Journal of Fish Biology (2004) **64,** 186–201
doi:10.1046/j.1095-8649.2004.00299.x, available online at http://www.blackwell-synergy.com

Phylogeography and intraspecific genetic variation of prochilodontid fishes endemic to rivers of northern South America

T. F. Turner*†, M. V. McPhee*, P. Campbell*‡ and K. O. Winemiller§

*Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque, New Mexico, 87131, U.S.A. and §Department of Wildlife and Fisheries Sciences, Texas A&M University, College Station, Texas, 77843, U.S.A.

(*Received 29 April 2003, Accepted 20 October 2003*)

Phylogeography and intraspecific genetic variation were studied in prochilodontids endemic to the Orinoco, Essequibo and Amazon River basins of northern South America. Portions of two protein-encoding mitochondrial (mt) DNA genes, ND4 and COI, were examined using single-strand conformational polymorphisms (SSCPs) and nucleotide sequencing. Phylogeographic analysis indicated that the geographically widespread *Prochilodus rubrotaeniatus* is paraphyletic, with individuals from the Orinoco sharing most recent common ancestry with co-occurring *Prochilodus mariae*. A second *Prochilodus rubrotaeniatus* clade was composed of haplotypes found in the Rio Cuyuní (Essequibo Basin) and tributaries of the Rio Negro (Amazon Basin). Intraspecific genetic analysis suggested that a complex set of processes have influenced patterns of genetic variation in prochilodontid lineages. *Prochilodus rubrotaeniatus* is monomorphic at both loci in the Rio Negro and probably recently colonized this basin from the Rio Essequibo. Only two of 55 *P. mariae* exhibited variant haplotypes, and both had resulted from non-synonymous changes in the ND4 region. These observations were counter to neutral expectation and consistent with the action of natural selection on the mitochondrion. Overall, these analyses implicate vicariance, demography and selection for driving diversification of prochilodontids in northern South America.

Key words: environmental gradient; genetic bottleneck; migration; natural selection; neotropics; vicariance.

INTRODUCTION

Understanding the processes that generate biodiversity in neotropical fresh waters is becoming increasingly important as these habitats are transformed by development and as aquatic species and communities are increasingly exploited for food to support a growing human population. South America is home to the most species-rich freshwater fish fauna in the world, and yet comparatively little is known about fish diversity or the evolutionary processes

†Author to whom correspondence should be addressed. Tel.: +1 505 277 7541; fax: +1 505 277 0304; email: turnert@unm.edu

‡Present address: Department of Biology, Boston University, Boston, MA 02215, U.S.A.

that generated this diversity. For some taxonomic groups, progress has been hampered by a lack of morphological variability, which has rendered evolutionary inference difficult at some spatial scales. In such cases, molecular studies have provided insight into the roles of large-scale biogeographic processes (Lovejoy & de Araújo, 2000; Lovejoy & Collette, 2001), and population-level phenomena (Sivasundar *et al*., 2001) for shaping genetic diversity in South American freshwater fishes.

To further understand the evolutionary forces that led to present-day patterns of ichthyofaunal diversity in rivers of northern South America, a molecular study on representatives of the morphologically conservative and widely distributed characiform family Prochilodontidae was conducted. This family comprises one of the most important inland commercial fisheries on the continent (Bonetto *et al*., 1981; Bayley & Petrere, 1989; Novoa, 1989; Welcomme, 1990), and most species are important detritivores and algivores in large river ecosystems (Bowen, 1983). Prochilodontidae includes 21 species in three genera; *Ichthyoelephas, Prochilodus*, and *Semaprochilodus* and members of the family occur in every major river basin in South America (Géry, 1977). *Prochilodus* is the most widespread and species-rich genus with 13 species (Castro, 1990). Because prochilodontids are abundant, widely distributed and highly migratory throughout their life history, the family has been suggested to be an important group for elucidating major biogeographic events that have shaped ichthyofaunal diversity in South America (Castro, 1990; Sivasundar *et al*., 2001).

A previous molecular phylogeographic study of *Prochilodus* species supported the importance of the family for phylogeographic analysis at a continental scale (Sivasundar *et al*., 2001). Sivasundar *et al*. (2001) evaluated phylogeographic relationships among and within species occupying major drainages of the continent including the Rio Paraná in Brazil, western Amazonas, Orinoco, and Magdalana Basins. Sampling effort focused most intensively on *Prochilodus lineatus* (Valenciennes) in the Rio Paraná. Phylogeographic analysis of two mtDNA loci revealed substantial genetic divergence among species that occupied distinct drainages. Moreover, gene tree topology was concordant with key geological events, leading Sivasundar *et al*. (2001) to conclude that species diversity in the family was greatly influenced by geological history of the South American continent. There was little evidence of population structure within the Rio Paraná Basin suggesting that gene flow is sufficient to preclude local population divergence within continuously distributed riverine habitats, despite the enormous size of the Paraná watershed.

In this study, a similar phylogeographic approach was used to evaluate processes that influenced genetic variation in prochilodontids across drainages in northern South America. Specifically, nucleotide sequence data at two protein-encoding mitochondrial (mt) DNA genes were gathered to evaluate phylogeographic relationships among populations and species of *Prochilodus* distributed across three Atlantic slope drainage systems in north-east South America: (i) the Rio Negro, a major tributary of the Amazon, which drains rivers of southern Venezuela and Colombia, (ii) the Rio Essequibo, which drains rivers of the northern Guyana shield and flows directly into the Atlantic Ocean; and (iii) the Rio Orinoco, which drains the north-western Guyana Shield and the Andes and llanos plains of western Venezuela and Colombia and

empties into the Atlantic Ocean north-west of the Essequibo. These rivers traverse a geologically complex landscape that appears to have exerted strong influence on patterns of genetic diversity of fishes in the region (Lundberg *et al.*, 1998; Lovejoy & de Araújo, 2000).

