical inferences in phylogeography. Mol Ecol 18: 1034-1047.

Chapter 3: Assessing species boundaries using multilocus species delimitation in a morphologically conserved group of Neotropical freshwater fishes, the *Poecilia sphenops* species complex (Poeciliidae)

Research Article

Assessing species boundaries using multilocus species delimitation in a morphologically conserved group of Neotropical freshwater fishes, the *Poecilia sphenops* species complex (Poeciliidae)

Justin C. Bagley[1]*, Fernando Alda[2], M. Florencia Breitman[3], Eldredge Bermingham[2], Eric P. van den Berghe[4], and Jerald B. Johnson[1,5]

1 Evolutionary Ecology Laboratories, Department of Biology, Brigham Young University, Provo, UT 84602, United States of America, **2** Smithsonian Tropical Research Institute, Balboa, Panamá, **3** Centro Nacional Patagónico (CENPAT-CONICET), U9120ACD, Puerto Madryn, Chubut, Argentina, **4** San Marcos, Carazo, Nicaragua, **5** Monte L. Bean Life Science Museum, Brigham Young University, Provo, Utah 84602, United States of America

Keywords: Bayesian species delimitation; Central America; coalescent; conservation; cryptic species; freshwater fishes; general mixed Yule-coalescent (GMYC); hybridization; incomplete lineage sorting; non-adaptive radiations; *Poecilia*; Poeciliidae; species trees; taxonomy

*Correspondence: Fax: +1 801 422 0090; E-mail: justin.bagley@byu.edu

Short Title: Multilocus species delimitation in *Poecilia* (43/70 characters with spaces)

Word count: ~11,555 (Title, Abstract, and text, minus Table/Figure legends and References)

Abstract (265/300 words)

Accurately delimiting species is fundamentally important for understanding species diversity and distributions and devising effective strategies to conserve biodiversity. However, species delimitation is problematic in many taxa, including 'non-adaptive radiations' containing morphologically cryptic lineages. Fortunately, coalescent-based species delimitation methods hold promise for objectively estimating species limits in such radiations, using multilocus genetic data. Using coalescent-based approaches, we delimit species and infer evolutionary diversification in a morphologically conserved group of Central American freshwater fishes, the *Poecilia sphenops* species complex. Phylogenetic analyses of multiple genetic markers (sequences of two mitochondrial DNA genes and five nuclear loci) from 10/15 species and genetic lineages recognized in the group support the *P. sphenops* species complex as monophyletic, with eight mitochondrial 'major-lineages' diverged by ≥2% pairwise genetic distances. From general mixed Yule-coalescent models, we discovered (conservatively) 10 species within our concatenated mitochondrial DNA dataset, 9 of which were strongly supported by subsequent multilocus Bayesian species delimitation and species tree analyses. Results suggested species-level diversity is underestimated and overestimated by at least ~15% in different lineages in the complex. Nonparametric statistics and coalescent simulations indicate genealogical discordance among our results has mainly derived from interspecific hybridization in the nuclear genome. However, mtDNA show little evidence for introgression, and our species delimitation results appear robust to effects of these processes. Overall, our findings support the utility of combining multiple lines of genetic evidence and broad phylogeographical sampling to discover and validate species using coalescent-based methods. Our study also highlights the importance of testing for hybridization *versus* incomplete lineage sorting, which aids inferring not only species limits but also evolutionary processes influencing genetic diversity.

Introduction

Species are widely used as fundamental units of analysis in biogeography, ecology, and evolutionary biology [1-3]. Species taxa also figure prominently in biodiversity assessments and conservation recovery programs [4]. Therefore, species delimitation, the practice of determining species boundaries and discovering new species, is of fundamental importance for understanding species diversity and distributions, and devising effective strategies to conserve biodiversity [5-7]. By contrast, inaccurately classifying individuals or populations to species could result in erroneous inferences in any analysis requiring *a priori* designation of species limits, such as comparative analyses of diversification [8,9], or misallocation of conservation resources and loss of species (e.g. under the U.S. Endangered Species Act of 1973; [5]).

Although species are universally recognized as metapopulation lineages distinct from other such aggregates ('general lineage concept', or GLC; [10-12]), determining which operational criteria should be used to assign individuals to species is a major problem in species delimitation. Independently applying operational criteria with different philosophical bases often yields incongruent species boundaries [5,13,14]. In turn, inconsistent application of operational species concepts creates unstable taxonomy, injecting taxonomic uncertainty into efforts at species enumeration e.g. [6]. In light of practical difficulties presented by applying alternative operational criteria, there is a growing consensus that multiple perspectives from different data-types or analyses are necessary to accurately delimit species, through 'integrative taxonomy', e.g. uniting classical morphology, phylogenetics, and ecological data and modeling [15-17].

The present surge of interest in integrative taxonomy has shifted biologists' focus away from using single operational criteria to sampling multiple lines of evidence, which ideally yields more robust species delimitations [5,17]. However, integrating morphology with genetic data is notoriously difficult in a variety of contexts. Some examples include: (1) morphologically

conserved, ‘non-adaptive radiations’ containing cryptic species [18-21]; (2) systems with high taxonomic uncertainty; (3) recently assembled communities of relatively young landforms, e.g. oceanic island archipelagos [22]; (4) rapid and recent adaptive radiations e.g. [23]; and (5) taxa with porous species boundaries e.g. [24]. In the former two cases, morphological methods often fail to detect cryptic species and are prone to underestimate species diversity [8]; thus, integrative taxonomic approaches combining morphology with other data will likely yield discordant inferences promoting subjective interpretations. Reliance on morphology can also produce spurious phylogenetic inferences due to disruptive natural selection or insufficient character variation [5,25]. In the latter three cases, speciation can be incomplete or in its early stages, yielding limited genetic variation and higher likelihood of gene tree discordance due to introgressive hybridization e.g. [26] or incomplete lineage sorting (ILS; e.g. [27]). Also in such cases, ‘DNA barcoding’ and single-locus gene trees may fail to establish clear phylogenetic support for fixed geographical differences in morphology e.g. [22].

Recently, the growth of methods for analyzing DNA sequence data in a coalescent-based framework capable of accounting for confounding processes such as ILS [28] has sparked a ‘Renaissance’ in empirical species delimitation (reviewed by [5,29]). Various coalescent-based methods are now available that address different goals in species delimitation, including *de novo* species discovery [30-34], species validation [25,35,36], and assignment of unknown individuals to species e.g. [37]. However, these methods are united in using algorithms modeling evolutionary processes, including likelihood and Bayesian analyses, to identify independent evolutionary lineages as distinct species based on multilocus data and species trees or ‘guide trees’ [29,34]. Indeed, the rapid growth of these methods owes partly to the incorporation of new methods for species tree inference using the multispecies coalescent e.g. [38,39], which has

also revolutionized phylogenetics [40]. Overall, the new wave of coalescent-based species delimitation methods greatly improves the rigor and objectivity of species delimitation, and holds promise for meeting the need for rapid biodiversity assessment and species descriptions [41-43] in light of the current global biodiversity crisis [44].

Although the field of coalescent-based species delimitation is in its infancy, its tools provide solutions to the problems of delimiting species in radiations at the extremes of morphological or genetic divergence (*sensu* [21], their Fig. 1; at least cases 2 and 4 above). For example, aside from delimiting species in "easy-delimitation" scenarios (e.g. deeply diverged lineages with small population sizes; [45]), coalescent-based methods have proven useful for resolving species limits in studies of more difficult cases of morphologically cryptic radiations including trapdoor spiders [46], cave fishes [20], kingsnakes [47], sun skinks [21] and water monitors [48]. In particular, the 'chimeric approach' developing preliminary species hypotheses using parametric or heuristic methods often applied to mitochondrial DNA (mtDNA), then validating these using Bayesian species delimitation with multiple genetic loci [25], appears to be a fruitful way forward (pioneered by Leaché & Fujita [43]; also see [20,21,49]). Under this approach, working hypotheses of species distributions are established and tested using multilocus data and methods taking ILS into account, and the results provide bases for subsequent tests of species morphological and ecological distinctiveness in an integrative taxonomy framework [5].

In this study, we use a coalescent-based chimeric approach to delimit species and expand on previous knowledge of the patterns and processes of diversification in a morphologically conserved radiation—livebearing freshwater fishes in the *Poecilia sphenops* species complex (family Poeciliidae) [50,51]. Despite being among the most common members of regional fish communities in the Mesoamerica biodiversity hotspot [52-54], species limits and taxonomy are

incompletely resolved in the group (reviewed by [55]). Here, we develop the most comprehensive geographical sampling and multilocus sequencing from across the distribution of the *P. sphenops* species complex to date, to delimit species and evaluate their evolutionary genetic relationships. Our objectives were (1) to develop preliminary species delimitation hypotheses using mtDNA; (2) to infer the species tree and timing of lineage diversification using relaxed molecular clocks; (3) to test species validity using multilocus Bayesian species delimitation; and (4) to test model fit and potential sources of gene tree discordance. We use our results to evaluate the validity of cryptic genetic lineages and nominal taxa currently recognized in this group, and to clarify species present distributions.

Materials and Methods

Systematic Background

The systematics of the genus *Poecilia* Bloch & Schneider 1801 has had a tumultuous history, with multiple changes since its initial description, including redescriptions and synonymizations. The currently accepted taxonomy of *Poecilia* recognizes four subgenera: *Limia*, *Pamphorichthys*, *Lebistes*, and *Poecilia*, also known as *Mollienesia* (*sensu* [56]). *Mollienesia* contains 15 to 25 species distributed from North to South America that fall into two species groups distinguished by differences in dorsal fin size and behavior—'sail-fin' and 'short-fin' species [50,51,56-58]. However, much taxonomic confusion in *Mollienesia* owes to their conserved morphology, which obscures interspecific variation; for example, diagnostic characters may overlap, and species display plasticity such that intraspecific phenotypic variance can outpace divergence between species [59,60]. Indeed, the morphologically conserved nature of *Mollienesia* led early workers to conclude that all short-fins represented 'races' or local variants of a single polytypic taxon, *P. sphenops* Valenciennes 1864, ranging from the Río

Grande drainage in northeastern Mexico to coastal Venezuela [61-63]. However, another more widely accepted view is that the short-fin group is composed of biological species with partially overlapping ranges that constitute the '*P. sphenops* species complex' [51,55,64,65].

The *P. sphenops* species complex is a monophyletic group of ~13 described species that is widely distributed along Atlantic and Pacific slopes throughout Mexico and the Central American Neotropics, from the Río Grande through Panama [51,55]. Some authors suggest that this complex can be further sub-divided into two sub-complexes: a '*P. sphenops* complex' including species from the Pacific slope of Mexico through Central America, and a '*P. mexicana* complex' including species from Atlantic coastal Mexico to Nicaragua [64,65]. Over their range, these two otherwise morphologically confusing complexes are distinguished in having tricuspid and unicuspid inner jaw teeth, respectively [50,55]. Also, recent molecular phylogenetic and biogeographic analyses of the group recovered the two complexes as well-supported mitochondrial clades within the *P. sphenops* species complex [55].

Table S1 summarizes the proposed taxonomic arrangements, tooth morphology, and currently recognized geographical distributions of species in the *P. sphenops* species complex. Although some species (e.g. *P. catemaconis* in Lake Catemaco, Mexico) are local endemics with restricted distributions, several others (e.g. *P. sphenops*) have relatively large ranges and occur along Atlantic and Pacific slopes (Table S1). Indeed, the large distribution of some species hinders taxonomic identification because intraspecific morphological gradients or local differentiations are common, and this has been hypothesized to promote character displacement when taxa in the complex occur in sympatry with one another [66].

Ethics Statement

Permission to undertake fieldwork for this study was obtained through permits issued to

JCB and JBJ in Nicaragua by MARENA (Ministerio de Ambiente y Recursos Naturales; DGPN/DB-IC-009-2012; DGPN/DB-21-2012) and in Costa Rica by SINAC-MINAET (Ministerio de Ambiente Energía y Telecomunicaciones; Resolución No. 030-2010-SINAC, Resolución No. 134-2012-SINAC). New specimens were obtained through these collections under Brigham Young University Institutional Animal Care and Use Committee (IACUC) approval #12-0701. By contrast, numerous samples were obtained through government-authorized fieldwork conducted in our previous studies ([55,67]; supplementary Data S1).

Taxon Sampling and Sequencing

We sampled populations of *Poecilia* through field expeditions conducted in Central America, and from the fish tissue archives of our laboratories, the STRI Neotropical Fish Collection (NFC-STRI) and the Monte L. Bean Life Science Museum Fish Collection (BYU). In total, we sampled 873 *Poecilia* individuals from 260 localities (Fig. 1; Data S1). We identified samples to species taxa based on their different combinations of morphology and geographic distributions, following published taxonomy and biogeography studies [50,55,68]. Voucher specimens are deposited at NFC-STRI and BYU.

Of the 13 described species in the *P. sphenops* species complex *sensu lato*, we sampled eight species, including two *P. sphenops* complex species (*P. catemaconis* and *P. sphenops*) and six *P. mexicana* complex species (*P. butleri*, *P. gillii*, *P. hondurensis*, *P. mexicana*, *P. orri*, and *P. salvatoris*) (Table S1). Additionally, we sampled two exclusive mtDNA lineages, or 'operational taxonomic units' (OTUs), identified from recent molecular phylogenetic analyses by Alda *et al.* [55]: "*sphenops*" sp. 1 from Honduras and Nicaragua and "*gillii*" sp. 2 from Rio Acla, Panama. Our own sampling was augmented with sequences of *P. sulphuraria*, *P. thermalis*, and the subspecies *P. mexicana limantouri* from previous studies (see below) and tested each of these

taxa as species-level OTUs. Taking the general lineage concept of species and reconsidering morphology and genetics using a phylogenetic criterion for species identification [11,12], we considered "*P. orri*" samples forming an exclusive genetic lineage from Rio Patuca, Honduras in [55] to be a novel OTU, or 'candidate species', that we refer to as *P.* sp. "Patuca". One motivation for this was that *P.* sp. "Patuca" males possess hooks on their gonopodia (anal fins modified into intromittent organs), whereas a lack of such hooks is a diagnostic character for *P. orri* [69]. We tested this hypothesis by also including in our analyses samples confidently assigned to *P. orri* from Roatan, the next major island adjacent to (~10 km from) the original type locality of *P. orri* at Bonacca Island off the northern Honduras coast [69]. Instead of rigorously evaluating species boundaries using morphological data, we used species diagnoses based on current taxonomy and our interpretation of published phylogenetic relationships as our null hypotheses. This study design amounts to testing hypotheses of species limits based on morphological (e.g. [70]) and/or phylogenetic criteria (genealogical or diagnostic, as in [10,12]) for empirical recognition of species. Our final dataset encompassed 10 out of 15 putative species-level lineages or OTUs recognized in the group (Table S1), including 10 described taxa (species and subspecies) and most of the geographic range of the complex. We also sampled four poeciliid outgroups: *P. latipinna*, *P. latipunctata* (Mexico), *Limia perugiae* (Hispaniola), and *P. caucana* (Panama) samples; yet we analyzed up to 15 outgroup taxa, including samples from genomic repositories, to obtain phylogenetic calibration points. Detailed outgroup data, including outgroups for each analysis, are given in the text or in supplementary Appendix S1.

We extracted whole genomic DNA from tissue samples using Qiagen DNeasy Tissue Kits (QIAGEN Sciences, Maryland, USA) and sequenced the protein-coding mitochondrial cytochrome *b* (cyt*b*) gene for every individual, except problematic *P. orri* and *P. salvatoris*

samples, using two primers flanking the gene, listed in Table 1. To obtain additional mtDNA characters for analysis, we sequenced the mtDNA cytochrome oxidase 1 (*cox1*) gene for individuals chosen to maximize geographic and phylogenetic coverage of mtDNA major-lineages, using fish 'barcode' primers (Table 1). In pilot analyses, *cox1* subsampling improved the mtDNA gene tree topology (data not shown); however, it appeared that sequencing every individual for *cox1* would not provide any added benefit, as expected when subsampling linked mitochondrial genes [24,71]. We also sequenced five nuclear DNA (nDNA) loci: ribosomal protein S7 (*RPS7*; introns 1 and 2 and exon 2); muscle-type lactate dehydrogenase (*ldh-A*); tyrosine-kinase class oncogenes, X-*src* and X-*yes*; and glycosyltransferase (*Glyt*). Because they showed restricted genetic variation and we could not sequence every individual for each locus, we sequenced the nuclear loci for subsamples chosen to maximize geographic and phylogenetic coverage, which we used for species tree and species validation analyses. We attempted to sequence the nuclear loci for 1–5 individuals from each major-lineage identified in our mtDNA-haplotype parsimony networks. With the exception of *ldh-A*, we amplified nuclear loci via nested polymerase chain reactions (PCR), as described in Table 1 and [72]. We purified PCR products using a Montage PCR 96 plate (Millipore, Billerica, MA, USA). Sequences were obtained via cycle sequencing with Big Dye 3.1 dye terminator chemistry using 1/16th reaction size and the manufacturer's protocol (Applied Biosystems, Foster City, CA, USA). We purified sequenced products using Sephadex^{TM} columns (G.E. Healthcare, Piscataway, NJ, USA) and ran them on an automated Applied Biosystems 3730*xl* capillary sequencer. We edited sequences using Sequencher v4.10.1 (Gene Codes Corporation, Ann Arbor, MI, USA). GenBank accession numbers are provided for all sequences in supplementary Data S1.

Mitochondrial DNA sequences contained no gaps and were aligned by visual inspection

in Sequencher; however, nuclear sequences were aligned in MAFFT v6.850 [73] using the local pair FFTS algorithm with a gap opening penalty of 1.53, a tree rebuilding number of 10, and MAXITERATE = 50. We used PHASE v2.1 [74,75] to determine the most probable pair of alleles for each of the nuclear loci, by resolving heterozygous sites. We ran PHASE in DnaSP v5.10 [76] for 100 iterations with thinning interval = 1 and 'burn-in' = 100. We ran three PHASE trials per locus to ensure consistency among phased allelic positions over the output probability threshold, and we used phased alleles in our analyses wherever possible (Appendix S1).

We collated four datasets used in our analyses. First, we created a 'full-cyt*b*' dataset of 941 *Poecilia* sequences by augmenting our database with 68 Mexican cyt*b* sequences (37 haplotypes) from Palacios *et al.* [77]; this increased our ingroup (with *P. sulphuraria*, *P. thermalis*, *P. butleri*, *P. mexicana mexicana*, and *P. m. limantouri*) and outgroup (*P. latipinna* and *P. latipunctata*) sampling. Using TCS v1.21 [78], we collapsed identical ingroup cyt*b* sequences into haplotypes, then generated a statistical parsimony network of ingroup haplotype clades (95% connection limit; data not shown) that we used as a basis for selecting individuals to sequence for subsampling at *cox1* and nuclear loci. A second 'concatenated mtDNA' dataset was comprised of 171 mtDNA cyt*b* (n = 155) and *cox1* (n = 115) subsamples spanning all mtDNA major-lineages, taxa, and OTUs that we sampled. Third, a 'concatenated nDNA' dataset contained 50 ingroup samples for up to 5 nuclear loci. Last, a fourth 'concatenated mtDNA + nDNA' dataset contained 80 ingroup (n = 50) and outgroup (n = 30) samples sequenced from 6 loci, including the mtDNA locus and up to 5 nuclear loci. We included sequences from [77] that formed exclusive mtDNA major-lineages in each dataset, except the concatenated nDNA dataset.

Neutrality and Recombination

We evaluated the selective neutrality of each mtDNA gene in our analysis using Hudson-Kreitman-Aguadé tests (HKA; [79]) in DnaSP, testing significance using 1000 coalescent simulations. We ran HKA tests using *P. caucana* sequences as outgroups, following [55]. We tested each nuclear locus for recombination using six tests implemented in RDP3 v3.44 [80] and described in Appendix S1. We also tested for recombination using 1000 coalescent simulations of the minimum number of recombination events (R_M), assuming the empirical per-gene level of recombination estimated in DnaSP. All parameters were simulated in DnaSP given mutation parameter θ ($=4N_e\mu$ for autosomal nuclear loci; for mtDNA, $\theta = 2N_{ef}\mu$). We considered evidence for recombination significant if a majority of the seven methods detected recombination events.

Gene Tree Analyses and Sequence Divergence

We estimated gene trees for *P. sphenops* species complex haplotypes and outgroup sequences in the concatenated mtDNA, concatenated nDNA (overall, and for each locus), and concatenated mtDNA + nDNA datasets using maximum-likelihood (ML) tree searches in GARLI v2.0 [81]. In GARLI, we partitioned the mtDNA data by codon position ({1+2}, 3) and the nDNA into data subsets by gene. We assigned each data subset its best-fit nucleotide substitution model (Table S2) selected using the decision-theory algorithm DT-ModSel [82], and we unlinked parameters across data subsets. We evaluated nodal support using 500 ML bootstrap pseudoreplicates, considering nodes with bootstrap proportions (BP) ≥70 well supported [83]. We also estimated gene trees, divergence times, and evolutionary parameters (e.g. substitution rates) for each locus using Bayesian inference analyses. To obtain an ultrametric time tree for species delimitation analyses below, we conducted a coalescent-dating analysis of the concatenated mtDNA dataset in BEAST v2.0.2 [84]. We linked tree and clock

models but partitioned the data into codon position subsets ({1+2}, 3) and unlinked site parameters across subsets. To ensure convergence, we ran three replicate searches (MCMC = 10^8, sampled every 4000 generations; burn-in = 10%) using relaxed, uncorrelated lognormal (ULN) molecular clocks. Birth-death tree priors were selected for each run, since this process is well suited for multispecies datasets with varying degrees of lineage divergence. We set uniform priors on ULN clock rates spanning protein-coding mitochondrial gene substitution rates for teleost fishes ('fish rate' = 0.017–0.14 × 10^{-8} substitutions/site/yr, per-lineage; refs. in [85,86]). Including *Poecilia* (subgenus *Limia*) outgroups in these analyses provided a calibration point constraining the split between *P.* (*L.*) *domicensis* from Cuba and *P.* (*L.*) *vittata* from Hispaniola to 17–14 million years ago (Ma), based on phylogenetic data [87] and dates for the geological separation of Cuba and Hispaniola, following [55] and refs. therein. We calibrated this node using a lognormal prior (mean in real space = 1, log standard deviation = 1.25, offset = 14). We used a similar calibration to constrain the tree's root age to 39.9 Ma with an extended tail (log standard deviation = 2.5), based on the oldest fossil poeciliids available from the Maiz Gordo and Lumbrera formations, Argentina [88]. We also estimated a gene tree for each nuclear locus in BEAST using short runs (MCMC = 20 million, sampled every 1000 generations; burn-in = 10%) specifying ULN clocks and birth-death tree priors. We summarized posterior parameter distributions and ensured that effective sample sizes (ESS) were >200 in Tracer v1.5 [89]. We summarized the posterior distribution of trees from each run by calculating a maximum clade credibility (MCC) tree annotated with median node ages from a sample of 5000 post-burn-in trees in TreeAnnotator v2.0.2 [84].

We estimated evolutionary sequence divergences between major-lineages in the concatenated mtDNA gene trees, and between distinct genetic lineages identified as species in

our species delimitation analyses below, using genetic distances. Mean among-clade *p*-distances were calculated in MEGA5 [90] as the number of base differences per site, averaged over all corresponding sequence pairs between groups in the full-cyt*b* dataset. We evaluated variance in the *p*-distances by estimating their standard errors using 500 bootstrap replicates. For comparison, we also estimated divergence between each of these ingroup clades and the two outgroup 'sail-fin' molly species (*P. latipinna* and *P. latipunctata*). We archived our sequence alignments and ML and Bayesian gene tree results in Dryad (doi:10.5061/dryad.XXX).

Coalescent-based Species Delimitation

We delimited species in the *P. sphenops* species complex using a multi-tiered Bayesian approach involving an initial species discovery step, followed by species validation. We base this 'chimeric' approach on previous studies [21,43,49], and recognition that the accuracy of species validation methods relies critically on accurate *a priori* species assignments, as well as guide trees (see below). First, we used the general mixed Yule-coalescent model (GMYC; [30,34]) to assign individuals to species and develop a preliminary set of hypothesized species limits. The GMYC identifies the transition point between speciational and coalescent branching processes on an ultrametric time tree derived from single-locus data [30]. Importantly, the model makes standard coalescent assumptions (neutrality, constant population size and mutation rate, no extinction) but no *a priori* assumptions about species boundaries. We used the Bayesian GMYC model implemented in the R package bGMYC [34] to discover species in the MCC tree from the concatenated mtDNA matrix. By accounting for phylogenetic error and allowing multiple threshold points across the tree (*cf.* [31]), bGMYC overcomes two main shortcomings of Pons *et al.*'s [30] original ML model. As bGMYC is prone to over-split trees containing identical alleles (i.e. zero-length branches) into species [34], we dropped any zero-length tips

from the MCC tree prior to analyses, then ran bGMYC using the single- and multiple-threshold models. For conservativeness and increased statistical power at species discovery (lower false positive, or Type I error, rate), we interpreted results as significant at a modified $\alpha = 0.10$ level. Tree depth heavily influences GMYC results so that transition points may not be detectable when speciation and coalescence rates are similar [34]. Thus, we checked speciation and coalescence rates in the MCC tree empirically using the python script "PTP.py" [91]. We also tested the assumption that the MCC tree contained two classes of branching processes, by performing likelihood-ratio tests comparing single- and multiple-threshold ML GMYC models against null models with one branching process (implying either that all tips are species, or the data represent a single species) in the R package SPLITS v2 [92].

Next, we used two Bayesian methods to validate and better infer the evolutionary history of the GMYC-delimited species: we estimated a multilocus species tree and divergence times and then independently tested the validity of each (originally mtDNA-inferred) species by estimating its Bayesian posterior probability (PP) on the species tree using only nuclear loci. We inferred the species tree and divergence times for the delimited species using the multispecies coalescent *BEAST method [39] implemented in BEAST. We ran *BEAST using all loci in the concatenated mtDNA + nDNA dataset and assigning individual sequences to 25 species, including delimited ingroup species plus 15 outgroup taxa (see Results, Appendix S1). Outgroups permitted setting two calibration points on the same nodes using lognormal priors identical to those in the calibrated BEAST analyses above. We ran *BEAST for five runs of 200 million generations each, sampling every 5000 generations, using Yule tree priors. Log files from each run were combined using LogCombiner v2.0.2 [84] and we visually checked the final log for proper MCMC convergence and mixing and ensured that ESS scores were >200 in

Tracer. Tree files were reduced in size and combined before a MCC tree was computed from 5000 post-burn-in trees in TreeAnnotator.

We tested the validity of the GMYC-delimited species using the Bayesian species delimitation method implemented in BP&P v2.1 [25], which uses a reverse-jump MCMC (rjMCMC) algorithm to generate marginal posterior probabilities for species-delimitation models using multilocus genetic data. BP&P accounts for gene tree variance and ILS, and calculates mutation-scaled population size (θ) and divergence time (τ) estimates. BP&P also assumes that no gene flow occurs following speciation, analogous to the biological species criterion of Mayr [93]. We ran BP&P on the concatenated mtDNA + nDNA dataset fully partitioned by gene, using the *BEAST species tree as a guide tree, and specifying a Dirichlet distribution ($\alpha = 2$) to account for variation in mutation rates among loci. Because BP&P is sensitive to the choice of priors [94], we assessed the impact of prior specification on our results by conducting runs using three different combinations of gamma-distributed priors for ancestral θ and root age (τ_0) [43]: large ancestral populations and deep divergences, $\theta \sim G(1, 10)$ and $\tau_0 \sim G(1, 10)$; small ancestral populations and shallow divergences, $\theta \sim G(2, 2000)$ and $\tau_0 \sim G(2, 2000)$; and a highly conservative prior with large ancestral populations and recent divergences, $\theta \sim G(1, 10)$ and $\tau_0 \sim G(2, 2000)$. We made three replicate runs (rjMCMC = 10^6; burn-in = 25,000) of each prior combination using algorithm 0 (default fine-tuning parameter, $\varepsilon = 15$) and algorithm 1 ($\alpha = 2$, $m = 1$). We conservatively accepted daughter lineages from nodes with speciation probabilities ≥0.95 across all three priors as strongly supported species.

Hybridization *Versus* Incomplete Lineage Sorting

Our analyses indicated several points of discordance between gene trees derived from different loci (see Results), which is often caused by hybridization-mediated introgression, or

ILS arising from the retention of ancestral polymorphisms [38,95]. Whereas these two confounding genetic processes are difficult to tease apart, a recent molecular study of the *P. sphenops* species complex by Alda *et al.* [55] inferred hybridization at the nuclear *RPS7* locus between two pairs of lineages in the complex that we also sampled in this study, *P. catemaconis*-*P. sphenops* and *P. mexicana*-"*gillii*" sp. 2, but no evidence for mtDNA hybridization. Thus, available data suggest that incongruences we observed among gene trees, particularly between mtDNA and nDNA gene trees, may be due to introgression in the nuclear genome. We conducted multiple analyses to determine whether the source of gene tree discordance was more likely due to gene flow *versus* ILS. First, we estimated the degree of exclusive ancestry of individuals of species as quantified by the genealogical sorting index (*gsi*; [96]). The *gsi* spans values normalized to the interval [0, 1], with 1 indicating monophyly, <1 indicating paraphyly, and 0 indicating non-exclusive ancestry in relation to other sampled species. We calculated *gsi* for delimited species based on ML gene trees derived from the concatenated mtDNA dataset, each nuclear locus, and the concatenated mtDNA + nDNA dataset. We also calculated an 'ensemble' *gsi* statistic (gsi_T) as the weighted sum of *gsi* across all five nuclear gene trees. Cummings *et al.* [96] showed that, by integrating across multiple loci, gsi_T has sufficient power to detect significant genealogical divergence well before monophyly is reached, even using small numbers of loci. Analyses were run on the *gsi* web server (http://www.genealogicalsorting.org) while assigning individuals to delimited species, and testing significance using 10^4 permutations.

Second, we used Joly *et al.*'s [95] method for detecting hybridization from species trees, as implemented in JML v1.0.2 [97]. JML uses posterior predictive checking to detect hybridization by testing the fit of a null model with no hybridization (but ILS) to sequence data, through simulations conducted on a posterior sample of species trees from *BEAST (thereby

accounting for phylogenetic error). We supplied JML with 1000 post-burn-in species trees from a *BEAST analysis consisting of five independent runs similar to those above (assigning individuals to delimited species, MCMC = 200 million, burn-in = 10%, birth-death tree priors, and a constant multispecies coalescent population function) but using ingroup samples. We then simulated gene trees and DNA sequence datasets on each species tree under a neutral coalescent model with no migration. For simulations, we specified ML estimates of model parameters from GARLI, evolutionary rates estimated in *BEAST, and appropriate heredity scalars (2 for nDNA, 0.5 for mtDNA) for each locus. We ran separate simulations drawing on ingroup mtDNA sequences from the concatenated mtDNA + nDNA dataset, plus the three nDNA loci with the most sampling (*ldh-A*, *RPS7*, and X-*src*). For each simulated dataset, we computed distributions of the minimum pairwise sequence distance between sequences of two species (*minDist*), a good predictor of hybridization events [95]. We evaluated fit of the ILS model (i.e. adequacy of *BEAST model fit to the data) by comparing *minDist* for the observed data to that of the simulated datasets, to calculate the probability that observed distances were due to hybridization. Using a one-tailed test, we rejected the ILS model at the $\alpha = 0.05$ level in favor of hybridization being the most likely explanation for observed DNA polymorphism patterns between species pairs [97]. For nDNA loci, we only considered significant results meaningful for taxa with observed sequence data, rather than simulated data alone (e.g. the case of clade 7), because while observed sequences are optional for JML an observed pair of aligned sequences is required to calculate exact probabilities of *minDist* values.

Results

Neutrality and Recombination

Based on HKA tests, DNA polymorphism levels in the mtDNA data were consistent with

expectations of neutral evolution, which was assumed in each of our analyses ($P > 0.05$; details in Appendix S1). Likewise, an outstanding majority of tests (91.4%) recovered no evidence for recombination in any of the nuclear loci analyzed (Appendix S1): six tests of each of five loci in RDP3 inferred a total of only three recombination signals (all in X-*yes*), and coalescent simulations showed no evidence of recombination based on R_M values ($P > 0.05$; Appendix S1).

Gene Tree Analyses and Sequence Divergence

The concatenated mtDNA dataset consisted of 1770 nucleotide base pairs (bp), including a 1086 bp fragment of cyt*b* and 684 bp of the partial *cox1* gene and flanking serine tRNA (Table S3). The ML gene tree derived from this dataset had a ln *L* of −12419.3689 and generally recovered well supported relationships among ingroup lineages, with BP >70% for most tip clades and internal nodes (Fig. 2). However, mtDNA lineages in the gene tree provided a variable fit to nominal taxonomy and currently recognized OTUs [51,55]. Haplotypes of *P. butleri*, *P. gillii*, "*gillii*" sp. 2, *P. hondurensis*, and *P.* sp. "Patuca" were recovered as highly supported monophyletic groups, and relationships among these lineages received moderate to high bootstrap support. Members of the *P. sphenops* complex *sensu stricto*, including *P. catemaconis*, *P. sphenops*, and "*sphenops*" sp. 1, were also monophyletic, although *P. sphenops* monophyly was poorly supported. By contrast, *P. mexicana* was polyphyletic, with samples from Rio Tipitapa, Nicaragua between Lake Managua and Lake Nicaragua (Fig. 1) recovered in a monophyletic group at the base of the complex *sensu lato*; and *P. orri* and *P. salvatoris* were each paraphyletic, nested within the principal *P. mexicana* clade. The position of the Tipitapa lineage was poorly resolved by mtDNA, and its sister relationship to all other *P. sphenops* species complex lineages received marginal support, yet given its genetic distinctiveness we refer to this *P. mexicana*-like lineage as a 'candidate species', *P.* sp. "Tipitapa". We also

recovered *P. thermalis* in a clade containing *P. sulphuraria*; however, these taxa shared identical cyt*b* haplotypes. For convenience of presentation and discussion, we identified eight mtDNA major-lineages (clades 1–8) in the gene tree differentiated by ≥2% mean among-clade *p*-distances (range 2.3–9.9%), which we visualized with distinct colors. We also identified 17 exclusive, moderate to strongly supported 'subclades' contained within these major-lineages (2-a to 8-j) in the mtDNA gene tree.

The BEAST relaxed clock analysis of the concatenated mtDNA dataset converged on a mean *L* of −12,590.73 and had good sampling properties (e.g. ESS > 316). From this run, we generated a MCC time tree (highest log clade credibility = −139.6855; Fig. S1A) that recovered ingroup relationships identical to the mtDNA ML gene tree, but with higher nodal support values (e.g. PP = 0.95–1 for most ingroup tip clades and internal nodes; Fig. 2). Unlike the mtDNA ML gene tree, however, we recovered one Lake Nicaragua tributary sample (172554) sister to other clade 2 samples with strong support (Fig. S1A). Nuclear genes in the concatenated nDNA dataset (3484 bp), and also in the concatenated mtDNA + nDNA dataset, were on average 685 bp long (range 191–967 bp), and averaged 59.6 variable characters, 45 parsimony informative characters, and 0.017 overall mean *d* based on *p*-distances (Table S3). Phylogenetic structuring in the ML gene tree derived from the concatenated mtDNA + nDNA dataset (ln *L* = −21,978.2011) mirrored relationships recovered in the concatenated mtDNA gene trees, except *P.* sp. "Tipitapa" was recovered sister to the *P. sphenops* complex *sensu stricto* (clade 2) with high support, clade 8-j was recovered in a monophyletic group with representatives of clades 8-a and 8-b, and while phylogeographical sub-structuring in clade 8-c was well supported the monophyly of clade 8-c itself was poorly supported (Fig. 3A). The concatenated nDNA gene tree was relatively less resolved than the other gene trees but also placed *P.* sp. "Tipitapa" sister

to clade 2 with moderate support, and strongly supported the monophyly of clades 1–4 (Fig. 3B). Evaluating each nuclear locus separately also indicated lower resolution, and along with varying degrees of genetic variation we observed differing degrees of species monophyly at different loci (Fig. S2; Table S3). Although different methods and datasets varied in the levels of support assigned to nodes in the tree, all of the analyses essentially identified the same major-lineages and recovered *P. orri*, *P. salvatoris*, and *P. thermalis* as paraphyletic (Fig. 2; Figs. S2 & S3).

Coalescent-based Species Delimitation

Separate bGMYC runs specifying different models gave very similar preliminary hypotheses of species boundaries, although the multiple-threshold model estimated finer groupings leading to slightly higher species diversity than the single-threshold model. Running the single-threshold model gave a pattern of 11 species that met our criteria (Fig. S3A), eight of which corresponded to mtDNA major-lineages identified using the gene tree and *p*-distances (Fig. 2). Similarly, the multiple-threshold model supported 14 species (Fig. S3B). Both models assigned species status to the single tip sample 172554 from clade 2 and sample 23082 from clade 5; however, such low allele sampling is non-optimal for bGMYC, and sample 172554 was consistently recovered within clade 2-b in the mtDNA ML gene tree analysis with strong support, so we conservatively considered only the subclades in these groups/clades defined by multiple individuals as potential species (subclades 2-a and 2-b; *cf*. [21]). Thus, we accepted a more conservative number and arrangement of clusters of 10 species with multiple individuals from the single-threshold bGMYC analysis as our preliminary species delimitation hypothesis. Rate calculations indicated that the GMYC results were unlikely to be confounded by proximal speciation and coalescence rates, which diverged widely (speciation rate per substitution, λ_s = 19.64; coalescent rate per substitution, λ_c = 508.67). Moreover, likelihood-ratio tests performed

in SPLITS confirmed that the two classes of branching processes assumed in the model were present in the tree (single-threshold test: null ln L = 637.54, max. ln L = 644.70, likelihood ratio = 14.32, $P < 0.01$; multiple-threshold test: null ln L = 637.54, max. ln L = 646.08, likelihood ratio = 16.93, $P < 0.01$).

The relaxed clock *BEAST species tree (mean L = −21,865.97, ESS = 1,382.69) inferred relationships among the *P. sphenops* species complex that were identical to those recovered in the concatenated mtDNA + nDNA ML tree, placing a strongly-supported monophyletic group containing clades 1 and 2 sister to all other members of the *P. sphenops* species complex *sensu lato* with strong support (Fig. S1B). Predictably, subsamples representing phylogeographic structuring within clades 2 and 8 were recovered as monophyletic. However, the monophyletic group containing clades 3–8 differed from the concatenated mtDNA gene tree in placing clade 6 sister to clades 5 + 7–8, rather than clade 5 sister to clades 6–8 (as in Figs. 2, 3A), although relationships among these clades were poorly supported. Based on the time to the most recent common ancestor (t_{MRCA}) estimated by *BEAST for the stem node splitting a *P. caucana* + 'short-fin' mollies clade and the ingroup, we inferred a maximally early-mid Miocene origin for the ancestral ingroup population in Central America [median age = 16.4 Ma, 95% highest posterior density (HPD) = 23.2–11.1]. Moreover, the ingroup t_{MRCA} indicated the diversification of the *P. sphenops* species complex *sensu lato* most likely began 17. 6–8.1 Ma (median age = 12.2) in the Miocene and continued to the present. *Poecilia* sp. "Tipitapa" was the oldest species (median age = 9.2 Ma, 95% HPD = 14.4–5.4), whereas *P. sphenops* complex clades 2-a and 2-b were the youngest delimited species, with a Plio-Pleistocene t_{MRCA} (median age = 2.4 Ma, 95% HPD = 4.7–0.54).

Running BP&P with algorithm 1 under priors reflecting different historical scenarios

strongly supported each of the 9 delimited species examined with high speciation probabilities (Fig. 4). However, the clade 8 crown nodes containing phylogeographical structuring between subclades 8-a–8-c and subclades 8-e–8-j received significant support from the models with large and small ancestral sizes and deep divergences (PP = 1), but no support from the small ancestral size, shallow divergence model (PP = 0). Quantitatively and qualitatively similar results were obtained in identical runs using algorithm 0 (Fig. S4). Given uncertainty in the internal nodes of our species tree, we also ran BP&P on the concatenated mtDNA + nDNA ML gene tree topology, and this yielded near-identical results. Thus, multilocus Bayesian species delimitation based on the present sampling strongly supports recognizing clades 1, 2-a, 2-b, and 3-8 as distinct species with 95% Bayesian posterior probability, but indicates that phylogeographical lineages within clade 8 receives substantial but not definitive support and cannot be treated as distinct species. Clade 7 was only evaluated in BP&P using mtDNA sequences from [77]; however, its monophyly and significant nodal support in the ML and Bayesian gene trees (Figs. 2 & 3), high Bayesian PP in the GMYC results, and the mtDNA *gsi* results below, indicate that clade 7 would likely have been strongly supported as a distinct species in BP&P had nDNA loci been available for this lineage.

Hybridization *Versus* Incomplete Lineage Sorting

Permutation tests of the *gsi* calculated from the mtDNA ML gene tree in Fig. 2 supported each bGMYC-delimited species as a monophyletic lineage in relation to other delimited species, with approximate to complete lineage sorting ($P < 0.001$; Table 3). Likewise, *gsi* tests supported all delimited species monophyly in the concatenated mtDNA + nDNA gene tree. We also detected significant genealogical divergence and sorting at different nDNA loci for most species, despite a lack of monophyly (18/27, or 67% of cases; Table 3). However, *gsi* values expectedly

fluctuated across nDNA loci, with values for loci with more variable characters tending to be higher, and with delimited species being consistently significantly sorted at *RPS7* and X-*src* but less consistently so at other loci (Tables 2 & 3). Still, all taxa with nuclear data had significant ensemble *gsi* scores (mean $gsi_T = 0.384$) across the nuclear gene trees ($P < 0.05$; Table 3).

We detected no instances of introgression in ingroup mtDNA based on 1000 coalescent simulations in JML. Thus, we conclude that post-speciation hybridization at mtDNA is unlikely, and that the multispecies coalescent model in *BEAST provides a good fit to the mtDNA data. Therefore, the mtDNA are also consistent with assumptions of BP&P [25]. By contrast, JML simulations consistently detected introgressed sequences between *P. butleri*-*P. catemaconis/sphenops* (clade 2-a) species pairs across all three nuclear loci examined, based on significant departures of observed *minDist* values from the posterior predictive distributions (*ldh-A*, $P = 0.001$; *RPS7*, $P = 0.001$; and X-*src*, $P = 0.001$; additional results in Appendix S1). These results suggest that the *BEAST model provides an inadequate fit to these three nuclear markers because it assumes that all gene tree discordance is due to ILS. Overall, our JML results indicate that the probability of obtaining para-/polyphyletic nDNA gene trees but monophyletic mtDNA gene trees is high, and that gene tree discordances observed in this study have likely resulted from hybridization instead of ILS in the nuclear genome. In particular, the low PP for the placement of *P. butleri* in the species tree (Fig. S2) seems likely due to hybridization.

Discussion

A growing number of empirical studies suggest that newly developed coalescent-based species delimitation methods [29] provide effective tools for delimiting species in morphologically conserved groups with cryptic species, using independent genetic loci e.g. [9,20,21,43,46-48]. Indeed, these methods are recommended to overcome the limited utility of

morphology to delimit species in these systems, e.g. few diagnostic characters distinguishing species [21,43,46]. One advantage of coalescent-based methods is that, whereas earlier species delimitation approaches based solely on phylogenetic criteria ('phylogenetic species concepts') required strict assumptions of monophyly and fixed allelic differences at one or more genetic loci (reviewed by [12]), coalescent species delimitation relaxes these constraints, given such patterns are not expected in multilocus datasets [26,38,98]. Thus, coalescent-based species delimitation methods can identify independently evolving lineages representing distinct species through probabilistic tests of alternative speciation hypotheses (e.g. different resolutions of species tree branches) while allowing for gene tree discordance and ILS (reviewed by [29]). Using a "chimeric approach" [46] combining coalescent methods for single-locus species discovery without assuming species boundaries *a priori* (i.e. Bayesian GMYC modeling; [34]), and Bayesian species delimitation using multiple independent loci (i.e. BP&P; [25]), we set out to delimit species and infer evolutionary relationships in a morphologically conserved group of Central American freshwater fishes, the *Poecilia sphenops* species complex [50,51]. Other studies have used similar approaches to delimit terrestrial and freshwater species, and served as bases for new species descriptions in several cases [9,20,43,46,47,49]. Yet ours is the first attempt to resolve taxonomic uncertainties in the Central American freshwater biota using coalescent-based species delimitation. Overall, our results provide compelling evidence for incongruence between genetically delimited species and nominal taxonomy indicating diversity is underestimated and overestimated in different lineages of the *P. sphenops* species complex, with important implications for taxonomy and conservation.

Species Delimitation in the *P. sphenops* Species Complex

Many previous systematic studies of poeciliid livebearing fishes, and of the *P. sphenops*

species complex in particular, have relied solely on classical morphology [50,51,59,60,63,68,99,100]. This has imposed an important limitation on studies of *Poecilia* , given the "confusingly variable" nature of morphology in the *P. sphenops* species complex [50], and that fishes in genus *Poecilia* (particularly subgenus *Mollienesia*) may exhibit ample intraspecific variation to swamp interspecific variation, especially at morphometric variables [59,60]. Indeed, after studying *Poecilia* including members of the *P. sphenops* species complex, Rivas [60] concluded, "there is considerable variation in [morphometric] characters individually, ontogenetically, seasonally, geographically, and environmentally and, therefore, they are of little or no value in distinguishing species" (our clarification in brackets). Also, very few meristic or external morphological characters are useful for diagnosing species in the *P. sphenops* species complex, except a handful of characters related to inner jaw tooth dentition, fin-ray counts, and preorbital head pores [50,59,99,100]. Perhaps not surprisingly, morphology-based taxonomy has been extremely confused in the group, with different authors synonymizing up to 34 taxa into *P. sphenops* [63] at one extreme, and recognizing at least six subspecies between *P. mexicana* [68] and *P. gillii* [59] at another. Aside from destabilizing taxonomy in the group, earlier morphological studies also suffered drawbacks of limited spatial sampling, and restricted taxonomic and phylogenetic perspectives focused on one species or species group e.g. [50,67].

Our results from applying coalescent models to genetic data from an extensive geographical sample of 8 of 13 species, one subspecies, and 2/2 molecular OTUs previously recognized in the *P. sphenops* species complex (Fig. 1; Table S1) strongly support at least 9 lineages as distinct 'species'. These include: (1) *P. butleri* (clade 6); (2) *P. catemaconis*/*sphenops* (including *P. catemaconis* and "*sphenops*" sp. 1 samples, clade 2-a); (3) *P. gillii* (clade 5-b and 5-c); (4) *P. hondurensis* (clade 4); (5) *P. mexicana* (clade 8); (6) *P.*

sphenops (clade 2-b); (7) clade 7, including multiple Mexican taxa; (8) the *P.* sp. "Tipitapa" lineage (clade 1), discovered in this study and identified in the field as *P. mexicana*; and (9) the "*gillii*" sp. 2 lineage (clade 3), initially discovered and identified in the field as *P. gillii* by Alda *et al.* [55]. Figure 4 summarizes the placement and inferred taxonomy of each lineage in the species tree, and Fig. 1 provides a map of each lineage's distribution in a regional context. Each species delimited by our full analysis is supported by multiple lines of evidence, including substantial nodal support in gene trees from the mtDNA or combined analyses, high Bayesian posterior probabilities (PP = 0.9–1) of conspecificity during bGMYC modeling, and high Bayesian speciation probabilities (PP = 0.95–1) in coalescent analyses using BP&P (Figs. 2, 3, 4, S1 & S4). Moreover, putative species are distinct from one another by ≥2% and more frequently ≥3% mean pairwise mtDNA genetic distances (Table 2) agrees with expectations derived from worldwide data on divergences between marine and freshwater fish species pairs [101]. That said, the 'cryptic' candidate species in clades 1 and 3 are highly distinct, being the only taxa except their sister lineages (clades 2 and 4) that are both deeply diverged from other lineages by ≥5% mtDNA genetic distances (Table 2) and strongly supported as monophyletic in all gene tree and species tree analyses (Figs. 2, 3, 4). We also inferred no instances of mtDNA hybridization and only one instance of nuclear hybridization involving either of these taxa (the *P. hondurensis–P.* sp. "Tipitapa" *RPS7* comparison in JML; Appendix S1), which if valid appears to represent ancient introgression since these species ranges are presently allopatric, separated by the Chortis Highlands of Nicaragua. In light of this, our findings demonstrate that species-level diversity within the *P. sphenops* species complex is underestimated by at least ~15%, relative to the 13 currently described species (Table S1).

However, we suspect that current diversity within the *P. sphenops* species complex is

underrepresented by our results, most likely in clades 2-a, 5, and 7. The two lineages lumped into clade 2-a, two lineages in clade 5 (*P.* sp. "Patuca" in 5-b, and 5-c), and two lineages in clade 7 (subclades 7-a and 7-b) respectively diverged from one another fairly recently ~0.86, ~1.14, and ~0.71 thousand years ago during the early-mid Pleistocene (Fig. S1). Given recent lineage divergences may cause Bayesian GMYC modeling to undersplit data into species (discussed below), bGMYC may have generated invalid species designations by lumping tips in into one species in these cases (Fig. 2). Clearly, resolving taxonomy in these clades will require additional sampling and analyses of multiple nuclear loci, and a coalescent approach similar to ours is recommended. Our ability to draw conclusions about clade 7 seems particularly limited, as we could not obtain analogous nuclear sequences for samples from Palacios *et al.* [77]. Still, [77]'s multilocus phylogeny recovered clades analogous to subclades 7-a and 7-b (with sub-structuring within 7-a), suggesting future analyses will likely recover these lineages as distinct species.

We have shown that earlier morphological treatments underestimated species-level diversity in the *P. sphenops* species complex and particularly within *P. mexicana* and *P. gillii* e.g. [59,102]. By contrast, our finding that *P. orri* and *P. salvatoris* are paraphyletic with respect to *P. mexicana*, nested within a larger clade otherwise exclusively comprised of *P. mexicana* in the mtDNA gene trees (Figs. 2 & S1), suggests nominal taxonomy likely overestimates diversity in clade 8, possibly by up to ~15%. *Poecilia orri* and *P. salvatoris* were recovered in a well-supported clade in the concatenated mtDNA + nDNA gene tree (Fig. 3A), but *P. salvatoris* was nested within *P. orri* sequences in this clade, and neither of these species was reciprocally monophyletic in our mtDNA or nDNA gene trees (Figs. 2, 3B). Thus, one or both of these taxa may not constitute distinct species, and this is also supported by the fact that neither taxon was

recovered as a distinct species during our species discovery analyses. In fact, bGMYC gave *P. orri* and *P. salvatoris* samples 95% Bayesian posterior probabilities of conspecificity with *P. mexicana* (Fig. S3). Therefore, we suggest that a formal taxonomic revision examining morphological and genetic data be undertaken to determine the status of these taxa.

Combining Species Discovery and Validation: Limitations and Sampling Considerations

Through our use of a "chimeric approach" [46] to coalescent-based species delimitation combining species discovery and validation methods, this study highlights key interactions between phylogenetic and statistical population genetic (coalescent) analyses typically integrated during such analyses e.g. [20,21,43]. Such integration is essential for statistically evaluating evolutionary patterns and processes at the species boundary, the interface between micro- and macroevolution [30]. The particular combination of developing preliminary species delimitation hypotheses through single-locus GMYC modeling, then testing these using multilocus Bayesian species tree and species validation analyses herein also has several strengths. For example, it accounts for gene tree discordance using the multispecies coalescent [39], avoids confounding gene trees with species trees, and also objectively arrives at *a priori* species assignments using coalescent methods implemented *before* conducting separate species validation analyses in BP&P [43,46]. Still, the multiple steps of such chimeric approaches are, overall, subject to several potential limitations, the most important of which we discuss below.

First, uncertainty associated with the topology and branch lengths of ultrametric phylogenies supplied for GMYC modeling can be high because trees are usually derived from single-locus mtDNA datasets. Thus, running GMYC models on a single phylogenetic point estimate could yield inaccurate results, leading to erroneous preliminary hypotheses of species limits [30,34]. Despite this, we consider our GMYC results reasonably accurate, because

bGMYC accounts for phylogenetic and modeling error by integrating over uncertainty in the parameters using Bayesian MCMC simulations. It is also important to note that we supplied bGMYC with a valid ultrametric MCC tree generated from a coalescent-dating analysis in BEAST using appropriate priors, including biogeographic and fossil calibration points (Figs. S1, S3). And our results seem unlikely to reflect confounding effects of branch length uncertainty: analyzing the concatenated mtDNA gene tree (Fig. 2) using a Bayesian application of a method, PTP [91], similar to GMYC but analyzing substitution patterns along gene trees with non-ultrametric branches gave species delimitations comparable to our bGMYC results (unpublished data; details in Appendix S1). This demonstrates that our mtDNA data are robust to the varying assumptions and quantitative approaches of different species discovery methods [30].

Second, recent lineage divergences are also problematic for Bayesian GMYC inference because they induce greater uncertainty into the model and are more likely to occur more recently than inferred threshold points [34]. In our study, the fact that multiple genetic lineages in clades 2-a, 5 and 7 were relatively recently diverged, more so than delimited species, suggests this situation may have caused bGMYC to inaccurately lump these lineages together. This would mean that bGMYC effectively treated actual species-level diversity as intraspecific genetic structuring in these clades. More sampling is necessary to test this hypothesis; however, it may be unrealistic to expect the youngest of these Pleistocene-evolved lineages to fare well in subsequent multilocus validation in BP&P: such recently evolved lineages may not have accumulated enough mutational differences to have high speciation probabilities.

Three additional limitations arise because Bayesian species delimitation using coalescent analyses in BP&P is subject to misspecifications of species limits, guide tree relationships, and model priors [43,94]. Due to the difficulty of confidently establishing species limits *a priori* in

non-adaptive radiations with uncertain taxonomy such as the *P. sphenops* species complex, it is essential that species discovery analyses used to set up BP&P runs be conducted as rigorously as possible [21]. Whereas, as noted above, we feel our bGMYC results are robust, our analyses do not permit us to know whether a multilocus species discovery step, e.g. employing Bayesian assignment tests as per [20,43], would have improved our initial hypotheses of species limits. However, our BP&P results do not seem susceptible to misspecifications of the guide tree or model priors. This is supported by the fact that running BP&P on the species tree as well as the topology from the concatenated mtDNA + nDNA gene tree (unpublished data) gave similar results, and that we obtained consistent results across priors. Based on coalescent theory and previous studies, priors specifying large ancestral θs and recent divergences (τ) are expected to favor the recovery of fewer species in BP&P [25,43]. Moreover, if multiple prior combinations support the one species delimitation while another prior scenario does not, then this may indicate that the data provide a poor fit to the latter prior, and *vice versa* e.g. [48]. Following Leaché & Fujita [43], we varied the prior distributions of population parameters estimated by BP&P by two orders of magnitude, and found that all models unambiguously supported the same nine species.

Last, coalescent-based species delimitation approaches, like all species delimitation methods [5], are subject to the peculiarities of each study's geographical, taxonomic, and character sampling strategies. Of particular concern are potentially negative effects of uneven sampling across distinct genetic lineages, as happens to be the case in our results (e.g. the large bias toward sampling *P. mexicana* in clade 8 *versus* other clades), or missing data on coalescent-based species delimitations. Our sampling is the most comprehensive for the complex to-date at multiple levels, and we sampled most species recognized in the *P. sphenops* species complex prior to this study (Table S1), except for four species with relatively restricted distributions,

known from only 1 to 2 drainage basins in subregions of Mexico (*P. chica*, *P. marcellinoi*, *P. maylandi*) and Belize (*P. teresae*). We acknowledge our phylogenetic inferences are therefore subject to potential effects of missing species. However, lacking some ingroup taxa does not impact coalescent-based inferences of distinct species e.g. [48]; rather, undetected variation from un-sampled species and populations can, at best, only influence inferred phylogeographical patterns and the positions of un-sampled taxa within gene trees and species trees in our data. Nevertheless, the wide geographical-sampling approach employed herein has permitted us to avoid pitfalls of more-limited taxon sampling, and to identify multiple independent evolutionary lineages, including two novel 'cryptic' species.

Hybridization *Versus* Incomplete Lineage Sorting

Identification of the species tree and species limits is a necessary prerequisite for understanding evolutionary genetic processes of hybridization-mediated introgression and incomplete lineage sorting, which are increasingly recognized in natural systems and thought to play a defining role influencing population genetic structure, speciation, and gene tree discordance [2,38,95]. Indeed, studies of these processes are vulnerable to the 'species problem', as they rely on defining species *a priori* before attempts are made to distinguish interspecific *versus* intraspecific processes [24]. Coalescent-based species delimitation provides a sound, objective basis for defining species for such analyses, which can provide important information reciprocally illuminating the nature of the species examined and conservation efforts [29]. Based on these methods, our study demonstrates a distinct pattern of nuclear, but not mitochondrial, hybridization and introgression, rather than ILS, as the main factor likely influencing gene tree discordance in the *P. sphenops* species complex. The presence of clear hybrid zones formed by post-speciation range expansion and secondary contact is a relatively common pattern in natural

populations [2,5,19], but is not indicated in our results. Instead, we infer that some admixture has occurred in the past between species that today are sympatric and/or allopatric, as evidenced by smaller minimum pairwise nuclear genetic distances than that expected from posterior predictive distributions generated using coalescent simulations on species trees in JML [95]. Evidence seems especially complete for *P. butleri* hybridization (e.g. with *P. sphenops*/*catemaconis* in clade 2-a), as we detected introgression between *P. butleri* and other taxa at all nuclear loci analyzed (Appendix S1). The available genetic evidence also apparently confirms previous morphological evidence for natural *P. butleri–P. sphenops* hybridization, including Schultz & Miller's [50] description of a hybrid *P. butleri* × *P.* "*sphenops*" individual. Moreover, whereas *P. mexicana* has traditionally been considered to hybridize rarely with other *Poecilia* [50,51], our results support hybridization between this very widespread species (clade 8, Fig. 1) and several other ingroup taxa (Appendix S1).

Whereas maternally inherited mtDNA genomes are thought to generally introgress more rapidly and therefore to present poor bases for single-locus phylogenetics in various taxa including some fishes [103], our results overwhelmingly support cytonuclear discordance indicating the opposite is true for the *P. sphenops* species complex. This finding agrees with the expectation that nuclear gene flow and hybridization should be higher in systems with female-based dispersal, which is somewhat counterintuitive but supported by theory and empirical review by Petit & Excoffier [104]. Therefore, we hypothesize that a pattern of sexual asymmetry prevails in the *P. sphenops* species complex, with female-biased dispersal promoting intraspecific gene flow that blocks interspecific mtDNA introgression (*cf.* [104], refs. therein). This is the most plausible explanation for the patterns in our results, and underscores a contributing factor as to why mtDNA provide an excellent basis for species delimitation in the

complex (as we have shown). In light of the above findings, that we observed consistency across our results and across species delimitation algorithms suggests that our species delimitations are robust to the effects of hybridization; we have also explicitly incorporated the effects of ILS during multiple modeling procedures, including species discovery and validation analyses.

Phylogenetics and Biogeography

Though a more detailed comparison of our phylogenetic results and those of previous studies is beyond the scope of this study, we note that our findings agree with and expand on previous molecular hypotheses of phylogenetic relationships, hence inferred biogeography and diversification patterns, in the *P. sphenops* species complex [55,58,67,77]. For example, within the complex *sensu lato*, previous molecular and morphological studies recognized two monophyletic sub-complexes that correlated well with inner jaw tooth morphology—the *P. sphenops* complex and *P. mexicana* complex [55,64] (Table S1). Likewise, our multilocus phylogenies support each of these sub-complexes as monophyletic (Figs. 3, 4, S1B). Morphological analyses are needed to determine whether the sub-complexes are reciprocally monophyletic, however, given we recover the undescribed species *P.* sp. "Tipitapa" with morphological affinities for *P. mexicana* but undocumented dentition patterns as sister to the *P. sphenops* complex. Similar to Alda *et al.* [55], we found it difficult to obtain strongly supported relationships at some internodes of our species tree (e.g. resolving relationships among clades 5–8), but results presented here and in [55] are congruent in suggesting that this has resulted from gene tree discordance caused by hybridization in the nuclear genome. By contrast, our six-gene dataset allowed us to obtain a species tree with better support for several relationships (with PP >80–90) than [55]'s species tree. Moreover, we present the first multilocus species tree analysis strongly supporting the monophyly of the *P. sphenops* species complex and relationships within

the *P. sphenops* complex (clades 1 and 2) (Figs. 4 & S1B).

There are several major biogeographical implications of this study that go hand-in-hand with the taxonomic implications discussed below. First, our results clarify the geographical range limits of several taxa and thus aid combating the "Wallacean shortfall", or gaps in our understanding of species distributions, in biodiversity studies (see [3]). Whereas others have considered *P. sphenops* to meet its southern range limit in eastern Guatemala or western Honduras [50,51,55,57], our results suggest that its range (e.g. of clade 2-b) extends further south, terminating at the lake district of Nicaragua, in Lake Nicaragua and its northern tributaries (Figs. 1, 4). Our results also clarify the distribution of 'true' *P. gillii* in clade 5 (the clade corresponding to the original type locality for this species, Rio Chagres; see [55]), which it no longer makes sense to consider as spanning from Guatemala to Panama and perhaps into Colombia e.g. [59,66,99,102]. Instead, we recommend researchers and managers to consider the range of *P. gillii* as extending mainly from Rio Playón Chico, Panama to Rio Parismina, Costa Rica on the Atlantic versant, and from Rio Bayano, Panama to around the western limit of the Rio Térraba basin, Costa Rica on the Pacific versant (Figs. 1, 2). As in [55], we also find *P. mexicana* (clade 8) to have a much wider geographical distribution than previously thought e.g. [50,51]; however, given the uncertain status of Mexican populations in clade 7, we consider *P. mexicana* to extend from at least the Lake Petén Itzá drainage, Guatemala southward to Rio Cuango, Panama on the Atlantic versant, and from Rio Goascorán (the El Salvador-Guatemala border) to the western Rio Bayano basin, Panama on the Pacific versant (Figs. 1, 2).

Second, the timing of diversification of the *P. sphenops* species complex inferred herein (Figs. 4, S1A) is congruent with the results of previous fossil- and biogeography-calibrated, multilocus divergence time analyses by Alda *et al.* [55]. Particularly, our results based on

expanded geographical and character sampling also show that lineage diversification has occurred *in situ* within Central America, and that all major-lineages diversified within the complex prior to the completion of the Isthmus of Panama, which connected North and South America ~3–1.8 Ma (reviewed in [105]). All nine delimited 'species' in our results fit this pattern (Fig. 4), which is consistent with emplacement of the ancestral population of the complex through dispersal into the region from outlying areas of North or South America well before the full development of the Central American Isthmus landscape. We also inferred that the *P. mexicana* and *P. sphenops* complexes initially speciated during the Miocene (17.8–8.1 Ma; Figs. S1, 4), whereas multiple analyses with slightly different calibrations in [55] place the most recent common ancestor of these lineages in a slightly earlier Oligocene-Miocene range (~38–13 Ma), but overlap with our age estimates. These results correspond well to the results of [106], thus multiple datasets are apparently converging on a similar picture of the evolution of this group. Yet our discovery and coalescent-dating of the origin of the 'cryptic' species *P.* sp. "Tipitapa" from Nicaragua provides a unique insight: *in situ* evolution of this species ~9.2 Ma (Figs. 4, S1A) correlates very closely with the origin of the Nicaraguan depression, which formed through southeast-northwestward opening of a rift valley between the Tortuguero lowlands of Costa Rica through the El Salvador Median Trough over 10–0 Ma (reviewed in [105]). This suggests that isolation in the Nicaraguan depression may have caused the initial divergence of this taxon.

Third, and more generally, we find evidence for both widespread and often-sympatric lineages (e.g. clades 2, 5, 8), as well as highly endemic lineages and phylogeographic units (e.g. clades 1, 2-a, 3, 4, and 5-a) (Figs. 1, 2). This suggests several contrasting biogeographical processes have been at play in shaping present-day distributions of species in the *P. sphenops* species complex. In particular, barriers between drainage basins (e.g. mountain ranges bounding

the Nicaraguan depression) have apparently generated prolonged genetic isolation facilitating the development of distinct populations and endemic species within some regions, e.g. isolation of clade 1 within the Rio San Juan basin. At the same time, dispersal barriers have been sufficiently negligible and time has been sufficiently great for some taxa, including *P. mexicana* and *P. gillii*, to obtain relatively extensive distributions across multiple biogeographical areas and physiographic provinces (reviewed in [105]), providing many opportunities for local adaptation and low levels of gene flow with sympatric congeners. These widespread lineages also inhabit very similar habitats [51,53,102], reflecting similar levels of phenotypic plasticity, and/or potentially large-scale ecological adaptation to similar environments. Extinctions have also undoubtedly played a role so that species that were once widespread now have widely disjunct, endemic populations; here, the principal case in point is the wide disjunction between the distributions of differentiated clades 5-c in Honduras, *versus* 5-a and 5-b largely restricted to Panama (Figs. 1, 2). Likely, intervening extinctions created such patterns as a result of the combined effects of marine transgressions, landscape evolution (e.g. orogeny), and climate change at different times in the past, but especially during regional and global upheavals in climate and sea levels during the Plio-Pleistocene (reviewed in [105]).

Taxonomic and Conservation Implications

Coalescent-based analyses such as those employed here should reduce investigator-driven biases in species delimitation, creating more stable and transparent taxonomy [29,43]. In making taxonomic interpretations based on our results, we follow a general lineage concept of species [10-12] and consider genealogical and statistical evidence from multiple unlinked genetic loci sufficient to diagnose independently evolving lineages representing distinct species [9,29,43]. This is considered best practice and is most consistent with recent progress in the

conceptualization of species [12]. However, we acknowledge that evidence from species distributions indicating geographical isolation (e.g. allopatric ranges; [21]) and evidence for fixed morphological or ecological differences relative to other species can also support independent lineages as valid species (*cf.* [14-17,47]), though such differentiation is less likely to be observed in morphologically cryptic taxa.

Our coalescent-based species delimitation results support the distinctiveness of several existing *Poecilia* species. Most of the 9 lineages within the *P. sphenops* species complex delimited as strongly supported species correspond exclusively to nominal taxa and thereby support their continued recognition as distinct species. Specifically, we recognize *P. butleri*, *P. hondurensis*, and *P. mexicana*, as distinct species, as presently defined, with the exception of considering *P. mexicana* to possess a more extensive range reaching Rio Bayano, Panama (Figs. 1 & 4, Table S1). Coalescent species delimitation also non-subjectively delimits at least two undescribed candidate species within *P. mexicana*, including one new species in clade 1, and two species within *P. gillii*, including the new species in clade 3, all of which are diagnosable based on molecular data including analyses of six independent loci. Figure 4 summarizes the placement of each of these lineages in the species tree, and Fig. 1 provides a map of each lineage's distribution in a regional context. Our interpretation that at least two species exist within *P. gillii* is conservative, given the species we consider 'true' *P. gillii* in clade 5 contains three sub-lineages, each of which was strongly supported in phylogenetic analyses of the mtDNA and concatenated mtDNA + nDNA datasets, though not delimited during GMYC species discovery analyses. Although *P. mexicana* and *P. gillii* vary substantially in pigmentation and dorsal fin coloration throughout their ranges [59,102], we are aware of very few morphological characters or ecological attributes distinguishing the two new candidate species within *P.*

mexicana, and aware of no such attributes distinguishing the two species within *P. gillii*. However, as these species are already strongly supported by multilocus molecular data, studies exploring their distributions, ecological niches, and morphology in further detail would provide additional support for their validity (*cf.* [43]). Thus, we recommend that a formal morphological description of each candidate species be undertaken, including an analysis of all related type material and morphological comparisons with closely related species.

Conclusions

Overall, our findings contribute to a growing appreciation of the utility of combining multiple lines of genetic evidence and broad phylogeographical sampling to discover and validate species limits using coalescent-based methods [29,43,92]. Our study also contributes to a more accurate accounting of the biodiversity and geographical distributions of *Poecilia* mollies (subgenus *Mollienesia*), as well as Central American freshwater fishes in general, through objectively delimiting species in the *P. sphenops* species complex using molecular data. The importance of testing for hybridization *versus* ILS on multilocus species trees is also highlighted by our results: distinguishing between these factors allowed us to infer not only species boundaries but also evolutionary processes influencing genetic diversity in the complex, as well as our inferences. In particular, our data support the hypothesis that cytonuclear discordance arises in this complex as a result of female-biased dispersal (although we cannot rule out at least some mtDNA introgression). We recommend additional sampling of *P. sphenops* species complex populations at additional unlinked genetic loci to further improve the taxonomy and biogeography of the group and achieve a phylogenetic analysis with more complete ingroup sampling; however, we highlight the importance of our findings to understanding the biogeographical processes influencing this group, as well as their significance for taxonomy and conservation.

Acknowledgments

We are grateful to E. Castro Nallar, J.T. Nelson, G.R. Reina, A.H. Smith, and E.P. van den Berghe for valuable assistance conducting fieldwork for this project in Costa Rica and Nicaragua. In addition, we thank K.A. Crandall and D. Shiozawa (Brigham Young University), J. Guevara Siquiera (SINAC-MINAET, San José, Costa Rica), E. Duarte (MARENA, Managua, Nicaragua) and E.P. van den Berghe (San Marcos, Nicaragua), and G.R. Reina (STRI) for help obtaining collecting and export permits. We thank A. Bentley (Kansas University Biodiversity Institute), M. Tobler (Oklahoma State University) and W.A. Matamoros (University of Southern Mississippi) for graciously donating ingroup and outgroup samples of *Poecilia* and *Limia* from Mexico, Honduras, and El Salvador, without which this study would not have been possible. We thank P.J. Unmack for providing primers and helpful suggestions for the molecular laboratory work, P.L. Wood Jr. for useful suggestions on species delimitation analyses, and S. Joly for assistance with running the JML analyses. D.S. Rogers, J.W. Sites Jr., P.J. Unmack, L.J. Welton, and P.L. Wood Jr. are also thanked for helpful suggestions on earlier drafts of this manuscript.

