

Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

## Molecular Phylogenetics and Evolution

journal homepage: [www.elsevier.com/locate/ympev](http://www.elsevier.com/locate/ympev)

## Short Communication

## Are rapids a barrier for floodplain fishes of the Amazon basin? A demographic study of the keystone floodplain species *Colossoma macropomum* (Teleostei: Characiformes)

Izeni Pires Farias<sup>a</sup>, Juan Pablo Torrico<sup>b</sup>, Carmen García-Dávila<sup>c</sup>, Maria da Conceição Freitas Santos<sup>a</sup>, Tomas Hrbek<sup>d,\*</sup>, Jean-François Renno<sup>b</sup>

<sup>a</sup> Universidade Federal do Amazonas (UFAM), Departamento de Ciências Biológicas, Laboratório de Evolução e Genética Animal, Mini Campus ICB, Av. Gen. Rodrigo Octávio Jordão Ramos, 3000 – Coroado, Cep 69.077-000, Manaus, AM, Brazil

<sup>b</sup> IRD UR 175/IBMB, Universidad Mayor San Andrés, Campus Cota Cota, La Paz, Bolivia

<sup>c</sup> Instituto de Investigaciones de la Amazonia Peruana (IIAP), Laboratorio de Biología Y Genética Molecular (LBMB), Av. Abelardo Quiñones km. 2.5, Iquitos, Peru

<sup>d</sup> University of Puerto Rico – Rio Piedras, Biology Department, San Juan, PR 00931, Puerto Rico

## ARTICLE INFO

## Article history:

Received 21 March 2010

Accepted 25 March 2010

Available online 30 March 2010

## Keywords:

Madeira River rapids  
Demographic history  
Genetic structure  
Phylogeography  
Mitochondrial DNA  
*Colossoma macropomum*

## ABSTRACT

We investigated demographic history and population structuring of *Colossoma macropomum* sampled from 14 localities in the Amazon basin and the Bolivian sub-basin; the two basins are separated by a series of 16 rapids. Although genetically differentiated, IMA analyses suggest non-zero bi-directional migration rates, and inter-basin divergence of approximately 17 thousand years ago. Analyses in BEAST indicated that Bolivian *C. macropomum* has been demographically stable except for a moderate population increase in the last 12 thousand years, while Amazonian *C. macropomum* has been experiencing demographic growth over the last 350 thousand years, resulting in approximately one order of magnitude increase in coalescent  $N_e$ .

© 2010 Elsevier Inc. All rights reserved.

## 1. Introduction

Most of the large rivers of the Amazon basin are free of major physical obstructions and barriers that would impede the movement of aquatic animals. A major exception to this pattern exists in the form of a series of rapids and cataracts on the Madeira River that separate the Bolivian sub-basin from the main Amazon basin. The Bolivian sub-basin encompasses the upper portion of the Madeira River drainage, and is separated from the main Amazon basin by a series of 16 rapids (Goulding et al., 2003). These rapids are known to form a distributional barrier for a number of species of fishes and aquatic vertebrates (Goulding, 1979), however, for some species such as the catfishes of the genera *Brachyplatystoma* and *Pseudoplatystoma*, and characids of the genus *Prochilodus* these rapids apparently pose no barrier to reproductive migrations. *Colossoma macropomum* is also found in both basins; however, ecological and fisheries data suggest that *C. macropomum* only undertake localized season movements for feeding and reproduction within lentic floodplain habitats, triggered by annual cycles of flooding and contraction of the floodplain (Araújo-Lima and Goulding,

1998), and therefore the Madeira River rapids are thought to present a major barrier to movement for *C. macropomum* between the Bolivian sub-basin and main Amazon basin.

Understanding the role that the rapids on the Madeira River play in disrupting genetic connectivity between aquatic organisms of the Bolivian Amazon and the main Amazon basins is fundamental for understanding the evolutionary dynamics that have and are shaping patterns of distribution of genetic diversity in the greater Amazon basins. *Colossoma macropomum* serves well as a model since it is a typical species associated with large river floodplains of the Amazon basin, and as most species occurring in this habitat is also assumed to be largely sedentary. Therefore, we investigate the role that the Madeira River rapids play in restricting gene flow between the basins, the strength and the directionality of this gene flow, and the relative importance of gene flow and genetic drift on the intra- and inter-basin evolutionary history of *C. macropomum*. Because coalescent based estimates of gene flow generally assume constant population size, we also investigated changes in effective population sizes over the coalescent history of *C. macropomum*, from the present through the Pleistocene. An additional motivation comes from recent studies suggesting effective population size expansions for a wide range of Amazonian organisms including passerine birds (Aleixo, 2004), leafcutter ants (Solomon et al.,

\* Corresponding author. Fax: +1 787 764 3875.

E-mail address: [hrbek@uprrp.edu](mailto:hrbek@uprrp.edu) (T. Hrbek).

2008), Amazonian manatees (Cantanhede et al., 2005) and piranhas (Hubert et al., 2007).

