Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America

E. BERMINGHAM* and A. P. MARTIN†
*Smithsonian Tropical Research Institute, Apartado 2072, Balboa, Republic of Panama, † Department of Biology, University of Nevada, Las Vegas, NV 89154–4004, USA

Abstract

Historical biogeography seeks to explain contemporary distributions of taxa in the context of intrinsic biological and extrinsic geological and climatic factors. To decipher the relative importance of biological characteristics vs. environmental conditions, it is necessary to ask whether groups of taxa with similar distributions share the same history of diversification. Because all of the taxa will have shared the same climatic and geological history, evidence of shared history across multiple species provides an estimate of the role of extrinsic factors in shaping contemporary biogeographic patterns. Similarly, differences in the records of evolutionary history across species will probably be signatures of biological differences. In this study, we focus on inferring the evolutionary history for geographical populations and closely related species representing three genera of primary freshwater fishes that are widely distributed in lower Central America (LCA) and northwestern Colombia. Analysis of mitochondrial gene trees provides the opportunity for robust tests of shared history across taxa. Moreover, because mtDNA permits inference of the temporal scale of diversification we can test hypotheses regarding the chronological development of the Isthmian corridor linking North and South America. We have focused attention on two issues. First, we show that many of the distinct populations of LCA fishes diverged in a relatively brief period of time thus limiting the phylogenetic signal available for tests of shared history. Second, our results provide reduced evidence of shared history when all drainages are included in the analysis because of inferred dispersion events that obscure the evolutionary history among drainage basins. When we restrict the analysis to areas that harbour endemic mitochondrial lineages, there is evidence of shared history across taxa. We hypothesize that there were two to three distinct waves of invasion into LCA from putative source populations in northwestern Colombia. The first probably happened in the late Miocene, prior to the final emergence of the Isthmus in the mid-Pliocene; the second was probably coincident with the rise of the Isthmus in the mid-Pliocene, and the third event occurred more recently, perhaps in the Pleistocene. In each case the geographical scale of the dispersion of lineages was progressively more limited, a pattern we attribute to the continuing development of the landscape due to orogeny and the consequent increase in the insularization of drainage basins. Thus, the fisheye view of LCA suggests a complex biogeographic history of overlaid cycles of colonization, diversification, sorting and extinction of lineages.

Keywords: dispersal, freshwater fishes, historical biogeography, mitochondrial DNA, neotropics, speciation

Introduction

Over the past 2–7 Myr there has been extensive intercontinental exchange of flora and fauna between North and South America across the isthmian bridge of

Correspondence: E. Bermingham, Smithsonian Tropical Research Institute, Unit 0948, APO AA 34002–0948, USA. Tel.: +01-507-2284339; Fax: +01-507-2280516; E-mail: eb@naos.si.edu

Panama, a phenomenon known as the Great American Interchange because of its importance for New World biogeography (Marshall *et al.* 1979; Stehli & Webb 1985). Freshwater fishes participated in the Great American Interchange (Miller 1966; Myers 1966; Bussing 1976, 1985a), although biogeographic studies of the regional fishes have been considerably less celebrated than the detailed and instructive studies of mammals (Marshall

et al. 1979; Webb 1985; Lundelius 1987; Marshall 1988; Marshall & Sempere 1993). Yet, because the dispersal of freshwater fishes is dependent on direct connections between drainage basins, and because the history of basin interconnections reflects the underlying geological development of landscapes, historical biogeographic analyses of freshwater fishes permit strong inference regarding the biotic and geological evolution of a region (Lundberg 1993).

Owing to its recent origins, lower Central America (LCA) provides a special opportunity to study the modern assembly and diversification of a biota. For this purpose, rivers provide a useful setting because they create repeated patterns of subdivision of taxa into discrete subpopulations and permit characterization of the dynamics of dispersal through a region and its influence on geographical differentiation and formation of new species. A difficulty with this approach, and a stumbling block for historical biogeography in general, has been the accurate assessment and phylogenetic ordering of independent evolutionary lineages. Species are typically used as the lowest level of biogeographic analysis and this tradition precludes robust inference for areas which are predominately inhabited by widespread species. The historical biogeographic interpretation of regions harbouring widespread species is one in which dispersal dominates the explanation of species' distributions. However, biological diversity is hierarchically structured at all levels; therefore, considerable information about the historical biogeography of a region can be obtained at levels below the species (i.e. among populations) provided that reliable estimates of population phylogenies can be obtained (Rosen 1978; Chernoff 1982; Bermingham & Avise 1986).

Inferences of historical biogeography emerge from demonstrations that multiple, independent taxa share a common history (Croizat *et al.* 1974; Rosen 1976, 1978; Platnick & Nelson 1978; Bermingham & Avise 1986; Brundin 1988; Avise 1992; Patton *et al.* 1994; Joseph *et al.* 1995). To fully explore the history of the LCA landscape requires testing the hypothesis that taxa with overlapping distributions share a common history. The null hypothesis is, of course, that each species has its own unique history and any apparent shared history across species has occurred by chance. There are two dimensions of shared earth history: geography and time. Thus, an ideal tool for inferring historical biogeography is one which provides a sound hypothesis of area relationships and permits the amount of change between terminal taxa to be at least approximately translated into time.

To characterize the relationships among areas, we have adopted mitochondrial DNA (mtDNA) as a tool for establishing phylogenetic relationships among individuals (e.g. Avise *et al.* 1979; Bermingham & Avise 1986;

Moritz *et al.* 1987; Patton *et al.* 1994). Individual taxa on the tree (mtDNA clades in our analyses) are replaced by locations, providing a description of the phylogenetic relationships among localities, a so-called area cladogram (Platnick & Nelson 1978; Nelson & Platnick 1981; Humphries & Parenti 1986). Mitochondrial DNA-based phylogeography has two basic contributions to make to the study of biological diversity. The first, descriptive component of phylogeographic analysis comprises the collection of the raw genetic data and phylogenetic analyses of these data. DNA sequence data provide information about the phylogenetic relationships, and genetic distances among, evolutionary lineages and permit inferences regarding the history of individual taxa within a region (e.g. Avise *et al.* 1979; Bermingham & Avise 1986; Avise 1994; Patton *et al.* 1994; Templeton *et al.* 1995; Joseph *et al.* 1995; Bermingham *et al.* 1996). The diversity of lineages within populations provides information concerning the history of that population, including estimates of its effective size (Felsenstein 1992; Fu 1994). The geographical distribution of lineages may be used to infer the history of a taxon's spread and the presence of continuing migration between populations (Slatkin 1989; Slatkin & Madison 1989; Hudson *et al.* 1992; Templeton *et al.* 1995). When nucleotide substitution rates are homogeneous with respect to taxon, lineage and time, one may also estimate the relative chronology of colonization and diversification events (Page 1991, 1996). The second component, comparative phylogeography, tests for consistencies in the evolutionary and distributional histories of taxa with respect to the particular geographical and ecological setting of a region, the timescale of phylogeographic events, intrinsic characteristics of taxa, and ages of populations (Bermingham & Avise 1986; Avise *et al.* 1987; Bermingham *et al.* 1992; Joseph *et al.* 1995). Phylogeographic information may also permit inferences about the extinction of taxa in the absence of fossil data. Thus, mtDNA provides an objective framework for the formal analysis of the link between population processes and biogeographic patterns of diversity.

Methods

The LCA region and its fishes

Our collections and phylogeographic analyses of neotropical freshwater fishes have focused on LCA, extending from the Choco region of north-western Colombia to the south shore of Lake Nicaragua and its outflow to the Atlantic, the San Juan River (Fig. 1). The modern LCA landscape is dominated by a geologically complex central cordillera reaching an elevation of 3000 m in Costa Rica and western Panama, then dipping
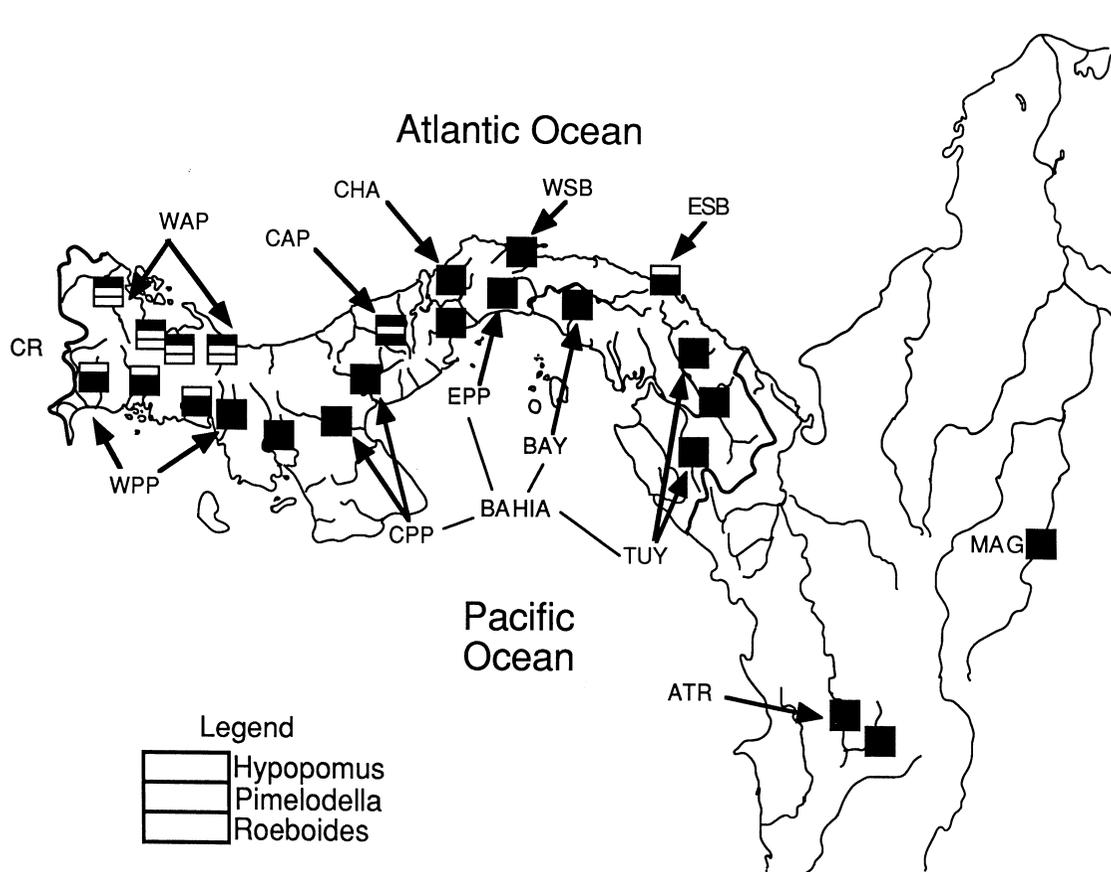
Fig. 1 Map showing the distribution of the freshwater fish genera *Hypopomus*, *Pimelodella*, and *Roeboides* in lower Central America (CA) and identification of the major drainage areas included in the historical biogeographic analysis of LCA. The shaded boxes refer to the distribution of the genera shown in the legend.

to 200 m in central Panama before rising again to approximately 600 m in eastern Panama. In general, the Pacific versant is more extensive than the Atlantic slope and this relationship reaches an extreme in eastern Panama, where the Atlantic coastal plain is only 5 km wide and rivers are rarely longer than 15 km. Notwithstanding some important exceptions, rivers of both slopes are moderately short and steep and all LCA rivers are subject to extreme variations in flow on both seasonal and daily timescales.