The study focused on three endemic prochilodontid species, *Prochilodus mariae* Eigenmann, *Prochilodus rubrotaeniatus* Jardine & Schomburgk and *Semaprochilodus kneri* (Pellegrin). These species share many life-history features and have overlapping geographic distributions in northern South America (Géry, 1977). *Prochilodus mariae* is broadly distributed in white-water rivers throughout the middle and lower reaches of the Orinoco Basin (downstream of Atures rapids; Fig. 1). *Prochilodus rubrotaeniatus* occurs in black-water and clear-water rivers draining the Guyana Shield within the upper Rio Negro (Amazon Basin), upper Orinoco, Caroní (tributary of lower Orinoco) and Cuyuní (Essequibo) river basins (Fig. 1). *Semaprochilodus kneri* is widely distributed in the Orinoco Basin where it is restricted to clear-water and black-water tributaries as adults, and its geographic range broadly overlaps *P. mariae*.

Nearly all prochilodontids undergo extensive migrations throughout their life history, and the study species are no exception. For example, *P. mariae* migrate from dry-season habitats in tributary rivers and streams that receive waters from Andes to wet-season spawning and feeding habitats in channels and floodplains of the lowlands where fish form spawning aggregations, and each female apparently spawns a single batch containing tens of thousands of small eggs (Saldaña & Venables, 1983). Larvae enter productive nursery habitats in the floodplains where they feed and grow for several months. At the onset of the dry season, young fish migrate back upstream and disperse within tributaries draining the Andes. In contrast to *P. mariae*, *P. rubrotaeniatus* of the Rio Caroní migrate to the floodplain region in the river's mid-reaches to spawn. *Prochilodus rubrotaeniatus* in the Guri Reservoir (constructed for generating hydroelectric power) spawn in or near floodplains of the Rio Caroní lying upstream from the reservoir (Alvarez *et al.*, 1986). In general, this highly

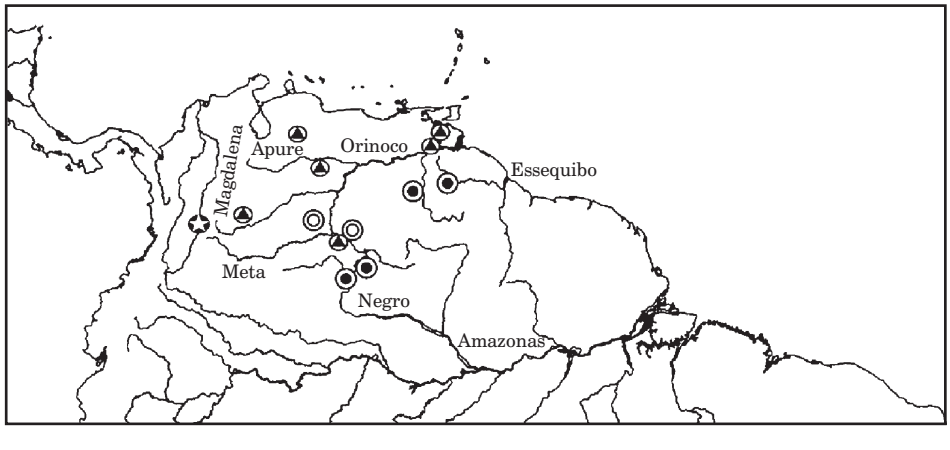

FIG. 1. Sampling localities for four species [*Prochilodus magdalenae* (✪), *Prochilodus mariae* (▲), *Prochilodus rubrotaeniatus* (◉) and *Semaprochilodus kneri* (◎)] of prochilodontids in rivers of northern South America.

migratory life history provides ample opportunity for dispersal and gene flow, and there is concomitantly little intra-basin population genetic structure expected in prochilodontids (Sivasundar *et al*., 2001).

Prochilodus mariae and *P. rubrotaeniatus* are the only members of the genus that occur in the region in eastern slope drainages of the Andes north of Amazonas. A fourth species, *Prochilodus magadalenae* Steindachner, occurs west of the Andes (the eastern Cordillera) in the Rio Magdalena basin in Colombia, and was included for purposes of estimating mutation rates for the mtDNA genes studied. Specifically, an accurately dated geological event (the final rise of the eastern Cordillera *c*. 10 million years ago; Lundberg *et al*., 1998) and genealogical methods based on coalescent theory were used to estimate key population parameters (genetic effective size N_e and mutation rate μ) for each lineage. These estimates were compared across lineages to lend additional insight into the timing and the nature of processes that have led to diversification of populations and species of prochilodontids in northern South America.

MATERIALS AND METHODS

SAMPLING AND CHARACTERIZATION OF GENETIC DIVERSITY

Caudal fin or gill arch tissues were sampled from 141 prochilodontid fishes at 13 localities in the Orinoco, Essequibo and Amazonas drainages in Venezuela, and the Rio Magdalena Basin in eastern Colombia, South America (Table I and Fig. 1). Tissues were preserved in 95% EtOH, and up to three whole voucher specimens at select localities were retained in 10% buffered formalin (Table I). Template DNA for polymerase chain reaction (PCR) was isolated using standard phenol/chloroform extraction and ethanol precipitation. Prochilodontid-specific PCR primers were developed for two protein-encoding mitochondrial (mt) DNA genes: NADH dehydrogenase subunit 4 (ND4) and cytochrome oxidase I (COI) (Table II). Single-strand conformational polymorphism analysis (SSCP) was employed to screen each individual for variability at these loci (Hayashi, 1991; Lessa & Applebaum, 1993). Representatives of each SSCP variant (haplotype) were sequenced directly (39 sequences for ND4, 21 sequences for COI) in both directions using the BIGDYE™ cycle sequencing kit and an ABI 377 automated sequencer, in accordance with the manufacturer's instructions. Fragments were assembled and aligned using the SEQUENCHER computer package for Macintosh.