References

1. Barraclough TG, Nee S (2001) Phylogenetics and speciation. Trends in Ecology and Evolution 16:391-399.
2. Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Massachusetts.
3. Lomolino M, Riddle BR, Brown JH (2012) *Biogeography*, 6th Edition. Sinauer Associates, Sunderland, Massachusetts.
4. Mace GM (2004) The role of taxonomy in species conservation. Philosophical Transactions of the Royal Society B 359:711-719.
5. Sites JW Jr, Marshall JC (2003) Delimiting species: a Renaissance issue in systematic biology. Trends in Ecology and Evolution 18:462-470.
6. Agapow PM, Bininda-Emonds ORP, Crandall KA, Gittleman JL, Mace GM, Marshall JC, Purvis A (2004) The impact of species concept on biodiversity studies. Quarterly Review of Biology 79:161-179.
7. Agapow PM (2005) Species: demarcation and diversity. In: *Phylogeny and Conservation* (eds Purvis A, Gittleman JL, Brooks T), pp. 57-75. Cambridge University Press, Cambridge, UK.
8. Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. Trends in Ecology and Evolution 22:148-155.
9. 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. doi:10.1093/sysbio/syt099.
10. Mayden RL (1997) A hierarchy of species concepts: the denouement in the saga of the species problem. In: *Species: the Units of Biodiversity* (eds Claridge MF, Dawah HA, Wilson MR), pp.381-424. Chapman & Hall Ltd, London.
11. de Queiroz K (1998) The general lineage concept of species, species criteria, and the process of speciation. In: *Endless Forms: Species and Speciation* (eds Howard DJ, Berlocher SH), pp. 57-75. Oxford University Press, New York.
12. de Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology*, 56, 879-886.
13. Mishler BD, Donoghue MJ (1982) Species concepts: a case for pluralism. Systematic Zoology 31:491-503.
14. Sites JW Jr, Marshall JC (2004) Operational criteria for delimiting species. Annual Review of Ecology, Evolution and Systematics 35:199-227.
15. Dayrat B (2005) Towards integrative taxonomy. Biological Journal of the Linnean Society 85:407-415.
16. 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 Royal Society of London B 106:12418-12423.

17. Padial J, Miralles A, De la Riva I, Vences M (2010) The integrative future of taxonomy. Frontiers in Zoology 7:16.
18. Gittenberger E (1991) What about non-adaptive radiation? Biological Journal of the Linnean Society 43:263-272.
19. Jockusch EL, Wake DB (2002) Falling apart and merging: diversification of slender salamanders (Plethodontidae: *Batrachoseps*) in the American West. Biological Journal of the Linnean Society 76:361-391.
20. Niemiller ML, Near TJ, Fitzpatick BM (2012) Delimiting species using multilocus data: diagnosing cryptic diversity in the southern cavefish, *Typhlichthys subterraneus* (Teleostei: Amblyopsidae). Evolution 66:846-866.
21. Barley AJ, White J, Diesmos AC, Brown RM (2013) The challenge of species delimitation at the extremes: diversification without morphological change in Philippine sun skinks. Evolution 67:3556-3572.
22. Bacon CD, McKenna MJ, Simmons MP, Wagner WL (2012) Evaluating multiple criteria for species delimitation: an empirical example using Hawaiian palms (Arecaceae: *Pritchardia*). BMC Evolutionary Biology 12:23.
23. Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, Seehausen O (2012) 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.
24. Willis SC, Macrander J, Farias IP, Ortí G (2012) Simultaneous delimitation of species and quantification of interspecific hybridization in Amazonian peacock cichlids (genus *Cichla*) using multi-locus data. BMC Evolutionary Biology 12:96.
25. 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.
26. Nosil P, Harmon LJ, Seehausen O (2009) Ecological explanations for (incomplete) speciation. Trends in Ecology and Evolution 24:145-156.
27. Maddison WP (1997) Gene trees in species trees. Systematic Biology 46:523-536.
28. Wakeley J (2008) *Coalescent Theory: An Introduction*. Roberts & Company Publishers, New York.
29. 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.
30. Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed

insects. Systematic Biology 55:595-609.

31. Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJG, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP (2009) Accelerated species inventory on madagascar using coalescent-based models of species delineation. Systematic Biology 58:298-311.
32. Carstens BC, Dewey TA (2010) Species delimitation using a combined coalescent and information-theoretic approach: an example from North American Myotis bats. Systematic Biology 59:400-414.
33. O'Meara BC (2010) New heuristic methods for joint species delimitation and species tree inference. Systematic Biology 59:59-73.
34. Reid NM, Carstens BC (2012) Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evolutionary Biology 12:196.
35. Ence DD, Carstens BC (2011) SpedeSTEM: a rapid and accurate method for species delimitation. Molecular Ecology Resources 11:473-480.
36. Camargo A, Morando M, Avila LJ, Sites JW (2012*a*) Species delimitation with ABC and other coalescent-based methods: a test of accuracy with simulations and an empirical example with lizards of the *Liolaemus darwinii* complex (Squamata: Liolaemidae). Evolution 66:2834-2849.
37. Abdo Z, Golding GB (2007) A step toward barcoding life: a model-based, decision-theoretic method to assign genes to preexisting species groups. Systematic Biology 56:44-56.
38. Degnan JH, Rosenberg MS (2009) Gene tree discordance, phylogenetic inference, and the multispecies coalescent. Trends in Ecology and Evolution 24:332-340.
39. Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution 27:570-580.
40. Edwards SV (2009) Is a new and general theory of molecular systematics emerging? Evolution 63:1-19.
41. Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP (2003) A plea for DNA taxonomy. Trends in Ecology and Evolution 18:70-74.
42. Vogler AP, Monaghan MT (2007) Recent advances in DNA taxonomy. Journal of Zoological Systematics and Evolutionary Research 45:1-10.
43. Leaché AD, Fujita MK (2010) Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proceedings of the Royal Society of London B 277:3071-3077.
44. Pimm SL, Russell GJ, Gittleman JL, Brooks TM (1995) The future of biodiversity. Science 269:347-350.
45. Camargo A, Avila LJ, Morando M, Sites JW Jr (2012*b*) Accuracy and precision of species trees: effects of locus, individual, and base pair sampling on inference of species trees in lizards of the

Liolaemus darwinii group (Squamata, Liolaemidae). Systematic Biology 61:272-288.

46. Satler JD, Carstens BC, Hedin M (2013) Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*). Systematic Biology. doi:10.1093/sysbio/syt041.
47. Myers EA, Rodríguez-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.
48. Welton LJ, Siler CD, Oaks JR, Diesmos AC, Brown RM (2013). Multilocus phylogeny and Bayesian estimates of species boundaries reveal hidden evolutionary relationships and cryptic diversity in Southeast Asian monitor lizards. Molecular Ecology 22:3495-3510.
49. Smith BT, Ribas CC, Whitney BM, Hernández-Baños, Klicka J (2013) Identifying biases at different spatial and temporal scales of diversification: a case study in the Neotropical parrotlet genus *Forpus*. Molecular Ecology 22:483-494.
50. Schultz RJ, Miller RR (1971) Species of the *Poecilia sphenops* complex (Pisces: Poeciliidae) in Mexico. Copeia 1971:282-290.
51. Miller RR (2005) *Freshwater Fishes of México*. The University of Chicago Press, Chicago.
52. Myers N, Mittermeier RA, Mittermeier CG, de Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature 403:853-858.
53. Smith SA, Bermingham E (2005) The biogeography of lower Mesoamerican freshwater fishes. Journal of Biogeography 32:1835-1854.
54. Matamoros WA, Kreiser BR, Schaefer JF (2012) A delineation of Nuclear Central America biogeographical provinces based on river basin faunistic similarities. Reviews in Fish Biology and Fisheries 22:351-365.
55. Alda FA, Reina RG, Doadrio I, Bermingham E (2013) Phylogeny and biogeography of the *Poecilia sphenops* species complex (Actinopterygii, Poeciliidae) in Central America. Molecular Phylogenetics and Evolution 66:1011-1026.
56. Miller RR (1975) Five new species of Mexican poeciliid fishes of the genera *Poecilia*, *Gambusia*, and *Poeciliopsis*. Occasional Papers of the Museum of Zoology of the University of Michigan 672:1-44.
57. Ptacek MB, Breden F (1998) Phylogenetic relationships among the mollies (Poeciliidae: *Poecilia*: *Mollienesia*) based on mitochondrial DNA sequences. Journal of Fish Biology 53:64-82.
58. Breden F, Ptacek MB, Rashed M, Taphorn D, Figueiredo CA (1999) Molecular phylogeny of the live-bearing fish genus *Poecilia* (Cyprinodontiformes: Poecilidae). Molecular Phylogenetics and Evolution 12:95-104.

59. Poeser FN (2003) *From the Amazon river to the Amazon molly and back again. The Taxonomy and Evolution of the genus Poecilia Bloch and Schneider, 1801*, University of Amsterdam, Amsterdam.
60. Rivas LR (1978) A new species of poeciliid fish of the genus *Poecilia* from Hispaniola, with reinstatement and redescription of *P. dominicensis* (Evermann and Clark). Northeast Gulf Science 2:98-112.
61. Hubbs CL (1926) Studies of the fishes of the order Cyprinodontiformes VI. Miscellaneous Publications of the University of Michigan Museum of Zoology 16:1-87.
62. Hubbs CL (1933) Species and hybrids of *Mollienesia*. Aquarium 1:263-268.
63. Rosen DE, Bailey RM (1963) The poeciliid fishes (Cyprinodontiformes), their structure, zoogeography, and systematics. Bulletin of the American Museum of Natural History 126:1-176.
64. Alpírez Quesada O (1971) Estudio sistemático del complejo *Poecilia sphenops* (Familia Poeciliidae) de Centroamérica en especial de las poblaciones de Costa Rica. Universidad de Costa Rica, San José.
65. Poeser FN (2011) A new species of *Poecilia* from Honduras (Teleostei: Poeciliidae). Copeia 2011:418-422.
66. Poeser FN (1998) The role of character displacement in the speciation of Central American members of the genus *Poecilia* (Poeciliidae). Italian Journal of Zoology 65:145-147.
67. Lee JB, Johnson JB (2009) Biogeography of the livebearing fish *Poecilia gillii* in Costa Rica: are phylogeographical breaks congruent with fish community boundaries? Molecular Ecology 18:4088-4101.
68. Menzel BW, Darnell RM (1973) Systematics of *Poecilia mexicana* (Pisces: Poeciliidae) in Northern Mexico. Copeia 1973:225-237.
69. Fowler HW (1943) A new poeciliid fish from Honduras. Notulae Naturae 117:1-3.
70. Cronquist A (1978) Once again, what is a species? In: *Biosystematics in Agriculture* (ed. Knutson LV), pp. 3-20. Allenheld O smun, Montclair, New Jersey.
71. Morando M, Avila LJ, Sites JW Jr (2003) Sampling strategies for delimiting species: genes, individuals, and populations in the *Liolaemus elongatus-kriegi* complex (Squamata: Liolaemidae) in Andean-Patagonian South America. Systematic Biology 52:159-185.
72. Unmack PJ, Bagley JC, Adams M, Hammer MP, Johnson JB (2012) Molecular phylogeny and phylogeography of the Australian freshwater fish genus *Galaxiella*, with an emphasis on dwarf galaxias (*G. pusilla*). PLoS One 7:e38433. doi:10.1371/journal.pone.0038433.
73. Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. Bioinformatics 9:286-298.
74. Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. American Journal of Human Genetics 73:1162-1169.

75. 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.
76. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452.
77. Palacios M, Arias-Rodriguez L, Plath M, Eifert C, Lerp H, Lamboj A, Voelker G, Tobler M (2013) The rediscovery of a long described species reveals additional complexity in speciation patterns of poeciliid fishes in sulfide springs. PLoS One 8:e71069. doi:10.1371/journal.pone.0071069.
78. Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657-1659.
79. Hudson RR, Kreitman M, Aguadé M (1987) A test of neutral molecular evolution based on nucleotide data. Genetics 116:153-159.
80. Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26:2462-2463.
81. Zwickl DJ (2006) *Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets Under the Maximum Likelihood Criterion*, The University of Texas, Austin, Texas.
82. Minin V, Abdo Z, Joyce P, Sullivan J (2003) Performance-based selection of likelihood models for phylogeny estimation. Systematic Biology 52:674-683.
83. Hillis DM, Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Systematic Biology 42:182-192.
84. Bouckaert R, Heled J, Kühnert D, Vaughan TG, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*. Available at: http://beast2.org/.
85. Waters JM, Burridge CP (1999) Extreme intraspecific mitochondrial DNA sequence divergence in *Galaxias maculatus* (Osteichthys: Galaxiidae), one of the world's most widespread freshwater fish. Molecular Phylogenetics and Evolution 11:1-12.
86. Burridge CP, Craw D, Fletcher D, Waters JM (2008) Geological dates and molecular rates: fish DNA sheds light on time dependency. Molecular Biology and Evolution 18:624-633.
87. Hamilton A (2001) Phylogeny of *Limia* (Teleostei: Poeciliidae) based on NADH dehydrogenase subunit 2 sequences. Molecular Phylogenetics and Evolution 19:277-289.
88. Pascual R, Bond M, Vucetich MG (1981) El Subgrupo Santa Bárbara (Grupo Salta) y sus vertebrados. Cronología, paleoambientes y paleobiogeografía. San Luis, Argentina: VIII Congreso Geologico Argentino, Actas III, 746-758.
89. Rambaut A, Drummond AJ (2013) Tracer v1.5. <http://beast.bio.ed.ac.uk/tracer>.

90. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28:2731-2739.
91. Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013) A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29:2869-2876.
92. Fujisawa T, Barraclough TG (2013) Delimiting species using single-locus data and generalized mixed Yule coalescent approach: a revised method and evaluation on simulated data sets. Systematic Biology. doi: 10.1093/sysbio/syt033.
93. Mayr E (1942) *Systematics and the Origin of Species, from the Viewpoint of a Zoologist*. Columbia University Press, New York.
94. Zhang C, Zhang D-X, Yang Z (2011) Evaluation of a Bayesian coalescent method of species delimitation. Systematic Biology 60:747-761.
95. Joly S, McLenachan PA, Lockhart PJ (2009) A statistical approach for distinguishing hybridization and incomplete lineage sorting. American Naturalist 174:e54-e70.
96. Cummings MP, Neel MC, Shaw K (2008) A genealogical approach to quantifying lineage divergence. Evolution 62:2411-2422.
97. Joly S (2012) JML: testing hybridization from species trees. Molecular Ecology Resources 12:179-184.
98. Hudson RR, Coyne JA (2002) Mathematical consequences of the genealogical species concept. Evolution 56:1557-1565.
99. Lucinda PHF, Reis RE (2005) Systematics of the subfamily Poeciliinae Bonaparte (Cyprinodontiformes: Poeciliidae), with an emphasis on the tribe Cnesterodontini Hubbs. Neotropical Ichthyology 3:1-60.
100. Greenfield DW (1990) *Poecilia teresae*, a new species of poeciliid fish from Belize, Central America. Copeia 1990:449-454.
101. Johns GC, Avise JC (1998) A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. Molecular Biology and Evolution 15:1481-1490.
102. Bussing WA (1998) *Freshwater Fishes of Costa Rica*, 2nd Edn. Editorial de la Universidad de Costa Rica, San José, Costa Rica.
103. Chan KMA, Levin SA (2007) Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. Evolution 59:720-729.
104. Petit RJ, Excoffier L (2009) Gene flow and species delimitation. Trends in Ecology and Evolution 24:386-393.
105. Bagley JC, Johnson (2014) Phylogeography and biogeogrpahy of the lower Central American

Neotropics: diversification between two continents and between two seas. Biological Reviews. doi:10.1111/brv.12076.

106. Hrbek T, Seckinger J, Meyer A (2007) A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. Molecular Phylogenetics and Evolution 43:986-998.

107. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD (2005) DNA Barcoding of Australia's fish species. Philosophical Transactions of the Royal Society B 360:1847-1857.

108. Quattro JM, Jones WJ (1999) Amplification primers that target locus-specific introns in actinopterygian fishes. Copeia 1999:191-196.

109. Chow S, Takeyama H (1998) Intron length variation observed in the creatine kinase and ribosomal protein genes of the swordfish *Xiphias gladius*. Fisheries Science 64:397-402.

110. Li C, Ortí G, Zhang G, Lu G (2007) A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. BMC Evolutionary Biology 7:44.

Tables

Table 1. PCR primers and annealing temperatures used to amplify mitochondrial and nuclear markers in this study.

Gene	Primer	Sequence (5'– to –3')	PCR steps§	T_A (annealing temperature, °C)	Reference
cyt*b*	L14725	GAYTTGAARAACCAYCGTTG	Single PCR	48	Hrbek *et al.* [106]
	H15982	CCTAGCTTTGGGAGYTAGG	Single PCR	48	Hrbek *et al.* [106]
cox1	FISH-F1	TCAACCAACCACAAAGACATTGGCAC	Single PCR	48–49	Ward *et al.* [107]
	FISH-R1	TAGACTTCTGGGTGGCCAAAGAATCA	Single PCR	48–49	Ward *et al.* [107]
ldh-A	LDHA6F2	GYGGAGAGCATCSWKAAGAACMTGC	Single PCR	48–49	Quattro & Jones [108]
	LDHA6R*	GCTSAGGAASACCTCRTCCTTCAC	Single PCR	48–49	Quattro & Jones [108]
RPS7	1F	TGGCCTCTTCCTTGGCCGTC	1st PCR	52	Chow & Takeyama [109]
	3R	GCCTTCAGGTCAGAGTTCAT	1st PCR	52	Chow & Takeyama [109]
	1F.2	CTCTTCCTTGGCCGTCGTTG	2nd PCR–1	52	Unmack *et al.* [72]
	2R.67	TACCTGGGARATTCCAGACTC	2nd PCR–1	52	Unmack *et al.* [72]
	2F.2.cat	GCCATGTTCAGTACCAGTGC	2nd PCR–2	52	Unmack *et al.* [72]
	3R.10	TCAGAGTTCATCTCCAGCTC	2nd PCR–2	52	Unmack *et al.* [72]
X-*src*	SRC.E7.1F	TGACAGACGTTTGTCCCGTACTGAAGC	1st PCR	52	Peter J. Unmack
	SRC.E10.endR	ATGAGKCGAGCCAGACCGAAATCAGC	1st PCR	52	Peter J. Unmack
	SRC.E8.1F	CTGAAGCCTGGCACCATGTC	2nd PCR	52	Peter J. Unmack
	SRC.E10.end2R	CCGAAATCAGCCACTTTACAMACCAG	2nd PCR	52	Peter J. Unmack
X-*yes*	Yes F1	GAGAGAATGAACTACATCCATAG	1st PCR	52	Peter J. Unmack
	Yes R1	GACCACACGTCTGATTTGATTGTGAA	1st PCR	52	Peter J. Unmack
	Yes F2	GACAACCTGGTCTGTAAGATCGC	2nd PCR	52	Peter J. Unmack
	Yes R2	GATTTGATTGTGAAGCGACCGTACA	2nd PCR	52	Peter J. Unmack
Glyt	Glyt_F559	GGACTGTCMAAGATGACCACMT	1st PCR	55	Li *et al.* [110]
	Glyt_R1562	CCCAAGAGGTTCTTGTTRAAGAT	1st PCR	55	Li *et al.* [110]

Glyt_F577	ACATGGTACCAGTATGGCTTTGT	2nd PCR	62	Li *et al.* [110]
Glyt_R1464	GTAAGGCATATASGTGTTCTCTCC	2nd PCR	62	Li *et al.* [110]

§Single PCR, only one PCR performed; 1st PCR or 2nd PCR, indicates the sequence in a nested set. Note also that a number *n* preceded by a dash in this column (e.g. "–2") indicates the *n*th second PCR step in a set of nested reactions.

Table 2. Mean pairwise genetic distances among 10 clades accepted as preliminary species hypotheses based on Bayesian general mixed Yule-coalescent (GMYC) results in Fig. 2.

	Clade 1	**Clade 2-a**	**Clade 2-b**	**Clade 3**	**Clade 4**	**Clade 5-a**	**Clade 5-b**	**Clade 6**	**Clade 7**	**Clade 8**
Clade 1	–	0.0080	0.0081	0.0087	0.0089	0.0087	0.0087	0.0082	0.0082	0.0082
Clade 2-a	0.085	–	0.0044	0.0069	0.0072	0.0071	0.0070	0.0072	0.0072	0.0068
Clade 2-b	0.082	0.031	–	0.0072	0.0075	0.0075	0.0076	0.0072	0.0074	0.0071
Clade 3	0.091	0.070	0.070	–	0.0065	0.0068	0.0066	0.0068	0.0067	0.0064
Clade 4	0.086	0.070	0.071	0.052	–	0.0064	0.0064	0.0066	0.0062	0.0057
Clade 5-a	0.094	0.071	0.073	0.060	0.057	–	0.0030	0.0062	0.0050	0.0051
Clade 5-b	0.096	0.074	0.077	0.060	0.058	0.013	–	0.0062	0.0053	0.0051
Clade 6	0.099	0.075	0.074	0.061	0.059	0.055	0.059	–	0.0054	0.0052
Clade 7	0.090	0.078	0.075	0.061	0.051	0.041	0.046	0.049	–	0.0031
Clade 8	0.089	0.071	0.072	0.056	0.048	0.040	0.043	0.046	0.023	–

Below the diagonal, mean among-clade *p*-distances based on the full-cyt*b* sequence database; above the diagonal, corresponding standard error values.

Table 3. Genealogical sorting index (*gsi*) scores and significance test results for GMYC-delimited species of the *Poecilia sphenops* species complex.

Delimited species	Concatenated mtDNA	nDNA loci						Concatenated mtDNA + nDNA
		Ldh-A	*RPS7*	X-*src*	X-*yes*	*Glyt*	Ensemble score (gsi_T)	
Clade 1, *P.* sp. "Tipitapa"	1**	0.039ns	0.488*	0.488*	–	–	0.203*	1*
Clade 2-b	1**	0.029ns	1**	1*	–	–	0.406*	1**
Clade 2-c, "*sphenops*" sp. 1	1**	0.083ns	1**	0.463*	0.472ns	1*	0.603**	1**
Clade 3, "*gillii*" sp. 2	1**	0.035ns	0.488*	1**	–	–	0.305*	1*
Clade 4, *P. hondurensis*	1**	0.024ns	1*	0.488*	–	–	0.302*	1*
Clade 5, *P. gillii*	1**	0.180ns	1**	0.551**	0.367*	0.250ns	0.470**	1**
Clade 6, *P. butleri*	1**	–	–	–	–	–	–	1*
Clade 7, "*limantouri*" clade	1**	–	–	–	–	–	–	1*
Clade 8, *P. mexicana* clade	1**	0.279*	0.688**	0.472**	0.336*	0.205ns	0.396**	1**

*$P < 0.05$; **$P < 0.001$; ns, not significant

Figure Legends

Figure 1. *Poecilia sphenops* species complex sampling localities and phylogeographical structuring throughout Central America. Sampling localities (dots) correspond to collections data in supplementary Data S1 and are colored according to phylogenetic clades or 'major-lineages' in Fig. 2 and the upper right legend. Some localities for clades 2-a and 6 are shown in the overview map (bottom left). The legend also lists clades corresponding to two monophyletic sub-complexes within the complex *sensu lato*, supported here (see Results) and in previous studies [55,64]. The locality for one sample whose phylogenetic position fluctuated during analyses (172554) is indicated on the map. Regional context is given by geopolitical boundaries (country names in red) and the continental divide (red line).

Figure 2. Results of bGMYC analysis for developing preliminary species delimitation hypotheses. Results presented are based on the concatenated mtDNA (cyt*b*, *cox1*, serine tRNA) dataset and represented on the gene tree resulting from maximum-likelihood (ML) analysis in GARLI. Nodal support values are ML bootstrap proportions (BP; ≥50%)/Bayesian posterior probabilities (PP; ≥0.95). Colored bars to the right of the phylogeny represent hypothesized species groupings based on ≥0.9 Bayesian posterior probability of conspecificity (calculated from Bayesian MCMC analysis of 100 post-burn-in trees from the concatenated mtDNA BEAST analysis), compared with bars demarcating clades meeting genetic distance thresholds (1–2%, 3%) and nominal taxonomy (NTAX).

Figure 3. Gene trees derived from maximum-likelihood analyses of the concatenated mtDNA + nDNA dataset (A) and the concatenated nDNA dataset (B) in GARLI. Numbers along branches indicate the level of nodal support from ML bootstrap

proportions (BP) ≥50%. Clade names at tips and colored bars representing delimited/nominal species correspond to those shown in Fig. 2.

Figure 4. Species tree inferred for the *P. sphenops* species complex showing speciation probabilities for each node. Bayesian speciation probabilities are posterior probabilities that a node is fully bifurcating and are shown for each node under each combination of priors in BP&P (top, large ancestral θ and deep root divergence, τ_0; middle, small ancestral θ and shallow τ_0; bottom, large ancestral θ and shallow τ_0). The red line distinguishes between species that were strongly supported (PP ≥0.95) using all three arbitrary prior combinations, and those with non-significant speciation probabilities. Results are presented for algorithm 1 runs and species that we could evaluate in BP&P. Fig. S4 shows speciation probabilities estimated using BP&P algorithm0.

Supporting Information

Data S1 Taxon list and locality (sub-population) details.

Appendix S1 Supplementary methods and results.

Table S1 Summary of the taxonomy, tooth morphology, and distributions of species the *Poecilia sphenops* species complex.

Table S2 DNA substitution models selected using DT-ModSel.

Table S3 Sequence attributes and DNA polymorphism levels in each of the datasets analyzed in this study, overall and by gene.

Fig. S1 BEAST MCC tree derived from the concatenated mtDNA dataset.

Fig. S2 Gene trees of nuclear DNA loci sampled for gene tree, species, tree, and species delimitation analyses. Each delimited species is coded the corresponding clade color from Fig. 2. Scale bars are in units of substitutions per site.

Fig. S3 Matching tree and posterior probability matrix from Bayesian general mixed Yule-coalescent bGMYC analyses. The phylogeny is the BEAST MCC tree (Fig. S1) and the tables at right of each tree provide sequence-by-sequence visualizations of the posterior probability that each species/sequence pair is conspecific. (A) Results from the single-threshold model. (B) Results from the multiple-threshold model.

Fig. S4 Species tree showing posterior probabilities of species for each node under each combination of priors using algorithm 0 in BP&P. The red line and nodal values correspond to the same criteria and priors described in Fig. 4.

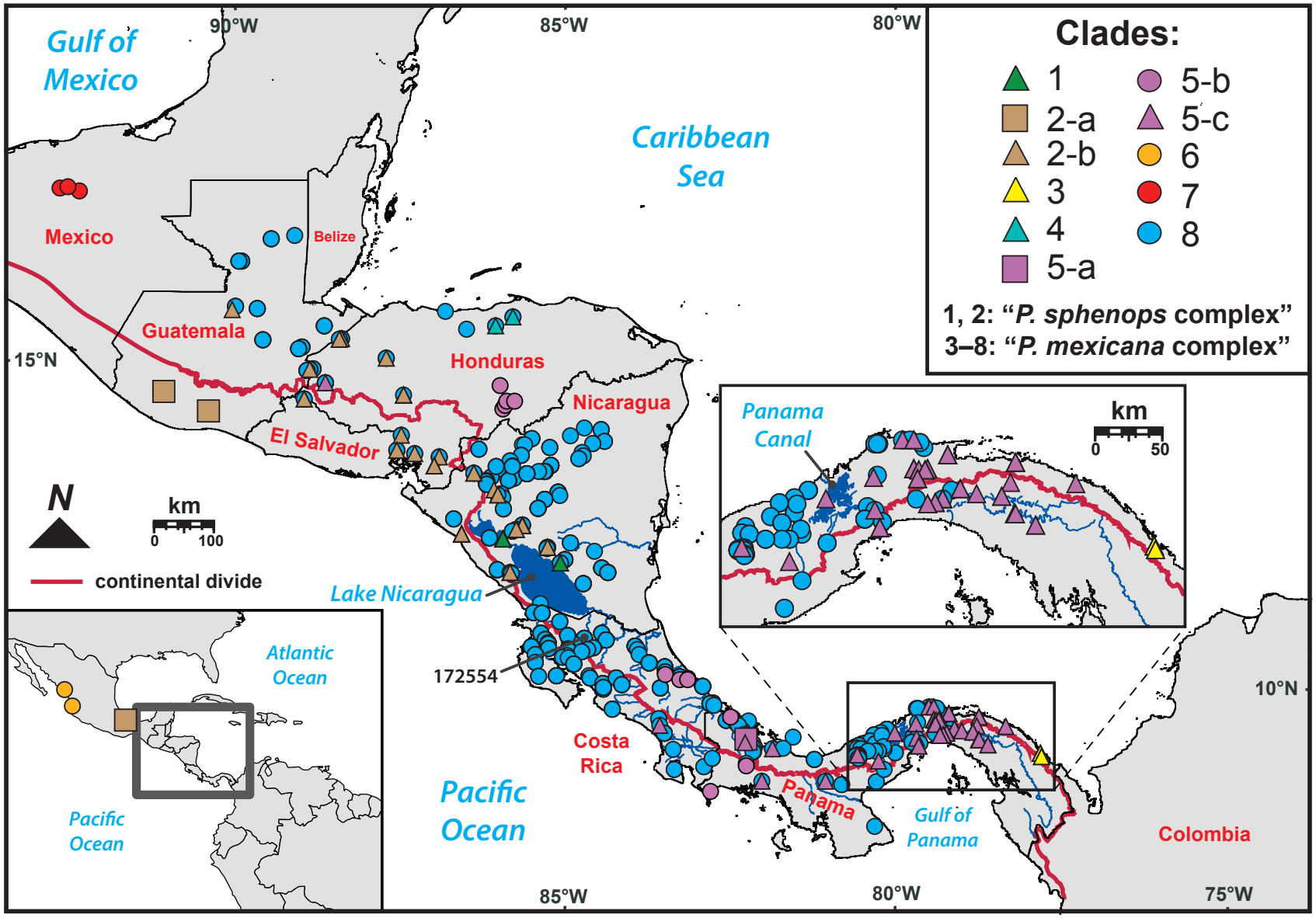
Gulf of Mexico
90°W
85°W
80°W
Caribbean Sea
Clades:
1
2-a
2-b
3
4
5-a
5-b
5-c
6
7
8
1, 2: "P. sphenops complex"
3–8: "P. mexicana complex"
Mexico
Belize
Guatemala
Honduras
Nicaragua
El Salvador
15°N
N
km
0 100
continental divide
Lake Nicaragua
172554
Panama Canal
km
0 50
Atlantic Ocean
Pacific Ocean
Pacific Ocean
Costa Rica
Panama
Gulf of Panama
Colombia
10°N
85°W
80°W
75°W

Figure 1

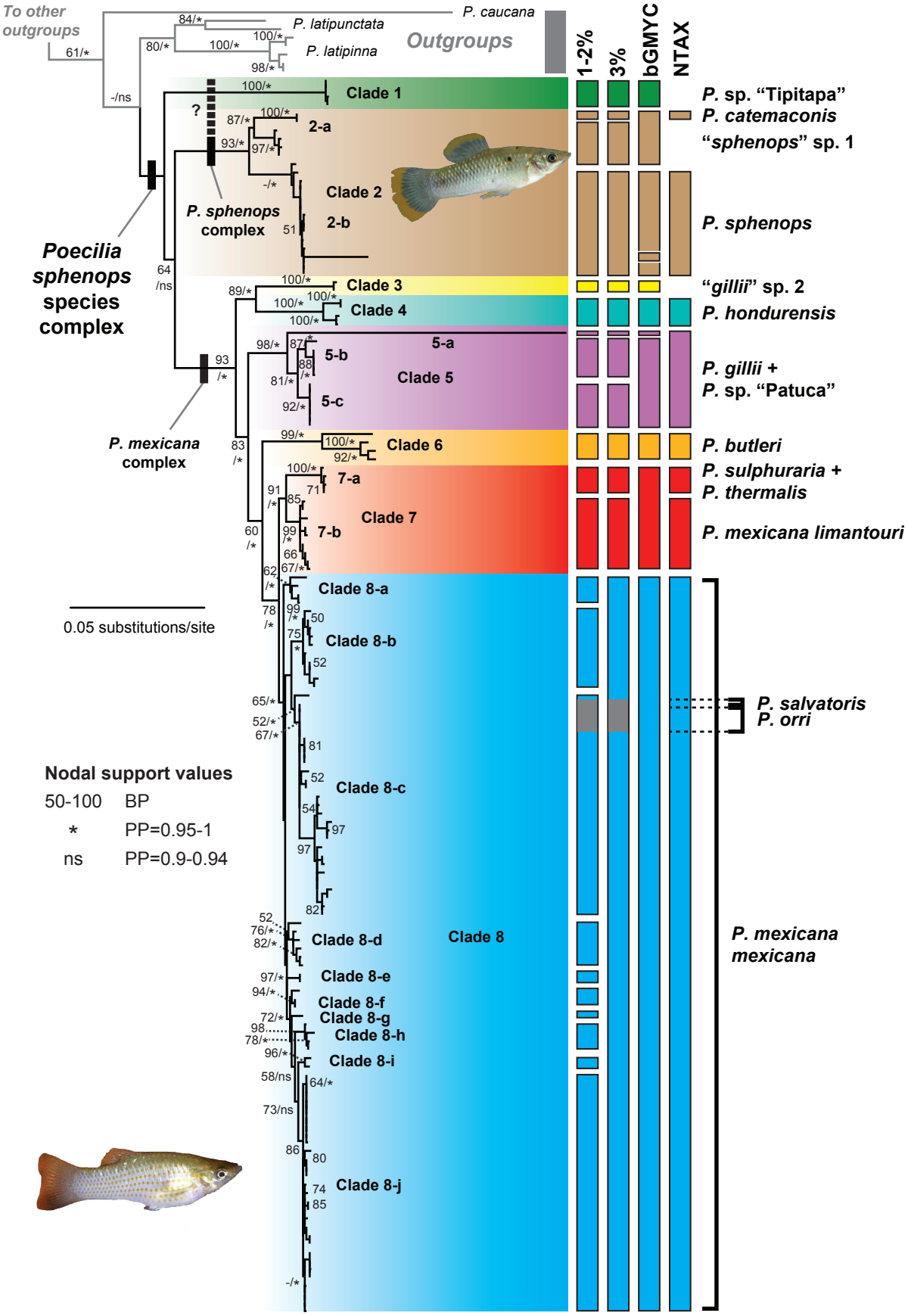
To other outgroups
P. caucana
P. latipunctata
P. latipinna
Outgroups
1-2%
3%
bGMYC
NTAX
Clade 1
2-a
Clade 2
2-b
P. sphenops complex
Poecilia sphenops species complex
Clade 3
Clade 4
5-a
5-b
5-c
Clade 5
P. mexicana complex
Clade 6
7-a
7-b
Clade 7
Clade 8-a
Clade 8-b
Clade 8-c
Clade 8-d
Clade 8-e
Clade 8-f
Clade 8-g
Clade 8-h
Clade 8-i
Clade 8-j
Clade 8
0.05 substitutions/site
Nodal support values
50-100 BP
* PP=0.95-1
ns PP=0.9-0.94
P. sp. "Tipitapa"
P. catemaconis
"sphenops" sp. 1
P. sphenops
"gillii" sp. 2
P. hondurensis
P. gillii +
P. sp. "Patuca"
P. butleri
P. sulphuraria +
P. thermalis
P. mexicana limantouri
P. salvatoris
P. orri
P. mexicana mexicana

Figure 2

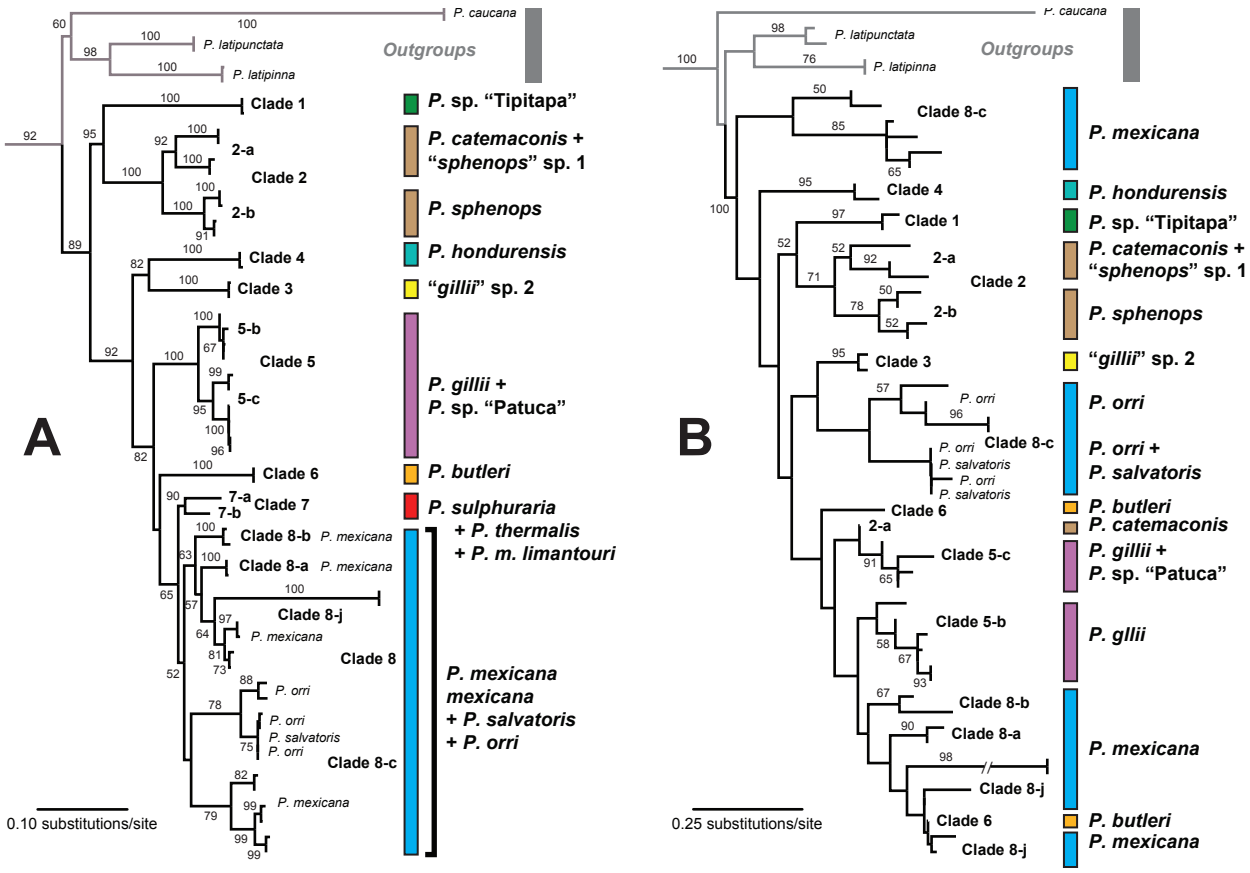
A
P. caucana
P. latipunctata
P. latipinna
Outgroups
Clade 1
2-a
Clade 2
2-b
Clade 4
Clade 3
5-b
Clade 5
5-c
Clade 6
7-a
7-b
Clade 7
Clade 8-b
P. mexicana
Clade 8-a
P. mexicana
Clade 8-j
P. mexicana
Clade 8
P. orri
P. orri
P. salvatoris
P. orri
Clade 8-c
P. mexicana
0.10 substitutions/site
P. sp. "Tipitapa"
P. catemaconis +
"sphenops" sp. 1
P. sphenops
P. hondurensis
"gillii" sp. 2
P. gillii +
P. sp. "Patuca"
P. butleri
P. sulphuraria
+ P. thermalis
+ P. m. limantouri
P. mexicana
mexicana
+ P. salvatoris
+ P. orri
B
P. caucana
P. latipunctata
P. latipinna
Outgroups
Clade 8-c
Clade 4
Clade 1
2-a
Clade 2
2-b
Clade 3
P. orri
Clade 8-c
P. orri
P. salvatoris
P. orri
P. salvatoris
Clade 6
2-a
Clade 5-c
Clade 5-b
Clade 8-b
Clade 8-a
Clade 8-j
Clade 6
Clade 8-j
0.25 substitutions/site
P. mexicana
P. hondurensis
P. sp. "Tipitapa"
P. catemaconis +
"sphenops" sp. 1
P. sphenops
"gillii" sp. 2
P. orri
P. orri +
P. salvatoris
P. butleri
P. catemaconis
P. gillii +
P. sp. "Patuca"
P. gllii
P. mexicana
P. butleri
P. mexicana

Figure 3

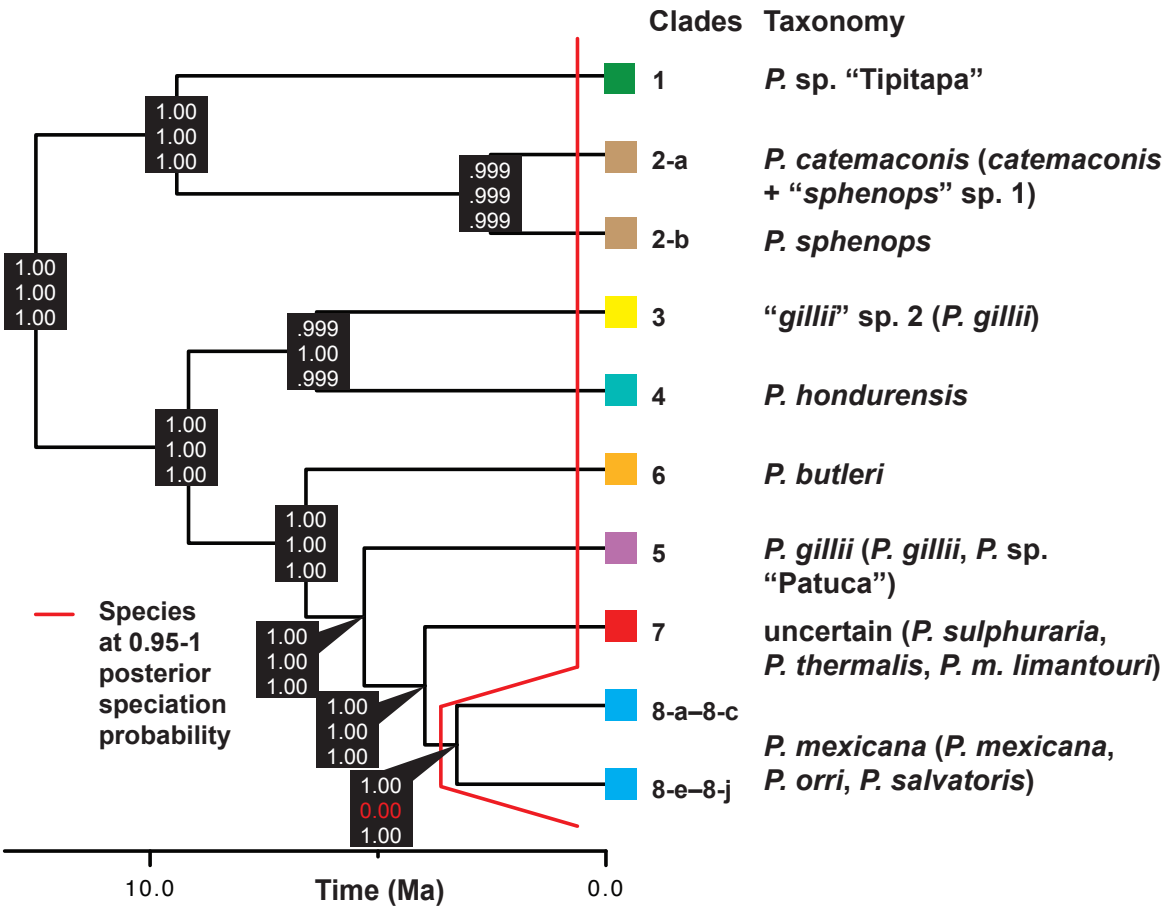
Clades
Taxonomy
1
P. sp. "Tipitapa"
2-a
P. catemaconis (catemaconis + "sphenops" sp. 1)
2-b
P. sphenops
3
"gillii" sp. 2 (P. gillii)
4
P. hondurensis
6
P. butleri
5
P. gillii (P. gillii, P. sp. "Patuca")
7
uncertain (P. sulphuraria, P. thermalis, P. m. limantouri)
8-a–8-c
8-e–8-j
P. mexicana (P. mexicana, P. orri, P. salvatoris)
Species at 0.95-1 posterior speciation probability
1.00
1.00
1.00
.999
.999
.999
1.00
1.00
1.00
.999
1.00
.999
1.00
1.00
1.00
1.00
1.00
1.00
1.00
1.00
1.00
1.00
0.00
1.00
10.0
Time (Ma)
0.0

Figure 4

Chapter 3 – Supplementary Material

Data S1 Taxon list and locality (sub-population) details.

Species	ID	Locality	Country	Latitude	Longitude	bGMYC-based taxonomy	Reference
***Belonesox belizanus* (OG)**	Bbeliz	–	–	N/A	N/A	–	Hrbek *et al.* (2007)
***Limia dominicensis* (OG)**	Ldom	–	–	N/A	N/A	–	Hrbek *et al.* (2007)
***Limia heterandria* (OG)**	Lheteran	–	–	N/A	N/A	–	This study
***Limia melanogaster* (OG)**	Lmela	–	–	N/A	N/A	–	Hrbek *et al.* (2007)
***Limia tridens* (OG)**	Ltrid	–	–	N/A	N/A	–	Hrbek *et al.* (2007)
***Limia vittata* (OG)**	LvitCU153, LvitC197	–	–	N/A	N/A	–	Doadrio *et al.* (2009)
***Micropoecilia picta* (OG)**	Ppict	–	Trinidad	N/A	N/A	–	Michael Tobler; this study
Poecilia* "*sphenops*" *sp.	8806	Rio Goascorán ~2.2 km southwest of Goascorán and 0.7 km south of El Amatillo	Honduras	13.58928	-87.76212	*P. sphenops*	This study
Poecilia butleri	Pbut	–	Mexico	N/A	N/A	*P. butleri*	Zuñiga-Vega *et al.* (2014)
P. butleri	Pbut1	Estero San Cristobal at San Blas, Nayarit	Mexico	21.54474	-105.27380	*P. butleri*	Tobler *et al.* (2011); Palacios *et al.* (2013)
P. butleri	Pbut4-0, Pbut1-0	Laguna de San Pedro Lagunillas at San Pedro Lagunillas, in Nayarit state	Mexico	21.21412	-104.73820	*P. butleri*	Palacios *et al.* (2013)
P. butleri	Pbut2	Rio Marabasco at Mexico Hwy 200, just east of Cihuatlan, Jalisco	Mexico	19.23647	-104.55400	*P. butleri*	Tobler *et al.* (2011); Palacios *et al.* (2013)
Poecilia catemaconis	Pcat-0, Pcat3	Lago Catemaco	Mexico	18.38327	-95.12275	*P. catemaconis* (*catemaconis* + "*sphenops*" sp. 1)	Palacios *et al.* (2013)
P. catemaconis	Pcatemac	Lago Catemaco	Mexico	N/A	N/A	*P. catemaconis*	This study
***Poecilia caucana* (OG)**	Pcauc	–	Trinidad	N/A	N/A	*P. caucana*	Michael Tobler; this study
Poecilia gillii	PG603.01, PG603.02	Lago Arenal	Costa Rica	10.47208	-84.76933	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG611.01–05	Quebrada Homiguera, trib. to Rio Tenorio	Costa Rica	10.69090	-85.08365	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG805.01–03	Quebrada La Canela	Costa Rica	9.851510	-84.52766	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG517.01, PG517.02	Rio Barrigones at Costa Rica Hwy 245, Osa Peninsula, Golfo Dulce	Costa Rica	8.593230	-83.42181	*P. mexicana*	Lee & Johnson (2009)
P. gillii	1205, 1206	Rio Cañas (trib. to Rio Tenorio) appxoimately 2 km southeast of town of Bebedero, ~10 km west of CA1	Costa Rica	10.34825	-85.16882	*P. mexicana*	This study
P. gillii	PG723.01	Rio Carrisal	Costa Rica	10.39501	-85.58688	*P. mexicana*	Lee & Johnson (2009)

P. gillii	PG806.01–05	Rio Centeno at Rd 131, approximately 1.6 km west of San Mateo	Costa Rica	9.941310	-84.53886	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG716.01–08	Rio Chimurria	Costa Rica	10.72740	-84.55823	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG612.01–05	Rio Chiquito	Costa Rica	10.43770	-84.86815	*P. mexicana*	Lee & Johnson (2009)
P. gillii	13308	Rio Ciruelas	Costa Rica	10.05913	-84.75919	*P. mexicana*	This study
P. gillii	PG808.01–06	Rio Congo at Pan American Hwy (CA 1), ~2.2 km east of Matapalo (drains to Golfo de Nicoya)	Costa Rica	10.23998	-84.99171	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG4814.01–02	Rio General	Costa Rica	9.38944	-83.66361	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PGB706.01–04, PG706.01–08	Rio Hatillo Viejo	Costa Rica	9.62311	-82.8552	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG703.01–05	Rio Herediana at Costa Rica Hwy 32, approximately 3 km northwest of Siquirres	Costa Rica	10.12416	-83.55616	*P. mexicana*	Lee & Johnson (2009)
P. gillii	2119	Rio Higueron	Costa Rica	10.34270	-85.07594	*P. mexicana*	This study
P. gillii	PG715.01–08	Rio Infiernito	Costa Rica	10.61801	-84.48418	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG724.01–15	Rio Irigary at Pan American Hwy (CA 1), at Irigary	Costa Rica	10.72340	-85.51038	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG636.01	Rio Isla Grande	Costa Rica	10.39300	-83.9682	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG617.01–03	Rio Javilla, trib. to Rio Cañas	Costa Rica	10.37208	-85.0974	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG602.01–05	Rio La Palma	Costa Rica	10.49875	-84.689	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG616.01–03	Rio Magdalena	Costa Rica	10.47945	-85.07811	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG722.01–16	Rio Marole at Costa Rica Hwy 21, Nicoya Peninsula, trib. to Golfo de Nicoya	Costa Rica	10.05828	-85.26201	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG807.01–05	Rio Naranjo, trib. to Rio Barranca, at CA 1, ~4.2 km northwest of Barranca	Costa Rica	10.02263	-84.73441	*P. mexicana*	Lee & Johnson (2009)
P. gillii	1231, 1232	Rio Nosara, approximately 5 km west of Costa Rica Rd 150, Nicoya Peninsula	Costa Rica	10.04833	-85.54520	*P. mexicana*	This study
P. gillii	PGB714.01, PGB714.02	Rio Nuevo	Costa Rica	8.64102	-82.95297	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG801.01–05	Rio Pacacua	Costa Rica	9.91960	-84.24130	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG710.01–07	Rio Parismina	Costa Rica	10.19771	-83.56873	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG4810.01–15	Rio Pejibaye	Costa Rica	9.15694	-83.57527	*P. mexicana*	Lee & Johnson (2009)
P. gillii	2170	Rio Pizote	Costa Rica	10.90839	-85.21126	*P. mexicana*	This study
P. gillii	PG719.01–16	Rio Queques, trib. to Rio Frio, which drains to Lago Nicaragua	Costa Rica	10.64481	-84.82223	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG809.01–05	Rio Rosales	Costa Rica	10.02978	-84.32581	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG608.01–05	Rio Sabalito, trib. to Lago Arenal (west side),1.5	Costa Rica	10.54858	-84.98080	*P. mexicana*	Lee & Johnson

		km southeast of Sabalito on Costa Rica Rd 142					(2009)
P. gillii	PG726.01–08	Rio Sabalo	Costa Rica	11.04283	-85.48921	*P. mexicana*	Lee & Johnson (2009)
P. gillii	2051	Rio Salama Nuevo, 1 km south of Carretera Interamericana (Costa Rica Hwy 2) and 1 km east of Finca Doce	Costa Rica	8.90425	-83.43932	*P. mexicana*	This study
P. gillii	2171	Rio San Juan	Costa Rica	10.90839	-85.21126	*P. mexicana*	This study
P. gillii	PG614.01, PG614.02	Rio Santa Rosa	Costa Rica	10.46113	-85.07438	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG725.01–08	Rio Sapoa just north of Costa Rica Hwy 4, approximately 2 km southeast of La Cruz	Costa Rica	11.04436	-85.61590	*P. mexicana*	Lee & Johnson (2009)
P. gillii	1245, 1246	Rio Sarapiquí	Costa Rica	10.47225	-83.99195	*P. mexicana*	This study
P. gillii	PG713.01–13	Rio Sarapiquí	Costa Rica	10.52455	-84.03133	*P. mexicana*	Lee & Johnson (2009)
P. gillii	1291	Rio Sixaola, ~6 km southeast of Bribri, 1 km south of Costa Rica Hwy 36	Costa Rica	9.59872	-82.80247	*P. mexicana*	This study
P. gillii	PG804.01–03	Rio Grande de Tárcoles at Costa Rica Rd 137, west of San Juan de Mata (El Llano)	Costa Rica	9.87980	-84.52780	*P. mexicana*	Lee & Johnson (2009)
P. gillii	2073	Rio Terraba	Costa Rica	9.28493	-83.64566	*P. gillii*	This study
P. gillii	PG708.01–08	Rio Toro, trib. to Rio Matina, just off Costa Rica Hwy 32 approximately 24 km west of Limon	Costa Rica	10.01678	-83.21021	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG708.03	Rio Toro, trib. to Rio Matina, just off Costa Rica Hwy 32 approximately 24 km west of Limon	Costa Rica	10.01678	-83.21021	*P. gillii* (*P.* sp. "Patuca")	Lee & Johnson (2009)
P. gillii	PG712.01–13, PG712.14–16	Rio Tortuguero	Costa Rica	10.25941	-83.81223	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG610.01–03	Small ditch at Hwy 927, trib. to Rio Corobici	Costa Rica	10.62406	-85.05811	*P. mexicana*	Lee & Johnson (2009)
P. gillii	2074	Trib. (Rio Peje?) to Rio General just east of Carretera Interamericana (Costa Rica Hwy 2)	Costa Rica	9.28493	-83.64566	*P. mexicana*	This study
P. gillii	PGB701.01–04, PG701.01–05	Trib. to Rio Reventazon at Rd 232 across from Angostura Lagoon	Costa Rica	9.87230	-83.63320	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PG707.01–08	Unnamed lagoon, trib. to Rio Bananito, ~11.2 km southeast of Limon on Costa Rica Hwy 32	Costa Rica	9.89258	-82.97228	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PGB702.01–04	Unnamed trib at Costa Rica Hwy 239, in between Las Lomas and Finca La Palma, ~0.25 km northwest of Finca La Palma	Costa Rica	9.53697	-84.38588	*P. mexicana*	Lee & Johnson (2009)
P. gillii	PGB704.01–04	Unnamed trib. to Rio Sixaola at Costa Rica Hwy 36, ~3.2 km northeast of Bribri	Costa Rica	9.63203	-82.81921	*P. mexicana*	This study
P. gillii	PG704.01–08	Unnamed trib. to Rio Sixaola, ~3 km northeast of Bribri, on Costa Rica Hwy 36	Costa Rica	9.63203	-82.81921	*P. mexicana*	Lee & Johnson (2009)
P. gillii	8343–44, 8355–58	Rio Ulúa at Cucuyagua	Honduras	14.65096	-88.88144	*P. mexicana*	This study
P. gillii	8350	Rio Ulúa at Cucuyagua	Honduras	14.65096	-88.88144	*P. gillii*	This study
P. gillii	8364, 8365	Rio Tio Higuito, trib. to Rio Motagua, at cuenca near Higuito	Honduras	14.83940	-89.16819	*P. mexicana*	This study
P. gillii	172455, 172459–62	Lago Jiloa	Nicaragua	12.21858	-86.31194	*P. mexicana*	This study
P. gillii	173003–08	Rio Caracol at Nicaragua Hwy 7, trib. to Rio Malacatoya	Nicaragua	12.35116	-85.88870	*P. sphenops*	This study

P. gillii	174214–16	Rio Estelí in Estelí	Nicaragua	13.10663	-86.35710	*P. mexicana*	This study
P. gillii	172439–46	Rio La Conquista	Nicaragua	11.72472	-86.18469	*P. mexicana*	This study
P. gillii	172867–68	Rio Malacatoya	Nicaragua	12.32661	-85.95552	*P. mexicana*	This study
P. gillii	173271–74	Rio Mayales	Nicaragua	12.05663	-85.40814	*P. mexicana*	This study
P. gillii	173336, 173337, 173339	Rio Mayales main stem	Nicaragua	12.06679	-85.40375	*P. mexicana*	This study
P. gillii	173340	Rio Mayales main stem	Nicaragua	12.06679	-85.40375	*P. sphenops*	This study
P. gillii	173341	Rio Mayales main stem	Nicaragua	12.06679	-85.40375	*P. mexicana*	This study
P. gillii	173343	Rio Mayales main stem	Nicaragua	12.06679	-85.40375	*P. sphenops*	This study
P. gillii	172539–42	Rio Ochomogo	Nicaragua	11.65663	-85.97319	*P. mexicana*	This study
P. gillii	174259–62	Rio Telica at Telica, just west (downstream) of Nicaragua Hwy 12, ~12 km north of Leon	Nicaragua	12.51656	-86.86542	*P. mexicana*	This study
P. gillii	172695, 172696, 172698–702	Rio Tipitapa	Nicaragua	12.20267	-86.10208	*P.* sp. "Tipitapa"	This study
P. gillii	172869	Rio Tipitapa/Laguna de Tisma	Nicaragua	12.32661	-85.95552	*P. mexicana*	This study
P. gillii	13417–20	Rio Viejo (afluente Lago de Managua)	Nicaragua	12.90702	-86.1283	*P. mexicana*	This study
P. gillii	167949	Trib. to Río Grande at La Trinidad	Nicaragua	12.97132	-86.2372	*P. sphenops*	This study
P. gillii	174327–31	Trib. to Rio Grande Viejo (Rio Viejoio) at Nicaragua Hwy 26 ~4.5 km northwest of El Jocote (between Estelí and León)	Nicaragua	12.89324	-86.17908	*P. mexicana*	This study
P. gillii	173020–22, 173024, 173026–27	Trib. To Rio Malacatoya at Teustepe ~27 km southwest of Boaco, Departamento Boaco	Nicaragua	12.41159	-85.79172	*P. sphenops*	This study
P. gillii	173273	Trib. to Rio Mayales about 6 km southwest of Juigalpa on Nicaragua Hwy 37	Nicaragua	12.05663	-85.40814	*P. mexicana*	This study
P. gillii	172553, 172554, 172558, 172559	Unnamed drainage ditch trib. 1 km north of Ochomogo	Nicaragua	11.67886	-85.98816	*P. sphenops*	This study
P. gillii	173459, 173460, 173464, 173465, 173468–70,	Unnamed trib. to Rio Rama south of Nueva Guinea	Nicaragua	11.67838	-84.45622	*P. mexicana*	This study
P. gillii	173680–85	Unnamed trib. to Lago Nicaragua	Nicaragua	11.74923	-84.55819	*P. mexicana*	This study
P. gillii	167955–58	Unnamed trib. to Rio Estelí, thus a trib. to Río Coco	Nicaragua	13.05866	-86.35114	*P. mexicana*	This study
P. gillii	168900–02	Unnamed trib. to Río Malacatoya at Teustepe, just off rd to Rama (Nicaragua Hwy 7), southwest of Boaco	Nicaragua	12.41825	-85.79299	*P. sphenops*	This study
P. gillii	173959, 173960	Unnamed trib. to Rio Ojucuapa at San Roque (Comarca Santa Marta) off unnamed road, ~2 km southwest of Hwy 25	Nicaragua	11.82942	-85.20479	*P. mexicana*	This study
P. gillii	173961, 173963	Unnamed trib. to Rio Ojucuapa at San Roque (Comarca Santa Marta) off unnamed road, ~2 km southwest of Hwy 25	Nicaragua	11.82942	-85.20479	*P.* sp. "Tipitapa"	This study
P. gillii	173964	Unnamed trib. to Rio Ojucuapa at San Roque (Comarca Santa Marta) off unnamed road, ~2 km southwest of Hwy 25	Nicaragua	11.82942	-85.20479	*P. mexicana*	This study
P. gillii	173966	Unnamed trib. to Rio Ojucuapa at San Roque (Comarca Santa Marta) off unnamed road, ~2 km	Nicaragua	11.82942	-85.20479	*P.* sp. "Tipitapa"	This study

		southwest of Hwy 25					
P. gillii	174109–14, 174116	Unnamed trib. to Rio Ojucuapa at San Roque (Comarca Santa Marta) off unnamed road, ~2 km southwest of Hwy 25	Nicaragua	11.87992	-85.13156	*P. mexicana*	This study
P. gillii	174511–15	Unnamed trib. to Rio Tamarindo, ~5 km northwest of El Tamarindo (around 20 km southeast of León)	Nicaragua	12.28114	-86.75105	*P. sphenops*	This study
P. gillii	16977	Two quebradas before Big Creek, Isla Colón, Bocas del Toro	Panama	9.35711	-82.25322	*P. mexicana*	This study
P. gillii	16423	Between Piriati and Quebrada Cali on the IAH to Darien	Panama	9.05266	-78.64700	*P. gillii*	This study
P. gillii	16197	Creek trib. to Rio Guarúmo near Punta Peña, Río Punta Agua Real	Panama	8.94608	-82.15711	*P. mexicana*	This study
P. gillii	6826	South side of Escudo de Veraguas, Bocas del Toro	Panama	9.10222	-81.56166	*P. mexicana*	This study
P. gillii	16419	First bridge east of Palenque, Colon	Panama	9.57052	-79.35833	*P. gillii*	This study
P. gillii	18620	Lago Gatún	Panama	9.15688	-79.96405	*P. gillii*	This study
P. gillii	16421	Nombre de Dios	Panama	9.57066	-79.43641	*P. gillii*	This study
P. gillii	6803	Old Gamboa Rd Creek	Panama	9.11000	-79.68000	*P. mexicana*	This study
P. gillii	12428	Quebrada by Almirante to Changuinola Rd	Panama	9.31469	-82.45036	*P. mexicana*	This study
P. gillii	18683	Quebrada Chiriquisito	Panama	8.68802	-82.29172	*P. mexicana*	This study
P. gillii	16149	Quebrada Congal	Panama	8.91411	-80.13405	*P. mexicana*	This study
P. gillii	11120, 11162	Quebrada El Nance, trib. to Rio Santa Maria	Panama	8.41316	-81.04850	*P. gillii*	This study
P. gillii	17114	Quebrada El Nance, trib. to Rio Santa Maria	Panama	8.41322	-82.04800	*P. gillii*	This study
P. gillii	17120	Quebrada El Nance, trib. to Rio Santa Maria	Panama	8.41322	-82.04800	*P. mexicana*	This study
P. gillii	16781–86	Quebrada en Mateo, at Pueblo Nuevo near Palmas Bellas	Panama	9.22580	-80.08588	*P. mexicana*	This study
P. gillii	18812, 18801	Quebrada Garay at Rambála-Almirante Rd, near Isla Pastores, Bocas del Toro	Panama	9.19575	-82.34311	*P. mexicana*	This study
P. gillii	16843	Quebrada Jobito	Panama	9.06355	-80.18802	*P. mexicana*	This study
P. gillii	15423, 15424	Quebrada Jobito	Panama	9.06394	-80.18597	*P. mexicana*	This study
P. gillii	10048	Quebrada La Candelaria, Rio Jobo, Rio Indio de Anton	Panama	9.13011	-80.17155	*P. mexicana*	This study
P. gillii	11517	Quebrada La Fe on new Punta Peña-Almirante Rd	Panama	9.06769	-82.29166	*P. gillii*	This study
P. gillii	16123, 16131	Quebrada Los Uveros	Panama	8.94719	-80.13825	*P. mexicana*	This study
P. gillii	7417, 7421, 7363, 7364	Quebrada Mandingo, Chagres	Panama	9.02530	-79.69890	*P. mexicana*	This study
P. gillii	7447, 7448	Quebrada Mandingo, Rio Velasque	Panama	8.96260	-79.59010	*P. gillii*	This study
P. gillii	18782, 18787	Quebrada Nigua	Panama	9.27894	-82.41525	*P. mexicana*	This study
P. gillii	12442	Quebrada on Almirante to Changuinola Rd	Panama	9.39686	-82.50058	*P. mexicana*	This study
P. gillii	12473	Quebrada at km 26 on Punta Peña Almirante Rd	Panama	9.06627	-82.29911	*P. mexicana*	This study
P. gillii	12481	Quebrada at km 34 on Punta Peña-Almirante Rd	Panama	9.11697	-82.29019	*P. mexicana*	This study
P. gillii	15532	Quebrada Platanal	Panama	8.87988	-80.27688	*P. mexicana*	This study
P. gillii	9030	Quebrada San Juan at Cuango	Panama	9.55080	-79.30920	*P. mexicana*	This study
P. gillii	15271	Quebrada Tolu	Panama	9.04113	-80.35497	*P. mexicana*	This study

P. gillii	15655	Quebrada Tortuguita	Panama	8.88127	-80.39088	*P. mexicana*	This study
P. gillii	18791, 18792	Quebrada Traicionera	Panama	9.13891	-82.30694	*P. mexicana*	This study
P. gillii	3871, 1736–38	Rio Acla	Panama	8.84306	-77.68361	*P.* "*gillii*" sp. 2	This study
P. gillii	4161, 4163	Rio Acla-Quebrada 2	Panama	8.79578	-77.67334	*P.* "*gillii*" sp. 2	This study
P. gillii	12290	Rio Aguas Claras	Panama	9.27380	-78.68133	*P. gillii*	This study
P. gillii	1118, 1119	Rio Anton	Panama	8.39680	-80.25851	*P. mexicana*	This study
P. gillii	18589	Rio Anton in Anton Valley	Panama	8.59719	-80.13775	*P. mexicana*	This study
P. gillii	3706, 3709, 3710	Rio Azucar	Panama	9.41167	-78.64583	*P. gillii*	This study
P. gillii	937, 938	Rio Bayano-B	Panama	9.17700	-78.74566	*P. gillii*	This study
P. gillii	4979, 4980, 4992, 4993	Rio Bongie	Panama	9.35990	-82.61000	*P. mexicana*	This study
P. gillii	16497, 16498	Rio Botija	Panama	8.81200	-80.57972	*P. mexicana*	This study
P. gillii	4796, 4799	Rio Caimito, Quebrada Mano de Piedra	Panama	8.85083	-79.96056	*P. mexicana*	This study
P. gillii	6885, 6886, 6888, 6890,	Rio Calovebora	Panama	8.74777	-81.22310	*P. mexicana*	This study
P. gillii	736	Rio Canaveral just south of Laguna Samani	Panama	8.92858	-81.71180	*P. mexicana*	This study
P. gillii	11625	Rio Canazas at Chiriqui Grande Rd, trib to Rio Guarumo	Panama	8.87333	-82.17444	*P. mexicana*	This study
P. gillii	18613	Rio Canita	Panama	9.20000	-78.91670	*P. gillii*	This study
P. gillii	16809, 16810, 16820, 16823	Rio Caño Rey	Panama	9.13358	-80.29422	*P. mexicana*	This study
P. gillii	15863, 15864	Rio Cardenas	Panama	9.00116	-79.57277	*P. mexicana*	This study
P. gillii	4533, 4334, 4536	Rio Cascajal ~2.5 km east of Portobelo	Panama	9.54722	-79.63040	*P. mexicana*	This study
P. gillii	12609, 12610, 12614	Rio Cascajal ~2.5 km east of Portobelo	Panama	9.54722	-79.60400	*P. mexicana*	This study
P. gillii	2956	Rio Cascajal ~5 km east of Portobelo	Panama	9.54642	-79.60625	*P. mexicana*	This study
P. gillii	16216, 16223, 16244	Rio Chagres	Panama	9.35963	-79.27877	*P. gillii*	This study
P. gillii	15184, 15185	Rio Chagres	Panama	9.36008	-79.32225	*P. gillii*	This study
P. gillii	16256, 16260	Rio Chagres	Panama	9.36727	-79.26055	*P. gillii*	This study
P. gillii		Rio Chagres	Panama	9.36727	-79.26055	*P. gillii*	This study
P. gillii	2887, 16935	Rio Chichebre	Panama	9.15888	-79.15638	*P. gillii*	This study
P. gillii	11373	Rio Chichebre	Panama	9.16083	-79.15416	*P. gillii*	This study
P. gillii	11374	Rio Chichebre	Panama	9.16083	-79.15416	*P. mexicana*	This study
P. gillii	18684	Rio Chiriqui	Panama	8.68802	-82.29172	*P.* sp. "Patuca"	This study
P. gillii	111, 112	Rio Chiriqui Viejo	Panama	8.76443	-82.82712	*P. mexicana*	This study
P. gillii	1347	Rio Cocle del Norte	Panama	8.81867	-80.55302	*P. mexicana*	This study
P. gillii	1367, 1368	Rio Cocle del Norte	Panama	8.81867	-80.55302	*P. gillii*	This study
P. gillii	16529	Rio Cocle del Norte	Panama	8.82147	-80.53355	*P. mexicana*	This study
P. gillii	12375, 12376	Rio Cricamola	Panama	8.91727	-81.87725	*P. mexicana*	This study
P. gillii	12377	Rio Cricamola	Panama	8.91727	-81.87725	*P. gillii*	This study
P. gillii	9395	Rio Cuango	Panama	9.51820	-79.28480	*P. mexicana*	This study

