2009-03-02

Biogeography of the Livebearing Fish Poecilia gillii in Costa Rica: Are Phylogeographic Breaks Congruent with Fish Community Boundaries?

Jared Benjamin Lee
Brigham Young University - Provo

BIOGEOGRAPHY OF THE LIVEBEARING FISH *POECILIA GILLII* IN COSTA RICA: ARE PHYLOGEOGRAPHIC BREAKS CONGRUENT WITH FISH COMMUNITY BOUNDARIES?

by

Jared B. Lee

A selected project submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Master of Science

Department of Biology
Brigham Young University
February 2009

BRIGHAM YOUNG UNIVERSITY GRADUATE COMMITTEE APPROVAL

of a selected project submitted by

Jared B. Lee

This selected project has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

2 March 2009
Date

Jerald B. Johnson, Chair

March 2, 2009
Date

Byron J. Adams

2 March '09
Date

Jack W. Sites, Jr.

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the selected project of Jared B. Lee in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

2 March 2009
Date

Jerald B. Johnson
Chair, Graduate Committee

Accepted for the Department

2 MAR 2009
Date

Keith A. Crandall
Department Chair

Accepted for the College

March 2, 2009
Date

Rodney J. Brown
Dean, College of Life Sciences

ABSTRACT

One of the original goals of phylogeography was to use genetic data to identify historical events that might contribute to breaks among biotic communities. In this study, we examine the phylogeography of a common livebearing fish (*Poecilia gillii*) from Costa Rica. Our goal was to see if phylogeographic breaks in this species were congruent with previously-defined boundaries among four fish community provinces. We hypothesized that if abiotic factors influence both community boundaries and genetic structuring in *P. gillii* then we would find four monophyletic clades within our focal species that were geographically separated along community boundary lines. Similarly, we expected to find most of the genetic variation *in P. gillii* partitioned among these four geographic regions. We generated DNA sequence data (mitochondrial cyt *b* and nuclear S7 small ribosomal subunit) for 260 individuals from 42 populations distributed across Costa Rica. We analyzed these data using phylogenetic (parsimony and likelihood) and coalescent approaches to estimate phylogenetic relationships among haplotypes, patterns of gene flow, and effective population size. Contrary to our expectations, we did not find four monophyletic groups that mapped cleanly to our geographic community provinces. However, one of our clades was restricted to a single province, suggesting that common earth history events could be responsible for both genetic structuring in *P. gillii* and fish community composition in this area. However, our results show a complex pattern of gene flow throughout other regions in Costa Rica where genetic structuring is not governed by community province boundaries.

ACKNOWLEDGEMENTS

We thank Byron Adams, Mark Belk, Keith Crandall, Jack Sites, and Peter Unmack for their help and input throughout various stages of this study. Justin Bagley, Carissa Jones, and Laura Scott provided help in the lab and reviewed earlier versions of this manuscript. We offer a special thanks to the following individuals for help with the field collections: John Aedo, Liesel Christensen, April Larson, Kevin and Brook Menzie, Dustin Patrick, Zac Lewis, Nina Laitinen, and Laura Pennock. We also thank Javier Guevara Sequeira from the Ministerio del Ambiente y Energía Sistema Nacional de Areas de Conservación in Costa Rica for his assistance in arranging for collecting permits and Ray K. Koplin for assistance generating Figure 1. Funding for this work was provided by Department of Biology and Monte L. Bean Life Science Museum at Brigham Young University, and by a major research instrumentation grant from the U.S. National Science Foundation (DBI-0520978).

Table of Contents

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Understanding what drives the formation of biotic communities remains a fundamental problem in ecology. Considerable attention has focused on whether communities are limited by ecological interactions that result in a predictable set of assembly rules versus the alternative that community formation is entirely a random process (Diamond 1975; Connor & Simberloff 1979). Interestingly, deviations from random patterns are almost always ascribed to biotic interactions, such as competition or predation (Toft 1985; Gillespie 2004). Much less attention has focused on the possibility that abiotic effects, including earth history events, might also contribute to contemporary community composition (Dunson & Travis 1991; Avise et al. 1987). However, it is well known that vicariant events can limit the distribution of some species and could therefore affect community assemblages (Ronquist 1997; Avise 2000).

Traditionally, biotic communities have been described by simply listing species present within a particular geographic region (Diamond 1975; Connell 1972) with much of this pioneering work focusing on islands. The practice of delineating community boundaries across contiguous land masses is less clear-cut, and usually relies on identifying geographic breaks in faunal composition and areas of endemism (Abell et al. 2008; Smith & Bermingham 2005). Classic examples of community assemblages determined this way can be found in inter-tidal organisms (Connell 1972), birds (Diamond 1975), and terrestrial plants (reviewed by Tilman 1988). More recently, biogeographers have focused on using species composition to describe freshwater fish communities both at global scales (Abell et al. 2008) and across smaller geographic ranges including Australia (Unmack 2001), Southeast Asia (Yap 2002) lower Central America (Smith & Bermingham 2005), South America (Hubert & Renno 2006), and

Europe (Reyjol et al. 2007). These biogeographic assessments can be viewed as communities in that they describe faunal assemblages in a given region. Although considerable effort has gone into describing patterns of community assembly, our understanding of factors that control these assemblages remains unknown in most systems.

Community ecologists could potentially benefit by considering the role of earth history events in defining the geographic limits of community boundaries. What we want to know is whether or not abiotic factors (such as orogenic events, volcanic activity, or glacial cycles) shape the composition of species within delineated communities. One way to approach this problem is to utilize the tools of phylogeography to identify natural geographic barriers to organismal movement. Several phylogeographic approaches can be used to reveal how populations have moved across a landscape through space and time (Avise 2000). While early work in phylogeography (Avise 1992, 1994) discusses how genetic boundaries in widespread species may coincide with described biogeographic boundaries, little work has followed up on actually testing for such breaks where boundaries have been described (Burton 1998). Where community provinces have been identified, we can test for congruence between phylogeographic breaks in gene flow within widespread species and geographic transitions from one species assemblage to another. In this way, widespread species that span the distribution of several community assemblages can be used to identify potential abiotic factors that might contribute to shifts from one community assemblage to another.

Here we identify a model system to explore the potential role of earth history events on freshwater fish assemblages in lower Central America. Smith and Bermingham (2005) identified seven fish communities throughout lower Central America (Costa Rica and Panama) using presence/absence data of all freshwater fishes from this

region. We recognize that communities can be defined at different scales; in this study we define communities in the broadest sense looking at regions of faunal assemblages. Several of the species within these assemblages have ranges that end near the geographic boundaries that delineate fish community provinces in Central America. However, some species are widespread and found in multiple community provinces. Here, we focus on one of these widespread species, a livebearing fish *Poecilia gillii*, to explore the potential role of vicariant events in determining fish community boundaries. We focus on this species in Costa Rica, an area that includes four of Smith and Bermingham's (2005) fish community provinces.

In this study, we focus on two objectives. First, we assess phylogeographic breaks within *P. gillii* to see if clades within this species correspond to the geographic distribution of Smith and Bermingham's (2005) fish community provinces in Costa Rica. A pattern of reciprocal monophyly in *P. gillii* from each of the four community provinces would suggest that vicariance could contribute to current community structure. Second, we reconstruct historic patterns of gene flow in *P. gillii* across Costa Rica. If vicariance is responsible for the population structure then we would expect to find low gene flow between each of the provinces. We find that contrary to our hypothesis, *P. gillii* in Costa Rica does not form four distinct clades coincident with the described community breaks. However, we do find that two phylogeographic breaks in *P. gillii* in the southwest of Costa Rica (the Herradura Divide and Southern Cordillera) are clearly congruent with the geographic transitions from Smith and Bermingham's (2005) assemblages. Hence, our data suggest that vicariant events could be responsible for some of the fish faunal transitions in Central America, but other factors must be invoked to explain additional fish community transition zones.

MATERIALS & METHODS

Study System

Smith &Bermingham (2005) described four freshwater fish communities that are geographically distributed in Costa Rica: San Juan, Bocas, Chiriqui, and Chiapas-Nicaragua (Figure 1). These fish communities are separated by four geographic boundaries which we label as follows: (1) Northern Cordillera; (2) Southern Cordillera, (3) Herradura Divide, and (4) Estrella Divide. The Northern and Southern Cordillera run from northwest to southeast along the central mountain range in Costa Rica; these boundaries mark the continental divide between the Caribbean and Pacific Oceans. The Northern Cordillera is composed of active and extinct volcanoes associated with orogeny of this range dating to the end of the Pleistocene (Coates & Obando 1996). Low valleys separate several of these volcanoes, creating passes where fish could potentially move between community provinces during river capture events. At the northern limit of this divide is a large plain which descends to Lake Nicaragua. Tectonic activity here could also result in river capture allowing fish to move between provinces. In contrast to the Northern Cordillera, the mountains along the Southern Cordillera are much higher in elevation and there are no low valleys, presenting a more formidable barrier. The Herradura Divide occurs in western Costa Rica and separates the Chiapas-Nicaragua Province from the Chiriqui Province. There is a prominent land formation here that creates a modest peninsula extending into the Gulf of Nicoya; rivers that drain this area flow northward and southward, such that the headwaters of these rivers are much closer than their mouths. Finally, the Estrella Divide is found in eastern Costa Rica and separates the San Juan from the Bocas provinces. Interestingly, this boundary has no distinguishing geographic features that separate the two fish provinces. It is also important to note that western Costa Rica is much drier than eastern Costa Rica, due to

the rain shadow effect of westward wind patterns in this part of the world. Hence, we might expect to see higher levels of isolation and fragmentation among *P. gillii* populations in the west than in the east.