## 2. Material and methods

### 2.1. Sample and data collection

The sample areas of *C. macropomum* collected in the main Amazon basin and Bolivian sub-basin on either side of the Madeira River rapids are shown in the Fig. 1. Muscle tissue was stored in 95% ethanol at ambient temperature until being processed in the laboratory. A total of 235 individuals were amplified: 160 samples were from the central Amazon basin (127 from Brazil, 33 from Peru), and 75 from the Bolivian sub-basin. Total genomic DNA was extracted using either a phenol–chloroform/proteinase K protocol (Sambrook et al., 1989) or DNeasy Tissue Kit (Qiagen). The mtDNA control region (d-loop, 733 bp) was amplified with the primers designed by Sivasundar et al. (2001) or Agnès et al. (2006). PCR reactions were carried out in a final volume of 25  $\mu$ l and contained 11.7  $\mu$ l of ddH<sub>2</sub>O, 3  $\mu$ l of MgCl<sub>2</sub> (25 mM), 2.5  $\mu$ l of dNTPs (10 mM), 2.5  $\mu$ l of 10 $\times$  buffer (100 mM Tris–HCl, 500 mM KCl), 2  $\mu$ l of each primer (2  $\mu$ M), 0.3  $\mu$ l of Taq DNA Polymerase (5 U/ $\mu$ l) and 1  $\mu$ l of DNA (concentration varied between 50 and 100 ng). For Sivasundar et al. (2001) primers, PCR conditions were as follows: denaturation at 92 °C for 35 s, primer annealing at 55 °C for 35 s, and primer extension at 72 °C for 90 s; these three steps were repeated 35 times. PCR conditions for primers from Agnès et al. (2006) were as follows: 120 s denaturation at 95 °C, 10 cycles starting at 66 °C with a 1 °C temperature decrement; finally 25 cycles with 60 s at 92 °C, 60 s at 56 °C and 90 s at 72 °C. Post-PCR extension was carried out for 5 min at 72 °C. Purification of products was done using an ethanol/salt precipitation.

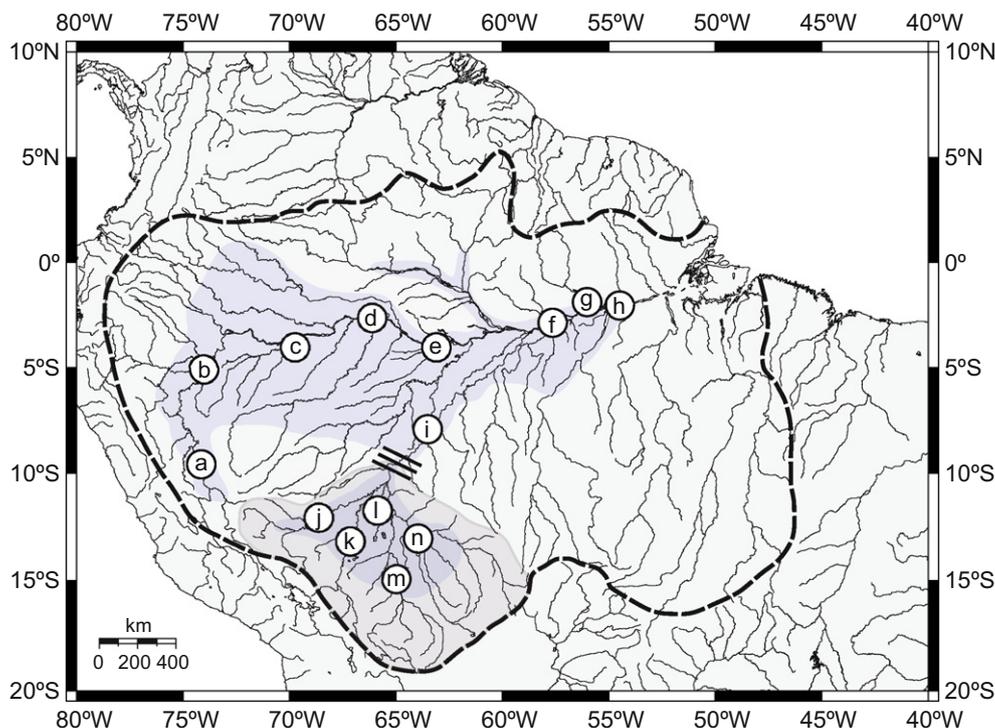
Purified PCR products of Brazilian samples were sequenced directly using the DYEamic ET Dye Terminator mix (GE Healthcare)

following manufacturer's protocol. Precipitated product were resolved on a MegaBACE 1000 automatic DNA analysis system (GE Healthcare) using the manufacturer's recommended settings. Bolivian and Peruvian samples were sequenced by MacroGen (Seoul, Korea). Sequences were edited and aligned by eye in the program BioEdit (Hall, 1999).

### 2.2. Data analysis

First level of analysis consisted of testing of genetic homogeneity among all sampled localities using an Analysis of Molecular Variance (AMOVA) (Excoffier et al., 1992). Since homogeneity was rejected, the samples were grouped into basin macro-regions where Peruvian (samples from Iquitos and Pucallpa) and Brazilian samples (Santarém, Oriximiná, Parintins, Coari, Tabatinga and Humaitá) comprised the Amazon basin macro-region, and the Madre de Dios, Beni, Yata, Mamoré, and Guapore samples comprised the Bolivia sub-basin macro-region. The two macro-regions are separated by the upper Madeira River rapids. We also estimated pair-wise  $\Phi_{ST}$  values between all pairs of sampling localities. In all cases of multiple tests, significance levels were adjusted using the sequential Bonferroni correction (Rice, 1989). Significance of observed differences was assessed non-parametrically via 10,000 permutations using uncorrected pair-wise differences between haplotypes. We also tested for a correlation between genetic and geographical distance of sampling localities. Significance of correlation was assessed with 10,000 random permutations. All AMOVA, pair-wise  $\Phi_{ST}$ , and correlation tests were performed in Arlequin 3.11 version (Excoffier et al., 2005).