PHYLOGEOGRAPHIC ANALYSIS

Phylogenetic analysis of mtDNA gene sequences was conducted using maximum parsimony (MP) and branch-and-bound searches in PAUP* (Swofford, 2001). All searches employed the following options: characters were unordered, taxa were added to starting trees at random with 10 replications per run, multi-state taxa were treated as uncertain, minimal trees were kept, tree swapping was conducted by tree bisection and reconnection, and zero length branches were collapsed. Each character set was resampled with replacement producing 1000 bootstrap replicates subjected to branch-and-bound searches. Phylogenies were generated for each fragment separately and for the two fragments combined. Tree topologies and bootstrap values were consistent across genes, thus, phylogenetic analysis of the combined data set is reported throughout the paper unless otherwise specified.

Genealogical relationships were also evaluated under a maximum-likelihood (ML) framework using PAUP*. All ML analyses employed the following options: heuristic searches with beginning trees determined by neighbour-joining, starting branch lengths

TABLE I. Species, sample sizes and sampling localities of fishes surveyed for variation in mtDNA genes. *, location not verified with GPS. MSB, Museum of Southwestern Biology, University of New Mexico; MCNG, Museo de Ciencias Naturales de Guanare, UNELLEZ, Estado Portuguesa, Venezuela

Species	*n*	Locality	Latitude and longitude	Drainage	Voucher deposition
P. mariae	10	Rio Portuguesa, VZ	09°07·50′ N; 67°45·93′ W*	Orinoco	
	10	Rio Chirere, VZ	07°47·11′ N; 67°13·87′ W	Orinoco	
	10	Puerto Ayacucho, VZ	05°39·70′ N; 67°37·68′ W*	Orinoco	
	10	Tucupita, VZ	09°05·63′ N; 62°04·68′ W*	Orinoco	
	10	Barancas, VZ	08°39·38′ N; 62°12·19′ W*	Orinoco	
	5	Rio Caparo, VZ	07°36·92′ N; 71°29·75′ W	Orinoco	MSB 48 481
P. magdalenae	4	Rio Magdalena, COL	10°43·78′ N; 73°23·33′ W*	Magdalena	
P. rubrotaeniatus	10	Rio Paragua, VZ	06°49·86′ N; 63°19·80′ W	Orinoco	MSB 48 491
	14	Rio Cuyuní, VZ	06°42·93′ N; 61°36·49′ W	Essequibo	MSB 48 492 MSB 48 493
	11	Rio Casiquiare, VZ	01°57·73′ N; 67°06·30′ W	Amazon (R. Negro)	MCNG 38 269
	5	Rio Siapa, VZ	02°04·69′ N; 66°08·44′ W*	Amazon (R. Negro)	MCNG 42 377
S. kneri	20	Rio Cinaruco, VZ	06°33·36′ N; 67°24·70′ W	Orinoco	MSB 48 486
	22	Puerto Ayacucho, VZ	05°39·70′ N; 67°37·68′ W	Orinoco	MSB 48 487

TABLE II. Primer pairs used for SSCP and nucleotide sequencing of portions of two protein-encoding mitochondrial genes, cytochrome oxidase I (COI) and NADH dehydrogenase subunit 4 (ND4), from prochilodontid fishes. PCR conditions were: reaction volume 10 μl, containing 1 μl (*c.* 50–100 ng) of sample DNA, 1·0 μl 10X reaction buffer, 200 μM each dGTP, dATP, dTTP, dCTP, 1·5 mM $MgCl_2$ and 0·1 units *Taq* DNA polymerase. PCR primer concentration was 0·8 μM with 0·1 μM of each primer end-labelled with γ^{33}-ATP. PCR consisted of 30 cycles of denaturation at 94° C for 30 s, annealing at 48° C for 30 s and extension at 72° C for 1 min

Primer	Sequence	Product length (bp)
COI–*Pma*–H	5′- cat ttc tca ata tca aac acc tt - 3′	301
COI–*Pma*–L	5′- ggc cat tat tgc tca tac tat t - 3′	
ND4–*Pma*–H	5′- gtt ctg ttt ggt tgc ctc ag - 3′	263
ND4–*Pma*–L	5′- ggg ttg aac ctc ctc taa cc - 3′	

were obtained using the Rogers-Swofford approximation method, tree swapping was conducted by tree bisection and reconnection, and zero length branches were collapsed. Maximum-likelihood model selection was conducted using likelihood-ratio testing (LRT) as implemented in the MODELTEST programme (Posada & Crandall, 1998). A second LRT was conducted to compare likelihood scores with a molecular clock enforced and without this constraint (Posada & Crandall, 2001). Failure to reject a molecular clock was interpreted as an absence of among-lineage rate heterogeneity.

The per-site mutation rate (μ) was estimated directly from the mtDNA data by determining branch lengths using the 'describe trees' feature in PAUP*. Branch lengths were evaluated for each locus separately, and then for the combined data set on the single ML tree generated from the combined data set. The expected number of mutations separating two clades was found by summing all branch lengths that intervened between clades of interest including tips. The sum of branch lengths was corrected for within-species variation in the ancestral population pool (Edwards & Beerli, 2000) using a method similar to that suggested by Nei & Li (1979). In this procedure, average branch lengths of tip clades were subtracted from the sum over all branch lengths.