Focal Species

In this study we use the livebearing fish species *Poecilia gillii* to examine phylogeographic breaks across Costa Rica. Beyond some early research on meta-population dynamics in this species (Chapman 1992; Chapman et al. 1991), little formal research has been conducted on *P. gillii*. This species is broadly distributed throughout Central America, extending from northern Colombia to Guatemala. It occupies a wide range of habitats, including coastal lagoons, small streams, and pools adjacent to larger rivers. It occurs at elevations from sea level up to 1200m (Bussing 1998 and personal observation). *Poecilia gillii* has a wide thermal tolerance for a tropical species found at temperatures from 19° to 37° C, feeding upon detritus and filamentous algae (Bussing 1998). This broad set of physiological and ecological tolerances contributes to the common occurrence of this species throughout its range.

Geographic sampling

We sampled 42 localities from throughout Costa Rica (Fig. 1, Table 1), including sites from each of the four fish community provinces identified in Costa Rica by Smith and Bermingham (2005). To ensure broad geographic sampling within each fish province, we sampled multiple sites and included representatives from each of the 17 freshwater hydrographic regions in Costa Rica identified by Bussing (1998). Fish were collected by seine net and preserved in ethanol during field collection trips made from 2005-2008, then transported to the laboratory and stored at room temperature prior to DNA

sequencing. Individuals included in our genetic analyses have been deposited as voucher specimens at the Monte L. Bean Life Science Museum at Brigham Young University (Table 1).

Data collection

We extracted whole genomic DNA from tissue samples using a DNeasy Tissue Kit following the manufacturer's protocols (Qiagen, Valencia, CA). We used polymerase chain reaction (PCR) to amplify the entire mitochondrial cytochrome *b* gene (cyt *b*) using two primers (Glu31 5'-GTGACTTGAAAAACCACCGTT-3' and Poec.15931 5'-CCCTCGACTTTCGGTTTACAAG -3') that flanked the entire gene. When this failed to produce sufficient PCR product the gene was amplified in two halves using Glu31-HD and rainL484 (5'- CARTGAATCTGAGGAGGCTT -3')-Poec.15931. We also amplified a fragment of the nuclear S7 ribosomal protein gene (678 bp of the first intron) with a nested PCR design using primers 1F-3R (Chow & Hazama 1998) in the first reaction and 1F.RKR (5'-CTCTTCCCAGGCCGTCGTTG-3')-2R.PJU (5'-GACTCAAACTCGTCTGGCTT-3') in the second reaction; this nested PCR approach improves amplification of low copy nuclear sequences. Primer HD is from T. Schmidt as given in Dowling & Naylor (1997). All other primers (Glu31, rainL484, Poec.15931, 1F.RKR, and 2R.PJU) were provided by P. Unmack and R. Remington. Final concentrations for PCR components per 25 µL reaction were as follows: 25 ng template DNA, 0.25 µM of each primer, 0.625 units of Taq DNA polymerase, 0.1 mM of each dNTP, 2.5 µL of 10X reaction buffer and 2.5mM $MgCl_2$. We amplified the cyt *b* gene for 260 individuals using a standard thermal cycling profile (30 s 95° C, 30s 48° C, and 90s 72° C); this was repeated 35 times with 2 minutes initial denaturing at 95° C and a 7 minute final extension at 72° C. Amplification of the S7 sequences used the parameters

above except for 53° C as the annealing temperature and only 60 s polymerization time. We generated a haplotype network in the program TCS v1.21 (Clement et al. 2000) and selected 64 individuals to sequence for the S7 gene; these individuals represented the ancestral mitochondrial DNA (mtDNA) haplotypes in each clade and a subset of derived haplotypes. For our outgroup, we sequenced cyt *b* and S7 for *Poecilia latipinna*. Some sequencing reactions and clean up were performed using a Parallab 350 (Parallabs, Worcester, Massachusetts, USA). The other cycle-sequencing reactions were run with purified PCR products following the ABI manufacturers protocol (Applied Biosystems, Inc. Palo Alto, CA) and were purified using sephadex columns. Sequences were obtained using an Applied Biosystems 3730 XL automated DNA sequencer at the Brigham Young University DNA Sequencing Center. Chromatograms were edited in Sequencher 4.7 (Gene Codes, Ann Arbour, MI, USA) and aligned manually for both markers. We deposited all sequences on GenBank under accession numbers FJ446153-FJ446478.

Analyses

We used several analytical approaches to address the objectives outlined above. We first employed a traditional set of phylogenetic analyses (maximum parsimony and maximum likelihood). These analyses are useful in intraspecific studies to identify deep patterns of evolutionary divergence among well-supported clades. We also used a phylogenetic network reconstruction to resolve shallow relationships among closely related haplotypes, where a bifurcating tree is a poor representation of the mutation process (Crandall 1994). Together, these methods provide better phylogenetic resolution together than either could alone. To uncover historic patterns of gene flow and population structuring in *P. gillii*, we used three different phylogeographic approaches: analysis of molecular variance (AMOVA; Excoffier et al. 1992); theta (Θ) and migration

estimates (Beerli & Felsenstein 1999, 2001); and Bayesian skyline analysis (Drummond et al. 2005). Our motivation in using these different techniques was to uncover patterns of gene flow supported under different assumptions, providing a more robust set of conclusions. We present each of these methods in detail below, including justification for their use to address our basic questions about community province boundaries and historical patterns of gene flow in *P. gillii*.

We removed redundant sequences for both markers using MacClade v4.06 (Maddison & Maddison 2003) producing 71 unique haplotypes for cyt *b* and 30 unique haplotypes for S7. We estimated phylogenetic relationships among the haplotypes using both maximum parsimony and maximum likelihood criteria in the software program PAUP* (Swofford 1999). Our maximum parsimony analysis used equal weights under a heuristic search with 1000 random replicates and TBR branch swapping. Maximum likelihood analysis requires selecting an appropriate model of molecular evolution. We used the Akaike Information Criterion (AIC) in ModelTest v3.06 (Posada & Crandall 1998) to select the TrN+I+G model for cyt *b* and TVM+G model for S7. We used parameter estimates identified in ModelTest to run our maximum likelihood analysis implemented in the software package GARLI (Zwickl, 2006; http://www.bio.utexas.edu/faculty/antisense/garli/Garli.html) to estimate the relationships among haplotypes under a likelihood criterion. Nodal support was evaluated by 1000 bootstrap pseudo-replicates. We used the software program TCS v1.21 to create a minimum spanning network of the cyt *b* haplotypes using statistical parsimony with a 95% probability that no multiple substitutions had occurred.

For the S7 dataset, we identified two indels in the sequences obtained. Each indel was scored as a separate character at the end of the original sequence with two character states, one representing the presence and the other the absence of the indel (Graham et al.

2000; Kelchner 2000). However, we noted some variation within each of the indels that we wanted to consider. Hence, for each nucleotide position within the indel that exhibited variation we coded a separate character with the two states present in the indel and the ambiguity code for the individuals with the indel. After this coding was completed we excluded the indels from further analyses and the resulting sequences were subjected to the phylogenetic analyses described above to estimate the relationships among haplotypes.

To determine patterns of genetic structuring in *P. gillii* we used an AMOVA approach, treating our four community provinces as distinct groups. Our localities were distributed among these groups as follows: (1) Chiriqui (localities 1-8), (2) Chiapas-Nicaragua (localities 9-23), (3) San Juan (localities 24-40), and (4) Bocas (localities 41-42). We ran our AMOVA in the software package Arlequin v3.1 (Excoffier et al. 2005; Excoffier et al. 1992) and partitioned the total genetic variation into differences among the four provinces (F_{CT}), among the localities within provinces (F_{SC}), and among localities across the four provinces (F_{ST}). This allowed us to test the *a priori* prediction that the factors shaping community province boundaries would be reflected in patterns of genetic structuring within *P. gillii*. We also evaluated gene flow by estimating migration rates between the four community provinces using the MCMC method implemented in the software package Lamarc v2.1 (Kuhner 2006); we conducted three independent replicate runs with a search strategy of 10 initial chains and two long final chains. The initial chains were performed with 500 samples with a sampling interval of 20 (10,000 steps), using a burn-in of 1000 samples for each chain. The final chains were conducted with the same burn-in and sampling intervals, but with 10,000 samples (200,000 steps).

