Journal of Zoology

Journal of Zoology. Print ISSN 0952-8369

Population history of the Amazonian one-lined pencilfish based on intron DNA data

M. J. Sistrom[1]*, N. L. Chao[2] & L. B. Beheregaray[1]

1 Molecular Ecology Laboratory, Department of Biological Sciences, Macquarie University, Sydney, Australia
2 Departamento de Ciências Pesqueiras, Universidade Federal do Amazonas, Manaus, AM, Brazil

Keywords
phylogenetics; biogeography; conservation genetics; speciation; cryptic species; Rio Negro; freshwater fish.

Correspondence
Luciano B. Beheregaray, Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia. Fax: + 61 2 9850 8245
Email: luciano.beheregaray@bio.mq.edu.au

*Current address: Mark Sistrom, Evolutionary Biology Unit, South Australian Museum, Adelaide, SA, Australia.

Editor: Jean-Nicolas Volff

Received 10 July 2008; revised 11 March 2009; accepted 11 March 2009

doi:10.1111/j.1469-7998.2009.00576.x

Abstract

The evolutionary history of Amazonian organisms is generally poorly understood. This is particularly true for small floodplain fish species that show reduced dispersal capabilities. The one-lined pencilfish *Nannostomus unifasciatus* (family Lebiasinidae) is a small fish found in flooded forests of the Rio Negro Floodplain (RNF) in central Amazonia, Brazil. We used a large number of samples collected throughout the species distribution in the RNF and in its headwaters and DNA data from the second intron of the S7 ribosomal protein to reconstruct the phylogeography of *N. unifasciatus*. Two markedly distinct phylogroups of *N. unifasciatus* were detected in the RNF. Although these lineages are largely allopatric, they remain reproductively isolated in regions where they overlap, suggesting cryptic speciation in this group in the Rio Negro basin. Coalescent-based statistical methods suggest that the history of these populations was dominated by a Miocene fragmentation of the species in the headwaters of the basin that originated the two phylogroups, followed by recent events of demographic and range expansions in the floodplain. This pattern is discussed within the context of the geomorphologic history of the region, especially geotectonics, and of marine incursions. Our results match the predictions of the palaeogeography hypothesis of speciation and outline the usefulness of an intron DNA marker to reconstruct population history of a central Amazonian fish. Because *N. unifasciatus* is harvested commercially in the ornamental fishery of the RNF, the two species and limited dispersal capacity between nearby populations identified here are also important to develop sound management decisions in the region.

Introduction

The Amazon basin is a hotspot of diversity for many organisms, particularly freshwater fishes. The basin hosts *c*. 3200 described freshwater fish species (Vari & Malabarba, 1998; Chao, 2001; Junk & Soares, 2001) and potentially up to 4800 species are yet to be described in the region (Vari & Malabarba, 1998). This represents the highest diversity of freshwater fish species in any river basin in the world (Chao & Prang, 1997; Lundberg *et al*., 2000). Very little is known about species relationships, population ecology, population history and phylogenetic assemblages in Amazonia (Da Silva & Patton, 1998; Kress *et al*., 1998; Beheregaray, 2008). Of particular interest is the lack of information about the processes which have resulted in the high levels of biodiversity found in the region. Hypotheses explaining the high levels of biodiversity in the tropics generally relate to low rates of extinction and/or high rates of speciation (Moritz *et al*., 2000). Hypotheses favoring low rates of extinction state that the tropics are stable and old relative to temperate and boreal systems and thus have accumulated a large number of paleoendemics (Moritz *et al*., 2000). Hypotheses favoring high rates of speciation include the riverine barrier hypothesis (Wallace, 1876), the refugia hypothesis (Haffer, 1969), the marine incursion 'museum' hypothesis (Hoorn, 1993; Fjeldså, 1994; Lovejoy, Bermingham & Martin, 1998), the gradient hypothesis (Endler, 1977), and the palaeogeography hypothesis (Räsänen, Salo & Kalliola, 1987; Da Silva & Patton, 1998).

Studies that use genetic data to understand the history of populations (i.e. phylogeographic studies) have the potential to discriminate competing hypotheses of speciation for Amazonian organisms. Although these studies are relatively scarce in the Amazon basin (Beheregaray, 2008), phylogeographic surveys have provided evidence for the involvement of tectonics, riverine barriers and marine incursions in shaping genetic structure and diversification of Amazonian organisms (Da Silva & Patton, 1998; Aleixo, 2006; Elmer, Davila & Lougheed, 2007; Hubert *et al*., 2007). Here we contribute to the understanding of population history and speciation in Amazonian organisms by reconstructing the phylogeography of a small fish species from the Rio Negro basin, in central Amazonia.

The Rio Negro is one of the largest tributaries of the Amazon River and one of the largest rivers in water volume in the world (Clapperton, 1993; Latrubesse & Franzinelli, 2005). It drains a basin of over 600 000 km^2 and extends through parts of Colombia, Venezuela and Brazil (Chao, 2001; Latrubesse & Franzinelli, 2005). The Rio Negro is constrained by a number of tectonic influences. These tectonic constraints take the form of large faults and arches related to transcurrent movements, which affected Quaternary stone lines (Hoorn, 1993; Lundberg *et al*., 1998; Latrubesse & Franzinelli, 2005). An important tectonic arch in the region is the Váupes, an arch located in the headwaters of the Rio Negro that emerged during the Tertiary and caused the separation between the Orinoco and the Amazon basin during the late Miocene (Hoorn, 1993). Sediment loads seen in the Rio Negro today are inconsistent with levels of sedimentation found in the alluvial floodplains of the river. As such it has been suggested that like the Solimões-Amazon River, the Rio Negro experienced high levels of vertical accretion during the Holocene, becoming a blocked river valley (Lundberg *et al*., 1998; Latrubesse & Franzinelli, 2005). This recent blocking of the river valley with sediment has created a persisting backwater effect on the river, creating large areas of standing water along its lower reaches and influenced the development of its large floodplain and characteristic anabranching form (Lundberg *et al*., 1998; Parolin *et al*., 2004; Latrubesse & Franzinelli, 2005).

The Rio Negro has a high level of fish diversity, with over 1000 fish species in the region (Chao, 2001). Many of these are small species that show a life-cycle associated to the flooded forest environment and the seasonal fluctuation of water level in the floodplain. One of these is our study species, the one-lined pencilfish *Nannostomus unifasciatus* (family Lebiasinidae). Pencilfish (genus *Nannostomus*) are found in many flooded forests, rivers and streams of the Amazon basin; however, *N. unifasciatus* appears to be found exclusively in flooded forest habitat (Weitzman & Coob, 1975; L. B. Beheregaray & N. L. Chao, pers. obs.). Records of this species in Brazil, Bolivia and Colombia (Catarino *et al*., 2004) probably do not represent a single species given the wide geographic distance separating these regions, the limited dispersal capability of the species, and the potential for morphologically cryptic species in the Amazon (Lundberg *et al*., 2000; Beheregaray & Caccone, 2007). Because the type-locality of *N. unifasciatus* is the Rio Negro Floodplain (RNF), we refer the species in the strict sense for the RNF. *Nannostomus unifasciatus* and other small floodplain fishes have been harvested commercially in large numbers in the ornamental fishery (about 40 millions $year^{-1}$), which has provided a sustainable income for the local riverine people for over six decades (Chao & Prang, 1997). As it is the case for other ornamental fish species in the RNF, information about levels of population structure and connectivity needed to develop sustainable management strategies are non-existing (Chao & Prang, 1997; Norris & Chao, 2002).