COALESCENT ANALYSIS OF INTRASPECIFIC VARIATION

An analytical method based on coalescent theory of gene genealogies was employed to estimate the parameter Θ for each species and lineage (where DNA polymorphism was observed), as implemented in the computer programme MIGRATE (Beerli & Felsenstein, 2001; Beerli, 2002). For a neutrally evolving, haploid locus in a panmictic population of stable size,

$$\Theta = 2N_{e(f)}\mu, \tag{1}$$

where $N_{e(f)}$ is the long-term female genetic effective population size (Avise, 2000) and μ is the per-site mutation rate (Kingman, 1982). MIGRATE runs were set as follows: searches were initiated with a starting value of Θ based on F_{ST}, nucleotide frequencies were set to observed values, transition: transversion ratio and gamma shape parameter (α) were set to values determined by MODELTEST, and searches included 10 short chains with 1000 steps each and four long chains with 20 000 steps each. MIGRATE analysis was conducted for each species and each locus separately.

Estimates of Θ and μ were substituted into equation 1 to compute long-term, inbreeding (female) effective population size, $N_{e(f)}$. If sex-ratio is unity, then long-term N_e for the entire population is $2N_{e(f)}$. This estimation method is only valid if marker loci are evolving in a neutral fashion. The null hypothesis that mtDNA was evolving neutrally

was tested by first evaluating substitution patterns at each gene using the MEGA 2.1 software package (Kumar *et al.*, 2001), and then by testing for deviation from neutral expectation using Tajima's (1989) *D* and Fu & Li's (1993) *D** statistics as implemented in the DNASP software package (Rozas & Rozas, 1999).

RESULTS

Characterization of SSCPs and subsequent nucleotide sequencing of variable haplotypes revealed substantial variation in both genes when compared across four prochilodontid species. For COI, 43 out of 301 sites were variable and 12 were parsimony informative. Slightly more variation was revealed at ND4, where 48 out of 263 sites were variable and 43 were parsimony informative. Most substitutions were at two- and four-fold degenerate sites and were synonymous. A single non-synonymous substitution was identified in COI, which resulted in a polar (threonine) amino acid shared by *P. magdalenae*, *P. mariae* and *P. rubrotaeniatus* (Orinoco Basin only) and a nonpolar (isoleucine) amino acid present in all other species and localities. Of four non-synonymous substitutions observed at ND4, two resulted in changes that affected polarity. A change from a nonpolar (proline) to polar (serine) amino acid was identified in a single *P. mariae* haplotype in the Rio Chirere, and a second change resulted in a single *P. magdalenae* haplotype possessing a nonpolar (alanine) rather than a polar (threonine) amino acid. The alignment contained no insertions or deletions, and no stop codons were identified in the data set. Nucleotide sequences were deposited in Genbank under accession number AY339 827–AY339 840 for ND4 and AY339 841–AY339 846 for COI, respectively.

Phylogenetic analysis of the combined data set resulted in a single most-parsimonious tree (Fig. 2) that provided strong support (bootstrap values >70%) for reciprocal monophyly of *P. magdalenae* and all other *Prochilodus* sampled. *Prochilodus rubrotaeniatus* was paraphyletic, with one clade composed of two divergent haplotypes from tributaries of the Essequibo, and a single haplotype identified in tributaries of the Rio Negro (Amazon) nested within the Essequibo clade (Fig. 2). The sister group to this clade included *P. rubrotaeniatus* from the Rio Paragua and the white-water dwelling *P. mariae* (Fig. 2).

Model selection procedures indicated that the DNA substitution model of Hasegawa *et al.* (1985) with gamma distributed among-site rate heterogeneity (*i.e.* HKY + G) best explained nucleotide sequence evolution of prochilodontid mtDNA genes. The transition: transversion ratio was estimated to be 10·5612, the gamma shape parameter was 0·0167 and nucleotide frequencies were A = 0·2761, C = 0·2974, G = 0·1411 and T = 0·2854. The ML tree topology generated under this model was identical to the MP topology (Fig. 2). Likelihood-ratio testing indicated that the molecular clock could not be rejected for either locus separately (COI, d.f. = 9, $P = 0{\cdot}98$; ND4, d.f. = 9, $P = 0{\cdot}83$) or for the combined data set (d.f. = 9, $P = 0{\cdot}85$), indicating little or no mutational rate heterogeneity among lineages of genus *Prochilodus*.

Evaluation of corrected branch lengths based on the ML tree under a HKY + G model indicated that the ND4 region was evolving at least twice as fast as the COI region. For ND4, *P. magdalenae* was 10·39% divergent from other *Prochilodus* species, whereas COI was 5·39% divergent. Taking both loci

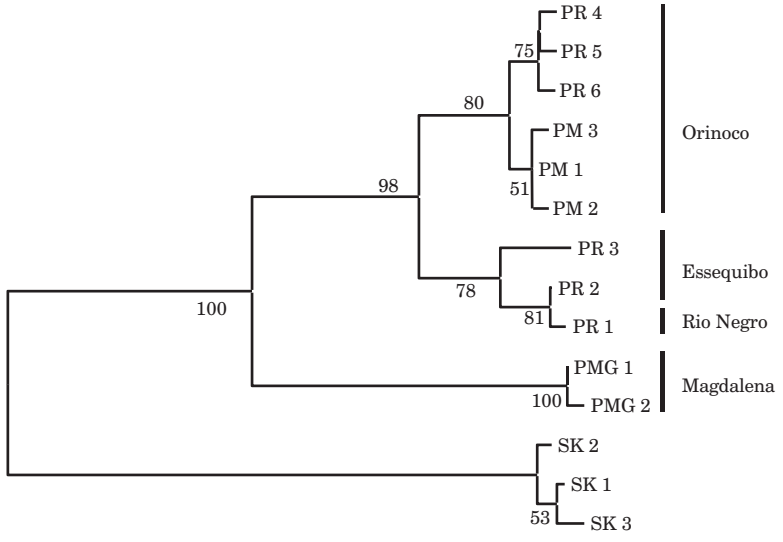

FIG. 2. The single most-parsimonious (MP) tree generated using combined COI and ND4 nucleotide sequence data (564 bp). Bootstrap percentages (from 1000 replicates) >50% are reported on branches. Branch lengths are proportional to the number of substitutions. OTUs correspond to unique haplotypes identified in *Semaprochilodus kneri* (SK), *Prochilodus magdalenae* (PMG), *Prochilodus mariae* (PM) and *Prochilodus rubrotaeniatus* (PR). Sample drainages for the haplotypes are indicated.

