Molecular Ecology (2007) **16**, 2488–2503 doi: 10.1111/j.1365-294X.2007.03338.x

Isolation by distance and Pleistocene expansion of the lowland populations of the white piranha *Serrasalmus rhombeus*

NICOLAS HUBERT,*†§ FABRICE DUPONCHELLE,*‡ JESUS NUÑEZ,*† ROSARIO RIVERA,† FRANÇOIS BONHOMME§ and JEAN-FRANÇOIS RENNO*†§

**U.R. 175 Institut de Recherche pour le Développement (IRD), GAMET, BP 5095, 361 rue JF Breton, 34196 Montpellier Cedex 05, France, †Instituto de Biología Moleculary Biotecnología, Universidad Mayor de San Andres, La Paz, Bolivia, ‡Unidad de Limnologia y Recursos Aquaticos, Universidad Mayor de San Simon, Cochabamba, Bolivia, §Laboratoire Génome, Populations, Interactions, Adaptation, CNRS-IFREMER-Université Montpellier II, UMR 5171, SMEL, 1 Quai de la Daurade, 34200 Sète, France*

Abstract

The genetic variability and distribution of Amazonian fish species have likely been influenced by major disturbance events in recent geological times. Alternatively, the great diversity of aquatic habitat in the Amazon is likely to shape ongoing gene flow and genetic diversity. In this context, complex patterns of genetic structure originating from a joint influence of historical and contemporary gene flow are to be expected. We explored the relative influence of Pleistocene climatic fluctuations and current water chemistry on the genetic structure of a piranha, *Serrasalmus rhombeus*, in the Upper Amazon by the simultaneous analysis of intron length polymorphism and mitochondrial DNA sequences. The Madeira river is well suited for that purpose as it is characterized by a great diversity of water types, the presence of one of the largest floodplain of the Amazon and the potential occurrence of two Pleistocene refuges. We found evidence of genetic structure even at a small geographical scale (less than 10 km), indicating that the floodplain is not a homogenizing factor promoting interdrainage dispersal in *S. rhombeus*. Likewise, the hierarchical genetic structure inferred was correlated to geographical distance instead of habitat characteristic. Our results also support the hypothesis that the area underwent population expansion during the last 800 000 years. In addition, a higher level of genetic diversity was found in the samples from the putative Aripuanã refuge. The present findings suggest that Pleistocene refuges contributed significantly to the colonization of the lowlands in the Upper Amazon valley during the Pleistocene.

Keywords: Characidae, museum hypothesis, Neotropics, nonequilibrium, population expansion, refuge hypothesis

Received 6 November 2006; revision received 20 February 2007; accepted 5 March 2007

Introduction

Distinguishing between contemporary and past gene flow is a task of primary importance in phylogeographical studies as it can potentially provide valuable information about the history of a species (Templeton *et al*. 1995; Knowles 2000). As ecosystems are dynamic by nature (Brown & Gibson 1983; Myers & Giller 1988), environmental disturbances often produce rapid changes of niche availability leading to species range expansions and contractions (Hewitt 2000, 2004). On the other hand, contemporary gene flow is influenced by extant landscape structure and environmental conditions regulating genetic connectivity among natural populations (Castric *et al*. 2001; Costello *et al*. 2003). Thus, heterogeneous patterns of genetic variability are expected to be commonplace throughout the species' range (e.g. Hutchinson & Templeton 1999; Turgeon & Bernatchez 2001; Garnier *et al*. 2004; Kuchta & Tan 2005). Consequently, extant

Correspondence: Nicolas Hubert, Station Méditerranéenne de l'Environnement Littoral, 1 Quai de la Daurade, 34200 Sète, France. Fax: 04 67 46 33 99; E-mail: myloplus@excite.com

genetic structure found in nature often results from the joint influence of past and contemporary environments on gene flows (e.g. Templeton *et al*. 1995).

Genetic structures of animal and plant species inhabiting the Neotropical region are not an exception to this rule (Lundberg *et al*. 1998; Saint-Paul *et al*. 2000). However, previous studies on the impact of palaeoecological events in shaping extant biological diversity have largely focused on the influence of marine incursions (Fjeldså 1994; Roy *et al*. 1997; Nores 1999; Hubert & Renno 2006) and Andean foreland dynamics (Räsänen *et al*. 1990, 1992; Patton *et al*. 1994; Hoorn *et al*. 1995; Patton & Da Silva 1998; Hubert *et al*. in press). Although obviously relevant for understanding the origin of the Amazonian biota, such studies mainly focused on the origin of taxonomic diversity during the Late Miocene (but see Aleixo 2004; Grazziotin *et al*. 2006; Noonan & Gaucher 2006). Consequently, the relative influence of Pleistocene climatic fluctuations on species demographic histories and genetic variability among natural populations remains poorly explored.

The refuge hypothesis predicts that Pleistocene climatic fluctuations limited the amount of suitable habitats for plants and animals during glacial times by the contraction of the tropical rainforest (Haffer 1969; Prance 1982; Whitemore & Prance 1987). Following this, populations outside the patch of rainforest refuges during glacial times should harbour evidence of severe bottlenecks while those from the refuges might be expected to harbour higher levels of genetic diversity. Although these predictions are logically deduced for terrestrial biota, the consequences of these environmental changes are less obvious for aquatic organisms. Phylogeographical studies recently revealed that the Upper Amazon was colonized during the last 4 million years (Myr) and only during the last 1 Myr for some tributaries (Aleixo 2004; Hubert *et al*. in press). Thus, the Pleistocene rainforest refuges should have been of great importance for the colonization of the upper part of the Amazon valley.

The Madeira, which originates in the Andean flank, is one of the largest tributaries of the Amazon. It hosts a great diversity of aquatic habitats due to the simultaneous influence of the Brazilian shield, the tertiary sediments of the lowlands and the Andes (Guyot 1993). Consequently, the tributaries of the Madeira vary according to their turbidity and nutrient contents and are considered as white, clear or black waters (Sioli 1984). Annual rainfall cycles further enhance habitat heterogeneity due to the seasonal inundation of one of the largest floodplains of the Amazon (Guyot *et al*. 1999). Thus, it could be expected that habitat heterogeneity increases genetic structure while the floodplain may act as a homogenizing factor making interbasin dispersal easier within the Upper Madeira. In addition, the palaeoecological context of the region is of further interest given at least two Pleistocene refuges were previously hypothesized in the area (Whitemore & Prance 1987). Moreover, previous phylogeographical results emphasized that the current topology of the Madeira was established during the last 2 Myr, and that the piranha species colonized the area afterwards (Hubert *et al*. in press).

In this context, we investigated the relative impact of both palaeoecological events and current habitat in shaping genetic structure and gene flow among lowland populations of the white piranha, *Serrasalmus rhombeus*. This species is abundant throughout the watershed of the Madeira (Lauzanne *et al*. 1991) and constitutes a suitable model for testing the relative influence of Pleistocene climatic changes and habitat heterogeneity in shaping gene flow and genetic diversity in Amazonian fishes. Hence, we address the following issues for *S. rhombeus* in the Madeira river: (i) according to the refuge hypothesis, populations from the putative rainforest refuges should harbour higher levels of genetic diversity than elsewhere; (ii) if the rainforest refuges prompted the colonization of the Madeira river, the oldest lineages and higher mean divergence should be found in the haplotypes sampled from the populations of the refuges; (iii) if contemporary gene flow is shaped by the environment, a relationships should be found between gene flow and some hydrological characteristics such as water chemistry. Although mitochondrial DNA (mtDNA) is useful for historical inferences due to a higher evolutionary rate than nuclear DNA (nDNA) (Birky *et al*. 1989), its maternal inheritance can lead to partial discrepancies with nDNA (e.g. Arnaud-Haond *et al*. 2003; Lemaire *et al*. 2005). Hence, we assessed simultaneously allelic diversity for several nDNA loci as well as nucleotide diversity within mtDNA sequences to test the aforementioned predictions. Finally, we discuss the results in the light of the palaeoecological and current ecological context of the region.

Materials and methods

Hydrological context and sampling

The Madeira is the second largest tributary of the Amazon (area 1.37×10^6 km^2) after the Solimões (2.24×10^6 km^2). Its headwaters are characterized by a marked annual cycle of rainy and dry seasons responsible for multipeaked floods in the Andean tributaries. The downstream pulse is stored in the Bolivian floodplain, which is one of the largest of the Amazon, with a potential flood extension of 0.15×10^6 km^2 (Guyot *et al*. 1999). Water chemistry and hydrological typology in this system is characterized by several types of water depending on the relative contribution of the Brazilian shield, the Tertiary sediments of the lowlands and the Andes (Sioli 1975; Guyot 1993). The rocks of the Brazilian shield are quartzitic and erosion is generally slow, so the rivers draining them are generally poor in sediments and crystalline, which is characteristic of the clear waters. The

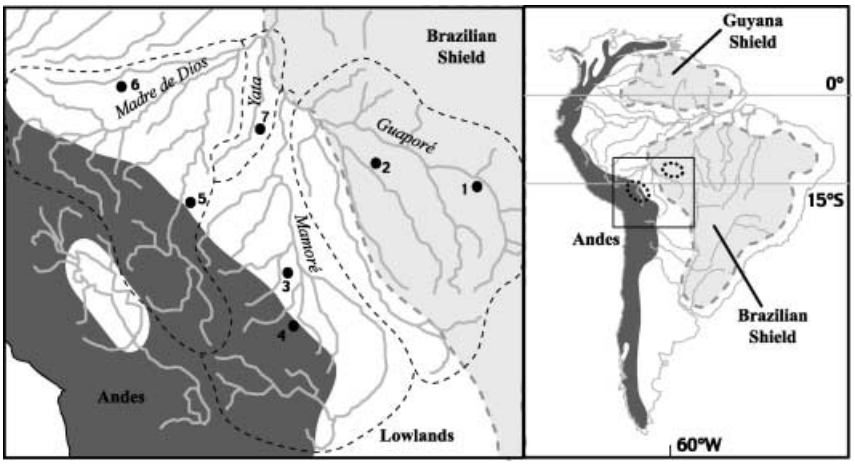

Fig. 1 Location of *Serrasalmus rhombeus* samples, geomorphologic and palaeoecological context in the Upper Madeira watershed. Circle, sampling localities within the Madeira watershed; dashed lines; Pleistocene refuges nearby the Bolivian Amazon (Whitemore & Prance 1987), namely Béni (left, Andean flanks) and Aripuanã (right, Brazilian shield).