Incomplete lineage sorting, due to large effective population size or lack of sufficient time, could potentially explain patterns of shared haplotypes or lack of

monophyly among the four provinces. To account for this we used Θestimates from the software program Lamarc v2.1 (Kuhner 2006) as a measure of effective population size to better understand population structure. We also ran a Bayesian skyline plot in the program Beast v1.4 (Drummond & Rambaut 2007) to estimate the population size over time. We ran this analysis separately in each of the four community provinces to test if *P. gillii* in any of these regions had experienced population growth, decline, or stability over time. We ran a number of initial short runs of 10^5 generations to optimize the settings that were then used to conduct multiple longer runs with a chain length of $2x10^6$; population size parameters were set to sample every 1000 iterations after an initial burn-in of 20^5. Under these conditions, the MCMC process performed well, achieving good stationarity and yielding large effective sample sizes (>200) as determined in the software program Tracer v1.4 (Rambaut & Drummond 2007). Individuals from the Bocas Province did not achieve stationarity due to small sample size, hence these results are not presented.

RESULTS

Of 260 individual *P. gillii* sampled for cyt *b*, we found 71 unique haplotypes (labeled H01-H71, see Table 2). Only four of these haplotypes were shared across fish community boundaries (one across the Herradura Divide and three across the Northern Cordillera). We found 192 variable nucleotide positions (out of 1140 characters), and of these 96 were parsimony informative. Maximum uncorrected sequence divergence among all haplotypes was 4.76%. However, this high level of divergence was due to just two haplotypes (59 and 65); when these haplotypes were removed sequence divergence

dropped to 2.42%. The nucleotide base frequencies were A = 0.26, T = 0.31, C = 0.29, and G = 0.14.

The 64 individuals sampled for the S7 region produced 30 unique haplotypes (labeled nH1-nH30, see Table 3). As with the cyt *b* data, four haplotypes were shared across fish community boundaries (again, one across the Herradura Divide and three across the Northern Cordillera), however, they were not the same individuals or the same populations as seen in the cyt *b* dataset. We found 40 variable nucleotide positions (out of 615 characters) of which 20 were parsimony informative. The nucleotide base frequencies were A = 0.26, T = 0.30, C = 0.20, and G = 0.24.

Phylogenetic relationships

We did not find four distinct clades that correspond to the four fish community provinces. Rather, our phylogenetic analysis showed a clear break between two haplotypes from the San Juan Province (Clade I, Figure 2) and a large clade containing the remaining 69 haplotypes; this clade contained several well-supported subclades (Clades II-IX). None of the nine clades we identified contained all of haplotypes from a single community province except Clade III, which contains all haplotypes from the Chiriqui Province and one haplotype shared between the Chiriqui and Chiapas-Nicaragua provinces. Maximum parsimony and maximum likelihood analyses produced similar trees for the cyt *b* data set. Maximum parsimony analysis produced 30 equally parsimonious trees of 628 steps (CI = 0.776; RI = 0.898; RC = 0.697). Maximum likelihood analysis produced a topology similar to the consensus parsimony analysis with a log likelihood (-lnl) score = 3135.81397 (Figure 2). Support for the major nodes between the two analyses was consistent and differences existed only at the tips of the trees. Both criteria revealed that very few of the nodes showed greater than 90 bootstrap

support. Our network analyses showed two haplotype networks did not connect at the 95% parsimony level based on statistical parsimony: one network consisted of just two haplotypes (59 and 65); a second network included the remaining 69 haplotypes (results not shown). This larger network showed the same relationships found in the phylogenetic trees where the haplotypes fall out by community province.

Our phylogenetic analysis of the nuclear S7 data also did not produce trees with four clades that corresponded to the four provinces. Like the cyt *b* already presented, the two criteria produced similar topologies with differences at the tips while the major clades were the same and few nodes showed greater than 90 bootstrap support. Maximum parsimony analysis produced 2316 equally parsimonious trees of 44 steps (CI=0.932; RI=0.958; RC=0.893). Maximum likelihood produced a single tree with a –lnL score = 1157.68301. The tree consists of four major clades (Figure 3) from the 30 haplotypes. Clade I consists of a single haplotype (H29) from the San Juan Province and Clade II has both a San Juan Province and the single Bocas Province haplotype as sister taxa. Clade III contains all of the haplotypes for the Chiriqui Province with two more from the Chiapas-Nicaragua Province as a smaller subclade and one shared haplotype between the two communities nested within all the other Chiriqui haplotypes. The largest clade, Clade IV, is made up of mostly the San Juan Province haplotypes with one haplotype unique to the Chiapas-Nicaragua Province (H11) and three more haplotype shared between the two communities.

Phylogeographic patterns

Our analysis of molecular variance (AMOVA) revealed some evidence for genetic subdivision across the four community provinces. The largest proportion of the variation (39.09%) was explained among localities within provinces (F_{SC} = 0.5818, p <

0.01). Differences among the four provinces accounted for 32.81% of the variation (F_{CT} = 0.3380, $p < 0.01$) and differences among localities across the four provinces accounted for 28.10% of the variation (F_{ST} = 0.28501, $p < 0.01$).

Our coalescent-based migration estimates between each of the fish community provinces revealed patterns consistent with our phylogenetic estimates (Table 4). *Poecilia gillii* within the Chiriqui Province showed virtually no migration into both the Bocas and San Juan provinces and no migration in the opposite direction. Our analyses indicated that there is movement across the Herradura divide between populations in the Chiapas-Nicaragua and Chiriqui provinces. This is consistent with the observation of a single shared haplotype between these two regions for both our cyt *b* (H12, see Figure 2) and S7 results (nH07, Figure 3); the shared cyt *b* haplotype was found at the tip of the network (results not shown) suggesting it is a younger haplotype than those found in the interior of the network.

Migration estimates for the other three provinces (Table 4) show the San Juan and Chiapas-Nicaragua provinces have high levels of movement (>100) both ways across the Northern Cordillera. The fact that three shared haplotypes found in both cyt *b* and S7 datasets are shared between these two communities supports the high levels of gene flow between these two provinces. The migration rates further show that there is also high levels of movement (243.1896) from the San Juan Province into the Bocas Province, but a much smaller amount of movement (35.3665) the opposite direction. But our phylogenetic analyses revealed no shared haplotypes between these two provinces at all. However, haplotype 71 clusters with other haplotypes from San Juan (specifically, H66, H67, H68 which are all from the adjacent drainage across the Estrella divide), being very different from the other two haplotypes found in the Bocas Province (H69 and H70).

Our Bayesian skyline analyses revealed stable population sizes through time for each of the three community provinces examined here until the recent past, at which point each group shows a decline in population size (Figure 4). However, fish from the Chiapas-Nicaragua and San Juan provinces followed this decline with a rapid increase in effective population size (Figure 4x,c), a pattern not found in Bocas Province. The Θ values from Lamarc for each community show San Juan with the largest effective population size and Bocas with the smallest (Table 4). The other two communities have intermediate values that are closer to San Juan than Bocas.

DISCUSSION

Our primary focus in this study was to determine if *P. gillii* in Costa Rica showed phylogeographic breaks concordant with boundaries that delineate fish community provinces (Smith & Bermingham 2005). We also wanted to examine historical patterns of gene flow in *P. gillii* across the province boundaries. Contrary to our initial expectations, we did not find four distinct clades that mapped to the four community provinces. Our results do show that patterns of gene flow vary across each of the four boundaries examined here, suggesting that different factors could contribute to population structuring in different parts of this species' range. Here, we examine patterns of genetic structuring in *P. gillii* and explore what these data tell us about the role of abiotic factors on fish community composition.

Do P. gillii clades correspond to community provinces?

Only one of the four biogeographic provinces that we examined was associated with a single clade of *P. gillii*. All haplotypes (for both cyt *b* and S7) from the Chiriqui Province in southwestern Costa Rica are recovered in a single clade (Figure 1; Clade III

in Figures 2 and 3). This supports our hypothesis that the Southern Cordillera and the Herradura Divide are both important barriers to gene flow in *P. gillii*. The Southern Cordillera separates the Chiriqui from the San Juan and Bocas provinces and probably limits fish movement due to its high elevation (Coates & Obando 1996). The Herradura Divide contains several geo-physical features that might explain its importance as barrier to fish movement. A small mountain at this barrier runs westward from the Southern Cordillera to the Pacific Ocean, tapering off at a peninsula which juts slightly into the Gulf of Nicoya; there is also a very short continental shelf in this area relative to other parts of Costa Rica. These topographic features create a situation where the headwaters of adjacent rivers across this barrier are much closer to each other than their respective mouths. This would make river capture at the headwaters the most likely mode of fish movement. However, this area falls within the rain shadow in Costa Rica, creating low water conditions in rivers for a substantial part of each year. Combined, these features likely explain both the genetic break in *P. gillii* and the distributional limit of several other freshwater fishes as documented in Smith and Bermingham (2005).