In this study we used a large number of samples of *N. unifasciatus* collected from throughout the species distribution in the RNF and from its remote headwater region to reconstruct the phylogeography of the species. Our DNA data are based on sequence variation of an intron nuclear marker, the S7 ribosomal protein intron 2 (S72). The application of intron markers in phylogeographic studies is a relatively novel approach compared with using data from the more traditional and popular mitochondrial DNA (mtDNA) (Avise, 1998; Beheregaray, 2008). Owing to the relatively slow mutation rate seen in introns compared with mtDNA (Friesen, 2000; Caccone *et al*., 2004; but see Cooke & Beheregaray, 2007 for an exception), they have the potential to provide good resolution in scenarios influenced by ancient biogeographic events (Avise, 1994; Freisen, 2000). We predict that the main ancient geological events in the Rio Negro basin during the Tertiary have resulted in historical vicariance of *N. unifasciatus* populations, with subsequent population expansion into the more recently formed floodplain. By conducting fine-scale sampling in a species with restricted dispersal, we can also interpret phylogeographic patterns within the context of key hypotheses of speciation for rainforest-dependent organisms and, as such, contribute to our understanding of biotic diversification in Amazonia. Finally, we use information from phylogeographic patterns in *N. unifasciatus* to propose management strategies relevant to the sustainable development of its ornamental fishery in the RNF.

Materials and methods

Field sampling

Samples were collected by LBB in the middle and upper RNF in 2002 and 2003 and in the headwaters region in 2005. A total of 403 samples were collected from 20 Rio Negro tributaries in central Amazonia, Brazil (Fig. 1 and Table 1). Fish were collected in the flooded forest (igapó) with hand held dip-nets and a small piece of muscle tissue was taken from behind the dorsal fin of each sample. No samples from the ornamental trade were used. Tissue samples were preserved in 95% ethanol. Voucher specimens were deposited in the fish collection of the Universidade Federal do Amazonas (Brazil).

Molecular methods

DNA extractions were carried out using a modified salting out technique (Sunnucks & Hales, 1996). The S72 intron was amplified using Exon Primed Intron Crossing PCR. Initially primers developed by Chow & Hazama (1998) were used. However as a result of difficulties in amplification, a new forward primer for S72 was designed (5′-GGCATTTCCCAGGTAAGATCA) based on *Nannostomus* data from initial amplifications and sequencing and GenBank data. The PCR amplification was conducted with a program consisting of 2 min at 94 °C, followed by 30 cycles at 94 °C for 30 s, primer annealing at 65 °C for 30 s and extension at 72 °C for 45 s, followed by a final extension step of 72 °C. All PCRs were carried out using an Eppendorf Mastercycler Gradient PCR cycler (Eppendorf, Hamburg,

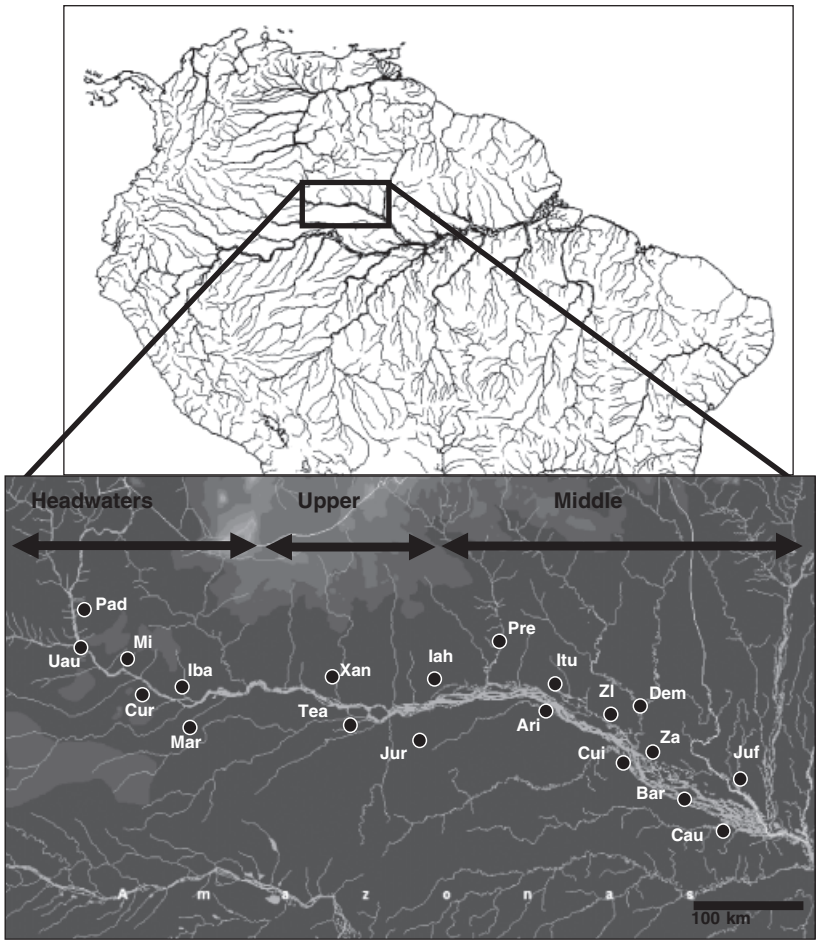

Figure 1 The Rio Negro region in northern South America, indicating the location of the study area and sampling sites of *Nannostomus unifasciatus*.

Germany). Reactions were carried out in 20 μL volumes, consisting of 0.2 μL of Qiagen Taq DNA polymerase (2 U), 0.5 mM of each dNTP, 1.8 mM of $MgCl_2$, 1 × PCR buffer (50 mM Tris/HCl, pH 8.3), 2.4 pmol of each primer and 150 ng of genomic DNA. Purification of PCR products excised from 2% agarose gel was conducted using the Ultraclean™ 15 DNA purification kit (Mo Bio Laboratories Inc., Carlsbad, CA, USA).

We use the single-stranded confirmation polymorphism technique (SSCP) following Sunnucks *et al*. (2000) to screen 15 individuals from each tributary for sequence variation. Two populations were run per gel, making for a total of 10 gels. This was undertaken by conducting the PCR reaction, as previously described in the presence of 0.05 μL α^{33}P dATP (10 mCi mL^{-1}) per sample. SSCP was run on non-denaturing polyacrylamide gels at constant 4 °C, at 4 W for 4 h. Gels were then dried and autoradiographed using Amersham Biosciences Hyperfilm™ High Performance Autoradiography Film. Gels were visually examined independently and heterozygotic and homozygotic individuals were easily distinguished due to the distinct increase in band number shown in heterozygotes. A representative sample of each homozygotic phenotype from each gel was sequenced in an Applied Biosystems 3130xl Genetic Analyser (Applied Biosystems, Carlsbad, CA, USA) following the manufacturer's directions. Representatives from heterozygote gel phenotypes were selected for cloning and subsequent sequencing. Cloning was undertaken using an Invitrogen TOPO TA Cloning Kit (Invitrogen Corporation, Carlsbad, CA, USA) to produce sequence data for both strands of S72. The process of cloning was needed because it was not possible to resolve the allele combinations based simply on a visual inspection of SSCP gels (see Sunnucks *et al*., 2000 for an example). All sequences were aligned and edited using Sequencher 4.1 (Gene Codes Corp., Ann Arbor, MI, USA).