Table 1 Locality information for *S. rhombeus* sampling sites. Population numbers correspond to Fig. 1. Sample size corresponds to the number of individual analysed for intron length polymorphism and mtDNA sequences

					Sample size		
Basin	Population		River	Specific locality	Introns	mtDNA	Latitude/longitude
Guaporé	1		Itenez	Bella Vista	30	12	13.5225S/61.5553W
	2	a	San Martin	California	20	5	13.2959S/63.5559W
		b		Bahia Sala	11	2	13.3345S/63.4421W
		c		San Joaquim	8	2	13.4167S/63.5038W
		d		Blanco	10	5	13.3375S/63.7246W
Mamoré	3	a	Isiboro	Santa Maria	16	6	15.3592S/65.0413W
		b		Santa Clara	8	2	15.3235S/65.0015W
		c		Assicussa	9	5	15.2757S/65.0201W
	4	a	Ichilo	Laguna Tambaquí	11	6	17.0171S/64.6398W
		b		Rio Izarsama	20	3	17.0408S/64.6873W
Madre de Dios	5	a	Béni	Laguna Fernande	10	3	14.2728S/67.4720W
		b		Laguna Gringo	5	3	14.3451S/67.4915W
	6	a	Manuripi	Lago Bay	17	3	11.9530S/68.6558W
		b		Manchester	10	7	11.4725S/67.9908W
	7	a	Yata	Yata 3	12	5	11.1173S/65.6668W
		b		Yata 5	46	5	11.6255S/65.6577W

Andes are young and erosion is intense. Thus, the rivers draining the Andean flank carry turbid, white waters rich in sediments (Sioli 1975). By contrast, the Tertiary sediments of the lowlands are poor in nutrients and often associated with rainforest. Thus, the water is generally crystalline, but coloured by the humic acids from the forest, which is characteristic of the black waters (Sioli 1975).

The headwaters of the Madeira can be separated into four major systems with distinct hydrological typology (Fig. 1) (Guyot 1993). The Guaporé River is the only tributary draining almost exclusively the Brazilian shield and is characterized by clear waters. By contrast, the Mamoré and Madre de Dios Rivers originate in the Andes and their main channels are characterized by white waters. However, these designations mask considerable heterogeneity, and small lowland tributaries with black water are frequently encountered in the lowlands along the main channel of the Madre de Dios and Mamoré Rivers. Finally, the Yata is a small tributary, remarkable for its central position and its black lowland waters. In addition, two Pleistocene refuges have been reported in the area (Fig. 1): the Béni located in the lowlands near the Andean flanks, and the Aripuanã in the Brazilian shield (Whitemore & Prance 1987). The Aripuanã refuge was located in the Guaporé watershed, while the Béni occurred in the headwaters of the Madre de Dios watershed.

Seven rivers were sampled between September 2002 and June 2003 (Fig. 1; Table 1) using hooks, lines and gillnets. For all the individuals sampled, 1 cm^3 of muscle were preserved in 96% ethanol solution. In the Guaporé, specimens

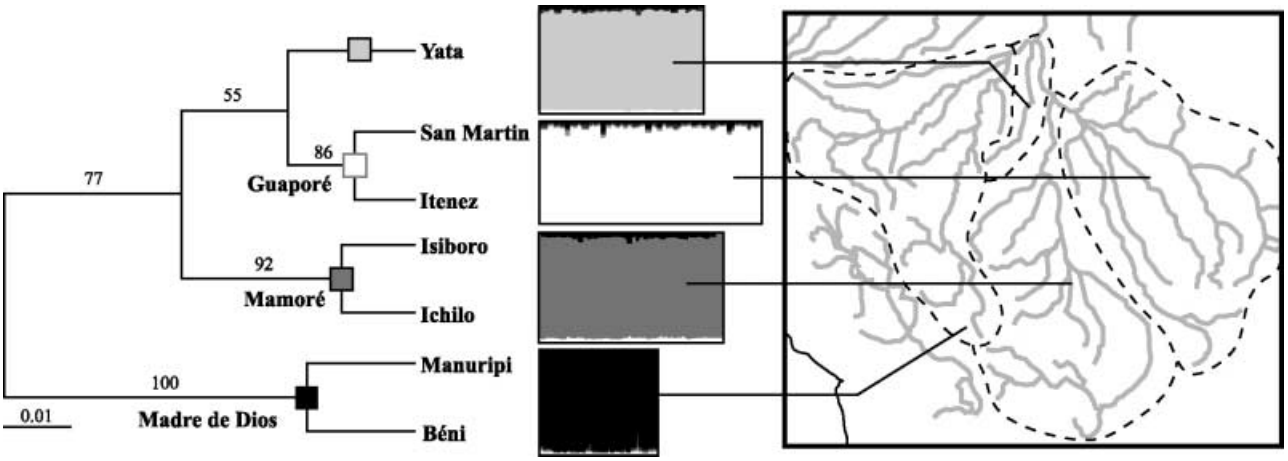

Fig. 2 UPGMA phenogram relating the seven populations sampled in the Upper Madeira watershed using Cavalli-Sforza & Edwards's (1967) chord distance. Bootstrap values are based on 100 replicates. Histograms represent the probability of each individual from the Guaporé, Yata, Mamoré and Madre de Dios originating from the tributary in which they were captured, as inferred using the assignment procedure in STRUCTURE.

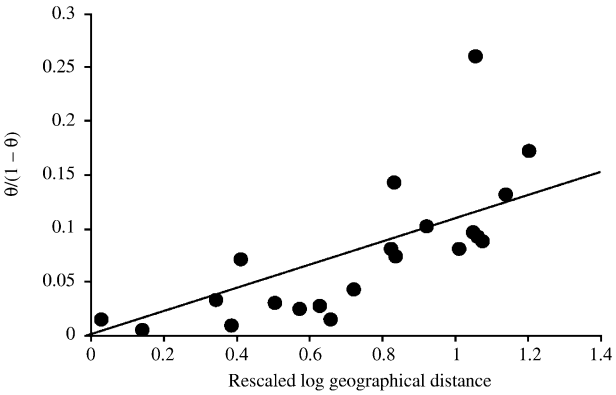

Fig. 3 Isolation-by distance relationships based on data from nuclear loci, between all population pairs, where $\theta/(1-\theta)$ was regressed over the logarithm of geographical distance through the main channel of the rivers following Rousset (1997). The regression of the isolation by distance relationship is $\theta/(1-\theta) = 0.113$ (rescaled log geographical distance), with $r^2 = 0.52$ and Mantel's (1967) statistic $Z = 11.27$ ($P = 0.0005$).

from clear water sites in the headwater (Fig. 1,) and the lower course (Figs 1 and 2) were sampled. In the Mamoré, specimens from two white water tributaries originating in the Andean flank were sampled (Figs 1, 3 and 4) while both a white water (Figs 1 and 5) and clear water tributary (Figs 1 and 6) were prospected in the Madre de Dios. A single black water site was sampled in the Yata (Figs 1 and 7).

Mitochondrial and nuclear markers

Genomic DNA was isolated from ethanol-preserved muscle tissues with the DNAeasy Tissue Kit (QIAGEN). Sequences of 980 bp from the mitochondrial control region, including the 3′ flanking tRNA genes (tRNA Thr and tRNA Pro), were amplified using the primer CR22U: 5′-TGGTTTAGTA-CATATTATGCAT-3′ (Hubert *et al*. in press) and F-12R: 5′-GTCAGGACCATGCCTTTGTG-3′ (Sivasundar *et al*. 2001). PCRs were performed in 50 μL volumes including 13.5 μL of genomic DNA, 3 U of *Taq* DNA polymerase, 5 μL of *Taq* 10x buffer, 3 μL of $MgCl_2$ (25 mM), 4 μL of dNTP (5 mM) and 3 μL of each primer (10 μM). For the control region, PCR conditions were as follows: 94 °C (5 min), 10 cycles of 94 °C (1 min), 66 to 56 °C decreasing of 1 °C per cycle (1 min, 30 s), 72 °C (2 min), 25 cycles of 94 °C (1 min), 56 °C (1 min, 30 s), 72 °C (2 min), followed by 72 °C (5 min). The nucleotide sequence data have been deposited in GenBank (Accession nos EF078921–EF078961).