The Estrella Divide in southeastern Costa Rica (Figure 1) is not a strong barrier to gene flow in *P. gillii*, despite its importance as a transitional zone from one fish community to another (Smith & Bermingham 2005). Although we did not find shared *P. gillii* haplotypes across this hypothetical barrier, *P. gillii* collected within the Bocas biogeographic province had haplotypes that were recovered in two different parts of our tree (Figure 2). Two mtDNA haplotypes from the Bocas Province form a basal part of our phylogeny and were distinct from all other clades (Clade II, Figure 2). One mtDNA haplotype and one nuclear DNA S7 haplotype grouped with haplotypes found in the Matina drainage, located directly north of the Estrella Divide (Clade V, Figure 2 & Clade II, Figure 3). Our migration estimates show high gene flow into the Bocas Province from

the San Juan Province, and low gene flow out of the Bocas Province. We note that except for drainage divides, there are no obvious geographical features limiting gene flow between these two biogeographic regions. Our inability to detect genetic structuring in *P. gillii* across this hypothetical barrier, coupled with the lack of a pronounced geo-physical barrier to fish movement, suggests the change in fish species composition in this region (described by Smith & Bermingham 2005) is likely governed by biotic interactions rather than earth history events.

Genetic structuring in *P. gillii* across the Northern Cordillera (delineating the Chiapas-Nicaragua and San Juan provinces) is less pronounced than that found across the Southern Cordillera. Three haplotypes for both the cyt *b* and S7 markers are shared among populations on each side of the hypothetical Northern Cordillera barrier, but all other haplotypes are unique to their respective provinces (Table 2). Smith and Bermingham (2005) noted only modest faunal turnover across the Northern Cordillera region, leading them to conclude this mountain range has not impeded the dispersal of all freshwater fishes. *Poecilia gillii* has likely moved across this barrier via a set of low passes found between major volcanic mountains that comprise much of this range or across the lowlands at the northernmost extent of these mountains (Figure 1). A recent study by Jones and Johnson (2009) examined historic levels of gene flow in a freshwater fish (*Xenophallus umbratilis*) which is found on both sides of this hypothetical barrier. Unlike the pattern we show for *P. gillii*, they found no shared haplotypes across the Northern Cordillera. However, populations found on either side of this divide were reciprocally monophyletic sister groups. We might expect the most likely place for gene flow across the Northern Cordillera barrier to be across the lowland in northern Costa Rica; yet cases of shared haplotypes across the drainages suggest another possibility (Tables 2 & 3). Cytochrome *b* haplotype 18 (H18) and S7 haplotypes 8 and 10 (nH08

and nH10) are found in localities in the southern region of the Northern Cordillera. These haplotypes are found in localities that straddle an area marked by mountains with no gaps, making potential fish movement via headwater capture difficult (Table 1, Figure 1). Mitochondrial haplotype 24 comes from localities in the central region of the Northern Cordillera, an area marked by low gaps providing areas for potential fish movement via headwater capture. Hence, our results suggest that there could be several areas of historic connection across the Northern Cordillera mountain range.

Incomplete lineage sorting

We recognize that incomplete lineage sorting could account for shared haplotypes observed between provinces. We considered this possibility as an alternative to recent gene flow in *P. gillii*, and found that in most cases incomplete lineage sorting was not well supported by our data. For example, haplotypes that pre-date the formation of boundaries should be older than more recently derived haplotypes unique to each province. However, we found that most of our shared haplotypes (H12, H18, H24) occurred at the tips of the haplotype networks (not shown), suggesting that they were more recently derived (Templeton et al. 1992); only cyt *b* haplotype (H13) appears to be ancestral. Incomplete lineage sorting might also be plausible if this species showed a consistently high effective population size over time (Figure 4). However, our data show that effective population in each of the four provinces has declined over time, and in two cases has recovered with an increase to current size. This bottleneck-like effect should result in a loss of rare haplotypes accelerating patterns of lineage sorting. Hence, although incomplete lineage sorting is plausible, it does not appear to be likely in this system.

Phylogeography and community ecology

What can phylogeographic patterns in *P. gillii* tell us about the nature of fish community provinces in Central America? The striking congruence between phylogeographic breaks in *P. gillii* and fish community province boundaries located at the Herradura Divide and the Southern Cordillera suggest that common vicariant events could explain both patterns. Although we cannot know with certainty what historical factors might contribute to these patterns, mountain barriers to fish movement as described above appear to be good candidates. Interestingly, we also found regions of discordance between our *P. gillii* data and fish community province boundaries. *Poecilia gillii* show evidence of gene flow across both the Northern Cordillera boundary and the Estrella Divide. One conclusion that we can take from this is that vicariant barriers to fish movement alone cannot account for the abrupt shift in fish fauna described across these boundaries by Smith and Bermingham (2005). Hence, community ecologists interested in understanding how these fish communities have assembled might look to biotic factors as potential drivers of community composition. Such inferences could be strengthened by evaluating phylogeographic breaks in several additional fish species with broad distributions that largely overlap with *P. gillii*.

Several researchers have called for a greater dialogue between community ecology and molecular ecology, emphasizing that these disciplines can mutually benefit by sharing tools and perspectives in addressing compelling questions (Vamosi et al. 2009; Emerson and Gillespie 2008, Webb et al. 2006, Webb et al. 2002). However, such discussions have focused primarily on using species-level phylogenetic trees in community ecology (Vamosi et al. 2009; Emerson and Gillespie 2008). We show that within-species phylogeography might also be beneficial in community ecology by pointing to cases where abiotic factors may or may not contribute to community composition. Specifically,

our results point to potential barriers that may have isolated the Chiriqui Province during the past three million years, contributing to the high levels of freshwater fish endemism now found there (Smith & Bermingham 2005). Moreover, in contrast to traditional biogeographic approaches that focus on narrowly distributed or endemic species to draw inference about vicariant events (Humphries & Parenti 1999; Morrone & Crisci 1995; Nelson & Platnick 1981), our study demonstrates the utility of common, widespread species for studies exploring and testing biogeographic patterns. A comparative phylogeographic study using multiple widespread taxa in this region would further clarify the processes occurring at each of the boundaries, and probably better aid ecologists in understanding the potential role of abiotic factors in shaping the distribution of freshwater fishes. While our study does not settle the long-standing debate over community assembly being a random or non-random process, it does suggest that phylogeography is an additional tool available for understanding community formation.

REFERENCES

Abell R, Thieme ML, Revenga C et al. (2008) Freshwater ecoregions of the world: a new map of biogeographic untis for freshwater biodiversity conservation. *BioScience*, 58, 403-414.

Avise JC (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography*, 36, 3-15.

Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Massachusetts.

Avise JC (1994) Molecular markers, natural history, and evolution. Chapman and Hall, New York.

Avise JC (1992) Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos*, 63, 62-76.

Avise JC, Arnold J, Ball RM et al. (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, 18, 489-522.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *PNAS*, 98, 4563-4568.

Beerli P, Felsenstein J (1999) Maximum likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, 152, 763-773.

Burton RS (1998) Intraspecific phylogeography across the Point Conception biogeographic boundary. *Evolution*, 52, 734-745.

Bussing WA (1998) *Freshwater fishes of Costa Rica*, 2nd edn. Editorial de la Universidad de Costa Rica, San Jose, Costa Rica.

Chapman LJ (1992) Variation in the structure of *Poecilia gillii* populations. *Copeia*, 1992, 908-914.

Chapman LJ, Kramer DL, Chapman CA (1991) Population dynamics of the fish *Poecilia gillii* (Poeciliidae) in pools of an intermittent tropical stream. *Journal of Animal Ecology*, 60, 441-453.

Chow S, Hazama K (1998) Universal PCR primers for S7 ribosomal protein gene introns in fish. *Molecular Ecology*, 7, 1255-1256.

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

Coates AG, Obando JA (1996) The geologic evolution of the Central America Isthmus. In: Evolution and Environment in Tropical America (eds. Jackson JBC, Budd AF, Coates AG), pp. 21–56. Chicago University Press, Chicago, IL.

Connell JH (1972) Community interactions on marine intertidal shores. *Annual Review of Ecology and Systematics*, 3, 169-192.

Crandall KA (1994) Intraspecific cladogram estimation: accuracy at higher levels of divergence. *Systematic Biology*, 43, 222–235.