Data analysis

Genetic variation and neutrality tests

Nucleotide diversity (π) and the mean number of pairwise differences between each allele was calculated for each tributary sample. Tajima (1989) and Fu & Li's (1993) tests for neutrality, based on the infinite-site model without recombination, were performed on all S72 sequence data including deletions as the fifth character. Empirical distributions were calculated using coalescent simulations for nucleotide diversity, Tajima's *D*'s and Fu's *F*'s. From these

Table 1 Tributaries sampled for *Nannostomus unifasciatus* including locality abbreviation, geographic co-ordinates and margin of the Rio Negro collected

Tributary	Abbreviation	Geographic co-ordinates	Region/margin
Jufari	Juf	00.59.40S 62.06.10W	Midwaters/left
Caurés	Cau	01.19.01S 62.24.54W	Midwaters/right
Zamula	Za	00.51.57S 62.46.22W	Midwaters/left
Baruri	Bar	00.53.35S 63.03.38W	Midwaters/right
Demini	Dem	00.23.40S 62.51.17W	Midwaters/left
Zalala	Zl	00.39.59S 63.00.32W	Midwaters/left
Cuiuni	Cui	00.46.09S 63.10.40W	Midwaters/right
Arirahá	Ari	00.26.54S 63.41.43W	Midwaters/right
Itu	Itu	00.26.00S 63.07.00W	Midwaters/left
Preto	Pre	00.06.40S 64.05.03W	Midwaters/left
Iahá	Iah	00.23.47S 64.36.26W	Upper/left
Jurubaxi	Jur	00.33.07S 64.48.06W	Upper/right
Arixaná	Xan	00.21.50S 62.11.51W	Upper/left
Tea	Tea	00.32.59S 65.15.13W	Upper/right
Niuá-mirim	Mi	00.08.24S 66.54.91W	Headwaters/left
Curicuriari	Cur	00.13.35S 66.24.58W	Headwaters/right
Marié	Mar	00.26.32S 66.24.58W	Headwaters/right
Vaupés	Vau	00.04.47N 67.24.13W	Headwaters/right
Paduá	Pad	00.12.23N 67.19.23W	Headwaters/left
Ibará	Iba	00.20.15S 66.35.04W	Headwaters/left

distributions, 95% confidence intervals were given for each statistic. All calculations were performed using DnaSP4.10 (Rozas *et al*., 2003).

Phylogenetic and genealogical relationships

We used two methods to infer relationships among S72 alleles. Firstly, a Bayesian phylogenetic analysis was conducted using MrBayes3.1 (Ronquist & Huelsenbeck, 2003). For this analysis, Markov Chain Monte Carlo sampling was conducted every 1000th generation until the standard deviation of split frequencies was below 0.01. A burn in period equal to 25% of the total generations was required in order to summarize the parameter values and trees. Parameter values were assessed based on 95% credibility levels to ensure the analysis had run for a sufficient number of generations. For this analysis we used the congeneric and morphologically similar *Nannostomus eques* as outgroup, a species proposed as sister to *N. unifasciatus* (Weitzman, 1978). Secondly, genealogical relationships were investigated by constructing a haplotype network using the software TCS 1.21 (Clement, Posada & Crandall, 2000). Deletions were included as a fifth character due to the relatively low variability of the data.

Population structure and history

Genetic differentiation between populations was investigated using comparisons of pairwise fixation values (Tajima, 1993) using Arlequin 3.01 (Schneider *et al*., 2000). A Tamura correction allowing for variable base frequencies, variable transversion frequencies and variable transition frequencies was implemented in the calculation of pairwise difference (θ_{ST}) values. The parameters of the Tamura correction are the most similar to the TrN+I model suggested by Modeltest that Arlequin 3.01 offers. We also investigated hierarchical patterns of genetic structure by performing an AMOVA in Arlequin 3.01 (Excoffier *et al*., 1992) using a Tamura correction and 10 000 permutations. The hierarchical groups tested were: (1) All populations as a single group; (2) All populations on the left margin of the Rio Negro versus all populations on the right margin; (3) All populations found in the headwaters versus those found in the middle and lower Rio Negro.

To test for a pattern of genetic structure consistent with isolation by distance of tributary populations along the floodplain, Mantel tests (Smouse, Long & Sokal, 1986) were undertaken using Arlequin 3.01. Distances along watercourses were approximated using Google Earth (2005). Mantel tests were conducted for (1) all tributary samples; (2) all tributary samples on the left margin of the Rio Negro; (3) all tributary samples on the right margin of the Rio Negro. The tests separating tributary samples on each margin were conducted to assess if the main channel of the Rio Negro, which can be over 30 km wide in some regions, acts as a barrier for gene flow between populations on each margin. This possibility is consistent with field observations because pencilfish are not found in the channels of the Rio Negro, but only on its tributary streams in the presence of flooded forest habitat.

The possibility of historical demographic expansion events associated with the formation of the extensive RNF was examined for each lineage independently by mismatch analysis (Rogers & Harpending, 1992) implemented in Arlequin 3.01 using a parametric bootstrap approach. Analyses were carried out for all sampled populations of *N. unifasciatus* in the RNF and for all individuals identified as lineage A and all individuals identified as lineage B in the genealogical and phylogenetic analyses (details in 'Results').

With the aim to assess competing evens that have influenced the development of phylogeographic patterns in *N. unifasciatus*, we used Nested Clade Phylogeographic Analysis (NCPA). This method uses information provided by nested genealogies to infer historical and recurrent events that have influenced population structure (Templeton, 1998). Using the haplotype network produced in TCS, loops in the network were resolved based on assumptions of coalescence (Templeton, 1998). Clades were then nested and GeoDis2.5 (Posada, Crandall & Templeton, 2000) was used to estimate statistically significant events that were interpreted using Templeton's updated inference key found at (http://darwin.uvigo.es/software/geodis.html). The NCPA approach has recently received criticism related to high false-positive rates documented in some studies (Petit, 2008). This opinion has been reconsidered by several authors (e.g. Garrick *et al*., 2008; Templeton, 2008) who strongly support the notion that NCPA provides a useful phylogeographic approach with no currently computationally feasible substitute. In addition, they suggest that NCPA's results are more valuable when interpreted cautiously and validated in combination with other methods and geomorphological data (Garrick *et al*., 2008), which is the strategy we used in our study.

Estimation of divergence time

In order to obtain an approximation of the time since divergence between the two groups detected (lineages A and B, details in 'Results'), the time of coalescence to the most recent common ancestor (MRCA) of these groups was estimated. For this we estimated a series of population parameters such as g (the exponential growth parameter in units of μ^{-1}) and current population size at a given substitution rate (μ) using Lamarc 2.1 (Kuhner, 2006). The appropriate transition:transversion (ti:tv) ratio (ti:tv = 1.3392) and α (g) values ($g = 0.2680$) were calculated using Modeltest version 2.2 (Posada & Crandall, 1998). This value was then used as an initial estimate of Θ. An average of eight recently published intron substitution rates (3.8875×10^{-10}) (Alter, Rynes & Palumbi, 2007) was used as an initial estimate.