together, *P. magdalenae* was 7·12% divergent from other *Prochilodus* species. Nucleotide divergence between *P. magdalenae* and the remainder of *Prochilodus* species implies that the average per-site mutation rates (μ) of these mtDNA protein encoding genes is *c.* $0{\cdot}52 \times 10^{-8}$ for the ND4 region, $0{\cdot}27 \times 10^{-8}$ for COI region and $0{\cdot}36 \times 10^{-8}$ for both genes, provided that the final rise of the eastern Cordillera was *c.* 10 million years ago.

Intraspecific patterns of genetic diversity differed considerably among lineages. For example, *P. mariae* exhibited no variation in the COI region, and 52 out of 55 individuals sampled at six geographically widespread localities in the Orinoco possessed haplotype PM1 (Table III). Two rare haplotypes (PM2 and PM3; Fig. 2) were identified at Rio Chirere, and both resulted from single, non-synonymous mutations in the ND4 region, one resulting in a change in polarity (haplotype PM3). Six unique haplotypes were identified in *P. rubrotaeniatus* and none were shared across basins (Table III). In the Rio Paragua, *P. rubrotaeniatus* exhibited three unique haplotypes (PR4, PR5 and PR6) that differed by two mutational steps from one another, with all of the variation occurring in the ND4 region. The Rio Cuyuní in the Essequibo Basin possessed two haplotypes (PR2 and PR3; Fig. 2.) that were separated by four mutational steps. Haplotype PR1 was found in all fishes in the Rio Negro and differed from PR2 by a single mutational step in the ND4 region. *Semaprochilodus kneri* possessed three haplotypes (SK1, SK2 and SK3) that were present at relatively high frequencies in both sampling localities in the Orinoco Basin. Haplotype (*h*) and nucleotide diversities (Θ) ranged from zero to 0·60 and from zero to 0·005, respectively (Table IV). Substitution of Θ and locus-specific μ into equation 1

TABLE III. Distribution of composite (ND4 and COI sequences combined) haplotypes tabulated across prochilodontid species and sampling localities in northern South America. Haplotype designations correspond to Table I and Fig. 2

		Composite haplotype										
Species	Locality	PM1	PM2	PM3	PR1	PR2	PR3	PR4	PR5	PR6	SK1	SK2	SK3
P. mariae	Rio Portuguesa	10											
	Rio Chirere	8	1	1									
	Pto. Ayacucho	10											
	Tucupita	10											
	Barancas	10											
	Rio Caparo	5											
P. rubrotaeniatus	Rio Paragua							7	2	1			
	Rio Cuyuní					1	11						
	Caño–Tumereno					1	2						
	Rio Casiquiare				10								
	Rio Siapa				5								
											3	10	4
S. kneri	Rio Cinaruco												
	Pto. Ayacucho										4	15	2

TABLE IV. Summary statistics for mtDNA protein-encoding genes by species, population and mtDNA gene locus. The total number of individuals sampled (n), the number of unique haplotypes identified (i), haplotype diversity (h), estimates of parameter Θ from MIGRATE (Beerli, 2002) and long-term genetic effective population size (N_e) are given. Per-site mutation rates for estimation of long-term N_e were $\mu_{(ND4)} = 0{\cdot}52 \times 10^{-8}$ and $\mu_{(COI)} = 0{\cdot}27 \times 10^{-8}$ (see text)

		ND4					COI				
Species	Locality	n	i	h	Θ	N_e	n	i	h	Θ	N_e
P. mariae	Rio Portuguesa	10	1	0·000	0·000	—	10	1	0·000	0·000	—
	Rio Chirere	10	3	0·378	0·00159	305 770	10	1	0·000	0·000	—
	Pto. Ayacucho	10	1	0·000	0·000	—	10	1	0·000	0·000	—
	Tucupita	10	1	0·000	0·000	—	10	1	0·000	0·000	—
	Rio Caparo	5	1	0·000	0·000	—	5	1	0·000	0·000	—
P. rubrotaeniatus	Rio Paragua	10	3	0·511	0·00368	707 692	10	1	0·000	0·000	—
	Rio Cuyuní	14	2	0·264	0·00106	203 846	14	2	0·264	0·00032	118 518
	Rio Casiquiare	11	1	0·000	0·000	—	11	1	0·000	0·000	—
	Rio Siapa	5	1	0·000	0·000	—	5	1	0·000	0·000	—
S. kneri	Rio Cinaruco	17	3	0·603	0·00149	286 538	19	1	0·000	0·000	—
	Pto. Ayacucho	21	3	0·467	0·00505	971 154	21	1	0·000	0·000	—

yielded estimates of the long-term (female) genetic effective population size, $N_{e(f)}$ for each locality (Table IV).