P. gillii	9020	Rio Cuango, ~2 km southwest of Playa Chiquita	Panama	9.55080	-79.30920	*P. mexicana*	This study
P. gillii	2875, 2879	Rio Garrapata	Panama	9.22670	-79.03027	*P. gillii*	This study
P. gillii	18627	Rio Gatuncillo	Panama	9.30816	-79.63335	*P. gillii*	This study
P. gillii	642	Rio Guarúmo	Panama	8.87250	-82.18933	*P. mexicana*	This study
P. gillii	12521	Rio Guarúmo	Panama	9.00000	-82.18333	*P. mexicana*	This study
P. gillii	15445, 15446	Rio Guasimo	Panama	8.99125	-80.27441	*P. mexicana*	This study
P. gillii	16043, 16045	Rio Guasimo	Panama	8.99133	-80.27433	*P. mexicana*	This study
P. gillii	3610, 3615	Rio Ipeti	Panama	8.97944	-78.50556	*P. gillii*	This study
P. gillii	16615, 16616	Rio La Jacinta	Panama	8.96863	-80.52950	*P. mexicana*	This study
P. gillii	11203	Rio Mamoni	Panama	9.22361	-79.09222	*P. mexicana*	This study
P. gillii	1320–22	Rio Mandinga at Golfo de San Blas	Panama	9.46995	-79.12415	*P. gillii*	This study
P. gillii	16722, 16723	Rio Membrillar	Panama	9.17388	-80.18500	*P. mexicana*	This study
P. gillii	18720	Rio Mensabé, Azuero Peninsula	Panama	7.70497	-80.27813	*P. mexicana*	This study
P. gillii	15281	Rio Miguel de la Borda	Panama	9.04113	-80.35497	*P. mexicana*	This study
P. gillii	16479, 16483	Rio Moreno	Panama	8.76666	-80.53613	*P. mexicana*	This study
P. gillii	15319	Rio Moreno	Panama	8.77941	-80.53447	*P. mexicana*	This study
P. gillii	9780, 9788, 9791	Rio Moreno on road to Coclesito	Panama	8.77450	-80.52783	*P. mexicana*	This study
P. gillii	2703, 2704	Rio Pacora, ~3 km west of Paso Blanco off Av. Jose Agustin Arango (~1.6 km above highway)	Panama	9.12306	-79.26250	*P. gillii*	This study
P. gillii	11337	Rio Parti	Panama	9.05600	-78.65950	*P. gillii*	This study
P. gillii	15849	Rio Pedro Miguel	Panama	9.08066	-79.62508	*P. mexicana*	This study
P. gillii	15854	Rio Pedro Miguel just off Av. Madden	Panama	9.08066	-79.62508	*P. gillii*	This study
P. gillii	16272, 16300	Rio Piedras	Panama	9.30155	-79.33152	*P. gillii*	This study
P. gillii	16438–40	Rio Piedras	Panama	9.37422	-79.35866	*P. gillii*	This study
P. gillii	18632	Rio Playa Alta, ~3 km east of Nombre de Dios	Panama	9.57069	-79.43613	*P. mexicana*	This study
P. gillii	6846, 6847	Rio Robalo	Panama	9.04055	-82.28583	*P. mexicana*	This study
P. gillii	23082	Rio San Bartolo	Panama	8.28333	-82.85000	*P.* sp. "Patuca"	This study
P. gillii	9768, 9773, 9774, 9776	Rio San Juan at Coclecito	Panama	8.80416	-80.58083	*P. mexicana*	This study
P. gillii	12345	Rio San Pedro at San Pedro	Panama	8.72680	-80.21938	*P. gillii*	This study
P. gillii	16957, 16958	Rio San San (Quebrada just east of Rio Negro) 1 km west of Finca 6	Panama	9.47563	-82.53805	*P.* sp. "Patuca"	This study
P. gillii	1292	Rio Sixaola	Panama	9.59872	-82.80247	*P. mexicana*	This study
P. gillii	4969	Rio Teribe at El Silencio, confluence of Teribe and Changuinola Rivers	Panama	9.37000	-82.54000	*P. mexicana*	This study
P. gillii	15365	Rio Toabré	Panama	8.91544	-80.50058	*P. mexicana*	This study
P. gillii	16679, 16686	Rio Toabre at Quebrada Patatilla	Panama	8.91533	-80.50066	*P. mexicana*	This study
P. gillii	16695, 16718	Rio Toabre at Quebrada Tortuguita	Panama	8.87861	-80.39047	*P. mexicana*	This study
P. gillii	4518, 4519	Rio Tranca	Panama	9.13805	-79.20916	*P. gillii*	This study
P. gillii	16094, 16096	Rio Uracillo	Panama	8.88086	-80.21983	*P. mexicana*	This study
P. gillii	2851	Rio Utive	Panama	9.16417	-79.34000	*P. mexicana*	This study

P. gillii	15409	Rio Victoria	Panama	8.92499	-80.55138	*P. mexicana*	This study
P. gillii	16562, 16578	Rio Victoria	Panama	8.92513	-80.55172	*P. mexicana*	This study
P. gillii	17945	Rio Zahino, trib. to Rio Viento Frio, ~2.5 km west of Palenque	Panama	9.57055	-79.38250	*P. mexicana*	This study
P. gillii	2589	Side Lagoon 1hr upstream of Rio Playon Chico	Panama	9.26611	-78.22556	*P. gillii*	This study
P. gillii	12459	Small creek at km 41 on Punta Peña Almirante Rd	Panama	9.14752	-82.31766	*P. mexicana*	This study
P. gillii	11540	Small creek on Punta Peña-Almirante Rd	Panama	9.01666	-82.30933	*P. gillii*	This study
P. gillii	16414	Stream between Sardinilla and Salamanca	Panama	9.32644	-79.61188	*P. mexicana*	This study
P. gillii	6380	Stream trib. to Rio Guarúmo near Punta Peña	Panama	8.92853	-82.18027	*P. mexicana*	This study
P. gillii	16199	Trib. to Rio Guarúmo near Punta Peña	Panama	8.87594	-82.17461	*P. mexicana*	This study
P. gillii	15330, 15331	Unnamed trib.	Panama	8.80463	-80.53327	*P. mexicana*	This study
P. gillii	16536	Unnamed trib.	Panama	8.82147	-80.53355	*P. mexicana*	This study
P. gillii	6379	Unnamed trib. to Rio Guarúmo just off Hwy 10 1.2 km north of Punta Peña (draining to Chiriqui Lagoon, Bocas del Toro)	Panama	8.92853	-82.18027	*P. mexicana*	This study
P. gillii	3141, 3148	Unnamed trib. to Rio Santa Maria, Azuero Peninsula	Panama	8.35278	-80.79923	*P. mexicana*	This study
Poecilia hondurensis	8520	Quebrada de Chicho between communities of Achiote and Cholomena	Honduras	15.53480	-86.21170	*P. hondurensis*	This study
P. hondurensis	8568	Rio Taujica at Taujica	Honduras	15.68100	-85.93930	*P. hondurensis*	This study
***Poecilia hondurensis* - *P.* "*spheops*" sp. in Alda *et al.* (2013)**	8479	Rio Camalote 1 km north of Macuelizo near Guatemala-Honduras border	Honduras	15.32656	-88.66264	*P. sphenops*	This study
***P. hondurensis* - *P.* "*spheops*" sp. in Alda *et al.* (2013)**	8859	Rio Goascorán at Caridad	Honduras	13.8277	-87.69480	*P. sphenops*	This study
***P. hondurensis* - *P.* "*spheops*" sp. in Alda *et al.* (2013)**	8409	Rio Naco at Ulúa	Honduras	15.34147	-88.62480	*P. sphenops*	This study
***Poecilia latipinna* (OG)**	Plat4, Plat5, Plati-0	–	–	N/A	N/A	*P. latipinna*	Palacios *et al.* (2013)
***P. latipinna* (OG)**	PlatJBLee	–	Costa Rica	N/A	N/A	–	Lee & Johnson (2009)
***P. latipinna* (OG)**	8112	–	Mexico	N/A	N/A	*P. latipinna*	This study
***P. latipinna* (OG)**	Platip	Ciudad Mante, Rio Panuco Mexico	Mexico	N/A	N/A	–	Hrbek *et al.* (2007)
***Poecilia latipunctata* (OG)**	Platipun	–	Mexico	N/A	N/A	*P. latipunctata*	Michael Tobler; this study
Poecilia mexicana	8722	Quebrada Carrizal, trib. to Rio Patuca, at Terrero Blanco	Honduras	14.41290	-86.04006	*P.* sp. "Patuca"	This study
P. mexicana	8770–72	Quebrada San Jose, trib. to Rio Guayape, at San Francisco de Becerra	Honduras	14.62763	-86.14237	*P.* sp. "Patuca"	This study
P. mexicana	8676–78	Rio Guayambre, trib. to Rio Patuca, La Cieniguita	Honduras	14.28718	-86.10293	*P.* sp. "Patuca"	This study
P. mexicana	8607	Rio Medina, trib. to Rio Aguan, at Coyoles Centrales	Honduras	15.48380	-86.66600	*P. mexicana*	This study
P. mexicana	8747	Rio Tepemechin, trib. to Rio Patuca, at Tepemechin	Honduras	14.40185	-85.93010	*P.* sp. "Patuca"	This study

P. mexicana	8618–20	Rio Yojoa at Ulúa	Honduras	15.03480	-87.92870	*P. mexicana*	This study
P. mexicana	Pmmex34, PmmxCDA	Cueva del Azufre, Tabasco state	Mexico	17.43843	-92.77476	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex31, PmmxPysp	Puyacatengo Springs, in Tabasco state	Mexico	17.45800	-92.88900	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	PmmxIxt2-0, PmmxIxt3, PmmxIxta	Rio Ixtapangajoya, Chiapas state	Mexico	17.49500	-92.99800	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex29	Rio Nututun, Palenque	Mexico	17.48416	-91.97376	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex28	Rio Oxolotan, Tapijulapa	Mexico	17.46443	-92.77430	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex25	Rio Pichaculco, Arroyo Rosita	Mexico	17.48500	-93.10400	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex14	Rio Pichucalco, Baños del Azufre	Mexico	17.55200	-92.99900	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex15, Pmmex16	Rio Pichucalco, Baños del Azufre	Mexico	17.55225	-92.99859	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex33	Rio Puyacatengo, Baños del Azufre	Mexico	17.55225	-92.99859	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex21–23	Rio Puyacatengo, La Lluvia	Mexico	17.46400	-92.89500	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex19, Pmmex20	Rio Puyacatengo, La Lluvia, Puyacatengo Springs	Mexico	17.45761	-92.88892	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex24	Rio Puyacatengo, Rio Pichucalco, La Joya, Santa Ana	Mexico	N/A	N/A	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex17, Pmmex18, Pmmex32	Rio Puyacatengo, Vicente Gurrero Lerma	Mexico	17.51008	-92.91448	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex26, Pmmex27	Rio Tacotalpa, Arroyo Bonita	Mexico	17.42685	-92.75213	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	Pmmex30	Rio Tacotalpa, Arroyo Tres	Mexico	17.48400	-92.77600	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	PmmxNSS2, PmmxNSSm–0	Tributary to Rio Ixtapangajoya, Chiapas state	Mexico	17.51000	-92.98000	*P. mexicana*	Palacios *et al.* (2013)
P. mexicana	13887	Rio Pantasma	Nicaragua	13.34163	-85.95636	*P. mexicana*	This study
P. mexicana	168807–14	Unnamed trib. to Lago de Apanás	Nicaragua	13.11843	-86.01022	*P. mexicana*	This study
P. mexicana	168168–74	Unnamed trib. to Rio Grande de Matagalpa, at stream just west of km marker 226	Nicaragua	13.25769	-85.45440	*P. mexicana*	This study
P. mexicana	168815–22	Unnamed trib. to Rio Grande de Matagalpa, east of Waslala (road to Siuna)	Nicaragua	13.35260	-85.35108	*P. mexicana*	This study
P. mexicana	167943, 168899	Unnamed trib. to Río Malacatoya at Teustepe, just off rd to Rama (Nicaragua Hwy 7), southwest of Boaco	Nicaragua	12.41825	-85.79299	*P. mexicana*	This study
Poecilia mexicana limantouri	PmlmPep	Pepeyocatitla, in Hidalgo	Mexico	20.91282	-98.39198	uncertain (*P. sulphuraria, P. thermalis*, *P. m. limantouri*)	Palacios *et al.* (2013)

P. mexicana limantouri	Pmlim6–11	Rio Soto la Marina at Mexican Hwy 180, Soto la Marina, Tamaulipas	Mexico	23.76063	-98.20607	uncertain (*P. sulphuraria, P. thermalis*, *P. m. limantouri*)	Palacios *et al.* (2013)
P. mexicana limantouri	PmlSnPe	San Pedro, in Hidalgo state	Mexico	N/A	N/A	uncertain (*P. sulphuraria, P. thermalis*, *P. m. limantouri*)	Palacios *et al.* (2013)
Poecilia orri	08.2060, 08.2061, 08.2066	Island of Roatan, Bay Islands, off northern coast of Honduras	Honduras	N/A	N/A	*P. mexicana*	N/A
***Poecilia reticulata* (OG)**	Pret	–	Trinidad	N/A	N/A	*P. reticulata*	Michael Tobler; this study
***P. reticulata* (OG)**	4289	Guanapo River	Trinidad	10.61743	-61.39936	*P. reticulata*	This study
***P. reticulata* (OG)**	4290	Guanapo River	Trinidad	10.61743	-61.39936	*P. reticulata*	This study
Poecilia salvatoris	Psalv	–	Honduras	N/A	N/A	*P. salvatoris*	Michael Tobler; this study
***Poecilia* sp.**	168888–90	Río Barbilla at province road just west of B-Line, a few km west of the road to Matina (dirt road, off northeast side of big road to Puerto Limon)	Costa Rica	10.04416	-83.33383	*P. mexicana*	This study
***Poecilia* sp.**	168891	Río Barbilla at province road just west of B-Line, a few km west of the road to Matina (dirt road, off northeast side of big road to Puerto Limon)	Costa Rica	10.04416	-83.33383	*P*. sp. "Patuca"	This study
***Poecilia* sp.**	168892–95	Río Barbilla at province road just west of B-Line, a few km west of the road to Matina (dirt road, off northeast side of big road to Puerto Limon)	Costa Rica	10.04416	-83.33383	*P. mexicana*	This study
***Poecilia* sp.**	168696–702	Rio Cabuyo at unnamed province road to la Reserva Biologica Lomas Bardudal	Costa Rica	10.48961	-85.38555	*P. mexicana*	This study
***Poecilia* sp.**	168800–06	Rio Diriá at CA1 approximately 2-3 km north of Santa Cruz, Nicoya Peninsula	Costa Rica	10.26677	-85.59261	*P. mexicana*	This study
***Poecilia* sp.**	167887, 16788, 167890–94,	Rio Liberia at outskirts of Liberia, Nicoya Peninsula	Costa Rica	10.62745	-85.43412	*P. mexicana*	This study
***Poecilia* sp.**	167904–06, 167908–10	Rio Salto at CA1 southeast of Liberia, Nicoya Peninsula	Costa Rica	10.56106	-85.39192	*P. mexicana*	This study
***Poecilia* sp.**	167895–902	Rio Sardinal at Sardinal, on Costa Rica Rd 151 approximately 5 km from 21	Costa Rica	10.51508	-85.65166	*P. mexicana*	This study
***Poecilia* sp.**	168792–98	Rio Tempisque on road between Guardia and Comunidad, Nicoya Peninsula	Costa Rica	10.57220	-85.39192	*P. mexicana*	This study
***Poecilia* sp.**	167933, 16734	Unnamed trib. to Río Tempisque drainage approximately 2 km south of Belén, Nicoya Peninsula	Costa Rica	10.39076	-85.59045	*P. mexicana*	This study
***Poecilia* sp.**	7957, 7958, 7960	Arroyo Comiston at La Pasion	Guatemala	16.55440	-90.19270	*P. mexicana*	This study
***Poecilia* sp.**	8084	Arroyo Sal Si Puedes, Belize	Guatemala	16.95730	-89.35930	*P. mexicana*	This study
***Poecilia* sp.**	7826, 7827, 7833	Las Conchas, Pueblo Canhuinic	Guatemala	15.83227	-90.33377	*Poecilia* sp.	This study
***Poecilia* sp.**	8181, 8184–86	Rio Amatillo at Lago Izabal near Venta de El Amatillo	Guatemala	15.53910	-88.89830	*P. mexicana*	This study
***Poecilia* sp.**	8288–90, 8294	Rio Chaguacal, trib. to Rio Polochic	Guatemala	15.31617	-89.85556	*P. mexicana*	This study
***Poecilia* sp.**	8241, 8245, 8246, 8248	Rio Dona Maria, trib. to Rio Motagua, near Motagua	Guatemala	15.20910	-89.24810	*P. mexicana*	This study

***Poecilia* sp.**	7995, 7996, 7999	Rio La Pasion	Guatemala	16.55060	-90.23010	*P. mexicana*	This study
***Poecilia* sp.**	8222, 8230, 8232	Rio Lobo, trib. to Rio Motagua, near Motagua	Guatemala	15.18160	-89.29940	*P. mexicana*	This study
***Poecilia* sp.**	7906, 7914, 7915, 7925	Rio Sebol at Finca Sebol	Guatemala	15.80630	-89.94480	*P. mexicana*	This study
***Poecilia* sp.**	7839, 7854, 7855	Rio San Simon	Guatemala	15.84110	-90.28920	*P. mexicana*	This study
***Poecilia* sp.**	8111, 8113, 8114	Zona Militar Estanque, locality approximated	Guatemala	16.90290	-89.72929	*P. mexicana*	This study
***Poecilia* sp.**	8719, 8720	Quebrada Carrizal, trib. to Rio Patuca, at Terrero Blanco	Honduras	14.41290	-86.04006	*P.* sp. "Patuca"	This study
***Poecilia* sp.**	8534, 8541	Quebrada de Chicho between Comunidades de Achiote and Cholomena	Honduras	15.53480	-86.21170	*P. mexicana*	This study
***Poecilia* sp.**	8705, 8706	Quebrada de Mercado at Rio Guayambre, trib. to Rio Patuca, near La Cruz	Honduras	14.35633	-86.06521	*P.* sp. "Patuca"	This study
***Poecilia* sp.**	8901–03, 8906	Quebrada Las Marias, trib. to Rio Choluteca, at Orocuina	Honduras	13.34770	-87.17430	*P. sphenops*	This study
***Poecilia* sp.**	8774	Quebrada San Jose at Guayape San Francisco de Becerra	Honduras	14.62763	-86.14237	*P.* sp. "Patuca"	This study
***Poecilia* sp.**	8475	Rio Camalote 1 km north of Macuelizo near Guatemala-Honduras border	Honduras	15.32656	-88.66264	*P. sphenops*	This study
***Poecilia* sp.**	8470, 8471	Rio Camalote 1 km north of Macuelizo near Guatemala-Honduras border	Honduras	15.32656	-88.66264	*P. mexicana*	This study
***Poecilia* sp.**	8872	Rio Chiquito, trib. to Rio Nacaome, in Nacaome 200 m downstream from CA 1	Honduras	13.54170	-87.47880	*P. sphenops*	This study
***Poecilia* sp.**	8873, 8875	Rio Chiquito, trib. to Rio Nacaome, in Nacaome 200 m downstream from CA 1	Honduras	13.54170	-87.47880	*P. mexicana*	This study
***Poecilia* sp.**	8917	Rio Chiquito, trib. to Rio Choluteca, just off Carretera N-85 ~3 km west of Apacilagua, near Orocuina	Honduras	13.48270	-87.09900	*P. sphenops*	This study
***Poecilia* sp.**	8914–16	Rio Chiquito, trib. to Rio Choluteca, just off Carretera N-85 ~3 km west of Apacilagua, near Orocuina	Honduras	13.48270	-87.09900	*P. mexicana*	This study
***Poecilia* sp.**	8463, 8465, 8466	Rio El Sauce (Amarillo), trib. to Rio Santa Rita	Honduras	14.86603	-89.06783	*P. mexicana*	This study
***Poecilia* sp.**	8375	Rio El Sauce (Amarillo), trib. to Rio Motagua, at Copán Ruins	Honduras	14.85589	-89.12355	*P. sphenops*	This study
***Poecilia* sp.**	8372–74	Rio El Sauce (Amarillo), trib. to Rio Motagua, at Copán Ruins	Honduras	14.85589	-89.12355	*P. mexicana*	This study
***Poecilia* sp.**	8858, 8860	Rio Goascorán at Caridad	Honduras	13.82770	-87.69480	*P. mexicana*	This study
***Poecilia* sp.**	8823	Rio Goascorán ~2.2 km southwest of Goascorán and 0.7 km south of El Amatillo	Honduras	13.58928	-87.76212	*P. sphenops*	This study
***Poecilia* sp.**	8805, 8807, 8815	Rio Goascorán ~2.2 km southwest of Goascorán and 0.7 km south of El Amatillo	Honduras	13.58928	-87.76212	*P. mexicana*	This study
***Poecilia* sp.**	8636–39	Rio Humuya, trib. to Rio Ulúa, at Comayagua	Honduras	14.45370	-87.65230	*P. sphenops*	This study
***Poecilia* sp.**	8311, 8312, 8314, 8316	Rio Lempa at Nueva Ocotepeque	Honduras	14.39417	-89.20816	*P. sphenops*	This study
***Poecilia* sp.**	8408, 8411	Rio Naco, trib. to Rio Ulúa	Honduras	15.34147	-88.62480	*P. mexicana*	This study
***Poecilia* sp.**	8558, 8565	Rio Taujica at Taujica	Honduras	15.68100	-85.93930	*P. mexicana*	This study
***Poecilia* sp.**	8628	Rio Yojoa at Ulúa	Honduras	15.03480	-87.92870	*P. sphenops*	This study

***Poecilia* sp.**	8362, 8363	Rio Tio Higuito, trib. to Rio Motagua, at cuenca near Higuito	Honduras	14.83940	-89.16819	*P. mexicana*	This study
***Poecilia* sp.**	PslBan3, PslBan2, PslBaños	Baños del Azufre	Mexico	17.55200	-92.99900	uncertain (*P. sulphuraria, P. thermalis*, *P. m. limantouri*)	Palacios *et al.* (2013)
***Poecilia* sp.**	PslLaGl3, PslLaGl1.0	La Gloria	Mexico	17.53201	-93.01513	uncertain (*P. sulphuraria, P. thermalis*, *P. m. limantouri*)	Palacios *et al.* (2013)
***Poecilia* sp.**	Psp2S.0, Psp1T.0	Suchiapa, in Chiapas	Mexico	16.61077	-93.08451	*P. catemaconis* (*catemaconis* + "*sphenops*" sp. 1)	Palacios *et al.* (2013)
***Poecilia* sp.**	13876	Quebrada Venquilla, trib. to Rio Pantasma	Nicaragua	13.33811	-85.94880	*P. mexicana*	This study
***Poecilia* sp.**	14172	Rio Babasca, trib. to Rio Tuma	Nicaragua	13.25611	-85.54452	*P. mexicana*	This study
***Poecilia* sp.**	14722	Rio Cardenas, trib. to Lago Nicaragua	Nicaragua	11.19033	-85.51783	*P. mexicana*	This study
***Poecilia* sp.**	14137	Rio Ceperna	Nicaragua	13.93861	-84.82472	*P. mexicana*	This study
***Poecilia* sp.**	14145, 14146	Rio Chico Smith	Nicaragua	13.55500	-84.86333	*P. mexicana*	This study
***Poecilia* sp.**	14100	Rio Danli	Nicaragua	13.82027	-85.04444	*P. mexicana*	This study
***Poecilia* sp.**	13313–15	Rio de las Calabazas	Nicaragua	12.67075	-86.09138	*P. mexicana*	This study
***Poecilia* sp.**	13921–24	Rio El Cua	Nicaragua	13.51294	-85.80986	*P. mexicana*	This study
***Poecilia* sp.**	14256	Rio El Guineo, trib. to Rio Labú	Nicaragua	13.50305	-84.84472	*P. mexicana*	This study
***Poecilia* sp.**	13666	Rio Espavel, trib. to Rio Chimalate	Nicaragua	12.01288	-84.66830	*P. mexicana*	This study
***Poecilia* sp.**	13429–31	Río Estelí	Nicaragua	13.09797	-86.36033	*P. mexicana*	This study
***Poecilia* sp.**	13838, 13845, 13846	Rio Jiguina, trib. to Lago Apanas	Nicaragua	13.15058	-85.92922	*P. mexicana*	This study
***Poecilia* sp.**	14041–43	Rio Juan Blanco	Nicaragua	13.77111	-85.64833	*P. mexicana*	This study
***Poecilia* sp.**	14070–72	Rio Kum	Nicaragua	13.63583	-85.36500	*P. mexicana*	This study
***Poecilia* sp.**	14060	Rio La Lana	Nicaragua	13.67611	-85.79611	*P. mexicana*	This study
***Poecilia* sp.**	13450, 13451, 13456	Rio Macuelizo	Nicaragua	13.60944	-86.47483	*P. mexicana*	This study
***Poecilia* sp.**	13997, 13998	Rio Milan	Nicaragua	13.42472	-85.98500	*P. mexicana*	This study
***Poecilia* sp.**	13986–88	Rio Orosal	Nicaragua	13.33335	-86.20416	*P. mexicana*	This study
***Poecilia* sp.**	14131, 14132	Rio Pia	Nicaragua	13.73222	-84.51472	*P. mexicana*	This study
***Poecilia* sp.**	13972–74	Rio Sacramento	Nicaragua	13.29133	-86.18027	*P. mexicana*	This study
***Poecilia* sp.**	13969–71	Rio San Gabriel	Nicaragua	13.17450	-86.28527	*P. mexicana*	This study
***Poecilia* sp.**	14110–12	Rio Sangsangwas	Nicaragua	13.91666	-84.56333	*P. mexicana*	This study
***Poecilia* sp.**	14154, 14155	Rio Santa Rita	Nicaragua	13.46000	-84.91444	*P. mexicana*	This study
***Poecilia* sp.**	14195, 14196	Rio Yaoska, trib. to Rio Tuma	Nicaragua	13.26233	-85.43922	*P. mexicana*	This study
***Poecilia* sp.**	14231, 14232	Rio Yaoya	Nicaragua	13.69575	-84.69794	*P. mexicana*	This study
***Poecilia* sp.**	167935–42	Trib. to Rio Grande de Matagalpa at Puente de Tierra Azul on road to Rio Blanco (town)	Nicaragua	12.68476	-85.54708	*P. mexicana*	This study
***Poecilia* sp.**	168159–66	Trib. to unnamed trib. of the Río Grande de Matagalpa, ~32 km west of Rio Blanco (town) on	Nicaragua	12.82341	-85.44279	*P. mexicana*	This study

		road between Matagalpa and Rio Blanco (town)					
***Poecilia* sp.**	14294	Unnamed trib. (Rio Blanco?) ~5 km southwest of Rio Blanco (town at Nicaragua Hwy 21B)	Nicaragua	12.87886	-85.21286	*P. mexicana*	This study
***Poecilia* sp.**	167960, 167962, 167964	Unnamed trib. to Lago Managua between Estelí and León	Nicaragua	13.22797	-86.55272	*P. mexicana*	This study
***Poecilia* sp.**	167927, 167930–32, 168708–10	Unnamed trib. to Lago Nicaragua just east of mile marker km 238 on road to San Miguelito, Chontales	Nicaragua	11.50538	-84.83956	*P. mexicana*	This study
***Poecilia* sp.**	167961, 167963	Unnamed trib. to Quebrada de Pedernal, 300 m south of Pedernal	Nicaragua	13.22797	-86.55272	*P. sphenops*	This study
***Poecilia* sp.**	167952–54, 167959	Unnamed trib. to Rio Estelí, thus a trib. to Río Coco	Nicaragua	13.05866	-86.35114	*P. mexicana*	This study
***Poecilia* sp.**	168175–82	Unnamed trib. to Rio Grande de Matagalpa, west of La Mora and slightly further west of La Dalia	Nicaragua	13.22058	-85.72626	*P. mexicana*	This study
***Poecilia* sp.**	168896–68	Unnamed trib. to Rio Malacatoya at Teustepe, just off road to Rama (Nicaragua Hwy 7), southwest of Boaco	Nicaragua	12.41825	-85.79299	*P. mexicana*	This study
***Poecilia* sp.**	168151–58	Unnamed trib. to Rio Tuma northwest of Rio Blanco (town; flowing from Mt. Musun)	Nicaragua	12.93613	-85.23434	*P. mexicana*	This study
***Poecilia* sp.**	167928, 16729	Unnamed trib./headwater of Rio Mico northeast of San Pedro de Lovago	Nicaragua	12.13630	-85.04597	*P. mexicana*	This study
Poecilia sphenops	7731	Rio Pachipa, San Antonio Suchitepeque	Guatemala	14.50340	-91.40770	*P.* "*sphenops*" sp. 1	This study
P. sphenops	7780	Rio Sinacapa, 9 km from Guanagasapa	Guatemala	14.20190	-90.70760	*P.* "*sphenops*" sp. 1	This study
P. sphenops	4303	Rio Agua Buena	Honduras	15.76611	-86.99889	*P. mexicana*	This study
P. sphenops	Psphe35	Rio Coatzocoalcos, Arroyo Prieto	Mexico	N/A	N/A	*P. catemaconis* (*catemaconis* + "*sphenops*" sp. 1)	Tobler *et al.* (2011); Palacios *et al.* (2013)
P. sphenops	Psphe36	Rio La Venta, Rio Ninguillo, Pomposa Castellano, Valle Morelos	Mexico	N/A	N/A	*P. catemaconis* (*catemaconis* + "*sphenops*" sp. 1)	Tobler *et al.* (2011); Palacios *et al.* (2013)
P. sphenops	13327	Rio Grande de Matagalpa just off CA 1, ~0.5 km southwest of Sebaco	Nicaragua	12.84516	-86.10272	*P. mexicana*	This study
Poecilia sulphuraria	Psulp38–41	La Gloria, Rio Pichucalco	Mexico	17.53201	-93.01513	uncertain (*P. sulphuraria, P. thermalis*, *P. m. limantouri*)	Tobler *et al.* (2011); Palacios *et al.* (2013)
Poecilia thermalis	PthrL3.0, PthrL2.0, PthrLa.0	La Esperanza large spring	Mexico	17.51100	-92.98300	uncertain (*P. sulphuraria, P. thermalis*, *P. m. limantouri*)	Palacios *et al.* (2013)
P. thermalis	PthrS3.0, PthrS2.0, PthrSm.0	La Esperanza small spring	Mexico	17.51100	-92.98000	uncertain (*P. sulphuraria, P. thermalis*, *P. m. limantouri*)	Palacios *et al.* (2013)
***Xiphophorus helleri* (OG)**	Xhell	–	Mexico	N/A	N/A	–	Hrbek *et al*. (2007)
***Xiphophorus maculatus* (OG)**	Xmac	–	Mexico	N/A	N/A	–	Hrbek *et al.* (2007)

Abbreviations: N/A, not available; OG, outgroup; sp., species; trib., tributary.

References

1. Doadrio I, Perea S, Alcaraz L, Hernandez N (2009) Molecular phylogeny and biogeography of the Cuban genus *Girardinus* Poey, 1854 and relationships within the tribe Girardinini (Actinopterygii, Poeciliidae). *Molecular Phylogenetics and Evolution*, 50, 16-30.
2. Hrbek T, Seckinger J, Meyer A (2007) A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. *Molecular Phylogenetics and Evolution*, **43**(3), 986-998.
3. Lee JB, Johnson JB (2009) Biogeography of the livebearing fish *Poecilia gillii* in Costa Rica: are phylogeographical breaks congruent with fish community boundaries? *Molecular Ecology*, **18**, 4088-4101.
4. Ornelas-García CP, Dominguez-Dominguez O, Doadrio I (2008) Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evolutionary Biology*, **8**, 340.
5. Palacios M, Arias-Rodriguez L, Plath M, Eifert C, Lerp H, Lamboj A, Voelker G, Tobler M (2013) The rediscovery of a long described species reveals additional complexity in speciation patterns of poeciliid fishes in sulfide springs. *PLoS One*, **8**, e71069.
6. Tobler M, Palacios M, Chapman LJ, Mitrofanov I, Bierbach D, Plath M, Arias-Rodriguez L, Garcia de Leon FJ, Mateos M (2011) Evolution in extreme environments: replicated phenotypic differentiation in livebearing fish inhabiting sulfidic springs. *Evolution*, **65**, 2213-2228.

Appendix S1 Supplementary methods and results.

Taxon sampling and sequencing, and outgroup details

Here, we provide additional sequence data and outgroup descriptions relevant to our analyses but not listed in the main text. As noted in the manuscript, our individual analyses utilized up to 21 tips representing outgroup taxa, and the sample data for each outgroup are given in table format in supplementary Data S1. However, the pool of outgroup taxa spanned 23 'outgroup' tips representing 15 nominal poeciliid taxa, including (1) *Poecilia caucana*, the sister taxon to the members of the *P. sphenops* species complex (based on analyses by Alda *et al.* [1]); the sail-fin mollies (2) *P. latipinna* and (3) *P. latipunctata*; the South American guppies (4) *Micropoecilia picta* and (5) *Poecilia reticulata*; the Mexican swordtails (6) *Xiphophorus helleri* and (7) *X. maculatus*; the Central American Pike Killifish, (8) *Belonesox belizanus*; and six species of fishes from the genus *Limia*, a closely related genus whose members were formerly included within *Poecilia* subgenus *Limia*: (9) *L. dominicensis*, (10) *L. melanogaster*, (11) *L. melanonotata*, (12) *L. tridens*, (13) *L. vittata*, (14) *L. heterandria*, and (15) *L. perugiae*. GenBank numbers for the sequences we used to represent these outgroup taxa are provided in Data S1. A total of 21 outgroup samples representing the first 13 of these outgroup taxa were used in phylogenetic analyses of our 'concatenated mtDNA' dataset, including the BEAST relaxed clock analysis whose results are presented in supplementary Fig. S1A. However, our 'concatenated nDNA' dataset included only 7 outgroup tips representing the following five species: *P. latipinna*, *P. latipunctata*, *M. picta*, *P. reticulata*, and *L. perugiae*; thus, concatenated gene trees from analyses of this dataset included up to five outgroup species, though only *Poecilia* are shown in the resulting figures (e.g. Fig. 3). As mentioned in the main text, our *BEAST analyses of the 'concatenated mtDNA + nDNA' dataset included outgroup samples from 15 species. These 15 outgroup lineages consisted of all 15 of the outgroup species listed above; again, most of these outgroups except selected *Poecilia* were pruned from the trees resulting from such analyses before finalizing our figures (e.g. Fig. S1B).

Regarding PHASE analyses, *Glyt* and X-*yes* alignments could not be completely resolved due to multiple positions with >2 variants per position, so we analyzed phased alleles for all other loci and coded ambiguities in the *Glyt* and *X-yes* alignments as missing. Iterative analyses (e.g. using all six loci in BEAST) using alignments for which we had arbitrarily resolved ambiguities in the data for these two loci did not give results that were significantly different

than those presented in the manuscript (data not shown).

Neutrality and recombination

Consistent with expectations of neutral evolution, which was assumed in all of our analyses, Hudson-Kreitman-Aguadé tests [2] were non-significant for the full-cyt*b* dataset (N = 938 ingroup sequences, $\chi^2 = 0.027$, $P = 0.87$), the cyt*b* matrix from the concatenated mtDNA dataset (N = 134 ingroup sequences, $\chi^2 = 0.046$, $P = 0.83$), and the *cox1* matrix from the concatenated mtDNA dataset (N = 111 ingroup sequences, $\chi^2 = 0.058$, $P = 0.81$). One *P. caucana* outgroup sample was used in each HKA test.

For additional insight into neutrality and in an attempt to cross-validate the HKA test results, we also ran McDonald & Kreitman tests [3] in DnaSP on the same mtDNA datasets. One *P. caucana* outgroup sample was used in each MK test. Similar to HKA tests, the MK tests also supported neutrality of the mtDNA data in analyses of the full-cyt*b* dataset (N = 938 ingroup sequences, alpha = 0.11, $P_{\text{Fisher}} = 0.818$; $G = 0.067$, $P = 0.796$) and the *cox1* matrix from the concatenated mtDNA dataset (N = 111 ingroup sequences, alpha = 1.00, $P_{\text{Fisher}} = 1.00$; *G*-test could not be performed). However, an MK test on the cyt*b* matrix from the concatenated mtDNA dataset was significant (N = 134 ingroup sequences, alpha = 0.723, $P_{\text{Fisher}} = 0.0026$; $G = 8.823$, $P = 0.0030$). To further evaluate whether the non-neutral signal in the concatenated mtDNA cyt*b* matrix may have resulted from past population dynamics, e.g. population genetic bottlenecking or expansion, rather than selection we conducted additional coalescent simulations on this dataset in DnaSP (again, testing significance with 1000 simulations) using the neutrality statistics Fu's F_S and R_2 [4]. We estimated a negative value of F_S for this dataset indicating potential past population growth, but this result was non-significant (mean $F_S = -1.102$, 95% confidence interval = [−14.855, 11.437], $P = 0.593$). However, a positive and significant R_2 value (mean $R_2 = 0.088$, 95% confidence intervals = [0.047, 0.145], $P < 0.001$) indicated that the non-neutral signal in the concatenated mtDNA cyt*b* dataset may owe to past population expansion rather than selection. In view of these results, further testing using analogous coalescent simulations of Fay & Wu's [5] *H* in DnaSP was used to evaluate whether the hypothesis of positive selection could be ruled out. The results were consistent with the interpretation that the MK and neutrality test results reported above for the concatenated mtDNA cyt*b* dataset were not influenced by positive selection, e.g. due to hitchhiking (mean $H = -1.350$, 95% confidence interval = [−77.191, 28.226], $P = 0.350$). Overall, the various mtDNA analyses

above indicate that all of the mtDNA data used in this study are selectively neutral, though likely influenced by historical demographic fluctuations. As a result, it seems worthwhile to delve further into this issue with analyses targeted at better understanding the historical demography of the *P. sphenops* species complex.

We ran seven different tests for recombination on each of the nuclear loci analyzed in this study. Six tests were run on each of the loci simultaneously in the program RDP3 (citation in main text; default parameters unless stated otherwise, in parentheses below) using different algorithms for recombination detection, including the original RDP method ([6]; window size = 30), the GENCONV Local method [7], the RecScan/Bootscan method ([8]; window size = 100, step size = 20,500 bootstrap replicates), the MaxChi Local method [9], the Chimaera method [10], and the 3seq method [11]. Among all 30 tests run RDP3 (six algorithms run on each of 5 loci), we recovered evidence for only three unique events corresponding to three recombination signals, which were only discovered by three methods (MaxChi, Chimaera, 3seq) when analyzing the X-*yes* dataset. All other tests conducted in RDP3 for X-*yes* and the other loci inferred zero events/signals.

Runs of our seventh test for recombination using coalescent simulations in DnaSP assumed intermediate levels of recombination (*R*, per gene) and empirical mutation parameter θ (per gene). The simulations were run based on an implementation of the coalescent based on Hudson [12], and DnaSP obtained the estimated *R*-values using the method of Hudson [13], whereas observed RM estimates were obtained using equations in Hudson & Kaplan [14]. Estimated values of *R* used in the simulations were, by gene, as follows: *ldh-A*, 0.499; *RPS7*, 0.399; X-*src*, 4.099; X-*yes*, 0.001; *Glyt*, 47.299. Empirically estimated minimum numbers of recombination events (R_M) calculated directly from the data were, by gene, as follows: *ldh-A*, 0; *RPS7*, 5; X-*src*, 9; X-*yes*, 10; *Glyt*, 5. The results of these coalescent simulations were non-significant, indicating less recombination than expected. Specifically, the probabilities of recombination being less than or equal to the minimum number of events (*P*-values) were each non-significant at the test level ($\alpha = 0.05$): *ldh-A*, $P = 0.87$; *RPS7*, $P = 1.00$; X-*src*, $P = 1.00$; X-*yes*, $P = 1.00$; *Glyt*, $P = 0.58$.

Coalescent-based species delimitation

As noted in the Discussion section, we conducted additional analyses to evaluate the potential effects of phylogenetic branch lengths and their uncertainty on our GMYC species

discovery analysis. Specifically, we tested whether an algorithm similar to our bGMYC modeling analysis, but not relying heavily on phylogenetic branch lengths, gave comparable results to our preliminary species delimitation hypotheses shown in Fig. 2. To accomplish this, we delimited species on our data using a Bayesian implementation of the PTP method [15], where PTP stands for "Poisson tree process" used to model speciation rates under this method. PTP was suitable for our purposes because it analyzes substitution patterns along gene tree branches without requiring an ultrametric topology, and without utilizing branch length information [15]. It also happens to be fast and intuitive to implement. Bayesian PTP analyses were run on the concatenated mtDNA ML gene tree from GARLI shown in Fig. 2, which, importantly, is similar to the MCC tree that we analyzed in all of our bGMYC runs (Fig. S1). We ran our PTP analysis using the "bPTP.py" python script implemented on the PTP web server (http://species.h-its.org/ptp/). As noted in the text, we found that Bayesian PTP gave species delimitations that were nearly identical to our bGMYC-delimited species (data not shown). Thus, we conclude that our mtDNA data are robust to species delimitation using methods with and without taking branch lengths into account. However, it remains unclear whether and to what extent uneven sampling across distinct species lineages may have influenced these species discovery analyses. Although evaluating such properties of the data and the bPTP algorithm are beyond the scope of the present study, we expect that these topics will be addressed using simulations in future studies.

Evolutionary rates estimated in *BEAST, used in JML analyses

The main *BEAST [16] analysis in BEAST v2.0.2 [17] described in the main text employed relaxed clocks for all loci and two fossil/biogeographical calibration points. Based on five independent runs conducted during this analysis, we inferred the following evolutionary substitution rates for each locus: concatenated mtDNA, 0.005656; *ldh-A*, 0.001768; *RPS7*, 0.0009142; X-*src*, 0.0004724; X-*yes*, 0.005094; *Glyt*, 0.0002984. Each of these rates is a mean estimate in units of substitutions per site per million years (subs/site/myr), per lineage. Note the mtDNA rate fell within the uniform 'fish rate' prior set on the locus for this analysis (0.0017–0.014 subs/site/myr), as expected.

We ran a second *BEAST analysis, again based on five independent runs, specifying independent relaxed clocks for each locus, but no calibration points. This analysis was

conducted in order to facilitate simulation analyses in JML [18], and calibration points were not available because only ingroup taxa in the *P. sphenops* species complex were included in the analysis, as warranted by our JML analyses. From this second *BEAST analysis, we inferred the following mean relative rate estimates for each locus: concatenated mtDNA, 1.049; *ldh-A*, 0.652; *RPS7*, 0.282; X-*src*, 0.166; X-*yes*, 0.473; *Glyt*, 0.08112. In contrast to the rates reported above, these values are relative evolutionary rates. Values from this set of relative rates were supplied to JML during coalescent simulations used to test for hybridization versus incomplete lineage sorting in the *ldh-A*, *RPS7*, and X-*src* loci.

Full/additional JML results

Full JML results for nDNA loci MINUS CLADE 7 (due to lack of observed sequence data for nDNA loci, which is required to calculate exact *minDist* probabilities):

By contrast, JML simulations detected introgressed nuclear *ldh-A* sequences between *P. gillii-P. hondurensis*, *P. butleri-P. catemaconis/sphenops* (clade 2-a), *P. mexicana-P. catemaconis/sphenops*, *P. mexicana-P. hondurensis*, and *P. mexicana-P. butleri* species pairs (*minDist P* = 0.001–0.048); introgressed *RPS7* sequences between *P. hondurensis-P.* sp. "Tipitapa", *P. hondurensis-P. catemaconis/sphenops*, *P. hondurensis-P. sphenops* (clade 2-b), *P. butleri-P. catemaconis/sphenops*, *P. mexicana-P. catemaconis/sphenops*, and *P. mexicana-P. butleri* species pairs (*minDist P* = 0.001–0.032); and introgressed X-*src* sequences between the *P. butleri-P. catemaconis/sphenops* species pairs (*minDist P* = 0.001).

By contrast, JML simulations detected introgressed nuclear *ldh-A* sequences between *P. gillii-P. hondurensis*, *P. butleri-P. catemaconis/sphenops* (clade 2-a), *P. mexicana-P. catemaconis/sphenops*, *P. mexicana-P. hondurensis*, and *P. mexicana-P. butleri* species pairs (*P* = 0.001–0.048); introgressed *RPS7* sequences between *P. hondurensis-P.* sp. "Tipitapa", *P. hondurensis-P. catemaconis/sphenops*, *P. hondurensis-P. sphenops* (clade 2-b), *P. butleri-P. catemaconis/sphenops*, *P. mexicana-P. catemaconis/sphenops*, and *P. mexicana-P. butleri* species pairs (*P* = 0.001–0.032); and introgressed X-*src* sequences between the *P. butleri-P. catemaconis/sphenops* species pairs (*P* = 0.001).

References

1. Alda FA, Reina RG, Doadrio I, Bermingham E (2013) Phylogeny and biogeography of the Poecilia sphenops species complex (Actinopterygii, Poeciliidae) in Central America. Molecular Phylogenetics and Evolution 66:1011-1026.
2. Hudson RR, Kreitman M, Aguadé M (1987) A test of neutral molecular evolution based on nucleotide data. Genetics, 116:153-159.
3. McDonald JH, Kreitman M (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. Nature, 351:652-654.
4. Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. Molecular Biology and Evolution, 19:2092-2100.
5. Fay JC, Wu CI (2000) Hitchhiking under positive Darwinian selection. Genetics, 155:1405-1413.
6. Martin D, Rybicki E (2000) RDP: detection of recombination amongst aligned sequences. Bioinformatics, 16:562-563.
7. Padidam M, Sawyer S, Fauquet CM (1999). Possible emergence of new geminiviruses by frequent recombination. Virology, 265:218-225.
8. Martin DP, Posada D, Crandall KA, Williamson C (2005) A modified BOOTSCAN algorithm for automated identification of recombinant sequences and recombination breakpoints. AIDS Research and Human Retroviruses, 21:98-102.
9. Maynard Smith J (1992). Analyzing the mosaic structure of genes. Journal of Molecular Evolution, 34:126-129.
10. Posada D, Crandall KA (2001). Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. Proceedings of the National Academy of Sciences of the United States of America, 98:13757-13762.
11. Boni MF, Posada D, Feldman MW (2007). An exact nonparametric method for inferring mosaic structure in sequence triplets. Genetics, 176:1035-1047.
12. Hudson RR (1990) Gene genealogies and the coalescent process. Oxford Surveys in Evolutionary Biology, 7:1-44.
13. Hudson RR (1987) Estimating the recombination parameter of a finite population model without selection. Genetic Research, 50:245-250.
14. Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics, 111:147-164.
15. Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013) A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29:2869-2876.
16. Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. Molecular

Biology and Evolution 27:570-580.

17. Bouckaert R, Heled J, Kühnert D, Vaughan TG, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST2: A software platform for Bayesian evolutionary analysis. PLoS Computational Biology. Available at: http://beast2.org/.
18. Joly S (2012) JML: testing hybridization from species trees. Molecular Ecology Resources 12:179-184.

Table S1 Summary of the taxonomy, tooth morphology, and distributions of species the *Poecilia sphenops* species complex.

Taxon name	Common name	Taxonomic hypotheses			Status ([3,5,6,9]; this study)	Inner jaw teeth	Biogeography	
		§Rosen & Bailey [1]	†Schultz & Miller [2]; Miller [3]	¶Alpírez Quesada [4]			Versant	Distribution [2–9]
Described species and subspecies								
Poecilia butleri* Jordan 1889	Pacific Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Valid	uni.	Pacific	Mexico to El Salvador
P. catemaconis* Miller 1975	Catemaco Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Valid	uni.	Atlantic	Lake Catemaco, Mexico
***P. chica* Miller 1975**	Dwarf Molly	*P. sphenops*	*P. sphenops* species complex	*P. sphenops* complex	Valid	tri.	Pacific	Basins of Cuetzamala River and Purificación in Jalisco, Mexico
P. gillii* (Kner 1863)	Gill's Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Valid	uni.	Atlantic, Pacific	Atlantic versant from Guatemala to Colombia, along the Pacific versant from Guatemala to the Terrába River, Costa Rica, and from the Grande River to the Bayano River in Panama
P. hondurensis* Poeser 2011	Honduras Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Valid	uni.	Atlantic	Caribbean drainages of Honduras
***P. marcellinoi* Poeser 1995**	Molly	*P. sphenops*	*P. sphenops* species complex	*P. sphenops* complex	Valid	tri.	Pacific	Ilopango Lake basin, El Salvador
***P. maylandi* Meyer 1983**	Balsas Molly	*P. sphenops*	*P. sphenops* species complex	*P. sphenops* complex	Valid	tri.	Pacific	Balsas River basin and Aguililla River, Mexico
P. mexicana* Steindachner 1863	Shortfin Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Valid	uni.	Atlantic, Pacific	Atlantic versant from northeastern Mexico to Costa Rica and in the Rio Tamarindo, in the Pacific slope of Nicaragua
P. m. mexicana* Steindachner 1863	Shortfin Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Synonym of *P. mexicana*; however, Menzel & Darnell [5] have recommended subspecies rank and suggested it intergrades with another subspecies (*P. m. limantouri*) in eastern Mexico	uni.	Atlantic, Pacific	Rio Cazones south (at least) to Rio Jamapa system in eastern Mexico
P. m. limantouri* Jordan & Synder 1901	Limantour's Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Synonym of *P. mexicana* [1]; however, others have recommended	uni.	Atlantic	Southern Rio Grande and Rio San Fernando headwaters, south to

					subspecies rank and suggested it intergrades with *P. m. mexicana* in eastern Mexico [5]			Pánuco River, Mexico
P. orri* Fowler 1943	Mangrove Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Valid	uni.	Atlantic	Western coasts of Yucatan Peninsula southeast to northern Honduras
***P. salvatoris* Regan 1907**	Salvador Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Valid	uni.	Pacific	El Salvador
P. sphenops* Valenciennes 1846	Mexican Molly	*P. sphenops*	*P. sphenops* species complex	*P. sphenops* complex	Valid	uni.	Atlantic, Pacific	Atlantic slope of Mexico from the Palma Sola River to the Grijalva River basin, and along the Pacific slope from the Rio Verde basin into Guatemala
P. sulphuraria* (Álvarez 1948)	Sulphur Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Valid	uni.	Atlantic	Baños del Azufre, near Teapa, Tabasco, Mexico
***P. teresae* Greenfield 1990**	Mountain Molly	*P. sphenops*	*P. sphenops* species complex	*P. mexicana* complex	Valid	uni.	Atlantic	Mountain Pine Ridge, Mayan Mountain Range, Belize
Molecular operational taxonomic units (OTUs)								
"*sphenops*" sp. 1*	–	–	–	–	Candidate species ([6]; this study)	–	Atlantic, Pacific	Rio Goascorán and Rio Ulúa, Honduras
"*gillii*" sp. 2*	–	–	–	–	Candidate species ([6]; this study)	–	Atlantic	Rio Acla, Panama
***P.* sp. "Patuca"**	–	–	–	–	Considered part of the "*P. gillii*" lineage, clade 5 (this study)	–	Atlantic	Rio Patuca basin, Honduras
***P.* sp. "Tipitapa"**	–	–	–	–	Candidate species (this study)	–	Atlantic	Rio Tipitapa and northern Lake Nicaragua tributaries, Rio San Juan basin, Nicaragua

Asterisks placed by taxon names indicate nominal taxa or molecular OTUs previously recognized in the *P. sphenops* species complex by other authors, and that we also sampled in our study. This table also presents data from [6] on differences in inner jaw tooth morphology displayed among taxa from the species complex (uni., unicuspid; tri., tricuspid).

[§]Interpretation recognizing the existence of a single, polytypic species; this study synonymized various taxa under *P. sphenops*.

[†]Interpretation recognizing multiple species forming a single "*P. sphenops* species complex."

¶Interpretation recognizing two species groups or complexes, the "*P. sphenops* complex" and the "*P. mexicana* complex", within the *P. sphenops* species complex *sensu lato*.

References

1. Rosen DE, Bailey RM (1963) The poeciliid fishes (Cyprinodontiformes), their structure, zoogeography, and systematics. Bulletin of the American Museum of Natural History 126:1-176.
2. Schultz RJ, Miller RR (1971) Species of the *Poecilia sphenops* complex (Pisces: Poeciliidae) in Mexico. Copeia 1971:282-290.
3. Miller RR (2005) *Freshwater Fishes of México*. The University of Chicago Press, Chicago.
4. Alpírez Quesada O (1971) Estudio sistemático del complejo *Poecilia sphenops* (Familia Poeciliidae) de Centroamérica en especial de las poblaciones de Costa Rica. Universidad de Costa Rica, San José.
5. Menzel BW, Darnell RM (1973) Systematics of *Poecilia mexicana* (Pisces: Poeciliidae) in Northern Mexico. Copeia 1973:225-237.
6. Alda FA, Reina RG, Doadrio I, Bermingham E (2013) Phylogeny and biogeography of the *Poecilia sphenops* species complex (Actinopterygii, Poeciliidae) in Central America. Molecular Phylogenetics and Evolution 66:1011-1026.
7. Matamoros WA, Kreiser BR, Schaefer JF (2012) A delineation of Nuclear Central America biogeographical provinces based on river basin faunistic similarities. Reviews in Fish Biology and Fisheries 22:351-365.
8. Bussing WA (1998) *Freshwater Fishes of Costa Rica*, 2nd Edn. Editorial de la Universidad de Costa Rica, San José, Costa Rica.
9. Poeser FN (2011) A new species of *Poecilia* from Honduras (Teleostei: Poeciliidae). Copeia 2011:418-422.

Table S2 DNA substitution models selected using DT-MODSEL.

DNA dataset	Partition	*n*	bp	Best model	Analysis
Full-cyt*b*					
	All cyt*b* (analyzed together, or as $N = 260$ haplotypes)	941	1086	TVM+Γ+*I*	TCS, DNASP
Concatenated mtDNA					
	1st + 2nd codon pos.	168	590	HKY+Γ+*I*	GARLI, MrBayes, BEAST, DNASP
	3rd codon pos.	168	1180	GTR+Γ	GARLI, MrBayes, BEAST, DNASP
Concatenated nDNA					
	ldh-A	50(42)	191	JC+*I*	GARLI, BPP
	RPS7	50(44)	1158	K80+Γ+*I*	GARLI, BPP
	X-*src*	50(45)	518	K80+*I*	GARLI, BPP
	X-*yes*	50(20)	833	HKY+*I*	GARLI, BPP
	Glyt	50(21)	915	K80+*I*	GARLI, BPP
Concatenated mtDNA + nDNA					
	mtDNA 1st + 2nd codon pos.	80	590	TVM+Γ+*I*	GARLI, MrBayes, BEAST (*BEAST)
	mtDNA 3rd codon pos.	80	1180	TrN+Γ+*I*	GARLI, MrBayes, BEAST (*BEAST)
	mtDNA 1st + 2nd codon pos., reduced dataset*	50	590	HKY+Γ	BEAST (*BEAST)
	mtDNA 3rd codon pos., reduced dataset*	50	1180	TrN+Γ	BEAST (*BEAST)
	ldh-A	80(42)	191	JC+Γ+*I*	GARLI, MrBayes, BEAST (*BEAST)
	RPS7	80(44)	967	HKY+Γ	GARLI, MrBayes, BEAST (*BEAST)
	X-*src*	80(45)	518	K80+*I*	GARLI, MrBayes, BEAST (*BEAST)
	X-*yes*	80(20)	833	HKY+Γ	GARLI, MrBayes, BEAST (*BEAST)
	Glyt	80(21)	915	K80+*I*	GARLI, MrBayes, BEAST (*BEAST)

Model selection analyses using the decision theory algorithm in DT-MODSEL [1] supported different best-fit models of DNA evolution for different datasets, including datasets filtered by codon partitions. We preferred DT-MODSEL for our substitution model selection analyses, rather than other model selection software, because DT-MODSEL has been shown to recover models that yield superior ML branch lengths relative to other comparable programs [1]. This table lists model selection results for each dataset analyzed in this study, as well as the analyses that each dataset (thus molecular model, wherever possible) was used in. Symbols and abbreviations: Γ, gamma-distributed rate variation; bp, number of nucleotide base pairs; DNASP, DNA polymorphism and neutrality statistics

analyses conducted in the program by the same name; *I*, parameter representing proportion of invariable sites; ML, maximum likelihood phylogenetic analyses estimating gene trees; *n*, sample size (numbers correspond to sequence alignment sizes, except for multilocus datasets the numbers in parentheses are sample sizes for each locus). Asterisks in the "Partition" column denote mitochondrial datasets that are reduced versions of the 86-taxon mtDNA dataset, and which we used (along with the nDNA in the concatenated mtDNA + nDNA dataset) in the *BEAST analyses ran to create a posterior distribution of species trees with appropriate tips for analysis in JML.

References

1. Minin V, Abdo Z, Joyce P, Sullivan J (2003) Performance-based selection of likelihood models for phylogeny estimation. Syst Biol 52: 674-683.

Table S3 Sequence attributes and DNA polymorphism levels in each of the datasets analyzed in this study, overall and by gene.

DNA dataset	Partition	*n*	bp	Variable characters (%)	Parsimony informative characters (%)	Overall mean *d* (s.e.)
Full-cyt*b*						
	All cyt*b* ingroup sequences	941	1086	341 (31.40)	301 (27.72)	0.026 (0.0020)
Concatenated mtDNA		168				
	1st + 2nd codon positions, ingroup only	147	1180	213 (18.05)	131 (11.10)	N/A
	3rd codon position, ingroup only	147	590	233 (39.49)	209 (35.42)	N/A
	All cyt*b*	155	1086	460 (42.36)	403 (37.11)	0.064 (0.0031)
	All *cox1*	115	684	206 (30.12)	128 (18.71)	0.032 (0.0030)
Concatenated nDNA						
	ldh-A	50 (42)	191	19 (9.94)	14 (7.32)	0.025 (0.0098)
	RPS7	50 (44)	1158	141 (12.18)	126 (10.88)	0.018 (0.0020)
	X-*src*	50 (45)	518	33 (6.37)	25 (4.83)	0.0095 (0.0022)
	X-*yes*	50 (20)	833	89 (10.68)	53 (6.36)	0.025 (0.0030)
	Glyt	50 (21)	915	30 (3.28)	14 (1.53)	0.0056 (0.0011)
Concatenated mtDNA + nDNA*						
	Concatenated mtDNA	80	1770	661 (37.34)	660 (37.29)	N/A
	mtDNA 1st + 2nd codon positions	80	1180	334 (28.31)	333 (28.22)	N/A
	mtDNA 3rd codon position	80	590	327 (55.42)	327 (55.42)	N/A

This table presents results from DNA sequence analyses in MEGA5 [1]. 'Overall mean *d*' values are estimates of the average evolutionary distance over all sequence pairs, calculated using *p*-distances (base differences per site), with their standard errors (s.e.) shown in parentheses based on 500 bootstrap pseudoreplicates. Sites with <95% site coverage were eliminated prior to *d* calculations. N/A, not available.

*The concatenated nDNA dataset was also analyzed along with these mtDNA in analyses of the 'concatenated mtDNA + nDNA' dataset, and their results are not duplicated here because they are the same as above.

References

1. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28:2731-2739.

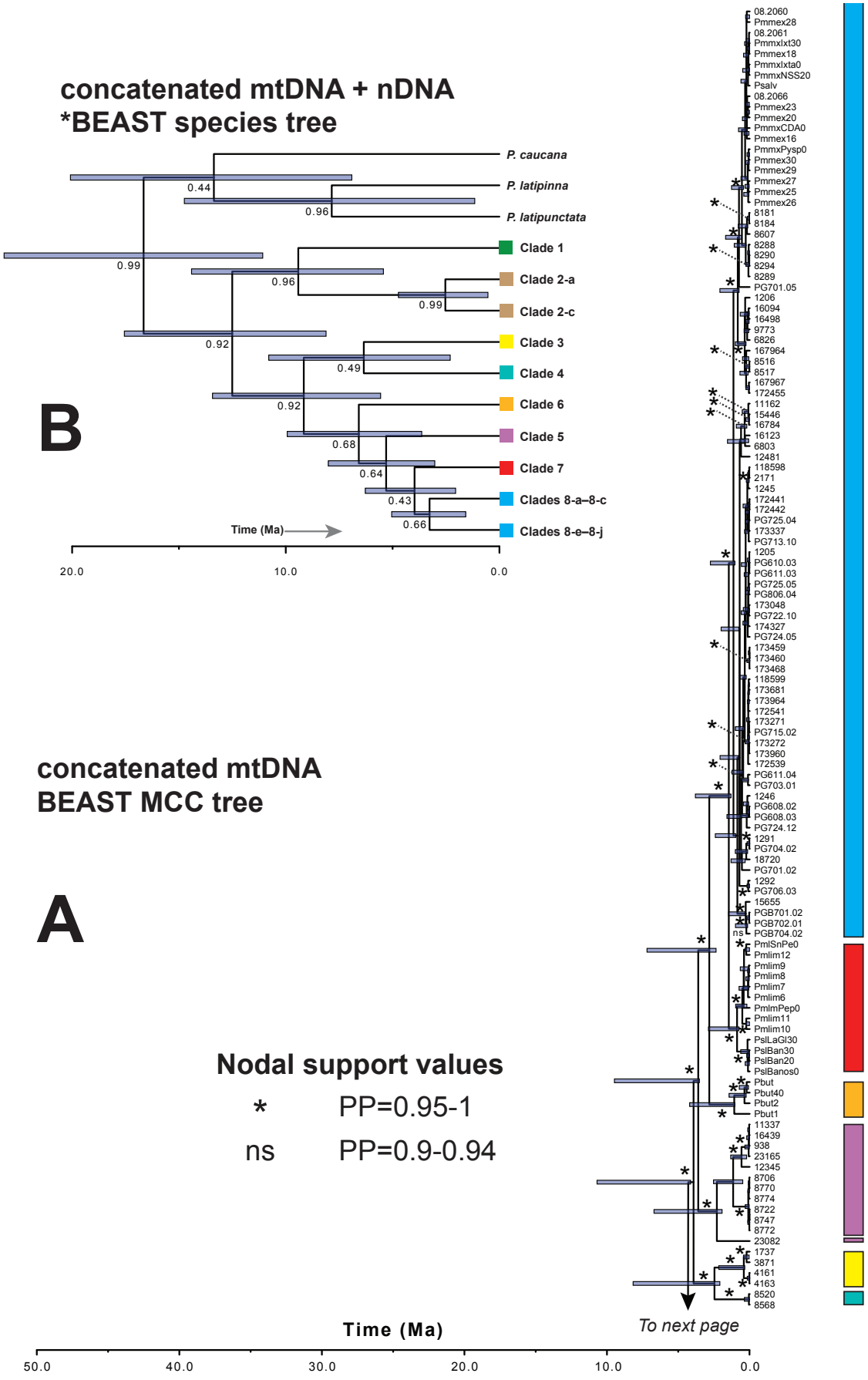
concatenated mtDNA + nDNA
*BEAST species tree
B
P. caucana
P. latipinna
P. latipunctata
Clade 1
Clade 2-a
Clade 2-c
Clade 3
Clade 4
Clade 6
Clade 5
Clade 7
Clades 8-a–8-c
Clades 8-e–8-j
0.44
0.96
0.99
0.96
0.99
0.92
0.49
0.92
0.68
0.64
0.43
0.66
Time (Ma)
20.0
10.0
0.0
concatenated mtDNA
BEAST MCC tree
A
Nodal support values
* PP=0.95-1
ns PP=0.9-0.94
Time (Ma)
50.0
40.0
30.0
20.0
10.0
0.0
To next page

Figure S1

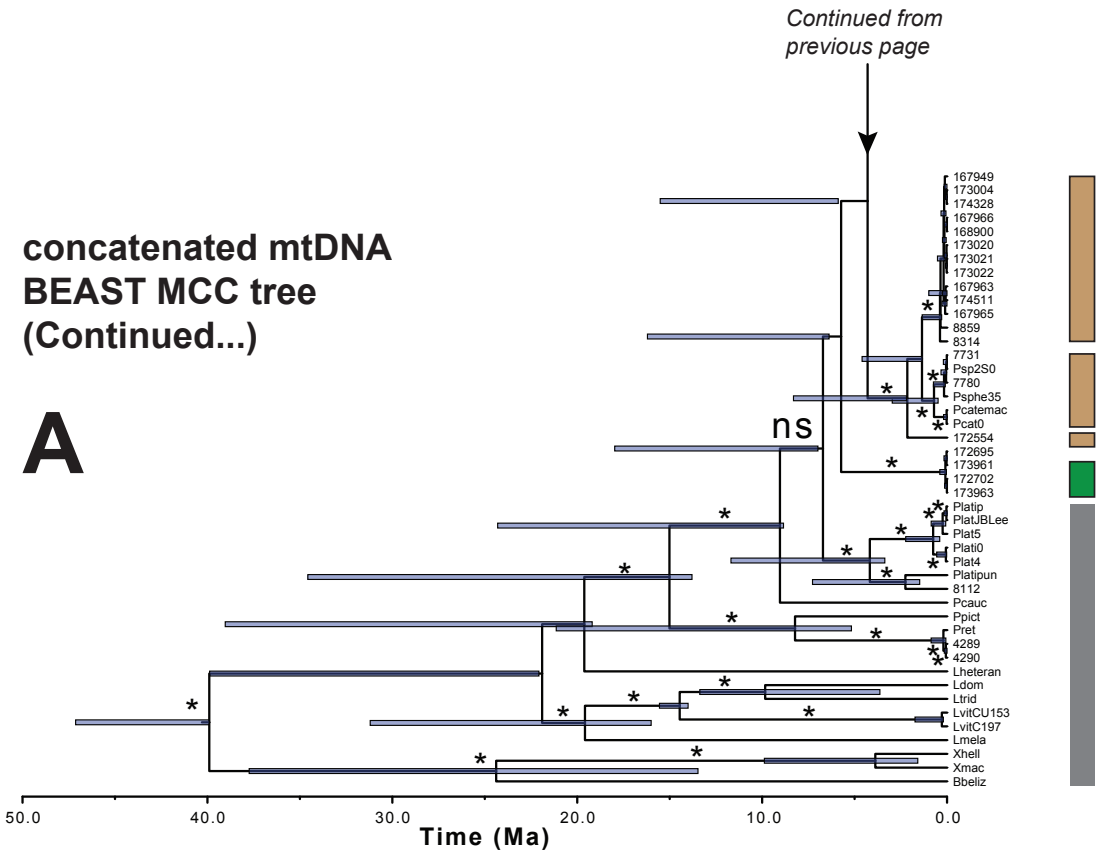

Figure S1 *Continued*.