Connor EF, Simberloff D (1979) The assembly of species communities: chance or competition? *Ecology*, 60, 1132-1140.

Diamond JM (1975) Assembly of species communities In: Ecology and evolution of communities (ed. Cody ML and Diamond JM) pp. 342-444. Harvard University Press, Cambridge.

Dowling TE, Naylor GJP (1997) Evolutionary relationships of minnows in the genus *Luxilus* (Teleostei: Cyprinidae) as determined from cytochrome *b* sequences. *Copeia*, 1997, 758-765.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*,7, 214.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, 22, 1185-1192.

Dunson WA, Travis J (1991) The role of abiotic factors in community organization. *The American Naturalist*, 138, 1067-1091.

Emerson BC, Gillespie RG (2008) Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology and Evolution*, 23, 619-630.

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.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, 47-50.

Gillespie R (2004) Community assembly through adaptive radiation in Hawaiian spiders. *Science*, 303, 356-359.

Graham SW, Reeves PA, Burns ACE, Olmstead RG (2000) Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. *International Journal of Plant Science*, 161, S83-S96.

Hubert N, Renno JF (2006) Historical biogeography of South American freshwater fishes. *Journal of Biogeography*, 33, 1414-1436.

Humphries CJ, Parenti LR (1999) Cladistic Biogeography: interpreting patterns of plant and animal distributions. Oxford University Press, New York.

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: In press.

Kelchner SA (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Gardens*, 87, 482-498.

Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics*, 22, 768-770.

Maddison DR, Maddison WP (2003) MacClade 4: Analysis of phylogeny and character evolution. Version 4.06. Sinauer Associates, Sunderland, Massachusetts.

Morrone JJ, Crisci JV (1995) Historical biogeography: introduction to methods. *Annual Review of Ecology and Systematics*, 26, 373-401.

Nelson G, Platnick NI (1981) Systematics and biogeography: cladistics and vicariance. Columbia University Press, New York.

Posada D, Crandall KA (1998) MODELTEST: Testing the model of DNA substitution. *Bioinformatics*, 14, 817-818.

Rambaut A, Drummond AJ (2007) Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Reyjol Y, Hugueny B, Pont D et al. (2007) Patterns in species richness and endemism of European freshwater fish. *Global Ecology and Biogeography*, 16, 65-75.

Ronquist F (1997) Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, 46, 195-203.

Smith SA, Bermingham E (2005) The biogeography of lower Mesoamerican freshwater fishes. *Journal of Biogeography*, 32, 1835-1854.

Swofford DL (1999) *PAUP*: Phylogenetic Analysis Using Parsimony (*And Other Methods), Version 4.0b2*. Sinauer Associates, Sunderland, Massachusetts, USA.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, 132, 619-633.

Tilman D (1988) Plant strategies and the dynamics and structure of plant communities. Princeton University Press, Princeton, New Jersey.

Toft CA (1985) Resource partitioning in amphibians and reptiles. *Copeia*, 1985, 1-21.

Unmack PJ (2001) Biogeography of Australian freshwater fishes. *Journal of Biogeography*, 28, 1053-1089.

Vamosi SM, Heard SB, Vamosi JC, Webb CO (2009) Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular Ecology*, 18, 572-592.

Webb CO, Losos JB, Agrawal AA (2006) Integrating phylogenies into community ecology. *Ecology*, 87, S1-S2.

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ (2002) Phylogenies and community ecology. *Annual Review of Ecology & Systematics*, 33, 475-505.

Yap SY (2002) On the distributional patterns of Southeast-East Asia freshwater fish and their history. *Journal of Biogeography*, 29, 1187-1199.

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

TABLE 1. Collection data for localities sampled in this study. Names and locality numbers correspond to those used in Tables 2 and 3 and in Figure 1. Sample size indicates the number of individuals collected from each locality. Museum voucher numbers are for fish accessioned into the Monte L. Bean Life Science Museum, Brigham Young University.

Locality	Name	GPS Coordinates	Sample Size	BYU museum numbers
1	Rio Nuevo	8° 38'27.70" N 82° 57'10.70" W	2	235222-235223
2	Rio Barrigones	8° 35'35.64"N 83° 25'18.54" W	2	235220-235221
3	Rio Pejibaye	9° 9'25.00" N 83° 34'31.00" W	15	109934-109948
4	Rio General	9° 23'22.00" N 83° 39'49.00" W	2	108769-108770
5	Rio Hatillo Viejo	9° 18'12.20" N 83° 54'40.60" W	4	235224-235227
6	Rio Savegre	9° 22'23.20" N 84° 1'31.50" W	4	235362-235635
7	Finco la Palma	9° 32'13.10" N 84° 23'9.20" W	4	235216-235219
8	Rio Chires	9° 34'57.60" N 84° 24'43.10" W	4	235358-235361
9	Quebrada La Canela	9° 51'5.44" N 84° 31'39.58" W	3	129610-129612
10	Rio Tarcoles	9° 52'47.28" N 84° 31'40.08" W	3	129607-129609
11	Rio Pacacua	9° 55'10.56" N 84° 14'28.68" W	5	129602-129606
12	Rio Rosales	10° 1'47.24" N 84 ° 19'32.95" W	5	129613-129617
13	Rio Centeno	9 ° 56'28.75 N 84 ° 32'19.90 W	5	129618-129622
14	Rio Naranjo	10° 1'21.50" N 84° 44'3.91" W	5	129623-129627
15	Rio Congo	10° 14'23.93" N 84° 59'30.16" W	6	129628-129633
16	Rio Javilla	10° 22'19.50" N 85° 5'50.64" W	3	118871-118872, 234740
17	Rio Santa Rosa	10° 27'40.08" N 85° 4'27.78" W	2	118868, 119013
18	Rio Magdalena	10° 28'46.02" N 85° 4'41.22" W	3	118869-118870, 119014
19	Small ditch	10° 37'26.64" N 85° 3'29.22" W	3	118862-118863, 119010
20	Quebrada Homiguera	10° 41'27.24" N 85° 5'1.14" W	5	118864-118865, 119011, 235127, 235129
21	Rio Marole	10° 3'29.82" N 85° 15'43.26" W	16	235212-235215, 235445-235456
22	Rio Carrisal	10° 23'42.06" N 85° 35'12.78" W	1	235211

23	Rio Irigary	10° 43'24.24" N 85° 30'37.38" W	15	235208-235210, 235457-235468
24	Rio Sapoa	11° 2'39.72" N 85° 36'57.24" W	8	235469-235476
25	Rio Sabalo	11° 2'34.20" N 85° 29'21.18" W	8	235204-235207, 235477-235480
26	Rio Queques	10° 38'41.34" N 84° 49'20.04" W	16	135258-135259, 235200-235203, 235435-235444
27	Rio Sabalito	10° 32'54.90" N 84° 58'50.88" W	5	118860-118861, 119009, 235356-235357
28	Rio Chiquito	10° 26'15.72" N 84° 52'5.34" W	5	118866-118867, 119012, 235354-235355
29	Lake Arenal	10° 28'19.50" N 84° 46'9.60" W	2	118859, 119008
30	Rio La Palma	10° 29'55.50" N 84° 41'20.40"W	5	118857-118858, 119007, 119773, 119776
31	Rio Infiernito	10° 37'4.86" N 84° 29'3.06" W	8	235196-235197, 235423-235428
32	Rio Chimurria	10° 43'39.60" N 84° 33'31.44" W	8	235198-235199, 235429-235434
33	Rio Sarapiqui	10° 31'28.38" N 84° 1'52.80" W	13	235192-235195, 235414-235422
34	Rio Isla Grande	10° 23'34.86" N 83° 58'5.40" W	1	232258
35	Rio Tortuguero	10° 15'33.90" N 83° 48'44.04" W	15	235188-235191, 235402-235409, 235411-235413
36	Rio Parismina	10° 11'51.78" N 83° 39'7.44" W	7	135233, 235396-235401
37	Rio Herediana	10° 7'27.00" N 83° 33'22.20" W	5	235186-235187, 235369-235371
38	Rio Reventazon	9° 52'20.28" N 83° 37'59.52" W	5	235184-235185, 235366-235368
39	Rio Toro	10° 1'0.42" N 83° 12'36.78" W	8	235182-235183, 235390-235395
40	Unnamed lagoon	9° 53'33.30" N 82° 58'20.22" W	8	235180-235181, 235384-235389
41	Rio Carbon	9° 37'23.22" N 82° 51'18.72" W	8	235178-235179, 235378-235383
42	Rio Sixaola	9° 37'55.32" N 82° 49'9.18" W	8	235176-235177, 235372-235377

Table 2. Haplotypes generated from the cytochrome *b* sequences presented by locality (see Figure 1). Localities are identified by the biogeographic province in which they occur (following Smith & Bermingham 2005). The number in each cell indicates the number of individuals from a particular locality with a particular haplotype; empty cells denote zero, and shaded rows show haplotypes that are shared between provinces.