We used the program IMA (Hey and Nielsen, 2007) to simultaneously estimate the coalescent effective population sizes (from the parameter  $\theta$ ) of *C. macropomum* in the Bolivian sub-basin, the main Amazon basin and of the ancestral *C. macropomum* population, migration rates between the basins ( $m$ ) and their significance,



**Fig. 1.** Distribution (light purple) and sampling localities of *Colossoma* collected in the Amazon and Bolivian basins. The Amazon basin is outlined with a dashed line, while the Bolivian basin is shaded in grey. Localities are (a) Pucallpa, (b) Iquitos, (c) Tabatinga, (d) Tefé, (e) Coari, (f) Parintins, (g) Oriximiná, (h) Santarém, (i) Humaitá, (j) Manuripi River, (k) Beni River, (l) Yata River, (m) Secure River, and (n) Guaporé River.

and the age of the most recent ancestor of the Bolivian and Amazon basin *C. macropomum* ( $\tau$ ). After preliminary runs which were necessary to obtain realistic upper and lower bound values for priors for each of the five parameters estimated, we performed nine independent replicates, holding priors identical but starting from different starting points, to verify convergence. Each run consisted of 1 million initial burn-in samples, followed by 10 million samples of the parameter space with every 1000th sample being recorded. Ten independent chains were run per replicate, and swap rates between changes were set to range between 90% and 50%. Trees from the independent replicates were combined, the posterior probabilities of parameters  $\theta$ ,  $m$  and  $\tau$  were estimated, and then were converted into demographic parameters: the coalescent effective population size of *C. macropomum* in each basin, the number of immigrants and emigrants to and from each basin, the time of divergence of the two basins, the ancestral effective population size, and  $2Nm$  for each basin. To make these conversions, we assumed a generation time of 3 years; this estimate is based on smallest size at maturity occurring at 50–55 cm (Goulding and Carvalho, 1982; Isaac et al., 1996) and length/age relationship estimated from Bertalanffy's model by Isaac and Ruffino (1996). We also assumed a conservative estimate of the neutral substitution rate of  $2.0 \times 10^{-8}$  mutations per site per year for the control region; this estimate is an average derived from a wide range of teleost species (Donaldson and Wilson, 1999; Sato et al., 2003). Sato et al. (2003) estimated  $2.2\text{--}4.3 \times 10^{-8}$  substitutions per site per year, while Donaldson and Wilson (1999) estimated  $1.8 \times 10^{-8}$  substitutions per site per year. As recently pointed out by Ho et al. (2005), substitution rate estimates tend to be underestimates of the actual mutation rate due to not taking into account high frequency, short-lived mutations, which in turn would result in estimates of lower effective population sizes, and earlier demographic events; therefore our estimates should represent conservative upper bounds.

Using the program IMA we also specifically tested the hypotheses of no gene flow from the Amazon basin to the Bolivian basin ( $m_1 = 0, m_2 > 0$ ), no gene flow from the Bolivian basin to the Amazon basin ( $m_1 > 0, m_2 = 0$ ), and no gene flow between both basins ( $m_1 = 0, m_2 = 0$ ). Selection among these alternate models and the null model of unconstrained gene flow ( $m_1 > 0, m_2 > 0$ ) was carried out using Bayes Factors (Kass and Raftery, 1995). The use of Bayes Factors is justifiable under the assumptions that the marginal priors for both hypotheses are the same (Pericchi, 2005), and in a case of nested hypotheses comparisons are analogous to a hierarchical likelihood ratio test (Hey and Nielsen, 2007).

Genetic connectivity within and between basins were evaluated further within the maximum-likelihood framework in the program MIGRATE version 2.4.4 (Beerli and Felsenstein, 2001). To establish initial  $\theta$  and  $M$  parameter values, we ran 10 short chains, sampling each chain 10,000 times. We then ran six long chains, sampling each chain 1,100,000 times, and discarding the first 100,000 samples as burn-in. Search of parameter space was improved through adaptive swapping among chains. Convergence between chains was assessed using the Gelman–Rubin criterion (Gelman and Rubin, 1992). MIGRATE analyses were repeated five times with different random start seeds (i.e. different random starting conditions) to assess convergence.

We used the program BEAST (Drummond and Rambaut, 2007) to obtain independent estimates of the diversity parameter  $\theta$ , the time of the most recent ancestor of the Bolivian and Amazon basin *C. macropomum* ( $\tau$ ), and to investigate patterns of changes in effective population size changes throughout the coalescent history of *C. macropomum* in each basin. Changes in effective population sizes were inferred via Bayesian skyline plots. All analyses were run for 50 million generations with a burn-in of 5 million initial samples, sampling every thousands topology. To account for differences in observed rates of transitions and transversions, we used the

HKY85 (Hasegawa et al., 1985) model of molecular evolution with a portion of sites assumed to be invariant and the mutation rate in the rest being gamma distributed, and a relaxed molecular clock (Drummond et al., 2006). Model of molecular evolution was selected in the program MODELTEST (Posada and Crandall, 1998). For initial analyses, priors were left at default values, and we sampled over 20 time periods ( $m = 20$ ). Priors were re-adjusted based on preliminary analyses. Final analyses were repeated three times to assure convergence of estimated parameters, and the three independent runs were combined. The absolute time scale of the Bayesian skyline plot, the estimates of the divergence of the Amazon and Bolivian basins and the time to the most recent common ancestor within each basin was based on the mutation rate of  $2.0 \times 10^{-8}$  substitutions per site per year (see justification above). The program BEAST (Drummond and Rambaut, 2007) also allows the estimation of posterior probabilities of specific demographic scenarios. We estimated the posterior probabilities of the hypotheses of constant population size, logistic growth and exponential growth for our data, and selected among alternate hypotheses using Bayes Factors (Pericchi, 2005; Suchard et al., 2001).