The analysis was repeated 10 times for the dataset in order to ensure the stability of the parameter estimates. Each run was conducted using 10 short Monte Carlo chains of 200 steps each and two long chains of 20 000 steps each with a sampling increment of 20. Lamarc generated a random topology for initial searching. Mean values of Θ, g and current population size calculated from the combined Lamarc output were then used to estimate the time to the MRCA of S72 alleles using the methodology specified in Wares & Cunningham (2001). An estimation that coalescence occurs when N_e reaches 1% of its current size was assumed. Using then the equation: $N_t/100 = \Theta e^{-(g\mu)t}$ where N_t is the effective population size at any time t in the past (Kuhner, Yamato & Felsenstein, 1998), t was derived as an estimate of the time since divergence of lineages A and B.

Results

Genetic variation and neutrality

Out of the total sample size from SSCP gels, 258 individuals were homozygotic and 57 were heterozygotic. The S72 in *N. unifasciatus* was 149 bp in length and 14 unique alleles with 17 polymorphic sites were detected (GenBank accession numbers for the data are FJ263890–FJ263903). The mean number of pairwise differences was 1.892 ± 0.595 and the nucleotide diversity (π) was 0.01295 ± 0.00123. Samples from headwater tributaries showed no within site variability, where as the floodplain had a range of within site variability from 0 to 3.66 mean pairwise differences (Table 2). Tajima's (1989) test for neutrality was not significant ($D = -0.8207$, $P = 0.45$). Both Fu's F's (F's $= -0.49$, $P = 0.12$) and Fu and Li's F^* ($F^* = 0.81$, $P = 0.32$) statistics were also not significant. This pattern suggests that the S72 data does not deviate from a model of neutral evolution, a result also observed by Cooke & Beheregaray (2007) for their S72 intron data.

Genealogical relationships

The Bayesian tree is shown in Fig. 2 with Bayesian posterior probabilities indicated beneath each node. The results of this analysis are consistent with the unrooted haplotype network shown in Fig. 3. Both methods show that *N. unifasciatus* is represented in the Rio Negro by two divergent groups of alleles, named lineages A and B here. The putatively ancestral alleles in each group (alleles 1 and 2; Fig. 3) were manually linked in the haplotype network in order to depict

Table 2 Pairwise nucleotide difference and nucleotide diversity values of S72 intron data for population samples of *Nannostomus unifasciatus*

Tributary sample	Sample size	No. of alleles	Mean no. of nucleotide differences	Nucleotide diversity
Vau	15	1	0	0
Pad	13	1	0	0
Mi	15	1	0	0
Cur	10	1	0	0
Iba	11	1	0	0
Mar	13	1	0	0
Xan	10	3	3.66 ± 2.03	0.026 ± 0.002
Tea	12	3	0.77 ± 0.57	0.005 ± 0.001
Jur	6	2	2.41 ± 1.37	0.004 ± 0.001
Iah	10	1	0	0
Pre	11	2	0.18 ± 0.03	0.001 ± 0.001
Ari	9	1	0	0
Itu	13	2	0.39 ± 0.32	0.003 ± 0.001
Za	10	1	0	0
Cu	12	1	0	0
Dem	11	3	1.08 ± 0.94	0.007 ± 0.002
Bar	9	3	3.62 ± 1.95	0.026 ± 0.004
Zl	10	1	0	0
Cau	13	2	0.85 ± 0.58	0.013 ± 0.001
Juf	11	1	0	0

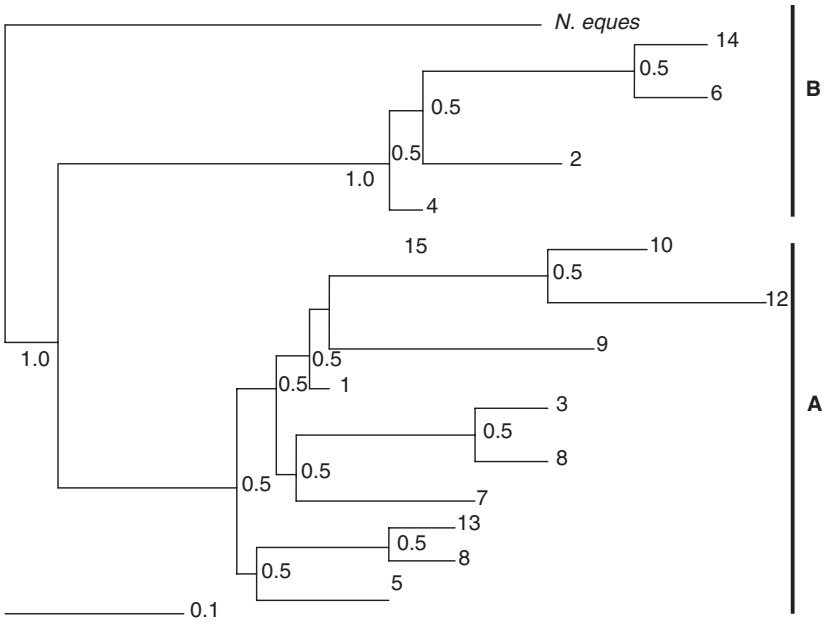

Figure 2 Bayesian tree of S72 alleles in *Nannostomus unifasciatus*. Numbers at the ends of branches represent individual alleles. Numbers at tree forks represent Bayesian posterior probability values. Two distinct clades, labeled A and B are seen in the tree. *Nannostomus eques* was used as outgroup.

their evolutionary distance. Marked sequence divergence exists between the two lineages (ranging between 2.7 and 3.6%). In contrast, each lineage is composed of highly related S72 alleles (maximum sequence divergence = 8.7%). Allele 2, the putative ancestral allele for lineage B, is only found in the headwaters, whereas allele 1 (the putative ancestral allele for lineage A) is found throughout the distribution of the species in the Rio Negro basin. However, this is the only lineage A allele found in the headwaters. This pattern suggests that the oldest populations of both detected lineages are located in the headwaters of the catchment. The two lineages show a complex but mostly allopatric distribution (Fig. 4). Nonetheless, individuals from the two lineages co-occur in some tributaries in the floodplain (in the tributaries Arixaná and Baruri, Fig. 4). None of the 57 heterozygotic individuals identified had alleles from both lineages, even when found in the same tributary, which strongly suggests reproductive isolation between the two lineages.

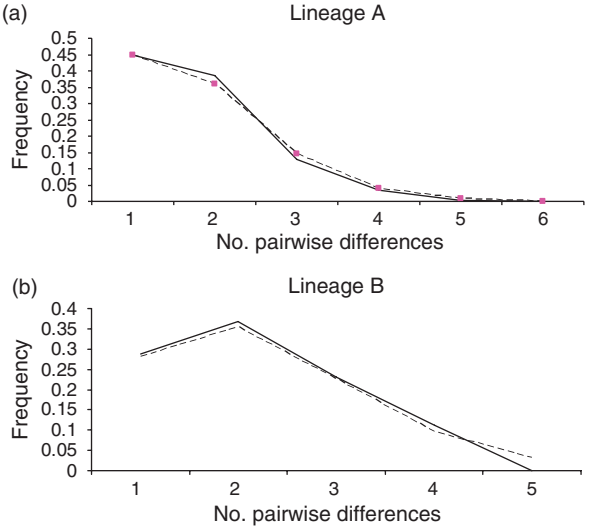

Figure 3 Mismatch analysis of the two detected clades of *Nannostomus unifasciatus*. The solid lines represent observed frequency of nucleotide differences between pairs of individuals, the dashed line represents the frequency expected under a model of demographic expansion.