Tajima's (1989) and Fu & Li's (1993) tests for deviation from neutral expectation were not significant at the nominal $\alpha = 0{\cdot}05$ level for *P. rubrotaeniatus* at any locality where genetic variation was observed [range Tajima's (1989) $D = -0{\cdot}56$ and $-0{\cdot}12$; Fu & Li's (1993) $D^* = 0{\cdot}17$ to $0{\cdot}94$, in the Orinoco and Essequibo clades, respectively]. Values for *S. kneri* were not significant for both tests [Tajima's (1989) $D = 1{\cdot}50$; Fu & Li's (1993) $D^* = 1{\cdot}03$]. *Prochilodus mariae* exhibited a significant result for Fu and Li's (1993) test ($D^* = -2{\cdot}57$, $P < 0{\cdot}05$) when variation in the ND4 region was tested, but Tajima's (1989) D statistic ($= -1{\cdot}45$) was not significant.

DISCUSSION

PHYLOGEOGRAPHY OF PROCHILODUS IN NORTHERN SOUTH AMERICA

Previous morphological (Castro, 1990, 1993) and molecular studies (Sivasundar *et al.*, 2001) have indicated a major role for geographic isolation of major South American river basins and structuring diversity and promoting speciation in prochilodontid fishes. Analysis of both data types revealed evidence for reciprocally monophyletic groups (usually distinct species) restricted to single drainage systems, and little evidence of genetic structure within drainages, with a few notable exceptions (*e.g. Prochilodus britskii* Castro, 1993).

The present phylogeographic analysis supports the idea that geographic isolation of lineages across major drainages is an important part of the speciation process in these abundant and highly migratory fish species. The deep split of *P. magdalenae* is consistent with an early vicariant separation following the final rise of the eastern Cordillera. Reciprocally monophyletic groups within *P. rubrotaeniatus*, one restricted to the Orinoco Basin (that shared recent common ancestry with *P. mariae*), and the other in the Essequibo and Amazon Basins were observed. Taken together, these results are consistent with the following hypothesis of river relationships: {Magdalena, [Orinoco, (Essequibo, Amazonas)]}. This agrees closely with the hypothesis proposed by Sivasundar *et al.* (2001) based on a similar analysis of protein- (ATPase 6 & 8) and non-protein-encoding (D-loop) mtDNA genes for these and other prochilodontid species. Observed paraphyly of *P. rubrotaeniatus* warrants further study and could indicate the presence of at least two cryptic species within this taxon.

If evolution under a molecular clock is assumed and the uplift of the eastern Cordillera is dated correctly, then it is possible to provide an approximate estimate of the timing of other coalescent events indicated by the mtDNA genealogy. Comparison of branch lengths (corrected for intralineage variation as above) indicated that the separation of the Orinoco and Essequibo plus Rio Negro clades occurred *c.* 3·0 million years ago. *Prochilodus rubrotaeniatus* and *P. mariae* in the Orinoco shared a common ancestor *c.* 800 000 years ago, and haplotypes from the Rio Negro and Essequibo *P. rubrotaeniatus* clades shared common ancestry probably <100 000 years ago.

The finding of a relatively recent connection between the Orinoco and Amazon Basins through the Essequibo is not surprising, considering that this migration pathway was proposed to explain genetic diversity in similarly distributed freshwater needlefishes (Lovejoy & de Araújo, 2000). Interestingly, Lovejoy & de Araújo (2000) reported an average p-distance of 0·035 between cytochrome-b haplotypes sampled from freshwater needlefishes from the lower Orinoco and the lower Amazon. This value is only slightly larger than the corrected sequence divergence for *P. rubrotaeniatus* between the Orinoco and Amazon (0·022), but is 40-fold larger than average divergence between the Essequibo and Amazon (0·0009). Thus, the connection between the Essequibo and Amazon appears to have continued up until the recent past. An alternative hypothesis that was not supported by the present data is a connection between the Orinoco and Amazon through the Rio Casiquiare. At present, the Casiquiare captures part of the upper Orinoco and connects to the Amazon through the Rio Negro. If this connection provided an avenue for gene flow in *P. rubrotaeniatus*, then the Orinoco and Rio Negro lineages would be expected to share haplotypes in common, or at least exhibit evidence of most recent common ancestry with respect to the Essequibo clade (Slatkin & Maddison, 1989).

INTRASPECIFIC GENETIC VARIATION

Under a neutral model of molecular evolution where equilibrium between mutation and genetic drift has been established, the level of genetic diversity in a panmictic population is expected to be proportional to the product of the genetic effective population size (provided it is constant through time) and the mutation rate (Kingman, 1982). Because prochilodontids usually exhibit large adult numbers and migratory life histories, high genetic diversity and very large genetic effective population sizes were expected for each species and each lineage surveyed. This is because large genetic effective size and high migration rates minimize the effect of genetic drift as a force that lowers intra-lineage genetic diversity. Results of intraspecific analysis of *P. rubrotaeniatus* (Orinoco lineage) and *S. kneri* were consistent with expectations of diversity at a neutral locus evolving in large, stable and highly migratory populations. Estimates of N_e based on ND4 data ranged from 10^5 to 10^6, which are consistent with estimates obtained from coalescent analysis of mtDNA coding regions surveyed in *P. lineatus* (ATPase 6 & 8 data from Sivasundar *et al*., 2001).

Intraspecific analysis revealed extremely low levels of haplotype diversity for *P. mariae* across its geographic range in the Orinoco and for the Rio Negro lineage of *P. rubrotaeniatus*. Examination of parameters and assumptions of the neutral model suggests three factors that could, theoretically, account for low genetic diversity in these lineages: (1) mutation rates are very low such that insufficient time has passed to accumulate nucleotide substitutions among haplotypes; (2) census numbers have fluctuated dramatically over time and historical genetic bottlenecks have resulted in depletion of genetic variation in some lineages (Grant & Bowen, 1998); (3) the study loci are the direct targets of natural selection or are linked to loci under selection, which has operated to reduce genetic diversity in some lineages (Wayne & Simonsen, 1998). For

purposes of discussion, each of these factors is discussed below in turn, but they are not necessarily mutually exclusive.