Unrooted phylogenetic relationships among the mtDNA haplotypes were inferred by maximum-likelihood and Bayesian inference methods. Both analyses were performed under the HKY85 model of molecular evolution (Hasegawa et al., 1985) with a portion of sites assumed to be invariant and the mutation rate in the rest being gamma distributed as determined in the program MODELTEST (Posada and Crandall, 1998). Maximum-likelihood (ML) topology was inferred in the program Treefinder (Jobb et al., 2004), and Bayesian inference topology was inferred in the program MrBayes 3.1 (Ronquist and Huelsenbeck, 2003).

### 3. Results

A total of 733 bases were sequenced for 235 *C. macropomum* specimens. Uncorrected pair-wise sequence divergence between *C. macropomum* haplotypes ranged from 0.13% to 1.5%. Sequences are available from GenBank (Accession Nos. GQ375614–GQ375745).

Analysis of Molecular Variance revealed significant allelic differences among sampled localities ( $\Phi_{ST} = 0.0755, P < 0.001$ ); however, based on examination of pair-wise  $\Phi_{ST}$ , significant contrasts involved between basin comparisons, but not within basin comparisons. Exceptions to this pattern were pair-wise comparisons with the Humaita sampling locality which were significant neither for Bolivian sub-basin nor the main Amazonian basin comparisons. A hierarchical analysis testing the effect of the upper Madeira River rapids on *C. macropomum* population structure resulted in the inference of differentiation between the main Amazon basin and the Bolivian sub-basin drainages ( $\Phi_{CT} = 0.1209, P < 0.001$ ), but no differentiation among localities within each region ( $\Phi_{SC} = 0.0755, P = 0.090$ ). In light of these results, we performed AMOVA and Mantel test analyses for the Amazon basin and the Bolivian sub-basin only. Results showed no population structuring ( $\Phi_{ST} = 0.0156, P = 0.094$ ) and no correlation between genetic and geographic distances ( $r = 0.1711; P = 0.196$ ) in the Amazon basin. Likewise, there was no significant structuring of localities ( $\Phi_{ST} = 0.0153, P = 0.262$ ), and no correlation between geographic and genetic distances ( $r = -0.3490; P = 0.751$ ) within the Bolivian basin.

To investigate the levels and significance of gene flow between the main Amazon basin and the Bolivian sub-basin, we performed an analysis in IMA (Hey and Nielsen, 2007). IMA analyses indicated that gene flow between basins was non-trivial; bi-directional gene flow was a significantly better hypothesis than any of the constrained models (Table 1). The rate of migration between the basins was similar ( $m_1 = 0.2161$  and  $m_2 = 0.3256$ ; Table 2).

**Table 1**  
Test of alternate hypotheses in the IMA model.

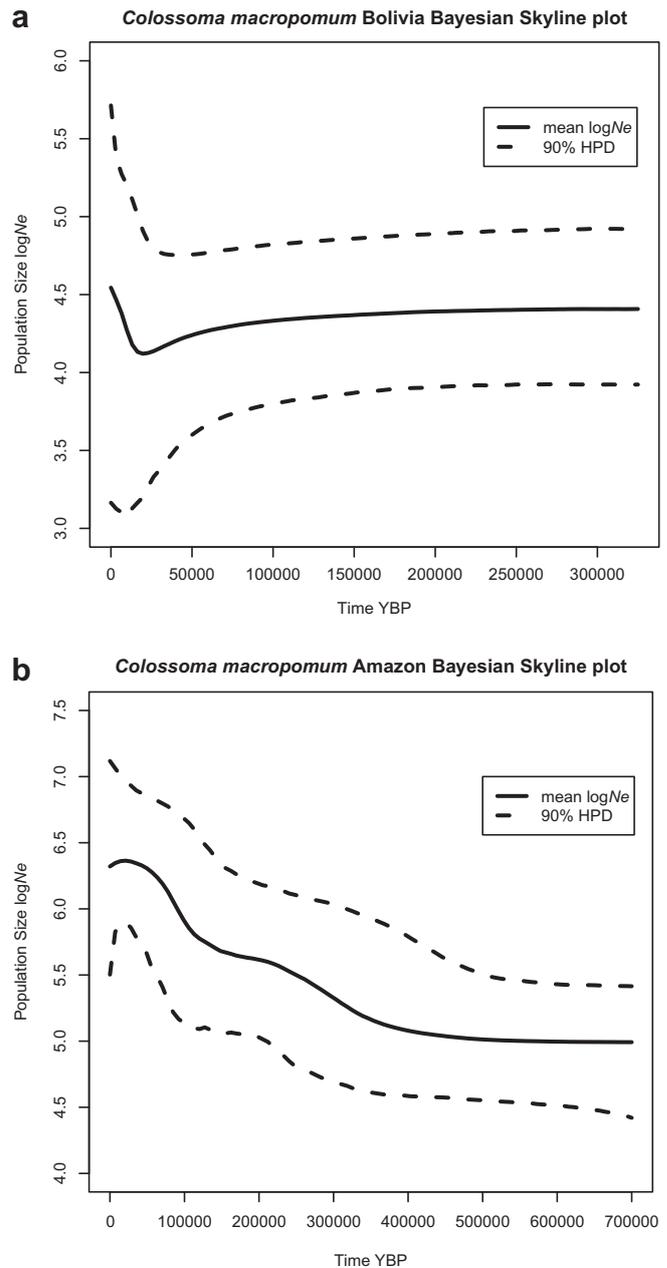
Model	log(P)	-2Δlog(P)	df	p
$\theta_1 \theta_2 \theta_A m_1 m_2$	-12.7222			
$\theta_1 \theta_2 \theta_A m_1 = m_2$	-148.922	272.400	1	<0.0001
$\theta_1 \theta_2 \theta_A m_1 m_2 = 0$	-149.404	273.364	1*	<0.0001
$\theta_1 \theta_2 \theta_A m_1 = 0 m_2$	-247.123	468.802	1*	<0.0001
$\theta_1 \theta_2 \theta_A m_1 = m_2 = 0$	-253.095	480.746	2*	<0.0001
$\theta_1 = \theta_2 \theta_A m_1 m_2$	-215.528	405.612	1	<0.0001
$\theta_1 = \theta_2 \theta_A m_1 = m_2$	-227.621	429.797	2	<0.0001
$\theta_1 = \theta_2 \theta_A m_1 = m_2 = 0$	-312.292	599.139	3*	<0.0001
$\theta_1 = \theta_A \theta_2 m_1 m_2$	-142.449	259.453	1	<0.0001
$\theta_1 = \theta_A \theta_2 m_1 = m_2$	-153.117	280.790	2	<0.0001
$\theta_1 = \theta_A \theta_2 m_1 = m_2 = 0$	-258.991	492.537	3*	<0.0001
$\theta_1 \theta_2 = \theta_A m_1 m_2$	-142.092	258.740	1	<0.0001
$\theta_1 \theta_2 = \theta_A m_1 = m_2$	-152.765	280.085	2	<0.0001
$\theta_1 \theta_2 = \theta_A m_1 = m_2 = 0$	-279.789	534.134	3*	<0.0001
$\theta_1 = \theta_2 = \theta_A m_1 m_2$	-228.067	430.690	2	<0.0001
$\theta_1 = \theta_2 = \theta_A m_1 = m_2$	-236.555	447.666	3	<0.0001
$\theta_1 = \theta_2 = \theta_A m_1 = m_2 = 0$	-318.188	610.931	4*	<0.0001