Idh-A
08.2060
08.2066
11337
173337
PGB70102
938
12345
172455
4161
8290
8181
8520
8607
173964
8706b
167949
7780
8859
172702
08.2060b
08.2066b
172455b
12345b
PGB70102b
173337b
8706
8747b
8747
11337b
938b
118598
118598b
8520b
167949b
8859b
7780b
172702b
173964b
4161b
8290b
8181b
8607b
0.003
RPS7
08.2060
08.2060b
08.2066
08.2066b
4161
4161b
11337
938
12345
11337b
938b
12345b
Pbut
8706
8747b
8747
8706b
172455
172455b
173337
173964
173337b
173964b
PGB70102
PGB70102b
167949
8859
167949b
8859b
7780
7780b
Pcatemac
172702
172702b
8181
8290
8181b
8290b
8607
8607b
8520
8520b
118598
118598b
0.005
X-src
08.2060
08.2060b
11337
938b
938
11337b
12345
12345b
118598
118598b
173337
173337b
173964b
8747
172455
PGB70102
PGB70102b
173964
08.2066
8607
Psalvb
08.2066b
Psalv
8607b
8181
8290
8181b
8290b
8520
8520b
167949
167949b
Pbut
7780
7780b
8859
8859b
172702
172702b
4161
4161b
172455b
8706
8706b
8747b
0.0008
X-yes
08.2060
118598
173964
173337
PGB70102
12345
172455
8706
8747
11337
938
172702
4161
167949
7780
8859
8520
8181
8290
8607
0.003
Glyt
08.2060
08.2066
8747
4161
172702
8706
11337
938
172455
118598
173964
173337
PGB70102
12345
7780
8181
8290
167949
8859
8520
8607
0.0004

Figure S2

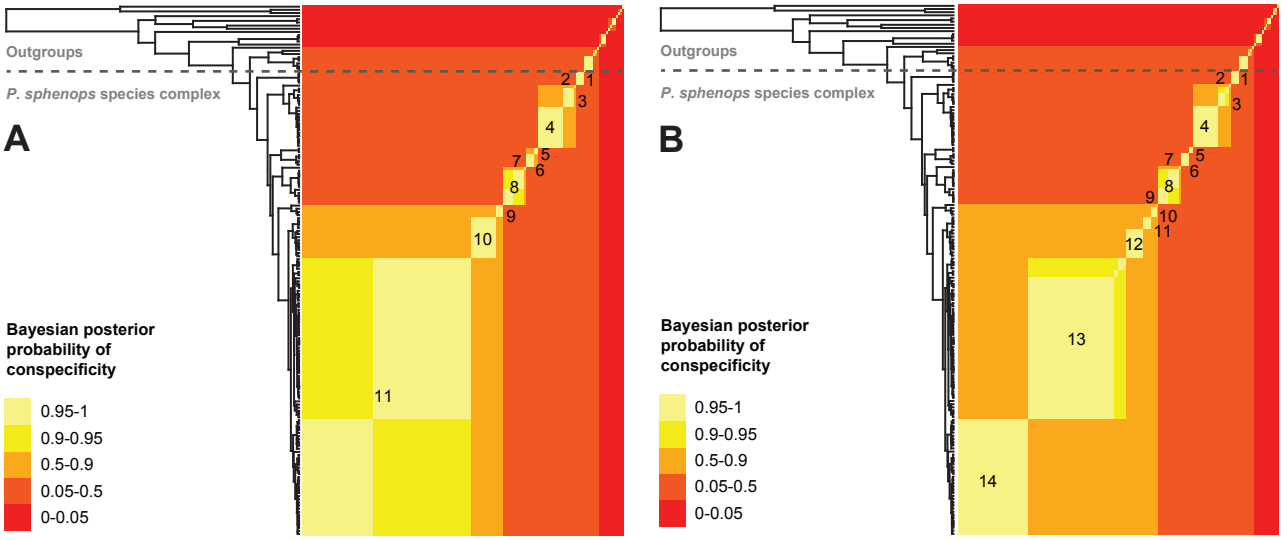
Outgroups
P. sphenops species complex
A
Bayesian posterior probability of conspecificity
0.95-1
0.9-0.95
0.5-0.9
0.05-0.5
0-0.05
1
2
3
4
5
6
7
8
9
10
11
Outgroups
P. sphenops species complex
B
Bayesian posterior probability of conspecificity
0.95-1
0.9-0.95
0.5-0.9
0.05-0.5
0-0.05
1
2
3
4
5
6
7
8
9
10
11
12
13
14

Figure S3

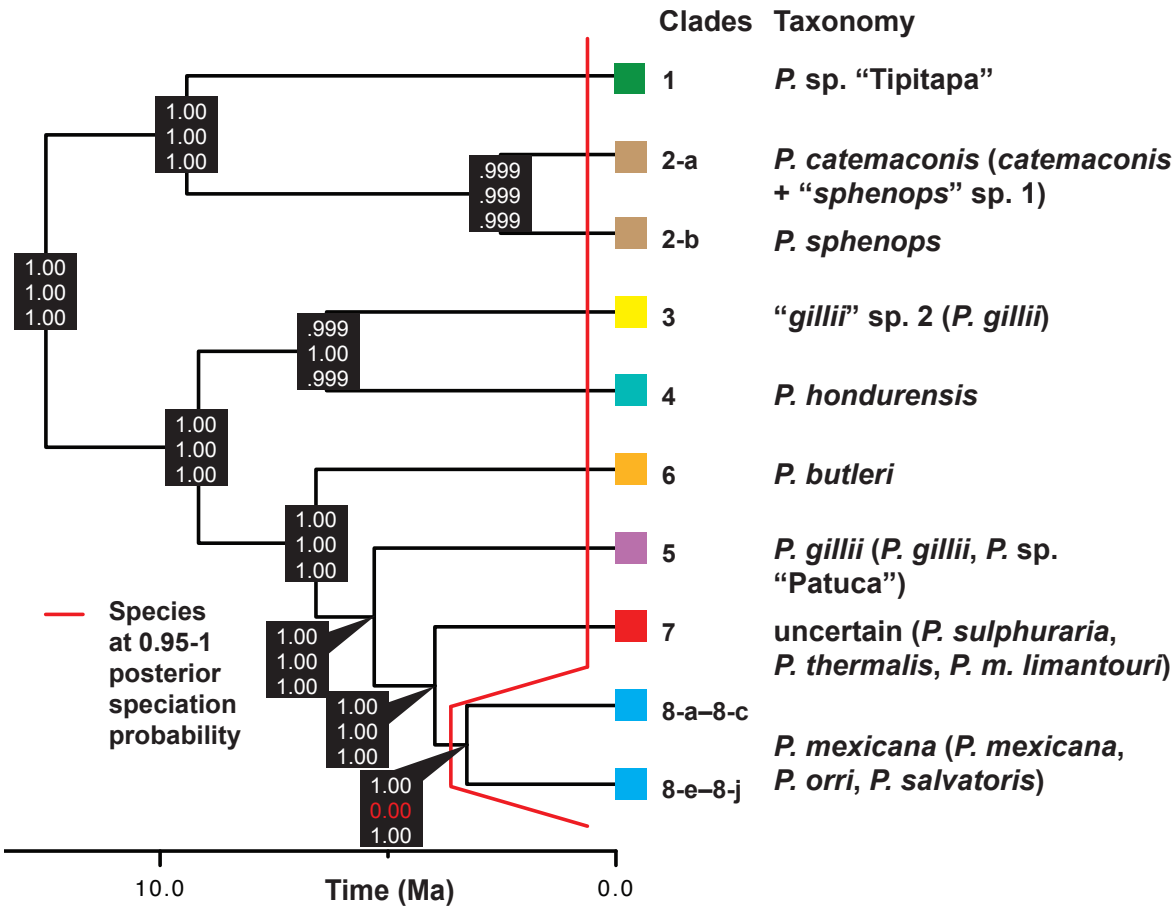
Clades
Taxonomy
1
P. sp. "Tipitapa"
2-a
P. catemaconis (catemaconis + "sphenops" sp. 1)
2-b
P. sphenops
3
"gillii" sp. 2 (P. gillii)
4
P. hondurensis
6
P. butleri
5
P. gillii (P. gillii, P. sp. "Patuca")
7
uncertain (P. sulphuraria, P. thermalis, P. m. limantouri)
8-a–8-c
8-e–8-j
P. mexicana (P. mexicana, P. orri, P. salvatoris)
Species at 0.95-1 posterior speciation probability
1.00
1.00
1.00
.999
.999
.999
1.00
1.00
1.00
.999
1.00
.999
1.00
1.00
1.00
1.00
1.00
1.00
1.00
1.00
1.00
1.00
0.00
1.00
10.0
Time (Ma)
0.0

Figure S4

Chapter 4: Effects of tectonism, sea levels, and drainage rearrangements on Neotropical fish diversification: a comparative phylogeographical test using eight Central American freshwater fish lineages

Effects of tectonism, sea levels, and drainage rearrangements on Neotropical fish diversification: a comparative phylogeographical test using eight Central American freshwater fish lineages

Justin C. Bagley[1*], M. Florencia Breitman[1], and Jerald B. Johnson[1,2]

[1]*Evolutionary Ecology Laboratories, Department of Biology, 4102 LSB, Brigham Young University, Provo, UT 84602, USA*

[2]*Monte L. Bean Life Science Museum, 645 E 1430 N, Provo, UT 84602, USA*

***Correspondence:** Justin C. Bagley, Fax: +1 801 422 0090; E-mail: justin.bagley@byu.edu

Keywords: approximate Bayesian computation (ABC), Central America, comparative phylogeography, drainage basins, ecological niche modeling (ENM), freshwater fishes

Running Title: Central American comparative phylogeography (43 characters w/spaces)

Word count: 13,992 (Title, Keywords, Abstract, and main text; excluding Table/Figure legends and References)

Abstract (250/250 words)

Drainage basins provide the fundamental context for evolutionary and biogeographical processes of diversification in freshwater fishes. However, it remains unclear whether species in the superdiverse Neotropical freshwater fish assemblage responded similarly to historical drainage-controlling processes of tectonism, sea level change, and drainage rearrangements. We used a comparative phylogeographical analysis of eight freshwater fish species/genera codistributed in Central America to test for shared evolutionary responses predicted by four drainage-based hypotheses of Neotropical fish diversification—the 'tectonic vicariance', 'marine vicariance', 'continental shelf width', and 'cross-cordillera exchange' hypotheses. Our approach integrated phylogeographic analyses of 2,091 mitochondrial cytochrome *b* sequences with paleodistribution modeling-based tests (54–231 occurrence records/lineage) for ancestral lineage codistribution, i.e. congruent Pleistocene range dynamics. Our results revealed highly variable spatial-genetic structuring (including taxon-specific patterns) but similar paleodistributional responses to Pleistocene climate change among lineages, with areas of historical range stability and overlap. That such incongruent phylogeographical architectures have arisen despite overlapping ancestral distributions suggests multiple routes to community assembly. Consistent with this, approximate Bayesian computation model averaging supported both simultaneous and asynchronous pulses of diversification across congruent genetic breaks in the upper San Carlos River (two lineages) and Sixaola River (three lineages), where divergences were mostly correlated to Neogene sea levels and continental shelf width. Seven focal lineages also displayed spatially congruent evidence for past drainage connections across the continental divide at the Guanacaste Cordillera. Overall, our results support complex biogeographical patterns illustrating the variable influence of species responses to historical drainage-controlling processes on Neotropical fish diversification.

Introduction

Ecological and evolutionary studies of freshwater fishes have contributed much to understanding adaptive radiation, phenotypic evolution, community assembly, and historical biogeography (Banarescu 1992; Poff & Allan 1995; Bermingham & Martin 1998; Seehausen 2006; Elmer & Meyer 2011). This largely owes to the ecological diversity and island-like nature of hydrological networks, in which fish populations are isolated by marine and terrestrial habitats at multiple spatial scales, creating distinct adaptive and biogeographical patterns (e.g. Seehausen 2006; Banarescu 1992). Particularly because they trace landscape evolution, drainage basins promote biotic diversification while also capturing geological history and providing a context for evolutionary and biogeographical processes to act (Bermingham & Martin 1998; Unmack 2001). Drainages are thus the principal factor shaping fish distributions (Gilbert 1980), and obligate freshwater fishes are highly informative for historical biogeography because they depend on inland drainage connections for dispersal and gene flow (e.g. Smith & Bermingham 2005). Evolutionary studies of freshwater fishes thus allow us to infer not only the historical connections among drainages, but also the evolutionary histories of their inhabitants.

Neotropical North and South America are of great interest for exploring the effects of historical processes on freshwater fish diversification (Lundberg *et al*. 1998; Bermingham & Martin 1998; Hubert & Renno 2006). These areas harbor the greatest diversity of freshwater fish species worldwide (Reis *et al*. 2003; Albert *et al*. 2011) and experienced complex geological changes during the Late Cenozoic (Hoorn *et al*. 2010; Bagley & Johnson 2014a). Given Neotropical areas escaped wide glaciations that affected temperate regions (e.g. Hewitt 2000), three main determinants of drainage area and isolation *other than* ice sheet cover that, with paleoclimatic changes (e.g. Bush *et al*. 1992), have likely critically shaped the distributions and genetic diversity of Neotropical freshwater fishes include: (i) tectonic (mountain-building)

processes at orogenic belts of the Central American volcanic arc (CAVA) and Andes and their effects on back-arc (foreland) basins (e.g. Bussing 1976; Lundberg *et al*. 1998; Hubert & Renno 2006; Ribeiro 2006); (ii) the interplay between drainages and fluctuating eustatic sea levels of the Miocene and Plio–Pleistocene (e.g. Bermingham & Martin 1998; Lovejoy *et al*. 1998; Jones & Johnson 2009); and (iii) drainage rearrangements caused by river reversal and river capture events (e.g. Menezes *et al*. 2008; Albert & Reis 2011). Unfortunately, determining the past effects of these abiotic controls on drainage geometry and fish distributions is difficult in large Neotropical areas. For example, the ancient Cretaceous–Miocene ages and wide areas of the Amazon Basin (6.92 million km^2) and Brazilian Shield (~6 million km^2) have provided ample time and space for the superimposition of multiple shifts in drainages and species ranges due to historical processes, yielding complex fish biogeographical patterns (e.g. Hubert & Renno 2006; Albert & Reis 2011; Albert *et al*. 2011). Studies examining the effects of historical drainage-controlling processes might therefore be more profitable if focused on smaller geological units where the likelihood of rare events is reduced. Insight into causal mechanisms underlying the assembly and diversification of Neotropical freshwater fish communities is also limited because regional-scale comparative phylogeographical perspectives have only recently become available and have evaluated relatively limited subsets of taxa (e.g. Bermingham & Martin 1998; Perdices *et al*. 2005; Reeves & Bermingham 2006). Despite illuminating many cryptic dispersal, vicariance, and range expansion events, these studies have largely focused on broad-scale processes of colonization (but see Bagley & Johnson 2014a,b), rather than testing if specific historical processes predict post-colonization diversification patterns (Jones & Johnson 2009). Thus, while general evolutionary patterns are emerging (Albert *et al*. 2011; Bagley & Johnson 2014a), it remains unclear whether, and how, species in the Neotropical freshwater fish

assemblage have responded similarly to historical drainage-controlling processes.

Central America (CA; Fig. 1) presents excellent opportunities for evolutionary studies of the effects of drainage-controlling processes on Neotropical fish diversification. This is due to its smaller size (533,726 km^2; ~7.7% of Amazon Basin area), recent and complex geological history, high freshwater fish diversity, and continental margins of varying susceptibility to changing sea levels (Coates & Obando 1996; Bermingham & Martin 1998; Smith & Bermingham 2005; Bagley & Johnson 2014a,b). With 525 species and up to 59.2% within-region endemicity, CA displays high freshwater fish alpha diversity, which proportional to area is three times higher than that of the Amazon River (Albert *et al*. 2011; Matamoros *et al*. 2014). Analyses of CA freshwater fishes also have identified areas of endemism defining biogeographic 'provinces' of characteristic fish communities within areas of shared drainage history (e.g. Miller 1966; Myers 1966; Bussing 1976; Smith & Bermingham 2005). The most exhaustive study to date recovered 10 fish biogeographic provinces in CA (Fig. 1d; Matamoros *et al*. 2014); yet, all such regionalization studies infer marked species turnover across drainage divides suggesting historical drainage-controlling processes heavily shaped CA freshwater fish distributions.

In this study, we use a comparative phylogeographical analysis of molecular data from eight freshwater fish species and genera (hereafter, 'lineages') codistributed in CA to test for shared evolutionary responses predicted by four drainage-based hypotheses of Neotropical fish diversification. Using several lineages as replicates in our tests of the hypotheses allows us (i) to avoid pitfalls of single-species biogeographic inference, and (ii) to partially circumvent the issue of coalescent stochasticity influencing single-locus patterns (e.g. Edwards & Beerli 2000). The null expectation is that our focal lineages will, by chance, exhibit multiple responses to historical events (Bermingham & Martin 1998; Sullivan *et al*. 2000; Bagley & Johnson 2014b). An

important but frequently overlooked assumption of comparative phylogeography is that presently coexisting species were codistributed in the past (Carstens & Richards 2007; Marske *et al*. 2012). Thus, we evaluated this assumption in the first step of our analysis by testing for congruent Pleistocene range dynamics and range overlap using paleodistribution modeling (e.g. Carstens & Richards 2007; Waltari *et al*. 2007). The paleodistribution models, in turn, allow us to identify areas of historical habitat stability and thereby to evaluate potential contributions of ancestral distributions to observed phylogeographic patterns. Below, we outline our main hypotheses in the context of the regional geological and biogeographical setting, and we summarize their predicted genetic consequences in Table 1.

(1) Tectonic vicariance hypothesis

All CA lands occur near active plate margins, where tectonic processes of uplift, volcanism, and faulting formed major landforms and modern drainage basins over Miocene to present (Rogers *et al*. 2002; Coates *et al*. 2004; Bagley & Johnson 2014a). The volcanic Chortis highlands of Nuclear CA (Guatemala, Honduras, and Nicaragua) formed and uplifted ~19–3.8 million years ago (Ma; Rogers *et al*. 2002), while most volcanic cordilleras of Lower CA (Costa Rica and Panama) uplifted <7 Ma (Coates & Obando 1996; Marshall *et al*. 2003; Coates *et al*. 2004; Bagley & Johnson 2014a). Subsequently, CA rivers became deeply entrenched in the Plio–Pleistocene. Volcanic eruptions often cause local and downstream extinctions as lava or debris deposit atop streams (McDowall 1996), and this would have caused vicariant isolation in CA freshwater fishes. One of the best-documented examples of this process occurs at the southern Choluteca-Tárcoles province boundary (Fig. 1d), where Pleistocene volcanic flows descended from the CAVA to the Pacific, covering the Tárcoles River (Fig. 1b; Marshall *et al*. 2003). These flows ran through the Tárcoles gorge beside the Herradura block, a high headland ~1600

m above present sea level [asl] that historically restricted fish movement (Lee & Johnson 2009). Subsequent fish isolation would have been maintained by tectonic uplift of the Herradura block, as indicated by the many fish species whose Pacific distributions terminate at the eastern Tárcoles drainage divide within this block (Bussing 1976, 1998; Smith & Bermingham 2005; Matamoros *et al*. 2014). We thus propose a 'tectonic vicariance hypothesis' predicting that tectonism and volcanic flows during the geologic evolution of CA promoted vicariance, allopatric isolation, and speciation in CA freshwater fishes within and between coasts, by isolating drainages on either side of CAVA ranges and causing local extinctions along CAVA and the Tárcoles River.

Sea level hypotheses: (2) marine vicariance and (3) continental shelf dispersal

Eustatic sea-level oscillations occurred in the Neogene then intensified in frequency during Pleistocene glacial cycles <2.6 Ma (e.g. Lambeck *et al*. 2002). These events undeniably impacted Neotropical fresh waters (e.g. Irion 1984; Coates & Obando 1996; McNeill *et al*. 2000; Hoorn *et al*. 2010) and altered drainages in at least two ways relevant to the Neogene–recent diversification of CA freshwater fishes. First, during warm periods of the Late Miocene–Pliocene, eustatic sea levels reached highstands of +25–50 m asl around ~7–5 Ma (possibly multiple peaks) and ~3.5–3 Ma (Haq *et al*. 1987; McNeill *et al*. 2000; Coates *et al*. 2004; Miller *et al*. 2005) that formed marine embayments in the Nicaraguan depression, Tortuguero lowlands, and Motagua Fault Zone until the Pliocene (modeled in Fig. 1a,c; Bussing 1976; Coates & Obando 1996; Bagley & Johnson 2014a,b). Ensuing highstands during Pleistocene interglaciations were much lower (Miller *et al*. 2005) until a +22 m asl event ~450 ka (550–390 ka; Hearty *et al*. 1999), which would have inundated coasts during marine isotope stage 11, extirpating lowland freshwater fishes and reinforcing genetic isolation in mid-upper river

reaches. Indeed, several studies of freshwater fishes and other Neotropical and Caribbean vertebrates support Neogene–recent diversification during marine incursions (e.g. Bermingham & Martin 1998; Perdices *et al*. 2002, 2005; Montoya-Burgos 2003; Jones & Johnson 2009; Barker *et al*. 2012). Thus, we advance a 'marine vicariance hypothesis' proposing that late Plio–Pleistocene sea-level highstands repeatedly caused vicariant isolation of freshwater fish populations in drainages. Second, Pleistocene glaciations lowered sea levels, and a 110–135 m drop is documented for the Last Glacial Maximum (LGM) ~21 ka (Lambeck *et al*. 2002) that would have allowed river braiding and anastomosis over exposed continental shelf, possibly re-connecting isolated fish populations (Unmack *et al*. 2012, 2013; Bagley *et al*. 2013). We thus apply the 'continental shelf width hypothesis' of Unmack *et al*. (2013) to CA, a hypothesis that has never been tested *a priori* in Neotropical freshwater fishes. This hypothesis predicts that drops in sea level during Pleistocene glaciations promoted (i) gene flow causing lower interdrainage genetic divergences in regions with wider continental shelf, and/or (ii) isolation within or between coastal drainages in areas with narrow continental shelf.

(4) Cross-cordillera exchange hypothesis

Drainage rearrangements through river captures (river section displacements to adjacent drainages) and episodic "wet connections" due to swamps/flooding on low drainage divides are two key mechanisms of interdrainage dispersal and vicariance in freshwater fishes (e.g. Bishop 1995; Burridge *et al*. 2008b). In CA, distributional data suggest that dispersals between adjacent drainages in the same versant have occurred through river captures (e.g. Bussing 1976; Smith & Bermingham 2005). Biogeographic, geologic, and genetic lines of evidence also indicate that Neogene–recent drainage connections allowed fish to move across CA cordilleras, between versants. For example, between the Atlantic Mosquitia-San Juan/Bocas and Pacific Choluteca-

Tárcoles provinces (Fig. 1d), 14 of 45 "primarily Atlantic forms" of fishes also inhabit Pacific rivers, resulting in high species covariation across the CAVA (Bussing 1976; Smith & Bermingham 2005). Lower San Juan basin headwater rivers also reversed across the CAVA then were captured by the Tárcoles River and rerouted to the Pacific coast in the Pleistocene (Marshall *et al*. 2003). In addition, mitochondrial DNA (mtDNA) phylogeography has revealed dispersals across the CAVA in catfishes, tetras, livebearers, synbranchid eels, and cichlids; and across the San Blas Range (central Panama) in catfishes, knifefishes, and cichlids (Bermingham & Martin 1998; Perdices *et al*. 2002, 2005; Reeves & Bermingham 2006; Jones & Johnson 2009; McCafferty *et al*. 2012). We thus formalize a 'cross-cordillera exchange hypothesis' predicting that cross-cordillera headwater exchanges caused (i) dispersal and range expansion with insufficient time for cladogenesis, thus low genetic divergences; or (ii) vicariance across drainage divides, with sufficient time and isolation for local evolution of reciprocal monophyly.

Materials and methods

Study area and sampling

The study area spans the CA Isthmus, from the Motagua fault zone of Guatemala, southeast to the Darién isthmus, Panama (<523,000 km^2). To evaluate our hypotheses, we inferred the phylogeographic histories of eight CA freshwater fish lineages from three families (Poeciliidae, Characidae, and Cichlidae) with diverse habitat preferences and feeding ecologies (Table 2). We chose these lineages for study because (i) they are endemic to CA; (ii) their distributions are reasonably well known; and (iii) their ranges overlap in multiple drainages and biogeographic provinces, making them suitable for comparative phylogeography (Bussing 1998; Miller *et al*. 2005; Bagley & Johnson 2014b). Indeed, the focal taxa range from sea level to 1270 m asl (co-occurring mainly 25–540 m asl) across major drainages from Honduras to Panama on the

Atlantic versant, and mainly from central Nicaragua through western Panama along the Pacific versant (Table 2 and Data S1).

We sampled populations of the focal lineages through field expeditions to CA, and from the fish tissue archives of the Monte L. Bean Life Science Museum Fish Collection (MLBM), STRI Neotropical Fish Collection (STRI), and private collections (Data S1). Specimens (whole or tissue samples) were preserved in 95% ethanol or DMSO in the field, and voucher specimens are deposited at MLBM and STRI. In total, our sampling included 2,091 individuals, with n = 93–761 individuals per focal lineage. We provide detailed information on ingroup and outgroup sampling in Data S1 and Appendix S1. We generated new DNA sequence data for six of eight lineages. However, we obtained our final *Astyanax* dataset by augmenting mtDNA from Ornelas-García *et al*. (2008) with new sequences generated from 67 Nicaraguan and Costa Rican specimens. The *Xenophallus umbratilis* (monotypic; hereafter, "*Xenophallus*") dataset mostly comprised sequences from Jones & Johnson (2009), to which we added 11 sequences from a Lake Nicaragua tributary. And we used the *P. mexicana* dataset from our recent analysis of species delimitation in the *P. sphenops* species complex (Bagley *et al*. in revision).

Ecological niche modeling

Testing the assumption that our focal lineages evolved as part of a longstanding assemblage or "evolutionary cohort" that tracked regional environmental changes through time (*sensu* Carstens & Richards 2007), requires being able to predict how species ranges have potentially changed in the past. Fossil data are lacking for our taxa, so we tested this assumption by reconstructing paleodistribution models for each focal lineage using paleoclimatic data and ENMs generated using the maximum entropy model in MaxEnt 3.3.3k (Phillips *et al*. 2006). We based our ENM analyses on global coverages of 19 bioclimatic predictor-variables (Appendix S1) with spatial

resolutions of 30 arc-seconds (1 km^2), downloaded from WorldClim (http://www.worldclim.org/; Hijmans *et al*. 2005). We obtained data layers for present-day climates (1950-2000) and analogous paleoclimatic data layers that we manipulated to have the same spatial resolution. Data layers reconstructing LGM environments ~22 ka were based on the Community Climate System Model (CCSM3; Collins *et al*. 2006; Otto-Bliesner *et al*. 2006a), and Last Interglaciation (LIG) layers for ~140–120 ka were derived from climate simulations of Otto-Bliesner *et al*. (2006b). Prior to MaxEnt analyses, we generated new datasets specific to each lineage by clipping the original three layer-sets to the approximate spatial extent of each lineage, using country-level mask shapefiles developed in ArcMap 10 (Environmental Systems Research Institute, Redlands, CA). The data layers included temperature, precipitation, and seasonality variables. These variables are often highly cross-correlated; in such cases, objectively eliminating redundant variables, or producing new layers capturing variability in multiple correlated predictor-variables, makes model testing and interpretation more straightforward (Elith *et al*. 2011) and can reduce the possibility of model over-fitting (Warren & Seifert 2011). Thus, we attempted to improve our ENM inferences by using ENMtools (Warren *et al*. 2010) to calculate Pearson correlations between pairs of the 19 predictor variables, and then retaining only a single variable when two were correlated at $r > 0.9$, giving preference to variables measuring minima or maxima rather than average environmental conditions (Shepard & Burbrink 2008). This left us with 14 variables that we used in our final MaxEnt analyses (Appendix S1). Our MaxEnt runs drew on 54–231 georeferenced occurrences for each focal lineage (JCB, unpublished data), collated from our field collections and species record searches within the Global Biodiversity Information Facility (http://www.gbif.org/) database. We reprojected ENMs output by MaxEnt onto LGM and LIG layers and interpreted areas with high predicted-

Pleistocene bioclimatic suitability as most-likely areas of past distribution. We evaluated the models using a threshold-independent measure of performance, the area under the Receiver Operating Characteristic curve (AUC) statistic, where scores closer to 1 (maximum) indicate higher predictive ability and AUC > 0.5 indicates better-than-random model prediction (Elith *et al*. 2006). Paleodistribution modeling analyses were analogous to those in Bagley *et al*. (2013), with minor changes above and in Appendix S1 adapting procedures to this study system.

Laboratory methods and sequencing

We extracted whole genomic DNA from tissue samples using Qiagen DNeasy Tissue Kits (QIAGEN Sciences, Maryland) and sequenced the protein-coding mtDNA cytochrome *b* (cyt*b*) gene for every individual using different primers flanking the gene for lineages from different genera or families, and different annealing temperatures, as shown in (Table S1). We purified polymerase chain reaction (PCR)-amplified products using a Montage PCR 96 plate (Millipore, Billerica, MA). Sequences were obtained via cycle sequencing with Big Dye 3.1 dye terminator chemistry using 1/16th reaction size and the manufacturer's protocol (Applied Biosystems, Foster City, CA). We purified sequenced products using Sephadex^{TM} columns (G.E. Healthcare, Piscataway, NJ) and ran them on an automated Applied Biosystems 3730*xl* capillary sequencer. We edited sequences while viewing electropherograms in Sequencher 4.10.1 (Gene Codes Corporation, Ann Arbor, MI). Sequences contained no gaps and were aligned visually in Sequencher, and in Geneious 5.5.7 (Biomatters Ltd., New Zealand). Cyt*b* alignments for each focal lineage were collapsed into unique haplotypes in DnaSP 5.10 (Librado & Rozas 2009).

Gene trees and patterns of spatial phylogeographic congruence

We tested for congruent spatial-genetic patterns among lineages consistent with our hypotheses (Table 1) by evaluating multiple lines of evidence for genetic breaks in the study area. First, we

reconstructed intraspecific gene trees for ingroup cyt*b* haplotypes and outgroup sequences using phylogenetic maximum-likelihood (ML) and Bayesian inference analyses. We performed maximum-likelihood tree searches in GARLI 2.0 (Zwickl 2006). We partitioned the mtDNA data by codon position ({1+2}, 3); assigned each data subset the best-fit nucleotide substitution model (Table S2) selected using the decision-theory algorithm DT-ModSel (Minin *et al*. 2003); and unlinked parameters across data subsets. We evaluated nodal support using 500 ML bootstrap pseudoreplicates, considering nodes with bootstrap proportions ≥70 well supported (Hillis & Bull 1993). We independently estimated gene trees during our Bayesian coalescent-dating analyses, described below. We also generated statistical parsimony haplotype networks for each focal lineage using TCS 1.2.1 (Clement *et al*. 2000). Phylogroups with substantial bootstrap proportion or posterior probability (≥0.95) support in ML and Bayesian analyses, and that formed unique parsimony networks (95% connection limit), were considered 'clades' lending strong support for phylogeographical breaks. Breaks identified in this fashion were compared with local/regional geology and physiography (e.g. Coates & Obando 1996; Marshall 2007) to identify geographical barriers correlated to each break. For breaks that were spatially congruent across multiple lineages, we drew on our previous literature reviews (Bagley & Johnson 2014a,b) to identify correlations with geographically relevant tectonic and sea level events that were potentially causally linked to each break.

We estimated % divergence within each focal lineage and clade, and between clades (i.e. population pairs) split across each shared break identified above. Using MEGA 5.2.2 (Tamura *et al*. 2011), we calculated mean and maximum sequence divergences over all sequence pairs in the corresponding groups using *p*-distances (raw nucleotide differences) and Tamura & Nei (1993) genetic distances, which adequately describe DNA sequence evolution (e.g. Suchard *et al*. 2001).

However, we estimated Tamura-Nei distances corrected for base heterogeneity (d_{MTN}) using a gamma mutation rate distribution with $\alpha = 0.5$.

Coalescent-dating analyses and demographic parameters of community divergence

To assess variation in gene-tree depths and whether intra-lineage diversification coincided with timescales predicted by the hypotheses, we simultaneously estimated the times to the most recent common ancestor (t_{MRCA}), gene trees, and evolutionary parameters for each of the focal lineages using Bayesian coalescent-dating in BEAST 2.1.3 (Bouckaert *et al*. 2014). We linked tree and clock models but partitioned the data into codon position subsets ({1+2}, 3) and unlinked site parameters across subsets. To ensure convergence, we ran three replicate searches on each dataset (MCMC = 2×10^8, sampled every 4000 generations; burn-in = 10%) using relaxed, uncorrelated lognormal (ULN) clock models and birth-death tree priors. Owing to uncertainty of fish substitution rates, we set uniform priors on ULN clock rates spanning published mtDNA rates for teleost fishes ('fish rate' = $0.017–0.14 \times 10^{-8}$ substitutions/site/yr, per-lineage; Waters *et al*. 1999; Burridge *et al*. 2008a). By incorporating sequence data for outgroups, we were able to place multiple fossil or biogeographic calibration points on nodes during analyses of each study lineage; for conciseness, calibration details are given in Appendix S1. We summarized posterior distributions of parameters and ensured that effective sample sizes (ESS) were >200 in Tracer v1.5 (Rambaut & Drummond 2013). We summarized the posterior distribution of trees from each run by calculating a maximum clade credibility tree annotated with node ages from a sample of 5000 random post-burn-in trees in TreeAnnotator 2.1.3 (Bouckaert *et al*. 2014).

To explore the demographic history of populations split across shared breaks identified above, we estimated demographic parameters of population divergence using the "isolation with migration" model in IMa2 (Hey & Nielsen 2004; Hey 2010). We used IMa2 to estimate splitting

times (t) and migration rates (m_1, m_2; unless m priors were set to zero, as described in Appendix S1) between population pairs allowed to vary in size (θ_1, θ_2) under a Bayesian coalescent model. Where applicable, confidence intervals for θ and m parameters were used to help identify priors for tests of simultaneous diversification below. After verifying that priors selected based on three initial runs led to appropriate chain mixing and convergence properties, we made three final runs with 10 MCMC chains that we monitored for convergence after initial burn-ins of 1–5 million steps. Convergence was assumed when update rates reached >10%, ESS scores reached >50 for parameters (Hey & Nielsen 2007), and runs found similar parameter estimates. We converted t estimates to absolute time (T_{div}) using the equation $T_{div} = t/\mu$ (μ = mutation rate per gene per year) and three mutation rates: the "fast" 2% vertebrate mtDNA rate and "slow" 0.9% salmonid mtDNA rate in Bagley & Johnson (2014b), plus a "moderate" 1.57% rate representing the mean of the uniform 'fish rate' prior used in our BEAST analyses. We used our IMa2 results to test whether the timing of lineage diversification across shared breaks fit our hypotheses, and to identify historical events that best fit the observed divergences. We did this because (i) IMa2 allows population sizes and divergence times of the daughter lineages to vary independently, which is more biologically realistic than our constant-size BEAST model; and (ii) IMa2 estimates dates of population divergence, whereas BEAST dates gene divergences that will overestimate the timing of population structure (e.g. Edwards & Beerli 2000). To test whether the timing of population divergences across shared genetic breaks supported the hypotheses, we compared Bayesian posterior T_{div} distributions for each shared break to seven sea level and tectonic events from the literature that we could correlate to those breaks (details in Appendix S1). We rejected events whose dates fell outside the Bayesian 95% highest posterior densities (HPDs; i.e. credible intervals), which had Bayesian conditional probabilities less than 5% ($P <$

0.05), but we failed to reject events whose dates fell in the 95% HPDs ($P > 0.05$).

Testing for simultaneous diversification using ABC model averaging

Understanding whether shared earth history events have impacted genetic divergence within the CA freshwater fish assemblage consistent with our hypotheses (Table 1) requires testing whether congruent spatial-genetic divergences among focal lineages arose synchronously or not, i.e. testing for temporal congruence (Sullivan *et al*. 2000; Donoghue & Moore 2003; Bagley & Johnson 2014b). A shared history of responses to historical events across a geographical barrier is indicated by synchronous genetic divergences in multiple lineages. While one way to test temporal congruence is to directly compare gene divergences (e.g. t_{MRCA}s), this may lead to erroneous biogeographic inferences because variance in coalescent and mutational processes as a function of past effective population sizes (N_e) of different species causes stochastic gene tree patterns (Edwards & Beerli 2000; Riddle & Hafner 2006). Hierarchical approximate Bayesian computation (ABC) models in the msBayes pipeline (Hickerson *et al*. 2006, 2007) address this problem by testing for simultaneous diversification of multiple population-pairs diverged across a phylogeographic break, while accounting for among-population variation at demographic parameters that might otherwise obscure 'true' patterns of community history.

We tested for simultaneous diversification and estimated divergence times of population-pairs diverged at shared phylogeographic breaks in the Bocas province (Fig. 1d; see Results), under a finite sites coalescent model, in MTML-msBayes (Huang *et al*. 2011). The problem of selecting appropriate prior distributions is common in Bayesian inference, and studies have shown that ABC estimation is sensitive to the choice of priors on migration and divergence time parameters (Huang *et al*. 2011, refs. therein). Overly broad priors can also cause hyper-prior ABC samplers like msBayes to run inefficiently, biasing msBayes towards an inference of

simultaneous divergence (Oaks *et al*. 2013). Therefore, we implemented approaches formalized by Huang *et al*. (2011) and Hickerson *et al*. (2014) to overcome this issue: we initially used ABC model choice to compare the posterior probabilities of multiple candidate priors covering parameter space, and then we used ABC model averaging on the candidate priors to estimate the parameters of the MTML-msBayes model while allowing uncertainty in model selection.

We conducted coalescent simulations and tests for simultaneous divergence through a four-step procedure. First, we developed $K = 8$ different prior sets (i.e. model classes), $\{M_1, \ldots, M_8\}$ for ABC model choice and model averaging. Each model class consisted of one of three uniform priors for population divergence times (τ), and one of two uniform priors for the ancestral population size (θ_A) and daughter population size (θ_D) parameters (see Results table). These different models are treated as a set of models specified by a categorical model indicator parameter to be estimated using ABC. Second, we obtained $k = 5 \times 10^6$ random (simulated) samples from each model class specified by a discrete uniform hyper-prior distribution $P(M_K) = 1/8$, with each of the eight models simulating the data with equal probability. Third, we obtained the ABC joint posterior distribution using the default summary statistic vector (D) from MTML-msBayes and rejection sampling to identify the 1000 closest Euclidean distances between the observed summary statistics (D^*) for the data and D_i calculated from 4×10^7 random draws across all eight priors $\{M_1,\ldots, M_8\}$; similar to Hickerson *et al*. (2014) we ran rejection sampling in two steps. This procedure outputs the approximate posterior probabilities $[P(M_K/D)^{1000}]$ of the prior model classes, allowing ABC model choice (Hickerson *et al*. 2014). Thus, fourthly, we compared the approximate posterior probabilities of the model classes to identify the best-supported model (model with highest posterior support), and we used the hyper-posterior probability distributions of Ψ and Ω estimates from independent runs of the best-fit models for

interpretation. Steps 1–3 above also yielded estimates of the number of co-divergence times (Ψ; number of possible assignments of Y taxon-pairs [population-pairs] across Ψ events) and the dispersion index of population divergence times ($\Omega = Var[\tau]/E[\tau]$; the ratio of variance to the mean of the divergence times) that were weighted, by ABC model averaging, on the posterior probability of the eight prior model classes (Huang *et al*. 2011; Hickerson *et al*. 2014).

Following previous studies (e.g. Leaché *et al*. 2007; Bagley & Johnson 2014b), we conducted hypotheses testing by comparing the posterior probabilities for the expected values of the hyper-parameters under a 'null' scenario of asynchronous diversification (H_0: $\Psi > 1$, and $\Omega > 0.05$) against the alternative of simultaneous diversification (H_A: $\Psi = 1$, and $\Omega < 0.05$). We also evaluated support for these hypotheses by comparing B_{10} Bayes factors calculated under the parameter thresholds above while accounting for prior support for the hypotheses, using established criteria for B_{10} "weight of evidence" (Kass & Raftery 1995). During interpretation, we placed our confidence in Ω where Ψ and Ω conflicted, because Ω is more biogeographically relevant and has been shown to outperform Ψ in correctly rejecting simultaneous divergence, even over very recent (≥0.06N) coalescent timescales (Hickerson *et al*. 2014). Ω also correctly rejects simultaneous divergences with large or small sample sizes (Hickerson *et al*. 2007). We estimated community divergence times by converting model-averaged $E[\tau]$ estimates (which are in coalescent units of $4N_{ave}$ generations, where N is mean N_e) to absolute time (T_{div}) using the equation $T_{div} = E[\tau]\times(\theta_{ave}/\mu)$, where μ is the mutation rate per site per generation and θ_{ave} is the midpoint of the upper θ prior, and the three mutation rates used in our IMa2 conversions above.

Results

Ecological niche modeling

Across all of the focal lineages, ENMs of predicted present-day and Pleistocene distributions (i.e.

suitable habitats) produced mean test AUC values >0.75 with nominal standard deviations, indicating significantly better-than-random model predictions and limited variance among independent runs (mean AUC range = 0.783–0.974; standard deviation range = 0.012–0.075; Table S3). Indeed, models performed well in AUC tests, indicating that ENMs and paleodistribution models could significantly discriminate between presence and absence sites. Predicted present-day distributions of the focal lineages (Fig. 2) also provided good fit to the known distributions of focal lineages, including a broad overlap with occurrence datasets we compiled for each lineage and limited false positives mainly in three lineages (Appendix S1). Perhaps expectedly, given the phylogenetic and ecological diversity of the test lineages, contributions of different predictor-variables to the MaxEnt models varied among lineages and among timescales of the paleoenvironments selected for ENM reprojections (LGM versus LIG). Still, temperature seasonality (BIO4), precipitation of driest quarter (BIO17), and maximum temperature of warmest month (BIO5) consistently made the single greatest or second greatest percent predictive contributions to the MaxEnt models for most focal lineages (Appendix S1).

Comparing the present-day ENMs and the predicted Pleistocene paleodistributions of the focal lineages revealed four general patterns. First, suitable habitat areas predicted by the LGM paleodistribution models overlapped substantially with present-day model predictions for all lineages (Figs 1 and 2). Second, for all lineages except *P. annectens*, the area of bioclimatically suitable habitat predicted in the LIG models was much lower than that of present-day ENMs and LGM paleodistribution models (Fig. 2). However, Atlantic-coastal Costa Rica and the west-Pacific coastal areas of Panama were among the only areas that maintained moderate to high bioclimatic suitability during the LIG; and for *A. cultratus*, *Xenophallus*, *R. bouchellei*, and *Amatitlania*, these were the only regions with trace predictions of suitable LIG habitat (Fig. 2a, e,

g, h). The sole exception to this pattern, *P. annectens*, had stable predicted paleodistributions and larger areas of high-predicted habitat suitability (>0.95) during the LIG (Fig. 2f). Third, suitable Pleistocene habitat was only predicted across mountain ranges of the Chorotega volcanic front (Lower CA) in *P. amates* and *P. annectens*. Last, the Tortuguero lowlands, Costa Rica constituted the main area of bioclimatic stability (refugia) hence inferred range stability of the focal lineages through time, being an area of moderate-high model prediction in 92% (22/24) of models. This result suggested that the Tortuguero lowlands have been a stable, environmentally suitable area for most species throughout the late Pleistocene–recent (Fig. 2).

Gene trees and patterns of spatial phylogeographic congruence

Maximum-likelihood gene tree topologies and parsimony networks estimated for each focal lineage are shown in Fig. S2. Bayesian gene tree topologies from independent runs in BEAST were essentially identical to these thus are not presented. Clades reflecting patterns of reciprocal monophyly that were well supported by ML bootstrap proportions, Bayesian posterior probabilities, and parsimony networks are mapped across focal lineages in Fig. 3. These results illustrate highly variable spatial phylogeographical structuring among the focal lineages indicating overall spatial incongruence; however, two general patterns stand out. On one hand, the poeciliids *A. cultratus, P. amates*, *P. annectens*, and *Xenophallus*, and the characid genus *Astyanax* (overall and within *A. nicaraguensis*), contained moderate to deep phylogeographical structuring with 3 to 8 main clades (Fig. 3). These taxa had intraspecific pairwise mtDNA genetic divergences of up to p = ~4%–11% and d_{MTN} = ~4.5%–16% (Table 2 and Appendix S1). On the other hand, a second group of wide-ranging taxa (one from each family) showed remarkably limited population divergences; despite wide ranges crossing 3–10 biogeographic provinces, and ample collections (>100 individuals/each; Table 2), samples of *Amatitlania*, *P.*

mexicana, and *R. bouchellei* (within *Roeboides* spp.) grouped into single well-supported mtDNA clades (Figs 3 and S2) indicating a lack of longstanding genetic barriers between populations. These taxa expectedly also showed more limited genetic differentiation, with genetic divergences of mostly ~1%–3% (Table 2).

Several well-supported phylogeographic divergences across major CAVA mountain ranges and the Tárcoles drainage divide (Fig. 1) were consistent with predictions of the tectonic vicariance hypothesis. One such divergence was the moderately deep genetic split between the *Astyanax* Sixaola-NCA and Lagarto-Puntarenas clades (mean $p = 5.1\%$; mean $d_{MTN} = 6.0\%$), which spanned the CAVA in Costa Rica (Figs 3b and S2). Another vicariance pattern correlated with tectonically uplifted areas was the basal split over a large geographic distance (>400 km) between *P. amates* clade 1 from the León River, Honduras and all other *P. amates* clades, which were sampled from Nicaragua and Costa Rica (Fig. 3f). This ancient split was localized in the Atlantic versant, with genetic divergences of up to ~8.8%–11.4% (Table 2) across the central–eastern Chortis highlands (Fig. 1a). Also, the characid genus *Roeboides* displayed a deep ~6% mtDNA split between *R. bouchellei* and *R. bussingi* that far surpassed their intraspecific divergences (at most 0.13%–1.15%; Table 2) and defined a phylogeographic break correlated to the area between the Tárcoles River and Osa Peninsula, Costa Rica (Fig. 3h). A final break consistent with tectonic vicariance was the finer-scale genetic split between *P. annectens* clade 5 and clades 6 + 7 at Miravalles volcano in the Guanacaste Cordillera, Costa Rica (Fig. 3c).

Considering patterns of spatial-genetic congruence, roughly congruent population divergences associated with isolation in two drainages were shared by multiple focal lineages (Figs 3, 4, and S2). *Priapichthys annectens* and *Xenophallus* had slightly different but overlapping breaks signaling long-term isolation of upland populations in the upper San Carlos

River basin in Costa Rica ("San Carlos River break"; Fig. 4a). In *P. annectens*, the San Carlos River break separated deeply diverged upland clade 4 from clades 5–7 and also involved a prominent, east–west division in the Central Cordillera; however, the much lower elevation of the more derived clade 4 in the Tortuguero basin agrees with marine vicariance predictions. In *Xenophallus*, one portion of this break was also localized in the Central Cordillera, similar to *P. annectens*, but the entire upper San Carlos clade 3 was largely isolated from other clades in surrounding drainages around Nicaraguan depression and Tortuguero lowlands. Support for hypotheses at this break was ambiguous because patterns of reciprocal monophyly of populations agreed with predictions of the tectonic vicariance and marine vicariance hypotheses. A second phylogeographic break shared by *P. amates*, *P. annectens*, and *A. orthodus*/sp. (within *Astyanax* spp.) indicated historical isolation in the Sixaola River basin ("Sixaola River break"; Fig. 4b). This pattern of reciprocal monophyly across the western Sixaola drainage divide in an area with very narrow (~10 km) continental shelf and was replicated across multiple lineages, indicating strong support for the predictions of the continental shelf width hypothesis.

The only pattern of spatial phylogeographical structuring common to all eight focal lineages was that rather genetically homogeneous clades were distributed across highlands and cordilleras at many points along the CA Isthmus (Figs 3 and S2), indicating very strong support for the predictions of the cross-cordillera exchange hypothesis. The most striking result was that clades of seven focal lineages crossed the Guanacaste Cordillera in Costa Rica, with nearly genetically identical populations in both the Pacific Tempisque/Bebedero drainage and the southern portion of the Atlantic-draining San Juan River basin (at least between the Frio and San Carlos Rivers, if not all of the lower San Juan). Clades that shared this pattern included *A. cultratus* clade 1, *P. amates* clade 4, *P. annectens* clade 5, and *Xenophallus* clades 1 and 3, as

well as the single well-supported *Amatitlania* spp., *P. mexicana* (network parts A, L, and M) and *R. bouchellei* clades (Figs 3 and S2). A similar pattern consistent with cross-cordillera exchange across the central-southern Chortis highlands (Fig. 1a) was shared by the *P. mexicana* clade (network part A) and *Astyanax* spp. Lake Managua and Lake Nicaragua clades (Figs 3 and S2). Finally, *P. mexicana* was the sole clade with closely related haplotypes shared across the Panamanian Central Cordillera (network parts I and H) and San Blas Range (parts F and H; Figs 1a, 3, and S2).

Coalescent-dating analyses and demographic parameters of community divergence

The divergence times of our focal lineages (t_{MRCA}s; Table 3) and of population-pairs split across the shared phylogeographic breaks (Table 4 and Fig. S1) revealed considerable variation in the timing of diversification of CA freshwater fish lineages, consistent with an overall inference of temporal congruence. Divergence time results suggested that most lineages diversified since the Miocene to late Pleistocene, coincident with the timing of CA landscape evolution. The estimated t_{MRCA}s from BEAST, which date the deepest splits between all populations sampled within each focal lineage, had geometric mean values in this range, with a maximum of 11.2 Ma in the mid-late Miocene (*P. annectens*) and a minimum of 255 ka in the late Pleistocene (*A. nasutus* lineage of *Astyanax* spp.). Geometric mean t_{MRCA}s for the characid and cichlid genera, which ranged from 11.6 Ma to 7.7 Ma in the Miocene, were slightly higher than the species-level t_{MRCA}s for the five poeciliids (mean t_{MRCA} = 6.3 Ma; Table 3). Coalescent dates of population divergences at shared breaks estimated in IMa2 overlapped, indicating potentially simultaneous diversification. Divergence dates within *P. annectens* and *Xenophallus* across the San Carlos River break were particularly tight, with mean T_{div} of 2.1–1.9 Ma to 4.7–4.3 Ma based on applying different molecular-rate conversions to our IMa2 results (Table 4). The 95% HPDs for

estimates overlapped completely for these two San Carlos River break population-pairs. By contrast, mean divergence times of *A. orthodus*/sp. (within *Astyanax* spp.), *P. amates*, and *P. annectens* across the Sixaola River were more dispersed, ranging from T_{div} = 3.6–0.3 Ma up to 8.1–0.7 Ma in the IMa2 analyses (Table 4). The 95% HPDs for the *A. orthodus*/sp. and *P. amates* estimates overlapped with one another but not with those of the more ancient, Miocene–Pliocene *P. annectens* population splitting time (Table 4).

In our tests for correspondence between population divergence times and the timing of diversification predicted by the hypotheses, Bayesian posterior distributions of divergence times (T_{div}) across the shared breaks almost exclusively supported marine vicariance. For the San Carlos River break, we rejected late Pleistocene volcanism in the Central Cordillera (event iv, <800–300 ka) and ancient Miocene volcanism in the Sarapiquí Arc (event vi, 18–11.4 Ma), which our results showed respectively to post-date and pre-date the origin of population structuring. By contrast, San Carlos population divergence times correlated well with the mid-Pliocene sea-level highstand ~3.5–3 Ma (event ii), which fell in the Bayesian 95% HPDs (Table 4, Fig. 4A). Despite seemingly poorer correlation, however, we also failed to reject vicariance due to Pliocene volcanic activity ~5–4 Ma in the nearby Grifo Alto formation (event vi), based on the two "slow" mtDNA rates, but not the 2% rate (Table 4, Fig. 4A). For the Sixaola River break, we found that population divergence times were also correlated with sea-level highstands. In *A. orthodus*/sp. and *P. amates*, we rejected the mostly Pliocene uplift of the Talamanca Cordillera (event vii) as a potential explanation for this break, but we failed to reject the late-Pleistocene sea-level highstand of +22 m asl (event i) (Table 4, Fig. 4A). However, we failed to reject the late Miocene–Pliocene high-sea frequencies (event iii) or the Talamanca Cordillera event (event vii) using Sixaola population divergence time estimates for *P. annectens*, although

the sea-level event generally covered more of the T_{div} posterior densities than the latter.

Testing for simultaneous diversification using ABC model averaging

The results of our hierarchical ABC analyses in MTML-msBayes are shown in Figs 5 and S3 and Table 5, except the model-averaged divergence time estimates (community $E[\tau]$ and T_{div}), which were comparable to those of IMa2 but slightly younger, are given in Table 4. Through ABC model choice based on comparing approximate posterior probabilities of eight prior models ran for each shared break, we found that M_1 provided the best fit to the San Carlos River break data, while M_7 provided the best fit to the Sixaola River break data (Table 5). Posterior probabilities of the best-fit models ranged only 0.21–0.22, indicating low to moderate posterior support. However, this was attributable to a 'dilution effect', with prior model classes with slightly different distributions performing nearly equally well, only not as well as the best models; taking this into account, posterior support was higher for groups of models. In the San Carlos River break analysis, there was moderate support for related models M_1–M_5 with $P(M_1–M_5|D) = 0.84$, and when we analyzed population-pairs split across the Sixaola River break there was higher support for related models M_1–M_4 ($P(M_1–M_4|D) = 0.57$). Nevertheless, consistent with the overlapping population divergence time estimates for the corresponding lineages, results of best-fit models selected by ABC model choice mostly agreed with a single divergence event, rather than asynchronous divergences, across the San Carlos River break (Fig. S3 and Table 5). Point estimates of the Ψ and Ω hyper-parameters supported simultaneous diversification, with posterior modal $\Psi = 1$ and posterior Ω spiking near zero with improbable tail values extending out from the origin (yielding a strong modal signal of simultaneous divergences, but seemingly unreasonable mean estimates). The posterior of Ω for the San Carlos River population-pairs derived from linear regression contained zero in its 95% HPDs and supported simultaneous

diversification, albeit weakly, based on posterior probabilities and Bayes factors (San Carlos M_1 $P(\Omega < 0.05|D) = 0.73$, and $B_{10} = 2.75$ for $\Omega < 0.05$ versus $\Omega > 0.05$) (Table 5; Appendix S1). In contrast to the San Carlos results, the best-fit Sixaola River break model provided strong evidence overall for asynchronous divergences (e.g. mean $\Psi > 1$, mean and modal $\Omega > 0.05$), despite that posterior modal Ψ was still one and the posterior of Ω contained zero in their 95% HPDs. Indeed, while there was no support for simultaneous diversification across this break based on the Ω posterior derived from linear regression (Sixaola M_7 $P(\Omega < 0.05|D) = 0.19$, and $B_{10} = 0.23$ for $\Omega < 0.05$ versus $\Omega > 0.05$) there was strong posterior support for asynchronous divergences (Sixaola M_7 $P(\Omega > 0.05|D) = 0.81$, and $B_{10} = 4.35$ for $\Omega > 0.05$ versus $\Omega < 0.05$) (Table 5; Appendix S1). In both cases, Ψ posteriors derived from polychotomous regression provided more ambiguous or conflicting support for either hypothesis than modal Ω results (e.g. San Carlos M_1 $P(\Psi = 1|D) = 0.58$; Sixaola M_7 $P(\Psi = 1|D) = 0.70$, and $B_{10} = 4.69$ for $\Psi = 1$ versus $\Psi > 1$).

The model-averaged estimates of hyper-parameter values in the San Carlos River break model were similar to those of the best-fit models (Table 5). However, the main discordance between ABC model choice and model-averaging results was that model-averaged Ψ and Ω estimates for the Sixaola River break were consistent with two divergence events (e.g. posterior modal $\Psi = 2$, $\Omega >> 0.05$) whereas several aspects of the Ψ estimates from ABC model choice supported simultaneous diversification at this break (Fig. 5; Table 5). We relied on Bayes factor support for the Ω signal to resolve this conflict (and minor conflicts among San Carlos River break results above) and guide interpretation, due to properties of Ω and Ψ (see Materials and methods) and because posterior probabilities alone are often sensitive to prior choice. Therefore, we reject simultaneous diversification based on very weak Ω Bayes factor support ($B_{10} = 0.12$ for

$\Omega < 0.05$ versus $\Omega > 0.05$), but Bayes factors very strong support for the null hypothesis of asynchronous diversification across the Sixaola River break ($B_{10} = 8.01$ for $\Omega > 0.05$ versus $\Omega < 0.05$).

Discussion

Traditional historical biogeography and phylogeography studies often use descriptive approaches formulating *ad hoc* scenarios to explain observed biogeographical patterns, only after matching modern species distributions or genetic breaks to geographic barriers (e.g. reviewed by Avise 2000). However, such 'pattern-matching' approaches can lead to erroneous phylogeographical inferences, because locus- or species-specific variance in genetic processes can obscure 'true' biogeographic history (e.g. Edwards & Beerli 2000; Riddle & Hafner 2006). Fortunately, the statistical rigor of phylogeography has increased tremendously over the last decade, sparking a paradigm shift towards developing and testing realistic demographic models to explain genetic variation and elucidate scenarios leading to the observed patterns based on statistically discriminating among alternative hypotheses, or "statistical phylogeography" (Knowles 2009). Of particular import was the development of full-likelihood ML and Bayesian methods and "likelihood-free" ABC methods for estimating and selecting among single- or multi-taxon vicariance and dispersal models, while accounting for confounding effects of coalescent and mutational stochasticity (e.g. Hey & Nielsen 2004; Hey 2010; Hickerson *et al*. 2006, 2007, 2014). Recent studies also show that comparative phylogeography provides a framework for integrating various lines of evidence from geology, spatial population structure, coalescent-based statistical models, paleoclimatology, and paleodistribution modeling, to infer the otherwise cryptic responses of whole assemblages of species to geographical barriers and historical events

(Carstens & Richards 2007; Waltari *et al*. 2007; Burridge *et al*. 2008b; Huang *et al*. 2011; Marske *et al*. 2012; Bagley *et al*. 2014b).

In this study, we used comparative phylogeography to test explicit spatial, phylogenetic, and temporal predictions of four *a priori* hypotheses of the influence of drainage-controlling processes on Neotropical fish diversification in Central America (Table 1), which we outlined based on geological and biogeographical data. We tested for spatial congruence of ancestral distributions and phylogeographic patterns, and then used novel ABC model-averaging techniques (Huang *et al*. 2011; Hickerson *et al*. 2014) to test for simultaneous diversification of multiple lineages at shared breaks consistent with the hypotheses. While various studies have inferred the biogeographical history of CA taxa using comparative phylogeography (e.g. Bermingham & Martin 1998; Perdices *et al*. 2005; Reeves & Bermingham 2006; Bagley & Johnson 2014b), ours includes the broadest spatial/numerical and taxonomic sampling of any such study to date and is the first to use ABC model averaging to test for simultaneous diversification during the assembly of CA biotas. By revealing congruent phylogeographical divergences in space and time among multiple fish lineages, our results provide strong evidence that historical drainage-controlling processes consistent with the marine vicariance hypothesis and the cross-cordillera exchange hypotheses have broadly imprinted upon genetic structuring in the CA freshwater fish assemblage. However, we infer that these patterns have arisen within an overarching model of spatially incongruent diversification involving multiple routes to community assembly, and that species have responded differently or at different times to tectonic processes and constraints imposed by continental shelf width. Here, we discuss several key ways that our results support or reject different aspects of our hypotheses and shed new light on the

role of historical drainage-controlling processes in shaping freshwater fish diversification in the CA Neotropics.

(1) Idiosyncratic effects of tectonic vicariance on Central American freshwater fish diversification

The geological record of CA documents dynamic effects of Neogene–Quaternary tectonic uplift, CA volcanic arc activity, and coastward volcanic fallout and debris flows on the evolution of regional landscapes and drainages (e.g. Rogers *et al*. 2002; Marshall *et al*. 2003; Coates *et al*. 2004). Based on this record and previous biogeography studies, we outlined a tectonic vicariance hypothesis (see Introduction, Table 1) predicting CA freshwater fish populations have been genetically sundered within or across the volcanoes and cordilleras of the CAVA (in areas up to ~3400 m asl) and the eastern Tárcoles drainage divide, yielding distinct lineages on either side of these barriers. Our phylogeographic analyses revealed genetic signatures that were broadly consistent with the geographical mode and tempo of diversification predicted by this hypothesis in the livebearing fish *P. amates* and the two characid genera, *Astyanax* spp. and *Roeboides* spp. The deepest divergence in *P. amates* and the *Astyanax* Sixaola-NCA and Lagarto-Puntarenas divergence roughly correlated to mountainous CAVA areas in the Chortis highlands and Chorotega volcanic front, whereas the split between the two *Roeboides* species was associated in part with the Tárcoles River area (Figs 3 and S2). Although we could not map the latter split with high spatial resolution, these breaks point to associations between morphotectonic features and phylogeographical structure that, when combined with their divergence time estimates, suggest vicariance due to Miocene–Pleistocene tectonic processes has caused or maintained population divergence. Indeed, comparing the t_{MRCA}s across these breaks with the geological dates of related rock formations shows that the fish-genetic and rock records

overlap: geological dates for CAVA and lava and debris fans overlying the Tárcoles River fall within the t_{MRCA} Bayesian credible intervals in every case, albeit posterior densities were much wider for the *Roeboides* split, likely owing to our use of a single-locus marker and the very low numerical sampling for the eastern *R. bussingi* clade (Tables 2 and 4 and Appendix S1).

Nonetheless, while the above genetic breaks support potential tectonic-vicariance events, they are idiosyncratic to different lineages, geographical barriers, and biogeographic province boundaries (Figs 1d and 3). Thus, the consensus from these complex patterns is that they provide limited support for the tectonic vicariance hypothesis, because they reflect incongruent or "pseudoincongruent" histories (Cunningham & Collins 1994; Donoghue & Moore 2003; although we could not rigorously test for temporal incongruence). Instead of shared biogeographic histories of dispersal and vicariance in response to the same historical process, the most parsimonious interpretation of these findings is that, consistent with the null hypothesis of our study, these taxa experienced multiple, lineage-specific responses to different historical processes (Bermingham & Martin 1998; Sullivan *et al*. 2000; Bagley & Johnson 2014b). Such patterns most likely have arisen due to different timing of dispersal across these barriers, possibly linked to the varying ecological attributes (e.g. colonization abilities) of these lineages (Table 2; e.g. Bermingham & Martin 1998; Bagley & Johnson 2014b). That said, our most concentrated sampling focused on the southern Nicaraguan depression and surrounding highlands of southern Nicaragua and Costa Rica, and this may have biased our study towards inferring that the above patterns were lineage-specific. Future analyses comparing more fish species from the same communities could potentially recover congruent divergences providing an improved basis for evaluating the timing of vicariance or gene flow across these barriers, and their correspondence to geological events.

(2) Congruent spatial and temporal divergences supporting marine vicariance

The chronology of fluctuating eustatic sea levels (Haq *et al*. 1987; Lambeck *et al*. 2002; Miller *et al*. 2005) strongly suggests the configurations of CA drainage basins have been dynamic over Neogene–recent, and that sea levels have figured importantly in shaping CA freshwater fish demographic histories (Bermingham & Martin 1998; Smith & Bermingham 2005; Jones & Johnson 2009; Bagley & Johnson 2014a,b). Under the marine vicariance hypothesis, we predicted that marine highstands of the Plio–Pleistocene promoted vicariance causing freshwater fish populations to be genetically diverged in mid-upper river reaches, which is supported by geological studies, eustatic records, and digital elevation models (Fig. 1c) suggesting that CA lowlands were widely inundated by marine incursions of these periods (e.g. McNeill *et al*. 2000). The results of our gene tree and parsimony network tests for spatial congruence, Bayesian coalescent divergence time analyses, and ABC tests for temporal congruence overwhelmingly support all of the predictions of the marine vicariance hypothesis in the livebearers *P. annectens* and *Xenophallus*. Spatially congruent divergences separate the populations in the upper reaches of the San Carlos River from those at lower elevations in surrounding basins (Figs 3, 4A, and S2), suggesting a shared history of isolation in upland tributaries. These tributaries also appear to have served as refugia, given the deeply diverged upland clades are ancestral to at least two or three more shallowly coalescing clades, from which we infer they diversified since ~4.5–2 Ma in the mid-Pliocene to early Pleistocene (Table 4) by a single pulse of diversification (Table 5). Moreover, full-Bayesian and ABC model-averaged community divergence time estimates also rule out other events but show that the timing of divergence at this break coincides with marine incursions during the mid-Pliocene highstand ~3.5–3 Ma (Table 4), which we predicted *a priori* as a potential cause of vicariance (Table 1), and occurred through a single pulse of diversification (Fig. 4A and Table 5).

It is important, however, to critically evaluate our confidence in whether the mid-Pliocene sea-level highstand could actually have caused the observed genetic divergences in the San Carlos River. In doing so, we find that this event is both strongly supported by our genetic results as well as multiple biogeographic and geo-climatic studies. Previous comparative phylogeographical analyses of three CA freshwater fish lineages in Panama revealed deep divergences within the genera *Roeboides* and *Hypostomus* dated to the same mid-Pliocene highstand mentioned above, suggesting that Lower CA lands were emergent and had been colonized by freshwater fishes prior to this event (Bermingham & Martin 1998). Also, eustatic curves based on geological and proxy data indicate that the mid-Pliocene sea-level event was the last substantial +50 m asl highstand of the Neogene, and was followed by mostly modern or lower-than-modern sea levels during mid-late Pleistocene glacials until the very most recent well-supported marine highstand ~450 ka buoyed sea levels to ~+22 m asl (Haq *et al*. 1987; Hearty *et al*. 1999; Miller *et al*. 2005). Thus, the mid-Pliocene highstand event may have been the most recent wide marine incursion to affect patterns of genetic diversity in CA freshwater fishes before the low sea stands of the Pleistocene. Stratigraphic data from mid-Pliocene deposits of the Buenos Aires Reef and Quebrada Chocolate formations at Limón headland, Costa Rica (McNeill *et al*. 2000) also show that sea levels did, in fact, inundate lowlands in CA back-arc basins at that time and establish good correlations with the eustatic curves, especially that of Haq *et al*. (1987). Last, high-resolution topography data allow approximating the effects of sea level rise, with the implication being that past highstand events would have had relatively greater effects in tectonically uplifting areas such as CA. From a digital elevation model, we infer that isolation in the upper San Carlos was highly physically plausible, because substantial Nicaraguan depression and Tortuguero-lowland areas (except the Sarapiquí Arc, ~200 m asl) would be

inundated by +50 m sea-level rise today (collapse of East Antarctic ice sheet; Fig. 1c), and Pliocene CA lands were *lower* and likely did not possess marked relief as seen in the region today (Coates & Obando 1996; Coates *et al*. 2004; Marshall 2007).

The only patterns among our results disagreeing with the above inferences are that (i) the spatial configuration of the San Carlos break could also be interpreted as consistent with tectonic vicariance associated with volcanism in the Central Cordillera; and (ii) that, when using the two slower mtDNA rates, we could not statistically reject the possibility that one or both species diverged in the San Carlos River coincident with Pliocene volcanic eruptions ~5–4 Ma of the Grifo Alto formation (Marshall *et al*. 2003). The former issue seems irrelevant because our IMa2 results reject diversification linked to Central Cordillera volcanism <800-300 ka, which is younger than the boundary conditions set by the lower credible intervals of our population divergence time estimates. The latter issue is significant, however, as geological reconstructions show that the proto-San Carlos–Chirripó Rivers drained the Grifo Alto formation of the Aguacate Cordillera, which served as the CAVA continental divide in central Costa Rica until the Plio–Pleistocene (Marshall *et al*. 2003). However, Pliocene volcanism in headwaters of this ancient river system cannot explain the pattern of upland isolation observed in modern fish populations, because volcanic eruptions more likely caused these headwater fish populations to go extinct. Also, the CAVA continental divide migrated from the Aguacate range to its present position above the Costa Rican Central Valley in the early-middle Pleistocene, "ponding" the flow of the proto-Chirripó and nearby drainages (indicated by interbedded lacustrine silts in the Central Valley; Marshall *et al*. 2003) such that the former Aguacate headwaters of these rivers remained isolated on the Pacific versant. Any ancient Aguacate headwater populations would then have been destroyed by subsequent volcanic flows in the Central Valley and Tárcoles gorge

and thus could not have given rise to modern populations in the upper San Carlos River.

(3) Multiple divergences in drainages associated with narrow continental margins

In contrast to the marine vicariance hypothesis, applying the continental shelf width hypothesis to CA led us to predict that low seas during Pleistocene glaciations (i) allowed riverine connections promoting fish dispersal and gene flow over the continental shelf, which led to lower interdrainage genetic divergences in areas with wide continental shelf (Table 1; Bagley *et al*. 2013; Unmack *et al*. 2013). We also expected the opposite pattern, that (ii) populations would exhibit divergences across drainage divides in areas with very narrow continental margins, which should maintain population isolation despite lowered sea levels (Table 1; Unmack *et al*. 2013). We could not rigorously test for wide-shelf dispersal (i above), because our sampling (Fig. 3) was limited in the areas of CA with wide continental shelf (Fig. 1c): the Mosquito Coast (maximum width ~330 m), El Salvador–Nicaragua Pacific coast (maximum width ~75 m), and Gulf of Panama (maximum width ~124 m). Still, one clade per focal lineage was represented in each El Salvador–Nicaragua Pacific coast drainage, except multiple *Astyanax* spp. clades occurred in sympatry in the Negro River, Gulf of Fonseca and Ciruelas River, Costa Rica. From this, we hypothesized that recent dispersal or complete mitochondrial replacement of pre-existing lineages has occurred in these drainages (e.g. Bermingham & Martin 1998); however, more studies are needed to test this hypothesis.

Regarding the narrow-shelf prediction (ii above), our phylogenetic results revealed that *P. amates*, *P. annectens*, and *A. orthodus*/sp. displayed congruent spatial patterns of reciprocally monophyletic populations diverged across the western Sixaola River drainage divide. In this region, the continental shelf is only ~10 km wide (defining shelf margin at the –135 m bathymetric contour), making it the narrowest 'true' continental margin in CA. This finding

lends very strong support to the continental shelf width hypothesis and suggests that the lower reaches of the Sixaola River did not anastomose with adjacent rivers during the Pleistocene because opportunities for differential erosion of drainage divides or lateral flows between river basins were highly restricted by the narrow continental shelf (Unmack 2001). Instead, stable drainage geometry has apparently sustained population isolation, allowing genetic drift rather than gene flow to dominate fish demographic history in the Sixaola basin.

While the inferred co-divergences in the Sixaola River highlight a shared spatial pattern of drainage area history, it is critical to statistically assess temporal congruence at multi-taxon breaks to identify potential underlying causal mechanisms (Sullivan *et al*. 2000; Donoghue & Moore 2003; Leaché *et al*. 2007; Bagley & Johnson 2014b). Rigorous tests of temporal congruence that statistically account for variance in mutational and coalescent processes and demographic histories (e.g. differences in N_e) among lineages are needed, because it is not possible to distinguish among single or multiple divergences by examining gene divergences (t_{MRCA}s) alone (Edwards & Beerli 2000; Hickerson *et al*. 2006; Knowles 2009). We addressed this issue by testing for simultaneous diversification in the Sixaola River using recently developed tools for comparative phylogeographical modeling in the msBayes bioinformatics pipeline (Hickerson *et al*. 2006, 2007; Huang *et al*. 2011). We also used the latest methods accounting for uncertainty in prior selection, and we set our priors based on empirical estimates for population size and divergence time parameters from independent coalescent analyses in IMa2; thus, our ABC results unlikely reflect downward bias in the hyper-parameters due to Lindley's paradox, which may occur if msBayes priors are overly broad (Oaks *et al*. 2013).

Most results from ABC model choice and model averaging over candidate priors weighted by their posterior probabilities (Huang *et al*. 2011; Hickerson *et al*. 2014) supported the

hypothesis that fish populations became isolated across the Sixaola drainage divide during multiple pulses of diversification (Table 5; Fig. 5). Indeed, all ABC results for Ω (=$Var[\tau]/E[\tau]$) provided substantial to very strong Bayes factor support (Kass & Raftery 1995) for asynchronous divergence at this break (i.e. *two* divergence events, non-zero Ω). In addition, posterior estimates of Ψ (number of divergence events hyper-parameter) from ABC model averaging also supported asynchronous diversification. In contrast to this, best-fit model estimates of Ψ supported a *single* divergence event at the Sixaola break, and thus conflicted with our other findings (Fig. S3). That different values of hyper-parameter Ψ supported different conclusions complicates interpretation of these results. However, there are good statistical reasons to favor the results of Bayesian model averaging over any single-model inference, even in the case of the best-fit models identified in Table 5; for example, Bayesian model averaging can reduce mean squared error of parameter estimates and produce more meaningful results even if all model classes are misspecified (Hoeting *et al*. 1999). Moreover, we follow recommendations in Hickerson *et al*. (2014) to rely on the Ω signal, rather than that of Ψ, because Ω outperforms Ψ with "near-zero" error probabilities when maximum divergence time (τ_{max}) priors are ≥0.06N (156,250 generations). Indeed, Ψ only appears more likely to yield the correct interpretation when multiple divergences have occurred in response to different historical events that occurred close together and very recently, i.e. that were of LGM to Holocene age (Hickerson *et al*. 2014). And this situation seems not to apply to the Sixaola River break, where mean model-averaged community divergence times, $E[\tau]$, date to the Plio–Pleistocene (~1N–2N; Table 4; Fig. 5). Thus, the most reasonable interpretation of our ABC model averaging results is that they support multiple pulses of population divergence across the Sixaola drainage divide. However, assuming that our model averaging results are effectively "correct", then our results suggest that sole

reliance on one set of priors, or the single best-fit model prior in ABC analysis can lead to more conflicting or error-prone results. In this light, our results agree with what Hickerson *et al*. (2014) deem the "cautionary tale about the importance of model prior specification in ABC methods": our study and theirs suggest that ABC model averaging can improve inference and overcome insufficiency of hyper-prior samplers of high-dimensional data.