Mitochondrial protein-encoding loci appear to evolve slowly in prochilodontids (*c.* 0·5% per million years; Sivasundar *et al.*, 2001; this study) compared to other fish taxa (1·3% per million years; Bermingham *et al.*, 1997). Differences in mutation rates between the ND4 and COI regions have apparently exerted influence on levels of variation within lineages of prochilodontid fishes. The COI was variable only in a single lineage (*P. rubrotaeniatus*, Rio Essequibo) where Θ was three times smaller than that obtained for ND4. It was invariant in other lineages despite moderate levels of variability in the ND4 region. Thus, these observations suggest that low levels of variation in COI have resulted from a low mutational rate at this locus. Because COI and ND4 are tightly linked in the mitochondrial genome, it is unlikely that one locus could be the target of natural selection independent of the other locus.

Is it possible that mutation rates at ND4 vary among lineages of prochilodontid fishes, and that lineage specific rate heterogeneity explains differences in patterns of variability? In order to explain low diversity, it is necessary to postulate that mutational rates vary among lineages of prochilodontids, and more specifically, that rates have slowed in *P. mariae* and *P. rubrotaeniatus* (Rio Negro). Likelihood-ratio testing failed to reject a molecular clock for either locus, which implies that mutation rates have not slowed in these lineages with respect to other prochilodontids. Thus, lineage specific rate heterogeneity is not a likely explanation for low genetic diversity observed in these two lineages.

Low levels of genetic diversity within lineages may also be explained by historical fluctuations in population size. In this situation, a population is reduced to a small number at some time in the past such that genetic drift removes genetic variation much faster than it arises by mutation (Avise, 2000). This situation is consistent with the pattern of phylogeographic and intraspecific variation observed in *P. rubrotaeniatus* in the Rio Negro. The haplotype observed in the Rio Negro basin differs by a single substitution from a common Essequibo haplotype, suggesting close affinity of the Rio Negro and Essequibo lineages. Based on this observation, a demographic bottleneck resulting from a recent founding event (within the last 100 000 years) probably best explains the distribution of genetic variation observed in *P. rubrotaeniatus* in the Rio Negro. An important caveat to this interpretation is that purifying natural selection as a force that limits gene diversity cannot be ruled out. Because ND4 was invariant in this lineage, it was not possible to test for deviation from neutral expectation. In principle, it is possible to distinguish a bottleneck from the action of purifying selection by examining patterns of genetic variation at an unlinked (*i.e.* nuclear) locus. Genetic drift should act uniformly across loci, which leads to the prediction that an unlinked locus should also exhibit low levels of genetic diversity in the Rio Negro if the founder-event situation is correct.

There is weak evidence that natural selection may have shaped mtDNA diversity in *P. mariae*. Very low genetic diversity and significantly negative Fu and Li's (1993) D^* were observed in this species, which together support an interpretation of purifying selection acting on the mitochondrion (Przeworski *et al.*, 1999). If selection is truly operating on the mitochondrion, it may be

associated with substantial environmental differences between habitats occupied by *P. rubrotaeniatus* and *P. mariae* in the Orinoco Basin. *Prochilodus rubrotaeniatus* is restricted to the Rio Caroní, a major black-water tributary of the Orinoco. Black-waters are highly acidic (pH 4·5–6·0) from leached humic and tannic compounds and are generally nutrient poor. *Prochilodus mariae* does not occur in the Caroní system but rather is restricted to the middle and lower Orinoco and east flowing tributaries, which are nutrient-rich, white-water systems characterized by high conductivity, neutral pH and frequently high suspended sediment loads. White-water habitats may impose purifying selection on the mitochondrion, which would imply that the observed rare, non-synonymous, variants are deleterious. This proposal is tentative for at least three reasons. First, there are very few total mutations within and among lineages of *Prochilodus*, which severely limits statistical power of tests for deviation from neutrality (Wayne & Simonsen, 1998). Second, statistical rejection of neutrality is not the same thing as demonstrating selection (Wayne & Simonsen, 1998). Third, it cannot be established unequivocally that the ancestor of *P. mariae* was a black-water form.

In conclusion, phylogeographic analysis indicated that vicariant processes associated with separation of major drainage basins over that last 10 million years played a critical role for diversification of prochilodontids over broad geographic regions. Inclusion of intraspecific analysis based on the coalescent theory of gene genealogies, however, revealed that demographic and possibly selective forces may also be important in some prochilodontid lineages (Moritz *et al*., 2000; Smith *et al*., 2001). The combination of phylogeography with robust coalescent methods offers the exciting prospect of uncovering details about the forces that shape genetic and ultimately ichthyofaunal diversity in neotropical fresh waters.

Institutional and field support was provided by D. Taphorn (Univ. Exper. Llanos Occidentales, Venezuela), A. Barbarino-Duque and N. Ortiz (Inst. Nac. Invest. Agropequarias, Venezuela), H. Lopéz, C. Layman and J. Walther (Texas A & M Univ.), and H. Julio (Univ. Est. Maringá, Brazil). G. Rosenberg and the Molecular Biology Facility at the University of New Mexico provided technical and laboratory support. H. Lopéz, G. Moyer and J. Wares provided helpful suggestions on the manuscript. *Prochilodus magdalenae* samples and helpful suggestions were provided by G. Ortí (University of Nebraska). Curatorial services were provided by A. Snyder in the Museum of Southwestern Biology. Funding was provided by the National Geographic Society. Specimens were collected under permit 0305 issued by the Ministerio de Ambiente y Recursos Naturales Renovables, Venezuela.

References

Alvarez, E., Balbas, L., Massa, I. & Pacheco, J. (1986). Aspectos ecologicos del embalse Guri. *Interciencia* **11,** 325–333.