\* Indicates models that are not fully nested.

However, coalescent effective population sizes differed vastly between basins (main Amazon basin  $N_e = 1.44 \times 10^6$ ; Bolivian sub-basin  $N_e = 2.25 \times 10^4$ ), and the estimated coalescent effective population size of the ancestral population was  $1.75 \times 10^5$  female individuals (Table 2). Due to the large differences in coalescent effective population sizes, gene flow into the main Amazon basin was much higher ( $2N_e m = 112.79$ ) than gene flow into the Bolivian sub-basin ( $2N_e m = 2.58$ ). Gene flow between the two basins inferred in the program MIGRATE version 2.4.4 (Beerli and Felsenstein, 2001) was  $2N_e m = 2.78$  from the main Amazon basin to the Bolivian sub-basin, and  $2N_e m = 129.28$  from the Bolivian sub-basin to the main Amazon basin. The estimated divergence time of *C. macropomum* from these two regions was 17.4 thousand years (Table 2). Although the IMA model assumes Fisher–Wright equilibrium (Hey and Nielsen, 2007), exponential population growth appears not to affect migration rate estimates, but may underestimate effective population sizes (Strasburg and Rieseberg, 2010).

Bayesian skyline plots showed a relatively stable but gradually decreasing population size of *C. macropomum* with a very recent demographic increase (approximately 12,000 years before present) in the Bolivian basin (Fig. 2a), while in the Amazon basin the population size of *C. macropomum* has remained constant until approximately 350,000 years ago after which point *C. macropomum* has been experiencing demographic growth (Fig. 2b). This demographic growth accelerated greatly in the last 125,000 years. Since the beginning of the initial growth phase, the population size of the Amazon basin appears to have increased by an order of magnitude, from approximately 0.1 to 2 million individuals.

Alternate demographic hypotheses were also tested via the comparison of posterior probabilities using Bayes Factors (Pericchi, 2005; Suchard et al., 2001). We concluded that demographic growth was a significantly better explanation than constant population size for the Amazon basin *C. macropomum* (exponential  $-\ln P(\text{model}|\text{data}) = 2904.717$  vs. constant  $-\ln P(\text{model}|\text{data}) = 2779.757$ ; BF = 54.270). Neither exponential growth nor constant population size were better demographic models for *C. macropomum* from the Bolivian basin over the coalescent history of this population (exponential  $-\ln P(\text{model}|\text{data}) = 1277.905$  vs. constant