The technique of amplification based on exon-primed intron-crossing polymerase chain reaction (EPIC-PCR) of Lessa (1992) and Palumbi & Baker (1994) was used for the detection of polymorphic nuclear introns. This technique consists in amplifying introns by using primers designed in the flanking exons to get a cross-amplification of the targeted intron. Hence, by using primers designed from the highly conserved exons, cross-amplifications are easiest and PCR artefacts such as null alleles are expected to be less frequent (Côrte-Real *et al*. 1994; Bierne *et al*. 2000). Since a substantial number of genes from the nuclear genome belong to multigenic families, various loci might be scored for a given pair of primers (Atarhouch *et al*. 2003; Hassan *et al*. 2003; Borsa *et al*. 2004). Recently, several intron loci were scored among the *Serrasalmus* species of the Madeira for molecular systematic purposes with several polymorphic loci detected for *S. rhombeus* in the area (Hubert *et al*. 2006). Following this previous result, seven loci were scored using the following primers: GPD2F/3R which amplifies intron 2 of the glyceraldehyde

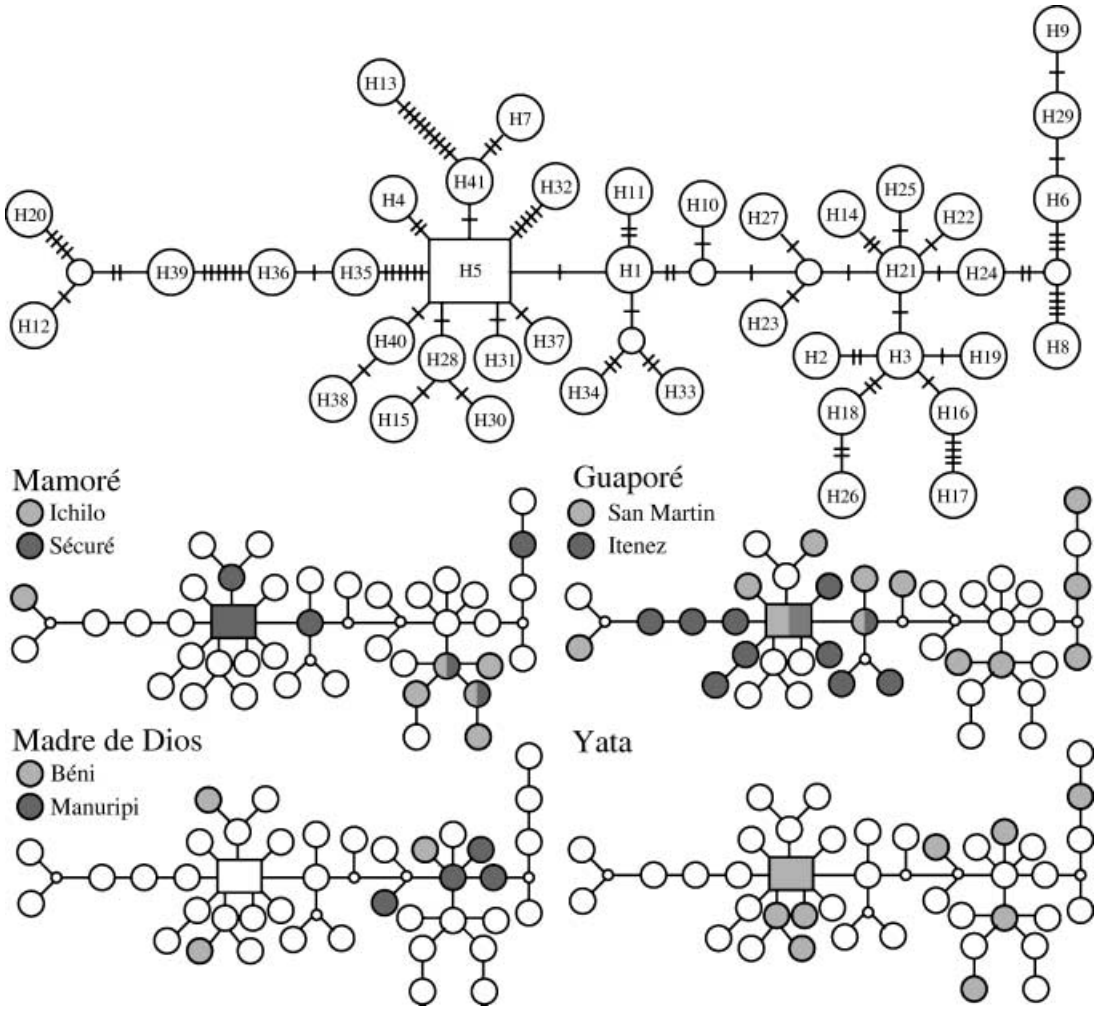

Fig. 4 Maximum parsimony network and haplotype distributions across the four mains tributaries of the Upper Madeira watershed. Size of the circles is not proportional to the haplotype frequency. Solid lines represent mutations. Haplotype H5 is the most likely ancestral haplotype. For each river, the haplotypes sampled are in grey, while missing haplotypes are in white.

3-phosphate dehydrogenase gene (Hassan *et al*. 2002); RPEX1F/2R which amplifies intron 1 of the S7 ribosomal protein (Chow & Hazama 1998); GH5F/6R which amplifies intron 5 of the Growth Hormone (Hassan *et al*. 2002); Am2b2F/3R which amplifies intron 2 of the Alpha Amylase (Hassan *et al*. 2002); PmOPSIF/R, designed from cDNA sequences of the Opsin (Bierne *et al*. 2000) and GnRH1F/R which amplify intron 1 of the Gonadotropin-releasing hormone 3 (Hassan *et al*. 2002). These loci were scored using the same protocols as in Hubert *et al*. (2006).

Analyses of nuclear population structure

Within-population genetic diversity was quantified as the number of alleles per locus for each population, mean number of alleles per population, proportion of polymorphic loci per population at the 0.99 level, observed heterozygosity (H_O) and the unbiased estimate of expected heterozygosity H_{NB} (Nei 1987). Deviation from Hardy–Weinberg equilibrium (HWE) in each river was tested using an approximation of an exact test based on a Markov chain iteration (Guo & Thompson 1992). The extent of population differentiation was quantified by Weir & Cockerham's (1984) estimator of F_{ST} (θ) and potential differences in allelic distributions were tested using an exact test (Raymond & Rousset 1995). The tests of HWE and allelic distribution were performed using GENEPOP 3.1 (Raymond & Rousset 1995). Critical significance levels for multiple testing were corrected following the sequential Bonferroni procedure (Rice 1989).

We assessed hierarchical population structure by computing the Cavalli-Sforza & Edward's (1967) chord distance (D_{CE}) and constructed a population phenogram using the UPGMA algorithm (Sneath & Sokal 1973). D_{CE} was preferred,

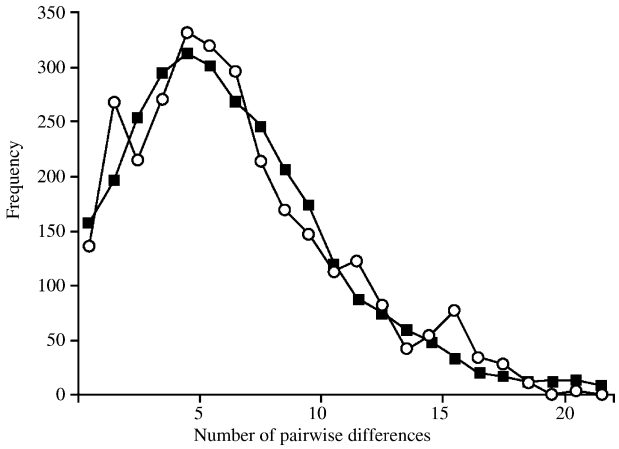

Fig. 5 Empirical mismatch distribution (open circles) compared to the expected distribution under a sudden expansion model (solid squares). The expansion model cannot be rejected ($P = 0.98$; $r = 0.005$).

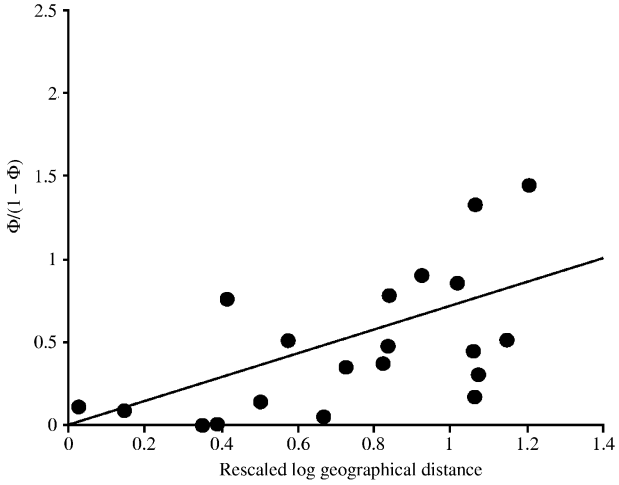

Fig. 6 Isolation-by distance relationship based on mtDNA data, between all populations pairs, where $\Phi/(1-\Phi)$ was regressed over the logarithm of geographical distance through the main channel of the rivers. The regression of the isolation by distance relationship is $\Phi/(1-\Phi) = 0.748$ (rescaled log geographical distance), with $r^2 = 0.15$ and Mantel's (1967) statistic $Z = 81.76$ ($P = 0.04$).