Localities (localities 1–23)

Haplotypes	Chiriqui								Chiapas-Nicaragua														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	2																						
2		2																					
3			1																				
4			12	1	4																		
5			1																				
6			1																				
7				1																			
8						3																	
9						1																	
10							3																
11							1																
12								4	1														
13									1					3									
14									1														
15										1	5	1											
16										1													
17										1							1						7
18												1											
19												2									1		
20												1											
21													4										
22													1										
23														2									
24															2		1	3				1	
25															1								
26															2								
27															1						14		
28																1			3	1			
29																1							
30																1							
31																				1			
32																				1			
33																				1			
34																				1			1
35																					1		
36																							6
37																							1

Localities (localities 24–42)

Haplotypes	San Juan																	Bocas	
	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
1																			
2																			
3																			
4																			
5																			
6																			
7																			
8																			
9																			
10																			
11																			
12																			
13	1																		
14																			
15																			
16																			
17																			
18													7	2					
19																			
20																			
21																			
22																			
23																			
24					2														
25																			
26																			
27																			
28																			
29																			
30																			
31																			
32																			
33																			
34																			
35																			
36																			
37																			

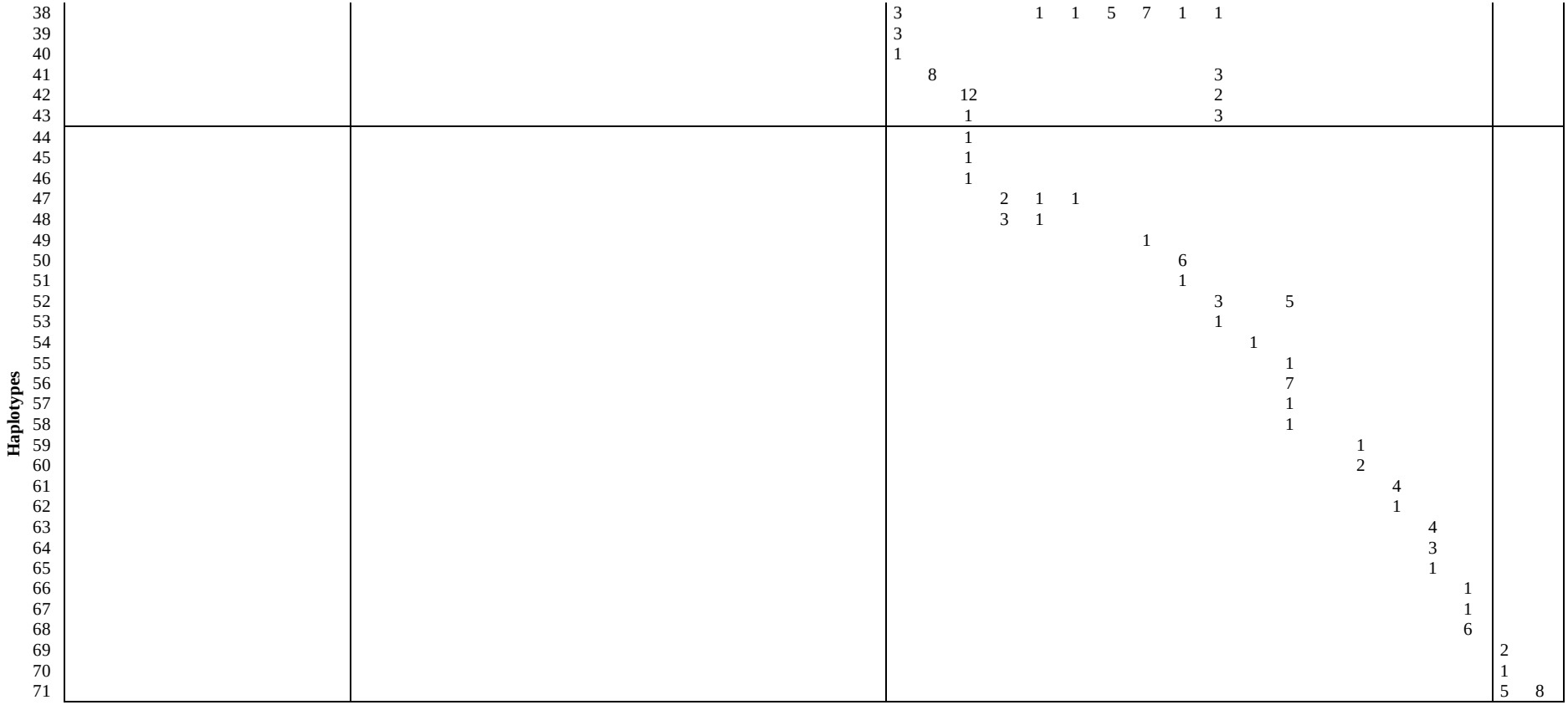

Haplotypes
38 3 1 1 5 7 1 1
39 3
40 1
41 8 3
42 12 2
43 1 3
44 1
45 1
46 1
47 2 1 1
48 3 1
49 1
50 6
51 1
52 3 5
53 1
54 1
55 1
56 7
57 1
58 1
59 1
60 2
61 4
62 1
63 4
64 3
65 1
66 1
67 1
68 6
69 2
70 1
71 5 8

TABLE 3. Haplotypes generated from the S7 sequences showing the locality number (see Figure 1) and name along with which province it came from. The numbers and shading are as described in Table 2.

Haplotypes	Localities: Chiriqui							Chiapas-Nicaragua										San Juan																	Bocas	
	1	2	3	4	5	6	7	9	11	12	14	15	18	19	20	21	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42
1	1	2																																		
2	1																																			
3			1		1																															
4			1	2																																
5					1																															
6						2																														
7							2			1				2	1																					
8								1																					1	1		2	1			
9									1				2			2																				
10										1																					1					
11											1																									
12												1																								
13																	1								1		1									
14																		1																		
15																		1																		
16																			2	1					1	1										
17																				1																
18																					1	1	1													
19																						1														
20																							1													
21																								1												
22																								1												
23																										1										
24																											1									
25																												1								
26																													1							
27																														1	1					
28																																	1			
29																																		2		
30																																			2	2

TABLE 4. Migration rates (individuals per generation) and theta estimates ($\Theta=2N_e\mu$) for each of the four fish community provinces calculated in Lamarc. Theta estimates are given in the second column for each of the respective provinces in the far left column. For the migration rates, provinces in the first column represent the source of the migrants and provinces on the top row represent the destination of the migration.

	Θ	**Chiriqui**	**Chiapas-Nicaragua**	**Bocas**	**San Juan**
Chiriqui	0.012832		78.81793	0.000107	0.000053
Chiapas-Nicaragua	0.007608	86.54454		40.52998	116.4214
Bocas	0.001052	0.000119	128.8017		35.36652
San Juan	0.015687	0.000116	182.8834	243.1896	

FIGURE LEGENDS

Figure 1. Map of Costa Rican rivers showing the boundaries (The Estrella Divide and Herradura Divide labeled in italics with arrows, the Northern Cordillera marked by a long dashed line, and the Southern Cordillera marked by a small dotted line) among Smith and Bermingham's (2005) fish community provinces (labeled in bold). Sampling localities are identified by numbered circles; locality numbers correspond to those presented in Table 1.

Figure 2. Maximum likelihood phylogram of 71 cytochrome *b* haplotypes of *Poecilia gillii* with nine major clades labeled in Roman numerals (as used in the text). Each haplotype is followed by its respective community province: Bocas (BO), Chiapas-Nicaragua (CN), Chiriqui (CH), and San Juan (SJ). Haplotypes shared between provinces are identified first by the province in which the haplotype is most common. Nodal support based on 1000 bootstrap pseudoreplicates.

Figure 3. Maximum likelihood phylogram of 30 S7 haplotypes of *Poecilia gillii* with four major clades labeled in Roman numerals (as used in the text), abbreviations and all other details are as in Figure 2.

Figure 4. Bayesian skyline plot for cytochrome *b* individuals showing change in effective population size over time in *P. gillii* in three different community provinces: (a) the Chiriqui Province, (b) the Chiapas-Nicaragua Province, and (c) the San Juan Province. Solid lines represent the mean effective population size and dotted lines represent the upper and lower 95% confidence intervals.

FIGURE 1.