There can be little doubt that the present-day physiography and hydrology of LCA play a role in the distribution patterns of LCA freshwater fishes. Of greater importance, however, is the geological evolution of LCA and the development and drainage relationships of the region's rivers. The geological history of LCA is uncertain from the Cretaceous, when North (NA) and South America (SA) first separated (Pindell & Barrett 1990), until the late Tertiary marking the genesis of modern LCA and ultimate reconnection of the South and North

American continents approximately 3–4 Myr BP (Coates & Obando 1996). Whether the continents were connected at any time in between (early to mid-Tertiary) remains unknown. The most vigorously defended interim connection is the proto-Antillean arc posited to have joined SA to nuclear CA and NA sometime prior to the Eocene (Rosen 1976, 1985; Pindell & Dewey 1982; Savage 1982; Sykes *et al*. 1982).

Primary freshwater fishes, defined as physiologically intolerant of marine conditions and considered to disperse via freshwater only (Myers 1938), provide little or no evidence of an early Tertiary corridor connecting SA to CA (Miller 1966; Myers 1966; but see Bussing 1985a). Thus we can centre our focus on the physiographic evolution of the modern LCA landscape, a process extending over the last 15 Myr (reviewed by Coates & Obando 1996). For freshwater fish dispersal and biogeography the following features of LCA geography and geology are of particular importance. (i) In the mid-Miocene, approximately 15 Mya, CA and SA were separated by an

ocean barrier of abyssal–bathyal depths in the region of present-day eastern Panama and the Colombian Choco. The remainder of LCA was a coastal shelf punctuated by small islands. (ii) By late Miocene the coastal shelf had shallowed considerably with a concomitant increase in emergent land. However, numerous shallow water connections remained between the Pacific Ocean and the Caribbean Sea and LCA remained an island landscape. (iii) In the Pliocene, 3.0–3.5 Mya, the isthmian terrestrial corridor linking SA and CA was completed and SA mammal fossils found in NA (Webb 1985; Lundelius 1987) and NA mammal fossils found in SA (Marshall *et al*. 1979; Marshall 1988) by 2.5–2.8 Mya bear testimony to the Pliocene movement of animals through LCA. (iv) Once completed the Central American isthmus may have been breached in central Panama during the late Pliocene as a result of eustatic sea-level changes (Cronin & Dowsett 1996).

Geological reconstructions of past landscapes are realized through a coarse-grained sampling of space and time. These reconstructions provide an evolutionary sequence of events that can be examined in the light of species distribution data and within a temporal framework estimated using molecular clocks. Thus, one aim of our research program is to use distribution patterns of freshwater fishes and gene trees to test the geological model of LCA formation.

Our analysis of the historical biogeography of LCA utilized three phylogenetically and ecologically divergent genera of neotropical primary freshwater fishes: *Roeboides*, *Pimelodella* and *Hypopomus*. Geographic evidence suggests that all three genera evolved in SA and colonized LCA from north-western Colombia (Miller 1966; Myers 1966; Bussing 1985a) and all are common stream inhabitants throughout most of Panama. The physiological similarities among these genera and their largely overlapping distributions make them suitable candidates to infer the historical biogeography of the LCA landscape. Moreover, the three genera represent three orders that are abundant and important components of neotropical diversity (Lundberg 1993).

Roeboides is a scale-eating characid with a widespread neotropical distribution. Bussing (1985a) suggested that Central American *Roeboides* contains species representing both the 'Old Southern Element' in CA, a distributional pattern that Bussing ascribes to colonization from 'the south in late Cretaceous to early Tertiary times' and the 'New Southern Element', again arriving from SA but sometime in the Pliocene. Our phylogeographic analysis of LCA *Roeboides* contained species assigned to both elements. *R. guatemalensis* represents Bussing's 'Old Southern Element' and *R. occidentalis* and *R. ilseae* (Bussing 1985b) comprise the 'New Southern Element' and are considered to be closely allied to *R. meeki* and *R.*

hildebrandi from the Atrato and San Juan rivers of the Colombian Choco region. Our study utilized mtDNA sequence data for 38 *Roeboides* specimens collected from 20 rivers extending from the Orinoco River in eastern Venezuela to the San Juan River along the northern border of Costa Rica (Table 1, Fig. 1). Geographic outgroup taxa included two *R. dayi* from two different drainage systems in Venezuela and one *R. magdelanae* from the Magdalena river in Colombia; phylogenetic outgroup taxa included *Roeboides* species from Peru and Argentina.

Hypopomus is a gymontiform that is commonly encountered in Panama streams and rivers, and is especially abundant in small tributaries. Its proclivity to bury itself diurnally may explain its tolerance of hypoxic conditions. In LCA there is a single described species (*H. occidentalis* = *H. brevirostris*), although fish inhabiting the extreme western Atlantic slope rivers are distinctly different from *Hypopomus* distributed elsewhere in Panama. Meek & Hildebrand (1916) note that the species is distributed on both slopes of Panama and throughout SA as far south as Argentina. Eigenmann (1922) noted that there is great variation within *Hypopomus* 'and it is possible that different species or varieties should be recognized.' Our preliminary studies of *Hypopomus* mtDNA phylogeography largely agree with Eigenmann's assessment. We analysed mtDNA sequence data for 16 *H. occidentalis* collected from 14 drainages extending from Venezuela to the Sixaola River which forms part of Costa Rica's Atlantic slope boundary with Panama (Table 2, Fig. 1). Geographic outgroup taxa included two *H. occidentalis* from the Maracaibo region in Venezuela; phylogenetic outgroup taxa included *Hypopomus* from Peru and Argentina.

Pimelodella is a small pimelodid catfish that is commonly encountered in schools or singly in clear streams with gravel or sandy substratum. Distinctive features of the fish include black lateral banding and the presence of poison glands at the base of the pectoral spines. *P. chagresi* is widely distributed in Panama on both the Pacific and Atlantic slopes (Meek & Hildebrand 1916; Loftin 1965) and its distribution extends into northern SA. Phylogenetic analysis of representative *P. chagresi* individuals indicates that this species is probably best considered a species complex, and within Panama there are two very divergent lineages (A. P. Martin & E. Bermingham, unpublished). Our study utilized mtDNA sequence data for 26 *P. chagresi* collected from 20 drainages extending from the Muyapa River in Venezuela to the Terraba drainage on the south-eastern Pacific slope of Costa Rica (Table 3, Fig. 1). Geographic outgroup taxa included two *P. chagresi* from Venezuela and one *P. chagresi* from the Magdalena river in Colombia; taxonomic outgroup taxa were represented

Table 1 Geographic distribution of specimens representing the characiform genus *Roeboides* and numbers of nucleotides sequenced for the mitochondrial ATPase 6,8 and partial COI genes

Species	STRI ID	Clade	Country	Area	Drainage	Locale ID	ATP	COI
R. guatemalensis	1147	rCa	Costa Rica	WPP	Coto	Rio Coloradita 1	841	593
R. guatemalensis	1148	rCa	Costa Rica	WPP	Coto	Rio Coloradita 2	841	593
R. guatemalensis	AM255	rCa	Panama	WPP	Chiriqui	Rio Platanal 1	840	564
R. guatemalensis	AM259	rCa	Panama	WPP	Chiriqui	Rio Platanal 2	841	637
R. guatemalensis	170	rCb	Panama	WPP	San Felix	Rio San Felix 1	839	474
R. guatemalensis	171	rCb	Panama	WPP	San Felix	Rio San Felix 2	839	474
*R. guatemalensis**	62	rCb	Panama	WPP	Tabasara	Rio Tabasara 1	840	472
R. guatemalensis	63	rCb	Panama	WPP	Tabasara	Rio Tabasara 2	840	472
R. occidentalis	1006	rD	Panama	CPP	La Villa	Rio La Villa 1	840	518
R. occidentalis	626	rD	Panama	CPP	San Pablo	Rio San Pablo 1	840	540
*R. occidentalis**	AM95	rD	Panama	CPP	Santa Maria	Rio Santa Maria	838	622
*R. occidentalis**	4795	rI	Panama	EPP	Caimito	Rio Caimito	842	634
R. occidentalis	583	rF	Panama	BAY	Bayano	Rio Bayano 1	840	576
*R. occidentalis**	584	rF	Panama	BAY	Bayano	Rio Bayano 2	841	576
*R. occidentalis**	AM194	rI	Panama	TUY	Tuira	Rio Pirre 1	841	628
R. occidentalis	AM205	rI	Panama	TUY	Tuira	Rio Pirre 2	840	624
R. guatemalensis	AM17	rEb	Panama	CAP	Cocle del Norte	Rio Tambo 1	841	651
R. guatemalensis	AM18	rEb	Panama	CAP	Cocle del Norte	Rio Tambo 2	841	579
R. guatemalensis	AM22	rEb	Panama	CAP	Cocle del Norte	Rio Tambo 3	841	570
*R. guatemalensis**	AM61	rEb	Panama	CAP	Cocle del Norte	Rio Juan Grande 1	841	653
R. guatemalensis	AM62	rEb	Panama	CAP	Cocle del Norte	Rio Juan Grande 2	841	640
*R. guatemalensis**	NIDN-60	rEa	Panama	CHA	Chagres	Rio Agua Salud 1	839	616
R. guatemalensis	NIDN-61	rEa	Panama	CHA	Chagres	Rio Agua Salud 2	839	453
R. guatemalensis	NIDN-66	rEa	Panama	CHA	Chagres	Rio Agua Salud 3	842	453
R. guatemalensis	NIDN-64	rEa	Panama	CHA	Chagres	Rio Agua Salud 4	842	625
R. guatemalensis	4526	rEa	Panama	CHA	Cascajal	Rio Cascajal	842	642
*R. guatemalensis**	321	rHb	Panama	WSB	Mandinga	Rio Mandinga 1	841	612
R. guatemalensis	322	rHb	Panama	WSB	Mandinga	Rio Mandinga 2	840	612
R. occidentalis	2591	rHa	Panama	WSB	Playon Chico	Rio Playon Chico 1	842	616
R. occidentalis	4898	rHa	Panama	WSB	Playon Chico	Rio Playon Chico 2	842	606
*R. occidentalis**	1740	rG	Panama	ESB	Acla	Rio Acla 1	841	587
R. occidentalis	1741	rG	Panama	ESB	Acla	Rio Acla 2	841	587
*R. meeki**	1495	rG	Colombia	ATR	Atrato	Rio Atrato 1	841	617
R. meeki	1496	rG	Colombia	ATR	Atrato	Rio Atrato 2	841	616
R. meeki	1498	rG	Colombia	ATR	Atrato	Rio Atrato 3	841	616
*R. magdalenae**	798	rOA	Colombia	MAG	Magdalena	Rio Magdalena	841	536
*R. dayi**	VZ95	rAb	Venezuela	VE	Maracaibo	Rio Muyapa	841	620
R. dayi	VZ117	rAd	Venezuela	VE	Orinoco	Rio Aguaro	841	620

*, indicates individuals selected for further analyses utilizing a single mtDNA lineage to represent each geographical area. STRI ID numbers refer to accessioned voucher specimens in the STRI fish collection.

by three *Pimelodella* species from Colombia, Guyana and Peru.