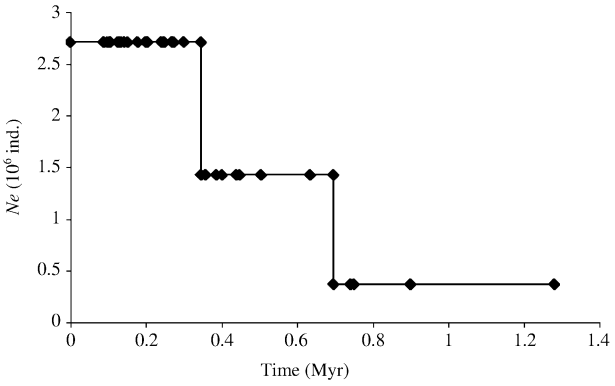

Fig. 7 Generalized skyline plot for the control region genealogy with ε = 0.0019 (the AIC estimate). The calibration relies on a rate of 0.0058 substitutions per site per Myr. See text for justification. The UPGMA phenogram was computed using the model TVM + I + Γ with the following parameters: base frequencies A = 0.31, G = 0.22, C = 0.16, T = 0.31, mutation rates A-C = 2.25, A-*G* = 13.59, A-T = 2.91, C-*G* = 0.00, C-T = 13.60, G-T = 1.00, proportion of invariable sites = 0.84, gamma distribution shape parameter = 0.86.

as it makes no assumption about population size and mutation rates. Bootstrap values were based on 100 replicates computed using PHYLIP 3.5 (Felsenstein 1993). To assess whether populations followed a pattern of isolation by distance in two dimensions, we plotted the pairwise genetic differentiation of the populations estimated by $\theta/(1-\theta)$ against the logarithm of the geographical distance through the main channel (Rousset 1997) and tested the relationship using the Mantel test (Mantel 1967) as implemented in ARLEQUIN 2.0 (Schneider *et al*. 2000). We further investigated the potential occurrence of population admixture and the presence of migrants in the samples by using the model-based clustering method described by Pritchard *et al*. (2000) to assign individuals to populations using multilocus genotype data, as implemented in the software STRUCTURE. This algorithm will compute the posterior probability of each individual being sampled from each of the populations defined a priori. Thus, it is possible to assess simultaneously the geographical origin of each individual and the relative contribution of migrants to local gene pools.

Finally, since populations of *S. rhombeus* in the Upper Madeira may originate from either the Aripuanã refuge from the Brazilian shield or the Béni refuge from the Andean flank (Whitemore & Prance 1987), we tested whether the observed pattern of genetic differentiation among drainages could be explained by the colonization from one of these putative refuges. Slatkin (1993) assessed two distinct models of colonization and predicted the expected pattern of genetic differentiation under conditions departing from migration/drift equilibrium. A gradual stepwise expansion is expected to result in greater divergence among the earlier founded populations than among the more recently founded ones (Good's model in Slatkin 1993). Thus, we tested if the colonization of the Upper Madeira originated from either the Béni or the Aripuanã refuge by plotting the estimated amount of gene flow from pairwise F_{ST}, namely $M = 1/4(1/F_{ST} - 1)$, against the logarithm of the geographical distance of the closest population from the putative origin (Slatkin 1993). If a stepwise and directional colonization had occurred, then a positive correlation between M and the logarithm of the geographical distance from the centre of origin will be observed.

Analyses of mitochondrial population structure and demography

Multiple alignments of the control region were performed using CLUSTAL W (Thompson *et al*. 1993), with three different opening to extending cost ratios: opening cost = 5 and extending cost = 4; opening cost = 15 and extending cost = 6 (default setting); opening cost = 20 and extending cost = 8, in order to detect potential alignment ambiguous sites (Gatesy *et al*. 1994). The genealogy of the control region was constructed following the statistical parsimony method of Templeton *et al*. (1992). This method estimates the maximum number of differences that result from single substitutions with a statistical confidence of 95% and implements calculations of root probabilities and relative haplotype ages following the predictions from the coalescent theory (Crandall & Templeton 1993). These calculations were performed using the TCS software (Clement *et al*. 2000). Alternative ambiguous connections resulting from homoplastic mutations were resolved by comparison with the maximum-parsimony (MP) tree obtained by heuristic searches as implemented in PAUP* version 4.0b10 (Swofford 2002).

Mean number of private and overall haplotypes, mean pairwise differences, haplotype diversity (h) and nucleotide diversity (π; Nei 1987) were computed for each population while Tajima's (1989) D test of neutrality was performed for the overall data set. The analysis of molecular variance (AMOVA; Excoffier *et al*. 1992) provided an estimate of the distribution of nucleotide diversity at three levels of population subdivision: among basins (CT); among rivers within basins (SC) and among individuals within rivers (ST). The correlation of alleles at each of the three hierarchical levels was assessed using the Φ-statistics (Excoffier *et al*. 1992) tested by 1000 permutations of individuals. In addition, Φ-statistics between pairwise comparisons were used to test for a potential isolation by distance pattern in the data by plotting $\Phi(1 - \Phi)$ against the logarithm of geographical distance through the main channel of the river. All the parameters of molecular diversity, Mantel tests, and AMOVA were computed using ARLEQUIN 2.0 (Schneider *et al*. 2000).

Demographic history was assessed by computing the pairwise mismatch distribution (Rogers & Harpending 1992). We tested for population expansion using the parameter of Roger (1995) model of sudden population expansion (r) as implemented in ARLEQUIN. We simulated the expected mismatch distribution in an expanding population following the coalescent approach, outlined by Slatkin & Hudson (1991) and Rogers & Harpending (1992), and further tested for significant fit of the model by the permutation procedure implemented in ARLEQUIN. Outputs of the mismatch distribution were then compared with those of the analysis of the generalized skyline plots (GSP; Pybus *et al*. 2000, 2001; Strimmer & Pybus 2001) as implemented in GENIE 3.0 (available at http://evolve.zoo.ox.ac.uk/software/genie). This method provides an estimate of the effective population size at different levels of a given genealogy using a coalescent approach. For this purpose, an ultrametric tree of the mtDNA control region sequences was constructed with the UPGMA algorithm (Sneath & Sokal 1973) as implemented in PAUP* 4.0b10 and using distances computed from the optimal model, as identified by the Akaike information criterion (AIC) implemented in MODELTEST 3.7 (Posada & Crandall 1998). The genealogy obtained was used for the GSP analysis, and a molecular calibration was used for rescaling the graphical output.

Results

Intron length polymorphism and population structure

Polymorphism at the seven intronic loci was previously detected within *Serrasalmus rhombeus*, and one to two polymorphic loci were obtained per primer pair with a diversity ranging from two to four alleles per locus, as previously reported (Hubert *et al*. 2006). The proportion of polymorphic loci per population ranged between 0.71 in the Madre de Dios samples (Béni and Manuripi) and 1 in the San Martin whereas the mean number of alleles ranged between 2.00 and 2.43 among populations (Table 2). Gene diversity varied between 0.02 and 0.73 per locus, whereas expected gene diversity (H_{NB}) ranged between 0.27 and 0.37 per population, with the eastern populations from the Guaporé showing the highest values (Table 2). The exact test of HWE showed a significant trend toward heterozygote deficiency in four populations (San Martin, Béni, Manuripi and Yata; Table 2) due to one locus in the Manuripi (GH5-1) and the Yata (GH5-2) and two loci in the San Martin (RPEX1-1 and RPEX1-3) and the Béni (GH5-1 and GnRH1-1). As several localities were sampled for each river, we checked if heterozygote deficiencies found in the San Martin, Manuripi, Béni and Yata might be attributed to small-scale geographical structure (less than 10 km). The exact test of HWE showed that none of the local samples harboured significant departure from HWE (Table 3).

The genetic differentiation was high between populations from different drainages with pairwise interbasin F_{ST} values ranging from 0.01 between the San Martin and the Yata to 0.207 between the Béni and Ichilo. By contrast, pairwise intrabasin F_{ST} values ranged from 0.01, between the San Martin and Itenez rivers from the Guaporé basin, and 0.026 between the Béni and Manuripi riversfrom the Madre de Dios basin (Table 4). Likewise, exact tests of allele distributions revealed that all pairwise comparisons involving localities from distinct drainages were highly significant, unlike pairwise comparisons within the same drainage (Table 4). This emphasized that individuals found in different rivers compose genetically distinct populations.

Table 2 Allelic frequencies at seven nuclear introns in seven locations from the Upper Madeira watershed. *N*, sample size; H_{NB}, unbiased expected heterozygosity (Nei 1987); H_O, observed heterozygosity; *p*, *P* values of the HWE exact test. $P_{(0.99)}$, proportion of polymorphic loci at the 0.99 level. NS, nonsignificant after sequential Bonferroni correction, * significant after sequential Bonferroni correction