Population structure, history and estimation of divergence time

Comparisons of pairwise population θ_{ST} values are shown in Table 3. Data were partitioned in this manner due to the fact that the groups are distributed both allopatrically and sympatrically in different parts of the RNF. High levels of genetic differentiation were detected in many population pairwise comparisons, especially for lineage A, in which headwater populations accounted for most of the divergence. For the AMOVA, when all populations were included in the analysis 60.07% of the variation accounted for by differences within populations and the correspondingly large 39.93% of the variation accounted for by differences between populations ($P<0.001$). On the other hand, no significant differences between groups composed of headwater and of middle/lower floodplain populations were detected by AMOVA ($P = 0.30$). AMOVA comparisons according to river margins also yielded non-significant results. This suggests that a simplistic model of allopatric divergence due to the Rio Negro acting as a barrier is not supported by the data. Different scenarios of isolation by distance were tested using Mantel tests but all resulted in non-significant results. As such, it is likely that *N. unifasciatus* populations do not display genetic structure associated

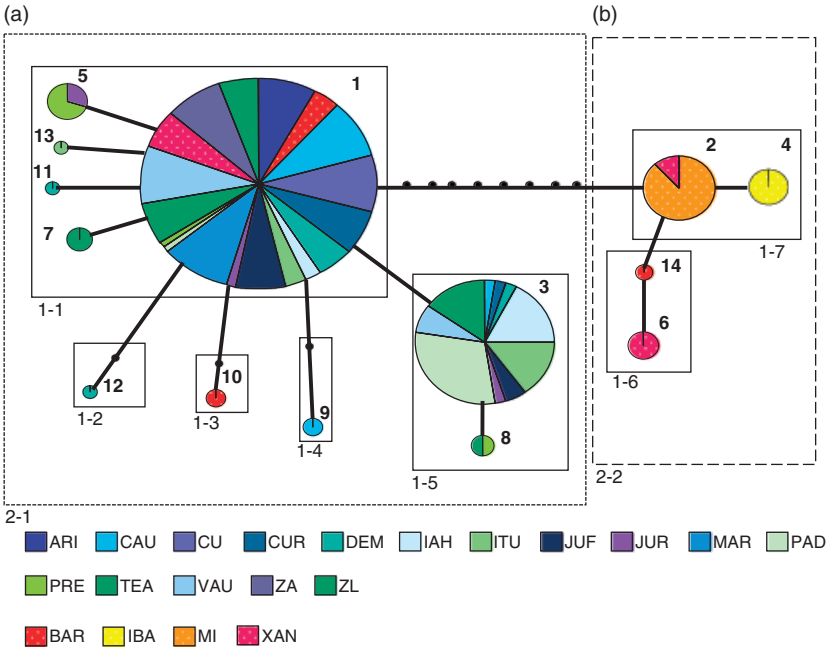

Figure 4 Nesting clade phylogeographic design for the haplotype network with loops resolved. Numbers in bold indicate alleles (numbered as in Fig. 1). The first smaller number indicates the nesting level, the second label the nested clade within its nesting level. Hatched segments represent samples collected from populations where lineage group B is found, solid colors represent samples collected from areas where only lineage group A was found.

with a simple model of geographical distance between tributaries.

In terms of demographic history, mismatch analysis resulted in non-significant *P* values (Table 4 and Fig. 3). This indicates that a model of population expansion cannot be rejected as a potential factor in the history of *N. unifasciatus* in the RNF. For NCPA (Fig. 4 and Table 5), the process affecting the total cladogram (nested clade 3-1), and therefore the most ancient population process, was past fragmentation. It appears that this fragmentation was subsequently followed sometime later by marked events of range expansion, as suggested for clades 2-1 and 2-2. Inferences about fine-scale population processes are mostly constrained by the low variability of the S72 data. However the results suggest that both level 1 clades in Group B and some level 1 clades in Group A have experienced gene flow restriction because of the expansion of the two groups.

For the analysis of divergence time, the values for g and Θ estimated in Lamarc were positive for growth for all runs. This is important because it corroborates the inferences about demographic and range expansion detected for both lineages of *N. unifasciatus* made by mismatch analysis and NCPA. The strong signal for population growth in both lineages allowed us to estimate an approximate time since divergence of lineages A and B, which was in the late Miocene, around 8.4 Mya (8.434×10^6 years). The upper and lower confidence interval of this estimate were 20.102×10^6 years and 5.1553931×10^6 years, respectively.

Discussion

Our analysis of intron DNA data in the small one-lined pencilfish *N. unifasciatus* revealed two markedly distinct lineages in the Rio Negro basin, central Amazonia. These two groups (A and B) were detected by genealogical, phylogenetic and population genetic analyses (Tables 3–5 and Figs 2–4). Cloning and subsequent DNA sequencing of S7 heterozygote individuals showed that no hybridization occurs between groups A and B, even when they overlap in distribution. Given that these two groups occur in sympatry in at least two Rio Negro tributaries (Arixaná and Baruri), the pattern of intron DNA divergence reported here provides strong evidence that the two groups are reproductively isolated and that one lineage represents a previously undescribed cryptic species (Beheregaray & Caccone, 2007; Bickford *et al.*, 2007; Elmer *et al.*, 2007; Pfenninger & Schwenk, 2007). In addition, our intron results are in agreement with data based on microsatellite markers (Beheregaray *et al.*, 2004; S. Corrigan *et al.*, unpubl. data) and mtDNA sequences (L. B. Beheregaray *et al.*, unpubl. data) collected from the same populations that also indicate the presence of two reproductively isolated groups. Evidence for the existence of cryptic diversity in *N. unifasciatus* and other Amazonian fish species (Junk & Soares, 2001; Hubert *et al.*, 2007; Piggott, Chao & Beheregaray, in review) suggest that the diversity of the Amazonian ichthyofauna is much higher than previously reported by bioinventories and morphological studies.

Analysis of population genetic structure based on traditional methods (e.g. θ_{ST}, AMOVA and Mantel tests) did not suggest any obvious geographic pattern of differentiation. For instance, *N. unifasciatus* is a forest-dependent species that shows reduced dispersal, is not found in Rio Negro's main channel, and therefore could represent a suitable test for the riverine hypothesis (Wallace, 1876). The 'riverine' hypothesis proposes that large rivers can act as barriers to dispersal and promote allopatric divergence on different margins. However, partitioning our dataset according to

Table 3 Comparison of pairwise θ_{ST} values for populations of *Nannostomus unifasciatus* based on S72 data

Group A	Vau	Pad	Cur	Mar	Xan	Tea	Jur	Iah	Pre	Ari	Itu	Zl	CurR	Dem	Bar	Za	Cau	Juf
Vau																		
Pad	0																	
Cur	0.552**	0.528**																
Mar	0	0	0.528**															
Xan	1	1**	0.587**	1**														
Tea	0.278**	0.26**	0.473**	0.26**	0.933**													
Jur	0.661**	0.637**	0.458**	0.637*	0.969**	0.121												
Iah	0	0	0.486**	0	1**	0.23**	0.594**											
Pre	0.919**	0.912**	0.599**	0.912**	0.988**	0.432**	0.237	0.9**										
Ari	0	0	0.471**	0	1**	0.219**	0.577*	0	0.896**									
Itu	0.518*	0.491*	0.436**	0.491*	0.968**	0.193	0.301	0.443*	0.661**	0.425								
Zl	0	0	0.486**	0	1**	0.23**	0.594**	0	0.9**	0	0.443							
Cur	0.511**	0.49**	0.052	0.49**	0.559*	0.453**	0.439**	0.453**	0.568**	0.44**	0.417**	0.453**						
Dem	0.339**	0.319**	0.007*	0.319**	0.601**	0.309**	0.3**	0.285**	0.456**	0.273**	0.271**	0.285**	0.001**					
Bar	0.647**	0.62**	0.427**	0.62**	0.948**	0.398**	0.499**	0.57**	0.74**	0.55**	0.449**	0.57**	0.418**	0.291**				
Za	0	0	0.486**	0	1**	0.23**	0.594**	0	0.9**	0	0.443	0	0.453**	0.285**	0.57**			
Cau	0.576**	0.556**	0.003	0.556**	0.45**	0.517**	0.498**	0.522*	0.602**	0.509**	0.483**	0.522**	0.004	0.087	0.474**	0.522**		
Juf	0	0	0.501**	0	1**	0.24**	0.609*	0	0.905**	0	0.46*	0	0.466**	0.297**	0.588**	0	0.534**	

Group B	Mi	Iba	Xan	Bar
MI				
Iba	0			
Xan	0.826**	0.798**		
Bar	1**	1**	0.769**	

*$P<0.05$.