Defining mtDNA lineages and geographical areas

Phylogenies

Establishing strongly supported phylogenetic hypotheses for codistributed taxa is the most daunting obstacle facing biogeographers. This is because comparative biogeographic analysis can elucidate the pattern and process of geographical diversification only if the true tree is used to represent the phylogenetic relationships of taxa within each monophyletic evolutionary lineage (genera in our investigation). Our experimental focus is on mtDNA which, owing to a high rate of nucleotide substitution, low probablitity of intragenic recombination, and reduced effective population size relative to most nuclear loci, is a very useful tool for population-level phylogenetics (Avise 1994; Moore 1995).

The biogeographical analyses presented below utilized the complete nucleotide sequences of the slightly overlapping mitochondrial ATP synthase six and eight genes

Table 2 Geographic distribution of specimens representing the gymnotiform genus *Hypopomus* and numbers of nucleotides sequenced for the mitochondrial ATPase 6,8 and partial COI genes.

Species	STRI ID	Clade	Country	Area	Drainage	Locale ID	ATP	COI
*H. occidentalis**	100	hCa	Panama	WPP	Tabasara	Rio Tabasara	842	640
*H. occidentalis**	AM274	hCa	Panama	CPP	Cocle del Sur	Rio Zarati	842	555
*H. occidentalis**	NIDN-69	hF	Panama	EPP	Pacora	Rio Pacora	842	448
*H. occidentalis**	618	hCc	Panama	BAY	Bayano	Rio Bayano 1	842	638
H. occidentalis	621	hCc	Panama	BAY	Bayano	Rio Bayano 2	842	638
*H. occidentalis**	AM189	hCb	Panama	TUY	Tuira	Rio Pirre	842	579
H. occidentalis	3917	hCb	Panama	TUY	Tuira	Rio Ukupti	839	639
H. occidentalis	AM281	hB	Panama	WAP	Changuinola	Rio Changuinola	684	572
*H. occidentalis**	AM309	hB	Panama	WAP	Cricamola	Rio Cricamola	842	572
*H. occidentalis**	AM270	hD	Panama	CAP	Cocle del Norte	Rio Tambo	840	558
*H. occidentalis**	NIDN-30	hG	Panama	CHA	Chagres	Rio Frijolito	842	476
*H. occidentalis**	1648	hF	Panama	WSB	Mandinga	Rio Mandinga	842	588
H. occidentalis	1466	hE	Colombia	ATR	San Juan	Rio San Juan	840	643
*H. occidentalis**	1552	hE	Colombia	ATR	Atrato	Rio Atrato	842	588
*H. occidentalis**	VZ14	hA	Venezuela	VE	Maracaibo	Cano Taguara	840	647
H. occidentalis	VZ9	hA	Venezuela	VE	Maracaibo	Lago Maracaibo	838	642

*, indicates individuals selected for analyses utilizing a single mtDNA lineage to represent each geographical area. STRI ID numbers refer to accessioned voucher specimens in the STRI fish collection.

(ATPase 6,8; 842 bp) and partial cytochrome oxidase I (COI) gene sequence (mean = 558; median = 570; range = 425–651; Tables 1, 2 and 3). (MtDNA sequences have been deposited in GenBank [ATPase 6,8: AFO40388–389, AFO40391–392, AFO40394–409, AFO40411–418, AFO40421–423, AFO40425, AFO40422–457, AFO40490–527; COI: AFO40426–441, AFO40458–489, AFO40528–564, AFO40410]; primer sequences are available from E. Bermingham). Figures 2, 3 and 4 illustrate the phylogenetic hypotheses for *Roeboides*, *Hypopomus* and *Pimelodella*. For all three taxa, different methods of phylogenetic analysis yielded similar, although not identical, topologies (including weighted and unweighted parsimony analyses, results not shown). We chose to present the outcomes of the maximum likelihood analyses assuming a HKY model of evolution (Hasegawa *et al.* 1985) and LogDet analysis (Steel 1994) as implemented in PAUP* (test version 4.0d59, written by David L. Swofford). The first analysis provides distances corrected for site-to-site rate variation and the second approach corrects for nucleotide composition bias. The concordance of trees constructed using genetic distances corrected for different types of bias suggests that our analyses are not particularly sensitive to the assumptions of the different models of nucleotide substitution. For some analyses trees were pruned to include single mtDNA lineages (marked with asterisks in Tables 1, 2 and 3 and Figs 2, 3 and 4) representing each geographical area. As mtDNA clades were usually reciprocally monophyletic and unique to a drainage or a regionally cohesive group

of drainages, swapping operational taxonomic units (OTUs) with their geographical location to define area cladograms was straightforward.

Geographic areas

Two LCA landscape features are worthy of mention without reference to the genetic data. For two of the three genera, the LCA landscape terminates approximately at the border of Costa Rica and Panama. Of our test genera only *Roeboides* is widespread in Costa Rica (CR), a region we define as extending from the Barranca River on the central Pacific coast of Costa Rica to the Matina River on Costa Rica's central Atlantic slope. *Roeboides* is, however, absent from the Caribbean versant of western Panama, a distribution terminus it shares with *Pimelodella*. It is noteworthy that *Hypopomus* occurs in this western, Bocas del Toro region of Panama (WAP).

The remainder of the regions identified in Fig. 1 result from the composite view of the LCA landscape provided by our mtDNA-based phylogenetic hypotheses for *Roeboides*, *Hypopomus* and *Pimelodella* (Figs 2, 3 and 4). Our Atrato (ATR) region includes the rivers that drain the Choco region of north-western Colombia; the larger Atrato River flows to the Caribbean whereas the San Juan and the Baudo rivers are Pacific drainages. The mtDNA data provide some evidence of isolation between Caribbean and Pacific Choco drainages, but small mtDNA distances ($\approx 1\%$) would indicate that the separation is recent. From the Colombian Choco working west

Table 3 Geographic distribution of specimens representing the siluriform genus *Pimelodella* and numbers of nucleotides sequenced for the mitochondrial ATPase 6,8 and partial COI genes.

Species	STRI ID	Clade	Country	Area	Drainage	Locale ID	ATP	COI
Type A								
P. chagresi	2022	pAFa	Costa Rica	WPP	Terraba	Rio Salama Nuevo 1	838	498
P. chagresi	2024	pAFa	Costa Rica	WPP	Terraba	Rio Salama Nuevo 2	841	541
P. chagresi	1173	pAFb	Costa Rica	WPP	Coto	Rio Abrojo	841	425
P. chagresi	AM256	pAFd	Panama	WPP	Escarrea	Rio Escarrea 1	842	508
P. chagresi	AM258	pAFd	Panama	WPP	Escarrea	Rio Escarrea 2	842	460
*P. chagresi**	AM242	pAFd	Panama	WPP	San Felix	Rio San Felix	842	460
P. chagresi	1077	pAH	Panama	CPP	Tonosi	Rio Tonosi	842	466
*P. chagresi**	3172	pAH	Panama	CPP	Santa Maria	Rio Lajas	842	501
P. chagresi	AM47	pAH	Panama	CPP	Cocle del Sur	Rio Cocle del Sur	842	452
*P. chagresi**	274	pAH	Panama	EPP	Pacora	Rio Pacora	842	550
*P. chagresi**	3557	pAH	Panama	TUY	Tuira	Rio Chucunaque	842	550
P. chagresi	271	pAG	Panama	CHA	Chagres	Rio Frijoles 1	842	476
P. chagresi	272	pAG	Panama	CHA	Chagres	Rio Frijoles 2	842	452
*P. chagresi**	817	pAG	Panama	CHA	Chagres	Rio Frijolito	842	538
*P. chagresi**	1653	pAE	Panama	WSB	Mandinga	Rio Mandinga 1	842	531
P. chagresi	1654	pAE	Panama	WSB	Mandinga	Rio Mandinga 2	842	531
P. chagresi	3747	pAE	Panama	WSB	Azucar	Rio Azucar	842	537
P. chagresi	2605	pADd	Panama	ESB	Playon Chico	Rio Playon Chico	842	538
*P. chagresi**	1727	pADc	Panama	ESB	Acla	Rio Acla 1	837	470
P. chagresi	1728	pADc	Panama	ESB	Acla	Rio Acla 2	841	470
P. chagresi	1401	pADa	Colombia	ATR	Baudo	Rio Baudo	842	501
P. chagresi	877	pADb	Colombia	ATR	San Juan	Rio San Juan	839	511
*P. chagresi**	1567	pADc	Colombia	ATR	Atrato	Rio Atrato 1	841	536
*P. chagresi**	806	pAC	Colombia	MAG	Magdalena	Rio Magdalena	842	557
*P. chagresi**	VZ87	pABb	Venezuela	VE	Maracaibo	Rio Muyapa	841	538
P. chagresi	VZ139	pABc	Venezuela	VE	Aroa	Cano Caripial	840	537
Type B								
P. chagresi	AM46	pBb	Panama	CPP	Cocle del Sur	Rio Cocle del Sur	842	539
P. chagresi	AM75	pBb	Panama	CPP	Santa Maria	Rio Santa Maria	842	501
P. chagresi	3631	pBc	Panama	BAY	Bayano	Rio Ipeti	842	501
P. chagresi	AM234	pBc	Panama	TUY	Tuira	Rio Pirre 1	842	541
P. chagresi	AM235	pBc	Panama	TUY	Tuira	Rio Pirre 2	841	470
P. chagresi	1521	pBa	Colombia	ATR	Atrato	Rio Atrato 2	842	528

*, indicates individuals selected for analyses utilizing a single mtDNA lineage to represent each geographical area. STRI ID numbers refer to accessioned voucher specimens in the STRI fish collection.