	Populations						
	1-Itenez	2-San Martin	3-Isiboro	4-Ichilo	5-Béni	6-Manuripi	7-Yata
N	30	49	33	31	15	27	58
RPEX1-1							
1020	0.27	0.18	0.14	0.08	—	—	0.10
955	0.08	0.12	0.12	0.10	—	—	0.02
995	0.65	0.69	0.74	0.82	1.00	1.00	0.88
H_{NB}	0.51	0.47	0.42	0.31	—	—	0.22
H_O	0.47	0.33	0.36	0.26	—	—	0.21
p	0.375^{NS}	0.007*	0.124^{NS}	0.101^{NS}	—	—	0.590^{NS}
RPEX1-3							
745	0.98	0.97	1.00	1.00	1.00	1.00	1.00
755	0.02	0.03	—	—	—	—	—
H_{NB}	0.03	0.06	—	—	—	—	—
H_O	0.03	0.02	—	—	—	—	—
p	—	0.031*	—	—	—	—	—
GH5-1							
1120	0.32	0.23	0.17	0.21	0.70	0.52	0.33
1105	0.68	0.77	0.83	0.79	0.30	0.48	0.67
H_{NB}	0.44	0.36	0.28	0.34	0.44	0.51	0.44
H_O	0.23	0.35	0.33	0.23	0.07	0.07	0.41
p	0.014^{NS}	0.708^{NS}	0.559^{NS}	0.092^{NS}	0.002*	0.001*	0.765^{NS}
GH5-2							
475	1.00	0.99	0.92	0.94	0.90	0.94	0.92
480	—	0.01	0.08	0.06	0.10	0.06	0.08
H_{NB}	—	0.02	0.14	0.12	0.19	0.11	0.14
H_O	—	0.02	0.15	0.13	0.07	0.11	0.05
p	—	—	1^{NS}	1^{NS}	0.103^{NS}	1^{NS}	0.001*
Am2b2–3							
435	0.23	0.33	0.15	0.19	0.03	0.11	0.27
440	0.77	0.67	0.85	0.81	0.97	0.89	0.73
H_{NB}	0.36	0.44	0.26	0.32	0.07	0.20	0.39
H_O	0.20	0.29	0.12	0.26	0.07	0.15	0.33
p	0.025^{NS}	0.020^{NS}	0.012^{NS}	0.291^{NS}	—	0.267^{NS}	0.198^{NS}
PmOPSI-1							
390	0.65	0.54	0.38	0.15	0.67	0.39	0.54
410	0.35	0.46	0.62	0.85	0.33	0.61	0.46
H_{NB}	0.46	0.50	0.48	0.25	0.46	0.48	0.50
H_O	0.50	0.51	0.45	0.23	0.40	0.33	0.33
p	1^{NS}	0.708^{NS}	1^{NS}	0.491^{NS}	1^{NS}	0.123^{NS}	0.009^{NS}
GnRH1-1							
1035	0.25	0.38	0.12	0.15	0.23	0.26	0.27
1045	0.35	0.21	0.32	0.37	0.23	0.20	0.22
1065	0.10	0.13	0.29	0.32	0.23	0.15	0.16
1085	0.30	0.28	0.27	0.16	0.30	0.39	0.36
H_{NB}	0.73	0.73	0.74	0.72	0.77	0.73	0.73
H_O	0.63	0.55	0.70	0.74	0.47	0.59	0.67
p	0.326^{NS}	0.094^{NS}	0.107^{NS}	0.273^{NS}	0.006*	0.422^{NS}	0.131^{NS}
All loci							
H_{NB}	0.36	0.37	0.33	0.30	0.27	0.29	0.35
H_O	0.30	0.30	0.30	0.26	0.15	0.18	0.29
p	0.022^{NS}	0.003*	0.096^{NS}	0.11^{NS}	0.001*	0.003*	0.001*
$P_{(0.99)}$	0.86	1.00	0.86	0.86	0.71	0.71	0.86
Mean number of alleles	2.29	2.43	2.29	2.29	2.00	2.00	2.29

Table 3 HWE exact tests for local sampling sites from the San Martin, Manuripi, Béni and Yata rivers. *N*, sample size; H_{NB}, unbiased expected heterozygosity (Nei 1987); H_O, observed heterozygosity; *P*, *P* values of the HW exact test (Weir 1990). NS, nonsignificant after sequential Bonferroni correction, * significant after sequential Bonferroni correction

	N	H_{NB}	H_O	*P*
2-San Martin				
2a-California	20	0.36	0.33	0.20^{NS}
2b-Bahia Sala	11	0.42	0.26	0.11^{NS}
2c-San Joaquim	8	0.35	0.32	0.46^{NS}
2d-Blanco	10	0.34	0.25	0.13^{NS}
5-Manuripi				
5a-Lago Bay	17	0.27	0.21	0.18^{NS}
5b-Manchester	10	0.21	0.15	0.12^{NS}
6-Béni				
6a-Laguna Fernande	10	0.27	0.17	0.01^{NS}
6b-Laguna Gringo	5	0.29	0.11	0.06^{NS}
7-Yata				
7a-Yata 3	12	0.35	0.28	0.85^{NS}
7b-Yata 5	46	0.34	0.29	0.01^{NS}

In addition, the lowest F_{ST} values were found between geographically proximate locations whereas the highest values were found in the most distant locations (Table 4).

The population phenogram inferred from D_{CE} distances further illustrated the population grouping by drainage and emphasized that the hierarchical clustering was related to geographical distance since tributaries from the Mamoré (Isiboro and Ichilo) were more closely related to the neighbouring tributaries from the Guaporé (Fig. 2). Generally, the clusters were well supported. A significant correlation was found between genetic differentiation as measured by $F_{ST}/(1 - F_{ST})$, and the logarithm of the geographical distance indicating that the present structure follows an isolation-by-distance pattern (Fig. 3). The individual assignment obtained using STRUCTURE indicated that each individual was assigned a posterior probability greater than 0.8 of originating from the drainage in which they were sampled (Fig. 2). This result emphasized that no or little population admixture occurred and confirmed the strong genetic structure in relation with the geographical distance.

The geographical pattern of differentiation did not fit Good's model of stepwise expansion. The correlations between gene flow (*M*) and the logarithm of geographical distance (Log *D*) was not significant, neither for an expansion originating in the San Martin ($M = -0.0003 \text{ Log } D + 0.88$; Mantel test; $Z = 8499.3$; $P = 0.83$), nor in the Ichilo ($M = 0.0003 \text{ Log } D + 0.48$; Mantel test; $Z = 13788.8$; $P = 0.11$) nor in the Béni ($M = 0.0004 \text{ Log } D + 0.42$; Mantel test; $Z = 12222.4$; $P = 0.07$) despite that a positive correlation between *M* and Log*D* was found for the Béni and Ichilo as expected under the Good's model.

Mitochondrial DNA variability and phylogeographical pattern

The three alignment parameter sets provided similar results indicating that the control region sequences did not contain any alignment ambiguous sites. The alignment of the 75 samples included 963 bp out of which three displayed an insertion-deletion of 1 bp. The parsimony network inferred included 91 steps, and the number of steps connecting haplotypes ranged from one to 10 steps (Fig. 4). Haplotype

Table 4 Estimates of pairwise F_{ST} (Weir & Cockerham 1984) and *P* values of the exact test of allelic distribution among populations (Raymond & Rousset 1995) (below diagonal), and stream channel distance in kilometres (above diagonal). NS, nonsignificant after sequential Bonferroni correction, * significant after sequential Bonferroni correction

	Guaporé			Mamoré		Madre de Dios	
	1-San Martin	2-Itenez	7-Yata	3-Ichilo	4-Isiboro	5-Béni	6-Manuripi
San Martin	—	570	730	1430	700	1140	1130
Itenez	$\theta = 0.004$ $P = 0.482^{NS}$	—	960	1550	1020	1250	1360
Yata	$\theta = 0.010$ $P = 0.002$*	$\theta = 0.015$ $P = 0.01$*	—	1450	820	750	880
Ichilo	$\theta = 0.084$ $P < 0.001$*	$\theta = 0.116$ $P < 0.001$*	$\theta = 0.081$ $P < 0.001$*	—	510	1430	1420
Isiboro	$\theta = 0.032$ $P < 0.001$*	$\theta = 0.042$ $P < 0.001$*	$\theta = 0.031$ $P < 0.001$*	$\theta = 0.015$ $P = 0.246^{NS}$	—	1650	1140
Béni	$\theta = 0.125$ $P < 0.001$*	$\theta = 0.093$ $P < 0.001$*	$\theta = 0.067$ $P < 0.001$*	$\theta = 0.207$ $P < 0.001$*	$\theta = 0.147$ $P < 0.001$*	—	930
Manuripi	$\theta = 0.075$ $P < 0.001$*	$\theta = 0.077$ $P < 0.001$*	$\theta = 0.025$ $P = 0.002$*	$\theta = 0.090$ $P < 0.001$*	$\theta = 0.069$ $P < 0.001$*	$\theta = 0.026$ $P = 0.152^{NS}$	—

Table 5 Frequency and distribution of the control region haplotypes. Shared haplotypes are in bold

	Guaporé		Mamoré		Madré de dios			
Haplotype	1-Itenez	2-San Martin	3-Isiboro	4-Ichilo	5-Béni	6-Manuripi	7-Yata	Total
1	1	1	2					4
2		1						1
3		2	7	1			1	11
4		1						1
5	2	2	1				2	7
6		1						1
7		1						1
8		1						1
9		1						1
10		1						1
11		1						1
12		1						1
13					4			4
14					1			1
15					1			1
16			1	2				3
17				1				1
18				1				1
19				3				3
20				1				1
21						7		7
22						1		1
23						1		1
24						1		1
25							1	1
26							1	1
27							1	1
28							1	1
29			2				1	3
30							1	1
31							1	1
32	1							1
33	1							1
34	1							1
35	1							1
36	1							1
37	1							1
38	1							1
39	1							1
40	1							1
41			1					1
total	12	14	14	9	6	10	10	75

sharing was very low with only five of the 41 haplotypes (12.2%) existing in more than one basin or population (Fig. 4; Table 5). Only haplotype 1 was shared among San Martín, Iténez and Isiboro, haplotype 3 among San Martín, Ichilo, Isiboro and Yata; the ancestral haplotype 5 (the root of the haplotype network; Fig. 4) among San Martín, Iténez, Isiboro and Yata, haplotype 16 in the Ichilo and Isiboro, and haplotype 29 in the Isiboro and Yata. It is worth noting that only ancestral haplotypes at nodes connecting several lineages were shared while all the haplotypes at the tips were found in a single drainage (Fig. 4; Table 5).