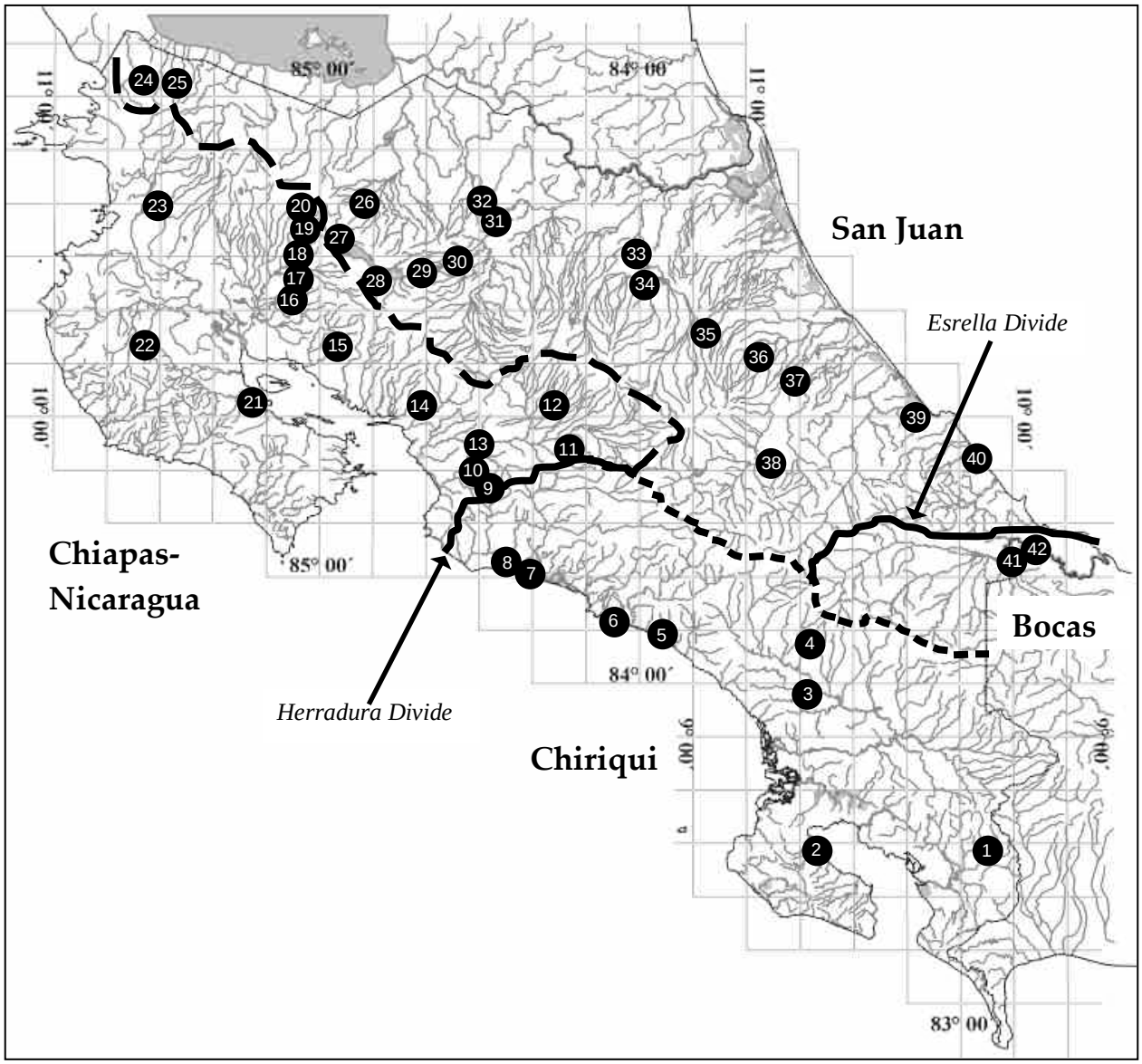

85° 00´
84° 00´
11° 00´
10° 00´
9° 00´
San Juan
Esrella Divide
Chiapas-
Nicaragua
Bocas
Herradura Divide
Chiriqui
83° 00´
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

FIGURE 2.

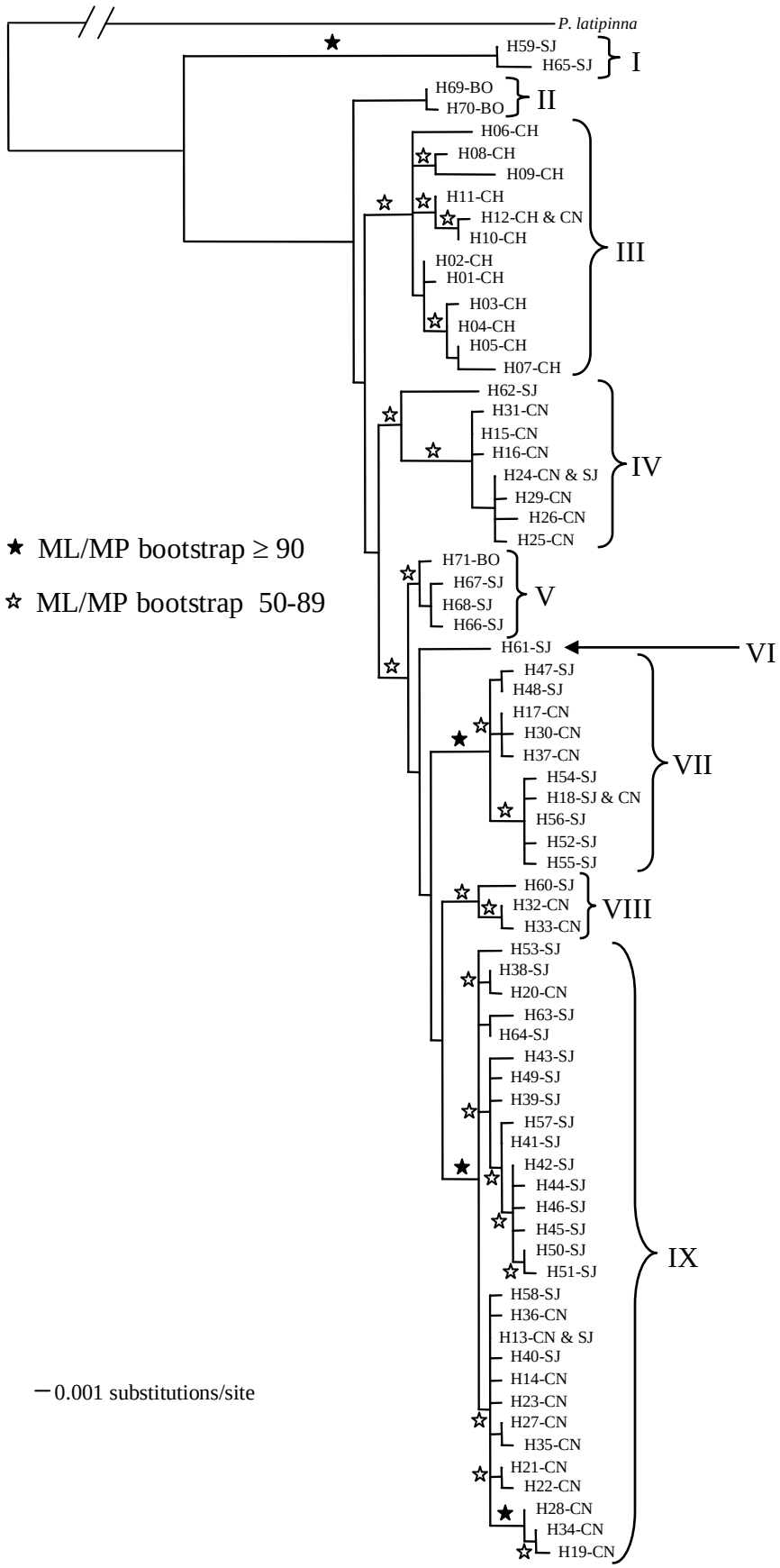

P. latipinna
H59-SJ
H65-SJ
I
H69-BO
H70-BO
II
H06-CH
H08-CH
H09-CH
H11-CH
H12-CH & CN
H10-CH
H02-CH
H01-CH
H03-CH
H04-CH
H05-CH
H07-CH
III
H62-SJ
H31-CN
H15-CN
H16-CN
H24-CN & SJ
H29-CN
H26-CN
H25-CN
IV
H71-BO
H67-SJ
H68-SJ
H66-SJ
V
H61-SJ
VI
H47-SJ
H48-SJ
H17-CN
H30-CN
H37-CN
H54-SJ
H18-SJ & CN
H56-SJ
H52-SJ
H55-SJ
VII
H60-SJ
H32-CN
H33-CN
VIII
H53-SJ
H38-SJ
H20-CN
H63-SJ
H64-SJ
H43-SJ
H49-SJ
H39-SJ
H57-SJ
H41-SJ
H42-SJ
H44-SJ
H46-SJ
H45-SJ
H50-SJ
H51-SJ
H58-SJ
H36-CN
H13-CN & SJ
H40-SJ
H14-CN
H23-CN
H27-CN
H35-CN
H21-CN
H22-CN
H28-CN
H34-CN
H19-CN
IX
★ ML/MP bootstrap ≥ 90
☆ ML/MP bootstrap 50-89
0.001 substitutions/site

FIGURE 3.