along the Atlantic slope, data identify the following regions: eastern San Blas (ESB; Acla River, distinct from ATR in our analyses owing only to the absence of *Hypopomus*), western San Blas (WSB; Playon Chico and Mandinga Rivers), Chagres (CHA; Cascajal and Chagres Rivers), central Atlantic Panama (CAP; Cocle del Norte River) and western Atlantic Panama (WAP; Cana, Cricamola, Changuinola and Sixaola Rivers). Working west along the Pacific slope, the mtDNA phylogeographic data distinguish the Tuira (TUY; Tuira and Chucunaque Rivers), Bayano (BAY; Bayano and Ipeti Rivers), eastern Pacific Panama (EPP; Pacora, Caimito and Capira Rivers),

central Pacific Panama (CPP; Anton, Cocle del Sur, Santa Maria, La Villa, and San Pablo Rivers) and western Pacific Panama (WPP; Tabasara, Chiriqui Grande, Pajonal and Coto, CR Rivers). For some analyses and discussion we have subsumed the TUY, BAY, EPP and CPP regions into an area named BAHIA to recognize a dispersal pathway along the Bay of Panama during marine regressions (Loftin 1965). In summary, the phylogeographic data have subdivided the LCA landscape into 12 areas whereas prevailing taxonomy identified a single area if based on *Pimelodella* and *Hypopomus* and three to four areas based on *Roeboides*.
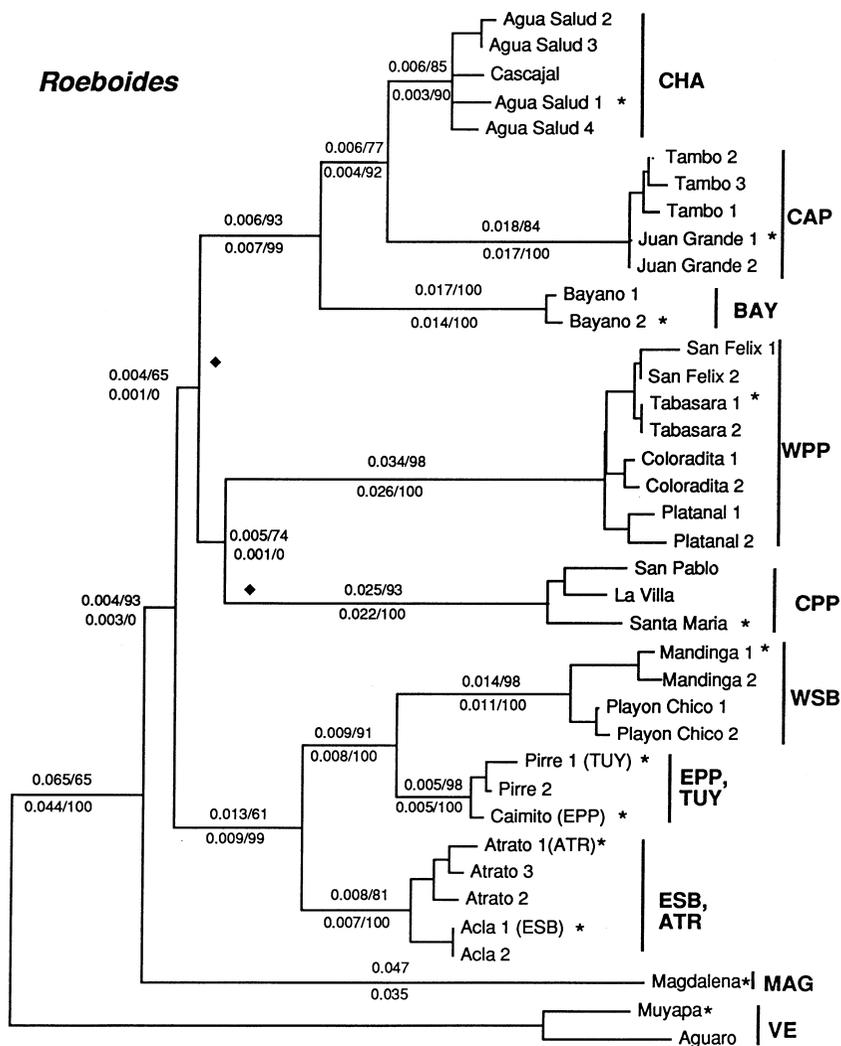
Roeboides

Fig. 2 Maximum likelihood (ML) tree for LCA collections of the characiform genus *Roeboides* based on approximately 1400 nucleotide bases representing the full mitochondrial ATP synthase six and eight (ATPase 6,8) and partial cytochrome oxidase (COI) genes. Numbers above the line are HKY distances (Hasegawa *et al.* 1985) followed by reliability indices as determined using the PUZZLE program (Strimmer & von Haeseler 1996). Numbers below the line are LogDet distances (Steel 1994; Lockhart *et al.* 1994) and bootstrap confidence levels determined using PAUP* (v.4.0.0d59). Terminal taxa marked by asterisks indicate the single individual used to represent each geographical area in COMPONENT analysis (Page 1993). The diamonds indicate a branch and its alternate placement when the PUZZLE and LogDet trees differ in topology. Area abbreviations and locations are described in the text of the proposal and are displayed in Fig. 1.

Results

Tests of shared history

Our phylogenetic analyses of fish representing three genera sampled across a broad geographical area identified distinct operational taxonomic units (OTUs), or mtDNA evolutionary lineages of relevance to the study of LCA historical biogeography. In some cases, the phylogenetic hypotheses permit a finer subdivision of the LCA landscape than presented here but sample availability centred our focus on the named geographical areas. Generally, the long branch lengths leading to terminal taxa (mtDNA clades) permitted their identification in greater than 90% of the bootstrapped trees (Figs 2, 3 and 4), yet the relationship among clades was often poorly resolved owing to very short internal branches. Overall, the gene trees for each species suggest that there are some lineages which

appear to have diverged from each other in a brief period of time, and others for which there is detectable hierarchical structure in the data.

In the case of *Pimelodella chagresi*, it was clear that there were two very distinct mtDNA taxa (differing by more than 8% mtDNA sequence divergence). In this case, the two divergent groups of lineages were treated as a distinct 'types' (A and B), and OTUs were identified within each type (Fig. 4; A. P. Martin & E. Bermingham, unpublished). *Pimelodella* type A was observed in ATR and across the breadth of *Pimelodella*'s distribution in Panama excepting BAY. *Pimelodella* type B has been found only in ATR and the BAHIA region of Panama. We have analysed 157 *Pimelodella chagresi* across LCA using a PCR–RFLP assay that distinguishes between types A and B. Our RFLP results (not shown; A. P. Martin & E. Bermingham, unpublished) permit us to conclude that
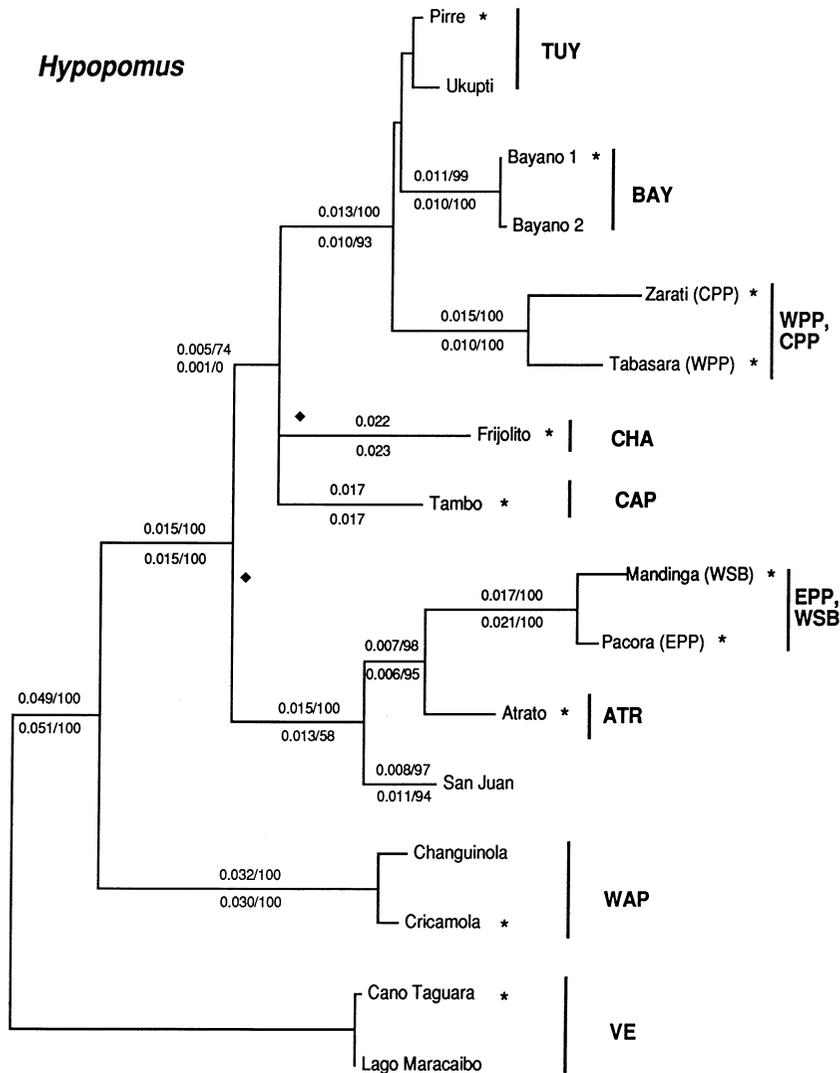
Fig. 3 Maximum likelihood (ML) tree for LCA collections of the gymnotiform genus *Hypopomus* based on approximately 1400 nucleotide bases representing the full mitochondrial ATP synthase six and eight (ATPase 6,8) and partial cytochrome oxidase (COI) genes. Conventions as in Fig. 2.

type B is almost certainly restricted to the ATR and BAHIA areas; however, we can not rule out the presence of type A in BAY owing to the relatively small numbers of *Pimelodella* analysed from that region. In ATR and BAHIA we have collected *Pimelodella* A and B types from the same drainages. An mtDNA-based phylogenetic analysis of additional *Pimelodella* species indicated that *P. chagresi* represents a paraphyletic assemblage of mtDNA lineages (results not shown; A. P. Martin & E. Bermingham, unpublished); thus in the ensuing biogeographic analyses, we used only the monophyletic *Pimelodella* type A.

With one exception, recent genetic exchanges among areas were inferred only for Pacific slope rivers draining into the Bay of Panama (our BAHIA area). The sole exception is an apparent cross-Cordillera movement of *Hypopomus* lineages involving EPP and WSB rivers that

originate in the same general area. In most cases, the posited dispersal events appear to have replaced pre-existing mtDNA lineages and thus these events effectively reset the biogeographic history of connectedness among the particular drainages.