Among this set of 41 haplotypes, the mean pairwise nucleotide difference among basins was 6.00 ± 2.89 (Table 6). Within basins, there were on average 12.75 haplotypes (range = 7–23), out of which 9.25 on average were private (range = 6–18), with a mean pairwise difference of 5.39 ± 1.02. Within populations, the mean number of haplotypes was 7.28 (range = 3–12) with on average 5.14

Table 6 Haplotype diversity and divergences among basins, populations and all samples, and results of the analysis of molecular variance as implemented in ARLEQUIN (Schneider *et al*. 2000). The percentage column indicates the amount of total variance explained by each of the hierarchical levels. Φ-statistics are analogous to Wright's *F*-statistics and identify the correlation of alleles at each of the hierarchical levels. *P* values were obtained by comparison of observed values with those generated by random permutation; NS, nonsignificant ($P > 0.05$); *, significant ($P < 0.001$)

	Mean number of haplotypes	Mean number of private haplotypes	Mean pairwise differences	Variance component	
				%	Φ
Among basins, Φ_{CT}	44	—	6.00 ± 2.89	0	0^{NS}
Among rivers, within basins, Φ_{SC}	12.75 (7–23)	9.25 (6–18)	5.39 ± 1.02	30.44	0.30*
Among individuals, within rivers, Φ_{ST}	7.28 (3–12)	5.14 (1–9)	4.67 ± 2.12	69.66	0.28*

Table 7 Haplotype diversity (*h*), nucleotide diversity (π) and mean pairwise differences among the samples of the present study

	N	*h*	π	Mean pairwise differences
Guaporé				
Itenez (1)	12	0.99 ± 0.04	0.006 ± 0.004	5.80 ± 2.99
San Martin (2)	14	0.98 ± 0.04	0.006 ± 0.004	6.09 ± 3.08
Mamoré				
Isiboro (3)	14	0.75 ± 0.11	0.003 ± 0.002	3.10 ± 1.71
Ichilo (4)	9	0.89 ± 0.09	0.005 ± 0.003	4.78 ± 2.58
Madre de Dios				
Béni (5)	6	0.60 ± 0.22	0.007 ± 0.005	7.13 ± 3.90
Manuripi (6)	10	0.53 ± 0.18	0.001 ± 0.0007	0.80 ± 0.63
Yata (7)	10	0.98 ± 0.05	0.005 ± 0.003	5.00 ± 2.66

(range = 1–9) private haplotypes and a mean pairwise nucleotide difference reaching 4.67 ± 2.12. The AMOVA indicated that almost 70% of the overall variance was due to individual differences within rivers compared with 30% due to river differences within basin (Table 6). The allelic correlation was significant within basins and within rivers, providing evidence of a strong genetic structure at these two levels ($\Phi_{SC} = 0.30$, $P < 0.001$; $\Phi_{ST} = 0.28$, $P < 0.001$). Finally, the diversity indices indicated that the eastern population from the Guaporé harboured a higher haplotypic diversity ($h = 0.98 \pm 0.04$ in San Martín and 0.99 ± 0.04 in Iténez; Table 7), higher nucleotide diversity ($\pi = 0.006 \pm 0.004$ in San Martín and Iténez; Table 7) and higher mean number of pairwise differences (5.80 ± 2.99 in San Martín and 6.09 ± 3.08 in Iténez).

The mismatch distribution displayed a unique mode at five mutational steps (Fig. 5). The observed pattern matched almost perfectly the simulated distribution following a sudden expansion model ($P = 0.98$; $r = 0.005$), and Tajima's *D* test provided evidence of significant departure from neutrality ($D = -1.62$; $P = 0.04$). The population genetic differentiation followed a pattern of isolation by distance as the correlation between $\Phi/(1 - \Phi)$ and the geographical distance was found to be significant (Fig. 6; Mantel test, $Z = 81.76$, $P = 0.04$). However, since the Φ-statistic includes the number of mutational steps between haplotypes, such a pattern may arise either because of an increased number of mutations in relation with geographical distance, or a gradual variation in haplotype frequencies. The regression was further tested with pairwise Φ-statistics, computed using the haplotype frequencies only, and was not significant (Mantel test; $Z = 23.68$; $P = 0.09$). This result emphasized that isolation by distance was related to an increase in the haplotype divergence among populations related to the geographical distance.

The AIC indicated that the TVM + I + Γ (Rodríguez *et al*. 1990) better fitted the mtDNA control region than other models and was used for constructing the UPGMA ultrametric tree (Fig. 7). Demographic inference from the GSP based on the UPGMA phenogram was consistent with the mismatch distribution (Fig. 7). Phylogenetic evidence suggests that the evolution of the control region sequences among *Serrasalmus* is clock-like following a substitution rate of 0.0058 substitutions per site per Myr (Hubert *et al*. in press). Following the inference obtained with the GSP, the expansion occurred during the Pleistocene climatic fluctuations between 800 000 and 400 000 years. In addition, the GSP analysis provided an estimate of a maternal effective size of 370 000 females at 700 000 years and 2.7 million females at the end of the expansion (Fig. 7).

Discussion

Genetic structure and habitat

The predicted consequences of the occurrence of varied water typologies and of a large floodplain were not verified. The high genetic structure observed even at small geographical scales indicates that the floodplain does not act as a homogenizing factor making interbasin dispersal easier within the Upper Madeira. However, this may not be a generalized trend of the fish communities in the area. *Serrasalmus* species reproduce during the wet season (Llamas

& Godinho 1996; Agostinho 2003), depositing adhesive eggs in the allochthonous vegetation of the flooded forest (Rizzo & Bazzoli 1991; Rizzo *et al*. 2002; Guimarães & Quagio-Grassiotto 2004) and are one of the very few Characiform species to provide parental care (Géry 1977; Uetanabaro *et al*. 1993; Llamas & Godinho 1996; Agostinho 2003). Thus, dispersal during the earliest stages of the life cycle is very limited since larvae and juveniles stay in the flooded forest instead of moving to the main channel of the river as do other species (Barthem & Gouding 1997; Bialetzki *et al*. 2002). The present genetic structure seems to corroborate this trend; moreover, a similar genetic pattern was previously observed in another care-providing species, *Cichla monoculus*, in the Upper Madeira (Renno *et al*. 2006).

Previous results at the community level emphasized that *Serrasalmus rhombeus* is a typical member of clear water communities, having less affinity for white water than other piranha species (Saint-Paul *et al*. 2000). Following this, clear water rivers may be expected to harbour populations of larger size and higher genetic diversity than white water rivers. Surprisingly, we found no evidence of a genetic clustering related to water type in the Upper Madeira watershed. For instance, the hierarchical genetic structure showed that individuals from the clear waters of the Manuripi were genetically more closely related to individuals from the white waters of the Béni than to individuals from the black waters of the neighbouring Yata tributary. Likewise, differences in genetic diversity across the Madeira populations were not related to water type as the samples from the clear waters of the Manuripi harboured a lower genetic diversity than the samples from the white waters of the Béni. Although the migration rate seems to be at least very low, we cannot rule out the possibility that limited statistical power may have hidden some preferential patterns of dispersal depending on water turbidity and chemistry.

Isolation by distance and nonequilibrium populations

This study provided new evidences about the role of historical events in shaping extant genetic diversity in *S. rhombeus*. The intron length polymorphism indicated that the populations of *S. rhombeus* were genetically highly structured and that individuals sampled from different drainages constituted genetically distinct populations. This result was confirmed by the distribution of mtDNA haplotypes, as only the oldest haplotypes were shared between drainages, suggesting that haplotype sharing was only due to historical connection among populations. However, the slope of the isolation-by-distance regression for nDNA was low, suggesting high levels of gene flow among populations (Rousset 1997; Hutchinson & Templeton 1999). By contrast, exact tests of HWE, and assignment of individuals to populations, indicated that no population admixture may be detected in the samples suggesting strong genetic structuring and low levels of gene flow, even at small geographical scales. Likewise, the lack of significant correlation between genetic and geographical distance for mtDNA, when haplotype divergence was not taken into account, further suggested that the present pattern was shaped by historical connection rather than ongoing gene flow. Additionally, the strong heterogeneity in genetic diversity found in mtDNA among populations and to some extent in the intron length polymorphism suggested that populations of *S. rhombeus* are not at migration/drift equilibrium. In this context, the levels of intrapopulation diversity may be more reflective of the allelic diversity in founding populations of differential abundance, rather than contemporary census (Castric *et al*. 2001). Furthermore, the high maternal effective size inferred from mtDNA sequences suggests a very large effective population size, which would contribute to the maintenance of the genetic variability established through the colonization of the area (Austerlitz *et al*. 1997).

The way populations are founded may be responsible for the establishment and maintenance of isolation-by-distance pattern in nonequilibrium populations (Slatkin 1993; Turgeon & Bernatchez 2001). For instance, we cannot discard this scenario as a population expansion was evidenced by the mismatch distribution and GSP analysis. Unfortunately, we failed to detect a signature of progressive recolonization in nDNA data, either from the Béni or the Aripuanã refuge, despite increased *M* with the geographical distance from the Béni, in agreement with the expectation of Good's model. However, given the relatively small number of comparisons we used to apply it, we cannot rule out the possibility that limited statistical power was responsible for this outcome. Alternatively, a radiation model can apply here as it can potentially produce a pattern of isolation by distance providing that the time of the radiation was long enough (Slatkin 1993). However, indices of genetic diversity from both nuclear and mitochondrial DNA across populations argue that rivers from the Guaporé drainage harbour higher levels of genetic diversity, suggesting that the populations from this drainage were founded earlier than elsewhere. This finding suggests that the populations from the Guaporé were more stable and larger than in other drainages during the Pleistocene climatic fluctuations, an observation consistent with a significant influence of the Aripuanã refuge in the area.