One final point of note is that the above msBayes outcome agrees very well with the IMa2 divergence time estimates, the latter of which seemed to suggest two clusters of divergence times at the Sixaola River break (Table 4). In particular, our IMa2 results for the *P. annectens* Sixaola break were the only divergence times that were consistent with Talamanca Cordillera uplift ~5.5–3.5 Ma and that had 95% HPDs that did not overlap with other co-diverged taxa. Thus, inferring multiple pulses of diversification in msBayes might be said to have been expected; however, we reiterate that the only way to arrive at rigorous inferences of temporal congruence or incongruence is through statistically estimating divergence times and their variances using statistical tests such as those offered in msBayes. Indeed, while single-taxon isolation with migration models assume a single divergence event, the "borrowing strength" of hierarchical ABC tests of simultaneous versus asynchronous divergence hypotheses in msBayes allows more of the information content of the data to be used than analyzing each population-pair separately in IMa2 (Hickerson *et al*. 2006, 2007). ABC estimates of divergence times in msBayes also are also at least as good as IMa2's (albeit IMa2 is less computationally efficient) and produce within-population-pair parameter estimates with lower mean squared error when ran in a multi-taxon model (Hickerson *et al*. 2006). The ability of msBayes to generate accurate estimates of community divergence times $E[\tau]$ (similar to T_{div} from IMa2) during multi-taxon analyses has also been validated by simulation tests of the estimators, which show that msBayes

estimates $E[\tau]$ better when using a single mtDNA gene than with 4–8 loci because of rate heterogeneity among loci; indeed, lower root mean squared error values are only obtained by using ≥16 loci (Huang *et al*. 2011; Hickerson *et al*. 2006; but adding more loci always improves Ω estimation accuracy). We cannot, however, rule out the possibility that our IMa2 or msBayes θ or T_{div} estimates have been influenced by other factors that we did not or could not model using mtDNA alone, including unsampled population structure, recombination, or error in the nucleotide substitution model (both programs are limited to the Hasegawa-Kishino-Yano model, which could potentially bias results). Our estimates of T_{div} and $E[\tau]$, migration rates, and θs might also be improved with the addition of multiple, independent loci; and we consider our study to provide an improved base of samples and mtDNA sequences from which to conduct additional multi-taxon multi-locus analyses in the future.

(4) Congruent patterns of dispersal across the Guanacaste Cordillera

Orogeny and landscape evolution provide opportunities for cross-cordillera dispersal and faunal exchange of freshwater organisms through headwater river capture and river reversal events, with distinct genetic consequences (e.g. Burridge *et al*. 2006, 2008b). Thus, we formalized a 'cross-cordillera exchange hypothesis' predicting cross-cordillera headwater exchanges have caused (i) dispersal/range expansion too recently for cladogenesis, or (ii) vicariance and local evolution of reciprocally monophyly across drainage divides in CA freshwater fishes. The effects of cross-cordillera dispersal on CA freshwater fish populations have been addressed in passing in *ad hoc* discussions of fish phylogeography (e.g. Perdices *et al*. 2002; Reeves & Bermingham 2006) and distributional data (Bussing 1976; Smith & Bermingham 2005), but had never been tested as an *a priori* hypothesis prior to this study. Our phylogeographic analyses revealed genetic signatures consistent with our predictions, with shallowly coalescing clades

distributed across four major CA cordilleras: the central-southern Chortis highlands, Guanacaste Cordillera, and Panamanian Central Cordillera and San Blas Range (Figs 3 and S2). These findings strongly support the cross-cordillera exchange hypothesis, and the low genetic divergences we observe across these ranges are most consistent with prediction i above.

In the most striking inference of cross-cordillera exchange in our study, eight clades in seven focal lineages displayed genetic patterns consistent with recent dispersal(s) across the Guanacaste Cordillera, Costa Rica (Figs 3 and S2). This pattern was previously recovered in a study of *Xenophallus* livebearers by Jones & Johnson (2009), who hypothesized that it resulted from dispersal across the continental divide from the Atlantic Zapote River (San Juan River tributary) to the Pacific Tempisque/Bebedero River drainage, due to headwater river capture caused by volcanic activity in the Guanacaste Cordillera. However, our study is unique from the above, because we extended their *Xenophallus* dataset and combined it with data from seven codistributed fish lineages to examine community-level responses of freshwater fishes to geological history in the Guanacaste Cordillera and other areas; thus we were able to make stronger inferences of multi-taxon dispersal across the continental divide. We agree with Jones & Johnson (2009), however, that these Guanacaste Cordillera results unlikely reflect ancient vicariance or alternative dispersal scenarios (e.g. recent colonization of Atlantic and Pacific versants without crossing the continental divide). Among geomorphological mechanisms of interdrainage fish movement (reviewed by Burridge *et al*. 2008b), our results are best explained by recent headwater river capture events or episodic "wet connections" between proximate headwater rivers. In particular, our findings agree with Bussing's (1976) hypothesis that the distance (~200 m) between Tempisque/Bebedero and Sapoa River headwaters at 400 m asl on the western slope of Orosí–Cacao volcano, and the swamp barrier separating upper Bebedero

(Rio Blanco) and Negro River (a Lake Nicaragua tributary) headwaters at 560 m asl, permitted exchanges of freshwater fish communities between versants. That connections opened intermittently between these rivers across the cordillera seems inevitable; however, freshwater fish movements across the Guanacaste continental divide could have happened through different events or routes, especially given the marked contrast in relief between towering volcanoes (~1400–2000 m asl) and surrounding lowlands (up to 200–400 m asl), with large gaps between the volcanic cones (Marshall 2007) that could have allowed stream captures in either direction. Nevertheless, Guanacaste Cordillera activity dates to the Quaternary, suggesting that if volcanism caused river captures between the lower San Juan and Tempisque/Bebedero basins then the resulting dispersal events should date no earlier than the Quaternary. Given lineages sharing the cross-Guanacaste Cordillera dispersal pattern in our study are generally more widely distributed along the Atlantic versant, we hypothesize that past dispersal between versants in this area most likely proceeded unidirectionally, from the Atlantic to the Pacific coast. Recurrent bi-directional gene flow might also explain the observed pattern of low genetic divergences, but future studies will be required to evaluate these competing migration models in greater detail.

Our phylogeographic results also supported a shared pattern of co-dispersals across the central-southern Chortis highlands in the wide-ranging *P. mexicana* network subclade part A (which spans from Guatemala to Panama) and *Astyanax* spp. Lake Managua and Lake Nicaragua clades (Figs 3 and S2). In these clades, mtDNA haplotypes from upper San Juan tributaries to Lakes Managua and Nicaragua and other central-southern Chortis drainages (e.g. Matagalpa River) are closely related to those from Pacific Nicaragua and Gulf of Fonseca drainages (Figs 1b and S2). This agrees with earlier traditional biogeography studies that identified fish species ranges overlapping in the Atlantic and Pacific coasts, including Bussing (1976), who posited that

historical freshwater connections likely existed between Gulf of Fonseca drainages (including Estero Real swampland) and Atlantic-coastal Ulúa and Patuca Rivers (Honduras) and Lake Managua (San Juan drainage). Similar to Bussing, our results indicate that historical connections (not present today) existed across the Nicaraguan depression, between the Atlantic drainages of the south-central Chortis plateau (especially the Matagalpa and Coco Rivers) and the Pacific drainages of Honduras and Nicaragua. This also suggests that the small Los Marabíos Cordillera (on the floor of the Nicaraguan depression), which originated through Quaternary volcanism (Marshall 2007), has not been an important barrier to fish movement. Notably, the cross-Los Marabíos Cordillera pattern was found in all three lineages sampled across this range, indicating that this pattern is potentially more general to the CA freshwater fish assemblage and should be expected in other, as-yet unsampled species.

Finally, *P. mexicana* was the only lineage in our study with identical or closely related haplotypes distributed across the Central Cordillera and San Blas Range, Panama (Figs 3 and S2). Genetic divergences within *P. mexicana* across these cordilleras are much younger than the history of the underlying strata (which have been subaerial since ~12–7 Ma in the Miocene; Coates & Obando 1996; Coates *et al*. 2004) and at 0%–0.3% are most consistent with recent, late Quaternary–Holocene dispersal events (assuming the 1.57% 'fish rate'; see Materials and methods). Previous phylogeography studies of "*Bryconamericus* '*emperador*'" and "*Brycon* east" tetra lineages, and *Andinoacara coeruleopunctatus* cichlids, recovered a similar pattern of cross-cordillera exchange across the Central Cordillera (e.g. zero to limited genetic divergences), indicating recent headwater river capture events caused by tectonic activity at El Valle volcano (Reeves & Bermingham 2006; McCafferty *et al*. 2012). Thus, levels of genetic divergence within multiple lineages from three different families reconcile well with the very recent

Quaternary–Holocene record of El Valle volcanic activity, since ~200 ka. However, it is only the pattern within *P. mexicana* subclade part H (haplotype 160, Data S1; not part I) that involved rivers (e.g. Indio River) that drain the slopes of El Valle's large caldera, as in the tetras and cichlids. In contrast, this pattern in part I alleles (e.g. haplotype 155, Data S1) may have arisen through another dispersal pathway or reflect an artifact of genetic processes such as incomplete lineage sorting. The *A. coeruleopunctatus* study, as well as previous studies of *Rhamdia guatemalensis* catfishes and *Hypostomus* spp. knifefishes also supported cross-cordillera exchanges similar to those in *P. mexicana* across the San Blas Range (Bermingham & Martin 1998; Perdices *et al*. 2002; McCafferty *et al*. 2012). Thus, the consensus from genetic data from four lineages of freshwater fishes strongly suggests that headwater captures have facilitated fish movement between eastern Pacific Panama drainages and western San Blas drainages with proximate headwaters in the same general Chagres–Tuira River area. And this inference is highly plausible because the drainage divides in this area are rather low and at mostly ~200-500 m asl constitute among the lowest continental-divide areas in Panama, and in CA at large.

Other biogeographers have attributed cross-cordillera river captures *between* versants limited impact on fish faunal composition, in general (e.g. Bishop 1995), and specifically in CA because relatively species-poor upland CA fish communities result in limited pools of potential colonists (e.g. Smith & Bermingham 2005; Matamoros *et al*. 2014). By contrast, our results from the most extensive comparative phylogeographic analysis in the region to date, including comparisons with published studies, demonstrate that headwater river captures and other means of cross-cordillera dispersal (episodic swamp connections or river reversals) have altered drainage arrangements across the continental divide, and thus exerted important general

influences on patterns of Neotropical fish distributions and genetic diversity in the CA freshwater fish assemblage.

(5) Overall paleodistributional congruence but phylogeographic incongruence

An important assumption of comparative phylogeography is that presently coexisting species were codistributed in the past; however, this assumption was rarely tested in phylogeography until geospatial tools for ENM, and paleo-environmental data, became more advanced and widely available (e.g. Hijmans *et al*. 2005; Carstens & Richards 2007; Knowles 2009; Marske *et al*. 2012). Indeed, ENMs provide invaluable predictions not only of species present-day distributions (and ecological tolerances, or environmental-space "niches"), but also of their potential paleodistributions through ENM hindcasting (re-projection) on paleoclimatic data layers (e.g. Elith *et al*. 2006; Waltari *et al*. 2007; Warren *et al*. 2010). Our paleodistribution modeling results revealed evidence for similar Pleistocene range dynamics, with range overlap and relative range stability (especially over LGM–recent) among all eight of the focal lineages (Fig. 2), providing strong support for the assumption of ancestral codistribution. Due to niche overlap in the ENMs across time slices and species, we inferred that the Tortuguero lowlands of Costa Rica have been the most ecologically stable area throughout the late Pleistocene (from LIG to present; Fig. 2); thus, we hypothesize that this area formed an important bioclimatic refugium for CA freshwater fishes during Pleistocene climate changes. These results agree with evidence that montane regions of CA experienced ~5–8°C cooling and drying during the LGM and probably other Pleistocene glaciations, and that the highest Talamanca Cordillera peaks were covered by small glaciers (Bush *et al*. 1992; refs. in Bagley & Johnson 2014a), indicating that at least some fish populations could have experienced climatically-driven size reduction or extinction during glaciations (as hypothesized for temperate European assemblages; Hewitt

2000). Indeed, habitats in the Talamanca Cordillera were predicted to be suitable through glaciations and interglaciations in only two out of eight focal lineages (livebearers *P. annectens* and *P. amates*; Fig. 2). By contrast, the Tortuguero lowlands experienced less extreme changes in temperature (~2°C cooling) during the LGM, as revealed by the paleoclimatic reconstructions we used in our MaxEnt analyses (CCSM3; Otto-Bliesner *et al*. 2006a) as well as independent palynological evidence and climatic simulations (refs. in Bagley & Johnson 2014a). Given these changes did not substantially alter the focal species predicted ranges during the LGM, this suggests that CA freshwater fishes unlikely live near their lower temperature limits (critical thermal minima), as we previously inferred from comparing physiological data and paleoclimatic models for the southeastern North American livebearing fish *Heterandria formosa* (Bagley *et al*. 2013).

Another key finding from our paleodistribution models was that they predicted that bioclimatically suitable habitats for most lineages within the Tortuguero lowlands were restricted to smaller, even trace-predicted areas during the warmer-than-recent LIG (Fig. 2), which we suggest indicates a congruent pattern of (active or passive) range contraction to one or multiple refugia. At first, this seems strikingly counterintuitive, because phylogeography studies have heavily emphasized reconstructing how distributions of temperate species in the Northern Hemisphere were affected by an opposite scenario of LIG range stability, followed by southward (or downslope) range contractions to LGM refugia, and then post-LGM range expansions (e.g. Hewitt 2000; Bagley *et al*. 2013, refs. therein). However, the congruent, multi-species range-shift scenario supported by our results can be simply explained by well-established adaptive patterns in aquatic biology. Despite living in among the warmest environments today and thus being thermally adapted to higher temperatures, tropical freshwater fishes and shallow marine

invertebrates live in systems where diurnal temperature fluctuations readily approach their upper thermal limits (critical thermal maxima; e.g. reviewed by Ficke *et al*. 2007). Global climate models predict that temperatures were ~2°C warmer in CA and other tropical regions, and that global sea levels were raised by ~3.4 m during the LIG (Otto-Bliesner *et al*. 2006b). These small increases in temperature during the LIG and other warmer periods of the past (or even future climate change projections) could easily cause daily thermal maxima to exceed species thermal limits (refs. in Ficke *et al*. 2007), especially during periods of CA dry-season drought. However, our results suggest the possibility that fish populations survived in 'microrefugia' (trace predicted areas) that may have been too fine-scale to be predicted by our models where available data layers were more coarse than the spatial scale of these microhabitats (see also Shepard & Burbrink 2008; Bagley *et al*. 2013). Also, we cannot rule out the possibility that fishes sought depth refugia, which we did not model in our analyses because high-resolution predicted paleo-environmental river depth data are not available for the Pleistocene and therefore could not be modeled across all three time slices.

Although the main goal for our paleodistribution modeling analyses was testing the assumption of ancestral codistribution while inferring congruent Pleistocene range dynamics, visually comparing paleodistribution models with patterns of genealogical lineages can help determine whether species have coexisted as a single "evolutionary cohort" *sensu* Carstens & Richards (2007) that likely experienced a longstanding history in the same local communities, despite climatic fluctuations. Visual comparison of ENM and paleodistribution results with patterns of genetic diversity also allows testing for relationships between genetic diversity and the relative bioclimatic stability versus instability of areas, which can indicate whether the genetic predictions of ENM refugial or recolonized-area models (respectively) are met (e.g.

Carnaval *et al*. 2009). Similar to the Carstens & Richards (2007) approach, we consider defining evolutionary cohorts to require three levels of phylogeographic congruence: spatial-genetic, temporal, and paleodistributional; and from groups of species that meet these more stringent criteria we can strongly infer that patterns of genetic variation have arisen from concerted geographical and genetic responses among taxa to environmental history in a region. In light of the overarching pattern of spatial phylogeographic incongruence but paleodistributional congruence uncovered herein, our results suggest that very different broad-scale phylogeographical patterns (e.g. in *P. mexicana* versus *P. amates*) have arisen from overlapping ancestral distributions. Thus our focal lineages do not, on the whole, represent an evolutionary cohort, but likely became assembled in modern communities through different routes, which intersected during climatically favorable periods. This interpretation is similar to that of recent comparative phylogeographical analyses of four dead-wood specialist beetles from New Zealand temperate forests, from which Marske *et al*. (2012) inferred overlapping climatic refugia during Pleistocene glaciations, but highly species-specific range transformations during recolonization out of those refugia. However, similar to the beetles, although our results do not identify a broad community-scale evolutionary cohort, inferred range overlap in an LGM refugium lends very strong support to the idea that CA fish communities persisted in these areas during glacial periods. This has important implications, because during the last million years glacial periods have dominated global climate patterns in 100-kyr cycles, only being interrupted by short ~11-18 kyr warmer interglaciation periods, such as those of the LIG and present-day (Lambeck *et al*. 2002). Thus, we extrapolate from our results that after colonizing the study area our focal lineages may have persisted in CA with ranges that were relatively similar to their present-day ranges for very long periods, which suggests that patterns of recolonization and range expansion

would only be indicated where lineages were substantially impacted by historical vicariance and dispersals facilitated by interactions between drainages and eustatic sea-level fluctuations, or mechanisms of drainage rearrangement and cross-cordillera exchange. We therefore hypothesize that the incongruent or idiosyncratic phylogeographical responses we have documented more likely owe to species-specific responses to these phenomena, rather than Pleistocene climatic fluctuations. That said, the overlapping paleodistributions and evidence for spatially and temporally congruent phylogeographical divergences in multiple lineages across shared breaks indicate concerted responses to environmental change consistent with the interpretation that these lineages formed Plio–Pleistocene evolutionary cohorts within the San Carlos and Sixaola Rivers.

Acknowledgements

We thank F. Alda (Smithsonian Tropical Research Institute [STRI], Tulane University), E. Bermingham (STRI, Patricia and Phillip Museum of Science), A. Bentley (Kansas University Biodiversity Institute and Natural History Museum), P. Chakrabarty and C. McMahan (Louisiana State University), Spencer J. Ingley (Brigham Young University [BYU]), W.A. Matamoros (Southern Mississippi University), and M. Tobler (Kansas State University) for providing fish tissue or DNA samples used in this work. We are also grateful to E. Castro Nallar, J.B. Lee, J.T. Nelson, G.R. Reina, A.H. Smith, and E.P. van den Berghe for valuable assistance conducting fieldwork for this project. Costa Rican sampling was conducted under SINAC-MINAET (Ministerio de Ambiente Energía y Telecomunicaciones) Resolución 030-2010-SINAC and Resolución 134-2012-SINAC, and Nicaraguan sampling was conducted under appropriate MARENA (Ministerio del Ambiente y los Recursos Naturales) permits DGPN/DB/DAP-IC-0008-2010, DGPN/DB-IC-009-2012 and DGPN/DB-21-2012. We thank J. Guevara Siquiera

(San José, Costa Rica), Edilberto Duarte (Managua, Nicaragua), Eric van den Berghe (San Marcos, Nicaragua), and G.R. Reina (Ciudad de Panama, Panama) for help issuing collecting and export permits. We also benefitted from discussions and comments on earlier drafts of the manuscript from B.J. Adams, M.C. Belk, K.A. Crandall, and P.J. Unmack. Research was funded by a BYU Mentoring Environment Grant (to J.B.J.) and Graduate Research Fellowship (to J.C.B.) award, a US National Science Foundation Doctoral Dissertation Improvement Grant (DEB-1210883), and stipend support from US National Science Foundation grant OISE-PIRE 0530267 (to J.B.J.). This work was submitted in partial fulfillment by J.C.B. of the requirements for the degree of PhD at Brigham Young University.

References

Albert JS, Petry P, Reis RE (2011). Major biogeographic and phylogenetic patterns. In: *Historical Biogeography of Neotropical Freshwater Fishes* (eds Albert JS, Reis RE), pp. 21-57. University of California Press, Berkeley, CA.

Albert JS, Reis RE (2011) Introduction to Neotropical freshwaters. In: *Historical Biogeography of Neotropical Freshwater Fishes* (eds Albert JS, Reis RE), pp. 3-20. University of California Press, Berkeley, CA.

Bagley JC, Alda F, Breitman MF, van den Berghe E, Bermingham E, Johnson JB (in revision) Assessing species boundaries using multilocus species delimitation in a morphologically conserved group of Neotropical freshwater fishes, the *Poecilia sphenops* species complex (Poeciliidae). *PLoS One*.

Bagley JC, Johnson (2014a) Phylogeography and biogeography of the lower Central American Neotropics: diversification between two continents and between two seas. *Biological Reviews*, **89**, 767-790.

Bagley JC, Johnson JB (2014b) Testing for shared biogeographic history in the lower Central American freshwater fish assemblage using comparative phylogeography: concerted, independent, or multiple evolutionary responses? *Ecology and Evolution*, **4**, 1686-1705.

Bagley JC, Sandel M, Travis J, Lozano-Vilano M de L, Johnson JB (2013) Paleoclimatic modeling and phylogeography of least killifish, *Heterandria formosa*: insights into Pleistocene expansion-contraction dynamics and evolutionary history of North American Coastal Plain freshwater biota. *BMC Evolutionary Biology*, **13**, 223.

Banarescu P (1992) *Zoogeography of Fresh Waters, vol. 2, Distribution and Dispersal of Freshwater Animals in North America and Eurasia*. AULA-Verlag, Wiesbaden, Germany.

Barker BS, Rodríguez-Robles JA, Aran VS, Montoya A, Waide RB, Cook JA (2012) Sea level, topography and island diversity: phylogeography of the Puerto Rican Red-eyed Coqui, Eleutherodactylus antillensis. *Molecular Ecology*, **21**, 6033-6052.

Bermingham E, Martin AP (1998) Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology*, **7**, 499-517.

Bishop P (1995) Drainage rearrangement by river capture, beheading and diversion. *Progress in Physical Geography*, **19**, 449-473.

Bouckaert R, Heled J, Kühnert D, Vaughan TG, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, **10**(4), e1003537. Available at: http://beast2.org/.

Burridge CP, Craw D, Fletcher D, Waters JM (2008a) Geological dates and molecular rates: fish DNA

sheds light on time dependency. *Molecular Biology and Evolution*, **18**, 624-633.

Burridge CP, Craw D, Jack DC, King TM, Waters JM (2008b) Does fish ecology predict dispersal across a river drainage divide? *Evolution*, **62**(6), 1484-1499.

Burridge CP, Craw D, Waters JM (2006) River capture, range expansion, and cladogenesis: the genetic signature of freshwater vicariance. *Evolution*, **60**, 1038-1049.

Bush MB, Piperno DR, Colinvaux PA, De Oliveira PE, Krissek LA, Miller MC, Rowe WE (1992) A 14,300-yr paleoecological profile of a lowland tropical lake in Panama. *Ecological Monographs*, **62**, 251-275.

Bussing WA (1976) Geographic distribution of the San Juan ichthyofauna of Central America with remarks on its origin and ecology. In: *Investigations of the ichthyofauna of Nicaraguan lakes* (ed. Thorson TB), pp. 157-175. University of Nebraska, Lincoln, NE.

Bussing WA (1998) *Freshwater Fishes of Costa Rica*, 2nd Edition. Editorial de la Universidad de Costa Rica, San José, Costa Rica.

Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C (2009) Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. *Science*, **323**, 785-789.

Carstens BC, Richards CL (2007) Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution*, **61**(6), 1439-1454.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657-1659.

Coates AG, Collins LS, Aubry MP, Berggren WA (2004) The geology of the Darien, Panama, and the late Miocene-Pliocene collision of the Panama arc with northwestern South America. *Geological Society of America Bulletin*, **116**, 1327-1344.

Coates AG, Obando JA (1996) The geologic evolution of the Central American isthmus. In: *Evolution and Environment in Tropical America* (eds Jackson JBC, Budd AF, Coates AG), pp. 21-56. University of Chicago Press, Chicago, IL.

Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JJ, Henderson TB, Kiehl JT, Large WG, McKenna DS, Santer BD, Smith RD (2006) The community climate system model version 3 (CCSM3). *Journal of Climatology*, **19**, 2122-2143.

Cunningham CW, Collins T (1994) Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. In: *Molecular Ecology and Evolution: Approaches and Applications* (eds Schierwater B, Streit B, Wagner GP, DeSalle R), pp. 405-433. Birkhauser Verlag, Basel, Switzerland.

Donoghue MJ, Moore BR (2003) Toward an integrative historical biogeography. *Integrative and Comparative Biology*, **43**, 261-270.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839-1854.

Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A *et al*. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**(2), 129-151.

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

Elmer KR, Meyer A (2011) Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends in Ecology adn Evolution*, **26**, 298-306.

Ficke AD, Myrick CA, Hansen LJ (2007) Potential impacts of global climate change on freshwater fisheries. *Reviews in Fish Biology and Fisheries*, **17**, 581-613.

Gilbert CR (1980) Zoogeographic factors in relation to biological monitoring of fish. In: *Biological Monitoring of Fish* (eds Hocutt CH, Stauffer JR Jr), pp. 309-339. Lexington Books, Lexington.

Haq BU, Hardenbol J, Vail PR (1987) Chronology of fluctuating sea levels since the Triassic. *Science*, **235**, 1156-1167.

Hearty PJ, Kindler P, Cheng H, Edwards RL (1999) A +20 m middle Pleistocene sea-level highstand (Bermuda and the Bahamas) due to partial collapse of Antarctic ice. *Geology*, **27**, 375-378.

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

Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905-920.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747-760.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Science of the USA*, **104**, 2785-2790.

Hickerson MJ, Stahl EA, Lessios HA (2006) Test for simultaneous divergence using approximate Bayesian computation. *Evolution*, **60**(12), 2435-2453.

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

Hickerson MJ, Stone GN, Lohse K, Demos TC, Xie X, Landerer C, Takebayashi N (2014) Recommendations for using msBayes to incorporate uncertainty in selecting an ABC model prior: a response to Oaks et al. *Evolution*, **68**, 284-294.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**(15), 1965-1978.

Hillis DM, Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology*, **42**, 182-192.

Hoeting JA, Madigan D, Raftery AE (1999) Bayesian model averaging: a tutorial. *Statistical Science*, **14**, 382-417.

Hoorn C, Wesselingh FP, ter Steege H, Bermudez MA, Mora A, Sevink J, Sanmartín I, Sanchez-Meseguer A, Anderson CL, Figueiredo JP, Jaramillo C, Riff D, Negri FR, Hooghiemstra H, Lundberg J, Stadler T, Särkinen T, Antonelli A (2010) Amazonia through time: Andean uplift, climate change, landscape evolution and biodiversity. *Science*, **330**, 927-931.

Huang W, Takebayashi N, Qi Y, Hickerson MJ (2011) MTML-msBayes: Approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. *BMC Bioinformatics*, **12**, 1.

Hubert N, Renno J-F (2006) Historical biogeography of South American freshwater fishes. *Journal of Biogeography*, **33**, 1414-1436.

Irion G (1984) Sedimentation and sediments of Amazonian rivers and evolution of the Amazonian landscape since Pliocene times. In: *The Amazon: Limnology and Landscape Ecology of a Mighty Tropical River and its Basin* (ed. Sioli H), pp. 201-214. Dr. W. Junk Publishers, The Hague.

Jones CP, Johnson JB (2009) Phylogeography of the livebearer *Xenophallus umbratilis* (Teleostei: Poeciliidae): glacial cycles and sea level change predict diversification of a freshwater tropical fish. *Molecular Ecology*, **18**, 1640-1653.

Kass RE, Raftery AE (1995) Bayes Factors. *Journal of the American Statistical Association*, **90**, 773-795.

Knowles LL (2009) Statistical phylogeography. *Annual Review in Ecology, Evolution, and Systematics*, **40**, 593-612.

Lambeck K, Esat TM, Potter EK (2002) Links between climate and sea levels for the past three million years. *Nature*, **419**, 199-206.

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.

Lee JB, Johnson JB (2009) Biogeography of the livebearing fish *Poecilia gillii* in Costa Rica: are phylogeographical breaks congruent with fish community boundaries? *Molecular Ecology*, **18**, 4088-4101.

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

Lovejoy NR, Bermingham E, Martin AP (1998) Marine incursion into South America. *Nature*, **396**, 421-

422.

Lundberg JG, Marshall LG, Guerrero J, Horton B, Malabra MCLS, Wesselingh F (1998) The stage for Neotropical diversification: a history of tropical South American rivers. Phylogeny and classification of Neotropical fishes (eds Malabra LR, Reis RE, Vari RP, Lucena ZM, Lucena CAS), pp. 13-48. EDIPUCRS, Porto Alegre, Rio Grande do Sul.

Marshall JS (2007) The geomorphology and physiographic provinces of Central America. In: *Central America: Geology, Resources and Hazards* (eds Bundschuh J, Alvarado GE), pp. 1-51. Taylor & Francis, Philadelphia, PA.

Marshall JS, Idleman BD, Gardner TW, Fisher DM (2003) Landscape evolution within a retreating volcanic arc, Costa Rica, Central America. *Geology*, **31**, 419-422.

Marske KA, Leschen RAB, Buckley TR (2012) Concerted versus independent evolution and the search for multiple refugia: comparative phylogeography of four forest beetles. *Evolution*, **66**(6), 1862-1877.

Matamoros WA, McMahan CD, Chakrabarty P, Albert JS, Schaefer JF (2014) Derivation of the freshwater fish fauna of Central America revisited: Myers' hypothesis in the twenty-first century. *Cladistics*, **2014**, 1-12.

McCafferty SS, Martin A, Bermingham E (2012) Phylogeographic diversity of the lower Central American cichlid *Andinoacara coeruleopunctatus* (Cichlidae). *International Journal of Evolutionary Biology*, **2012**.

McDowall RM (1996) Volcanism and freshwater fish biogeography in the northeastern North Island of New Zealand. *Journal of Biogeography*, **23**(2), 139-148.

McNeill DF, Coates AG, Budd AF, Borne PF (2000) Integrated paleontologic and paleomagnetic stratigraphy of the upper Neogene deposits around Limon, Costa Rica: a coastal emergence record of the Central American Isthmus. *GSA Bulletin*, **112**(7), 963-981.

Menezes NA, Ribeiro AC, Weitzman SH, Torres RA (2008) Biogeography of Glandulocaudinae (Teleostei: Characiformes: Characidae) revisited: Phylogenetic patterns, historical geology and genetic connectivity. *Zootaxa*, **1726**, 33-48.

Miller KG, Kominz MA, Browning JV, Wright JD, Mountain GS, Katz ME, Sugarman PJ, Cramer BS, Christie-Blick N, Pekar SF (2005) The phanerozoic record of global sea-level change. *Science*, **310**, 1293-1298.

Miller RR (1966) Geographical distribution of Central American freshwater fishes. *Copeia*, **1996**(4), 773-802.

Miller RR, Minckley WL, Norris SM (2005) *Freshwater Fishes of México*. The University of Chicago Press, Chicago.

Minin V, Abdo Z, Joyce P, Sullivan J (2003) Performance-based selection of likelihood models for

phylogeny estimation. *Systematic Biology*, **52**, 674-683.

Montoya-Burgos J-I (2003) Historical biogeography of the catfi sh genus *Hypostomus* (Siluriformes: Loricariidae), with implications on the diversifi cation of Neotropical ichthyofauna. *Molecular Ecology*, **12**, 1855-1867.

Myers GS (1938) Fresh-water fishes and West Indian zoogeography. *Annual Report Smithsonian Institution*, **1937**, 339-364.

Myers GS (1966) Derivation of the freshwater fish fauna of Central America. *Copeia*, **4**, 766-773.

Oaks JR, Sukumaran J, Esselstyn JA, Linkem CW, Siler CD, Holder MT, Brown RM (2013) Evidence for climate-driven diversification? A caution for interpreting ABC inferences of simultaneous historical events. *Evolution*, **67**, 991-1010.

Ornelas-García CP, Dominguez-Dominguez O, Doadrio I (2008) Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evolutionary Biology*, **8**, 340.

Otto-Bliesner BL, Brady EC, Clauzet G, Tomas R, Levis S, Kothavala Z (2006a) Last Glacial Maximum and Holocene climate in CCSM3. *Journal of Climate*, **19**, 2526-2544.

Otto-Bliesner BL, Marshall SJ, Overpeck JT, Miller GH, Hu A, CAPE Last Interglacial Project members (2006b) Simulating arctic climate warmth and icefield retreat in the last interglaciation. *Science*, **311**, 1751-1753.

Perdices A, Bermingham E, Montilla A, Doadrio I (2002) Evolutionary history of the genus *Rhamdia* (Teleostei: Pimelodidae) in Central America. *Molecular Phylogenetics and Evolution*, **25**, 172-189.

Perdices A, Doadrio I, Bermingham E (2005) Evolutionary history of the synbranchid eels (Teleostei: Synbranchidae) in Central America and the Caribbean islands inferred from their molecular phylogeny. *Molecular Phylogenetics and Evolution*, **37**, 460-473.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecolical Modeling*, **190**(3-4), 231-259.

Poff NL, Allan JD (1995) Functional organization of stream fish assemblages in relation to hydrological variability. *Ecology*, **76**, 606-627.

Rambaut A, Drummond AJ (2013) Tracer v1.5. <http://beast.bio.ed.ac.uk/tracer>.

Reeves RG, Bermingham E (2006) Colonization, population expansion, and lineage turnover: phylogeography of Mesoamerican characiform fish. *Biological Journal of the Linnean Society*, **88**, 235-255.

Reis RE, Kullander SO, Ferraris CJ (2003) *Checklist of the Freshwater Fishes of South and Central America*. EDIPUCRS, Porto Alegre, Rio Grande do Sul.

Ribeiro AC (2006) Tectonic history and the biogeography of the freshwater fishes from the coastal

drainages of eastern Brazil: An example of faunal evolution associated with a divergent continental margin. *Neotropical Ichthyology*, **4**, 225-246.

Riddle BR, Hafner DJ (2006) A step-wise approach to integrating phylogeographic and phylogenetic biogeographic perspectives on the history of a core North American warm deserts biota. *Journal of Arid Environments*, **66**, 435-461.

Rogers RD, Karason H, van der Hilst R (2002) Epirogenic uplift above a detached slab in northern Central America. *Geology*, **30**, 1031-1034.

Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society of London B*, **273**, 1987-1998.

Shepard DB, Burbrink FT (2008) Lineage diversification and historical demography of a sky island salamander, *Plethodon ouachitae*, from the interior highlands. *Molecular Ecology*, **17**, 5315-5335.

Smith SA, Bermingham E (2005) The biogeography of lower Mesoamerican freshwater fishes. *Journal of Biogeography*, **32**, 1835-1854.

Suchard MA, Weiss RE, Sinsheimer JS (2001) Bayesian selection of continuous-time Markov chain evolutionary model. *Molecular Biology and Evolution*, **18**, 1001-1013.

Sullivan J, Arellano E, Rogers DS (2000) Comparative phylogeography of mesoamerican highland rodents: concerted versus independent response to past climatic fluctuations. *American Naturalist*, **155**, 755-768.

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

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731-2739.

Unmack PJ (2001) Biogeography of Australian freshwater fishes. *Journal of Biogeography*, **28**, 1053-1089.

Unmack PJ, Bagley JC, Adams M, Hammer MP, Johnson JB (2012) Molecular phylogeny and phylogeography of the Australian freshwater fish genus *Galaxiella*, with an emphasis on dwarf galaxias (*G. pusilla*). *PLoS One*, **7**, e38433.

Unmack PJ, Hammer MP, Adams M, Johnson JB, Dowling TE (2013) The role of continental shelf width in determining freshwater phylogeographic patterns in south-eastern Australian pygmy perches (Teleostei: Percichthyidae). *Molecular Ecology*, **22**(6), 1683-1699.

Waltari E, Hijmans RJ, Peterson AT, Nyari AS, Perkins SL, Guralnick RP (2007) Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS One*, **2**(7):e563.

Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental

niche models. *Ecography*, **33**, 607-611.

Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**, 335-342.

Waters JM, Burridge CP (1999) Extreme intraspecific mitochondrial DNA sequence divergence in *Galaxias maculatus* (Osteichthys: Galaxiidae), one of the world's most widespread freshwater fish. *Molecular Phylogenetics and Evolution*, **11**, 1-12.

Zwickl DJ (2006) *Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets Under the Maximum Likelihood Criterion*. The University of Texas, Austin, TX.

Tables

Table 1 Predictions of four hypotheses of Neotropical fish diversification and the methods used to test them in this study

	1. Tectonic vicariance hypothesis	2. Marine vicariance hypothesis	3. Continental shelf width hypothesis	4. Cross-cordillera exchange hypothesis	Method used
Historical processes:	Vicariance due to tectonic processes at mountain ranges and volcanoes of the CAVA and the continental divide, and at the eastern Tárcoles River drainage divide	Vicariance during marine incursions during Miocene–Pleistocene sea-level highstands	Pleistocene–recent sea-level low stands causing: i. Interdrainage dispersal in areas with wide continental shelf; ii. Isolation within or between coastal drainages in areas with narrow continental shelf	Headwater drainage rearrangements causing: i. Range expansion, dispersal, with insufficient time for cladogenesis; ii. Vicariance across drainage divides (e.g. mountainous cordilleras)	–
Reciprocal monophyly of populations:	Across mountain ranges and volcanoes (volcanic domes, fallout/flow zones) of CAVA and the continental divide, and eastern Tárcoles drainage divide	Upland versus lowland clades located between upland freshwater refugia and recolonized lowland populations	i. Limited or no divergence in areas with wide continental shelf due to dispersal (gene flow); ii. Between drainage basins in regions with narrow continental shelf	i. If range expansion, clades distributed across cordilleras with limited or no genetic divergence; ii. If vicariance, clades/sub-clades shallowly to deeply diverged across mountainous cordilleras	ML (GARLI) and Bayesian (BEAST) gene tree reconstruction
Timing of diversification:	Mid-Miocene–Quaternary, but specific to different geographical barriers. a. CAVA: Nuclear	The same for all CA Caribbean and Pacific lowlands: Late Miocene–Quaternary eustatic sea-level highstands 7–5 Ma,	N/A; anytime, but populations may exhibit Pleistocene–recent coalescent histories if repeated glacial low-sea stands	N/A; anytime, but possibly coincident with tectonic periods of cordillera (e.g. CAVA) formation, as in the tectonic vicariance	Coalescent divergence-dating analyses in BEAST and IMa2 (shared breaks); tests for simultaneous diversification (see

	CA: Chortis highlands, 19–3.8 Ma; Lower CA: Guanacaste Cord., Quaternary; Tilarán Cord., Quaternary; Costa Rica Central Cord., <800–300 ka; Grifo Alto form. (Aguacate Cord.), 5–4 Ma; Sarapiquí Arc mtns., 18–11.4 Ma; Talamanca Cord., main uplift phase 5.5–3.5 Ma; b. Volcanic debris fans over Tárcoles River, 1.7–0.3 Ma (Tivives, Intracañon, Avalancha, and Orotina formations)	3.5–3 Ma, and 550–390 ka; older or deeper coalescing upland alleles, and more recent or shallowly coalescing lowland population histories	permitted recurrent gene flow and secondary contact in areas with wider continental shelf	hypothesis (e.g. San Carlos River (Atlantic) to Tárcoles River (Pacific) divergences across the Central Cordillera would date to tectonism in the cordillera ~800–300 ka)	below)
Ancestral populations are located:	N/A; either side of CAVA cordilleras or volcanoes, and eastern Tárcoles drainage divide	Upland freshwater refugia	N/A; one or multiple coastal drainages	N/A; either side of cordilleras	ML (GARLI) and Bayesian (BEAST) gene tree reconstruction
Derived populations are located:	N/A; either side of CAVA cordilleras or volcanoes, and eastern Tárcoles drainage divide	Coastal plain populations (lower elevations than upland clades)	N/A; one or multiple coastal drainages	N/A; either side of cordilleras	ML (GARLI) and Bayesian (BEAST) gene tree reconstruction
Simultaneous divergences:	Yes	Yes	No	Yes	Multi-taxon tests for simultaneous vs. non-

			simultaneous diversification across shared genetic breaks (MTML-msBayes)

Consistent with the focus of this study, predictions are specified in the context of Central America, rather than the entire North and South American Neotropics (see references and discussion in text and Appendix S1). Abbreviations: CA, Central America; CAVA, Central America volcanic arc; Cord., Cordillera; form., formation; ML, maximum-likelihood; mtns., mountains; N/A, not applicable.

Table 2 List of focal lineages (species and genera) used to test hypotheses of Neotropical fish diversification in this study, and their taxonomic, DNA sequence, and ecological characteristics

Focal lineage	Family	n	bp	Max. % div. (p / d_{MTN})	No. clades	Habit	Elevation (m asl)	Pacific range	Atlantic range
1. *Alfaro cultratus*	Poeciliidae	355	1140	4.01 / 4.50	5	Bp, I	0–300	Ni, CR	Ho, Ni, CR, Pa
2. *Phallichthys amates*	Poeciliidae	93	1140	8.80 / 11.36	4	Bp, O	0–500	Ni, CR	Ho, Ni, CR, Pa
3. *Poecilia mexicana*	Poeciliidae	761	1086	3.04 / 3.35	1	Bp, O	0–1220	ES, Gu, Ho, N, CR, Pa	Gu, Ho, Ni, CR, Pa
4. *Priapichthys annectens*	Poeciliidae	101	1140	11.15 / 15.72	7	Bp, O	25–1270	–	CR
5. *Xenophallus umbratilis*	Poeciliidae	180	1140	5.35 / 6.36	3	Bp, O	35–590	CR	Ni, CR
6. *Astyanax* spp.	Characidae	153	1140	6.35 / 7.58	8	Bp, O	0–1000	ES, Gu, Ho, Ni, CR, Pa	Gu, Ho, Ni, CR, Pa
A. nasutus	Characidae	16	1140	0.70 / 0.72	1	Bp, O	1–100?	–	ES, Ho, Ni
A. nicaraguensis	Characidae	58	1140	1.86 / 1.96	3	Bp, O	1–100?	–	Ni, CR
A. orthodus/sp.	Characidae	2	1140	0.00 / 0.00	1	Bp, O	1–60	–	Ni?, CR
7. *Roeboides* spp.	Characidae	108	1151	6.00 / 6.53	2	B, H/D	0–610	Ni, CR	Ni, CR, Pa
R. bouchellei	Characidae	98	1151	1.15 / 1.11	1	B, H/D	0–610	Ni, CR	Ni, CR
R. bussingi	Characidae	10	1151	0.13 / 0.27	1	B, H/D	0–100?	CR, Pa	–
8. *Amatitlania* spp.	Cichlidae	326	1140	2.77 / 3.01	1	Bp, O	0–540	ES, (Gu,) Ho, Ni, CR, Pa	G, Ho, Ni, CR, Pa

Taxonomic family; alignment characteristics (n, sample size; bp, sequence length); mtDNA cyt*b* genetic divergences calculated as maximum pairwise percent divergence (Max. % div.) between all samples based on *p*-distances and modified Tamura-Nei distances (d_{MTN}); number (no.) of mtDNA phylogeographic clades (this study); general ecological habits and elevational range (Bussing 1998; Miller *et al*. 2005; this study); and distributional limits on Pacific and Atlantic versants are shown for all eight focal lineages. Abbreviations: B, benthic; bp, nucleotide base pairs; Bp, benthopelagic; CR, Costa Rica; D, detritivore; ES, El Salvador; Gu, Guatemala; H, herbivore; Ho, Honduras; I, insectivore; N/A, not applicable; Ni, Nicaragua; O, omnivore; Pa, Panama.

Table 3 Divergence times for the eight focal lineages, estimated as times to the most recent common ancestor (t_{MRCA}s) using relaxed molecular clock dating analyses in BEAST with multiple calibration points

Focal lineage	t_{MRCA} **(Ma)**	**95% HPDs (Ma)**
1. *Alfaro cultratus*	3.664	[2.220, 5.751]
2. *Phallichthys amates*	9.739	[5.063, 16.216]
5. *Poecilia mexicana*	1.686	[1.137, 2.354]
4. *Priapichthys annectens*	11.241	[6.886, 17.547]
5. *Xenophallus umbratilis*	5.106	[2.879, 8.416]
6. *Astyanax* spp.	11.313	[8.289, 14.964]
A. nasutus	0.255	[0.008, 0.552]
A. nicaraguensis	0.588	[0.215, 1.179]
A. orthodus/sp.	1.398	[0.445, 2.802]
7. *Roeboides* spp.	7.707	[1.426, 21.277]
R. bouchellei	2.848	[0.830, 6.301]
R. bussingi	0.529	[0.103, 1.496]
8. *Amatitlania* spp.	11.605	[8.624, 15.028]

Divergence times estimated as geometric mean t_{MRCA}s and their 95% highest posterior densities (HPDs) in units of millions of years ago (Ma) are presented from combined BEAST runs that each employed uncorrelated lognormal relaxed clock models and multiple fossil or biogeographic calibration points described in Appendix S1.

Table 4 Population divergence time estimates for population-pairs split across the shared phylogeographic breaks identified in this study

	IMa2				**MTML-msBayes**			
	t	**T_{div}, 2% rate (Ma)**	**T_{div}, 1.57% rate (Ma)**	**T_{div}, 0.9% rate (Ma)**	***E*[*τ*]**	**T_{div}, 2% rate (Ma)**	**T_{div}, 1.57% rate (Ma)**	**T_{div}, 0.9% rate (Ma)**
San Carlos River break								
Priapichthys annectens	21.83	1.915	2.439	4.255	–	–	–	–
95% HPDs	[11.120, 31.320]	[0.975, 2.747]	[1.243, 3.500]	[2.168, 6.105]	–	–	–	–
Xenophallus umbratilis	24.19	2.122	2.703	4.715	–	–	–	–
95% HPDs	[9.060, 36.740]	[0.795, 3.223]	[1.012, 4.105]	[1.766, 7.162]	–	–	–	–
Model-averaged community estimate	–	–	–	–	0.603	1.392	1.773	3.093
95% HPDs	–	–	–	–	[0.298, 0.872]	[0.688, 2.013]	[0.876, 2.564]	[1.529, 4.473]
Sixaola River break								
Astyanax orthodus/sp.	3.655	0.321	0.408	0.712	–	–	–	–
95% HPDs	[0.000, 9.995]	[0.000, 0.877]	[0.000, 1.117]	[0.000, 1.948]	–	–	–	–
Phallichthys amates	9.138	0.802	1.021	1.781	–	–	–	–
95% HPDs	[1.867, 15.640]	[0.164, 1.372]	[0.209, 1.748]	[0.364, 3.049]	–	–	–	–
Priapichthys annectens	41.36	3.628	4.622	8.062	–	–	–	–
95% HPDs	[24.870, 57.150]	[2.182, 5.013]	[2.779, 6.386]	[4.848, 11.140]	–	–	–	–
Model-averaged community estimate	–	–	–	–	0.300	1.216	1.550	2.703
95% HPDs	–	–	–	–	[0.113, 0.461]	[0.458, 1.869]	[0.584, 2.381]	[1.018, 4.154]

Estimated mean mutation-scaled population splitting times (*t*; divergence per gene) and population divergence times (T_{div}) of each lineage are reported from IMa2. Mean estimated community divergence times across each shared break are also given in global

coalescent units ($E[\tau]$, in units of $4N_{ave}$) and in absolute time (T_{div}) are from MTML-msBayes, where $E[\tau]$ were derived by averaging across eight model classes in Table 5 while weighting by their relative posterior probabilities using ABC model averaging. Population and community divergence times were converted to millions of years ago (Ma) using three rates of mtDNA evolution and conversions (from t and $E[\tau]$) discussed in the text. Bayesian 95% highest posterior densities (HPDs) for estimates are given in brackets.

Table 5 Model comparisons and parameter estimates from ABC model averaging analyses in MTML-msBayes

Pop. set	Prior	$P(\tau)$	$P(\theta_D)$	$P(\theta_A)$	$P(M_K\|D)^{1000}$	Ψ mode	Ψ mean	Ω mode	Ω mean [95% HPDs]
San Carlos River break (Y = 2)						1	1.350	0.000619	0.0436 [0.000, 0.514]
	M_1	**~U(0, 2.09)**	**~U(0, 0.031)**	**~U(0, 0. 5)**	**<u>0.2100</u>**	**1**	**1.430**	**0.00293**	**0.0770 [0.000, 0.402]**
	M_2	~U(0, 2.09)	~U(0, 0.031)	~U(0, 1.0)	0.1572	–	–	–	–
	M_3	~U(0, 4.18)	~U(0, 0.031)	~U(0, 0.5)	0.1564	–	–	–	–
	M_4	~U(0, 4.18)	~U(0, 0.031)	~U(0, 1.0)	0.1366	–	–	–	–
	M_5	~U(0, 2.09)	~U(0, 0.062)	~U(0, 0.5)	0.1777	–	–	–	–
	M_6	~U(0, 2.09)	~U(0, 0.062)	~U(0, 1.0)	0.0235	–	–	–	–
	M_7	~U(0, 4.18)	~U(0, 0.062)	~U(0, 0.5)	0.0883	–	–	–	–
	M_8	~U(0, 4.18)	~U(0, 0.062)	~U(0, 1.0)	0.0504	–	–	–	–
Sixaola River break (Y = 3)						2	2.211	0.219	0.253 [0.000, 0.606]
	M_1	~U(0, 1.85)	~U(0, 0.054)	~U(0, 0.5)	0.1543	–	–	–	–
	M_2	~U(0, 1.85)	~U(0, 0.054)	~U(0, 1.0)	0.1552	–	–	–	–
	M_3	~U(0, 3.70)	~U(0, 0.054)	~U(0, 0.5)	0.1620	–	–	–	–
	M_4	~U(0, 3.70)	~U(0, 0.054)	~U(0, 1.0)	0.0990	–	–	–	–
	M_5	~U(0, 1.85)	~U(0, 0.108)	~U(0, 0.5)	0.1513	–	–	–	–
	M_6	~U(0, 1.85)	~U(0, 0.108)	~U(0, 1.0)	0.0580	–	–	–	–
	M_7	**~U(0, 3.70)**	**~U(0, 0.108)**	**~U(0, 0.5)**	**<u>0.2187</u>**	**1**	**1.444**	**0.165**	**0.171 [0.000, 0.504]**
	M_8	~U(0, 3.70)	~U(0, 0.108)	~U(0, 1.0)	0.0014	–	–	–	–

We ran MTML-msBayes analyses on two different population (pop.) sets, each with Y number of population-pairs, given at left. This table shows $K = 8$ prior model classes, $\{M_1, \ldots, M_8\}$, run for each of the analyses, characterized by different uniform prior distributions $P(\tau)$, $P(\theta_D)$, $P(\theta_A)$, and their approximate posterior probabilities $P(M_K|D)^{1000}$ based on 1000 accepted simulated draws from 4×10^7 random draws from the eight prior models. For each analysis (pop set.), results of the best-supported model are given in bold face with its posterior probability underlined, and mode and mean Ψ and Ω estimates (with 95% highest posterior densities [HPDs] around their means) are given as output from one run for that model (see text for hyper-parameter details). Model-averaged hyper-parameter estimates derived from all eight prior models are given in the first row of each section.

Figure Legends

Fig. 1 Maps of Central America showing major physiographic elements (A), hydrological features (major rivers and lakes) (B), a sea level model of lowland areas potentially inundated by marine incursions (from 250 m digital elevation model; dashed line, –135 m bathymetric contour; C), and fish biogeographic provinces (D) in the study area. Dotted lines in panel A outline distinct physiographic provinces of Marshall (2007), and major geologic faults and units (e.g. Chortis highlands). The 10 fish biogeographic provinces depicted in panel D are from Matamoros *et al*. (2014) and divide along drainage divides including the continental divide (red line). Colored relief is also shown based on digital elevation layers derived from NASA Shuttle Radar Topography Mission (SRTM) image PIA03364.

Fig. 2 Predicted paleodistributions based on ecological niche modeling analyses. MaxEnt reconstructions of the present-day (left) and Pleistocene Last Glacial Maximum (LGM; middle) and Last Interglaciation (LIG; right) paleodistributions are mapped with Azimuthal-Equidistant projections for each of the focal lineages: (A) *A. cultratus*; (B) *Astyanax* spp.; (C) *P. mexicana*; (D) *P. amates*; (E) *Amatitlania* spp.; (F) *P. annectens*; (G) *Xenophallus*; (H) and *R. bouchellei*. Colors represent logistic ENM scores ranging from 0 (dark blue; bioclimatically unsuitable areas) to 1 (100% bioclimatic suitability).

Fig. 3 Summary maps showing sampling localities and the geographical distributions of well-supported genetic lineages (clades) within all eight lineages of Central American freshwater fishes analyzed in this study. Sampling localities are represented in different color and marker styles corresponding to different clades that were consistently recovered across gene tree and

parsimony network analyses of mtDNA cyt*b* haplotypes (Fig. S2). (A) *P. mexicana*; (B) *Astyanax* spp.; (C) *P. annectens*; (D) *Xenophallus*; (E) *A. cultratus*; (F) *P. amates*; (G) *Amatitlania* spp.; (H) and *Roeboides* spp.

Fig. 4 Posterior distributions of divergence times (t_{MRCA}s) of Central American freshwater fish lineages across two geographic barriers corresponding to phylogeographic breaks in this study. Graphs at left display marginal posterior probabilities estimated in BEAST for (A) the San Carlos River break, and (B) the Sixaola River basin break (Matina–Sixaola drainage divide), and their correlation with the timing of seven regional historical events that were geographically relevant to each break, including sea-level highstands (left) and geological events (right) (details in Appendix S1, Discussion). Maps at right show the approximate geographical positions of each shared break (thick lines) and corresponding topographic features.

Fig. 5 Results of tests for simultaneous diversification using hierarchical ABC model averaging in MTML-msBayes. Left panels show comparisons of the prior versus posterior distributions of the number of divergence events (hyper-parameter Ψ), while panels at right display the joint hyper-posterior probability distributions of the mean divergence times, $E[\tau]$ (community divergence time), and Ω. Results were obtained from ABC model averaging across eight prior models (Table 4) based on ABC-rejection using local linear regression and are presented for separate analyses of population-pairs diverged across (A) the San Carlos River break and (B) the Sixaola River break. Best-fit model MTML-msBayes results are presented in Fig. S3.

Supporting Information

Data S1 Sampling localities and GenBank numbers for each lineage sampled

Appendix S1 Supplementary methods and results

Table S1 PCR primers and annealing temperatures used to amplify the mitochondrial cytochrome *b* (cyt*b*) gene for each focal lineage in this study

Table S2 Best-fit models of DNA substitution estimated in DT-ModSel

Table S3 Mean MaxEnt model AUC scores and their standard deviations (s.d.)

Figure S1 Posterior distributions of parameters estimated in IMa2 for each population-pair split across shared geographic breaks.

Figure S2 MtDNA gene trees estimated in GARLI and parsimony networks estimated using TCS for each focal lineage, showing well-supported phylogeographic clades. For comparison, geographical distributions of the main clades for each focal lineage are mapped over modern topography, similar to Fig. 3.

Figure S3 Results of best-fit models of community divergence across the San Carlos River break and Sixaola River break identified by ABC model choice in MTML-msBayes. Panels show comparisons of the prior versus posterior distributions of the dispersion index of divergence times Ω (=$Var[\tau]/E[\tau]$, where τ = divergence time).

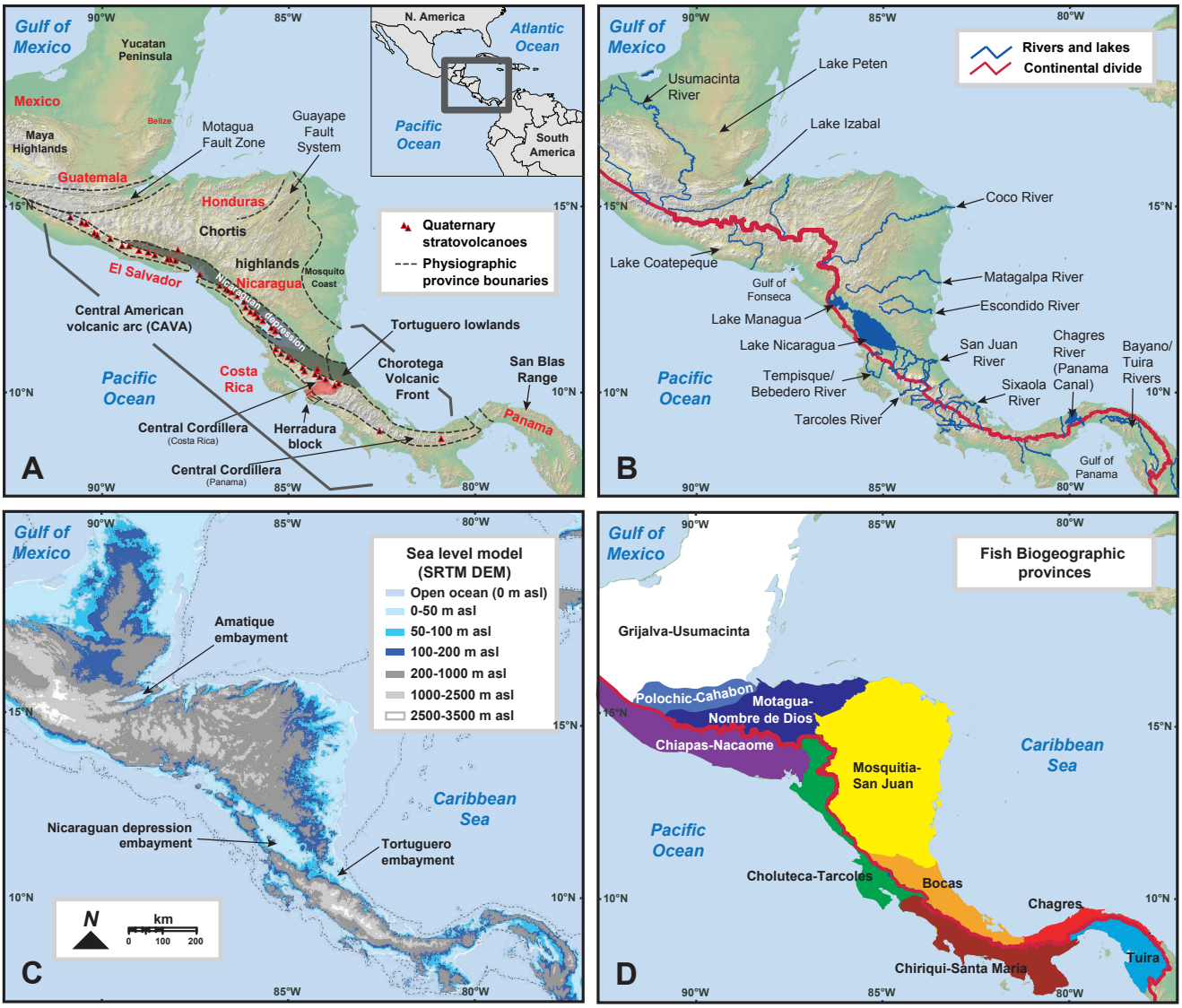
A
Gulf of Mexico
Yucatan Peninsula
Mexico
Belize
Maya Highlands
Motagua Fault Zone
Guayape Fault System
Guatemala
Honduras
Chortis
highlands
El Salvador
Nicaragua
Mosquito Coast
Nicaraguan depression
Central American volcanic arc (CAVA)
Pacific Ocean
Costa Rica
Tortuguero lowlands
Chorotega Volcanic Front
San Blas Range
Panama
Central Cordillera (Costa Rica)
Herradura block
Central Cordillera (Panama)
N. America
Atlantic Ocean
Pacific Ocean
South America
Quaternary stratovolcanoes
Physiographic province bounaries
B
Rivers and lakes
Continental divide
Usumacinta River
Lake Peten
Lake Izabal
Coco River
Lake Coatepeque
Gulf of Fonseca
Matagalpa River
Escondido River
Lake Managua
Lake Nicaragua
San Juan River
Chagres River (Panama Canal)
Bayano/ Tuira Rivers
Tempisque/ Bebedero River
Sixaola River
Tarcoles River
Gulf of Panama
C
Sea level model (SRTM DEM)
Open ocean (0 m asl)
0-50 m asl
50-100 m asl
100-200 m asl
200-1000 m asl
1000-2500 m asl
2500-3500 m asl
Amatique embayment
Caribbean Sea
Nicaraguan depression embayment
Tortuguero embayment
N
km
0 100 200
D
Fish Biogeographic provinces
Grijalva-Usumacinta
Polochic-Cahabon
Motagua-Nombre de Dios
Chiapas-Nacaome
Mosquitia-San Juan
Caribbean Sea
Pacific Ocean
Choluteca-Tarcoles
Bocas
Chagres
Tuira
Chiriqui-Santa Maria

Figure 1

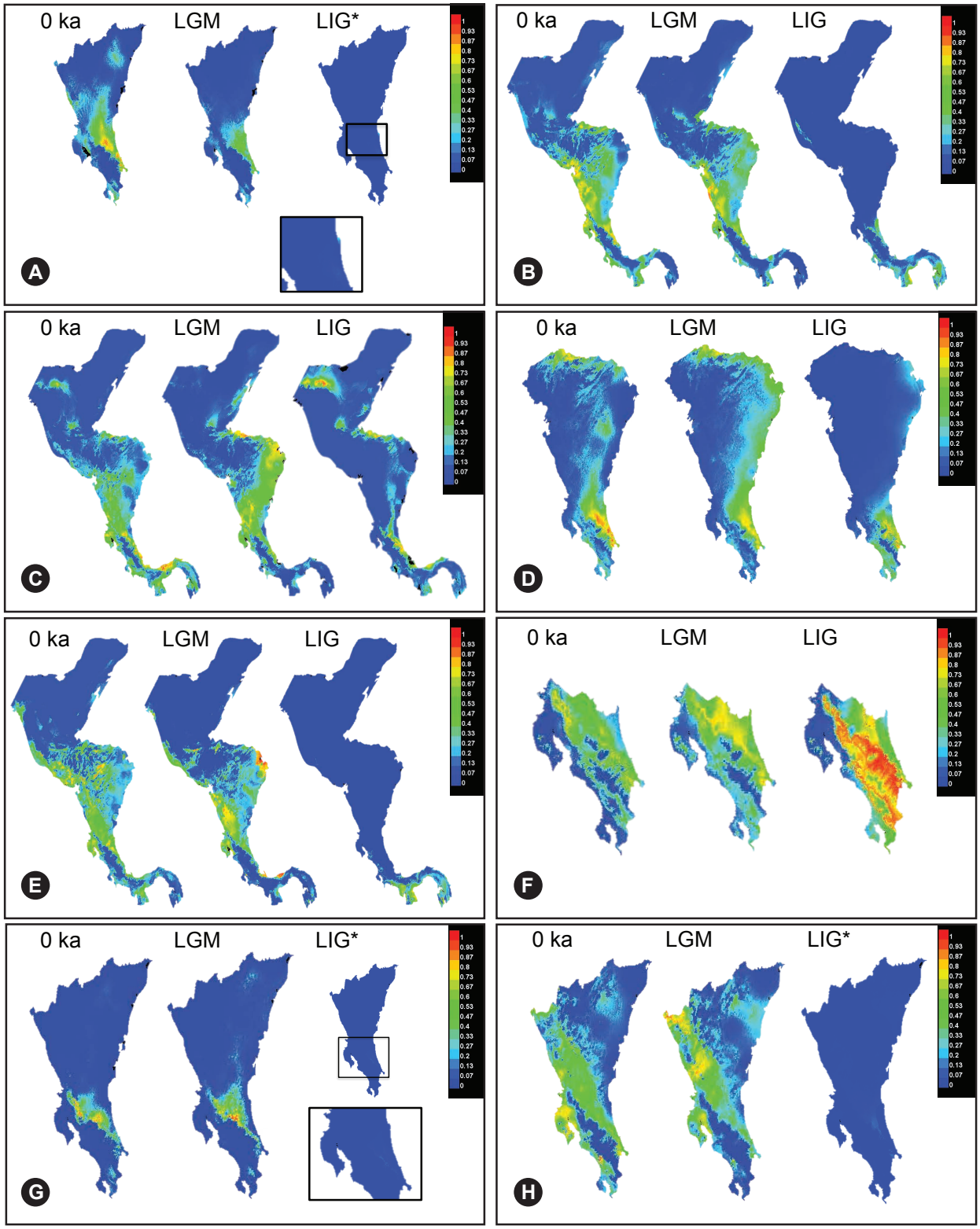
0 ka
LGM
LIG*
A
0 ka
LGM
LIG
B
0 ka
LGM
LIG
C
0 ka
LGM
LIG
D
0 ka
LGM
LIG
E
0 ka
LGM
LIG
F
0 ka
LGM
LIG*
G
0 ka
LGM
LIG*
H

Figure 2

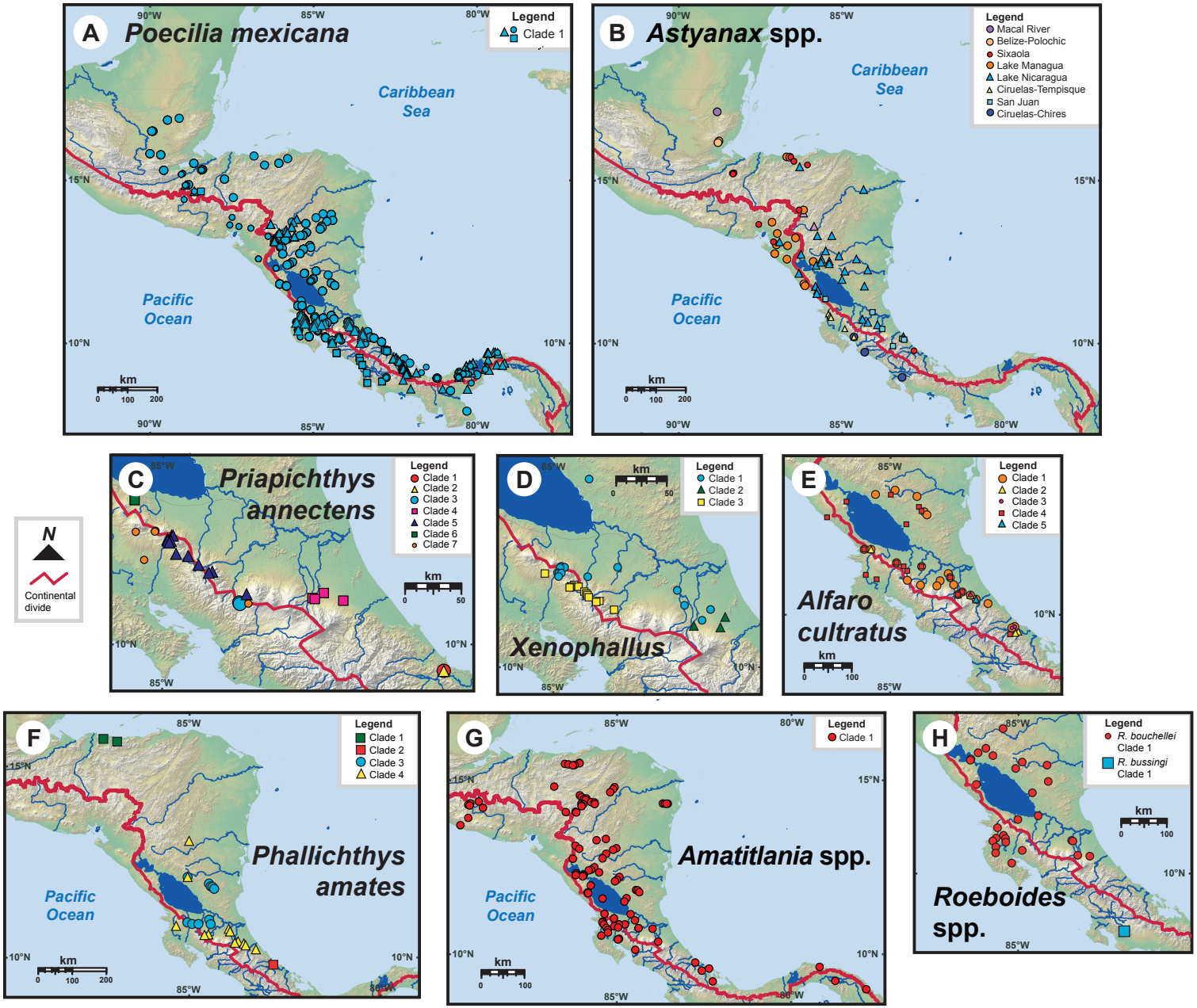
A
Poecilia mexicana
Legend
Clade 1
Caribbean Sea
Pacific Ocean
B
Astyanax spp.
Legend
Macal River
Belize-Polochic
Sixaola
Lake Managua
Lake Nicaragua
Ciruelas-Tempisque
San Juan
Ciruelas-Chires
C
Priapichthys annectens
Legend
Clade 1
Clade 2
Clade 3
Clade 4
Clade 5
Clade 6
Clade 7
D
Xenophallus
Legend
Clade 1
Clade 2
Clade 3
E
Alfaro cultratus
Legend
Clade 1
Clade 2
Clade 3
Clade 4
Clade 5
N
Continental divide
F
Phallichthys amates
Legend
Clade 1
Clade 2
Clade 3
Clade 4
G
Amatitlania spp.
Legend
Clade 1
H
Roeboides spp.
Legend
R. bouchellei
Clade 1
R. bussingi
Clade 1
km

Figure 3

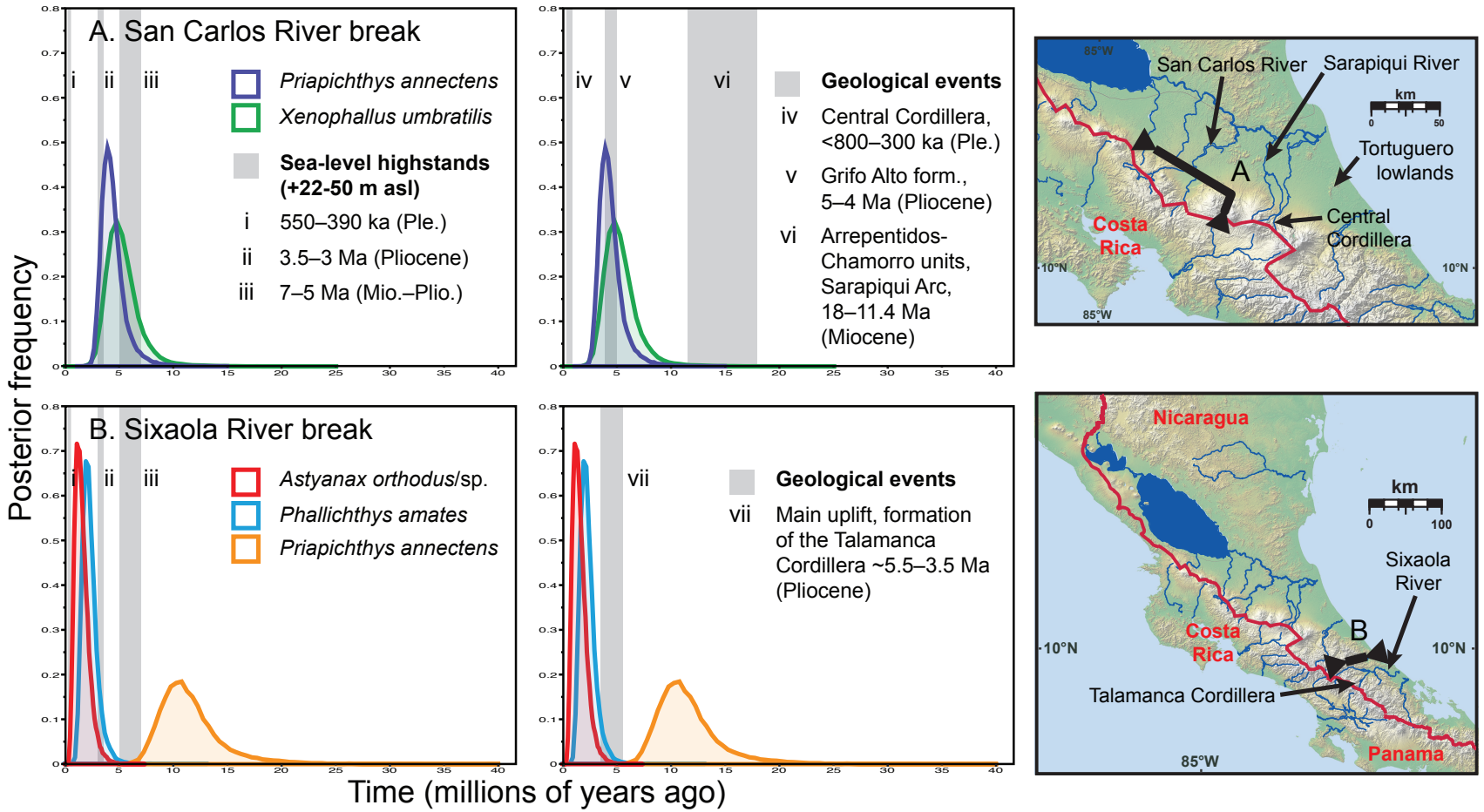
A. San Carlos River break
i ii iii
Priapichthys annectens
Xenophallus umbratilis
Sea-level highstands (+22-50 m asl)
i 550–390 ka (Ple.)
ii 3.5–3 Ma (Pliocene)
iii 7–5 Ma (Mio.–Plio.)
iv v vi
Geological events
iv Central Cordillera, <800–300 ka (Ple.)
v Grifo Alto form., 5–4 Ma (Pliocene)
vi Arrepentidos-Chamorro units, Sarapiqui Arc, 18–11.4 Ma (Miocene)
B. Sixaola River break
i ii iii
Astyanax orthodus/sp.
Phallichthys amates
Priapichthys annectens
vii
Geological events
vii Main uplift, formation of the Talamanca Cordillera ~5.5–3.5 Ma (Pliocene)
Posterior frequency
Time (millions of years ago)
San Carlos River
Sarapiqui River
Tortuguero lowlands
Central Cordillera
Costa Rica
Nicaragua
Sixaola River
Talamanca Cordillera
Panama
km
85°W
10°N

Figure 4

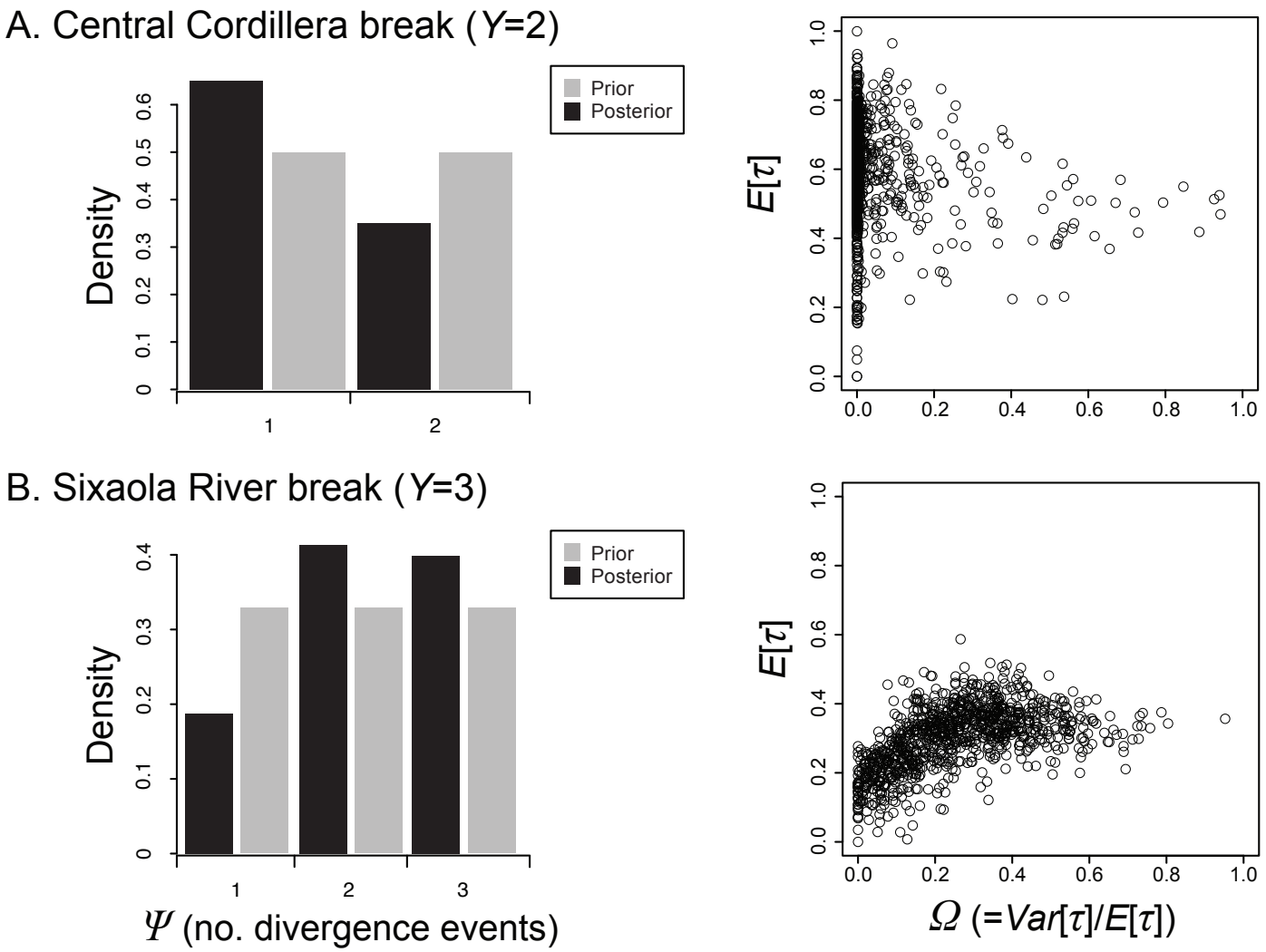
A. Central Cordillera break (Y=2)
Prior
Posterior
Density
0
0.1
0.2
0.3
0.4
0.5
0.6
1
2
E[τ]
0.0
0.2
0.4
0.6
0.8
1.0
B. Sixaola River break (Y=3)
Prior
Posterior
Density
0
0.1
0.2
0.3
0.4
1
2
3
Ψ (no. divergence events)
E[τ]
0.0
0.2
0.4
0.6
0.8
1.0
Ω (=Var[τ]/E[τ])

Figure 5

Chapter 4 – Supplementary Material

Data S1 Sampling localities and GenBank numbers for each lineage sampled.