Tests of shared history, path-length analyses

1a. If genetic distances are additive and there is shared, hierarchical structure between the different genera then we should expect to see a significant positive correlation between the genetic distance matrices. The normalized statistic Z (Mantel 1967), as implemented in NTSYS-pc (Rohlf 1993), provided the correlation coefficient between pairs of ML distance matrices based on the genus trees pruned to include a single mtDNA lineage per area. The matrix comparisons were based on individuals
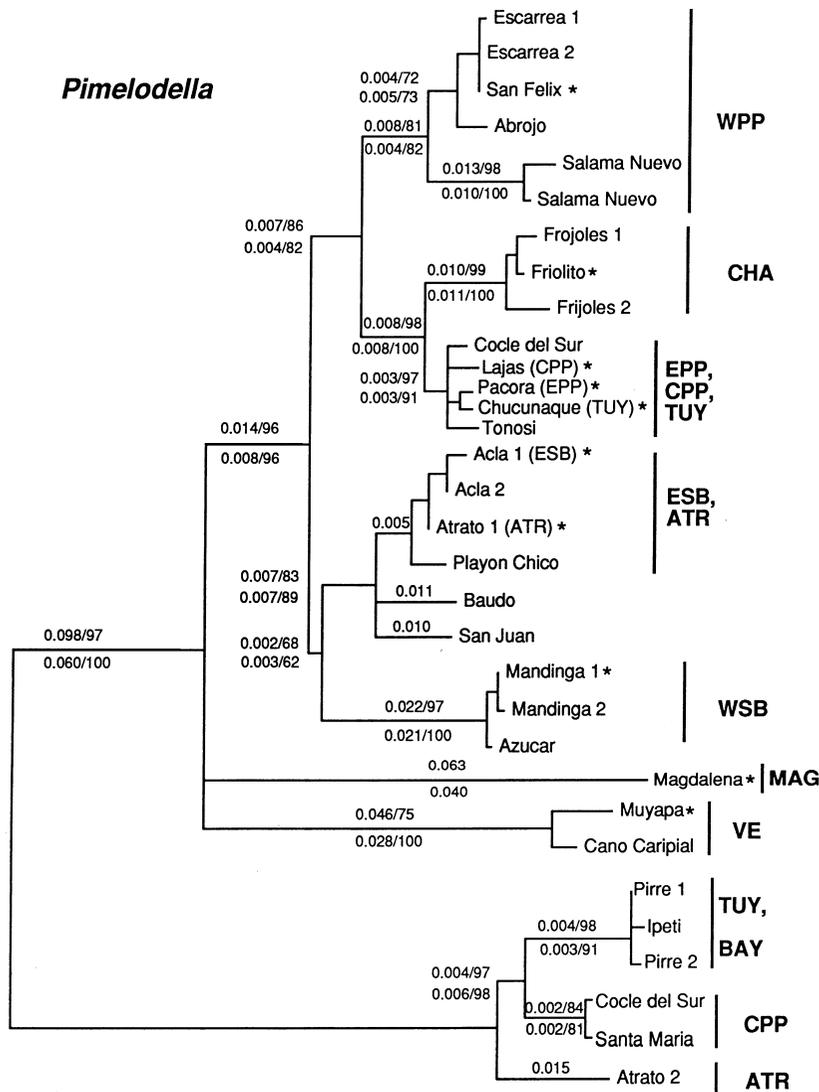
Pimelodella

Fig. 4 Maximum likelihood (ML) tree for LCA collections of the siluriform genus *Pimelodella* based on approximately 1400 nucleotide bases representing the full mitochondrial ATP synthase six and eight (ATPase 6,8) and partial cytochrome oxidase (COI) genes. Conventions as in Fig. 2.

representing the nine areas shared in common between all species (Table 4). The significance of the correlation coefficients was obtained from the table of critical values for the comparison of two independent cladograms (Table 1, Lapointe & Legendre 1992). For nine taxa, a critical value of 0.58 is significant at the 0.01 level and thus our results demonstrate shared history among the three genera tested (Table 4).

1b. To represent the relative period of temporal isolation among regions, 'mtDNA endemicity' was calculated by dividing branch tip to nearest-neighbour internode distances by tip-to-root-node distance. Branch lengths were estimated using a HKY maximum likelihood model constrained by a molecular clock assumption. This estimate provides a relative age of haplotypes within drainages (measured in terms of numbers of nucleotide substitutions accumulated). For each area we summed the

minimum branch lengths across taxa as a measure of the degree that individual areas have been isolated from other drainages (Fig. 5).

Examination of branch lengths for lineages provides an informative picture of differences in the relative degree of isolation of drainage basins. In some drainages, there are relatively long branch lengths for all three species, and

Table 4 Mantel (1967) test of the correlation between the genetic distance matrices of *Roeboides*, *Hypopomus* and *Pimelodella*

Matrix 1	Matrix 2	r	P value
Roeboides	*Hypopomus*	0.72828	P < 0.01
Roeboides	*Pimelodella*	0.74376	P < 0.01
Hypopomus	*Pimelodella*	0.75953	P < 0.01
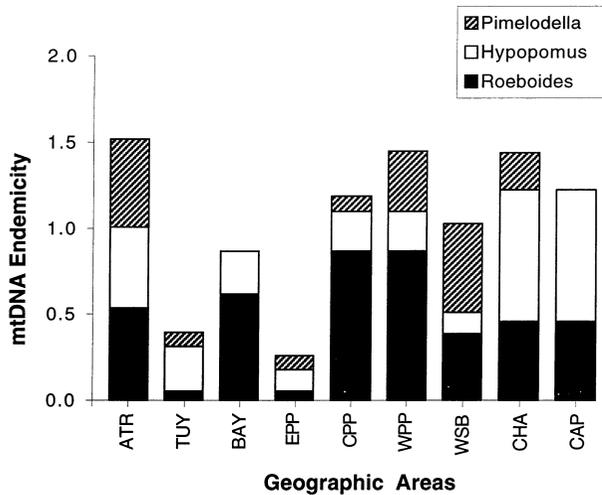
Fig. 5 Stacked histogram illustrating relative branch lengths of lineages inhabiting areas for the three species. Branch lengths are measured as proportional to the total length from the tip of the tree to the root node and were estimated assuming ultrametricity of the data. The adjusted branch lengths can be interpreted as a proxy for time, and therefore represent a measure of endemicity, assuming that endemicity is positively related to time of isolation.

this is reflected in a high composite mtDNA endemicity (e.g. ATR and WPP, Fig. 5). At the other extreme, there are areas which harbour recent lineages for all three species reflected in a relatively low composite endemicity index (e.g. TUY and EPP, Fig. 5). In a few cases, there are areas which harbour relatively ancient lineages for one taxa, and recent lineages for another taxa (e.g. CPP and CHA, Fig. 5). An a posteriori test indicates that there is a significant difference in mtDNA endemicity between the areas which drain into the Bay of Panama (BAHIA) compared to a group including all other areas (two-tailed *t*-test, $P = 0.048$). In addition, the rivers of Bocas del Toro (WAP)

harbour an endemic mtDNA lineage representing *Hypopomus*, the only one of the three genera present in WAP.

Tests of shared history, topology comparisons

2a. Kishino & Hasegawa (1989) log-likelihood ratio tests were performed for the set of three minimum length trees and each one of the data matrices. An HKY model of evolution was assumed in each case (Hasegawa *et al*. 1985). The log-likelihood test determines whether a particular tree is a significantly worse explanation of the data than an alternative tree, given a specific probabilistic model of sequence evolution. As can be determined from Table 5, alternative area cladograms were a significantly worse explanation of a particular genera's mtDNA data than was it's own area cladogram. The simple interpretation of these results is that the tests indicate a lack of shared history across genera; however, alternative subtrees tested against pruned data did return a signal of shared history. These results are not explored further here owing to their general similarity to the agreement subtree and COMPONENT results discussed below. Lapointe & Legendre (1992) indicated that topological and path-length comparisons are quite different and need not yield the same results. This is clearly the case in our study in which the log-likelihood ratio tests provided no evidence of the shared history indicated by the matrix correlations.

2b. Agreement subtrees (Fig. 6) were determined in PAUP* for the *Roeboides*, *Hypopomus* and *Pimelodella* area cladograms. Three subtrees were found and varied in the three possible pairwise combinations of sister group pairings for the Pacific slope areas WPP, CPP and BAY. All trees identified the Caribbean slope regions ATR and WSB as sister areas. The signal of historical association among areas draining the same slope of the continental divide, although not unanticipated, registered against an

Table 5 Kishino–Hasegawa likelihood ratio tests of alternative trees constructed using distance as the optimization criteria

Data	Tree	– lnL	Diff – lnL	SD (diff)	T	P
Hypopomus	*Hypopomus*	3322.90114	(best)			
	Roeboides	3503.21167	180.31053	21.98616	8.2011	< 0.0001**
	Pimelodella	3482.00564	159.10450	19.80102	8.0352	< 0.0001**
Roeboides	*Hypopomus*	3009.42252	55.92873	14.95610	3.7395	< 0.0002**
	Roeboides	2953.49380	(best)			
	Pimelodella	2990.59814	37.10	10.48591	3.5385	0.0004**
Pimelodella	*Hypopomus*	1991.12649	33.62611	13.15192	2.5567	< 0.0107**
	Roeboides	1989.96431	32.46393	13.44941	2.4138	< 0.0160**
	Pimelodella	1957.50038	(best)			

**, significantly different at $P < 0.05$. Compared for areas WPP, CPP, EPP, TUY, CHA, WSB, ATR & VE
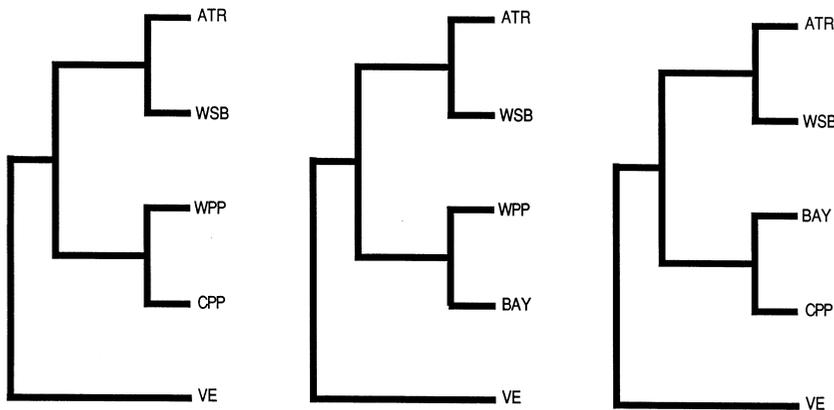
Fig. 6 The three agreement subtrees determined from analysis of *Roeboides*, *Hypopomus* and *Pimelodella* area cladograms.

evolutionary backdrop muted by finer-scale, phylogeographic differences among the genera.

2c. We used COMPONENT (Page 1993) as our principal analytical method for investigating shared history across *Roeboides*, *Hypopomus* and *Pimelodella*. The nearest-neighbour interchange (NNI) heuristics search option was utilized to infer the relationships among areas from the gene trees (Figs 2, 3 and 4) for the three genera. The minimizing criteria selected were leaves added and missing areas

Fig. 7 LCA area cladogram used to represent the historical association among areas shown in Fig. 1. The area cladogram was estimated from the mtDNA phylogenies of *Roeboides*, *Hypopomus* and *Pimelodella* (Figs 2, 3 and 4) using the NNI heuristic search option in COMPONENT (Page 1993). Owing to the primitive absence of *Roeboides* and *Pimelodella* in WAP, the basal position of that LCA area is purely a result of the ancestral position of the WAP mtDNA lineages in the *Hypopomus* gene tree. When more LCA taxa are added to our historical biogeographic analyses we anticipate that WAP will occupy a more derived position in the area cladogram as predicted by geography.

were treated as primitively absent from the distribution of each genus. NNI yielded three equally parsimonious area cladograms that differed only in the relative placement of MAG and WAP. We reasoned that MAG and WAP reversed positions (or grouped together in the third area cladogram) owing to the fact that we are missing *Hypopomus* from the Magdalena whereas *Pimelodella* and *Roeboides* are absent in WAP. Thus of the three cladograms, we utilized the one that resolved the Magdalena as basal to all areas in LCA (Fig. 7) rather than the consensus tree which left MAG and WAP as unresolved basal LCA areas. It should be noted that the position of WAP in the area cladogram owes only to its basal position in the *Hypopomus* gene tree and its presumed primitive absence from the distribution, and thus phylogenies, of *Roeboides* and *Pimelodella*.