**$P<0.01$.

Table 4 Summary of the results of mismatch analysis

Mismatch analysis	SSD	*P*	*R*	*P*
All populations	0.032931	0.61	0.052873	0.86
Lineage group A	0.000921	0.18	0.079541	0.30
Lineage group B	0.029877	0.20	0.211045	0.46

SSD, sum of squared deviations; *R*, harpending's raggedness index.

Table 5 Inferences based on a nested clade phylogeographic analysis of the *Nannostomus unifasciatus* S72 intron dataset

Clade	Chain of inference	Inferred population process
1-1	1-2-3-4-No	Restricted gene flow with isolation by distance
1-5	1-2-3-4-No	Restricted gene flow with isolation by distance
1-6	1-2-3-4-9-No	Past fragmentation
1-7	1-2-3-5-6-7-Yes	Restricted gene flow with some long distance dispersal.
2-1	1-2-11-12-13-14-No	Between contiguous range expansion and long distance colonization
2-2	1-2-11-12-No	Contiguous range expansion
Total cladogram	1-2-3-4-9-No	Past fragmentation

river margin did no yield any discernible or significant pattern of structure. This suggests that the Rio Negro has not acted as a strong barrier to gene flow in this forest fish. The other possibility, that populations were structured following a simple model of isolation by geographic distance, was extensively tested using Mantel tests but yielded non-significant results. Lateral dispersal (via the floodplain) in this forest fish is expected to be associated with the annual inundation cycle of the Rio Negro floodplain (Winemiller, 1993; Marshall, Forsberg & Thome-Souza, 2008). On the other hand, one should expect reduced habitat connectivity for *N. unifasciatus* in the steeper sided valleys and fragmented flooded forests of the headwater regions of the Rio Negro. We did find a lack of allelic diversity and higher genetic structure (including fixation of distinct alleles) in headwater populations, which is consistent with reduced habitat connectivity in this environment. Nonetheless, we did not find any significant pattern of structure by conducting comparisons between headwater versus floodplain populations. Altogether, these results suggest a more complex geographic scenario driving the diversification of *N. unifasciatus* in the Rio Negro.

In contrast with results based on traditional analyses of genetic structure, our analysis of population history based on coalescent statistical methods provided strong support for a complex biogeographic scenario of divergence in *N. unifasciatus*. This scenario combines both old and recent well understood and well characterized geomorphological events (Clapperton, 1993; Lundberg, 1998; Latrubesse & Franzinelli, 2005) that account surprisingly well for the structure observed in our study. As described below, most of these events did not leave obvious geographic features in the contemporary landscape of the Rio Negro (Lundberg, 1998; Latrubesse & Franzinelli, 2005), which partially explains the reduced efficiency of traditional analyses when testing for geographic associations of populations in our dataset.

The most ancient population process inferred in the history of *N. unifasciatus* in the Rio Negro was past fragmentation (Fig. 4 and Table 5). Information derived from the genealogical arrangements of S72 alleles and their relative levels of divergence strongly suggest that the oldest populations are located in the headwaters of the basin. The region today represented by the headwaters of the Rio Negro underwent dramatic geological, climatic and geomorphic change during the Miocene. The final stages of uplift of the northern Andes, the youngest mountain range in the Andes system, are of late Miocene age (Clapperton, 1993; Lundberg *et al*., 1998). These final stages resulted in major re-organization of the drainage patterns of Amazonian fluvial systems and established several well-documented paleoarches that bound inter-cratonic segments of the foreland basin (Lundberg *et al*., 1998). Of particular interest is the location of the Vaupés arch in the headwater of the Rio Negro. We propose that the late Miocene uplift of this arch, which separated the Orinoco from the Negro river systems around 11 Mya (Lundberg *et al*., 1998), acted as a vicariant event isolating headwater populations of *N. unifasciatus* in two lineage groups (Fig. 5). Paleoarches have been implied to play an important role in population differentiation of other Amazonian organisms (Da Silva & Patton, 1998; Elmer *et al*., 2007; Hubert *et al*., 2007). Our suggestion is consistent with results of the coalescent method implemented in Lamarc used to estimate time to MRCA. This method places the divergence of the two groups of *N. unifasciatus* in the late Miocene, around 8.4 Mya. Interestingly, this date considerably predates the environmental fluctuations thought to be associated with the creation of Pleistocene rainforest refugia in Amazonia (Haffer, 1969; Moritz *et al*., 2000) and also the formation of the current

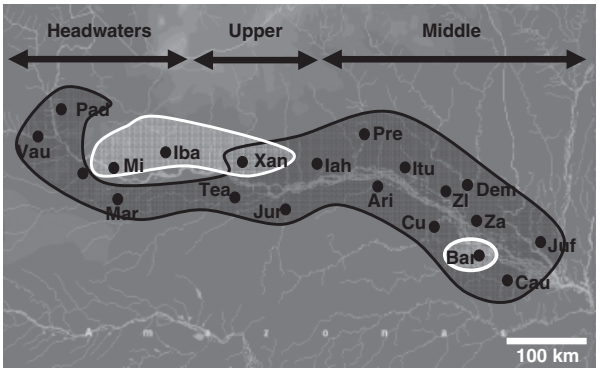

Figure 5 Geographic distribution of the two groups identified based on S72 intron data. Lineage group A is shown in black and lineage group B in white. Note that in the headwaters the groups are allopatrically distributed but in the middle and lower floodplain there is some overlap of the two groups.

drainage pattern of the Rio Negro (Clapperton, 1993; Lundberg *et al*., 1998; Latrubesse & Franzinelli, 2005).

Another potential scenario of divergence for the two groups relates to the eustatic sea level changes and subsequent marine incursions that occurred in northern Amazon during the Miocene (Clapperton, 1993; Lovejoy *et al*., 1998; Lundberg *et al*., 1998). These marine incursions are thought to have fragmented the Amazon basin considerably, confining forest taxa to highland refugia (e.g. Lovejoy *et al*., 1998; Hall & Harvey, 2002; Aleixo, 2006). However, the precise extent and persistence of fragmentation caused by incursion of marine or estuarine waters into the RNF area are difficult to infer in light of the dynamic nature of the topology of the northern Amazonian region during this period (Clapperton, 1993; Hoorn, 1993; Lundberg *et al*., 1998). It is chronologically feasible to propose that a combination of one or more marine incursions that restricted freshwater fishes to higher altitude areas and the arch uplift that separated headwater tributaries could account for the prolonged fragmentation of *N. unifasciatus* populations and subsequent divergence of lineages A and B.