Sampling localities for *Amatitlania spp*.

Species	Site No.	Locality	Country	Latitude	Longitude	Cyt*b* haplotype (sample IDs)
Amatitlania coatepeque	35	Lago Coatepeque	El Salvador	13.88156	-89.52486	Hap 31 (LSUMZ-F 2316–19, LSUMZ-F 2322)
Amatitlania kanna	91	Rio Sixaola	Costa Rica	9.59872	-82.80247	Hap 106 (STRI 209), 107 (STRI 210)
A. kanna	93	Rio Changuinola	Panama	9.42267	-82.39791	Hap 109 (STRI 291)
A. kanna	94	Rio Bongie	Panama	9.35990	-82.61000	Hap 110 (STRI 2677)
***A. kanna* ("*Archocentrus panamensis*")**	92	Rio Mandinga	Panama	9.46995	-79.12415	Hap 108 (STRI 1658)
***A. kanna* ("*Archocentrus panamensis*")**	95	Rio Parti	Panama	9.05600	-78.65950	Hap 111 (STRI 913)
***A. kanna* ("*Archocentrus panamensis*")**	96	Rio Róbalo	Panama	9.04056	-82.28583	Hap 112 (STRI 6861)
***A. kanna* ("*Archocentrus* sp. Aff. *Panamensis*")**	97	Rio Ukupti	Panama	8.81000	-77.74000	Hap 113 (STRI 3901)
Amatitlania nigrofasciata	16	Chalatenango, Municipio Citela, Departamento Chalatenango	El Salvador	14.36911	-89.21261	Hap 15 (LSUMZ-F 2548)
A. nigrofasciata	19	Laguna Metapan, Turicentro Las Flores, Departamento Santa Ana	El Salvador	14.31567	-89.45919	Hap 15 (LSUMZ-F 2492), Hap 18 (LSUMZ-F 2491)
A. nigrofasciata	30	Lago Güija	El Salvador	14.24792	-89.48417	Hap 15 (LSUMZ-F 2479), Hap 23 (LSUMZ-F 2482), Hap 24 (LSUMZ-F 2480), Hap 25 (LSUMZ-F 2478), Hap 26 (LSUMZ-F 2481)
A. nigrofasciata	31	Rio Tilapa, at Municipio Tegutle, Departamento Chalatenango	El Salvador	14.17969	-89.08911	Hap 27 (LSUMZ-F 2558)
A. nigrofasciata	36	under bridge in Santa Emilia, Departamento Sonsonate	El Salvador	13.66947	-89.75775	Hap 32 (LSUMZ-F 2296), Hap 33 (LSUMZ-F 2293–2295, LSUMZ-F 2297)
A. nigrofasciata	1	Rio Monga at Saba	Honduras	15.51647	-86.23713	Hap 1 (H06, H08, H16), Hap 2 (H03), Hap 3 (H09, H10), Hap 4 (H01), Hap 5 (H05, H07)
A. nigrofasciata	2	Rio Monga upstream	Honduras	15.49786	-86.22645	Hap 1 (H22, H23, H26, H28, H30, H34, H38, H39, H46), Hap 6 (H27)
A. nigrofasciata	3	Rio Agalteca	Honduras	15.48326	-86.66544	Hap 1 (H77, H78, H80, H83, H84?, H85, H88, H90), Hap 7 (H76)
A. nigrofasciata	4	Rio Chupa	Honduras	15.46569	-86.55347	Hap 1 (H104, H108, H99), Hap 3 (H102), Hap 8 (H96–H98, H103, H105, H107)
A. nigrofasciata	5	Rio Jaguaca	Honduras	15.44734	-86.37263	Hap 1 (H63, H64, H68, H71, H73, H74), Hap 3 (H65, H66), Hap 9 (H72)
A. nigrofasciata	6	Rio San Francisco at Carretera Saba-Yoro, ~1 km southwest of El Juncal and ~18 km east of Olanchito	Honduras	15.43657	-86.42625	Hap 1 (H44, H49, H53, H55, H57, H58, H60), Hap 3 (H42), Hap 9 (H59)

A. nigrofasciata	7	Creek near El Zamorano, trib. to Rio Guayape, Departamento Francisco Morazan	Honduras	14.80848	-86.98155	Hap 10 (H301, H302)
A. nigrofasciata	32	Rio Los Almendros	Honduras	14.06847	-86.34785	Hap 11 (H154, H157, H161, H162, H165, H167), Hap 28 (H149)
A. nigrofasciata	34	Piedra Ancha approximately 4.5 km northwest of Danli	Honduras	14.05522	-86.59303	Hap 11 (H122, H124, H127, H131, H133), Hap 29 (H145), Hap 30 (H144)
Amatitlania siquia	68	Rio Sabalo	Costa Rica	11.04283	-85.48922	Hap 74 (MLBM 174605–07)
A. siquia	69	Rio Frio en Los Chiles	Costa Rica	11.03004	-84.71821	Hap 35 (MLBM 160874), Hap 74 (MLBM 160875), Hap 75 (MLBM 160873), Hap 76 (MLBM 160876)
A. siquia	70	Rio Pizote	Costa Rica	10.90839	-85.21126	Hap 35 (STRI 2144), Hap 77 (STRI 2145)
A. siquia	71	Trib. to Rio Medio Queso on dirt road 2.5 km east of Costa Rica Hwy 35, approximately 25 km south-southeast of Los Chiles	Costa Rica	10.84717	-84.59366	Hap 35 (MLBM 160985–91)
A. siquia	72	Rio Tempisquito	Costa Rica	10.81467	-85.54390	Hap 78 (MLBM 160580), Hap 79 (MLBM 160581), Hap 80 (MLBM 160582, MLBM 160585), Hap 81 (MLBM 160583), Hap 82 (MLBM 160584), Hap 83 (MLBM 160586)
A. siquia	73	Rio Tempisquito	Costa Rica	10.78553	-85.55441	Hap 80 (MLBM 166475), Hap 81 (MLBM 166502), Hap 84 (MLBM 166471), Hap 85 (MLBM 166472), Hap 86 (MLBM 166473), Hap 87 (MLBM 166474)
A. siquia	74	Unnamed trib. to Rio Zapote near Bijagua	Costa Rica	10.73145	-85.05530	Hap 74 (MLBM 174616–17), Hap 88 (MLBM 174615)
A. siquia	75	Rio Celeste at Viejo Oeste Bar, just off Costa Rica Hwy 4, approximately 2.5 km southeast of Katira	Costa Rica	10.73125	-84.88818	Hap 75 (MLBM 160703–04)
A. siquia	76	Trib. to Rio Bijagua	Costa Rica	10.72422	-85.06640	Hap 74 (MLBM 174621)
A. siquia	77	Rio Irigaray in Irigaray at CA 1, approximately 2.5 km west of Canas Dulces	Costa Rica	10.72340	-85.51038	Hap 43 (MLBM 174611–12)
A. siquia	78	Quebrada Arena, trib. to Rio Blanco, near south side of Fortuna	Costa Rica	10.67217	-85.19942	Hap 89 (MLBM 161136)
A. siquia	79	Rio Gata, 2 km off road between Cañas and Upala	Costa Rica	10.63925	-85.08725	Hap 90 (MLBM 174608)
A. siquia	80	Rio Liberia at dirt road on outskirts of Liberia	Costa Rica	10.62745	-85.43412	Hap 91 (MLBM 166477), Hap 92 (MLBM 166478-83)
A. siquia	81	Trib. to Rio Infernito	Costa Rica	10.61802	-84.48418	Hap 93 (MLBM 161128–30)
A. siquia	82	Trib. to Rio Infernito	Costa Rica	10.61802	-84.48418	Hap 93 (MLBM 167239–41)
A. siquia	83	Trib. to Rio Toro north of Golfito	Costa Rica	10.59927	-84.06908	Hap 94 (MLBM 161114), Hap 95 (MLBM 161115)
A. siquia	84	Rio Salto	Costa Rica	10.56106	-85.39192	Hap 91 (MLBM 159045), Hap 96 (MLBM 159046)
A. siquia	85	Rio Cabuyo at unnamed province road to the Reserva	Costa Rica	10.48961	-85.38555	Hap 97 (MLBM 168039), Hap 98 (MLBM 168040), Hap 99 (MLBM 168041), Hap 100 (MLBM 168042), Hap 101 (MLBM 168043–46)

		Biologica Lomas Bardudal				
A. siquia	86	Rio Carrisal	Costa Rica	10.39502	-85.58688	Hap 90 (MLBM 174618)
A. siquia	87	Rio Lajas	Costa Rica	10.30552	-85.05733	Hap 90 (MLBM 174610)
A. siquia	88	Rio Canas	Costa Rica	10.28452	-85.08912	Hap 102 (MLBM 161731)
A. siquia	89	Trib. to Brazo del Sucio, near where it drains into Rio Chirripó, just off Carretera Braulio Carillo (Costa Rica Hwy 4), approximately 15 km west of Guapiles	Costa Rica	10.21872	-83.90466	Hap 35 (MLBM 160278–80), Hap 103 (MLBM 160277)
A. siquia	90	Rio San Rafael, trib. to Rio Jesus Maria	Costa Rica	9.97428	-84.57959	Hap 104 (STRI 2106), Hap 105 (STRI 2107)
A. siquia	27	Lago Güija	El Salvador	14.27360	-89.52613	Hap 22 (MNCN 184828)
A. siquia	8	Rio Patuca, Reserva Biologica Tawhka	Honduras	14.79288	-85.19438	Hap 11 (LSUMZ-F 4296)
A. siquia	9	Rio Patuca, Reserva Biologica Tawhka	Honduras	14.73172	-85.23652	Hap 11 (LSUMZ-F 4311)
A. siquia	10	Rio Cuyamel, trib. to Rio Patuca, Reserva Biologica Tawhka	Honduras	14.64152	-85.32090	Hap 11 (LSUMZ-F 4330)
A. siquia	11	Rio Jalan, trib. to Rio Patuca, approximately 1 km west of Bijagual	Honduras	14.55003	-86.22005	Hap 12 (LSUMZ-F 4180)
A. siquia	12	Rio Patuca	Honduras	14.49652	-86.16668	Hap 13 (LSUMZ-F 4235)
A. siquia	13	Rio Patuca	Honduras	14.43980	-86.05311	Hap 11 (LSUMZ-F 4049)
A. siquia	14	Rio Patuca	Honduras	14.43534	-86.10938	Hap 11 (LSUMZ-F 4247)
A. siquia	15	Rio Patuca	Honduras	14.41638	-86.05173	Hap 11 (LSUMZ-F 3941), Hap 14 (LSUMZ-F 3992)
A. siquia	17	Rio Patuca	Honduras	14.35073	-85.88889	Hap 16 (LSUMZ-F 4284)
A. siquia	18	Rio Patuca	Honduras	14.34202	-85.78539	Hap 17 (LSUMZ-F 4097)
A. siquia	20	Rio Patuca	Honduras	14.31537	-85.80143	Hap 11 (LSUMZ-F 4110)
A. siquia	21	Rio Patuca	Honduras	14.30967	-86.17146	Hap 13 (LSUMZ-F 4220)
A. siquia	26	Rio Patuca	Honduras	14.28901	-86.12000	Hap 21 (LSUMZ-F 4005)
A. siquia	28	Rio Patuca	Honduras	14.25099	-86.16664	Hap 11 (LSUMZ-F 4144)
A. siquia	29	Rio Patuca	Honduras	14.25036	-86.16634	Hap 11 (LSUMZ-F 4178)
A. siquia	33	Rio Patuca	Honduras	14.06795	-86.34697	Hap 11 (LSUMZ-F 4128)
A. siquia	22	Creek at community of Samilaya, Departamento RAAN	Nicaragua	14.30911	-83.71714	Hap 11 (LSUMZ-F 2801–03), Hap 19 (LSUMZ-F 2804), Hap 20 (LSUMZ-F 2805)

A. siquia	23	Creek at bridge close to Auyatara, Departamento RAAN	Nicaragua	14.30511	-83.63761	Hap 11 (LSUMZ-F 2839–42)
A. siquia	24	Creek at bridge outside community of Samilaya, Departamento RAAN	Nicaragua	14.30297	-83.70406	Hap 20 (LSUMZ-F 2814)
A. siquia	25	Creek at bridge outside community of Panua, Departamento RAAN	Nicaragua	14.30242	-83.67614	Hap 11 (LSUMZ-F 2829–30)
A. siquia	37	Unnamed trib. to Rio Las Vallas at stream just west of km marker 226; this is a trib. of Rio Yaosca, which drains into Rio Tuma, a major tributary of Rio Grande de Matagalpa	Nicaragua	13.25769	-85.45440	Hap 10 (MLBM 166527–28)
A. siquia	38	Rio Sinecapa, trib. to Lago Managua, just off Nicaragua Hwy 26 at La Empalme, located between Estelí and León (approximately 30–40 km west of Estelí)	Nicaragua	12.67376	-86.42606	Hap 34 (MLBM 166494), Hap 35 (MLBM 166495, MLBM 116496, MLBM 116498–116500), Hap 36 (MLBM 166497, MLBM 166501)
A. siquia	39	Unnamed trib. to Rio Grande de Matagalpa, west of La Mora and slightly further west of La Dalia	Nicaragua	13.22058	-85.72626	Hap 37 (MLBM 166529–33)
A. siquia	40	Unnamed trib. to Rio Estelí, thus a trib. to Rio Coco	Nicaragua	13.05866	-86.35114	Hap 38 (MLBM 166488–93)
A. siquia	41	Unnamed trib. to Rio Tuma east of Rio Blanco (town), Matagalpa province	Nicaragua	12.98351	-85.13827	Hap 10 (MLBM 166507–09), Hap 39 (MLBM 166504–06, MLBM 166534), Hap 40 (MLBM 166503),
A. siquia	42	Trib. to Rio Grande at La Trinidad	Nicaragua	12.97132	-86.23720	Hap 35 (MLBM 166516), Hap 38 (MLBM 166487, MLBM 166510), Hap 41 (MLBM 166511–15, MLBM 166517–18)
A. siquia	43	Unnamed trib. to Rio Tuma northwest of Rio Blanco (town; flowing from Mt. Musun)	Nicaragua	12.93613	-85.23434	Hap 10 (MLBM 168886, N/A), Hap 42 (MLBM 166519–21)
A. siquia	44	Trib. to unnamed trib. of the Rio Grande de Matagalpa, approximately 32 km west of Rio Blanco (town) on road between Matagalpa and Rio Blanco (town)	Nicaragua	12.82341	-85.44279	Hap 43 (MLBM 166522), Hap 44 (MLBM 166523–24, MLBM 166526), Hap 45 (MLBM 166525)
A. siquia	45	Trib. to Rio Olama, trib. to Rio Grande de Matagalpa, at Puente de Tierra Azul on road to Rio Blanco (Nicaragua Hwy 9)	Nicaragua	12.68476	-85.54708	Hap 46 (MLBM 168884–85)
A. siquia	46	Trib. to Rio Malacatoya in Teustepe just off road to Rama (Nicaragua Hwy 7), southwest of Boaco and approximately 9 km northwest of Laguna Presa	Nicaragua	12.41825	-85.79299	Hap 35 (MLBM 166476)
A. siquia	47	Rio Caracol at Nicaragua Hwy 7, trib. to Laguna Presa, the lake	Nicaragua	12.35116	-85.88870	Hap 47 (MLBM 172998)

		formed by the damming of Rio Malacatoya				
A. siquia	49	Rio Tecolostote at Nicaragua Hwy 7 approximately 2 km north of Tecolostote	Nicaragua	12.26707	-85.65119	Hap 35 (MLBM 173083), Hap 49 (MLBM 173084)
A. siquia	50	Lago Xiloa	Nicaragua	12.21392	-86.31652	Hap 35 (1000009–11, 100014, 100016), Hap 50 (1000012), Hap 51 (1000015)
A. siquia	51	Lago Jiloa	Nicaragua	12.21381	-86.31597	Hap 35 (MLBM 172646)
A. siquia	52	Lago Xiloa at principal entrance, Departamento Managua	Nicaragua	12.21136	-86.31692	Hap 35 (LSUMZ-F 2858, LSUMZ-F 2860), Hap 52 (LSUMZ-F 2862), Hap 53 (LSUMZ-F 2859)
A. siquia	53	Rio Tipitapa at CA 1 in Tipitapa	Nicaragua	12.20267	-86.10208	Hap 35 (MLBM 172832), Hap 55 (MLBM 172834), Hap 57 (MLBM 172831)
A. siquia	54	Entre Trapicne at edge of Lago Managua on north highway, Departamento Managua	Nicaragua	12.17631	-86.11278	Hap 36 (LSUMZ-F 2982)
A. siquia	55	Unnamed trib. at km marker 18.5 (from center of Managua) on CA1, just southwest of town of Tipitapa	Nicaragua	12.17289	-86.11618	Hap 35 (MLBM 168035, MLBM 168871)
A. siquia	56	Unnamed trib./headwater of Rio Mico northeast of San Pedro de Lovago	Nicaragua	12.13630	-85.04597	Hap 35 (MLBM 168877, MLBM 168879–82), Hap 59 (MLBM 168878, MLBM 158883)
A. siquia	57	Trib. to Rio Acoyapa at Nicaragua Hwy 7 approximately 3.2 km west of Santo Tomas	Nicaragua	12.05056	-85.12458	Hap 35 (MLBM 173798)
A. siquia	58	Unnamed trib. to Rio Acoyapa at Nicaragua Hwy 7 approximately 1.8 km west of Lovago	Nicaragua	12.00340	-85.18153	Hap 60 (MLBM 173789), Hap 61 (MLBM 173790), Hap 62 (MLBM 173791), Hap 63 (MLBM 173792)
A. siquia	59	Lago Nicaragua at El Rayo at the Rancho Azul Restaurant, Departamento Granada	Nicaragua	11.88250	-85.89642	Hap 35 (LSUMZ-F 2935–37, LSUMZ-F 2939), Hap 64 (LSUMZ-F 2934), Hap 65 (LSUMZ-F 2938)
A. siquia	60	Unnamed trib. to Rio Rama at Nicaragua Hwy 71, approximately 1.5 km south of Colonia El Corocito	Nicaragua	11.74923	-84.55819	Hap 66 (MLBM 173686)
A. siquia	61	Unnamed trib. to Rio Rama at Nicaragua Hwy 71 approximately 7.5 km northwest of Nueva Guinea	Nicaragua	11.73855	-84.51068	Hap 35 (MLBM 173735–36, MLBM 173740–42), Hap 67 (MLBM 173737, MLBM 173739), Hap 68 (MLBM 173738)
A. siquia	62	Unnamed trib. to Rio Zapote (trib to Rio Plata, which drains into Rio Rama) south of Nueva Guinea, approximately 1 km south of end of Nicaragua Hwy 71	Nicaragua	11.68127	-84.46019	Hap 69 (MLBM 173444–45)
A. siquia	63	Unnamed trib to Rio Ochomogo 300 m west of Nicaragua Hwy 1	Nicaragua	11.67886	-85.98817	Hap 35 (MLBM 172547)
A. siquia	64	Rio Ochomogo 300 m west of Nicaragua Hwy 1, near Paso	Nicaragua	11.65664	-85.97319	Hap 35 (MLBM 172545), Hap 70 (MLBM 172544)

		Real de Ochomogo				
A. siquia	65	Unnamed trib. to Lago Nicaragua just east of mile marker km 238 on road to San Miguelito, Chontales	Nicaragua	11.50538	-84.83956	Hap 35 (MLBM 168507, MLBM 168874–76)
A. siquia	66	Ometepe spring, Isla de Ometepe	Nicaragua	11.49692	-85.54821	Hap 35 (1000008), Hap 71 (1000001, 1000003, 1000004, 1000005), Hap 72 (1000002, 1000006, 1000007)
A. siquia	67	Lago Nicaragua at San Miguelito	Nicaragua	11.40046	-84.90489	Hap 35 (MLBM 168872), Hap 73 (MLBM 168873)
***Amatitlania centrarchus* (OG)**	48	Rio Malacatoya	Nicaragua	12.32661	-85.95553	Hap 48 (MLBM 172837)
Amatitlania siquia	53	Rio Tipitapa at CA 1 in Tipitapa	Nicaragua	12.20267	-86.10208	Hap 54 (MLBM 172829), Hap 56 (MLBM 172830), Hap 58 (MLBM 172833)

Abbreviations: "H" under "Cyt*b* haplotype (sample IDs)", field numbers for samples collected by Wilfredo Matamoros and/or Michael Tobler in Honduras; LSUMZ-F, LSU Museum of Natural Science tissue catalog number; MLBM, Brigham Young University Fish Collection (Monte L. Bean Museum); N/A, not available; OG, outgroup; sp., species; STRI, Smithsonian Tropical Research Institute Neotropical Fish Collection; trib., tributary.

Sampling localities for *Priapichthys annectens*

Species	Site No.	Locality	Country	Latitude	Longitude	Cyt*b* haplotype (sample IDs)
Priapichthys annectens	1	Unnamed trib. to Rio Haciendas at Costa Rica Hwy 4, approximately 7.5 km east of Santa Cecilia	Costa Rica	11.01920	-85.35766	Hap 31 (MLBM 160519, MLBM 160521, MLBM 160524, MLBM 160526), Hap 32 (MLBM 160523, MLBM 160525), Hap 36 (MLBM 160520, MLBM 160522)
P. annectens	2	Quebrada El Carmen, trib. to unnamed trib. of Rio Niño (Rio Pizote), at Costa Rica Rd 164 approximately 6 km south of Aguas Claras	Costa Rica	10.76972	-85.18712	Hap 33 (MLBM 170065–66), Hap 35 (MLBM 170061, MLBM 170067), Hap 37 (MLBM 170062), Hap 39 (MLBM 170063, MLBM 170064, MLBM 170068)
P. annectens	3	Trib. to south Fork of Rio Negro, at Costa Rica Rd 164, approximately 10 km north of La Fortuna, near "Guayabal"	Costa Rica	10.76277	-85.19325	Hap 29 (MLBM 154680–82, MLBM 154684, MLBM 154685, MLBM 154687), Hap 34 (MLBM 154683, MLBM 154686)
P. annectens	4	Upper Rio Colorado on dirt road just east of entrance to Rincon de la Vieja park	Costa Rica	10.76134	-85.35030	Hap 11 (MLBM 155610), Hap 17 (no ID, cllection 97-21, jar 30.2), Hap 24 (MLBM 155611), Hap 25 (MLBM 155614), Hap 28 (MLBM 155613), Hap 30 (no ID, collection 97-21)
P. annectens	5	Unnamed trib. to Rio Zapote near Bijagua ("Rio Bijagua")	Costa Rica	10.73145	-85.05530	Hap 1 (MLBM 153951–55)
P. annectens	6	Unnamed trib. to Rio Zapote near south side of Bijagua	Costa Rica	10.73142	-85.05535	Hap 1 (MLBM 148981, 148982)
P. annectens	7	Quebrada Peru at Costa Rica Hwy 6 approximately 4-5 km southwest of Bijagua	Costa Rica	10.70353	-85.08002	Hap 1 (MLBM 153938, MLBM 153939)
P. annectens	8	Quebrada Hormiguero at Costa Rica Hwy 6 approximately 5.5 km south of Bijagua	Costa Rica	10.69090	-85.08365	Hap 1 (MLBM 154897–900)
P. annectens	9	Rio Esquivetto, 2 km east of Costa Rica Hwy 6, off dirt road ~4.5 km south of Bijagua	Costa Rica	10.68718	-85.06670	Hap 1 (MLBM 154584, MLBM 154586–91), Hap 3 (MLBM 154585), Hap 41 (MLBM 154583)

P. annectens	10	Edwin's irrigation canal, trib. to Rio Tenorio, about 1 km east of Costa Rica Hwy 6 near Rio Naranjo	Costa Rica	10.68422	-85.07530	Hap 4 (MLBM 149150–52)
P. annectens	11	Unnamed trib. to Rio Corobici off dirt road approximately 700 m north of Tierras Morenas, west of Lago Arenal	Costa Rica	10.58198	-85.01952	Hap 2 (Pa1, Pa2)
P. annectens	12	Rio Quebrada guape, trib. to Lago Arenal, approximately 300 m north of Costa Rica Rd 142 near Establo Arenal	Costa Rica	10.56325	-84.92085	Hap 6 (Pa3)
P. annectens	13	River at CA 1 (Carretera Panamericana), just ~1 km west of Bagaces	Costa Rica	10.53117	-85.28066	Hap 29 (Pa15)
P. annectens	14	Unknown tributary to Lago Arenal near Hotel La Mansion Inn Arenal, approximately 30 km northeast of Tilarán on road to Fortuna (Costa Rica Rd 142)	Costa Rica	10.49220	-84.83580	Hap 26 (MLBM 154839, MLBM 154841)
P. annectens	15	Trib. to Rio Agua Caliente southwest of Volcan Arenal	Costa Rica	10.43505	-84.72340	Hap 40 (MLBM 149138)
P. annectens	16	La Vuelta del Borracho, trib. to Lago Arenal on El Fosforo-El Castillo Rd, southwest of Volcan Arenal	Costa Rica	10.42743	-84.75225	Hap 1 (MLBM 149212), Hap 5 (MLBM 149211)
P. annectens	17	Upper Rio Tortuguero 0.5 km off of Costa Rica Rd 249, 1 km north of San Rafael and approximately 5 km north of Guapiles	Costa Rica	10.25942	-83.81223	Hap 19 (Pa14)
P. annectens	18	Trib. to Brazo del Sucio, near where it drains into Rio Chirripó, just off Carretera Braulio Carillo (Costa Rica Hwy 4), approximately 15 km west of Guapiles	Costa Rica	10.21872	-83.90466	Hap 19 (MLBM 160297, MLBM 160298, MLBM 160302), Hap 22 (MLBM 160300), Hap 23 (MLBM 160301), Hap 38 (MLBM 160296)
P. annectens	19	Rio Corinto, trib. to Rio Chirripó, at Carretera Braulio Carillo (Costa Rica Hwy 32), approximately 7.5 km west of Guapiles	Costa Rica	10.21123	-83.88523	Hap 12 (MLBM 154011), Hap 20 (MLBM 154015), Hap 21 (MLBM 154017)
P. annectens	20	Trib. to Rio Parismina just off Costa Rica Hwy 32, approximately 4 km southeast of Guacimo	Costa Rica	10.19772	-83.65210	Hap 13 (MLBM 153924), Hap 14 (MLBM 153928), Hap 18 (MLBM 153926)
P. annectens	21	Rio Balsas at Costa Rica Rd 702, approximately 11–12 km north of San Ramon	Costa Rica	10.17295	-84.49815	Hap 7 (MLBM 161703, MLBM 161705–12), Hap 27 (MLBM 161704)
P. annectens	22	Headwater trib. to Rio Balsa at Costa Rica Rd 702 approximately 11 km north of San Ramon and only ~3-4 km north of Angeles Sur	Costa Rica	10.17207	-84.49802	Hap 1 (Pa8, Pa9), Hap 7 (Pa5), Hap 29 (Pa6, Pa7)
P. annectens	23	Trib. to Rio Sixaola	Costa Rica	9.62427	-82.83273	Hap 8 (MLBM 154895), Hap 9 (MLBM 154894), Hap 10 (MLBM 154892), Hap 15 (MLBM 154893), Hap 16 (MLBM 154896)

Abbreviations: MLBM, Brigham Young University Fish Collection (Monte L. Bean Museum); trib., tributary.

Sampling localities for *Alfaro cultratus*

Species	Site No.	Locality	Country	Latitude	Longitude	Cyt*b* haplotype (sample IDs)
Alfaro cultratus	1	Unnamed trib./headwater of Rio Mico northeast of San Pedro de Lovago	Nicaragua	12.13630	-85.04597	Hap 73 (MLBM 167995–98)
A. cultratus	2	Trib. to Rio Mayales	Nicaragua	12.05663	-85.40814	Hap 72 (MLBM 173303–10)
A. cultratus	3	Trib. to Rio Acoyapa	Nicaragua	12.05056	-85.12458	Hap 1 (MLBM 172431)
A. cultratus	4	Rio Mico	Nicaragua	12.02057	-84.64513	Hap 73 (MLBM 173416, MLBM 173418, MLBM 173419, MLBM 173422, MLBM 173424), Hap 74 (MLBM 173417, MLBM 173420, MLBM 173421, MLBM 173423)

A. cultratus	5	Lago Nicaragua at San Carlos, Departamento Rio San Juan	Nicaragua	11.12129	-84.77922	Hap 1 (FJ178773), Hap 2 (FJ178772)
A. cultratus	6	Unnamed trib. to Rio Rama at Nicaragua Hwy 71, approximately 13 km southeast of El Coral, just after turn-off for road to Talolinga	Nicaragua	11.82544	-84.60611	Hap 1 (MLBM 172423–30)
A. cultratus	7	Unnamed trib. to Rio Rama at Nicaragua Hwy 71, approximately 1.5 km south of Colonia El Corocito	Nicaragua	11.74923	-84.55819	Hap 1 (MLBM 173550, MLBM 173552), Hap 81 (MLBM 173549, MLBM 173553, MLBM 173554, MLBM 173556, MLBM 173557), Hap 82 (MLBM 173551)
A. cultratus	8	Unnamed trib. to Rio Rama at Nicaragua Hwy 71 approximately 7.5 km northwest of Nueva Guinea	Nicaragua	11.73855	-84.51068	Hap 77 (MLBM 173559, MLBM 173560, MLBM 173562), Hap 78 (MLBM 173561), Hap 79 (MLBM 173563, MLBM 173564, MLBM 173566), Hap 80 (MLBM 173565)
A. cultratus	9	Unnamed trib. south of Nueva Guinea	Nicaragua	11.68127	-84.46019	Hap 73 (MLBM 173454), Hap 75 (MLBM 173447–51, MLBM 173453), Hap 76 (MLBM 173452)
A. cultratus	10	Rio El Monje, trib. to Lago Managua	Nicaragua	11.63330	-86.30000	Hap 3 (FJ178774)
A. cultratus	11	Unnamed trib. to Lago Nicaragua just east of mile marker km 238 on road to San Miguelito, Chontales	Nicaragua	11.50538	-84.83956	Hap 66 (MLBM 167994), Hap 83 (MLBM 167991, MLBM 167993), Hap 84 (MLBM 167992)
A. cultratus	12	Rio Sapoa	Costa Rica	11.04437	-85.61590	Hap 1 (AcSAPF.19, AcSAPF.25, MLBM 147267, MLBM 147753, MLBM 147754, MLBM 147763, MLBM 147783, MLBM 147785, MLBM 147788), Hap 4 (AcSAPF.16), Hap 5 (AcSAPF.17), Hap 6 (AcSAPF.18), Hap 7 (AcSAPF.20, AcSAPF.24), Hap 8 (AcSAPF.26), Hap 9 (MLBM 147786, MLBM 147787), Hap 10 (Ac0376)
A. cultratus	13	Rio Sabalo	Costa Rica	11.04283	-85.48922	Hap 20 (AcSABF.1), Hap 21 (AcSABF.2, AcSABF.4), Hap 22 (AcSABF.3)
A. cultratus	14	Rio Salto, trib. to Rio Zapote, at Costa Rica Hwy 6, approximately 10 km north of Bijagua	Costa Rica	10.79823	-85.02327	Hap 11 (AcSALF.1), Hap 12 (AcSALF.2), Hap 13 (AcSALF.3), Hap 14 (AcSALF.11, AcSALF.4, AcSALF.5, MLBM 147151, MLBM 147152, MLBM 147155, MLBM 147156), Hap 15 (AcSALF.10, AcSALF.6, AcSALF.7, MLBM 147150, MLBM 147153, MLBM 147154), Hap 16 (AcSALF.8), Hap 17 (AcSALF.9), Hap 18 (AcSALF.13), Hap 19 (MLBM 146311)
A. cultratus	15	Rio Celeste at Viejo Oeste Bar, just off Costa Rica Hwy 4, approximately 2.5 km southeast of Katira	Costa Rica	10.73125	-84.88818	Hap 66 (MLBM 160710, MLBM 160711)
A. cultratus	16	Rio Chimurria at Costa Rica Hwy 35, near Santa Clara de Upala, approximately 25 km north of Boca Arenal	Costa Rica	10.72740	-84.55823	Hap 4 (AcCHIF.3, AcCHIF.4), Hap 19 (AcCHIF.7, MLBM 148624), Hap 25 (AcCHIF.1), Hap 26 (AcCHIF.2), Hap 27 (AcCHIF.6), Hap 28 (AcCHIF.9), Hap 29 (MLBM 148278), Hap 30 (MLBM 147281–84, MLBM, 148183), Hap 31 (MLBM 148265), Hap 32 (MLBM 147285), Hap 33 (MLBM 147286)
A. cultratus	17	Rio Venado at Costa Rica Rd 143, approximately 2.5 km south of San Rafael de Guatuso	Costa Rica	10.64482	-84.82223	Hap 23 (MLBM 147255–58, MLBM 148184, MLBM 148185), Hap 24 (MLBM 147259)
A. cultratus	18	Trib. to R. Toro north Golfito	Costa Rica	10.59927	-84.06908	Hap 68 (MLBM 161078, MLBM 161079)
A. cultratus	19	Rio Tempisque on road between Guardia and Comunidad, Nicoya Peninsula	Costa Rica	10.57220	-85.58477	Hap 1 (MLBM 167979–82)
A. cultratus	20	Trib. to Rio Sarapiquí on road ~2 km off end-Costa Rica Rd. 505, approximately 9 km northwest of	Costa Rica	10.52452	-84.03127	Hap 19 (AcSARF.7, MLBM 147625, MLBM 147627, MLBM 147629, MLBM 147795), Hap 34 (MLBM 147798), Hap 37 (AcSARF.3, AcSARF.4, MLBM 147624, MLBM 147793, MLBM 147796), Hap 41 (AcSARF.5), Hap 42 (AcSARF.6), Hap 43

		Puerto Viejo de Sarapiquí				(AcSARF.8), Hap 44 (MLBM 147628), Hap 45 (MLBM 147623, MLBM 147794), Hap 46 (MLBM 147626, MLBM 147792), Hap 47 (MLBM 147797)
A. cultratus	21	Rio Cabuyo at unnamed province road to the Reserva Biologica Lomas Bardudal	Costa Rica	10.48961	-85.38555	Hap 1 (MLBM 167984–86, MLBM 170081), Hap 71 (MLBM 167983, MLBM 170082, MLBM 170083)
A. cultratus	22	Quebrada Perez	Costa Rica	10.47352	-84.82225	Hap 4 (AcPERF.1)
A. cultratus	23	Rio Cuarto (or other trib. to R. Toro) in between Santa Isabel and Santa Rafael	Costa Rica	10.46667	-84.21667	Hap 67 (MLBM 160228–30, MLBM 160232, MLBM 160233)
A. cultratus	24	Isla Grande	Costa Rica	10.39302	-83.96817	Hap 19 (AcISLF.7, MLBM 146312–16), Hap 21 (AcISLF.5), Hap 22 (AcISLF.1), Hap 44 (AcISLF.6, AcISLF.8), Hap 48 (AcISLF.2–4), Hap 49 (MLBM 146317)
A. cultratus	25	Quebrada Piecueca	Costa Rica	10.38607	-84.57897	Hap 34 (MLBM 147744–49)
A. cultratus	26	Rio Cano Grande	Costa Rica	10.37278	-84.27823	Hap 36 (MLBM 146786, MLBM 146787, MLBM 146789, MLBM 146790)
A. cultratus	27	Upper Rio Tortuguero 0.5 km off of Costa Rica Rd 249, 1 km north of San Rafael and approximately 5 km north of Guapiles	Costa Rica	10.25942	-83.81223	Hap 4 (AcTORS.4–7), Hap 37 (AcTORS.8), Hap 50 (AcTORS.1), Hap 51 (AcTORS.2), Hap 52 (AcTORS.3)
A. cultratus	28	Rio Corinto	Costa Rica	10.21185	-83.88647	Hap 14 (MLBM 147149), Hap 20 (AcCORF.1, AcCORF.2, AcCORF.4), Hap 35 (MLBM 147144–46, MLBM 147148), Hap 37 (MLBM 147143, MLBM 147147, MLBM 147150)
A. cultratus	29	Trib. to Rio Parismina 200 m off of Costa Rica Hwy 32, approximately 3 km southeast of Guacimo	Costa Rica	10.19968	-83.65873	Hap 4 (MLBM 147274–76, MLBM 147278–80), Hap 10 (AcPARS.7–10), Hap 19 (AcPARS.11), Hap 20 (AcPARS.3, AcPARS.4), Hap 53 (AcPARS.1, AcPARS.2), Hap 54 (MLBM 147277)
A. cultratus	30	Rio Herediana	Costa Rica	10.12417	-83.55617	Hap 10 (AcHERS.9, MLBM 147260–66, MLBM 147765), Hap 19 (AcHERS.2, AcHERS.3, AcHERS.5, AcHERS.6), Hap 55 (AcHERS.4), Hap 56 (AcHERS.7), Hap 57 (AcHERS.8)
A. cultratus	31	Rio Barbilla (major trib. to Rio Matina) at Costa Rica Rd 805 (dirt road) just west of B-Line, 0.5 km from Costa Rica Hwy 32	Costa Rica	10.04416	-83.33383	Hap 69 (MLBM 167988, MLBM 167989), Hap 70 (MLBM 167987, MLBM 167990)
A. cultratus	32	Rio Carbon	Costa Rica	9.62312	-82.85520	Hap 38 (MLBM 147789/90), Hap 39 (MLBM 147791), Hap 40 (MLBM 146785)
A. cultratus	33	Trib. to Rio Sixaola	Costa Rica	9.62093	-82.85768	Hap 19 (AcSIXS.8–11, AcSIXS.13), Hap 38 (MLBM 147755 / 147756), Hap 52 (AcSIXS.15), Hap 58 (MLBM 147757, MLBM 147759, MLBM 147761-62), Hap 59 (AcSIXS.12), Hap 60 (MLBM 147321), Hap 61 (MLBM 147320), Hap 62 (MLBM 147319), Hap 63 (MLBM 147322), Hap 64 (MLBM 147324), Hap 65 (MLBM 147758, MLBM 147760)
A. cultratus	N/A	see Hrbek *et al.* (2007)	Costa Rica	N/A	N/A	Hap 1 (EF017531)

Abbreviations: MLBM, Brigham Young University Fish Collection (Monte L. Bean Museum); N/A, not available; trib., tributary.

Sampling localities for *Xenophallus umbratilis*

Species	Site No.	Locality	Country	Latitude	Longitude	Cyt*b* haplotype (sample IDs)
Xenophallus umbratilis	1	Unnamed trib. to Lago Nicaragua just east of mile marker km 238 on rd to San Miguelito,	Nicaragua	11.50538	-84.83956	Hap 1 (MLBM 167968–75, MLBM 167977, MLBM 167978), Hap 2 (MLBM 167976)

		Departamento Chontales				
X. umbratilis	2	Unnamed trib. to Rio Zapote near Bijagua	Costa Rica	10.73145	-85.05530	Hap 3 (Xu0621H1, Xu0621H2, Xu0621H5–H8), Hap 4 (Xu0621H3, Xu0621H4)
X. umbratilis	3	Rio Bijagua in field approximately 1 km east of Costa Rica Hwy 6 at Bijagua	Costa Rica	10.72773	-85.03127	Hap 3 (Xu0620H2, Xu0620H4–H8), Hap 5 (Xu0620H1), Hap 6 (Xu0620H3)
X. umbratilis	4	Rio Chimurria at Costa Rica Hwy 35 approximately 25 km north of Boca Arenal	Costa Rica	10.72736	-84.55817	Hap 7 (Xu0635H1, Xu0635H3–H8), Hap 8 (Xu0635H2)
X. umbratilis	5	Trib. to Rio Bijagua at Costa Rica Hwy 6 approximately 1.4 km southwest of Bijagua	Costa Rica	10.72422	-85.06640	Hap 3 (Xu0625H1–H3, Xu0625H6–H8), Hap 4 (Xu0625H4, Xu0625H5)
X. umbratilis	6	Quebrada Hormiguero at Costa Rica Hwy 6 approximately 5.5 km south of Bijagua	Costa Rica	10.69090	-85.08365	Hap 9 (Xu0611H1, Xu0611H4, Xu0611H8), Hap 10 (Xu0611H2, Xu0611H3, Xu0611H6, Xu0611H7), Hap 11 (Xu0611H5)
X. umbratilis	7	Rio Tenorio 1 km east of Costa Rica Hwy 6, off dirt road 5.5 km south of Bijagua	Costa Rica	10.68808	-85.07602	Hap 9 (Xu0630H4, Xu0630H6), Hap 12 (Xu0630H1–H3, Xu0630H5, Xu0630H7, Xu0630H8)
X. umbratilis	8	Rio Esquivetto, 2 km east of Costa Rica Hwy 6, off dirt road ~4.5 km south of Bijagua	Costa Rica	10.68718	-85.06670	Hap 9 (Xu0629H2), Hap 12 (Xu0629H1, Xu0629H3–H7), Hap 13 (Xu0629H8)
X. umbratilis	9	Quebrada Arena, trib. to Rio Blanco, at site on south side of Fortuna	Costa Rica	10.67217	-85.19942	Hap 14 (Xu0638H1–H8)
X. umbratilis	10	Rio Las Flores 200 m east of Costa Rica Hwy 6 approximately 4.8 km south of Bijagua	Costa Rica	10.65778	-85.09037	Hap 9 (XuFlores1–3), Hap 12 (XuFlores2, XuFlores5, XuFlores6), Hap 15 (XuFlores4)
X. umbratilis	11	Rio Venado at Costa Rica Rd 143, approximately 2.5 km south of San Rafael de Guatuso	Costa Rica	10.64482	-84.82223	Hap 16 (Xu0719H1), Hap 17 (Xu0719H2, Xu0719H5, Xu0719H6)
X. umbratilis	12	Trib. to Lago Arenal approximately 5.5 km west of Nuevo Arenal, about 1 km southwest of Eco Lodge Hotel	Costa Rica	10.56030	-84.94030	Hap 18 (Xu0632H1–H3, Xu0632H7, Xu0632H8), Hap 19 (Xu0632H4–H6)
X. umbratilis	13	Lago Arenal	Costa Rica	10.55970	-84.96970	Hap 18 (Xu0607H3, Xu0607H5, Xu0607H6, Xu0607H8), Hap 20 (Xu0607H1, Xu0607H2), Hap 21 (Xu0607H4), Hap 22 (Xu0607H7)
X. umbratilis	14	Quebrada Jilguero, tributary to Lago Arenal west of Arenal 10.6 miles from Tilarán	Costa Rica	10.55514	-84.90354	Hap 18 (Xu9829H1, Xu9829H3–H8), Hap 20 (Xu9829H2)
X. umbratilis	15	Lago Arenal just off Costa Rica Rd 142 near intersection between 142 and Tierras Morenas, at end of bay west of Tico Wind Windsurf Center	Costa Rica	10.54860	-84.98080	Hap 18 (Xu0631H1, Xu0631H3–H8), Hap 23 (Xu0631H2)
X. umbratilis	16	Trib. to Rio Sarapiquí on road ~2 km off end-Costa Rica Rd. 505, approximately 9 km northwest of Puerto Viejo de Sarapiquí	Costa Rica	10.52452	-84.03127	Hap 24 (Xu0713H1)
X. umbratilis	17	Lago Arenal at Costa Rica Rd 142 approximately 9.5 km southeast of Nuevo Arenal	Costa Rica	10.50645	-84.84600	Hap 18 (Xu0633H1–H4)
X. umbratilis	18	Lago Arenal	Costa Rica	10.50140	-84.84060	Hap 18 (Xu0604H1–H3)
X. umbratilis	19	Unknown tributary to Lago Arenal near Hotel La Mansion Inn Arenal, approximately 30 km northeast of Tilarán on road to Fortuna (Costa Rica Rd 142)	Costa Rica	10.49220	-84.83580	Hap 18 (Xu9821H2, Xu9821H3, Xu9821H5–H8), Hap 20 (Xu9821H1), Hap 25 (Xu9821H4)
X. umbratilis	20	Lago Arenal	Costa Rica	10.47360	-84.82220	Hap 18 (Xu0634H1–H4)
X. umbratilis	21	Trib. to Rio Agua Caliente southwest of Volcan Arenal	Costa Rica	10.43505	-84.72340	Hap 18 (Xu0717H1–H3), Hap 26 (Xu0717H4)

X. umbratilis	22	La Vuelta Del Borracho, trib. to Lago Arenal on El Fosforo-El Castillo Rd, southwest of Volcan Arenal	Costa Rica	10.42743	-84.75225	Hap 18 (Xu0718H1–H4)
X. umbratilis	23	Rio Isla Grande, trib. to Rio Chirripó	Costa Rica	10.39300	-83.96820	Hap 27 (Xu0636H1, Xu0636H6–H9), Hap 28 (Xu0636H2), Hap 29 (Xu0636H3), Hap 30 (Xu0636H4)
X. umbratilis	24	Rio Cariari, trib. to Upper Rio Tortuguero on south side of Cariari, just off Costa Rica Rd 247	Costa Rica	10.35530	-83.73750	Hap 31 (XuCar1Hco)
X. umbratilis	25	Quebrada Piecueca (San Carlos) just off Costa Rica Rd 702, ~1.6 km northwest of Tigra	Costa Rica	10.35170	-84.58810	Hap 32 (XuTigra1, XuTigra2)
X. umbratilis	26	Rio Jiménez at Costa Rica Rd 248 just west of Villa Franca	Costa Rica	10.28940	-83.61000	Hap 33 (XuJimez1, XuJimez2)
X. umbratilis	27	Upper Rio Tortuguero 0.5 km off of Costa Rica Rd 249, 1 km north of San Rafael and approximately 5 km north of Guapiles	Costa Rica	10.25942	-83.81223	Hap 35 (Xu0712H1, Xu0712H2, Xu0712H6–H8), Hap 36 (Xu0712H3–H5)
X. umbratilis	28	Rio Corinto, trib. to Rio Chirripó	Costa Rica	10.21190	-83.88650	Hap 37 (Xu0637H1–H8)
X. umbratilis	29	Trib. to Rio Parismina just off Costa Rica Hwy 32, approximately 4 km southeast of Guacimo	Costa Rica	10.19772	-83.65210	Hap 33 (Xu0710H3, Xu0710H5), Hap 38 (Xu0710H1, Xu0710H4, Xu0710H6, Xu0710H7)
X. umbratilis	N/A	See Hrbek *et al.* 2007	Costa Rica	N/A	N/A	Hap 39 (Numb)

Abbreviations: MLBM, Brigham Young University Fish Collection (Monte L. Bean Museum); N/A, not available; trib., tributary.

Sampling localities for *Phallichthys amates*

Species	Site No.	Locality	Country	Latitude	Longitude	Cyt*b* haplotype (sample IDs)
Phallichthys amates	9	Trib. to Rio Medio Queso on dirt road 2.5 km east of Costa Rica Hwy 35, approximately 25 km south-southeast of Los Chiles	Costa Rica	10.84717	-84.59366	Hap 7 (MLBM 161015, MLBM 161017, MLBM 161018, MLBM 161023, MLBM 161025, MLBM 161026), Hap 8 (MLBM 161016, MLBM 161021, MLBM 161024), Hap 9 (MLBM 161020), Hap 10 (MLBM 161022)
P. amates	10	Unnamed trib. to Rio Pizote at Costa Rica Rd 164 approximately 6 km south of Aguas Claras	Costa Rica	10.76972	-85.18712	Hap 11 (MLBM 161133), Hap 12 (MLBM 161134)
P. amates	11	Unnamed trib. to Rio Zapote near south side of Bijagua	Costa Rica	10.73142	-85.05535	Hap 13 (MLBM 148983–89)
P. amates	12	Rio Celeste at Viejo Oeste Bar, just off Costa Rica Hwy 4, approximately 2.5 km southeast of Katira	Costa Rica	10.73125	-84.88818	Hap 7 (MLBM 160701)
P. amates	13	Rio Chimurria at Costa Rica Hwy 35, near Santa Clara de Upala, approximately 25 km north of Boca Arenal	Costa Rica	10.72740	-84.55820	Hap 14 (Pha13)
P. amates	14	Rio Tempisque	Costa Rica	10.62745	-85.43412	Hap 15 (MLBM 167999, MLBM 168000)
P. amates	15	Trib. to Rio Toro north of Golfito	Costa Rica	10.59927	-84.06908	Hap 16 (MLBM 161080, MLBM 161081)
P. amates	16	Trib. to Rio Sarapiquí on road ~2 km off end-Costa Rica Rd. 505, approximately 9 km northwest of Puerto Viejo de Sarapiquí	Costa Rica	10.52452	-84.03127	Hap 16 (MLBM 148993, MLBM 148994), Hap 17 (MLBM 148990)
P. amates	17	Rio La Piedrita, trib. to Rio Burro, just off Costa Rica Rd 702, approximately 4-4.5 km southeast of La Fortuna/Barrio Pilo (neighborhoods at Rd 142)	Costa Rica	10.44456	-84.61420	Hap 18 (MLBM 161809)

P. amates	18	Trib. to Rio Agua Caliente southwest of Volcan Arenal	Costa Rica	10.43505	-84.72340	Hap 18 (MLBM 149139–43, MLBM 149145, MLBM 149146), Hap 19 (MLBM 149144)
P. amates	19	Upper Rio Tortuguero 0.5 km off of Costa Rica Rd 249, 1 km north of San Rafael and approximately 5 km north of Guapiles	Costa Rica	10.25942	-83.81220	Hap 16 (MLBM 149159, MLBM 149160, MLBM 149164, Pha6), Hap 20 (MLBM 149161), Hap 21 (MLBM 149163)
P. amates	20	Brazo del Sucio, trib. to Rio Chirripó	Costa Rica	10.21872	-83.90466	Hap 16 (MLBM 160281–92)
P. amates	21	Trib. to Rio Parismina 200 m off of Costa Rica Hwy 32, approximately 3 km southeast of Guacimo	Costa Rica	10.19968	-83.65870	Hap 16 (Pha2)
P. amates	22	Trib. to Rio Parismina just off Costa Rica Hwy 32, approximately 4 km southeast of Guacimo	Costa Rica	10.19772	-83.65210	Hap 22 (MLBM 149137)
P. amates	23	Rio Barbilla (major trib. to Rio Matina) at Costa Rica Rd 805 (dirt road) just west of B-Line, 0.5 km from Costa Rica Hwy 32	Costa Rica	10.04416	-83.33383	Hap 15 (MLBM 168003, MLBM 168004, MLBM 168010–12), Hap 23 (MLBM 168005), Hap 24 (MLBM 168006), Hap 25 (MLBM 168013), Hap 26 (MLBM 169940, MLBM 169941), Hap 27 (MLBM 169942), Hap 28 (MLBM 168009)
P. amates	24	Trib. to Rio Sixaola	Costa Rica	9.62093	-82.85768	Hap 29 (MLBM 149133), Hap 30 (MLBM 149131), Hap 31 (MLBM 149132)
P. amates	1	Tributary to Rio Lancetilla at Lancetilla Preserve, Departamento Atlantida	Honduras	15.74068	-87.45578	Hap 1 (1180 [LSUMZ 14522])
P. amates	2	Rio Santiago at community of San Rafael near La Masica, Departamento Atlantida	Honduras	15.66433	-87.08475	Hap 1 (1253 [LSUMZ 14564])
P. amates	3	Rio Grande de Matagalpa	Nicaragua	12.98351	-85.13827	Hap 2 (MLBM 168007)
P. amates	4	Unnamed trib. to Rio Acoyapa at Nicaragua Hwy 7 approximately 1.8 km west of Lovago	Nicaragua	12.00340	-85.18153	Hap 3 (MLBM 173793), Hap 4 (MLBM 173794, MLBM 173796)
P. amates	5	Unnamed trib. to Rio Rama at Nicaragua Hwy 71, approximately 8 km south of Puente Rio Rama	Nicaragua	11.82544	-84.60611	Hap 5 (MLBM 172434)
P. amates	6	Unnamed trib. to Rio Rama at Nicaragua Hwy 71, approximately 1.5 km south of Colonia El Corocito	Nicaragua	11.74923	-84.55819	Hap 5 (MLBM 173687)
P. amates	7	Unnamed trib. to Rio Rama at Nicaragua Hwy 71 approximately 7.5 km northwest of Nueva Guinea	Nicaragua	11.73855	-84.51068	Hap 4 (MLBM 173800–06, MLBM 173808), Hap 6 (MLBM 173815)
P. amates	8	Unnamed trib. to Rio Zapote (a trib. to Rio Plata, which drains into Rio Rama) south of Nueva Guinea, approximately 1 km south of end of Nicaragua Hwy 71	Nicaragua	11.68127	-84.46019	Hap 4 (MLBM 173472), Hap 7 (MLBM 173443)
P. amates	N/A	See Hrbek *et al.* (2007)	N/A	N/A	N/A	Hap 32 (Pamat)

Abbreviations: LSUMZ, LSU Museum of Natural Science catalog number; MLBM, Brigham Young University Fish Collection (Monte L. Bean Museum); N/A, not available; trib., tributary.