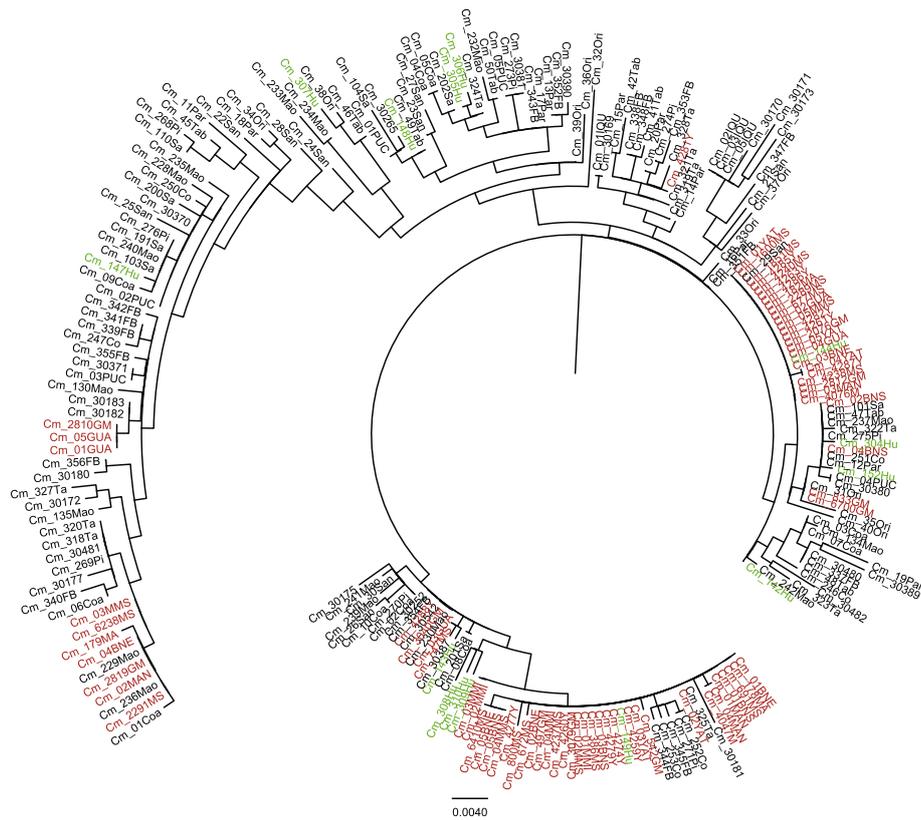
**Fig. 2.** Bayesian skyline plot estimated from Bolivian (a) and Amazon (b) samples. The thick solid line is the median estimate, and the dashed lines represent the 90% HPD limits. Note:  $N_e$  is on a log10 scale.

$-\ln P(\text{model}|\text{data}) = 1277.814$ ; BF = 0.040); and therefore the simpler model of constant population should be chosen.

An unrooted maximum-likelihood tree estimated using HKY85+I+G model of molecular evolution showed a star-like relationship among haplotypes with two major haplotype clades comprising the majority of the Bolivian samples (Fig. 3); this pattern was also recovered by the Bayesian inference analysis (results not shown). These two major haplotype clades are hereafter

**Table 2**  
Maximum-Likelihood Estimates (MLE) and the 90% Highest Posterior Density (HPD) intervals of demographic parameters.

Comparison	$\theta_1$	$\theta_2$	$\theta_A$	$m_1$	$m_2$	$\tau$	$N_{e1}$	$N_{e2}$	$N_{eA}$	$2N_{e1}m_1$	$2N_{e2}m_2$	t (years)
MLE	1016.1447	15.8542	123.1467	0.222	0.326	1.0175	1444053.74	22530.41	175003.84	112.792	2.584	17351.64
Lower 90% HPD	668.0893	7.0054	58.2550	0.042	0.002	0.7025	949422.04	9955.38	82786.21	-	-	11979.88
Upper 90% HPD	2142.9002	33.6994	211.6354	0.490	1.522	3.1075	3045276.55	47890.23	300755.17	-	-	52992.84



**Fig. 3.** Unrooted maximum-likelihood (ML) trees showing the relationship of *Colossoma* haplotypes estimated using the HKY85 model of molecular evolution. The geographical origin of haplotypes is identified by sample locality IDs (Brazilian samples: Santarém (SAN), Oriximiná (ORI), Parintins (PAR), Coari (COA), Tefé (TEF), Tabatinga (TAB) and Humaitá (HUM); Peru samples: Iquitos (IQU) and Ucayali River at Pucallpa (PUC); Bolivia samples: Madre de Dios basin (Manuripi River (MAN)), Bení Basin (Puerto Salinas (BNS) and Esperanza from Bení River (BNE)), Mamoré Basin (Secure River (MMS), Ichiro River (MMI), Yata River (YAT) and Guaporé River (GUA)). Localities in black are from the Amazon basin, in red from the Bolivian basin and in green from Humaitá (Amazon basin).

referred to as clades A and B. Clade B is almost entirely restricted to Bolivia but includes one sample from Santarém and another one from Oriximiná, both localities near the eastern-most distributional limit of *Colossoma*, and downstream of the confluence of the Madeira River with the Amazon River. In addition to these two major haplotype clades, six additional haplotypes occur in Bolivia and are nested among the other haplotypes of Peru and Brazil.

## 4. Discussion

### 4.1. Madeira rapids limiting gene flow

The Madeira River basin is the largest sub-basin of the Amazon River; at 1.4 million km<sup>2</sup>, its catchment area represents just over 20% of the whole Amazon River basin. The Madeira River contributes 50% of the total sediment carried by the Amazon, and 15–16% of the total annual discharge, making it arguably the most important tributary of the Amazon. The Madeira River basin is divided into two sections by a series of rapids; commencing just upstream of the Brazilian city of Porto Velho, there are 16 major rapids along a 360 km stretch between Porto Velho, Guajara Mirim and Riberalta. The Bolivian sub-basin upstream of the rapids comprises around 60% of the total area drained by the Madeira. The rapids correspond to regions where the Madeira breaks through and descends the pre-Cambrian Brazilian Shield. The current Madeira Basin watershed is relatively young, having taken on its present limits after the breach of the Purus Arch in the Late Miocene (~8 mya) (Lundberg et al., 1998). The Early to Mid Pliocene (4–3 mya) rise of the Fitzcarrald Arch gradually isolated and uplifted

the Bolivian basin (Campbell et al., 2001; Hoorn et al., 1995; Räsänen et al., 1987). This uplift resulted in the major shifts in drainage pattern such that all major rivers of the northern Bolivian Basin started to drain through the rapids region of upper Madeira (Campbell et al., 2001; Hoorn et al., 1995; Räsänen et al., 1987).