Reconciled trees were generated using the area cladogram (Fig. 7) and the mtDNA gene tree for each genus except that each area was represented by a single mtDNA lineage (asterisked individuals in Figs 2, 3 and 4 and Tables 1, 2 and 3). The reconciled trees are presented in Fig. 8 and several general points concerning the COMPONENT analysis are summarized here. First, we compared the fit of the mtDNA gene tree for each genus to 5000 random area cladograms and significance values were determined directly from the distribution of the 5000 random trees. In each of the three comparisons, one representing each genus, there was a significantly better fit of the mtDNA tree to the area cladogram pictured in Fig. 7 than to random cladograms ($P < 0.0001$). Second, for each of the reconciled genus trees showed in Fig. 8, it was possible to improve the fit of the mtDNA trees to the area cladogram using the 'prune every leaf' option to identify the leaf that, if pruned from the tree, would best improve the fit of the mtDNA data to the area cladogram. For *Roeboides* and *Hypopomus*, the first leaf identified was one of the BAHIA lineages (*Roeboides*: BAY; *Hypopomus*: TUY) lending additional support to the proposition that dispersal

has partially overwritten the history of the areas comprising BAHIA. The fit of the *Pimelodella* data to the area cladogram was improved by pruning WPP.

The reconciled trees resulting from COMPONENT analysis provide insight into putative dispersal and extinction events across the LCA landscape. Because reconciled trees consider only branching order, here we temper the results of the COMPONENT analysis through additional reference to the gene trees which provide insight to the relative chronology of diversification. We posit an early wave of *Roeboides* colonization that must have swept through Panama leaving extant but divergent mtDNA lineages in Pacific slope drainages representing CPP and WPP. Distribution data for *Roeboides* in CA indicate that this first wave was able to penetrate as far as Tehuantepec, Mexico (Bussing 1985a; Perez *et al.* 1993). Following this event, there was either another wave of colonization from putative source populations in north-western SA, or *Roeboides* back-colonized Panama from 'refugia' populations of the western Pacific slope. Lineages derived from this wave of colonization persist today in CAP, CHA and BAY and are generally unique, suggesting a moderately long history within the drainages. Finally, there may have been a more recent dispersal event from SA that was able to colonize

drainages of eastern Panama (geographically close to putative source populations) but failed to penetrate into western Pacific slope drainages or the central Atlantic slope drainages. The absence of multiple divergent lineages of *Roeboides* in the Atrato (and other rivers of northwestern Colombia [i.e. San Juan, Baudo, Magdalena]), suggests that a single evolutionary lineage was the source for each wave of colonization, and this may explain the derived placement of the Atrato in gene trees for the species.

In *Hypopomus* there is compelling evidence displayed in both the reconciled and gene trees for an early LCA colonization event. Lineages suspected to have been derived from this event persist today in the extreme western Atlantic slope rivers (WAP). The presence of *Hypopomus* in WAP, however, suggests that these fish spread across the emergent Isthmian landscape and subsequently disappeared, leaving a gap in the distribution (not demonstrated on the reconciled tree owing to the basal placement of WAP on the area cladogram). Following the hypothetical early dispersion of lineages, there was a second *Hypopomus* wave that either emerged from SA, or conversely spread from putative refugia population in extreme western Panama (or Costa Rica). Again, we
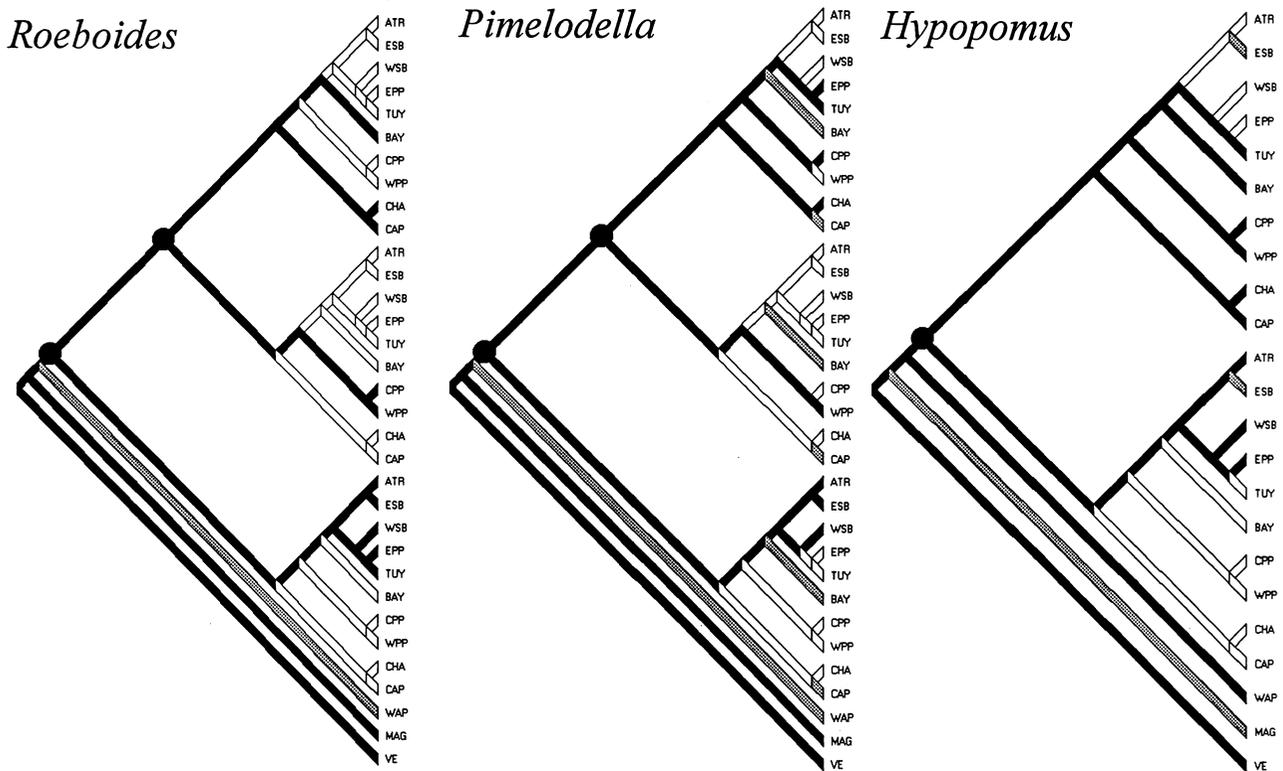
Fig. 8 Mitochondrial gene lineage trees reconciled with the best area cladogram pictured in Fig. 7. Dark lines are sampled lineages, stippled lines indicate that lineage was not sampled, white lines represent extinctions, and closed circles mark putative dispersal events.

favour the hypothesis of a second wave of colonization from SA. This second wave appears to have spread across the emergent Isthmian landscape, reaching western Panama, the remote Atlantic slope rivers and most BAHIA drainages. As in the case for *Roeboides*, many of these lineages are ancient, and may date to the late Pliocene (based on crude estimates using a molecular clock calibrated for fishes (Bermingham *et al.* 1997)). Finally, there was a third dispersion of mtDNA lineages out of north-western Colombia which spread into the near Atlantic slope region (WSB) and presumably from there across the cordillera into EPP. As we argued for *Roeboides*, the absence of multiple divergent lineages of *Hypopomus* in the Atrato (and other rivers of north-western Colombia [i.e. San Juan, Baudo, Magdalena]), suggests that a single evolutionary lineage was the source for each wave of colonization.

In *Pimelodella* there are clearly two very distinct mtDNA lineages that we recognize as distinct types. Nevertheless, the pattern of colonization and distribution of lineages for both *Pimelodella* mtDNA types considered together is remarkably similar to the other genera, except that there is no evidence for an early dispersal event. *Pimelodella* evidences a colonization episode that corresponds with the second event in the other two genera, both in terms of the penetration of lineages across the Isthmian landscape (*Pimelodella* is present in western Pacific slope rivers but its Atlantic versant distribution stops in CHA, rather than the more western CAP, as is the case for *Roeboides*) and the age of individual lineages (the branch lengths for some of these lineages are similar to branch lengths for the other two taxa). As in the other two genera, *Pimelodella* also indicates a more recent colonization episode, and again these lineages are only present in eastern Panama drainages, having failed to penetrate into the western Pacific slope rivers or the more remote Atlantic slope rivers. The *Pimelodella* gene tree clearly shows that evolutionary lineages participating in the respective colonization events were different, and both were sampled from putative source populations in north-western Colombia.

Discussion

We have described the historical relationships among the major drainage basins in LCA based on phylogenetic analysis of individual mtDNA lineages for three widespread taxa of primary freshwater fishes. Crucial to the study was the unambiguous and objective identification of mtDNA clades at levels below the currently recognized species. The mitochondrial lineages identified as OTUs formed the basis for inferring historical biogeography. Moreover, because we used the same metric (covariation of mtDNA sequences), the data allowed direct comparison across independent area cladograms and permitted estimation of the relative and absolute timing of events. Such a fine-scale inference of historical biogeography establishes a temporal and geographical framework for biotic diversification in LCA and lends itself well to what Rosen (1978) called reciprocal illumination between geological and biogeographic portraits of history. The story inferred from the fishes can fill in some detail missing from the geological record, and the geological record places the diversification of lineages within the context of large-scale events which have shaped the Isthmian landscape and thus constrain biological explanation (Grande 1985; Lundberg 1993).

Before we develop a detailed hypothesis regarding the diversification history of LCA freshwater fishes, several general results are particularly noteworthy. First, the prevailing taxonomy of *Roeboides*, *Pimelodella* and *Hypopomus* depicts a simple LCA landscape across which recent dispersal, gene flow and/or geography have provided little opportunity for diversification. Phylogeographic patterns described for the three genera challenge this view and within four of the five described species there are many well differentiated, reciprocally monophyletic mtDNA clades. Based on objective criteria, we discovered greater than 20 distinct evolutionary lineages. These OTUs formed the basis to test the hypothesis of shared history across taxa. The genetic distances between mtDNA clades also permit a rough estimate of the diversification chronology. Comparison to mtDNA ATPase and COI distances observed between species pairs of marine fishes putatively separated by the Pliocene rise of the Central American isthmus (Bermingham *et al.* 1997) would suggest that many LCA freshwater drainages harbour unique evolutionary lineages that originated at least 1–3 Mya. In a few cases, lineage age appears to slightly predate the Pliocene estimates for the emergence of the Panama isthmus (reviewed in Coates *et al.* 1992).