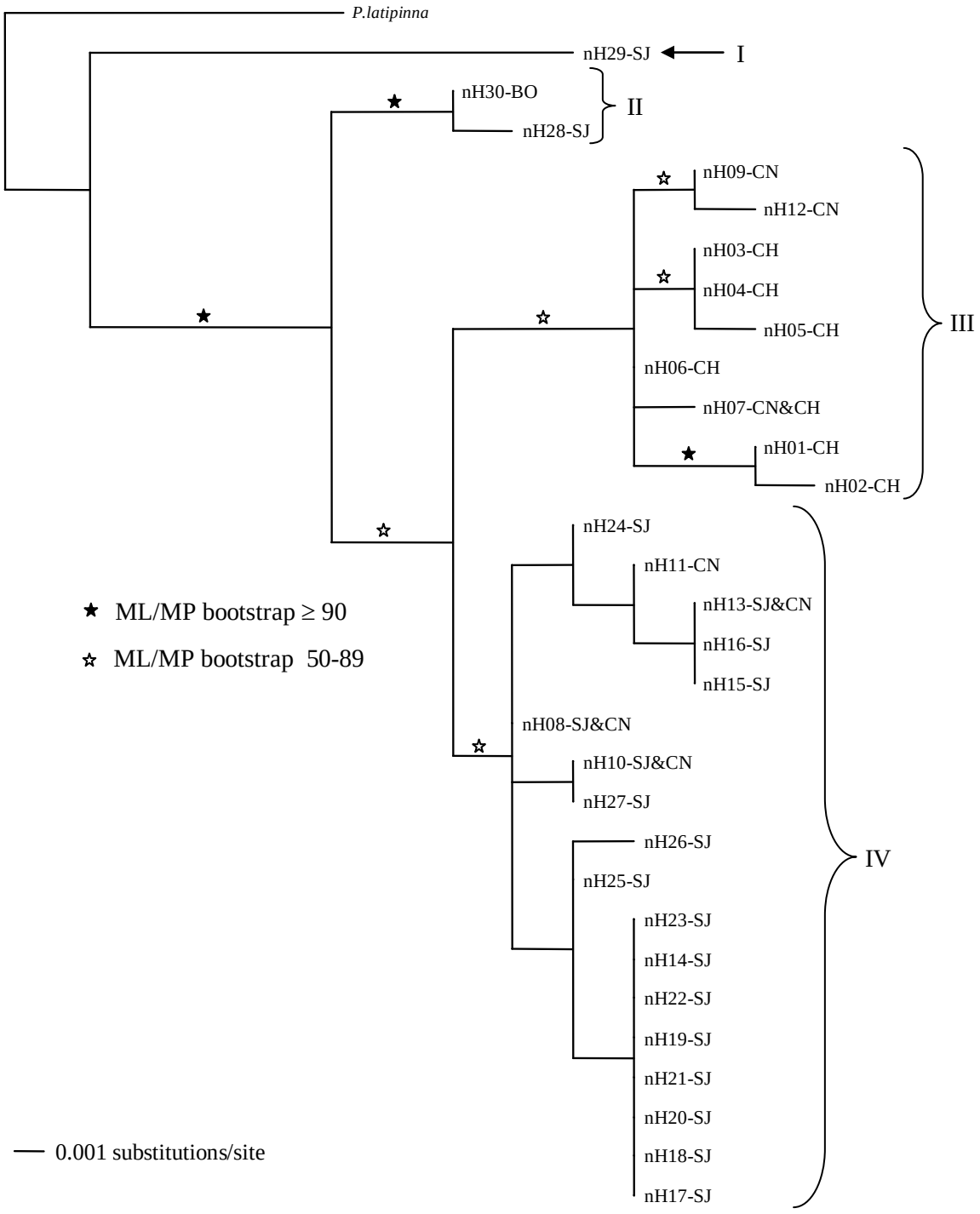
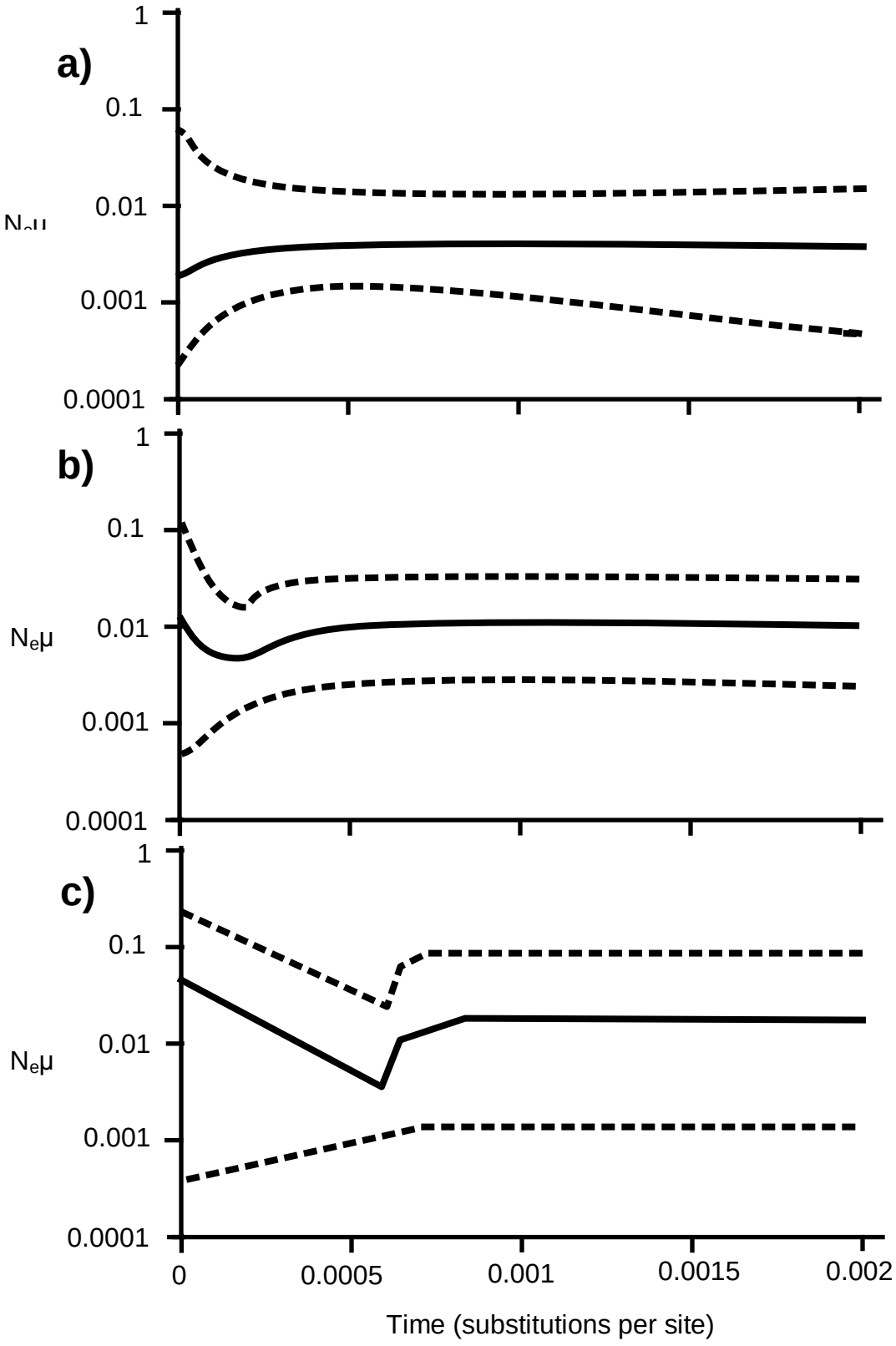

P.latipinna
nH29-SJ
I
nH30-BO
nH28-SJ
II
nH09-CN
nH12-CN
nH03-CH
nH04-CH
nH05-CH
nH06-CH
nH07-CN&CH
nH01-CH
nH02-CH
III
nH24-SJ
nH11-CN
nH13-SJ&CN
nH16-SJ
nH15-SJ
nH08-SJ&CN
nH10-SJ&CN
nH27-SJ
nH26-SJ
nH25-SJ
nH23-SJ
nH14-SJ
nH22-SJ
nH19-SJ
nH21-SJ
nH20-SJ
nH18-SJ
nH17-SJ
IV
ML/MP bootstrap ≥ 90
ML/MP bootstrap 50-89
0.001 substitutions/site

FIGURE 4.


a)
b)
c)
1
0.1
0.01
0.001
0.0001
$N_e\mu$
0
0.0005
0.001
0.0015
0.002
Time (substitutions per site)