Following allopatric divergence in the headwaters, our analyses recovered strong signal of population growth in the floodplain, as inferred by both mismatch analysis and parameters estimated using Lamarc. In addition, NCPA infers range expansions as population processes affecting the two lineage groups (Fig. 4 and Table 5). This indicates that, more recently, an increase in population size of *N. unifasciatus* in the floodplain associated with range expansion from the headwaters is likely to have occurred. Vertical accretion of very large quantities of sediment in the RNF occurred during the Holocene, which in turn resulted in the creation of the contemporary structure of the RNF (Latrubesse & Franzinelli, 2005). The filling in with sediment of the Rio Negro river valley and subsequent backwater effects between the Rio Negro and the Amazon directed the establishment of the extensive alluvial fans seen today in the RNF (Latrubesse & Franzinelli, 2005). This resulted in high level of habitat connectivity for fish species reliant on flooded forest habitat (e.g. Weitzman & Coob, 1975; Chao, 2001; Marshall *et al*., 2008). Thus, the largely contiguous habitat available in the floodplain during the Holocene has produced conditions conductive to expansion of *N. unifasciatus* throughout the RNF, as inferred in this study.

In summary, we suggest that geomorphological features that have shaped the evolution of *N. unifasciatus* in the RNF are no longer clearly evident in the contemporary landscape. The combination of genealogical methods that are essentially graphical in nature (e.g. NCA) and statistical phylogeographic approaches (e.g. coalescent simulations) have consistently indicated that historical fragmentation took place in the headwaters, followed by recent range expansion and exponential population growth in the floodplain region. Our reconstruction of population history in this forest-dependent species that coincides with known events of Earth history in Amazonia is important given the general lack of studies that can propose causative factors about mechanisms responsible for the high biodiversity of the region.

The existence of reproductively isolated cryptic species of *N. unifasciatus* implies that the two lineage groups have undergone distinct evolutionary histories and need to be managed independently (Ryder, 1986; Moritz, 1994). Exploitation of headwater regions should be avoided due to apparent lack of gene flow between populations. This should not be difficult because this is a remote region where traders involved with the pencilfish fishery do not operate. Other localities in the lower floodplain represent contiguously connected populations (except for the lineage B population in Baruri). As such, these areas represent a potentially sustainable fishing zone. Further studies based on markers that can reveal fine-scale genetic structure in *N. unifasciatus* (e.g. microsatellites, Beheregaray *et al*., 2004) are needed to clarify population structures for conservation management.

Acknowledgments

This study was funded by the Discovery program of the Australian Research Council (ARC grant DP0556496 to L.B. Beheregaray), by the ECOSAVE program of the Yale Institute of Biospheric Studies (YIBS) and by Macquarie University through a postgraduate research award to M. Sistrom. Logistics and local arrangements in Amazonia were supported in part by Project Piaba, a project funded by the Universidade Federal do Amazonas and the Brazilian National Council of Research and Technology, CNPq (process 46.6098/2001-4). Collection permit is under IBAMA number 12273-1 to N. Chao.

References

Aleixo, A. (2006). Historical diversification of floodplain forest specialist species in the Amazon: a case study with two species of the avian genus *Xiphorhynchus* (Aves: Dendrocolaptidae). *Biol. J. Linn. Soc.* **89**, 383–395.

Alter, S.E., Rynes, E. & Palumbi, S.R. (2007). DNA evidence for historic population size and past ecosystem impacts of gray whales. *Proc. Natl. Acad. Sci. USA* **104**, 15162–15167.

Avise, J.C. (1994). *Molecular Markers, Natural History, and Evolution*. New York: Chapman & Hall.

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

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

Beheregaray, L.B. & Caccone, A. (2007). Cryptic biodiversity in a changing world. *J. Biol.* **6**, 1–5.

Beheregaray, L.B., Schwartz, T.S., Möller, L.M., Call, D., Chao, N.L. & Caccone, G. (2004). A set of microsatellite DNA markers for the one-lined pencilfish *Nannostomus unifasciatus*, an Amazonian flooded forest fish. *Mol. Ecol. Notes* **4**, 333–335.

Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K. & Das, I. (2007). Cryptic

species as a window on diversity and conservation. *Trends Ecol. Evol.* **22**, 148–155.

Caccone, A., Gentile, G., Burns, C.E., Sezzi, E., Bergman, W., Ruelle, M., Saltonstall, K. & Powell, J.R. (2004). Extreme difference in rate of mitochondrial and nuclear DNA evolution in a large ectotherm, Galápagos tortoises. *Molecular phylogenetics and evolution* **31**, 794–798.

Catarino, M., Rodrigues, H., Paes, E. & Zuanon, J. (2004). Distribution and abundance of *Nannostomus unifasciatus* (Characiformes: Lebiasinidae) at Lago Amana, Amazonas, Brazil. *Int. Congr. Biol. Fish: Fish Commun. Fish* **1**, 215–218.

Chao, N.L. (2001). The fishery, diversity, and conservation of ornamental fishes in the Rio Negro Basin, Brazil – a review of Project Piaba (1989–99). In *Conservation and management of ornamental fish resources of the Rio Negro Basin, Amazonia, Brazil – Project Piaba*: 161–205. Chao, N.L., Petry, P., Prang, G., Sonneschien, L. & Tlusty, L. (Eds). Manaus: Editora da Universidade do Amazonas.

Chao, N.L. & Prang, G. (1997). Project Piaba – towards a sustainable ornamental fishery in the Amazon. *Aquarium Sci. Conserv.* **1**, 105–111.

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

Clapperton, C.M. (1993). *Quaternary geology and geomorphology of South America*. Aberdeen, Scotland, UK: Elsevier.

Clement, M., Posada, D. & Crandall, K. (2000). TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**, 1657–1660.

Cooke, G.M. & Beheregaray, L.B. (2007). Extremely high variability in the S72 intron of the Amazonian cardinal tetra (*Paracheirodon axelrodi*). *J. Fish Biol.* **71**, 132–140.

Da Silva, M.N.F. & Patton, J.L. (1998). Molecular phylogeography and the evolution and conservation of Amazonian mammals. *Mol. Ecol.* **7**, 475–486.

Elmer, K.R., Davila, J.A. & Lougheed, S.C. (2007). Cryptic diversity and deep divergence in an upper Amazonian frog, *Eleutherodactylus ockendeni*. *BMC Evol. Biol.* **7**, 247.

Endler, J.A. (1977). *Geographic variation, speciation, and clines*. New Jersey: Princeton University Press.

Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992). Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes Application to Human Mitochondrial DNA Restriction Data. *Genetics* **131**, 479–491.

Fjeldså, J. (1994). Geographical patterns for relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodivers. Conserv.* **3**, 207–226.

Freisen, V.L. (2000). Introns. In: *Molecular Methods in Ecology*: 274–290. Baker, A.J. (Ed). Oxford: Blackwell Science.

Fu, Y.-X. & Li, W.-H. (1993). Statistical tests of neutrality of mutations. *Genetics* **133**, 693–709.

Garrick, R.C., Dyer, R.J., Beheregaray, L.B. & Sunnucks, P. (2008). Babies and bathwater: a comment on the premature obituary for nested clade phylogeographic analysis. *Molecular Ecology* **17**, 1401–1403.

Haffer, J. (1969). Speciation in Amazonian forest birds. *Science* **165**, 131–137.