On average, Atlantic slope drainages appear to be more historically isolated and harbour greater mtDNA endemicity than Pacific slope drainages. The reduced endemicity of most Pacific slope drainages may result from increased dispersal potential along the coast relative to Atlantic slope drainages. Panama's two largest river systems, the Tuira and the Bayano, drain the eastern Pacific slope and enter the broad, shallow Gulf of Panama bordered to the west by Panama's Azuero peninsula and to the east by deep water off the north-western Colombia coast. Pleistocene glaciations exposed most of the Gulf of Panama (Golik 1968; Bartlett & Baghorn 1973; Fairbanks 1989) and the low shelf gradient would have facilitated fish dispersal through anastomosing lowland streams and swamps along the Pacific coast as far west as the Azuero peninsula (Loftin 1965). The positive correlation of river size

with fish population numbers and flood discharge volumes would predict a dominant role for the Tuira and Bayano basins as immigrant sources. Under this scenario it is likely that western Pacific Panama (WPP) retains its distinctive character as a result of isolation by distance and/or the effectiveness of the Azuero and Sona peninsulas as biogeographic barriers, particularly at low sea-level stands. It appears that over evolutionary time the eastern and central Pacific slope drainages (BAHIA) represent a set of populations that are intermittently severed and joined, a situation that can promote the accumulation of genetic diversity (Takahata 1993) but overwrites the earlier historical records.

A second general result is the repeated observation of short internodes joining mtDNA clades. Short internodes (and concomitant lack of strong hierarchical structure in the mtDNA data) suggest an approximately contemporaneous origin for many of the mitochondrial lineages. One explanation for this pattern is rapid dispersion of lineages across the emergent landscape relative to the rate of accumulation of substitutions which define clades. Put differently, the nucleotide substitutions which we use to infer the historical saga of LCA could not, in all cases, keep up with the spread of individuals moving across the embryonic Isthmus. Nevertheless, the rapid diversification (or speciation) for all three taxa is evidence of shared history.

In testing the hypothesis of shared history across the three genera, we discovered a single most parsimonious area cladogram that forms the basis of our working hypothesis for the historical biogeography of LCA (Fig. 7). In turn, reconciliation of the gene trees with the inferred area cladogram (Fig. 8) and interpretation of the gene trees themselves (Figs 2, 3 and 4) revealed interesting similarities and differences across our test genera (Fig. 8). Where the genera have recorded the events of history differently, the records appear to trace unique dispersal or extinction events. Therefore, for example, *Hypopomus* appears to have successfully dispersed across the continental divide so recently that some EPP individuals from the eastern Pacific slope are nearly identical to WSB individuals. By contrast, mtDNA lineages representing *Roeboides* and *Pimelodella* from EPP and WSB are genetically distinct. Despite indication of unique historical events, all three taxa have recorded relatively recent dispersal events among Pacific slope rivers which drain into the Bay of Panama. Thus, the idiosyncratic behaviour observed for each species upon close inspection disappears when the analysis focuses at a larger geographical scale.

All three reconciled trees (Fig. 7) suggest more than one invasion of lineages from putative source populations in north-western Colombia into the emergent LCA landscape. The *Hypopomus* gene tree registers the most basal LCA lineage in extreme western Panama (WAP),
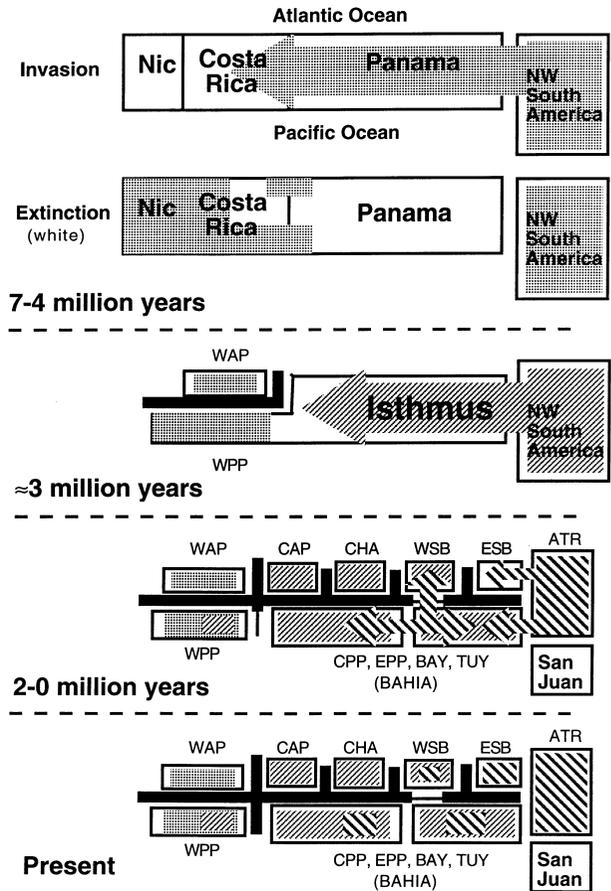
Fig. 9 Hypothetical model illustrating the history of the LCA region suggested by the analysis of mtDNA sequences from freshwater fishes. Arrows depict dispersal events and the three shading patterns represent the mtDNA phylogeographic legacy of the colonization episodes. Thick lines denote isolation of drainage areas. The approximate age of each episode is indicated. Area descriptions and codes are presented in the text and in the legend to Fig. 1.

and *Roeboides* mtDNA lineages in WPP and CPP are among the most divergent LCA genotypes observed. Rivers of western Panama are far from the putative source areas and suggest that these lineages may be relics of an early wave of colonization out of SA at the close of the Miocene (based on molecular clock estimates provided by geminate marine fishes (Bermingham *et al.* 1997)). The reconciled trees record a number of extinction events (lineage losses) erasing the tracks from this hypothetical early wave throughout much of eastern and central Panama.

We posit a second wave of colonization from putative source populations in north-western Colombia that followed closely on the heels of the postulated terminal Miocene event. Short internodes connecting many of the

lineages of the three genera suggest that dispersion from source populations occurred relatively rapidly, and the distribution of lineages suggests that dispersal was widespread reaching western Panama and the remote Atlantic slope rivers. This event probably coincided with the emergence of the Panama isthmus in the mid-Pliocene (Keigwin 1978, 1982; Coates *et al.* 1992; Coates & Obando 1996). Lineages derived from this wave of colonization persist today and are generally unique, suggesting a relatively long history within the drainages.

Finally, there is evidence for a more recent dispersal event that formed a clade consisting of eastern Panama rivers and the Choco region (ATR). Although it is possible that lineages back-colonized the Atrato from eastern Panama, we favour the hypothesis of a recent wave of colonization from north-western Colombia into Panama. Lineages which participated in this event have a limited distribution relative to the earlier waves of dispersal, suggesting that this event occurred after many of the drainage basins had become isolated by continued uplift and eventual stabilization of the landscape. In some cases, the new lineages appear to have replaced older lineages (examples in all three taxa). In other cases both old and new mtDNA lineages, such as *Pimelodella* type A and B, coexist in the same drainages. The restricted distribution and small genetic distances observed among *Pimelodella* type B lineages indicate that it is a very recent immigrant in LCA.

Our hypothesis is summarized in Fig. 9 and is qualitatively similar to the hypothesis of Central American freshwater fish colonization proposed by Bussing (1985a). However, there are significant differences in detail and the timing of the early LCA invasion by primary freshwater fish is clearly much more recent than suggested by Bussing (1985a) (see also Murphy & Collier 1996). Branch lengths for basal lineages suggest that they have been in place for $\approx$ 4–7 Myr based on comparison to divergences between geminate species of fishes separated by the Isthmus of Panama (Martin *et al.* 1992; Shulman & Bermingham 1995; Bermingham *et al.* 1997), rather than the early Tertiary estimates of Bussing for the 'Old Southern Element'. This hypothesis closely matches a record of divergence present in transisthmian, shallow-water snapping shrimp (Knowlton *et al.* 1993), suggesting that the rising isthmus explains the chronology of diversification in both marine shrimp and freshwater fishes. The putative age and observed distribution of the most divergent congeneric mtDNA lineages, particularly those observed for *Roeboides* and *Hypopomus*, coupled with the fact that our analyses focus on primary freshwater fishes, implies that there was an emergent LCA landscape prior to the geologically dated Pliocene event. In the light of the geological and climatic record, we posit a low gradient isthmian corridor that may have emerged as the Cocos Ridge began uplifting the Chorotega block coincident with dramatically reduced sea levels at the close of the Miocene. The following extreme high sea-level stand at the start of the Pliocene would have inundated the corridor and caused widescale LCA extinction of freshwater faunas and the genetic isolation of any LCA freshwater refuges (e.g. WAP). Soon thereafter, freshwater fish again colonized LCA as the Panama land-bridge emerged in the mid-Pliocene; in the time since then, more localized dispersal (e.g. stream captures) and the idiosyncratic biology of the fish has led to the complex, but interpretable, pattern of freshwater fish distribution that we observe in LCA today.

Our hypothesis indicates waves of dispersal of unique mtDNA lineages across the Isthmian landscape. The result is a complex tapestry of different evolutionary lineages (defined by mitochondrial DNA sequences). We do not know if different evolutionary lineages are able to reproduce, and thus whether there is a mixing of nuclear genes in the face of the mitochondrial lineage sorting that must have occurred in some drainages to explain the mtDNA lineage replacement registered in the phylogeographic record. Because we encountered multiple cases in which there appears to have been lineage sorting within drainages, our analyses indicate a need for improved geographical sampling of evolutionary lineages, analysis of nuclear genes, and more robust phylogenetic inference. Moreover, our hypothesis makes testable predictions that can be addressed using additional genera of neotropical fishes.

An important conclusion which can be drawn from our research is that the views of history afforded by historical biogeography have profound implications for understanding the composition of ecological communities and their stability over time (Ricklefs 1987; Cornell & Lawton 1992; Ricklefs & Schluter 1993a,b). Our results suggest that the freshwater communities inhabiting contemporary, isolated drainage basins of isthmian CA are not closed, equilibrium systems but are dynamic open systems subject to episodes of invasion and extinction. The vagaries of historical contingency are dampened in these systems, however, because the invading species and those becoming extinct are often the same. Only the genetic identity of the individuals differs.

Acknowledgements

people in Panama who permitted us to collect fish from the rivers of their respective Comarcas. They demonstrated superb local knowledge of the fish and provided us with excellent assistance in the field. Our collections from the Atrato, San Juan and Baudo drainages in the Colombian Choco would not have been possible without the assistance of Professor German Galvis from the Museo Nacional de Colombia. Professor Hernan Ortega and Fonchii Chang representing the Museo de Historia Natural, Universidad Nacional Mayor de San Marcos graciously helped us collect fishes from the Rio Manu, Peru and provided two neophytes an invaluable lesson in neotropical ichthyology. Our training was continued by Wild Bill and Myrna Bussing of the Museo de Zoologia, Universidad de Costa Rica who spent two weeks traversing and collecting in Costa Rica with us. Our thanks to Craig Moritz and an anonymous reviewer for their extensive review and suggested improvements to the manuscript. And last, but certainly not least, we would like to thank the laboratory group in Panama. Heidi Banford and Shawn McCafferty have collected throughout Panama with us. Gustavo Ybazeta and Nimia Gomez have sequenced the fruits of our collecting labours and Vijay Aswani, Shawn McCafferty and Jeff Hunt developed a database to help keep track of it all.

References

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 (1994) *Molecular Markers, Natural History and Evolution.* Chapman and Hall, New York, NY.

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

Avise JC, Giblin-Davidson C, Laerm J, Patton JC, Lansman RA (1979) Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis. Proceedings of the National Academy of Sciences USA*, **76**, 6694–6698.