Biogeographical implications

Phylogenetic studies of Amazonian biota previously emphasized that the populations and species from the lowlands of the upper Amazon valley originated more recently than others in the Brazilian and Guyana shields (Vari 1991; Da Silva & Patton 1993, 1998; Hall & Harvey

2002; Aleixo 2004; Hubert *et al*. in press). This finding was previously attributed to the Miocene marine incursion that occurred nearly 5 million years ago and provided considerable species range contractions for terrestrial and freshwater biota (Fjeldså 1994; Roy *et al*. 1997; Nores 1999). This isostatic change prompted a marine highstand of 100 m that considerably reduced the area of suitable habitat for freshwater fishes. Consequently, the colonization of the lowlands from the Upper Amazon valley occurred after the marine retreat during the last 4 Myr.

Recent phylogenetic data of the piranha from the genus *Serrasalmus* revealed that species and populations were established only recently in the Upper Madeira during the last million years (Hubert *et al*. in press). Therefore, the colonization of the area occurred during the Pleistocene climatic fluctuations (Haffer 1969; Prance 1982; Whitemore & Prance 1987). The result from the phylogeography of *S. rhombeus* in the area is consistent with this result, since a population expansion was detected here between 800 000 and 400 000 years. Additionally, the higher levels of mtDNA diversity in the populations from the Brazilian shield suggest that the modern mtDNA lineages originated mostly in the Aripuanã refuge. The present finding suggests that some of the forest refuges also served as freshwater refuges for aquatic biota, as previously reported for other fish species of the genus *Leporinus* in French Guyana (Renno *et al*. 1990, 1991). Consequently, the Pleistocene refuges might have contributed to the colonization of the lowlands from the Precambrian shields by creating opportunities for suitable establishment in stable areas. These midpoints might have been of crucial importance for subsequent colonization of the less stable areas during the climatic fluctuations of the Pleistocene. Phylogenetic studies of Neotropical fishes previously stated that an important diversification stage actually predates the Pleistocene and rejected the refuge hypothesis as a potential explanation for the origin of the South American fish diversity (Weitzman & Weitzman 1982; Montoya-Burgos 2003). Our results agree with the fact that the Pleistocene refuges probably did not contribute significantly to increasing the taxonomic diversity of freshwater fishes, but rather support a role in the origin of high levels of local diversity by buffering the impact of stochastic changes in niche availability and limiting the occurrence of species extirpation.

Acknowledgements

This research was supported by Institut de Recherche pour le Développement (IRD, France); Instituto de Biología Molecular y Biotechnología, La Paz (IBM y B, Bolivia), Instituto de Ecología, La Paz (Bolivia) and the GPIA Laboratory, Montpellier (France). Many thanks go to E. Lambert, N. Bierne, B. Guinand, K. Belkhir, J.D. Durand and K. Dawson for their support in the GPIA laboratory; G. Rodriguo, N. Mamani, V. Iñiguez and P. Torrico from the IBMB for laboratory supports and facilities. We wish to thank F. Carvajal, A. Parada, L. Torres, T. Yunoki for their precious help during field sampling and J. Pinto, R. Marín and M. Baudoin for their long-standing interests in our work. Many thanks go to D. Paugy and M. Legendre for their support. We also thank the two anonymous referees for their helpful comments about the manuscript.

References

Agostinho FR (2003) Reproductive aspects of piranhas *Serrasalmus spilopleura* and *Serrasalmus marginatus* into the upper Paraná river, Brazil. *Brazilian Journal of Biology*, **63**, 1–6.

Aleixo A (2004) Historical diversification of a terra-firme forest bird superspecies: a phylogeographic perspective on the role of different hypotheses of Amazonian diversification. *Evolution*, **58**, 1303–1317.

Arnaud-Haond S, Monteforte M, Blanc F, Bonhomme F (2003) Evidence for male-biased effective sex-ratio and recent step-by-step colonization in the bivalve *Pinctada mazatlantica*. *Journal of Evolutionary Biology*, **16**, 790–796.

Atarhouch T, Rami M, Cattaneo-Berrebi G *et al*. (2003) Primers for EPIC amplification of intron sequences for fish and other vertebrate population genetic studies. *BioTechniques*, **35**, 676–682.

Austerlitz F, Jung-Muller B, Godelle B, Gouyon PH (1997) Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology*, **51**, 148–164.

Barthem R, Gouding M (1997) *The Catfish Connection, Ecology, Migration, and Conservation of Amazon Predators*. Columbia University Press, New York.

Bialetzki A, Nakatani K, Sanches PV, Baumgartner G (2002) Spatial and temporal distribution of larvae and juveniles of *Hoplias* aff. *malabaricus* (Characiformes, Erythrinidae) in the upper Paraná river floodplain, Brazil. *Brazilian Journal of Biology*, **62**, 211–222.

Bierne N, Lehnert A, Bédier E, Bonhomme F, Moore S (2000) Screening for intron-length polymorphisms in penaeid shrimps using exon-primed intron-crossing (EPIC)-PCR. *Molecular Ecology*, **9**, 233–235.

Birky CW, Fuerst P, Maruyama T (1989) Organelle gene diversity under migration and drift: equilibrium expectations, approch to equilibrium, effects of heteroplasmic cells and comparison to nuclear genes. *Genetics*, **121**, 613–627.

Borsa P, Collet A, Durand JD (2004) Nuclear-DNA markers confirm the presence of two anchovy species in the Mediterranean. *Comptes Rendus Biologies*, **327**, 1113–1123.

Brown JH, Gibson AC (1983). *Biogeography*. Mosby, St Louis, Missouri.

Castric V, Bonney F, Bernatchez L (2001) Landscape structure and hierarchical genetic diversity in the brook charr, *Salvelinus fontinalis*. *Evolution*, **55**, 1016–1028.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, **32**, 550–570.

Chow S, Hazama K (1998) Universal PCR primers for S7 ribosomal protein gene in introns in fish. *Molecular Ecology*, **7**, 1247–1263.

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

Côrte-Real HBSM, Dixon DR, Holland PWH (1994) Intron-targeted PCR: a new approach to survey neutral DNA polymorphism in bivalve populations. *Marine Biology*, **120**, 407–413.

Costello AB, Down TE, Pollard SM, Pacas CJ, Taylor EB (2003) The influence of history and contemporary stream hydrology on the evolution of genetic diversity within species: an examination of microsatellite DNA variation in bull trout, *Salvelinus confluentus* (Pisces: Salmonidae). *Molecular Ecology*, **57**, 328–344.

Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics*, **134**, 965–969.

Da Silva MNF, Patton JL (1993) Amazonian phylogeography: mtDNA sequence variation in arboreal echimyid rodents (Caviomorpha). *Molecular Phylogenetics and Evolution*, **2**, 243–255.

Da Silva MNF, Patton JL (1998) Molecular phylogeography and the evolution and conservation of Amazonian mammals. *Molecular Ecology*, **7**, 475–486.

Excoffier L, Smouse P, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: an application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Felsenstein J (1993) PHYLIP: phylogeny inference package. Version 3.5c. Department of Genetics, University of Washington, Seattle, Washington.

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

Garnier S, Alibert P, Audiot R, Prieur B, Rasplus JY (2004) Isolation by distance and sharp discontinuities in gene frequencies: implications for the phylogeography of an alpine insect species, *Carabus solieri*. *Molecular Ecology*, **13**, 1883–1897.

Gatesy J, DeSalle R, Wheeler W (1994) Alignment ambiguous nucleotide sites and the exclusion of systematic data. *Molecular Phylogenetics and Evolution*, **2**, 152–157.

Géry J (1977) *Characoids of the World*. Tropical Fish Hobbyist Publications, Neptune City, New Jersey.

Grazziotin FG, Monzel M, Echeverrigaray S, Bonatto SL (2006) Phylogeography of the *Bothrops jararaca* complex (Serpentes: Viperidae): past fragmentation and island colonization in the Brazilian Atlantic forest. *Molecular Ecology*, **15**, 3969–3982.

Guimarães ACD, Quagio-Grassiotto I (2004) Ultrastructural aspects of oogenesis and oocyte primary growth in *Serrasalmus spilopleura* (Teleostei, Characiformes, Characidae). *Tissue and Cell*, **33**, 241–248.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportions for multiple alleles. *Biometrics*, **48**, 361–372.

Guyot JL (1993) *Hydrogéochimie des Fleuves de l'Amazonie Bolivienne*. Editions de l'ORSTOM, Collection Etudes et Thèses, Paris.

Guyot JL, Jouanneau JM, Wasson JG (1999) Characterisation of the river bed and suspended sediments in the Rio Madeira drainage basin (Bolivian Amazonia). *Journal of South American Sciences*, **12**, 401–410.

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

Hall JP, Harvey DJ (2002) The phylogeography of Amazonia revisited: new evidence from Rionid butterflies. *Evolution*, **56**, 1489–1497.

Hassan M, Lemaire C, Fauvelot C, Bonhomme F (2002) Seventeen new EPIC-PCR amplifiable introns in fish. *Molecular Ecology*, **2**, 334–340.

Hassan M, Harmelin-Vivien M, Bonhomme F (2003) Lessepsian invasion without bottleneck: example of two rabbitfish species (*Siganus rivulatus* and *Siganus luridus*). *Journal of Experimental Marine Biology and Ecology*, **291**, 219–232.

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

Hewitt G (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **359**, 183–195.

Hoorn C, Guerrero J, Sarmiento GA, Lorente MA (1995) Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. *Geology*, **23**, 237–240.