Hall, J.P.W. & Harvey, D.J. (2002). The phylogeography of Amazonia revisited: new evidence from riodinid butterflies. *Evolution* **56**, 1489–1497.

Hoorn, C. (1993). Marine incursions and the influence of Andean tectonics on the Miocene depositional history of northwestern Amazonia: results of a palynostratigraphic study. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **105**, 267–309.

Hubert, N., Duponchelle, F., Nunez, J., Garcia-Davila, C., Paugy, D. & Renno, J.F. (2007). Phylogeography of the piranha genera Serrasalmus and Pygocentrus: implications for the diversification of the neotropical ichthyofauna. *Mol. Ecol.* **16**, 2115–2136.

Junk, W.J. & Soares, M.G.M. (2001). Freshwater fish habitats in Amazonia: state of knowledge, management, and protection. *Aquat. Ecosyst. Health Mgmt.* **4**, 437–451.

Kress, W.J., Heyer, W.R., Acevedo, P., Coddington, J., Cole, D., Erwin, T.L., Meggers, B.J., Pogue, M., Thorington, R.W., Vari, R.P., Weitzman, M.J. & Weitzman, S.H. (1998). Amazonian biodiversity: assessing conservation priorities with taxonomic data. *Biodivers. Conserv.* **7**, 1577–1587.

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

Kuhner, M.K., Yamato, J. & Felsenstein, J. (1998). Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* **149**, 429–434.

Latrubesse, E.M. & Franzinelli, E. (2005). The late quaternary evolution of the Negro River, Amazon, Brazil: implications for island and floodplain formation in large anabranching tropical systems. *Geomorphology* **70**, 372–397.

Lovejoy, N.R., Bermingham, E. & Martin, A.P. (1998). Marine incursion into South America. *Nature* **396**, 421–422.

Lundberg, J.G. (1998). The temporal context for the diversification of neotropical fishes. In *Phylogeny and classification of neotropical fishes*: 49–68. Malabarba, L.R., Reis, R.E., Vari, R.P., Lucena, A.S. & Lucena, Z.M. (Eds). Porto Alegre, Brazil: Editora Universitaria – EDIPUCRS.

Lundberg, J.G., Kottelat, M., Smith, G.R., Stiassny, M.L.J. & Gill, A.C. (2000). So many fishes, so little time: an overview of recent ichthyological discovery in continental waters. *Ann. Mo. Bot. Gard.* **87**, 26–62.

Lundberg, J.G., Marshall, L.G., Guerrero, J., Horton, B., Malabarba, M.C. & Wesselingh, F. (1998). The stage for neotropical fish diversification: a history of tropical South American Rivers. In *Phylogeny and classification of neotropical fishes*: 13–48. Malabarba, L.R., Reis, R.E., Vari, R.P.,

Lucena, A.S. & Lucena, Z.M. (Eds). Porto Alegre, Brazil: Editora Universitaria – EDIPCURS.

Marshall, B.G., Forsberg, B.R. & Thome-Souza, M.J.F. (2008). Autotrophic energy sources for *Paracheirodon axelrodi* (Osteichthyes, Characidae) in the middle Negro River, Central Amazon, Brazil. *Hydrobiologica* **596**, 95–103.

Moritz, C. (1994). Defining ''evolutionary significant units'' for conservation. *Trends Ecol. Evol.* **100**, 373–375.

Moritz, C., Patton, J.L., Schneider, C.J. & Smith, T.B. (2000). Diversification of rainforest faunas: an integrated molecular approach. *Annu. Rev. Ecol. Syst.* **31**, 533–563.

Norris, S. & Chao, N.L. (2002). Buy a fish, save a tree – safeguarding sustainability in an Amazonian ornamental fishery. *Conserv. Pract.* **3**, 30–35.

Parolin, P., Adis, J., Rodrigues, W.A., Amaral, I. & Piedade, M.T.F. (2004). Floristic study of an igapo floodplain forest in Central Amazonia, Brazil (Taruma-Mirim, Rio Negro). *Amazoniana* **18**, 29–47.

Petit, R. (2008). The coup de grâce for the nested clade phylogeographic analysis? *Mol. Ecol.* **17**, 516–518.

Pfenninger, M. & Schwenk, K. (2007). Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evol. Biol.* **7**, 121.

Piggott, M., Chao, N.L. & Beheregaray, L.B. (in review). Three fishes in one: cryptic species in an Amazonian floodplain forest specialist, the hatchet-fish *Carnegiella marthae*. *Biol. J. Linn. Soc.*.

Posada, D. & Crandall, K.A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Posada, D., Crandall, K.A. & Templeton, A.R. (2000). GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Mol. Ecol.* **9**, 487–488.

Räsänen, M.E., Salo, J.S. & Kalliola, R. (1987). Fluvial perturbance in the western Amazon basin: regulation by long-term sub-Andean tectonics. *Science* **238**, 1398–1401.

Rogers, A.R. & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9**, 552–569.

Ronquist, F. & Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X. & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496–2497.

Ryder, O.A. (1986). Species conservation and systematics: the dilemma of subspecies. *Trends Ecol. Evol.* **1**, 9–10.

Schneider, S., Roessli, D. & Excoffier, L. (2000). *Arlequin: A software for population genetics data analysis*. Geneva: Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Smouse, P.E., Long, J.C. & Sokal, R.R. (1986). Multiple regression and correlation extensions of the mantel test of matrix correspondence. *Syst. Zool.* **35**, 627–632.

Sunnucks, P. & Hales, D.F. (1996). Numerous transposed sequences of mitochondrial cytochrome oxidase I–II in aphids of the genus Sitobion (Hemiptera: Aphididae). *Mol. Biol. Evol.* **13**, 510–524.

Sunnucks, P., Wilson, A.C.C., Beheregaray, L.B., Zenger, K., French, J. & Taylor, A.C. (2000). SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Mol. Ecol.* **9**, 1699–1710.

Tajima, F. (1989). Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* **123**, 585–596.

Tajima, F. (1993). Measurement of DNA polymorphism. In: *Mechanisms of molecular evolution Introduction to molecular paleopopulation biology*: 37–59. Takahata, N. & Clark, A. (Eds). Mishima, Japan: Japan Scientific Societies Press.

Templeton, A.R. (1998). Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Mol. Ecol.* **7**, 381–397.

Templeton, A.R. (2008). Nested clade analysis: an extensively validated method for strong phylogeographic inference. *Mol. Ecol.* **17**, 1877–1880.

Vari, R.P. & Malabarba, L.R. (1998). Neotropical ichthyology: an overview. In *Phylogeny and classification of neotropical fishes*: 1–11. Malabarba, L.R., Reis, R.E., Vari, R.P., Lucena, Z.M. & Lucena, A.S. (Eds). Porto Alegre, Brazil: Editoria Universitaria – EDIPUCRS.

Wallace, A.R. (1876). *The geographical distribution of animals*. New York: Hafner Publishing Company.

Wares, J.P. & Cunningham, C.W. (2001). Phylogeography and historical ecology of the North Atlantic intertidal. *Evolution* **55**, 2455–2469.

Weitzman, S.H. (1978). Three new species of fishes of the genus *Nannostomus* from the Brazilian states of Para and Amazonas. *Smithson. Contrib. Zool.* **263**, 1–14.

Weitzman, S.H. & Coob, J. (1975). A revision of the South American fishes of the genus *Nannostomus* Giinther (family Lebiasinidae). *Smithson. Contrib. Zool.* **186**, 1–36.

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