Bartlett AS, Barghoorn ES (1973) Phytogeographic history of the Isthmus of Panama during the past 12 000 years, a history of vegetation, climate and sea–level change, In: *Vegetation and Vegetational History of Northern Latin America* (ed. Graham A), pp. 203–299. Elsevier, New York.

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

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

Bermingham E, Rohwer S, Freeman S, Wood C (1992) Vicariance biogeography in the Pleistocene and speciation in North American wood warblers, a test of Mengel's model. *Proceedings of the National Academy of Sciences USA*, **89**, 6624–6628.

Bermingham E, Seutin G, Ricklefs RE (1996) Regional approaches to conservation biology: RFLPs, DNA sequences, and Caribbean birds. In: *Molecular Genetic Approaches in Conservation* (eds Smith TB, Wayne RK), pp. 104–124. Oxford University Press, NY.

Brundin LZ (1988) Phylogenetic biogeography. In: *Analytical Biogeography: An Integrated Approach to the Study of Animal and Plant Distributions* (eds Myers AA, Giller PS), pp. 343–370. Chapman & Hall, London.

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 (1985a) Patterns of distribution of the Central American ichthyofauna. In: *The Great American Biotic Interchange* (eds Stehli FG, Webb SD), pp. 453–473. Plenum Press, NY.

Bussing WA (1985b) *Roeboides ilseae*, n. sp., a new scale-eating characid fish from Costa Rica. *Revista Biologica Tropical*, **33**, 45–50.

Chernoff B (1982) Character variation among populations and the analysis of biogeography. *American Zoologist*, **22**, 425–439.

Coates AG, Jackson JBC, Collins LS *et al.* (1992) Closure of the Isthmus of Panama, the near-shore marine record of Costa Rica and western Panama. *Bulletin of the Geological Society of America*, **104**, 814–828.

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

Cornell HV, Lawton JH (1992) Species interactions, local and regional processes, and limits to the richness of ecological communitites, a theoretical perspective. *Journal of Animal Ecology*, **61**, 1–12.

Croizat L, Nelson G, Rosen DE (1974) Centers of origin and related concepts. *Systematic Zoology*, **23**, 265–287.

Cronin TM, Dowsett HF (1996) Biotic and oceanographic response to the Pliocene closing of the Central American Isthmus. In: *Evolution and Environment in Tropical America* (eds Jackson J, Budd AF, Coates AG), pp. 76–104. Chicago University Press, Chicago, Ill.

Eigenmann CH (1922) The fishes of western South America. Part 1 – The fresh-water fishes of northwestern South America, including Colombia, Panama and the Pacific slopes of Ecuador and Peru, together with an appendix upon the fishes of the Rio Meta in Colombia. *Memoirs of the Carnegie Museum*, **9**, 1–350.

Fairbanks RG (1989) A 17 000-year glacio-eustatic sea level record, influences of glacial melting rates on the Younger Dryas event and deep-ocean circulation. *Nature*, **342**, 637–642.

Felsenstein J (1992) Estimating effective population size from samples of sequences: inefficiency of pairwise and segregating sites as compared to phylogenetic estimates. *Genetical Research*, **59**, 139–147.

Fu Y-X (1994) Estimating effective population size or mutation rate using the frequencies of mutations of various classes in a sample of DNA sequences. *Genetics*, **138**, 1375–1386.

Golik A (1968) History of the holocene transgression in the Gulf of Panama. *Journal of Geology*, **76**, 497–507.

Grande L (1985) The use of paleontology in systematics and biogeography and a time control refinement for historical biogeography. *Paleobiology*, **11**, 234–243.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **21**, 160–174.

Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics*, **132**, 583–589.

Humphries CJ, Parenti L (1986) *Cladistic Biogeography*. Clarendon Press, Oxford.

Joseph L, Moritz C, Hugall A (1995) Molecular support for vicariance as a source of diversity in rainforest. *Proceedings of the Royal Society of London B*, **260**, 177–182.

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

Keigwin LD (1982) Isotopic paleoceanography of the Caribbean and east Pacific, Role of Panama uplift in Late Neogene time. *Science*, **217**, 350–353.

Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *Journal of Molecular Evolution*, **29**, 170–179.

Knowlton N, Weigt LA, Solorzano LA, Mills DK, Bermingham E (1993) Divergence in proteins, mitochondrial DNA, and reproductive compatability across the Isthmus of Panama. *Science*, **260**, 1629–1632.

Lapointe FJ, Legendre P (1992) Statistical significance of the matrix correlation coefficient for comparing independent phylogenetic trees. *Systematic Biology*, **41**, 378–384.

Lockhart PJ, Steel MA, Hendy MD, Penny D (1994) Recovering evolutionary trees under a more realistic model of sequence evolution. *Molecular Biology and Evolution*, **11**, 605–612.

Loftin HG (1965) *The Geographical Distribution of Freshwater Fishes in Panama*. PhD thesis, Florida State University, FL.

Lundberg JG (1993) African–South American freshwater fish clades and continental drift, problems with a paradigm. In: *Biotic Relationships Between Africa and South America* (ed. Goldblatt P), pp. 156–198. Yale University Press, New Haven, Connecticut.

Lundelius EL (1987) The North American Quaternary sequence In: *Cenozoic Mammals of North America* (ed. Woodburne MO), pp. 211–235. University of California Press, Los Angeles, CA.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

Marshall LG (1988) Land mammals and the Great American Interchange. *American Scientist*, **76**, 380–388.

Marshall LG, Butler RF, Drake RE, Curtis GH, Tedford RH (1979) Calibration of the Great American Interchange. *Science*, **204**, 272–279.

Marshall LG, Sempere T (1993) Evolution of the Neotropical Cenozoic land mammal fauna in its geochronologic, stratigraphic, and tectonic context. In: *Biological Relationships Between Africa and South America* (ed. Goldblatt P), pp. 329–392. Yale University Press, New Haven, Connecticut.

Martin AP, Naylor GJP, Palumbi SR (1992) Rates of mitochondrial DNA evolution in sharks is slow compared with mammals. *Nature*, **357**, 153–155.

Meek SE, Hildebrand SF (1916) The fishes of the freshwaters of Panama. *Field Museum of Natural History, Zoological Series*, **10**, 217–373.

Miller RR (1966) Geographical distribution of freshwater fish fauna of Central America. *Copeia*, **1966 (4)**, 773–802.

Moore WS (1995) Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution*, **49**, 718–726.

Moritz C, Dowling T, Brown W (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Annual Review of Ecology and Systematics*, **18**, 269–292.

Murphy WJ, Collier GE (1996) Phylogenetic relatioships within the Aplocheiloid fish genus *Rivulus* (Cyprinodontiformes, Rivulidae): Implications for Caribbean and Central American biogeography. *Molecular Biology and Evolution*, **13**, 642–649.

Myers GS (1938) Fresh-water fishes and West Indian zoogeography. *Smithsonian Report for 1937*, 339–364.

Myers GS (1966) Derivation of the freshwater fish fauna of Central America. *Copeia*, **1996**, 766–773.

Nelson G, Platnick N (1981) *Systematics and Biogeography/Cladistics and Vicariance*. Columbia University Press, New York.

Page RDM (1991) Clocks, clades, cospeciation, comparing rates of evolution and timing of cospeciation events in host–parasite assemblages. *Systematic Zoology*, **40**, 188–198.

Page RDM (1993) COMPONENT, *version 2.0*. The Natural History Museum, London.

Page RDM (1996) Temporal congruence revisited: comparison of mitochondrial DNA sequence divergence in cospeciating pocket gophers and their chewing lice. *Systematic Biology*, **45**, 151–167.

Patton J, da Silva MNF, Malcolm JR (1994) Gene genealogy and differentiation among spiny arboreal rats (Rodentia: Echimyidae) of the Amazon basin: a test of the riverine barrier hypothesis. *Evolution*, **48**, 1314–1323.

Perez HE, Dillanes MTG, Mata PF (1993) *Listados Faunísticos de México III. Los Peces Dulceacuícolas Mexicanos*. Universidad Nacional Autónoma de México, México.

Pindell JL, Barrett SF (1990) Geological evolution of the Caribbean region; a plate-tectonic perspective. In: *The Caribbean Region, The Decade of North American Geology* volume H (eds Dengo G, Case JE), pp. 405–432. Geological Society of America, Boulder Colorado.

Pindell J, Dewey JF (1982) Permo-Triassic reconstruction of western Pangea and the evolution of the Gulf of Mexico/Caribbean region. *Tectonics*, **1**, 179–211.

Platnick NI, Nelson G (1978) A method of analysis for historical biogeography. *Systematic Zoology*, **27**, 1–16.

Ricklefs RE (1987) Community diversity, relative roles of local and regional processes. *Science*, **235**, 167–171.

Ricklefs RE, Schluter D (1993a) Species diversity: Regional and historical influences. In: *Species Diversity in Ecological Communities, Historical and Geological Perspectives* (eds Ricklefs RE, Schluter D), pp. 1–10. University of Chicago Press, Chicago, IL.

Ricklefs RE, Schluter D (1993b) Species diversity, an introduction to the problem. In: *Species Diversity in Ecological Communities, Historical and Geological Perspectives* (eds Ricklefs RE, Schluter D), pp. 1–10. University of Chicago Press, Chicago, IL.

Rohlf FJ (1993) *NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, Version 1.80*. Exeter Software, Setauket, New York.

Rosen DE (1976) A vicariance model of Caribbean biogeography. *Systematic Zoology*, **24**, 431–464.

Rosen DE (1978) Vicariant patterns and historical explanation in biogeography. *Systematic Zoology*, **27**, 159–188.

Rosen DE (1985) Geological hierarchies and biogeographic congruence in the Caribbean. *Annuals Missouri Botanical Gardens*, **72**, 636–659.

Savage JM (1982) The enigma of the Central American herpeto-

fauna: dispersals or vicariance? *Annuals Missouri Botanical Gardens*, **69**, 464–547.

Shulman MJ, Bermingham E (1995) Early life histories, ocean currents, and the population genetics of Caribbean reef fishes. *Evolution*, **49**, 897–910.

Slatkin M (1989) Detecting small amounts of gene flow from phylogenies of alleles. *Genetics*, **121**, 609–612.

Slatkin M, Madison WP (1989) A cladistic measure of gene flow from the phylogenies of alleles. *Genetics*, **123**, 603–613.

Steel M (1994) Recovering a tree from the Markov model. *Applied Mathematical Letters*, **7**, 19–23.

Stehli FG, Webb SD (1985) *The Great American Biotic Interchange.* Plenum Press, New York.

Strimmer K, von Haesler A (1996) Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Molecular Biology and Evolution*, **13**, 964–969.

Sykes LR, McCann WR, Kafka AL (1982) Motion of Caribbean plate during the last 7 million years and implications for earlier Cenozoic movements. *Journal of Geophysical Research*, **87**, 10656–10676.

Takahata N (1993) Allelic geneology and human evolution. *Molecular Biology and Evolution*, **10**, 2–22.

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum. Genetics*, **140**, 767–782.

Webb SD (1985) Main pathways of mammalian diversification in North America. In: *The Great American Biotic Interchange* (eds Stehli FG, Webb SD), pp. 201–217. Plenum Press, New York.