The causal role of the Madeira rapids in differentiating populations from the Bolivian and Amazon basins has been proposed for such species as needlefishes *Potamorhaphis eigenmanni* (Lovejoy and Collette, 2001), the pink river dolphin *Inia geoffrensis* (Banguera-Hinestroza et al., 2002), and the giant Amazon River turtle *Podocnemis expansa* (Pearse et al., 2006). The Madeira River rapids also have been suggested to form a major barrier to migration for a number of fish species, and they limit the distribution of a number of other fish species (Goulding, 1979). However, large catfishes of the genera *Brachyplatystoma* and *Pseudoplatystoma* appear to pass through the rapids without difficulty, as do the migratory characins *Prochilodus nigricans* (Goulding, 1979). The effect of the rapids on *Colossoma macropomum* and other groups of fishes and vertebrates that occur upstream and downstream of the rapids is uncertain. For example, while it is generally assumed that *C. macropomum* is a relatively sedentary species occurring in the flooded forest of the Amazon and Bolivian basin, and only undergoing lateral migrations associated with feeding and reproduction during the high water season (Araújo-Lima and Goulding, 1998), Goulding (1979) reports that *C. macropomum* have been observed to pass through Teotônio, the largest of the Madeira River rapids. Interviews with local fishermen also indicate that *C. macropomum* scale, though not frequently the Madeira River rapids (IPF and TH pers. obs.).

Our results indicate that the Madeira River rapids do not form an absolute barrier to genetic exchange for *C. macropomum* of the

two basins. Migration rates are bi-directional, but not symmetrical ( $-2\Delta\log(P) = 272.400$  non-symmetrical vs. symmetrical migration rate;  $df = 1$ ;  $P < 0.001$ ); the migration rate from the main Amazon basin to the Bolivian sub-basin is greater ( $m = 3.26 \times 10^{-4}$ ) than from the Bolivian sub-basin to the main Amazon basin ( $m = 2.22 \times 10^{-4}$ ) (Table 2). Both migration rates are also significantly greater than zero (Bolivia to Amazon basin  $m = 0$  vs.  $m > 0$ ,  $-2\Delta\log(P) = 273.364$ ;  $df = 1$ ;  $P < 0.001$ ; Amazon to Bolivia basin  $m = 0$  vs.  $m > 0$ ,  $-2\Delta\log(P) = 468.802$ ;  $df = 1$ ;  $P < 0.001$ ) supporting the inference that even a major geological feature such as the Madeira River rapids do not impede significant levels of genetic exchange between floodplain adapted fishes.

When number of effective migrants exchanged between basin are estimated, both IMA and MIGRATE analyses suggest very high rates of downstream gene flow from the Bolivian sub-basin into the main Amazon basin ( $2N_e m \approx 100$ ), while rates of upstream gene flow from the main Amazon basin to the Bolivian sub-basin are nearly two orders of magnitude lower ( $2N_e m \approx 2$ ). This drastic difference in effective number of migrants results from the nearly two orders of magnitude difference in coalescent effective population sizes of the two basins. The net movement from the Bolivian basin to the Amazon basin is intuitively satisfying, and directional movement of *C. macropomum* has already been proposed. Reinert and Winter (2002) suggest that the relatively unexploited Bolivian population of *C. macropomum* acts as a source for the heavily exploited Brazilian population via passive transport of larval and juvenile individuals into downstream nursery areas in the Brazilian portion of the Amazon basin.

#### 4.2. Intra basin demographic histories

Prior to the demographic expansion in the Amazon basin, the effective population size of *C. macropomum* was approximately five times that of the *C. macropomum* in the Bolivian sub-basin. This estimate broadly corresponds to the relative proportion of potential habitat in the Bolivian and Amazon basins. The Bolivian sub-basin comprises approximately 10% of the total area of the Amazon basin; however, with the exception of their mouths *C. macropomum* does not occur in Amazonian rivers descending the Brazilian and Guiana Shields which drain approximately 40% of the Amazon basin. The Bolivian basin, therefore, comprises slightly less than 20% of the total area occupied by *C. macropomum*, and until the middle to late Pleistocene, the Bolivian basin had a *C. macropomum* population approximately 20% smaller than the rest of the Amazon basin.

Bayesian estimators of demographic processes implemented in the program BEAST suggest that *C. macropomum* of the Amazon basin experienced significant population growth in the Pleistocene in the last 350,000 years, followed by a modest population decrease in the last 10,000 years. During this time, the female coalescent effective population size has increased by approximately one order of magnitude to approximately 2.1 million individuals (Fig. 2b). The result obtained in BEAST is broadly compatible with the coalescent effective population size obtained in IMA (1.4 million individuals) for the Amazon basin. For the Bolivian basin, Bayesian skyline plots suggest a relatively stable but gradually decreasing population size with a very recent demographic increase (approximately 12,000 years before present), and a coalescent effective population size of 34 thousand individuals (Fig. 2a), while analyses in IMA suggest 22 thousand effective individuals. Furthermore, formal hypothesis testing under the assumption of different demographic histories rejects constant population size in favor of exponential growth for the Amazon basin, but not for the Bolivian basin. The lack of statistical power to detect population growth in the Bolivian sub-basin is most likely to be due to the relatively recent and modest (threefold) increase in size, and that throughout the recorded co-

alescent history of *C. macropomum* in the Bolivian sub-basin, its coalescent effective population size increased less than 50%.