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

Bayley, P. B. & Petrere, M. (1989). Amazon fisheries: assessment of methods, status, and management options. In *Proceedings of the International Large Rivers Symposium* (Dodge, D. P., ed.). *Canadian Special Publications in Fisheries and Aquatic Sciences* **106,** 385–398.

Beerli, P. & Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences USA* **98,** 4563–4568.
Bermingham, E., McCafferty, S. & Martin, A. (1997). Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. In *Molecular Systematics of Fishes* (Kocher, T. & Stepien, C., eds), pp. 113–126. New York: Academic Press.
Bonetto, A. A., Cerón, M. C. & Roldán, D. (1981). Nuevos aportes al conocimiento de las migraciones de peces en el Río Paraná. *Ecosur* **8,** 29–40.
Bowen, S. H. (1983). Detritivory in neotropical fish communities. *Environmental Biology of Fishes* **9,** 137–144.
Castro, R. M. C. (1990). Revisao taxonomica da familia Prochilodontidae (Ostariophysi, Characiformes). Doctoral Thesis, Universidade de Sao Paulo, Brazil.
Castro, R. M. C. (1993). *Prochilodus britskii*, a new species of prochilodontid fish (Ostariophysi: Characiformes), from the Rio Apiaca, Rio Tapajos system, Mato Grosso, Brazil. *Proceedings of the Biological Society of Washington* **106,** 57–62.
Edwards, S. V. & Beerli, P. (2000). Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* **54,** 1839–1854.
Fu, Y. X. & Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics* **133,** 693–709.
Géry, J. (1977). *Characoids of the World*. New Jersey: TFH Publications, Inc.
Grant, W. S. & Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity* **89,** 415–426.
Hasegawa, M., Kishino, H. & Yano, T. A. (1985). Dating of the human ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22,** 160–174.
Hayashi, K. (1991). PCR-SSCP: A simple and sensitive method for detection of mutations in genomic DNA. *PCR Methods and Applications* **1,** 34–38.
Kingman, J. F. C. (1982). The coalescent. *Stochastic Processes and their Applications* **13,** 235–248.
Kumar, S., Tamura, K., Jakobsen, I. B. & Nei, M. (2001). *MEGA2: Molecular Evolutionary Genetics Analysis Software*. Tempe, AZ: Arizona State University.
Lessa, E. P. & Applebaum, G. (1993). Screening techniques for detecting allelic variation in DNA sequences. *Molecular Ecology* **2,** 119–129.
Lovejoy, N. R. & de Araújo M. L. G. (2000). Molecular systematics, biogeography and population structure of neotropical freshwater needlefishes of the genus *Potamorrhaphis*. *Molecular Ecology* **9,** 259–268.
Lovejoy, N. R. & Collette, B. B. (2001). Phylogenetic relationships of new world needlefishes (Teleostei: Belonidae) and the biogeography of transitions between marine and freshwater habitats. *Copeia* **2001,** 324–338.
Lundberg, J. G., Marshall, L. G., Guerrero, J., Horton, B., Malabarba, M. C. S. L. & Wesselingh F. (1998). The stage for neotropical fish diversification: A history of tropical South American rivers. In *Phylogeny and Classification of Neotropical Fishes* (Malabarba, L. R., Reis, R. E., Vari, R. P., Lucena, Z. M. S. & Lucena, C. A. S., eds), pp. 13–48. Porto Alegre: Edipucrs.
Moritz, C., Patton, J. L., Schneider, C. J. & Smith, T. B. (2000). Diversification of rainforest faunas: An integrated molecular approach. *Annual Review of Ecology and Systematics* **31,** 533–563.
Nei, M. & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences USA* **76,** 5269–5273.
Novoa, D. F. (1989). The multispecies fisheries of the Orinoco river: development, present status, and management strategies. In *Proceedings of the International Large Rivers Symposium* (Dodge D. P., ed.). *Canadian Special Publications in Fisheries and Aquatic Sciences* **106,** 422–428.
Posada, D. & Crandall K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14,** 817–818.

Posada, D. & Crandall, K. A. (2001). Selecting models of nucleotide substitution: An application to human immunodeficiency virus 1 (HIV-1). *Molecular Biology and Evolution* **18,** 897–906.
Przeworski, M., Charlesworth, B. & Wall, J. D. (1999). Genealogies and weak purifying selection. *Molecular Biology and Evolution* **16,** 246–252.
Rozas, J. & Rozas, R. (1999). DnaSP version 3.1: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15,** 174–175.
Saldaña, J. & Venables, B. (1983). Energy compartmentalization in a migratory fish, *Prochilodus mariae* (Prochilodontidae), of the Orinoco River. *Copeia* **1983,** 617–623.
Sivasundar, A., Bermingham, E. & Orti, G. (2001). Population structure and biogeography of migratory freshwater fishes (*Prochilodus*: Characiformes) in major South American rivers. *Molecular Ecology* **10,** 407–417.
Slatkin, M. & Maddison, W. P. (1989). A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* **123,** 603–613.
Smith, T. B., Schneider, C. J. & Holder, K. (2001). Refugial isolation *v.* ecological gradients. *Genetica* **112–113,** 383–398.
Swofford, D. L. (2001). *Paup*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.1*. Sunderland, MA: Sinauer Associates.
Tajima, F. (1989). Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123,** 585–595.
Wayne, M. L. & Simonsen, K. L. (1998). Statistical tests of neutrality in the age of weak selection. *Trends in Ecology and Evolution* **13,** 236–240.
Welcomme, R. L. (1990). Status of fisheries in South American rivers. *Interciencia* **15,** 337–345.

Electronic Reference

Beerli, P. (2002). *MIGRATE: Documentation and Program, part of LAMARC*. Version 1.6 available via world wide web at URL = http://evolution.genetics.washington.edu/lamarc.html