Sampling localities for *Poecilia mexicana*

Species	Site No.	Locality	Country	Latitude	Longitude	Cyt*b* haplotype (sample IDs)
Poecilia sphenops	23	Rio Agua Buena	Honduras	15.76611	-86.99889	Hap 26 (STRI 4303)
P. sphenops	79	Rio Grande de Matagalpa	Nicaragua	12.84517	-86.10272	Hap 44 (STRI 13327)
P. mexicana	108	Rio Liberia at outskirts of Liberia	Costa Rica	10.62745	-85.43412	Hap 95 (MLBM 167890, MLBM 167891, MLBM 167893), Hap 106 (MLBM 167887, MLBM 167888, MLBM 167892,

						MLBM 167894)
P. mexicana	111	Rio Tempisque on road between Guardia and Comunidad, Nicoya Peninsula	Costa Rica	10.57220	-85.39192	Hap 95 (MLBM 168793, MLBM 168794, MLBM 168797), Hap 108 (MLBM 168792, MLBM 168795, MLBM 168798), Hap 109 (MLBM 168796)
P. mexicana	112	Rio Salto at CA1 southeast of Liberia, Nicoya Peninsula	Costa Rica	10.56106	-85.39192	Hap 110 (MLBM 167904, MLBM 167905, MLBM 167908–10), Hap 111 (MLBM 167906)
P. mexicana	115	Rio Sardinal at Sardinal, on Rd 151 approximately 5 km from 21	Costa Rica	10.51508	-85.65166	Hap 95 (MLBM 167896, MLBM 167901), Hap 108 (MLBM 167895, MLBM 167897), Hap 110 (MLBM 167898–900), Hap 116 (MLBM 167902)
P. mexicana	117	Rio Cabuyo at unnamed province road to la Reserva Biologica Lomas Bardudal	Costa Rica	10.48961	-85.38555	Hap 108 (MLBM 168698), Hap 110 (MLBM 168697, MLBM 168700–702), Hap 117 (MLBM 168696), Hap 118 (MLBM 168699)
P. mexicana	125	Unnamed trib. to Rio Tempisque drainage approximately 2 km south of Belén, Nicoya Peninsula	Costa Rica	10.39076	-85.59045	Hap 110 (MLBM 167933, MLBM 167934)
P. mexicana	129	Rio Diriá at CA1 approximately 2-3 km north of Santa Cruz, Nicoya Peninsula	Costa Rica	10.26677	-85.59261	Hap 95 (MLBM 168801), Hap 97 (MLBM 168804), Hap 108 (MLBM 168802, MLBM 168803), Hap 110 (MLBM 168800, MLBM 168805), Hap 121 (MLBM 168806)
P. mexicana	137	Rio Barbilla at province road just west of B-Line, a few km west of the road to Matina (dirt road, off northeast side of big road to Puerto Limon)	Costa Rica	10.04416	-83.33383	Hap 132 (MLBM 168888–90, MLBM 168892–95)
P. mexicana	17	Arroyo Sal Si Puedes, Belize	Guatemala	16.95730	-89.35930	Hap 18 (STRI 8084)
P. mexicana	18	Zona Militar Estanque, locality approximated	Guatemala	16.90290	-89.72929	Hap 19 (STRI 8111, STRI 8113, STRI 8114)
P. mexicana	19	Arroyo Comiston, trib. to Rio La Pasion	Guatemala	16.55440	-90.19270	Hap 21 (STRI 7957), Hap 22 (STRI 7958, STRI 7960)
P. mexicana	20	Rio La Pasion	Guatemala	16.55060	-90.23010	Hap 21 (STRI 7999), Hap 23 (STRI 7995, STRI 7996)
P. mexicana	21	Rio San Simon	Guatemala	15.84110	-90.28920	Hap 21 (STRI 7839, STRI 7854)
P. mexicana	22	Rio Sebol at Finca Sebol	Guatemala	15.80630	-89.94480	Hap 25 (STRI 7906, STRI 7914, STRI 7915, STRI 7925)
P. mexicana	25	Rio Amatillo at Lago Izabal near Venta de El Amatillo	Guatemala	15.53910	-88.89830	Hap 29 (STRI 8181, STRI 8184–86)
P. mexicana	30	Rio Chaguacal, trib. to Rio Polochic	Guatemala	15.31617	-89.85556	Hap 33 (STRI 8288–90, STRI 8294)
P. mexicana	31	Rio Dona Maria, trib. to Rio Motagua	Guatemala	15.20910	-89.24810	Hap 34 (STRI 8241), Hap 35 (STRI 8245, STRI 8246, STRI 8248)
P. mexicana	32	Rio Lobo, trib. to Rio Motagua	Guatemala	15.18160	-89.29940	Hap 34 (STRI 8232), Hap 35 (STRI 8222, STRI 8230)
P. mexicana	24	Rio Taujica at Taujica	Honduras	15.68100	-85.93930	Hap 27 (STRI 8558), Hap 28 (STRI 8565)
P. mexicana	26	Quebrada de Chicho between Comunidades de Achiote and Cholomena	Honduras	15.53480	-86.21170	Hap 30 (STRI 8534, STRI 8541)
P. mexicana	27	Rio Medina, trib. to Rio Aguán, at Coyoles Centrales	Honduras	15.48380	-86.66600	Hap 30 (STRI 8607)
P. mexicana	28	Rio Naco, trib. to Ulúa	Honduras	15.34147	-88.62480	Hap 31 (STRI 8408, STRI 8411)
P. mexicana	29	Rio Camalote, trib. to Rio Ulúa	Honduras	15.32656	-88.66264	Hap 31 (STRI 8470, STRI 8471)

P. mexicana	33	Rio Yojoa, trib. to Rio Ulúa	Honduras	15.03480	-87.92870	Hap 36 (STRI 8618, STRI 8620), Hap 37 (STRI 8619)
P. mexicana	34	Rio El Sauce (Amarillo) at Santa Rita	Honduras	14.86603	-89.06783	Hap 38 (STRI 8463, STRI 8465, STRI 8466)
P. mexicana	35	Rio El Sauce (Amarillo), trib. to Rio Motagua, near Copán Ruins	Honduras	14.85589	-89.12355	Hap 34 (STRI 8372–74)
P. mexicana	36	Tio Higuito, trib. to Rio Motagua, at cuenca near Higuito	Honduras	14.83940	-89.16819	Hap 39 (STRI 8362, STRI 8363)
P. mexicana	38	Rio Humuya, trib. to Rio Ulúa, at Comayagua	Honduras	14.45370	-87.65230	Hap 36 (STRI 8637)
P. mexicana	39	Rio Lempa at Nueva Ocotepeque	Honduras	14.39417	-89.20816	Hap 43 (STRI 8311, STRI 8316)
P. mexicana	42	Rio Goascorán at Caridad	Honduras	13.82770	-87.69480	Hap 45 (STRI 8858, STRI 8860)
P. mexicana	51	Rio Goascorán at Goascorán	Honduras	13.58928	-87.76212	Hap 53 (STRI 8805, STRI 8815), Hap 54 (STRI 8807)
P. mexicana	53	Rio Chiquito, trib. to Rio Nacaome, at Nacaome	Honduras	13.54170	-87.47880	Hap 55 (STRI 8873, STRI 8875)
P. mexicana	56	Rio Chiquito, trib. to Rio Choluteca, at Orocuina	Honduras	13.48270	-87.09900	Hap 32 (STRI 8914–16)
P. mexicana	4	Rio Puyacatengo, Banos del Azufre	Mexico	17.55225	-92.99859	Hap 5 (Pmmex33), Hap 6 (Pmmex15), Hap 7 (Pmmex16)
P. mexicana	5	Rio Pichucalco, Banos del Azufre	Mexico	17.55200	-92.99900	Hap 8 (Pmmex14)
P. mexicana	6	Rio Puyacatengo, Vicente Gurrero Lerma	Mexico	17.51008	-92.91448	Hap 1 (Pmmex17), Hap 5 (Pmmex32), Hap 6 (Pmmex18)
P. mexicana	7	Tributary to the Rio Ixtapangajoya, Chiapas state	Mexico	17.51000	-92.98000	Hap 1 (PmmxNSS2-0), Hap 9 (PmmxNSSm-0)
P. mexicana	8	Rio Ixtapangajoya, Chiapas state	Mexico	17.49500	-92.99800	Hap 8 (PmmxIxta-0), Hap 9 (PmmxIxt2-0, PmmxIxt3-0), Hap 10 (Pmmex22)
P. mexicana	9	Rio Nututun, Palenque	Mexico	17.48417	-91.97376	Hap 11 (Pmmex29)
P. mexicana	10	Rio Tacotalpa, Arroyo Tres	Mexico	17.48400	-92.77600	Hap 12 (Pmmex30)
P. mexicana	11	Rio Oxolotan, Tapijulapa	Mexico	17.46444	-92.77430	Hap 12 (Pmmex28)
P. mexicana	12	Rio Puyacatengo, La Lluvia	Mexico	17.46400	-92.89500	Hap 13 (Pmmex21, Pmmex22, Pmmex23)
P. mexicana	13	Puyacatengo Springs, in Tabasco state	Mexico	17.45800	-92.88900	Hap 5 (PmmxPysp-0, Pmmex31)
P. mexicana	14	Rio Puyacatengo, La Lluvia, Puyacatengo Springs	Mexico	17.45761	-92.88892	Hap 14 (Pmmex19, Pmmex20)
P. mexicana	15	Cueva del Azufre, Tabasco state	Mexico	17.43843	-92.77476	Hap 15 (PmmxCDA-0, Pmmex34)
P. mexicana	16	Rio Tacotalpa, Arroyo Bonita	Mexico	17.42685	-92.75213	Hap 16 (Pmmex26), Hap 17 (Pmmex27)
P. mexicana	N/A	Rio Puyacatengo, Rio Pichucalco, La Joya, Santa Ana; see Tobler *et al.* (2011)	Mexico	N/A	N/A	Hap 1 (Pmmex24)
P. mexicana	40	Rio Prinzapolka	Nicaragua	13.93861	-84.82472	Hap 44 (STRI 14137)
P. mexicana	41	Rio Prinzapolka	Nicaragua	13.91667	-84.56333	Hap 44 (STRI 14110–12)
P. mexicana	43	Rio Prinzapolka	Nicaragua	13.82028	-85.04444	Hap 44 (STRI 14100)
P. mexicana	44	Rio Coco	Nicaragua	13.77111	-85.64833	Hap 46 (STRI 14041–43)

P. mexicana	45	Rio Prinzapolka	Nicaragua	13.73222	-84.51472	Hap 44 (STRI 14131, STRI 14132)
P. mexicana	46	Rio Prinzapolka	Nicaragua	13.69575	-84.69794	Hap 44 (STRI 14231, STRI 14232)
P. mexicana	47	Rio Coco	Nicaragua	13.67611	-85.79611	Hap 47 (STRI 14060)
P. mexicana	48	Rio Coco	Nicaragua	13.67356	-85.76389	Hap 48 (STRI 13934)
P. mexicana	49	Rio Prinzapolka	Nicaragua	13.63583	-85.36500	Hap 49 (STRI 14070, STRI 14072), Hap 50 (STRI 14071)
P. mexicana	50	Rio Coco	Nicaragua	13.60944	-86.47483	Hap 48 (STRI 13450), Hap 51 (STRI 13451), Hap 52 (STRI 13456)
P. mexicana	52	Rio Prinzapolka	Nicaragua	13.55500	-84.86333	Hap 44 (STRI 14145, STRI 14146)
P. mexicana	54	Rio Coco	Nicaragua	13.51294	-85.80986	Hap 48 (STRI 13921–23)
P. mexicana	55	Rio Prinzapolka	Nicaragua	13.50306	-84.84472	Hap 44 (STRI 14256)
P. mexicana	57	Rio Grande de Matagalpa	Nicaragua	13.46000	-84.91444	Hap 44 (STRI 14154, STRI 14155)
P. mexicana	58	Rio Coco	Nicaragua	13.42472	-85.98500	Hap 48 (STRI 13997), Hap 56 (STRI 13998)
P. mexicana	59	Unnamed trib. to Rio Grande de Matagalpa, east of Waslala (road to Siuna)	Nicaragua	13.35260	-85.35108	Hap 44 (MLBM 168822), Hap 49 (MLBM 168815–21)
P. mexicana	60	Rio Coco	Nicaragua	13.34164	-85.95636	Hap 57 (STRI 13887)
P. mexicana	61	Rio Coco	Nicaragua	13.33811	-85.94881	Hap 48 (STRI 13876)
P. mexicana	62	Rio Coco	Nicaragua	13.33335	-86.20417	Hap 48 (STRI 13988), Hap 57 (STRI 13986, STRI 13987)
P. mexicana	63	Rio Coco	Nicaragua	13.29133	-86.18028	Hap 58 (STRI 13972–74)
P. mexicana	64	Rio Grande de Matagalpa	Nicaragua	13.26233	-85.43922	Hap 44 (STRI 14195, STRI 14196)
P. mexicana	65	Unnamed trib. to Rio Grande de Matagalpa, at stream just west of km marker 226	Nicaragua	13.25769	-85.45440	Hap 44 (MLBM 168168–74)
P. mexicana	66	Rio Grande de Matagalpa	Nicaragua	13.25611	-85.54453	Hap 59 (STRI 14172)
P. mexicana	67	Unnamed trib. to Lago Managua between Estelí and León	Nicaragua	13.22797	-86.55272	Hap 60 (MLBM 167960, MLBM 167962, MLBM 167964)
P. mexicana	68	Unnamed trib. to Rio Grande de Matagalpa, west of La Mora and slightly further west of La Dalia	Nicaragua	13.22058	-85.72626	Hap 61 (MLBM 168175–82)
P. mexicana	69	Rio Grande de Matagalpa	Nicaragua	13.17450	-86.28528	Hap 44 (STRI 13970, STRI 13971), Hap 62 (STRI 13969)
P. mexicana	70	Rio Grande de Matagalpa	Nicaragua	13.15058	-85.92922	Hap 44 (STRI 13838, STRI 13845, STRI 13846)
P. mexicana	71	Unnamed trib. to Lago de Apanás	Nicaragua	13.11843	-86.01022	Hap 44 (MLBM 168807–11, MLBM 168813, 168814), Hap 63 (MLBM 168812)
P. mexicana	73	Rio Coco	Nicaragua	13.09797	-86.36033	Hap 48 (STRI 13429–31)
P. mexicana	74	Unnamed trib. to Rio Estelí, thus a trib. to Rio Coco	Nicaragua	13.05866	-86.35114	Hap 44 (MLBM 167956), Hap 48 (MLBM 167958), Hap 64 (MLBM 167952–55, MLBM 167957, MLBM 167959)
P. mexicana	75	Unnamed trib. to Rio Tuma northwest of Rio Blanco (town; flowing from Mt. Musun)	Nicaragua	12.93613	-85.23434	Hap 65 (MLBM 168151–58)

P. mexicana	78	Rio Grande de Matagalpa	Nicaragua	12.87886	-85.21286	Hap 44 (STRI 14294)
P. mexicana	80	Trib. to unnamed trib. of the Rio Grande de Matagalpa, approximately 32 km west of Rio Blanco (town) on road between Matagalpa and Rio Blanco (town)	Nicaragua	12.82341	-85.44279	Hap 44 (MLBM 168159, MLBM 168161, MLBM 168162), Hap 67 (MLBM 168160), Hap 68 (MLBM 168163, MLBM 168164), Hap 69 (MLBM 168165), Hap 70 (MLBM 168166)
P. mexicana	81	Trib. to Rio Grande de Matagalpa at Puente de Tierra Azul on road to Rio Blanco	Nicaragua	12.68476	-85.54708	Hap 44 (MLBM 167938), Hap 71 (MLBM 167935, MLBM 167936, MLBM 167939, MLBM 167941), Hap 72 (MLBM 167937, MLBM 167940, MLBM 167942)
P. mexicana	82	Rio Grande de Matagalpa	Nicaragua	12.67075	-86.09139	Hap 44 (STRI 13313–15)
P. mexicana	84	Unnamed trib. to Rio Malacatoya at Teustepe, just off road to Rama (Nicaragua Hwy 7), southwest of Boaco	Nicaragua	12.41825	-85.79299	Hap 66 (MLBM 167943), Hap 73 (MLBM 168897), Hap 74 (MLBM 168898), Hap 75 (MLBM 168896)
P. mexicana	88	Unnamed trib./headwater of Rio Mico northeast of San Pedro de Lovago	Nicaragua	12.13630	-85.04597	Hap 78 (MLBM 167928, MLBM 167929)
P. mexicana	91	Rio Escondido	Nicaragua	12.01289	-84.66831	Hap 66 (STRI 13666)
P. mexicana	99	Unnamed trib. to Lago Nicaragua just east of mile marker km 238 on road to San Miguelito, Chontales	Nicaragua	11.50538	-84.83956	Hap 66 (MLBM 168708), Hap 87 (MLBM 167931, MLBM 168709), Hap 88 (MLBM 167927), Hap 89 (MLBM 167930, MLBM 168710), Hap 90 (MLBM 167932)
P. mexicana	100	Rio San Juan	Nicaragua	11.19033	-85.51783	Hap 87 (STRI 14722)
P. gillii	101	Rio Sapoa	Costa Rica	11.04437	-85.61590	Hap 44 (PG725.05), Hap 66 (PG725.02, PG725.03, PG725.08), Hap 88 (PG725.01, PG725.04, PG725.06), Hap 91 (PG725.07)
P. gillii	102	Rio Sabalo	Costa Rica	11.04283	-85.48922	Hap 92 (PG726.01–08)
P. gillii	103	Rio San Juan	Costa Rica	10.90839	-85.21126	Hap 87 (STRI 2171), Hap 92 (STRI 2170)
P. gillii	104	Rio Chimurria	Costa Rica	10.72740	-84.55823	Hap 66 (PG716.04), Hap 93 (PG716.01–06), Hap 94 (PG716.08)
P. gillii	105	Rio Irigaray in Irigaray at CA 1, approximately 2.5 km west of Canas Dulces	Costa Rica	10.72340	-85.51038	Hap 95 (PG724.01, PG724.02, PG724.06, PG724.09–12), Hap 96 (PG724.14), Hap 97 (PG724.13), Hap 98 (PG724.03–05, PG724.07, PG724.08, PG724.15)
P. gillii	106	Quebrada Homiguera	Costa Rica	10.69090	-85.08365	Hap 97 (PG611.02, PG611.03), Hap 99 (PG611.05), Hap 100 (PG611.01), Hap 101 (PG611.04)
P. gillii	107	Rio Queques	Costa Rica	10.64482	-84.82223	Hap 87 (PG719.01, PG719.03, PG719.05–08, PG719.10, PG719.12–16), Hap 102 (PG719.02), Hap 103 (PG719.04), Hap 104 (PG719.09), Hap 105 (PG719.11)
P. gillii	109	Small ditch	Costa Rica	10.62407	-85.05812	Hap 97 (PG610.01–03)
P. gillii	110	Rio Infiernito	Costa Rica	10.61802	-84.48418	Hap 66 (PG715.01–07), Hap 107 (PG715.08)
P. gillii	113	Rio Sabalito	Costa Rica	10.54858	-84.98080	Hap 112 (PG608.03, PG608.04), Hap 113 (PG608.01, PG608.02, PG608.05)
P. gillii	114	Rio Sarapiquí	Costa Rica	10.52455	-84.03133	Hap 66 (PG713.04), Hap 87 (PG713.06, PG713.08), Hap 92 (PG713.03, PG713.09, PG713.12), Hap 102 (PG713.02, PG713.10, PG713.13), Hap 114 (PG713.01, PG713.07, PG713.11), Hap 115 (PG713.05)
P. gillii	116	Rio La Palma	Costa Rica	10.49875	-84.68900	Hap 66 (PG602.01–05)

P. gillii	118	Rio Magdalena	Costa Rica	10.47945	-85.07812	Hap 110 (PG616.01–03)
P. gillii	119	Rio Sarapiquí	Costa Rica	10.47225	-83.99195	Hap 92 (STRI 1245)
P. gillii	119	Rio Sarapiquí	Costa Rica	10.47225	-83.99195	Hap 114 (STRI 1246)
P. gillii	120	Lago Arenal	Costa Rica	10.47208	-84.76933	Hap 66 (PG603.02), Hap 112 (PG603.01)
P. gillii	121	Rio Santa Rosa	Costa Rica	10.46113	-85.07438	Hap 95 (PG614.02), Hap 110 (PG614.01)
P. gillii	122	Rio Chiquito	Costa Rica	10.43770	-84.86815	Hap 66 (PG612.05), Hap 110 (PG612.03, PG612.04), Hap 112 (PG612.01), Hap 113 (PG612.02)
P. gillii	123	Rio Carrisal	Costa Rica	10.39502	-85.58688	Hap 110 (PG723.01)
P. gillii	124	Rio Isla Grande	Costa Rica	10.39300	-83.96820	Hap 119 (PG636.01)
P. gillii	126	Rio Javilla	Costa Rica	10.37208	-85.09740	Hap 97 (PG617.02), Hap 116 (PG617.03), Hap 120 (PG617.01)
P. gillii	127	Rio Canas	Costa Rica	10.34825	-85.16882	Hap 108 (STRI 1205), Hap 110 (STRI 1206)
P. gillii	128	Rio Higueron	Costa Rica	10.34270	-85.07594	Hap 110 (STRI 2119), Hap 114 (PG712.04)
P. gillii	130	Rio Tortuguero	Costa Rica	10.25942	-83.81223	Hap 114 (PG712.08, PG712.10, PG712.15, PG712.16), Hap 122 (PG712.01–03, PG712.05, PG712.09, PG712.11, PG712.14), Hap 123 (PG712.06), Hap 124 (PG712.07), Hap 125 (PG712.12)
P. gillii	131	Rio Congo	Costa Rica	10.23998	-84.99171	Hap 110 (PG808.01, PG808.03), Hap 121 (PG808.02), Hap 126 (PG808.04, PG808.05), Hap 127 (PG808.06)
P. gillii	132	Rio Parismina	Costa Rica	10.19772	-83.56873	Hap 128 (PG710.01–07)
P. gillii	133	Rio Herediana	Costa Rica	10.12417	-83.55617	Hap 128 (PG703.02, PG703.03), Hap 129 (PG703.01, PG703.05)
P. gillii	134	Rio Ciruelas	Costa Rica	10.05914	-84.75919	Hap 130 (STRI 13308)
P. gillii	135	Rio Morote at Costa Rica Hwy 21 approximately 6 km north of Carmona	Costa Rica	10.05828	-85.26202	Hap 108 (PG722.06), Hap 121 (PG722.01, PG722.02, PG722.04, PG722.05, PG722.07–16), Hap 131 (PG722.03)
P. gillii	136	Rio Nosara	Costa Rica	10.04833	-85.54520	Hap 110 (STRI 1231, STRI 1232)
P. gillii	138	Rio Rosales	Costa Rica	10.02979	-84.32582	Hap 66 (PG809.01), Hap 108 (PG809.02, PG809.03), Hap 128 (PG809.04), Hap 133 (PG809.05)
P. gillii	139	Rio Naranjo	Costa Rica	10.02264	-84.73442	Hap 44 (PG807.01, PG807.04, PG807.05), Hap 130 (PG807.02, PG807.03)
P. gillii	140	Rio Toro, trib. to Rio Matina, just off Costa Rica Hwy 32 approximately 24 km west of Limon	Costa Rica	10.01678	-83.21022	Hap 132 (PG708.06–08), Hap 134 (PG708.01, PG708.02, PG708.04, PG708.05)
P. gillii	141	Rio Centeno	Costa Rica	9.94132	-84.53886	Hap 135 (PG806.01, PG806.03–05), Hap 136 (PG806.02)
P. gillii	142	Rio Pacacua	Costa Rica	9.91960	-84.24130	Hap 133 (PG801.01–05)
P. gillii	143	Unnamed lagoon	Costa Rica	9.89258	-82.97228	Hap 137 (PG707.02–04, PG707.06–08), Hap 138 (PG707.01), Hap 139 (PG707.05)
P. gillii	144	Rio Grande de Tárcoles	Costa Rica	9.87980	-84.52780	Hap 95 (PG804.01), Hap 133 (PG804.02), Hap 140 (PG804.03)

P. gillii	145	Rio Reventazon	Costa Rica	9.87230	-83.63320	Hap 43 (PG701.05), Hap 141 (PGB701.01–04), Hap 142 (PG701.01–04)
P. gillii	146	Quebrada La Canela	Costa Rica	9.85151	-84.52766	Hap 44 (PG805.02), Hap 58 (PG805.01), Hap 141 (PG805.03)
P. gillii	147	Unnamed trib. to Rio Sixaola at Costa Rica Hwy 36 less than 1 km west of Catarata	Costa Rica	9.63203	-82.81922	Hap 143 (PGB704.01, PGB704.02, PGB704.04), Hap 144 (PGB704.03), Hap 145 (PG704.01–08)
P. gillii	148	Rio Hatillo Viejo	Costa Rica	9.62312	-82.85520	Hap 74 (MLBM 173339), Hap 79 (MLBM 173337), Hap 80 (MLBM 173336), Hap 81 (MLBM 173341)
P. gillii	149	Rio Sixaola	Costa Rica	9.59872	-82.80247	Hap 145 (STRI 1291), Hap 147 (STRI 1292)
P. gillii	156	Finco la Palma	Costa Rica	9.53697	-84.38589	Hap 151 (PGB702.01, PGB702.02), Hap 152 (PGB702.04), Hap 153 (PGB702.03)
P. gillii	159	Rio General	Costa Rica	9.38944	-83.66361	Hap 146 (PG4814.02), Hap 154 (PG4814.01)
P. gillii	165	Rio Peje, trib. to Rio General	Costa Rica	9.28493	-83.64566	Hap 156 (STRI 2074)
P. gillii	173	Rio Pejibaye	Costa Rica	9.15694	-83.57528	Hap 146 (PG4810.02, PG4810.03, PG4810.05–10, PG4810.12–15), Hap 162 (PG4810.01), Hap 163 (PG4810.04), Hap 164 (PG4810.11)
P. gillii	203	Rio Salama Nuevo	Costa Rica	8.90425	-83.43932	Hap 175 (STRI 2051)
P. gillii	223	Rio Nuevo	Costa Rica	8.64103	-82.95297	Hap 175 (PGB714.01, PGB714.02)
P. gillii	225	Rio Barrigones	Costa Rica	8.59323	-83.42182	Hap 178 (PG517.01, PG517.02)
P. gillii	36	Tio Higuito at cuenca Motagua / Higuito	Honduras	14.83940	-89.16819	Hap 39 (STRI 8364), Hap 40 (STRI 8365)
P. gillii	37	Rio Ulúa	Honduras	14.65096	-88.88144	Hap 39 (STRI 8343, STRI 8356), Hap 40 (STRI 8355, STRI 8358), Hap 41 (STRI 8357), Hap 42 (STRI 8344)
P. gillii	72	Rio Estelí in Estelí	Nicaragua	13.10663	-86.35710	Hap 64 (MLBM 174214–16)
P. gillii	76	Rio Viejo (afluente Lago de Managua)	Nicaragua	12.90703	-86.12831	Hap 44 (STRI 13417–20)
P. gillii	77	Trib. to Rio Grande Viejo on road between Estelí and León	Nicaragua	12.89324	-86.17908	Hap 32 (MLBM 174330), Hap 44 (MLBM 174327), Hap 66 (MLBM 174331)
P. gillii	83	Telica	Nicaragua	12.51656	-86.86542	Hap 32 (MLBM 174259–62)
P. gillii	84	Unnamed trib. to Rio Malacatoya at Teustepe, just off road to Rama (Nicaragua Hwy 7), southwest of Boaco	Nicaragua	12.41825	-85.79299	Hap 66 (MLBM 168899)
P. gillii	85	Rio Caracol	Nicaragua	12.35116	-85.88870	Hap 74 (MLBM 173007, MLBM 173008), Hap 76 (MLBM 173005, MLBM 173006)
P. gillii	86	Rio Malacatoya	Nicaragua	12.32661	-85.95553	Hap 73 (MLBM 172867, MLBM 172868)
P. gillii	87	Lago Jiloa	Nicaragua	12.21858	-86.31194	Hap 32 (MLBM 172459, MLBM 172460, MLBM 172462), Hap 77 (MLBM 172455, MLBM 172461)
P. gillii	89	Rio Mayales main stem	Nicaragua	12.06679	-85.40375	Hap 74 (MLBM 173339), Hap 79 (MLBM 173337), Hap 80 (MLBM 173336), Hap 81 (MLBM 173341)
P. gillii	90	Trib. to Rio Mayales about 6 km southwest of Juigalpa on Nicaragua Hwy 37	Nicaragua	12.05663	-85.40814	Hap 66 (MLBM 173271, MLBM 173272), Hap 74 (MLBM 173274), Hap 82 (MLBM 173273)
P. gillii	92	Unnamed trib.	Nicaragua	11.87992	-85.13156	Hap 66 (MLBM 174109–13, MLBM 174116), Hap 80 (MLBM 174114)

P. gillii	93	Unnamed trib.	Nicaragua	11.82942	-85.20479	Hap 32 (MLBM 173959), Hap 66 (MLBM 173960, MLBM 173964)
P. gillii	94	Unnamed trib. to Lago Nicaragua	Nicaragua	11.74923	-84.55819	Hap 66 (MLBM 173680–85)
P. gillii	95	Rio La Conquista	Nicaragua	11.72472	-86.18469	Hap 74 (MLBM 172443, MLBM 172444, MLBM 172446), Hap 79 (MLBM 172439), Hap 83 (MLBM 172440–42, MLBM 172445)
P. gillii	96	Unnamed drainage ditch trib. 1 km north of Ochomogo	Nicaragua	11.67886	-85.98817	Hap 66 (MLBM 172559), Hap 84 (MLBM 172558)
P. gillii	97	Unnamed trib. to Rio Rama drainage south of Nueva Guinea	Nicaragua	11.67839	-84.45622	Hap 78 (MLBM 173459, MLBM 173460, MLBM 173464, MLBM 173468–70), Hap 85 (MLBM 173465)
P. gillii	98	Rio Ochomogo	Nicaragua	11.65664	-85.97319	Hap 66 (MLBM 172539, MLBM 172541, MLBM 172542), Hap 86 (MLBM 172540)
P. gillii	150	Rio Playa Alta near Nombre de Dios	Panama	9.57069	-79.43614	Hap 149 (STRI 18632)
P. gillii	151	Rio Viento Frio	Panama	9.57056	-79.38250	Hap 40 (STRI 17945)
P. gillii	152	Rio Cuango	Panama	9.55080	-79.30920	Hap 40 (STRI 9020), Hap 150 (STRI 9030)
P. gillii	153	Rio Cascajal, upstream about 5 km east of Fuerte de San Lorenzo	Panama	9.54722	-79.60400	Hap 40 (STRI 12609, STRI 12610, STRI 12614)
P. gillii	154	Rio Cascajal, upstream about 5 km east of Fuerte de San Lorenzo	Panama	9.54722	-79.63040	Hap 40 (STRI 4533, STRI 4534, STRI 4536)
P. gillii	155	Rio Cascajal	Panama	9.54642	-79.60625	Hap 40 (STRI 2956)
P. gillii	157	Rio Cuango	Panama	9.51820	-79.28480	Hap 40 (STRI 9395)
P. gillii	158	Quebrada on Almirante-Changuinola road	Panama	9.39686	-82.50058	Hap 40 (STRI 12442)
P. gillii	160	Rio Changuinola	Panama	9.37000	-82.54000	Hap 149 (STRI 4969)
P. gillii	161	Rio Bongie, trib. to Rio Teribe	Panama	9.35990	-82.61000	Hap 149 (STRI 4979, STRI 4980, STRI 4992, STRI 4993)
P. gillii	162	Two quebradas before Big Creek, Isla Colón, Bocas del Toro	Panama	9.35711	-82.25322	Hap 155 (STRI 16977)
P. gillii	163	Stream between Sardinilla and Salamanca	Panama	9.32644	-79.61189	Hap 40 (STRI 16414)
P. gillii	164	Quebrada by Almirante-Changuinola road	Panama	9.31469	-82.45036	Hap 149 (STRI 12428)
P. gillii	166	Unnamed trib. to Quebrada Nigua, at Ruta Rambala-Almirante	Panama	9.27894	-82.41525	Hap 149 (STRI 18787), Hap 157 (STRI 18782)
P. gillii	167	Quebrada en Mateo	Panama	9.22581	-80.08589	Hap 158 (STRI 16786), Hap 159 (STRI 16781–85)
P. gillii	168	Rio Mamoni	Panama	9.22361	-79.09222	Hap 159 (STRI 11203)
P. gillii	170	Rio Membrillar	Panama	9.17389	-80.18500	Hap 160 (STRI 16723), Hap 161 (STRI 16722)
P. gillii	171	Rio Pacora	Panama	9.16417	-79.34000	Hap 149 (STRI 2851)
P. gillii	172	Rio Chichebre	Panama	9.16083	-79.15417	Hap 149 (STRI 11374)
P. gillii	174	Small creek on km 41 of Punta Pena - Almirante road	Panama	9.14753	-82.31767	Hap 149 (STRI 12459)

P. gillii	175	Rio Uyama	Panama	9.13892	-82.30694	Hap 132 (STRI 18792), Hap 149 (STRI 18791)
P. gillii	176	Rio Miguel de la Borda	Panama	9.13358	-80.29422	Hap 40 (STRI 16810, STRI 16820), Hap 160 (STRI 16809, STRI 16823)
P. gillii	177	Quebrada La Candelaria, Rio Jobo, Rio Indio de Anton	Panama	9.13011	-80.17155	Hap 41 (STRI 10048)
P. gillii	178	Quebrada on Km 34 at Punta Pena - Almirante road.	Panama	9.11697	-82.29019	Hap 157 (STRI 12481)
P. gillii	179	Rio Chagres	Panama	9.11000	-79.68000	Hap 165 (STRI 6803)
P. gillii	180	Escudo de Veraguas, Bocas del Toro	Panama	9.10222	-81.56167	Hap 166 (STRI 6826)
P. gillii	181	Rio Pedro Miguel	Panama	9.08067	-79.62508	Hap 149 (STRI 15849)
P. gillii	182	Quebrada on Km 26 at Punta Pena - Almirante road.	Panama	9.06628	-82.29911	Hap 149 (STRI 12473)
P. gillii	183	Quebrada Jobito	Panama	9.06394	-80.18597	Hap 40 (STRI 15423, STRI 15424)
P. gillii	184	Rio Indio	Panama	9.06356	-80.18803	Hap 160 (STRI 16843)
P. gillii	185	Quebrada Tolu	Panama	9.04114	-80.35497	Hap 40 (STRI 15271, STRI 15281)
P. gillii	186	Rio Róbalo	Panama	9.04056	-82.28583	Hap 149 (STRI 6846, STRI 6847)
P. gillii	187	Rio Chagres	Panama	9.02530	-79.69890	Hap 39 (STRI 7363, STRI 7364), Hap 40 (STRI 7417, STRI 7421)
P. gillii	188	Rio Cardenas	Panama	9.00117	-79.57278	Hap 40 (STRI 15863, STRI 15864)
P. gillii	189	Rio Guarumo	Panama	9.00000	-82.18333	Hap 149 (STRI 12521)
P. gillii	190	Rio Guasimo	Panama	8.99133	-80.27433	Hap 40 (STRI 16045), Hap 167 (STRI 16043)
P. gillii	191	Rio Guasimo	Panama	8.99125	-80.27442	Hap 167 (STRI 15446), Hap 168 (STRI 15445)
P. gillii	192	Rio La Jacinta	Panama	8.96864	-80.52950	Hap 41 (STRI 16615, STRI 16616)
P. gillii	193	Quebrada Los Uveros	Panama	8.94719	-80.13825	Hap 41 (STRI 16131), Hap 169 (STRI 16123)
P. gillii	194	Creek on Punta Peña, Rio Punta Agua Real	Panama	8.94608	-82.15711	Hap 149 (STRI 16197)
P. gillii	195	Rio Canaveral	Panama	8.92858	-81.71180	Hap 166 (STRI 736)
P. gillii	196	Rio Guarumo	Panama	8.92853	-82.18028	Hap 149 (STRI 6379), Hap 170 (STRI 6380)
P. gillii	197	Rio Victoria	Panama	8.92514	-80.55172	Hap 171 (STRI 16562), Hap 172 (STRI 16578)
P. gillii	198	Rio Victoria	Panama	8.92500	-80.55139	Hap 171 (STRI 15409)
P. gillii	199	Rio Cricamola	Panama	8.91728	-81.87725	Hap 149 (STRI 12376), Hap 173 (STRI 12375)
P. gillii	200	Rio Toabré	Panama	8.91544	-80.50058	Hap 174 (STRI 15365)
P. gillii	201	Rio Toabre at Quebrada Patatilla	Panama	8.91533	-80.50067	Hap 41 (STRI 16686), Hap 171 (STRI 16679)

P. gillii	202	Quebrada Congal	Panama	8.91411	-80.13406	Hap 169 (STRI 16149)
P. gillii	204	Quebrada Tortuguita	Panama	8.88128	-80.39089	Hap 156 (STRI 15655)
P. gillii	205	Rio Uracillo	Panama	8.88086	-80.21983	Hap 176 (STRI 16094, STRI 16096)
P. gillii	206	Quebrada Platanal	Panama	8.87989	-80.27689	Hap 176 (STRI 15532)
P. gillii	207	Rio Toabre at Quebrada Tortuguita	Panama	8.87861	-80.39047	Hap 177 (STRI 16695, STRI 16718)
P. gillii	208	Punta Peña	Panama	8.87594	-82.17461	Hap 149 (STRI 16199)
P. gillii	209	Rio Canazas at Chiriqui Grande road.	Panama	8.87333	-82.17444	Hap 149 (STRI 11625)
P. gillii	210	Rio Guarumo	Panama	8.87250	-82.18933	Hap 149 (STRI 642)
P. gillii	211	Rio Caimito	Panama	8.85083	-79.96056	Hap 132 (STRI 4796, STRI 4799)
P. gillii	212	Rio Cocle del Norte	Panama	8.82147	-80.53356	Hap 41 (STRI 16529), Hap 171 (STRI 16536)
P. gillii	213	Rio Cocle del Norte	Panama	8.81867	-80.55302	Hap 41 (STRI 1347)
P. gillii	214	Rio Botija	Panama	8.81200	-80.57972	Hap 41 (STRI 16497), Hap 171 (STRI 16498)
P. gillii	215	Rio Cascajal	Panama	8.80464	-80.53328	Hap 171 (STRI 15330, STRI 15331)
P. gillii	216	Rio Cocle del Norte	Panama	8.80417	-80.58083	Hap 39 (STRI 9773, STRI 9776), Hap 41 (STRI 9768, STRI 9774)
P. gillii	217	Rio Moreno	Panama	8.77942	-80.53447	Hap 41 (STRI 15319)
P. gillii	218	Rio Cocle del Norte	Panama	8.77450	-80.52783	Hap 132 (STRI 9788), Hap 149 (STRI 9791), Hap 171 (STRI 9780)
P. gillii	219	Rio Moreno	Panama	8.76667	-80.53614	Hap 41 (STRI 16479, STRI 16483)
P. gillii	220	Rio Chiriqui Viejo	Panama	8.76443	-82.82712	Hap 41 (STRI 112), Hap 132 (STRI 111)
P. gillii	221	Rio Calovebora	Panama	8.74778	-81.22310	Hap 137 (STRI 6885, STRI 6888–90), Hap 171 (STRI 6886)
P. gillii	222	Rio Chiriqui	Panama	8.68803	-82.29172	Hap 40 (STRI 18683)
P. gillii	224	Rio Anton	Panama	8.59719	-80.13775	Hap 39 (STRI 18589)
P. gillii	226	Rio Santa Maria	Panama	8.41322	-82.04800	Hap 40 (STRI 17120)
P. gillii	227	Quebrada El Nance, trib. to Rio Santa Maria	Panama	8.41317	-81.04850	Hap 160 (STRI 11162)
P. gillii	228	Rio Anton	Panama	8.39680	-80.25851	Hap 40 (STRI 1118, STRI 1119)
P. gillii	229	Rio Santa Maria	Panama	8.35278	-80.79923	Hap 132 (STRI 3141–48)
P. gillii	230	Rio Salado	Panama	7.70497	-80.27814	Hap 155 (STRI 18720)
P. gillii	169	Quebrada Garay	Panama	9.19575	-82.34311	Hap 132 (STRI 18812), Hap 155 (STRI 18801)
P. latipinna **(OG)**	1	Cape Fear, Wilmington, North Carolina	United States of	34.24200	-77.95521	Hap 2 (Plati-0)

			America			
***P. latipinna* (OG)**	2	Mounds Pool, St. Mark's Wildlife Refuge, Wakulla, Florida	United States of America	30.09728	-84.15385	Hap 3 (Plat4)
***P. latipinna* (OG)**	3	Brownsville, Texas	United States of America	25.88443	-97.47564	Hap 4 (Plat5)

Abbreviations: MLBM, Brigham Young University Fish Collection (Monte L. Bean Museum); N/A, not available; STRI, Smithsonian Tropical Research Institute Neotropical Fish Collection; trib., tributary.

Sampling localities for *Roeboidies* spp.

Species	Site No.	Locality	Country	Latitude	Longitude	Cyt*b* haplotype (sample IDs)
Roeboides bouchellei	13	Rio Frío	Costa Rica	11.03004	-84.71821	Hap 3 (MLBM 160859), Hap 11 (MLBM 160862)
R. bouchellei	14	Rio Haciendas	Costa Rica	10.95468	-85.13692	Hap 12 (MLBM 161766), Hap 13 (MLBM 161799), Hap 14 (MLBM 161800)
R. bouchellei	15	Rio Tempisquito	Costa Rica	10.78553	-85.55441	Hap 15 (MLBM 168840), Hap 16 (MLBM 168841, MLBM 168914, MLBM 168915), Hap 17 (MLBM 168842)
R. bouchellei	16	Rio Liberia at outskirts of Liberia	Costa Rica	10.62745	-85.43412	Hap 15 (MLBM 168999–9006)
R. bouchellei	17	Rio Tempisque on road between Guardia and Comunidad, Nicoya Peninsula	Costa Rica	10.57220	-85.58477	Hap 15 (MLBM 168904), Hap 16 (MLBM 168903, MLBM 168907), Hap 18 (MLBM 168905, MLBM 168906)
R. bouchellei	18	Rio Salto at CA1 southeast of Liberia, Nicoya Peninsula	Costa Rica	10.56106	-85.39192	Hap 16 (MLBM 168311–16, MLBM 168308, MLBM 168309)
R. bouchellei	19	Small stream feeding into Canal B-9	Costa Rica	10.52452	-84.03127	Hap 19 (MLBM 118369)
R. bouchellei	20	Trib. to Rio Sarapiquí	Costa Rica	10.52452	-84.03127	Hap 20 (MLBM 118312)
R. bouchellei	21	Rio Sardinal at Sardinal, on Costa Rica Hwy 151, approximately 5 km west of Costa Rica Hwy 21, Nicoya Peninsula	Costa Rica	10.51508	-85.65166	Hap 16 (MLBM 169128–30), Hap 21 (MLBM 169127)
R. bouchellei	22	Rio Cabuyo at unnamed province road to the Reserva Biologica Lomas Bardudal	Costa Rica	10.48961	-85.38555	Hap 16 (MLBM 168912, MLBM 168913, MLBM 169161–68)
R. bouchellei	23	Unnamed trib. to Rio Tempisque drainage approximately 2 km south of Belén, Nicoya Peninsula	Costa Rica	10.39076	-85.59045	Hap 16 (MLBM 168064, MLBM 168437, MLBM 168910, MLBM 168911)
R. bouchellei	24	Rio Lajas	Costa Rica	10.32095	-85.04733	Hap 16 (MLBM 160919, MLBM 160920)
R. bouchellei	25	Rio Diriá at CA1 approximately 2-3 km north of Santa Cruz, Nicoya Peninsula	Costa Rica	10.26677	-85.59261	Hap 16 (MLBM 168625)
R. bouchellei	26	Brazo del Sucio	Costa Rica	10.21872	-83.90466	Hap 3 (MLBM 160876)
R. bouchellei	27	Trib. to Rio Parismina	Costa Rica	10.19968	-83.65873	Hap 22 (MLBM 118298, MLBM 118302, MLBM 118305, MLBM 118311, MLBM 118314–16, MLBM 118318–20), Hap 24 (MLBM 118321)
R. bouchellei	28	Rio Morote at Costa Rica Hwy 21 approximately 6 km north of Carmona	Costa Rica	10.05828	-85.26202	Hap 25 (MLBM 118225)

R. bouchellei	1	Trib. to unnamed trib. of the Rio Grande de Matagalpa, approximately 32 km west of Rio Blanco (town) on road between Matagalpa and Rio Blanco (town)	Nicaragua	12.82341	-85.44279	Hap 1 (MLBM 168957, MLBM 168958)
R. bouchellei	2	Unnamed trib. to Rio Malacatoya at Teustepe, just off road to Rama (Nicaragua Hwy 7), southwest of Boaco	Nicaragua	12.41825	-85.79299	Hap 2 (MLBM 168627), Hap 3 (MLBM 168628–30, MLBM 168626)
R. bouchellei	3	Rio Malacatoya	Nicaragua	12.32661	-85.95553	Hap 2 (MLBM 168922), Hap 3 (MLBM 172858), Hap 4 (MLBM 172857)
R. bouchellei	4	Rio Cuisala	Nicaragua	12.26707	-85.65119	Hap 5 (MLBM 173085), Hap 6 (MLBM 173086)
R. bouchellei	5	Rio Tipitapa	Nicaragua	12.20267	-86.10208	Hap 3 (MLBM 172813–15)
R. bouchellei	6	Unnamed trib. at km marker 18.5 (from center of Managua) on CA1, just southwest of town of Tipitapa	Nicaragua	12.17289	-86.11618	Hap 3 (MLBM 170087)
R. bouchellei	7	Unnamed trib./headwater of Rio Mico northeast of San Pedro de Lovago	Nicaragua	12.13630	-85.04597	Hap 7 (MLBM 168916–21, MLBM 169047), Hap 8 (MLBM 169048)
R. bouchellei	8	Rio Mico	Nicaragua	12.07401	-84.53689	Hap 3 (MLBM 173376)
R. bouchellei	9	Unnamed trib.	Nicaragua	12.00340	-85.18153	Hap 9 (MLBM 173842)
R. bouchellei	10	Unnamed trib.	Nicaragua	11.73855	-84.51068	Hap 3 (MLBM 173688–90)
R. bouchellei	11	Unnamed trib. to Rio Ochomogo	Nicaragua	11.67886	-85.98817	Hap 3 (MLBM 172551)
R. bouchellei	12	Unnamed trib. to Lago Nicaragua just east of mile marker km 238 on road to San Miguelito, Chontales	Nicaragua	11.50538	-84.83956	Hap 10 (MLBM 169660)
Roeboides bussingi	29	Unnamed small stream feeds into Canal B-9, Puntarenas	Costa Rica	8.65508	-82.94633	Hap 19 (BelkRo4–7, MLBM 118360, MLBM 118361, MLBM 118365), Hap 26 (BelkRo1), Hap 27 (BelkRo2)

Abbreviations: MLBM, Brigham Young University Fish Collection (Monte L. Bean Museum); trib., tributary.

Sampling localities for *Astyanax* spp.

Species	**Locality**	**Country**	**Latitude**	**Longitude**	**Cyt*b* haplotype (sample IDs or GenBank numbers)**
Astyanax aeneus	Rio Mopan	Guatemala	N/A	N/A	Hap 57 (FJ439314)
A. aeneus	Crooked Tree A.	Guatemala or Mexico	N/A	N/A	Hap 58 (FJ439312)
A. aeneus	Rio Nuevo	Guatemala or Mexico	N/A	N/A	Hap 57 (FJ439311)
A. aeneus	Atoyac R.	Mexico	N/A	N/A	Hap 88 (FJ439336)
A. aeneus	Candelaria R.	Mexico	N/A	N/A	Hap 53 (FJ439220)
A. aeneus	Chacamax R.	Mexico	N/A	N/A	Hap 50 (FJ439192)
A. aeneus	Laguna Chinchancanab	Mexico	N/A	N/A	Hap 57 (AY177210)
A. aeneus	Cupatizio R.	Mexico	N/A	N/A	Hap 81 (FJ439203)
A. aeneus	El Carmen C.	Mexico	N/A	N/A	Hap 53 (FJ439247, FJ439248)

A. aeneus	Itzamatitlan R.	Mexico	N/A	N/A	Hap 84 (FJ439249)
A. aeneus	La Media Luna A.	Mexico	N/A	N/A	Hap 82 (FJ439226)
A. aeneus	Mamantel R.	Mexico	N/A	N/A	Hap 62 (FJ439223)
A. aeneus	Noc - Ac C.	Mexico	N/A	N/A	Hap 52 (FJ439181, FJ439182)
A. aeneus	Ojo de Agua S.	Mexico	N/A	N/A	Hap 63 (FJ439193)
A. aeneus	Parque Uruapan	Mexico	N/A	N/A	Hap 83 (FJ439215)
A. aeneus	Puente Nacional R.	Mexico	N/A	N/A	Hap 87 (FJ439198)
A. aeneus	Rio Grande	Mexico	N/A	N/A	Hap 49 (FJ439251), Hap 67 (FJ439194)
A. aeneus	Rio Papaloapan	Mexico	N/A	N/A	Hap 68 (FJ439197)
A. aeneus	Salado R.	Mexico	N/A	N/A	Hap 69 (FJ439250)
A. aeneus	San Antonio R.	Mexico	N/A	N/A	Hap 61 (FJ439221)
A. aeneus	Sian Ka´an C.	Mexico	N/A	N/A	Hap 59 (FJ439237)
A. aeneus	T´Sil A.	Mexico	N/A	N/A	Hap 60 (FJ439238)
A. aeneus	Tamazula R.	Mexico	N/A	N/A	Hap 80 (FJ439224), Hap 89 (FJ439252)
A. aeneus	Tolome R.	Mexico	N/A	N/A	Hap 85 (FJ439195), Hap 86 (FJ439196)
A. aeneus	Lago Catemaco	N/A	N/A	N/A	Hap 56 (FJ439187), Hap 70 (FJ439189), Hap 71 (FJ439243), Hap 74 (FJ439188), Hap 75 (FJ439190), Hap 76 (FJ439244)
A. aeneus	Lago Chalchoapan	N/A	N/A	N/A	Hap 65 (FJ439205)
A. aeneus	Chuniapan R.	N/A	N/A	N/A	Hap 54 (FJ439341), Hap 55 (FJ439342), Hap 73 (FJ439184)
A. aeneus	Cuetzalapan R.	N/A	N/A	N/A	Hap 56 (FJ439343)
A. aeneus	El Saltillo R.	N/A	N/A	N/A	Hap 64 (FJ439185)
A. aeneus	Salinas R.	N/A	N/A	N/A	Hap 66 (FJ439183)
A. aeneus	San Joaquin R.	N/A	N/A	N/A	Hap 77 (FJ439191), Hap 76 (FJ439245), Hap 78 (FJ439246), Hap 79 (FJ439242)
A. aeneus	Xoteapan R.	N/A	N/A	N/A	Hap 72 (FJ439186)
Astyanax belizanus	Amatillo R.	Guatemala	15.55402	88.93545	Hap 91 (FJ439272)
A. belizanus	Chahuacal R.	Guatemala	15.55402	-88.93545	Hap 92 (FJ439273)
A. belizanus	Lago Izabal	Guatemala	N/A	N/A	Hap 94 (FJ439318)
A. belizanus	Los Amates R.	Guatemala	15.26134	-89.08810	Hap 96 (FJ439321)
A. belizanus	Puente Virginia R.	Guatemala	15.43946	-88.95288	Hap 93 (FJ439271)

A. belizanus	Rio Morazan	Guatemala	N/A	N/A	Hap 97 (FJ439335)
A. belizanus	Rio San Pedro	Guatemala	N/A	N/A	Hap 95 (FJ439319)
A. belizanus	Staan Creek	Mexico	16.99660	-88.31333	Hap 91 (FJ439313)
***Astyanax bimaculatus* (OG)**	Argentina	Argentina	N/A	N/A	Hap 98 (FJ439334)
***Astyanax fascitus* (OG)**	Brazil	Brazil	N/A	N/A	Hap 206 (AY177205), Hap 207 (AY177206)
Astyanax hubbsi	Peñon Blanco	Mexico	N/A	N/A	Hap 104 (FJ439210)
A. hubbsi	San Juan S.	Mexico	N/A	N/A	Hap 101 (FJ439213)
A. hubbsi	El Ahuaje S.	Mexico	N/A	N/A	Hap 102 (FJ439229)
A. hubbsi	Tamasopo A.	Mexico	N/A	N/A	Hap 103 (FJ439236)
Astyanax mexicanus	Bobo R.	Mexico	N/A	N/A	Hap 111 (FJ439255)
A. mexicanus	Cariño de la Montaña R.	Mexico	N/A	N/A	Hap 128 (FJ439208), Hap 129 (FJ439211), Hap 130 (FJ439232)
A. mexicanus	Cuatro Cienegas A.	Mexico	N/A	N/A	Hap 126 (FJ439216)
A. mexicanus	Dos Arroyos R.	Mexico	N/A	N/A	Hap 110 (FJ439256)
A. mexicanus	El Cuarto S.	Mexico	N/A	N/A	Hap 132 (FJ439209), Hap 133 (FJ439218)
A. mexicanus	Falcon Dam	Mexico	N/A	N/A	Hap 123 (FJ439201)
A. mexicanus	Güemez R.	Mexico	N/A	N/A	Hap 120 (FJ439199)
A. mexicanus	Huichihuayan R.	Mexico	N/A	N/A	Hap 131 (FJ439228)
A. mexicanus	Jalpan Cave	Mexico	N/A	N/A	Hap 117 (FJ439344), Hap 135 (FJ439337)
A. mexicanus	Jalpan R.	Mexico	N/A	N/A	Hap 118 (FJ439346)
A. mexicanus	La Nutria R.	Mexico	N/A	N/A	Hap 124 (FJ439202)
A. mexicanus	Peñon Blanco	Mexico	N/A	N/A	Hap 134 (FJ439219)
A. mexicanus	San Bernabe S.	Mexico	N/A	N/A	Hap 125 (FJ439233)
A. mexicanus	San Juan S.	Mexico	N/A	N/A	Hap 108 (FJ439214), Hap 109 (FJ439227)
A. mexicanus	Santa Maria R.	Mexico	N/A	N/A	Hap 119 (FJ439225)
A. mexicanus	Chica Cave	N/A	N/A	N/A	Hap 113 (FJ439253), Hap 208 (AY639041)
A. mexicanus	El Limón R.	N/A	N/A	N/A	Hap 90 (FJ439258), Hap 116 (FJ439235)
A. mexicanus	El Nacimiento S.	N/A	N/A	N/A	Hap 105 (FJ439259)

A. mexicanus	Frio R.	N/A	N/A	N/A	Hap 112 (FJ439254)
A. mexicanus	Guayalejo R.	N/A	N/A	N/A	Hap 114 (FJ439206)
A. mexicanus	Mante Dam	N/A	N/A	N/A	Hap 121 (FJ439207), Hap 122 (FJ439200)
A. mexicanus	Molino Cave	N/A	N/A	N/A	Hap 211 (AY639046)
A. mexicanus	Oyul Dam	N/A	N/A	N/A	Hap 106 (FJ439257)
A. mexicanus	Pachon Cave	N/A	N/A	N/A	Hap 210 (AY639043)
A. mexicanus	Puente La Raya R.	N/A	N/A	N/A	Hap 115 (FJ439204)
A. mexicanus	Yerbaniz Cave	N/A	N/A	N/A	Hap 209 (AY639042)
Astyanax nasutus	Large dry riverbed area where Rio Los Quesos and Rio Los Encuentros meet, off Nicaragua Hwy 38, northeast of San Juan de Limay	Nicaragua	13.18695	-86.59486	Hap 137 (FJ439305)
A. nasutus	Rio Acuitanca in La Trinidad, Departamento Carazo	Nicaragua	11.73834	-86.33671	Hap 138 (FJ439302)
A. nasutus	Rio Atoya at Nicaragua Hwy 12 near Santa Maria, approximately 10 km northwest of El Viejo, Chinandega	Nicaragua	12.69237	-87.25532	Hap 140 (FJ439307)
A. nasutus	Rio Casares at dirt road at El Barranco, Departamento Carazo	Nicaragua	11.67420	-86.30531	Hap 47 (FJ439303)
A. nasutus	Rio Estero San Antonio at CA1 at Las Maderas (San Antonio)	Nicaragua	12.44898	-86.03985	Hap 142 (FJ439300), Hap 143 (FJ439301)
A. nasutus	Rio Telica at Nicaragua Hwy 12 near Telica, Leon	Nicaragua	12.51705	-86.86495	Hap 139 (FJ439306)
A. nasutus	Rio Villanueva at Nicaragua Hwy 24, ~4 km southwest of Villa Nueva, and 13 km south of Somotillo	Nicaragua	12.94597	-86.84813	Hap 141 (FJ439304)
Astyanax nicaraguensis	Main stem Rio San Juan near mouth/entry of Rio San Carlos	Costa Rica	10.78995	-84.19511	Hap 146 (FJ439294), Hap 151 (FJ439295)
A. nicaraguensis	Rio Barbilla at province road just west of B-Line, a few km west of the road to Matina (dirt road, off northeast side of big road to Puerto Limon)	Costa Rica	10.04416	-83.33383	Hap 161 (FJ439292)
A. nicaraguensis	Rio Puerto Viejo just off Costa Rica Hwy 4 approximately 4 km north of Buenos Aires	Costa Rica	10.38078	-83.97533	Hap 160 (FJ439293)
A. nicaraguensis	Rio Sabalo, near Nicaragua-Costa Rica border, approximately 25 km north on dirt road (taken near Santa Cecilia)	Costa Rica	11.15054	-85.47914	Hap 150 (FJ439291)
A. nicaraguensis	Rio Sarapiquí at Costa Rica Hwy 4 just south of Puerto Viejo de Sarapiquí	Costa Rica	10.44869	-84.01070	Hap 156 (FJ439217)
A. nicaraguensis	Trib. to Rio Infernito	Costa Rica	10.61802	-84.48418	Hap 162 (FJ439355)
A. nicaraguensis	Trib. to Rio Reventaza	Costa Rica	9.87230	-83.63320	Hap 146 (FJ439356), Hap 163 (FJ439358)
A. nicaraguensis	Trib. to Rio San Carlos, San Carlos Canton, near Angeles	Costa Rica	10.52224	-84.31873	Hap 155 (FJ439222)
A. nicaraguensis	Rio Brito (or trib.) off Nicaragua Hwy 62	Nicaragua	11.43526	-85.92319	Hap 12 (FJ439289)

	approximately 1 km east of Tola				
A. nicaraguensis	Rio Coco near Quilali at Nicaragua Hwy 51	Nicaragua	13.55911	-86.01535	Hap 164 (FJ439279)
A. nicaraguensis	Rio Compazagua, trib. to Rio Grande de Matagalpa, at Muy-Muy	Nicaragua	12.77711	-85.64406	Hap 154 (FJ439285)
A. nicaraguensis	Rio Gil Gonzalez trib. to Lago Nicaragua, at Carretera Panamericana / Nicaragua Hwy 2, at Los Viejitos	Nicaragua	11.53677	-85.89743	Hap 12 (FJ439359), Hap 147 (FJ439288)
A. nicaraguensis	Rio Grande de Matagalpa off big dirt road west of San Jose de Murra	Nicaragua	12.62368	-85.22708	Hap 27 (FJ439277), Hap 153 (FJ439283)
A. nicaraguensis	Rio Malacatoya	Nicaragua	12.32661	-85.95553	Hap 144 (FJ439298), Hap 152 (FJ439296)
A. nicaraguensis	Rio Mico of dirt road 600 m to south side of Nicaragua Hwy 7 (road to Rama), just east of El Recreo and ~13-14 km west of Rama	Nicaragua	12.17404	-84.31475	Hap 159 (FJ439299)
A. nicaraguensis	Rio Nauawas, trib. To Rio Siquia, ~14-16 km east on road that splits off to right/east from Nicaragua Rd 108 northeast of El Ayote	Nicaragua	12.52584	-84.67795	Hap 149 (FJ439280)
A. nicaraguensis	Rio Ochomogo 300 m west of Nicaragua Hwy 1, near Paso Real de Ochomogo	Nicaragua	11.65664	-85.97319	Hap 145 (FJ439287)
A. nicaraguensis	Rio Sinecapa, trib. to Lago Managua, just off Nicaragua Hwy 26 at La Empalme, located between Estelí and León (approximately 30-40 km west of Estelí)	Nicaragua	12.67376	-86.42606	Hap 157 (FJ439286)
A. nicaraguensis	Small trib. approximately 1 km west of Camoapa at Nicaragua Hwy 17	Nicaragua	12.37781	-85.53095	Hap 158 (FJ439297)
A. nicaraguensis	Trib. At north side tip of Lago Apanas, at Nicaragua Hwy 43	Nicaragua	13.25885	-85.91207	Hap 153 (FJ439281)
A. nicaraguensis	Trib. to Rio Chiquito southeast of Nueva Guinea, RAAS	Nicaragua	11.66814	-84.42363	Hap 148 (FJ439284)
A. nicaraguensis	Unnamed trib. just 500 m east of Santa Rita at Carretera Vieja a Leon, and ~1 km west of Los Cedros	Nicaragua	12.07224	-86.49273	Hap 136 (FJ439278)
A. nicaraguensis	Viejo R.	Nicaragua	12.51843	-85.77606	Hap 13 (FJ439290), Hap 153 (FJ439282)
Astyanax orthodus	Trib. to Rio Sixaola	Costa Rica	9.62093	-82.85768	Hap 165 (FJ439357), Hap 165 (MLBM 118297)
Astyanax petenensis	Lago Peten-Itza	Guatemala	N/A	N/A	Hap 166 (FJ439315), Hap 167 (FJ439316)
A. petenensis	Rio Cansis	Guatemala	N/A	N/A	Hap 169 (FJ439320)
A. petenensis	San Juan/Peten R.	Guatemala	N/A	N/A	Hap 168 (FJ439317)
A. petenensis	Candelaria R.	N/A	N/A	N/A	Hap 168 (FJ439325), Hap 173 (FJ439326), Hap 174 (FJ439327)
A. petenensis	Candelaria-Yalicar R.	N/A	N/A	N/A	Hap 178 (FJ439332, FJ439333)
A. petenensis	Rio Chajmaic	N/A	N/A	N/A	Hap 172 (FJ439322)
A. petenensis	Rio San Simon	N/A	N/A	N/A	Hap 175 (FJ439328), Hap 176 (FJ439329), Hap 177 (FJ439330)

A. petenensis	Semococh R.	N/A	N/A	N/A	Hap 170 (FJ439323), Hap 171 (FJ439324)
***Astyanax* sp. Novo 1**	La Guija A.	El Salvador	14.32890	-89.43731	Hap 186 (FJ439239), Hap 187 (FJ439240)
***A.* sp. Novo 1**	Pachipa R.	Guatemala	N/A	N/A	Hap 183 (FJ439262), Hap 184 (FJ439263)
***A.* sp. Novo 1**	Bolas R.	Mexico	N/A	N/A	Hap 185 (FJ439261)
***A.* sp. Novo 1**	Chifle R.	Mexico	N/A	N/A	Hap 193 (FJ439260)
***A.* sp. Novo 1**	El Sardinero R.	Mexico	N/A	N/A	Hap 192 (FJ439231)
***A.* sp. Novo 1**	Huehuetan R.	Mexico	N/A	N/A	Hap 190 (FJ439234), Hap 191 (FJ439230)
***A.* sp. Novo 1**	Jocotal A.	Mexico	13.66490	-88.42560	Hap 188 (FJ439241)
***A.* sp. Novo 1**	Pichoacan R.	Mexico	N/A	N/A	Hap 194 (FJ439264)
***Astyanax* sp. Novo 2**	Maquinas R.	Mexico	N/A	N/A	Hap 180 (FJ439266), Hap 200 (FJ439339)
***A.* sp. Novo 2**	La Palma R.	N/A	N/A	N/A	Hap 195 (FJ439265)
***Astyanax* sp. Novo 3**	Montebello A.	Mexico	16.10104	-91.73859	Hap 182 (FJ439268)
***Astyanax* sp. Novo 5**	Rio Chires at Costa Rica Rd 239	Costa Rica	9.58295	-84.41170	Hap 203 (FJ439352)
***Astyanax* sp. Novo 6**	Rio Irigaray in Irigaray at CA 1, approximately 2.5 km west of Canas Dulces	Costa Rica	10.72340	-85.51038	Hap 201 (FJ439354)
***A.* sp. Novo 6**	Rio Ciruelas at Costa Rica Hwy 1 (CA 1), approximately 1.8 km northwest of CA 1-Costa Rica Rd 144 intersection	Costa Rica	10.06025	-84.75861	Hap 199 (FJ439274)
***A.* sp. Novo 6**	Rio Colorado, trib. to Rio Tempisque, at Costa Rica Hwy 1 (CA1)	Costa Rica	10.66885	-85.48133	Hap 198 (FJ439275)
***Astyanax* sp. Novo 7**	Quebrada Arena, trib. to Rio Blanco, at site on south side of Fortuna	Costa Rica	10.67217	-85.19942	Hap 197 (FJ439353), Hap 202 (FJ439351), Hap 204 (FJ439349)
***A.* sp. Novo 7**	Rio Lagarto at Costa Rica Hwy 2 (Carretera Interamericana) approximately 2-3 km west of Rio Claro / Finca Rio Claro	Costa Rica	8.68180	-83.07630	Hap 197 (FJ439348)
***Astyanax* sp. Novo 8**	Pipeline Road (Camino del Oleoducto), near Panama Canal, just northwest of Gamboa off northeast side of Av Omar Torrijos Herrera	Panama	9.12835	-79.71527	Hap 48 (FJ439308)
***A.* sp. Novo 8**	Rio Chagres at Panama Hwy 9 bridge, off Panama-Colon Expressway	Panama	9.19298	-79.65170	Hap 48 (FJ439309)
***Astyanax* sp. Novo 9**	Cahabon R.	Mexico	N/A	N/A	Hap 179 (FJ439270)
***A.* sp. Novo 9**	Jeronimo R.	Mexico	N/A	N/A	Hap 179 (FJ439269)
***A.* sp. Novo 9**	Arroyo Sachicha R.	N/A	N/A	N/A	Hap 189 (FJ439331)
Astyanax aeneus	Rio Tempisquito	Costa Rica	10.78553	-85.55441	Hap 8 (MLBM 168855, MLBM 168858–60)
A. aeneus	Trib. to Rio Madre de Dios approximately 17 km	Costa Rica	10.07055	-83.38631	Hap 9 (MLBM 168825), Hap 10 (MLBM 168823), Hap 14 (MLBM

	east of Siquirres on Costa Rica Rd 32				168830)
A. aeneus	Rio Caracol	Nicaragua	12.35116	-85.88870	Hap 24 (MLBM 173000)
A. aeneus	Unnamed trib. to Rio Grande de Matagalpa, at stream just west of km marker 226	Nicaragua	13.25769	-85.45440	Hap 6 (MLBM 170009), Hap 20 (MLBM 170010), Hap 21 (MLBM 170007), Hap 23 (MLBM 170011), Hap 25 (MLBM 170012)
A. aeneus	Unnamed trib. to Rio Malacatoya at Teustepe, just off road to Rama (Nicaragua Hwy 7), southwest of Boaco	Nicaragua	12.41825	-85.79299	Hap 11 (MLBM 169913), Hap 13 (MLBM 169917)
A. aeneus	Unnamed trib. to Rio Murra at Nicaragua Hwy 19 about halfway between Boaco and Camoapa	Nicaragua	12.45513	-85.53541	Hap 18 (MLBM 173081), Hap 22 (MLBM 173078)
A. aeneus	Unnamed trib. to Rio Ojucuapa, trib. to Lago Nicaragua, in cow pasture just off road ~5.5 km northeast of El Guasimo	Nicaragua	11.82942	-85.20479	Hap 12 (MLBM 174008), Hap 16 (MLBM 174105), Hap 19 (MLBM 174108)
A. aeneus	Unnamed trib./headwater of Rio Mico northeast of San Pedro de Lovago	Nicaragua	12.13630	-85.04597	Hap 5 (MLBM 169143), Hap 15 (MLBM 169146), Hap 17 (MLBM 169144), Hap 18 (MLBM 169145)
***Astyanax* sp.**	Rio Lajas	Costa Rica	10.32095	-85.04733	Hap 7 (MLBM 160912)
***A.* sp.**	Rio Negro, Golfo de Fonseca	El Salvador	13.06210	-87.10649	Hap 26 (82627), Hap 33 (82619), Hap 34 (82624), Hap 43 (82628), Hap 46 (81020), Hap 47 (82623)
***A.* sp.**	Rio Pespire, trib. to Rio Nacaome (near El Salvador-Honduras border)	Honduras	13.68701	-87.34051	Hap 26 (82997), Hap 31 (82657), Hap 47 (83000)
***A.* sp.**	Rio Los Almendros, trib. to Rio Patuca	Honduras	14.06847	-86.34785	Hap 27 (81598), Hap 34 (80701, 80713), Hap 35 (80708), Hap 36 (81599), Hap 37 (80969), Hap 40 (81600), Hap 41 (81597), Hap 44 (80947)
***A.* sp.**	Rio Motagua-Chamelecón-Ulúa	Honduras	15.21000	-88.56881	Hap 28 (82951), Hap 29 (82964), Hap 30 (82932)
***A.* sp.**	Rio Danto, trib. to Rio Nombre de Dios	Honduras	15.73547	-86.78287	Hap 27 (80920), Hap 32 (81603), Hap 39 (81616), Hap 45 (80878)
***A.* sp.**	N/A	N/A	N/A	N/A	Hap 34 (TH0726), Hap 38 (TH0724), Hap 42 (81582)
***Brycon guatemalensis* (OG)**	Rio Sinecapa, trib. to Lago Managua, just off Nicaragua Hwy 26 at La Empalme, located between Estelí and León (approximately 30-40 km west of Estelí)	Nicaragua	12.67376	-86.42606	Hap 1 (168335N, 168338N, 168339N, 168342N), Hap 2 (168336N, 168337N, 168341N), Hap 3 (168340N)
***Roeboides salvadoris* (OG)**	El Salvador	El Salvador	N/A	N/A	Hap 212 (FJ439180)

Individuals with "A.", "C." and "R." abbreviations after locality names that generally lack geographical coordinate data and are listed with GenBank numbers in the right column are from Ornelas-García *et al.* (2008). Abbreviations: MLBM, Brigham Young University Fish Collection (Monte L. Bean Museum); N/A, not available; OG, outgroup; sp., species; TH, field numbers for samples collected by Wilfredo Matamoros and/or Michael Tobler in Honduras; trib., tributary.

References

1. Hrbek T, Seckinger J, Meyer A (2007) A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. *Molecular Phylogenetics and Evolution*, **43**(3), 986-998.
2. Ornelas-García CP, Dominguez-Dominguez O, Doadrio I (2008) Evolutionary history of the fish genus *Astyanax* Baird &

Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evolutionary Biology*, **8**, 340.

3. Tobler M, Palacios M, Chapman LJ, Mitrofanov I, Bierbach D, Plath M, Arias-Rodriguez L, Garcia de Leon FJ, Mateos M (2011) Evolution in extreme environments: replicated phenotypic differentiation in livebearing fish inhabiting sulfidic springs. *Evolution*, **65**, 2213-2228.

Appendix S1 Supplementary methods and results.

Additional details on hypotheses of Neotropical fish diversification for CA

We tested several non-mutually exclusive hypotheses of Neotropical freshwater fish diversification focused on the predicted genetic impacts of historical drainage-controlling processes. As discussed in the main text and shown in Table 1, we cast these hypotheses in the context of the Central American (CA) region and its freshwater fishes, rather than the North American and South American Neotropics as a whole. Others have already developed similar although less detailed hypotheses for South American freshwater fishes (e.g. Hubert & Renno 2006). Two of our hypotheses—the 'tectonic vicariance hypothesis' and 'marine vicariance hypothesis'—make explicit predictions with reference to the timing of vicariance events in specific areas where we hypothesized *a priori* that genetic breaks might be expected to occur (Table 1). However, the periods of diversification of CA freshwater fish lineages predicted by these two hypotheses overlap significantly. Here, we briefly discuss how these hypotheses can be distinguished, despite their similarities, based on contrasts in their predicted spatial patterns of reciprocal monophyly of populations (Table 1).

Tectonic vicariance mechanisms would predict population divergences across major mountain ranges and volcanoes, and across the Tárcoles River area (including Tárcoles gorge and the Herradura block; see text), consistent with the timing of geological events at these geographical barriers. Given major mountain ranges are predominantly northwest-trending in orientation throughout lower CA, we would expect genetic breaks to either cross these ranges or to parallel them, due to (post-colonization) vicariance related to specific volcanic eruptions and arc-perpendicular/parallel volcanic fallout (e.g. Fig. 1; Marshall 2007; Bagley & Johnson 2014a). However, in Nuclear CA, the CAVA is closer to the Pacific versant and volcanic regions (e.g. serranías) of the Chortis highlands are spread across hundreds of kilometers over the broad (>500 km wide) Chortis block, the geological block of intrusive rocks, sedimentary rocks, and Paleozoic–Neogene volcanic rocks imprinted atop the crystalline basement (continental crust) that underlies the NCA region (Rogers *et al.* 2002, 2007; Marshall 2007). The wider, complex geological formations in NCA make it more difficult to predict where we might expect vicariance events to have most likely occurred, and indicate that genetic breaks might be expected to occur over larger distances. However, the key event of the Miocene uplift of the Chortis highlands ~19–3.8 Ma (particularly volcanics overprinting the southern and central

Chortis terranes; "Miocene Ignimbrite province" of Rogers *et al.* 2002) suggests that we should expect vicariant isolation of widespread ancestral taxa to have occurred between the Guatemala–Honduras Atlantic coast and southern Nicaragua (Nicaraguan depression) to Costa Rica, or between the aforementioned Atlantic regions and Pacific El Salvador–Guatemala (if not localized at finer spatial scales at fault zones and volcanoes within Chortis itself). Given that the Miocene volcanics in southern Chortis extended essentially to the Pacific Ocean at the Gulf of Fonseca, we might also expect a phylogeographic break (or a remnant of one) across the Gulf of Fonseca. That is, unless changing sea levels and drainage divides during the Quaternary facilitated river anastomosis and gene flow across this region, erasing genetic signatures of earlier evolutionary divergences.

Compared with tectonic vicariance predictions, the marine vicariance hypothesis predicts that phylogeographic breaks will occur in a different spatial pattern involving genetic splits between clades from upland areas that served as freshwater refugia during sea-level highstands versus clades from recolonized lowland populations (Table 1). Evidence for refugia under this hypothesis would most likely take one of two forms, which might be considered sub-hypotheses under the marine vicariance hypothesis. First, divergence might have occurred between a coastal lineage and an older and possibly more genetically diverse lineage from an upland area further inland. However, there have been multiple opportunities for vicariance due to marine incursions during sea-level highstands of the Miocene–Pliocene, mid-Pliocene, and Pleistocene (Table 1). Thus, in this case, we would expect the pattern of vicariance to reflect the prevailing influence of the single highstand event that most heavily impinged on a given lineage (species), or the last among a series of highstands that impinged on that lineage. Or, secondly, divergence might occur between multiple upland and lowland clades in coastal drainages with ambiguous internal branching patterns among them, and short coalescent times leading to short poly/paraphyletic branching patterns with little or no genetic structure within each clade (Brunsfeld *et al.* 2001). In this latter case, we would expect to see a single timeframe for multiple colonization events, which we suspect would be more likely tied to late Pliocene–Pleistocene sea level fluctuations. Between these two sub-hypotheses of the more general marine vicariance hypothesis, we favor the first scenario because there is better consensus among geologists and paleoclimatologists that eustatic sea-level highstands of >20–50 m asl that could have significantly impacted the distributions and genetic variation of freshwater organisms occurred in the Miocene–Pliocene,

whereas the Pleistocene sea level record is known to have included very few significant sea-level highstands of more than 4–10 m asl (references and discussion in the main text). However, more recent events might be the only events it is possible to detect within some of our focal lineages, for example if they originated within or dispersed to the study area only relatively recently. Thus, we do not rule either of these scenarios out.

Taxon sampling and sequencing, and outgroups details

Here, we provide additional sampling, sequence data, and outgroup descriptions relevant to our analyses but not listed in the main text because space was prohibiting. We obtained samples throughout much of the study area for the three most widely distributed lineages—*Amatitlania* spp., *Astyanax* spp., and *Poecilia mexicana*, which commonly inhabit most habitats across major drainages in the region (e.g. Bussing 1998; Miller *et al.* 2005). However, our sampling for five lineages—*Alfaro cultratus*, *Phallichthys amates*, *Priapichthys annectens*, *Roeboides* spp., and *Xenophallus*—emphasized areas within the endemic ranges of these species between Nicaragua and Costa Rica. We sampled these five taxa wherever possible from Atlantic-coastal Nicaragua, across the lower Chortis highlands, spanning the lower Nicaraguan depression in many cases from the Great Lakes District to the Limón basin east of the Tortuguero lowlands, Costa Rica. Permission to undertake fieldwork for this study was obtained through permits issued to JCB and JBJ in Nicaragua by MARENA (Ministerio de Ambiente y Recursos Naturales; DGPN/DB-IC-009-2012; DGPN/DB-21-2012) and in Costa Rica by SINAC-MINAET (Ministerio de Ambiente Energía y Telecomunicaciones; Resolución No. 030-2010-SINAC, Resolución No. 134-2012-SINAC). New specimens were obtained through these collections under Brigham Young University Institutional Animal Care and Use Committee (IACUC) approval #12-0701.