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

Hubert N, Duponchelle F, Nuñez J, Garcia-Davila C, Paugy D, Renno JF (2006) Phylogeography of the piranha genera *Serrasalmus* and *Pygocentrus*: implications for the diversification of the Neotropical ichtyofauna. *Molecular Ecology*, doi:10.1111/j.1365-294X.2007.03267.x

Hubert N, Duponchelle F, Nuñez J, Rivera R, Renno JF (2006) Evidence of reproductive isolation among closely related *Serrasalmus* (Ostariophysii, Characidae) from the Upper Madeira river, Amazon, Bolivia. *Journal of Fish Biology*, **69A**, 31–51.

Hutchinson DW, Templeton AR (1999) Correlation of pairwise genetic and geographic measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898–1914.

Knowles LL (2000) Tests of Pleistocene speciation in montane grasshoppers (genus *Malanoplus*) from the sky islands of western North America. *Evolution*, **54**, 1337–1348.

Kuchta SR, Tan AM (2005) Isolation by distance and post-glacial range expansion in the rough-skinned newt, *Taricha granulosa*. *Molecular Ecology*, **14**, 225–244.

Lauzanne L, Loubens G, Le Guennec B (1991) Liste commentée des poissons de l'Amazonie bolivienne. *Revue D'hydrobiologie Tropicale*, **24**, 61–76.

Lemaire C, Versini JJ, Bonhomme F (2005) Maintenance of genetic differentiation across a transition zone in the sea: discordance between nuclear and cytoplasmic markers. *Journal of Evolutionary Biology*, **18**, 70–80.

Lessa EP (1992) Rapid survey of DNA sequence variation in natural populations. *Molecular Biology and Evolution*, **9**, 323–330.

Llamas IR, Godinho AL (1996) Reproduction in the piranha *Serrasalmus spilopleura*, a Neotropical fish with an unusual pattern of sexual maturity. *Environmental Biology of Fishes*, **45**, 161–168.

Lundberg JG, Marshall LG, Guerrero J, Horton B, Malabarba MCSL, Wesselingh F (1998) The stage for Neotropical fish diversification: a history of tropical South American rivers. In: *Phylogeny and Classification of Neotropical Fishes* (eds Malabarba LR, Reis RE, Vari RP, Lucena ZM, Lucena CAS), pp. 13–48. Edipucrs, Porto Alegre, Brazil.

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

Montoya-Burgos JI (2003) Historical biogeography of the catfish genus *Hypostomus* (Siluriformes: Loricariidae), with implications on the diversification of Neotropical ichthyofauna. *Molecular Ecology*, **12**, 1855–1867.

Myers AA, Giller PS (1988) *Analytical Biogeography: an Integrated Approach to the Study of Animals and Plant Distributions*. Chapman & Hall, London.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Noonan BP, Gaucher P (2006) Refugial colonisation and secondary contact in the dyeing poison frog *Dendrobates tinctorius*. *Molecular Ecology*, **15**, 4425–4435.

Nores M (1999) An alternative hypothesis to the origin of Amazonian bird diversity. *Journal of Biogeography*, **26**, 475–485.

Palumbi SR, Baker CS (1994) Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molecular Biology and Evolution*, **11**, 426–435.

Patton JL, Da Silva MNF (1998) Rivers, refuges and ridges: the geography of speciation of Amazonian mammals. In: *Endless Forms: Species and Speciation* (eds Berlocher S, Howard D), pp. 202–216. Oxford University Press, New York.

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

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

Prance GT (1982) *Biological Diversification in the Tropics*. Columbia University Press, New York.

Pritchard JK, Stephens M, Donnelly PJ (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pybus OG, Rambaut A, Harvey PH (2000) An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics*, **155**, 1429–1437.

Pybus OG, Charleston MA, Gupta S, Rambaut A, Holmes EC, Harvey PH (2001) The epidemic behavior of the hepatitis C virus. *Science*, **292**, 2323–2325.

Räsänen M, Neller R, Salo JS, Jungner H (1992) Recent and ancient fluvial deposition systems in the Amazonian foreland basin. *Peru*. *Geological Magazine*, **129**, 293–306.

Räsänen ME, Salo JS, Jungnert H, Pittman LR (1990) Evolution of the western Amazon lowland relief: impact of Andean foreland dynamics. *Terra Nova*, **2**, 320–332.

Raymond M, Rousset F (1995) GENEPOP (version 3.1): population genetics software for exact tests and ecumenism. *Journal of Heredity*, **86**, 248–249.

Renno JF, Berrebi P, Boujard T, Guyomard R (1990) Intraspecific genetic differentiation of *Leporinus friderici* (Anostomidae, Pisces) in French Guiana and Brazil: a genetic approach to the refuge theory. *Journal of Fish Biology*, **36**, 85–95.

Renno JF, Hubert N, Torrico JP *et al*. (2006) Phylogeography of *Cichla* (Cichlidae) in the upper Madera basin (Bolivian Amazon). *Molecular Phylogenetics and Evolution*, **41**, 503–510.

Renno JF, Machordom A, Blanquer A, Boursot P (1991) Polymorphism of mitochondrial genes in populations of *Leporinus friderici* (Bloch, 1974): intraspecific structure and zoogeography of the Neotropical fish. *Genetica*, **84**, 137–142.

Rice WR (1989) Analysing tables of statistical tests. *Evolution*, **43**, 223–225.

Rizzo E, Bazzoli N (1991) The zona pellucida of the Brazilian white piranha, *Serrasalmus brandtii* Reinhardt, 1874 (Pisces, Characidae): a cytological and cytochemical study. *Functional and Developmental Morphology*, **1**, 21–24.

Rizzo E, Sato Y, Barreto B, Godinho HP (2002) Adhesiveness and surface patterns of eggs in Neotropical freshwater teleosts. *Journal of Fish Biology*, **61**, 615–632.

Rodríguez F, Oliver JF, Marín A, Medina JR (1990) The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology*, **142**, 485–501.

Roger AR (1995) Genetic evidence for a Pleistocene population explosion. *Evolution*, **49**, 608–615.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of the pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Roy MS, Silva JMC, Arctander P, García-Moreno J, Fjeldså J (1997) The speciation of South American and African birds in montane regions. In: *Avian Molecular Evolution and Systematics* (ed. Mindell DP), pp. 325–343. Academic Press, New York.

Saint-Paul U, Zuanon J, Villacorta Correa MA, García M, Fabré NN, Berger U, Junk WJ (2000) Fish communities in central Amazonian white- and blackwater floodplains. *Environmental Biology of Fishes*, **57**, 235–250.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN version 2.0: a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

Sioli H (1975) Tropical rivers as expressions of their terrestrial environments. In: *Tropical Ecological Systems: Trends in Terrestrial and Aquatic Research* (eds Golley FB, Medina E), pp. 275–288. Springer Verlag, Berlin.

Sioli H (1984) The Amazon and its main affluents: hydrography, morphology of the river courses, and river types. In: *The Amazon: Limnology and Landscape Ecology of a Mighty Tropical River and its Basin* (ed. Sioli H), pp. 127–165. Monographiae Biologicae 56, The Netherlands.

Sivasundar A, Bermingham E, Ortí G (2001) Population structure and biogeography of migratory freshwater fishes (*Prochilodus*: Characiformes) in major South American rivers. *Molecular Ecology*, **10**, 407–417.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Sneath PHA, Sokal RR (1973) *Numerical Taxonomy*. W.H. Freeman, San Francisco, California.

Strimmer K, Pybus OG (2001) Exploring the demographic history of DNA sequences using the generalised skyline plot. *Molecular Biology and Evolution*, **18**, 2298–2305.

Swofford DL (2002) PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinaeur Associates, Sunderland, Massachusetts.

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

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

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

Thompson JD, Higgins DG, Gibson TJ (1993) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

Turgeon J, Bernatchez L (2001) Clinal variation at microsatellite loci reveals historical secondary intergradation between glacial races of *Coregonus artedi* (Teleostei: Coregoninae). *Evolution*, **55**, 2274–2286.

Uetanabaro M, Wang T, Abe AS (1993) Breeding behaviour of the red-bellied piranha, *Pygocentrus nattereri*, in nature. *Environmental Biology of Fishes*, **38**, 369–371.

Vari RP (1991) Systematics of the Neotropical characoid genus *Steidachnerina* Fowler (Pisces: Characoidei). *Smithsonian Contribution to Zoology*, **507**, 1–118.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Weir BS (1990) *Genetic data analysis: methods for discrete population analysis*. Sunderland, MA Sinauer.

Weitzman SH, Weitzman MJ (1982) Biogeography and evolutionary diversification in Neotropical freshwater fishes with comments on the refuge theory. In: *Biological Diversification in the Tropics* (ed. Prance GT), pp. 403–422. Columbia University Press, New York.

Whitemore TC, Prance GT (1987) *Biogeography and Quaternary History in Tropical America*. Oxford University Press, Oxford.

Nicolas Hubert obtained his PhD on the biogeography and the evolution of the Piranhas and this work is part of his PhD. Jesus Nuñez is a fish physiologist specialist in reproduction. Fabrice Duponchelle is a population biologist interested in the interactions between environment and life histories in fishes. François Bonhomme is currently director of the GPIA laboratory in Montpellier. Jean-François Renno is population geneticist. He spent 8 years in Africa working on the genetic of the millet's domestication and he is now the coordinator of the Amazonian part of the IRD project: Caractérization et Valorization de la Diversité Ichtyologique pour une Aquaculture Raisonnée (CAVIAR). Jean-François was also Nicolas's PhD advisor.