Recently, we conducted phylogenetic analyses in BEAST as part of a species delimitation analysis of *Poecilia mexicana* in the context of the evolution of the larger species complex of which it is a part, the *P. sphenops* species complex (Bagley *et al.* in revision). As in that paper, the BEAST analyses we used to estimate a t_{MRCA} and sub-clade divergence times within the *P. mexicana* lineage in the present study utilized 21 outgroup tips representing 13 nominal poeciliid outgroup taxa (family Poeciliidae; refs. in Bagley *et al.* in revision). These outgroups included (1) *Poecilia caucana*, which is considered the sister taxon to the members of the *P. sphenops* species complex based on analyses by Alda *et al.* (2013); the "sail-fin" mollies (2) *P. latipinna*

and (3) *P. latipunctata*; the South American guppies (4) *Micropoecilia picta* and (5) *Poecilia reticulata*; the Mexican swordtails (6) *Xiphophorus helleri* and (7) *X. maculatus*; the Central American Pike Killifish, (8) *Belonesox belizanus*; and five species of fishes from the genus *Limia*, a closely related genus whose members were formerly included within *Poecilia* subgenus *Limia*: (9) *L. dominicensis*, (10) *L. melanogaster*, (11) *L. melanonotata*, (12) *L. tridens*, and (13) *L. vittata*. GenBank numbers for the sequences we used to represent these outgroup taxa are provided in Data S1, and some of these sequences were generated in Bagley *et al.* (in revision). We do not present the trees resulting from the BEAST relaxed clock analysis using these outgroups, but the divergence time estimates resulting from this analysis are shown in Table 3.

We built our *Astyanax* spp. dataset by sequencing 54 new individual samples we collected from Costa Rica and Nicaragua and adding them to available cyt*b* sequences from Ornelas-Garcia *et al.* (2008). Ornelas-Garcia *et al.* (2008) generated 247 cyt*b* sequences for North American *Astyanax*, plus several samples of South American *A. fasciatus*, *A. bimaculatus*, and *Roeboides salvadoris* outgroups; however, while we considered the South American characid samples suitable outgroups, we did not consider all of their *Astyanax* samples suitable for our study because they included samples of species/lineages from outside of our study area. So, we only used samples from Ornelas-Garcia *et al.* (2008) that came from our study area. Specifically, we used 99 cyt*b* sequences from CA *Astyanax* samples they collected, and we excluded 148 of their samples of *A. aeneus*, *A. hubbsi*, *A. mexicanus*, *A.* sp. Novo 2, *A.* sp. Novo 4, *Brycon guatemalensis*, and *Bryconamericus scleroparius*.

DNA sequence variation

We calculated corrected Tamura & Nei (1993) genetic distances (d_{MTN}) within and among lineages and clades identified in our phylogenetic analyses in MEGA5 (Tamura *et al.* 2011). We did this despite that various models of DNA sequence evolution were selected as the best-fit models in DT-ModSel for different datasets (Table S2), and we could have opted to estimate model-corrected genetic distances for the corresponding best-fit models for different datasets; however, using different model corrections would have been more complex and would have produced results for which comparisons would be less straightforward. We considered calculating d_{MTN} distances for all groups to be a better method because it yielded distances that are comparable across lineages, especially because we calculated all distances from the same genetic marker (cyt*b*) using the same model, and the Tamura–Nei substitution model was the

best-fit model for several of our datasets (Table S2).

Calibrating our coalescent-dating analyses

As noted in the main text, we used multiple fossil or biogeographic calibration points in each of our BEAST coalescent-dating analyses, and we outline these here. Specifically, during analyses of the five focal poeciliid lineages, outgroup sampling permitted placement of up to three constraints as follows. (1) *Poecilia* (subgenus *Limia*) outgroup samples provided a calibration point constraining the split between *P.* (*L.*) *domicensis* from Cuba and *P.* (*L.*) *vittata* from Hispaniola to 17–14 million years ago (Ma), based on phylogenetic data (Hamilton 2001) and dates for the geological separation of Cuba and Hispaniola, following Alda *et al.* (2013; refs. therein). We set this constraint using a lognormal distribution (mean in real space = 1, sigma = 1.25, offset = 14). (2) Using a similar calibration, we constrained the basal *Pseudoxiphophorus* divergence to be between 11–5 Ma (mean in real space = 2.1, sigma = 1.25, offset = 5), following Agoretta *et al.* (2013). (3) We also used a lognormal calibration to constrain the tree's root age to 39.9 Ma with an extended tail (log standard deviation = 2.5, offset = 39.9), based on the oldest fossil poeciliids available from the Maiz Gordo and Lumbrera formations, Argentina (Pascual *et al.* 1981). We applied all three of these calibration points in our BEAST analyses of each poeciliid focal lineage.

We applied two biogeographic calibrations during independent BEAST analyses of each of the two focal characid lineages, the genera *Astyanax* and *Roeboides*, and characid outgroups. Specifically, similar to Ornelas-García *et al.* (2008), we applied (1) a uniform prior on the basal North-Central American *Astyanax* divergence (excluding South American *A. fasciatus?* and *A. bimaculatus?* and a *Roeboides salvador* outgroups) to 15–8 Ma corresponding to upper and lower age estimates for the isolation of the Maracaibo drainage basin, Venezuela; and (2) a normal distribution constraint on the divergence of the "Bravo-Conchos" *versus* "Panuco-Tuxpan" *Astyanax* lineages corresponding to the volcanic development of the Trans-Mexico Volcanic Belt between 7.5–3 Ma (mean = 5.25, sigma = 1.15) (Ferrari *et al.* 2005; Ornelas-García *et al.* 2008).

Last, for the analysis of our *Amatitlania* cichlid dataset plus 83 cichlid outgroups, we employed three biogeographic calibration points similar to those used by Chakrabarty (2006) and Říčan *et al.* (2013), including (1) a lognormal distribution constraining the diversification of heroine cichlids (within tribe Heroini) including 80 outgroup individuals plus the entire ingroup

sampling to 48.6–39.9 Ma (mean in real space = 4, sigma = 1.25, offset = 39.9); (2) the same Cuba–Hispaniola split as above, except applied to four species of *Nandopsis*; and (3) a normal distribution constraining the separation of the Orinoco and Magdalena drainage basins to 11.8-10.1 Ma, which we applied to two samples of *Caquetaia* (mean = 10.95, sigma = 0.45, offset = 0).

IMa2 methods

We ran IMa2 using isolation-with-migration models with one ancestral population and 2 extant populations, specifying adjacent clades (population pairs) split across the shared phylogeographic breaks we identified as the populations. Although IMa2 supersedes the original formulation of the program (IM) in being able to accommodate more than one ancestral population and more than two extant populations, we ran these more simple two-population models because more complicated multi-population models require very large amounts of data and samples to achieve good MCMC mixing and convergence and appropriate parameter estimates (Pinho & Hey 2010) and our analyses were based on the mtDNA locus.

In terms of methods, all of our IMa2 models used the Hasegawa-Kishino-Yano (HKY) DNA substitution model (Hasegawa *et al.* 1985; Palsbøll *et al.* 2004). This is the most appropriate evolutionary model implemented in IMa2 for DNA sequence data because it allows multiple substitutions as well as different transition and transversion rates. Moreover, other DNA substitution models selected for our datasets by DT-ModSel (supplementary Table S2) are not implemented in IMa2. In initial runs, we ran models for all population pairs in IMa2 in multiple independent runs with and without migration parameters (*m*), by setting the upper bounds of the migration priors to 1 or 0, respectively. We checked the output from regular isolation-with-migration runs (invoking flags "–m1 1 –m2 1"), and where posterior *m* values or their HPD ranges were consistent with effectively zero on-going gene flow (i.e. if peak likelihoods for *m* were at the origin/zero), we deemed the focal lineage/population pair a 'zero-migration' population pair and checked runs with zero migration (Hey 2005; Nielsen & Beaumont 2009). If runs specifying $m = 0$ (invoking flags "–m1 0 –m2 0") showed good convergence of θ and t parameter estimates, then we used the same prior settings from these 'zero-migration' runs in three final IMa2 runs, one of which we report the results from in the manuscript. Alternatively, if regular isolation-with-migration models estimated *m* values that did not center on zero, and if the priors produced runs with good convergence properties and

parameter estimates, then we used the same prior settings from these isolation-with-migration runs in three final IMa2 runs. For all of our final IMa2 runs (regardless of whether they estimated *m* parameters), we logged burn-in periods of 1 million steps (which were discarded) followed by usually 3 million post-burn-in steps, although longer runs with up to 5 million burn-in steps or up to 5 million post-burn-in steps were used if necessary. This procedure yielded reliable estimates of most parameters in most cases based on sufficient convergence and MCMC chain-swapping rates. In addition to procedures mentioned in the text (see Materials and Methods), we also judged convergence by looking for stable trendline plots and checking whether splitting times (*t*) of population pairs were updated at higher rates in higher-numbered chains, suggesting update rates were acceptable. All of these convergence-checking procedures are supported by the authors of the IMa2 algorithm in their publications and the user manual (e.g. Hey & Nielsen 2007; Hey 2011).

Regarding our IMa2 results, we found that it was uncommon for the posterior distributions of *t* estimates and other parameter estimates to peak at relatively low values, drop, and then converge to constant non-zero values, which was a common pattern we discussed in our previous manuscript on CA livebearing fishes (Bagley & Johnson 2014b). This pattern is a common issue with single-locus IM and IMa2 analyses that specify migration parameters to be estimated, because it is often difficult to exclude higher values of *t* when allowing migration in the model (Nielsen & Beaumont 2009). However, the peak likelihood indicates the value of *t* that is most likely and that the data do not support a pure, equilibrium migration model with no population divergence, where *t* would be infinite, and this simpler model (hypothesis) can still be effectively excluded by the data (Nielsen & Beaumont 2009).

As noted in the main text, We converted *t* estimates to absolute time (T_{div}) using the equation $T_{div} = t/\mu$ (where μ = per gene mutation rate) and three different values of μ, including the "fast" 2% vertebrate mtDNA rate and "slow" 0.9% fish mtDNA rate in Bagley & Johnson (2014b), as well as a "moderate" 1.57% rate representing the mean of the 'fish' rate prior used in our BEAST analyses. The reasoning behind this conversion, as outlined by Hey & Nielsen (2004; the paper describing the IM formulation of the isolation-with-migration software), is that *t* is estimated as a mutation-scaled population splitting time (multiplied by the mutation rate). Since the only major change in the algorithms between IM (Hey & Nielsen 2004) and IMa2 (Hey 2010) is that Hey's 2010 paper extended the algorithms to handle multiple ancestral and modern

populations, it is justifiable to calculate T_{div} from t as per above.

For analyses where we estimated migration, we also converted resulting m rate estimates to effective numbers of immigrants per generation using θ_1 and θ_2 estimated from IMa2, and we converted θs to effective population size (N_e) estimates using the equation $\theta = x4N\mu$, where the mutation rate scalar $x = 1$ and μ is the per-gene mutation rate (which should factor in generation time because coalescent models view time in terms of the number of generations). As a slight aside, IMa2 assumes a generation time of 1, and this fits our poeciliid focal lineages reasonably well (Winemiller 1993; see also Bagley *et al.* 2013); however, it is less clear whether this is the case, or how well this assumption works for our other focal lineages. Back to migration rates, several of the focal lineages that we analyzed in IMa2 were deemed to have 'zero-migration' population pairs, so m values were not estimated in their final runs, and hence were not converted or used in any subsequent analyses. Nevertheless, m values output by IMa2 are per gene rates scaled by the mutation rate (which hence is also per gene; Hey & Nielsen 2004), with m_1 (or $m0 > 1$) representing migration from population 1 to population 0 forward in time, and m_2 (or $m1 > 0$) representing migration from population 0 to population 1 forward in time, where population 0 and 1 are the modern populations in the model. These m values can be converted to the effective numbers of migrant gene copies per generation ($2NM$) by using the equation $2NM = 4N\mu \times m/2$ (Hey 2011). For example, the number of migrant gene copies received by population 1 per generation forward in time is $2N_1M = 4N_1\mu \times m_2/2$. Fortunately, IMa2 outputs the correct calculations of these values when appropriate per-gene mutation rates are given in the input files (with "–y1.0" invoked for single-locus datasets) and the "–p3" flag is invoked to print histograms in demographic values; we included these specifications in our IMa2 runs in order to tell the program to obtain and output the desired demographic values.

MTML-msBayes methods

Given that effective population size has the greatest effect on coalescent-based divergence times, it was critical that we use rigorous ranges of upper prior bounds in our msBayes analyses (the lower bounds of msBayes priors are often set to zero or near-zero values). As a result, we followed the authors' instructions (Hickerson 2014) and previously published msBayes analyses (e.g. Barber & Klicka 2010) and chose priors for the upper bounds (θ_{max}) current and ancestral N_e (where $\theta = 4N\mu$; and $N = N_e$, and μ = per site per generation mutation rate) based on empirical estimates of nucleotide diversity (π; Tajima 1983). In particular, for

each set of model classes, we used twice the within-population π (i.e. π_w) estimate for current populations as our standard prior (e.g. San Carlos prior model class M_1; Table 5), and then added a second prior model class with a value of current θ_{max} that doubled this value. To set priors on the ancestral theta multiplier ($\theta_{anc\text{-}max}$), we divided the average of Watterson's (1975) estimator of θ (θ_W) by mean π (i.e. where $\theta_{anc\text{-}max} = \theta_W/\pi$) and used this as our approximation for the upper bound of the ancestral theta multiplier. Similar to the above, we created a different prior model class by doubling this ancestral theta multiplier estimate. Last, we identified upper bounds for the population-pair divergence time parameters (τ) in the model by using empirical estimates and twice the empirical estimates of the mean τs, converted from the mean t estimates output by IMa2, as discussed above. By applying the above procedure to identify and specify all combinations of these uniform prior bounds, we set eight model classes for each analysis of each shared genetic break that we identified in our phylogeographic analysis (Table 5).

Ecological niche modeling (ENM) methods

As noted in the main text, we conducted ENM analyses on each of the focal lineages in MaxEnt (see main text). To build ENMs predicting locations of suitable habitat where species may occur, MaxEnt uses environmental-climatic data from sites where species have been sampled ('presence-only' data) and contrasts them against data extracted from pseudo-absence sites sampled from across the remaining modeled area (Elith *et al.* 2011). We considered MaxEnt ideal for our study because its predictive ability outperforms other methods (e.g. Elith *et al.* 2006) and MaxEnt predictions of species Pleistocene range dynamics correlate well with phylogeographic structuring (Waltari *et al.* 2007).

We predicted present-day geographic distributions of each focal lineage by generating an ENM in MaxEnt while specifying autofeatures, the default regularization multiplier parameter (1.0), and other basic settings (10^4 background points, logistic model with habitat suitability between 0 and 1). However, we increased the number of iterations from the default of 500 up to 5000 to ensure model convergence. We also averaged results over 10 replicate subsampling runs, each starting from a different random seed. During our MaxEnt runs, we used datasets of species geographical occurrence records that covered the known distributions of each lineage, including 72 occurrences of *A. cultratus*, 79 occurrences of *P. amates*, 231 occurrences of *P. mexicana*, 54 occurrences of *P. annectens*, 58 occurrences of *Xenophallus*, 72 occurrences of *Astyanax* spp., 62 occurrences of *R. bouchellei* (within *Roeboides* spp.), and 192 occurrences of

Amatitlania spp. These sampling levels were more than adequate given that sampling greater than or equal to 10 sites permits accurate niche model construction (Stockwell & Peterson 2002) and that we obtained MaxEnt models with good predictive ability in previous analyses of freshwater fishes based on similar sampling (Bagley *et al.* 2013). We generated response curves showing the impact of each variable alone on MaxEnt prediction, and we used 'multivariate environmental similarity surfaces' to assess extrapolation and potential effects of novel environments on predictions.

Our ENM analyses in MaxEnt drew upon the 19 bioclimatic data-layers in the WorldClim dataset (Hijmans *et al.* 2005) to model environments in the study area. For convenience and because space was prohibiting in the main text, we provide a list of these environmental-climatic predictor variables and their descriptions here, in Table A1 below. Based on our ENMtools analysis (see text), we reduced this dataset down to 14 variables with limited cross-correlations, by removing variables BIO10, BIO11, and BIO13–BIO15, which were highly correlated with (and thus redundant with) other predictor-variables. We then used the reduced 14-variable dataset as our source of temperature and precipitation data for our analyses.

Table A1 Bioclimatic environmental data variables used in this study.

Variable #	Code*	Name/Description
1	BIO1	Annual mean temperature
2	BIO2	Mean diurnal range (mean of monthly max. temp - mean of monthly min. temp)
3	BIO3	Isothermality [(BIO2 / BIO7)*100]
4	BIO4	Temperature seasonality (standard deviation*100)
5	BIO5	Maximum temperature of warmest month
6	BIO6	Minimum temperature of coldest month
7	BIO7	Temperature annual range (BIO5 - BIO6)
8	BIO8	Mean temperature of wettest quarter
9	BIO9	Mean temperature of driest quarter
10	BIO10	Mean temperature of warmest quarter
11	BIO11	Mean temperature of coldest quarter
12	BIO12	Annual precipitation
13	BIO13	Precipitation of wettest month
14	BIO14	Precipitation of driest month
15	BIO15	Precipitation seasonality (standard deviation of averages of weekly precipitation)
16	BIO16	Precipitation of wettest quarter
17	BIO17	Precipitation of driest quarter

18	BIO18	Precipitation of warmest quarter
19	BIO19	Precipitation of coldest quarter

The codes in this table are the official WorldClim codes for each variable. Variables shaded grey are those that were eliminated from the dataset during the data reduction step of our analysis, due to cross-correlations with other variables in the original (unclipped) data layers.

Additional ENM results and discussion

Regarding the results of our ENM analyses in MaxEnt, all models had good test AUC values, as mentioned in the text and shown in supplementary Table S3. However, two of the highest mean test AUC values corresponded to the ENM models predicting the distribution of *Xenophallus* during the LGM (mean test AUC = 0.974) and the LIG (mean test AUC = 0.874). While this would seem to indicate that the *Xenophallus* models provided essentially the best prediction and hence were superior to those for the other focal lineages, AUC values tend to be higher for species with narrow ranges such as *Xenophallus*, so that this more likely reflects an artifact of the AUC statistic (Phillips 2006). However, given that our paleodistribution modeling/ENM analyses involved a data reduction step in which we removed the most correlated and perhaps least important predictor-variables, this should mean that we could be more confident in the percent predictive contributions of the predictor-variables output by our MaxEnt analyses, because these values are sensitive to correlations between environmental variables (Phillips 2006).

Additional MTML-msBayes results and discussion

Despite different ranges in their mean population-pair divergence times, levels of variance in divergence times were similar across the San Carlos River break (M_1 $Var[\tau] = 0.099$; model-averaging $Var[\tau] = 0.107$) and the Sixaola River break (M_7 $Var[\tau] = 0.038$; model-averaging $Var[\tau] = 0.293$). Despite evidence for simultaneous diversification at the San Carlos River break in our study, we highlight here that we agree with Hickerson *et al.* (2014) that it is more realistic to interpret such inferences as a "pulse" of divergences related to a single event rather than one literal event. Others have also used this terminology (e.g. Barber & Klicka 2010).

References

Alda FA, Reina RG, Doadrio I, Bermingham E (2013) Phylogeny and biogeography of the *Poecilia sphenops* species complex (Actinopterygii, Poeciliidae) in Central America. *Molecular Phylogenetics*

and Evolution, **66**, 1011-1026.

Agoretta A, Dominguez-Dominguez O, Reina RG, Miranda R, Bermingham E, Doadrio I (2013) Phylogenetic relationships and biogeography of *Pseudoxiphophorus* (Teleostei: Poeciliidae) based on mitochondrial and nuclear genes. *Molecular Phylogenetics and Evolution*, **66**, 80-90.

Bagley JC, Alda F, Breitman MF, van den Berghe E, Bermingham E, Johnson JB (in revision) Assessing species boundaries using multilocus species delimitation in a morphologically conserved group of Neotropical freshwater fishes, the *Poecilia sphenops* species complex (Poeciliidae). *PLoS One*.

Bagley JC, Johnson (2014a) Phylogeography and biogeogrpahy of the lower Central American Neotropics: diversification between two continents and between two seas. *Biological Reviews*, **89**, 767-790.

Bagley JC, Johnson JB (2014b) Testing for shared biogeographic history in the lower Central American freshwater fish assemblage using comparative phylogeography: concerted, independent, or multiple evolutionary responses? *Ecology and Evolution*, **4**, 1686-1705.

Bagley JC, Sandel M, Travis J, Lozano-Vilano M de L, Johnson JB (2013) Paleoclimatic modeling and phylogeography of least killifish, *Heterandria formosa*: insights into Pleistocene expansion-contraction dynamics and evolutionary history of North American Coastal Plain freshwater biota. *BMC Evolutionary Biology*, **13**, 223.

Barber BR, Klicka J (2010) Two pulses of diversification across the Isthmus of Tehuantepec in a montane Mexican bird fauna. *Proceedings of the Royal Society of London B*. doi:10.1098/rspb.2010.0343.

Bussing WA (1998) *Freshwater Fishes of Costa Rica*, 2nd Edition. Editorial de la Universidad de Costa Rica, San José, Costa Rica.

Chakrabarty P (2006) Systematics and historical biogeography of Greater Antillean Cichlidae. *Molecular Phylogenetics and Evolution*, **39**, 619-627.

Ferrari L, Tagami T, Eguchi M, Orozco-Esquivel MT, Petrone CM, Jacobo-Albarrán J, López-Martínez M (2005) Geology, geochronology and tectonic setting of late Cenozoic volcanism along the southwestern Gulf of Mexico: the Eastern Alkaline Province revisited. *Journal of Volcanology and Geothermal Research*, **146**, 284-306.

Hamilton A (2001) Phylogeny of *Limia* (Teleostei: Poeciliidae) based on NADH dehydrogenase subunit 2 sequences. *Molecular Phylogenetics and Evolution*, **19**, 277-289.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160-174.

Hickerson MJ (2014) Simplified manual for multi-taxa multi-locus msBayes (MTML-msBayes). Biology Department, City College of New York–The Graduate Center, City University of New York, New York, NY. <http://hickerlab.wordpress.com/software-code-scripts/> Accessed 12 September 12 2014.

Hickerson MJ, Stone GN, Lohse K, Demos TC, Xie X, Landerer C, Takebayashi N (2014) Recommendations for using msBayes to incorporate uncertainty in selecting an ABC model prior: a response to Oaks et al. *Evolution*, **68**, 284-294.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**(15), 1965-1978.

Hubert N, Renno J-F (2006) Historical biogeography of South American freshwater fishes. *Journal of Biogeography*, **33**, 1414-1436.

Marshall JS (2007) The geomorphology and physiographic provinces of Central America. In: *Central America: Geology, Resources and Hazards* (eds Bundschuh J, Alvarado GE), pp. 1-51. Taylor & Francis, Philadelphia, PA.

Miller RR, Minckley WL, Norris SM (2005) *Freshwater Fishes of México*. The University of Chicago Press, Chicago.

Nielsen R, Beaumont MA (2009) Statistical inferences in phylogeography. *Molecular Ecology*, **18**, 1034-1047.

Ornelas-García CP, Dominguez-Dominguez O, Doadrio I (2008) Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evolutionary Biology*, **8**, 340.

Palsbøll PJ, Berube M, Aguilar A, Notarbartolo di Sciara G, Nielsen R (2004) Discerning between recurrent gene flow and recent divergence under a finite-site mutation model applied to North Atlantic and Mediterranean sea fin whale (*Balaenoptera physalus*) populations. *Evolution*, **58**, 670-675.

Pascual R, Bond M, Vucetich MG (1981) El Subgrupo Santa Bárbara (Grupo Salta) y sus vertebra- dos. Cronología, paleoambientes y paleobiogeografía. San Luis, Argentina. VIII Congreso Geologico Argentino, Actas III, 746-758.

Phillips SJ (2006) A Brief Tutorial on Maxent. AT&T Research. Available at: http://www.cs.princeton.edu/~schapire/maxent/tutorial/tutorial.doc. Last accessed 28 September 2014.

Pinho C, Hey J (2010) Divergence with gene flow: models and data. *Annual Review of Ecology, Evolution and Systematics*, 41, 215-230.

Říčan O, Piálek L, Zardoya R, Doadrio I, Zrzavý J (2013) Biogeography of the Mesoamerican Cichlidae (Teleostei: Heroini): colonization through the GAARlandia land bridge and early diversification. *Journal of Biogeography*, **40**, 579-593.

Rogers RD, Karason H, van der Hilst R (2002) Epirogenic uplift above a detached slab in northern Central America. *Geology*, 30, 1031-1034.

Rogers RD, Mann P, Emmet PA (2007) Tectonic terranes of the Chortis block based on integration of regional aeromagnetic and geologic data. In: *Geologic and Tectonic Development of the Caribbean Plate in Northern Central America* (ed. Mann P), pp. 65-88. Geological Society of America Special Paper 428.

Stockwell DRB, Peterson AT (2002) Effects of sample size on accuracy of species distribution models. *Ecolical Modeling*, **148**(1), 1-13.

Tajima F (1983) Evolutionary relationships of DNA sequences in finite populations. *Genetics*, **105**, 437-460.

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

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731-2739.

Waltari E, Hijmans RJ, Peterson AT, Nyari AS, Perkins SL, Guralnick RP (2007) Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS One*, 2(7):e563.

Watterson GA (1975) Number of segregating sites in genetic models without recombination. *Theoretical Population Biology*, **7**, 256-276.

Winemiller KO (1993) Seasonality of reproduction by livebearing fishes in tropical rainforest streams. *Oecologia* **95**, 266-276.

Table S1 PCR primers and annealing temperatures used to amplify the mitochondrial cytochrome *b* (cyt*b*) gene for each focal lineage in this study.

Focal lineage	Family	Primer	Sequence	T_A (°C)	Reference
All	All	Glu31	5'–TGRCTTGAAAAACCACCGTTGT–3'	48–49	Unmack *et al.* (2009)
		HD / INH	5'–GGGTTGTTTGATCCTGTTTCGT–3'		Schmidt *et al.* (1998)
All poeciliids (focal lineages 1–5 in Table 1)	Poeciliidae	L14725	5'–GAYTTGAARAACCAYCGTTG–3'	48	Hrbek *et al.* (2007)
		H15982	5'–CCTAGCTTTGGGAGYTAGG–3'		Hrbek *et al.* (2007)
All characids (focal lineages 6 and 7 in Table 1)	Characidae	Glu-F	5'–GAAGAACCACCGTTGTTATTCAA–3'	48–49	Zardoya & Doadrio (1998)
		Thr-R	5'–ACCTCCRATCTYCGGATTACA–3'		Zardoya & Doadrio (1998)
Amatitlania spp.	Cichlidae	Glu31	5'–TGRCTTGAAAAACCACCGTTGT–3'	48–49	Unmack *et al.* (2009)
		RF.Thr.48	5'–GCAGTAGGAGGGAATTTAACCTTCG–3'		Zardoya & Doadrio (1998)

The Glu31-HD primer pair amplifies the first 601 bp of cyt*b*, and typically does so well in all higher fishes; however, it was mainly used to obtain high-resolution sequences of the front end of the gene, rather than as our main forward primer. The other primer pairs consistently yielded good sequences of the complete cyt*b* gene, and sometimes the first 10–50 bp of the *Glu*-tRNA that follows the cyt*b* gene (thus we obtained 1151 bp sequences for some of our characid samples). Abbreviations: T_A, annealing temperature in units of degrees Celsius.

References

Hrbek T, Seckinger J, Meyer A (2007) A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. *Molecular Phylogenetics and Evolution*, **43**, 986-998.

Schmidt TR, Bielawski JP, Gold JR (1998) Molecular phylogenetics and evolution of the cytochrome b gene in the cyprinid genus *Lythrurus* (Actinopterygii: Cypriniformes). *Copeia*, **1998**, 14-22.

Unmack PJ, Bennin AP, Habit EM, Victoriano PF, Johnson JB (2009) Impact of ocean barriers, topography, and glaciation on the phylogeography of the catfish *Trichomycterus areolatus* (Teleostei: Trichomycteridae) in Chile. *Biological Journal of the Linnean Society*, **97**, 876-892.

Zardoya R, Doadrio I (1998) Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. *Proceedings of the Royal Society of London B*, **265**(1403), 1365-1372.

Table S2 DNA substitution models selected using DT-MODSEL.

Species/lineage	Gene/alignment	*n*	bp	Best model
Amatitlania spp.	Cyt*b* ingroup	322	1137	HKY+*I*
	460 cichlids cyt*b* (ingroup + outgroups)	460	1137	TrN+*Γ*+*I*
	460 cichlids cyt*b* 1st + 2nd codon position	460	758	TrN+*Γ*+*I*
	460 cichlids cyt*b* 3rd codon position	460	379	GTR+*Γ*+*I*
Alfaro cultratus	Cyt*b* ingroup	270	1140	TrN+*Γ*+*I*
	359 cyt*b* (ingroup + outgroups)	359	1140	TrN+*Γ*+*I*
	359 cyt*b* 1st + 2nd codon position	359	760	TrN+*Γ*+*I*
	359 cyt*b* 3rd codon position	359	380	GTR+*Γ*+*I*
Astyanax spp.	Cyt*b* ingroup	243	1140	TVM+*Γ*+*I*
	256 characins cyt*b* (ingroup + outgroups)	256	1140	K81uf+*Γ*+*I*
	256 characins cyt*b* 1st + 2nd codon position	256	760	HKY+*Γ*
	256 characins cyt*b* 3rd codon position	256	380	TIM+*Γ*
Priapichthys annectens	Cyt*b* ingroup	100	1140	TrN+*Γ*+*I*
	159 cyt*b* (ingroup + outgroups)	159	1140	TrN+*Γ*+*I*
	159 cyt*b* 1st + 2nd codon position	159	760	TrN+*Γ*+*I*
	159 cyt*b* 3rd codon position	159	380	TIM+*Γ*+*I*
Phallichthys amates	Cyt*b* ingroup	93	1140	TrN+*Γ*
	152 cyt*b* (ingroup + outgroups)	152	1140	TrN+*Γ*+*I*
	152 cyt*b* 1st + 2nd codon position	152	760	TrN+*Γ*+*I*
	152 cyt*b* 3rd codon position	152	380	TrN+*Γ*+*I*
Poecilia mexicana	Cyt*b* ingroup	761	1086	TrN+*Γ*+*I*
	853 cyt*b* (ingroup + outgroups)	853	1086	TrN+*Γ*+*I*
	853 cyt*b* 1st + 2nd codon position	853	724	TrN+*Γ*+*I*
	853 cyt*b* 3rd codon position	853	362	GTR+*Γ*
Roeboides spp.	Cyt*b* ingroup	108	1151	HKY+*I*
	155 characins cyt*b* (ingroup + outgroups)	155	1151	K81uf+*Γ*+*I*
	155 characins cyt*b* 1st + 2nd codon position	155	768	HKY+*Γ*+*I*
	155 characins cyt*b* 3rd codon position	155	383	TIM+*Γ*+*I*
Xenophallus umbratilis	Cyt*b* ingroup	180	1140	TrN+*I*
	237 cyt*b* (ingroup + outgroups)	237	1140	TrN+*Γ*+*I*
	237 cyt*b* 1st + 2nd codon position	237	760	TrN+*Γ*+*I*
	237 cyt*b* 3rd codon position	237	380	TIM+*Γ*+*I*

Model selection analyses using the decision theory algorithm in DT-MODSEL (Minin *et al.* 2003) supported different best-fit models of DNA evolution for different datasets, including datasets filtered by codon partitions. We preferred DT-MODSEL for our substitution model selection analyses, rather than other model selection software, because DT-MODSEL has been shown to recover models that yield superior ML branch lengths relative to other comparable programs (Minin *et al.* 2003). This table lists model selection results for each dataset analyzed in this study. Symbols and abbreviations: cyt*b*, mitochondrial cytochrome *b* gene; *I*, parameter representing proportion of invariable sites; *Γ*, gamma-distributed rate variation; bp, number of nucleotide base pairs; ML, maximum likelihood phylogenetic analyses estimating gene trees; *n*, sample size (numbers correspond to sequence alignment sizes, except for multilocus datasets the numbers in parentheses are sample sizes for each locus).

References

Minin V, Abdo Z, Joyce P, Sullivan J (2003) Performance-based selection of likelihood models for phylogeny estimation. *Systematic Biology*, **52**, 674-683.

Table S3 Mean MaxEnt model AUC scores and their standard deviations (s.d.).

	LGM		LIG	
Focal lineage	**mean**	**s.d.**	**mean**	**s.d.**
1. *Alfaro cultratus*	0.876	0.048	0.851	0.042
2. *Phallichthys amates*	0.847	0.026	0.840	0.053
3. *Poecilia mexicana*	0.836	0.025	0.829	0.033
4. *Priapichthys annectens*	0.794	0.027	0.802	0.041
5. *Xenophallus umbratilis*	0.974	0.012	0.874	0.022
6. *Astyanax* spp.	0.840	0.045	0.841	0.040
A. nasutus	–	–	–	–
A. nicaraguensis	–	–	–	–
A. orthodus/sp.	–	–	–	–
7. *Roeboides* spp.	–	–	–	–
R. bouchellei	0.826	0.058	0.783	0.075
R. bussingi	–	–	–	–
8. *Amatitlania* spp.	0.862	0.020	0.855	0.017

AUC stands for area under the receiving operator characteristic curve. This table gives test AUC values from MAXENT as means and their standard deviations for the two sets of paleodistribution modeling analyses we conducted for each focal lineage: ENM prediction of the present-day distribution of the focal lineage, followed by reprojection of the ENM onto paleoclimatic environmental reconstructions for the Pleistocene Last Glacial Maximum (LGM); and present-day ENM prediction, followed by reprojection of the ENM onto paleoclimatic environmental data for the Pleistocene Last Interglaciation (LIG). See additional details of the MAXENT analyses in the main text and Appendix S1.

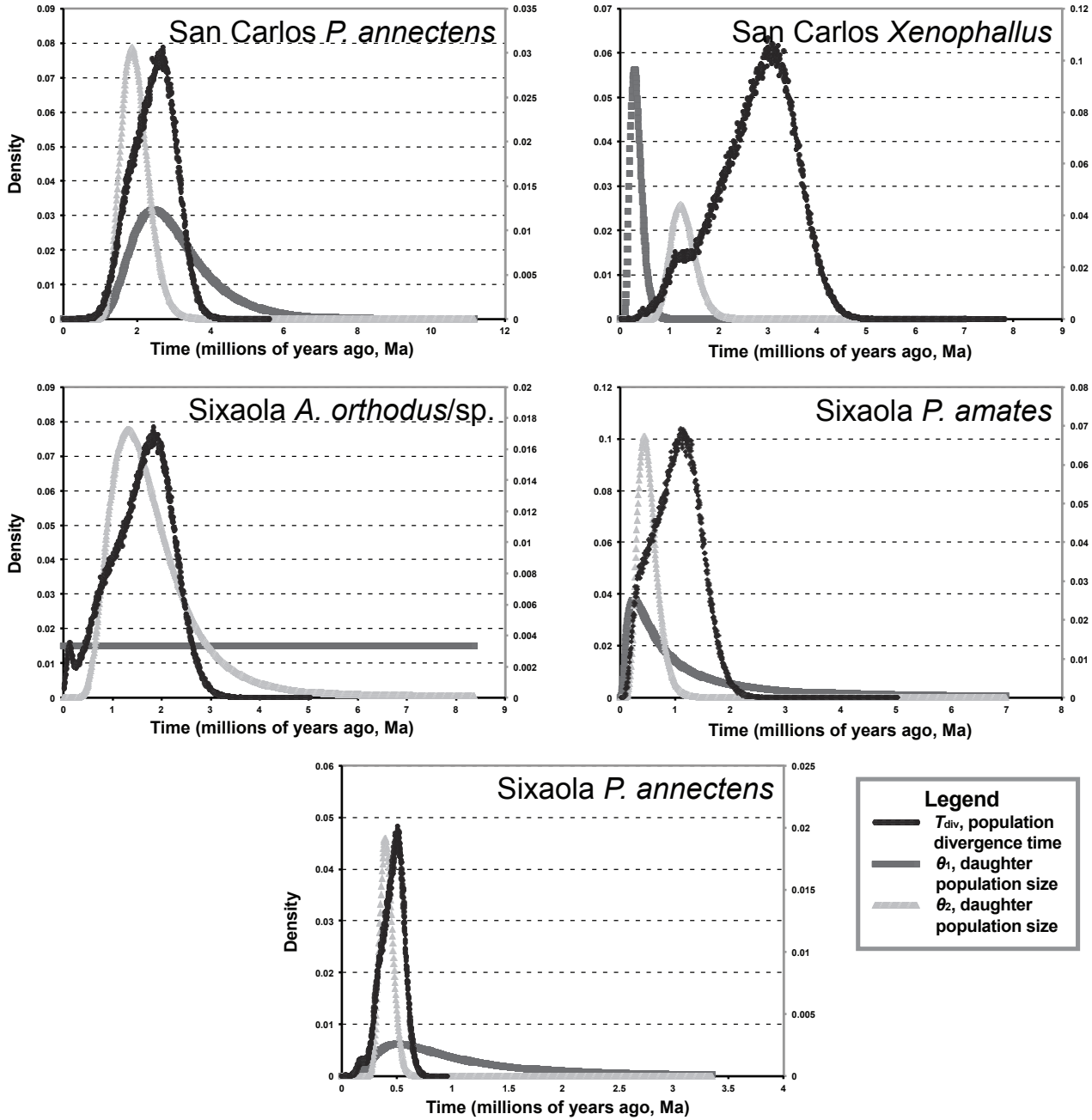
San Carlos P. annectens
San Carlos Xenophallus
Sixaola A. orthodus/sp.
Sixaola P. amates
Sixaola P. annectens
Density
Time (millions of years ago, Ma)
Legend
Tdiv, population divergence time
θ1, daughter population size
θ2, daughter population size

Figure S1

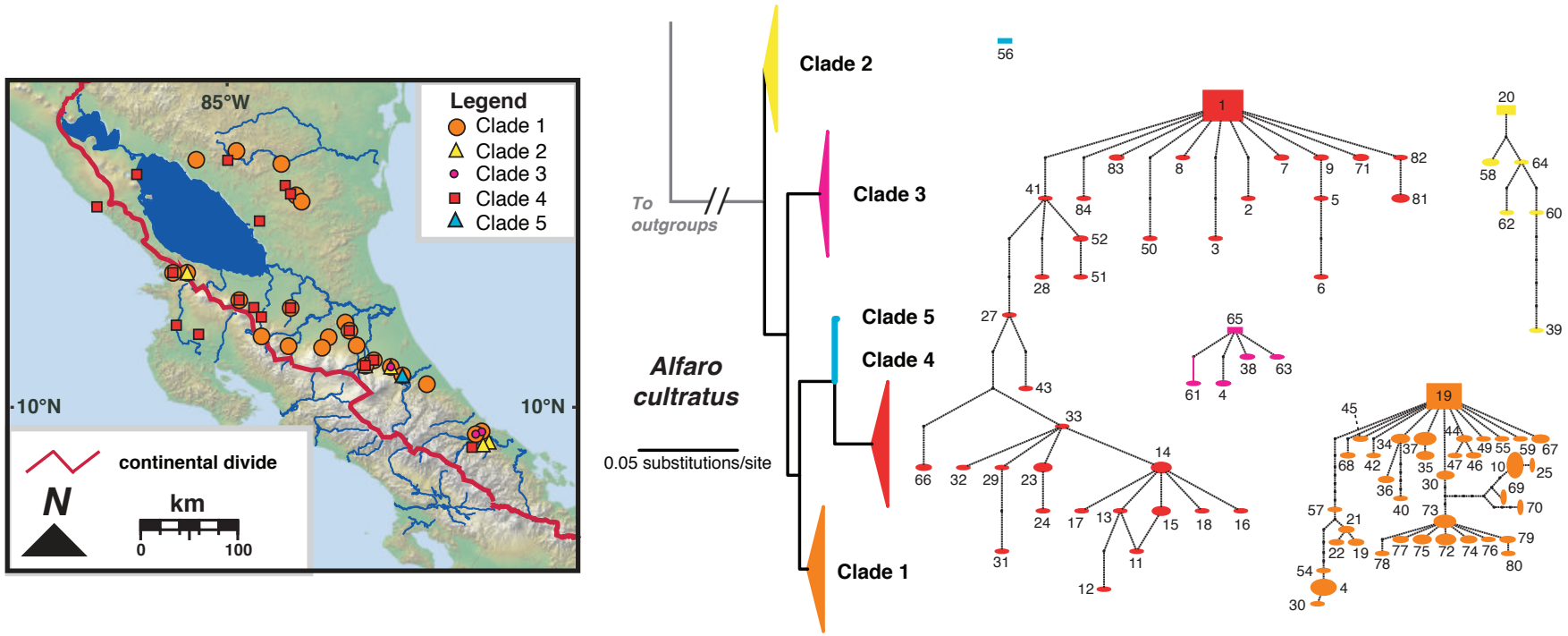

Figure S2 MtDNA gene trees estimated in GARLI and parsimony networks estimated using TCS for each focal lineage, showing well-supported phylogeographic clades. For comparison, geographical distributions of the main clades for each focal lineage are mapped over modern topography, similar to Fig. 3.

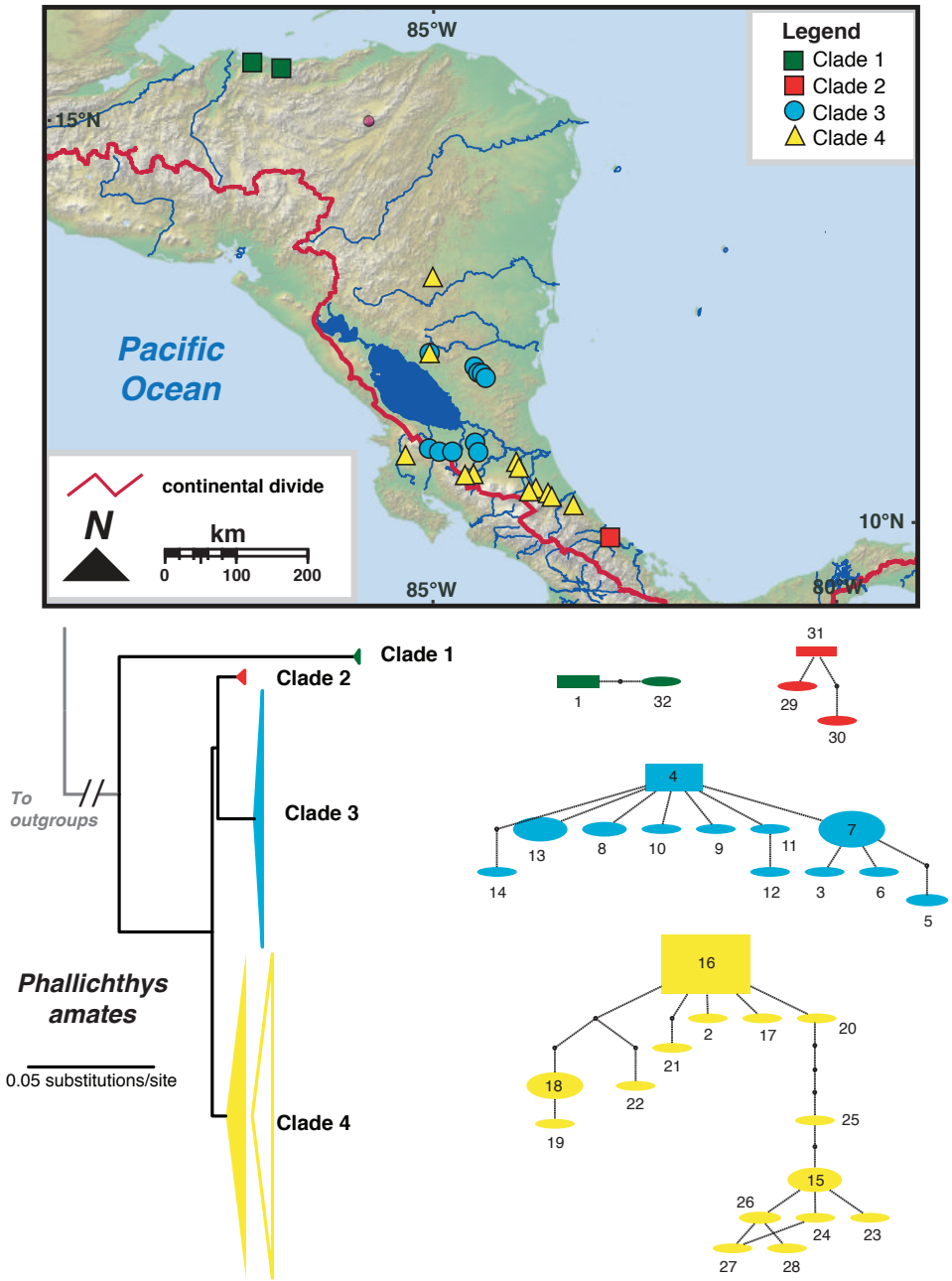

Figure S2 *Continued.*

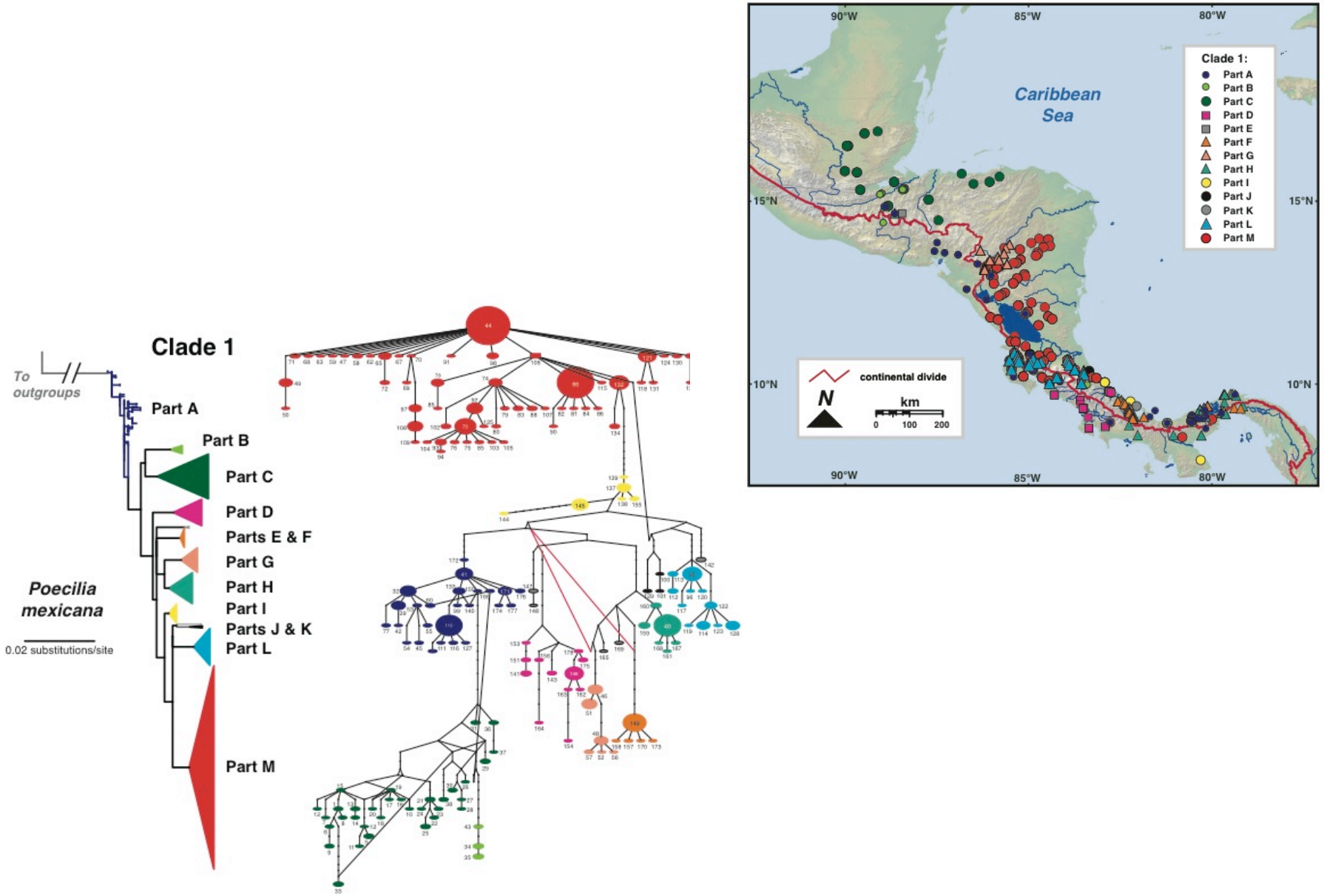

Figure S2 *Continued.*

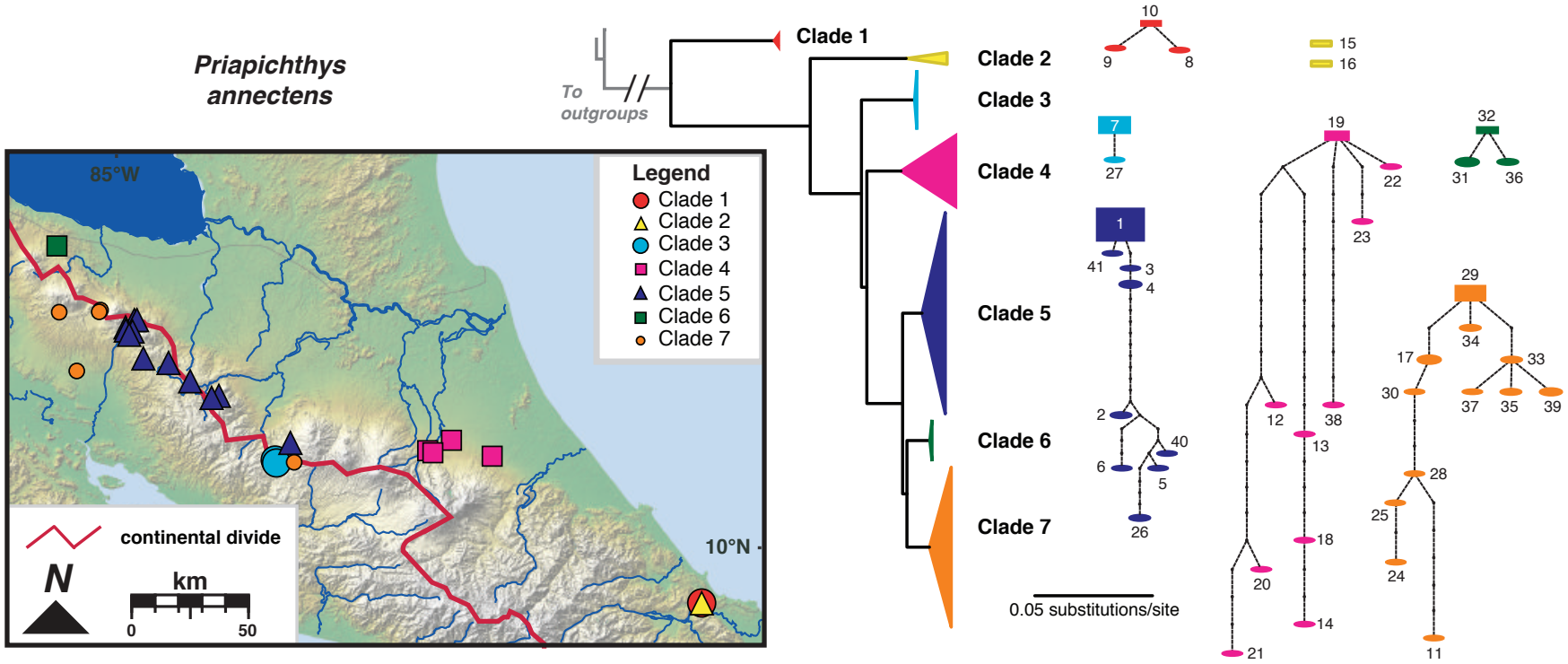

Figure S2 *Continued.*

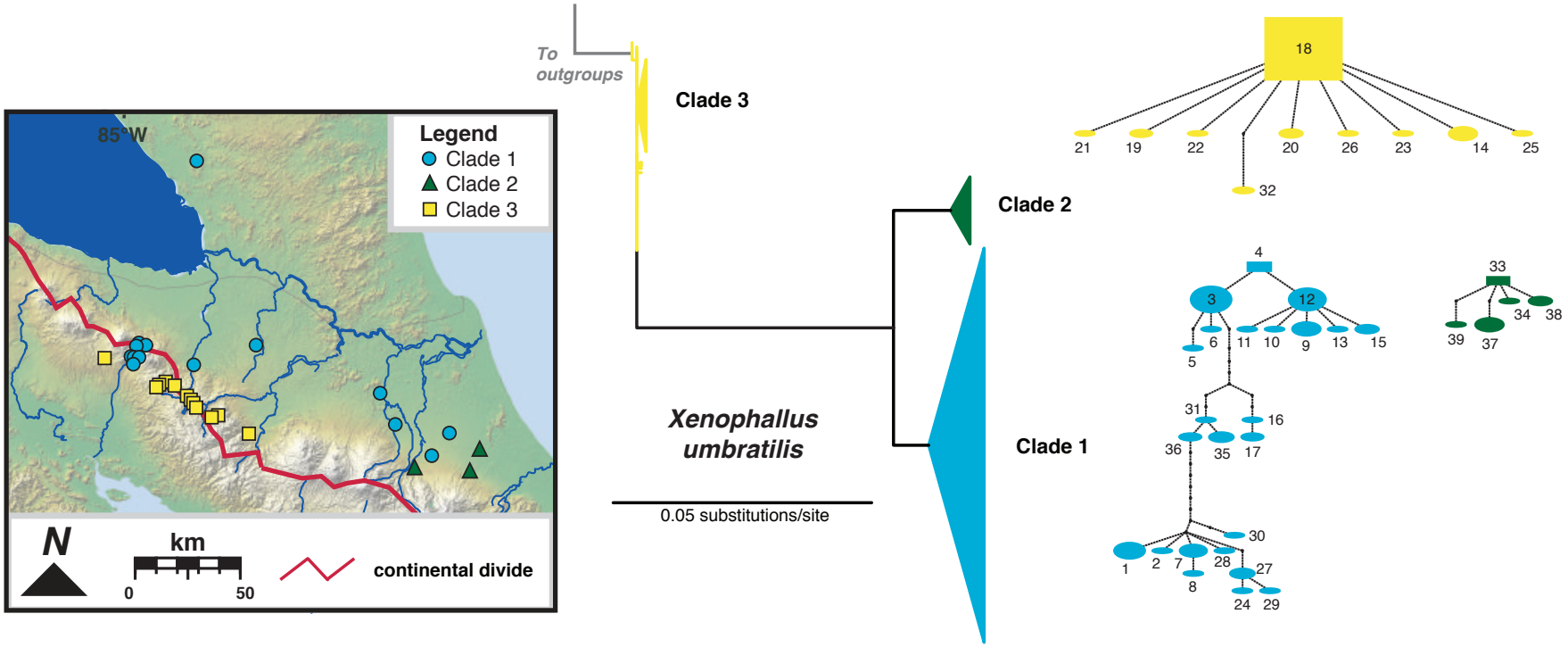

Figure S2 *Continued.*

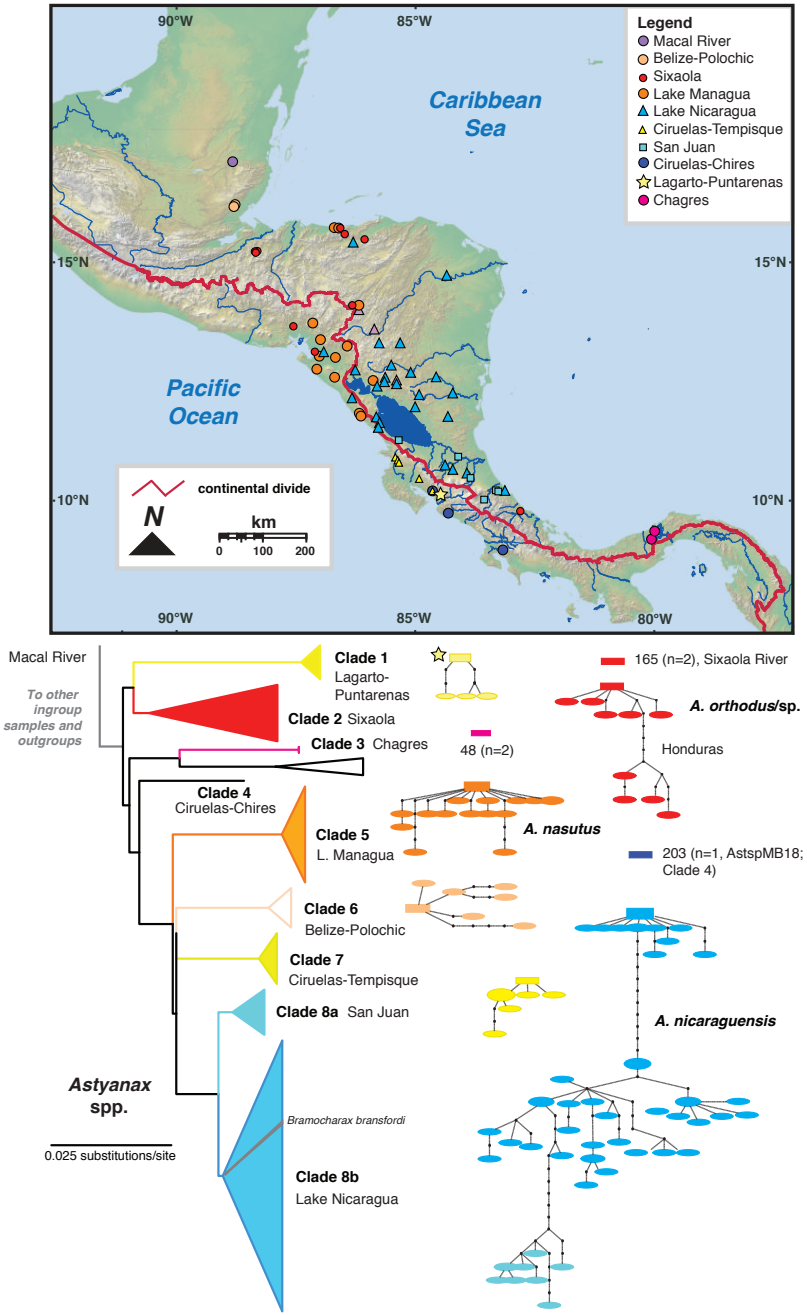

Figure S2 *Continued.*

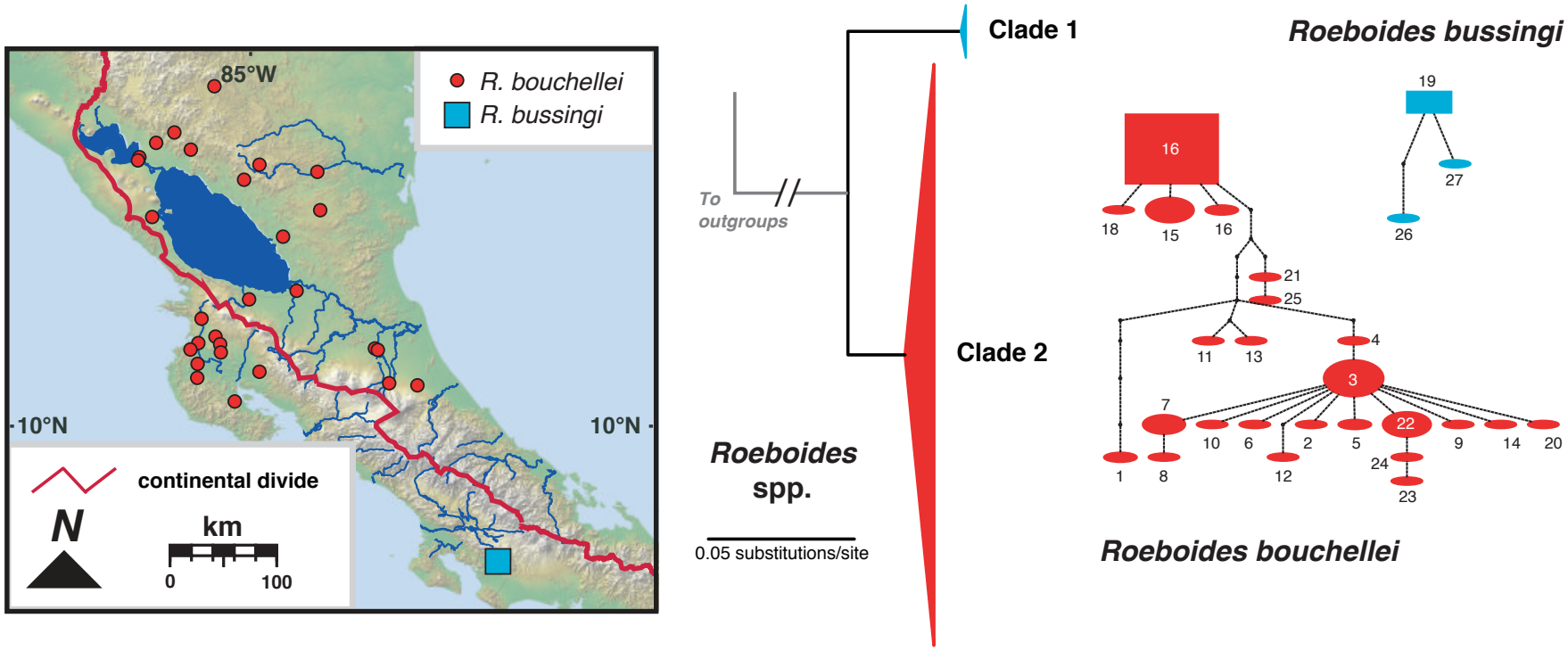

Figure S2 *Continued.*

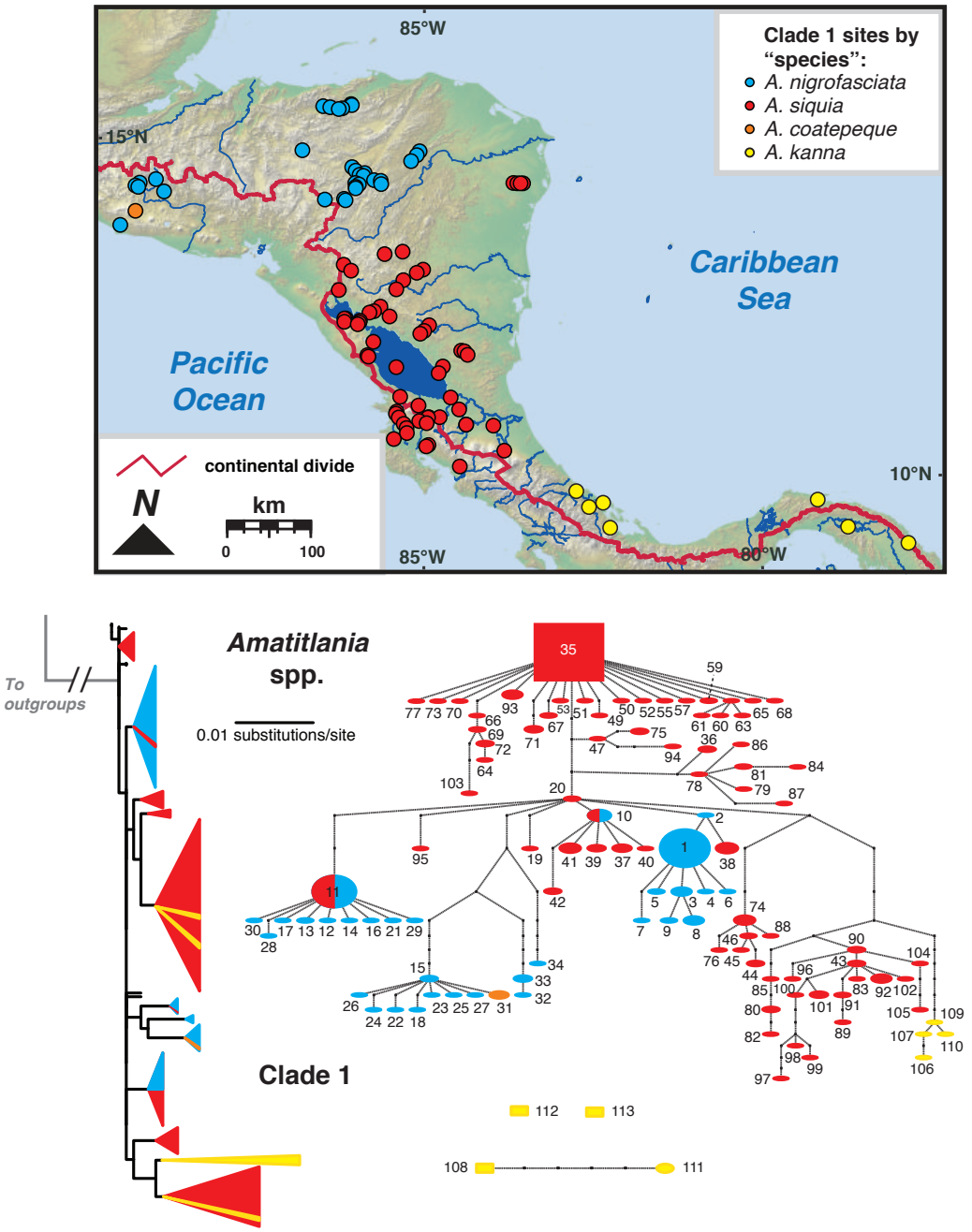

Figure S2 *Continued.*

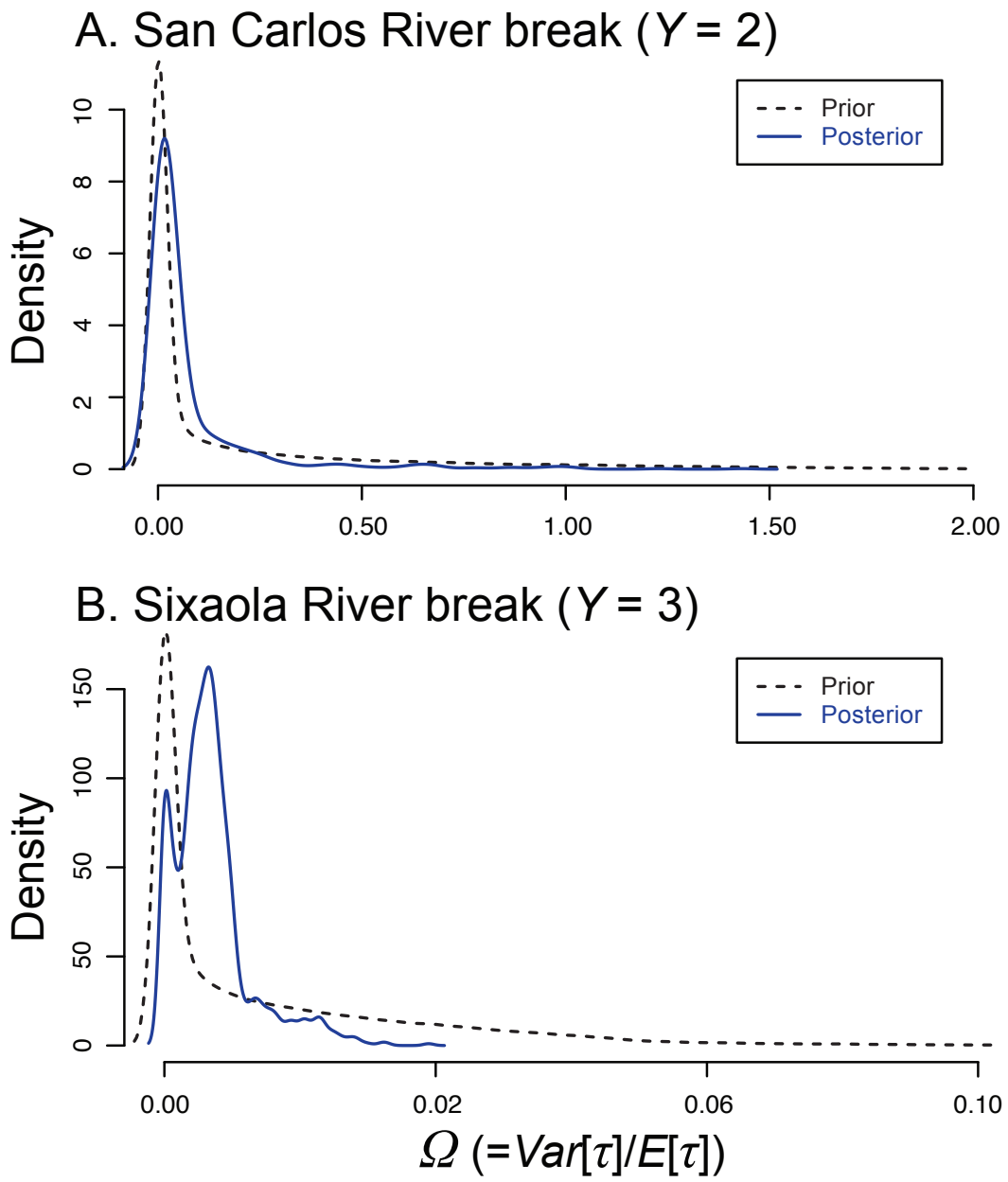
A. San Carlos River break (Y = 2)
Prior
Posterior
Density
0
2
4
6
8
10
0.00
0.50
1.00
1.50
2.00
B. Sixaola River break (Y = 3)
Prior
Posterior
Density
0
50
50
100
150
0.00
0.02
0.06
0.10
Ω (=Var[τ]/E[τ])

Figure S3