Demographic expansion events in the Amazon basin are inferred to have begun around 350,000 year ago with a second phase commencing around 125,000 years ago. These times broadly correspond to large climatic shifts, sea level changes, and potential marine transgressions that have occurred within the last 800,000 years (Hooghiemstra and Ran, 1994). The Pleistocene has seen large oscillations in the sea level relative to the present, most resulting in lower sea levels than the present, but others in higher sea levels than the present (Siddall et al., 2003, Crowley et al., 1986). Particularly large oscillations occurred in the last 1 million years (Billuops, 2004) and major high sea level rises occurred around 50,000 and 400,000 years ago (Liu and Herbert, 2004, Billuops, 2004). The observed differences in  $^{18}\text{O}/^{16}\text{O}$  ratios from present day ratios are indicative of around 100 m rise in sea level. Such a rise is sufficient to flood the entire eastern and central Amazon basin, and the eastern portions of the western Amazon basin (Nores, 1999). A more moderate temperature increase and associated sea level rise occurred in the last interglacial (ca. 130,000–115,000 ybp) with sea levels rising between 6 and 9 m above present levels (Kopp et al., 2009). A sea level rise could have either of two effects. It could have resulted in a marine incursion which would have destroyed freshwater habitat, but upon its retreat, it would have created new habitat suitable to colonization. Alternatively, a sea level rise would have backed up the Amazon River, creating a massive freshwater lake in the lower and central Amazon basin which would have provided new areas of lacustrine and floodplain habitat. The freshwater lake scenario is not unreasonable as the western Amazon basin was covered by a large freshwater lake in the Miocene (Marshall and Lundberg, 1996; Wesselingh et al., 2002), and the backing up of the Amazon occurs twice daily with incoming tides. There is no saltwater incursion; however, oceanic tides can be perceived up to 1000 km inland from the mouth of the Amazon (Goulding et al., 2003).

Pleistocene refugia and implicitly subsequent population expansions have been proposed for forest dwelling vertebrates of the Amazon basin (e.g. Ab'Saber, 1977; Haffer, 1969), and evidence from birds (Aleixo, 2004) and ants (Solomon et al., 2008) support these hypotheses, we lack information on the potential and possible effect of Pleistocene climate changes on aquatic faunas. Although our data suggest a correlation between demographic expansions and climate/sea level change, the generality of this pattern in aquatic vertebrates remains to be seen.

#### 5. Conclusions

Contrary to our initial expectations, the Madeira River rapids do not form a strong barrier to genetic exchange for *C. macropomum* of the two basins. Although specialists of lentic floodplains of large Amazonian rivers (Araújo, 1998) *C. macropomum* actively migrate upstream into the Bolivian basin. There is large amount of potentially passive gene flow from the Bolivian basin into the Amazon basin. At the same time gene flow between the two basins is a sufficiently weak evolutionary force to have allowed independent demographic trajectories to be followed by *C. macropomum* in the Amazon basin and the Bolivian sub-basin.

Recent plans by the Brazilian government to construct a dam at the site of the Santo Antônio rapid near Porto Velho and a second dam at the Jiráu rapid 130 km upstream of Porto Velho would disrupt the current population dynamics of *C. macropomum* and many other aquatic organisms. These dams are envisioned to create a navigable waterway for the export of soybeans from the southern Amazonian agricultural frontier via the Amazon River to international markets, and to provide hydroelectric power. This project will clearly have

serious consequences for the biodiversity of the Madeira River and its drainage system, altering and impacting major tracts of habitat, including habitat important for fishes, and disrupting ecological processes and population dynamics of many aquatic species.

## Acknowledgments

The idea of this publication was conceived during a RIIA (Red de Investigación sobre la Ictiofauna Amazónica) meeting in 2005 in Iquitos Peru. RIIA is a network between researchers from Brazil, Bolivia, Colombia, Peru, USA and France with a goal to try to understand the biology, biogeography and phylogeography of Amazonian fishes. This research was supported by the MCT/CNPq/PPG7 557090/2005-9, CNPq/CT-Amazonia 554057/2006-9 and CNPq/CT-Amazonia, the Instituto de Investigación en la Amazonia Peruana (IIAP, Peru), the Universidad Mayor San Andrés (UMSA – Bolivia), NSF HRD-0931659 and by the “Institut de Recherche pour le Développement” (IRD – France). Brazilian permits for field collection and molecular analyses were given by IBAMA/CGEN 11325-1. We thank María Doris Escobar L. for fisheries information, and Hernán López-Fernández and Ximena Velez-Zuazo for discussion.

## References

- Ab'Saber, A.N., 1977. Espaços ocupados pela expansão dos climas secos na América de Sul, po ocasião dos períodos glaciais quaternários. *Paleoclimas* 3, 1–18.
- Agnès, J.-F., Zent, F., Legros, O., Sellos, D., 2006. Phylogenetic relationships and phylogeography of the killifish species of the subgenus *Chromaphysomion* (Radda, 1971) in West Africa, inferred from mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* 40, 332–346.
- 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.
- Araújo-Lima, C.A.R.M., Goulding, M., 1998. Os Frutos do Tambaqui. *Ecologia, Conservação e Cultivo na Amazônia*. Sociedade Civil Mamirauá – MCT – CNPq, Belem, Brazil.
- Araújo, M.L.G., 1998. Biologia reprodutiva e pesca de *Potamotrygon* sp. C (Chondrichthyes–Potamotrygonidae), no médio rio Negro, Amazonas. *Biologia de Água Doce e Pesca Interior*. Universidade Federal do Amazonas (UFAM) and Instituto Nacional de Pesquisas da Amazônia (INPA). Manaus, AM, Brazil, p. 171.
- Banguera-Hinestroza, E., Cárdenas, H., Ruiz-García, M., Marmontel, M., Gaitán, E., Vázquez, R., García-Vallejo, F., 2002. Molecular identification of evolutionarily significant units in the Amazon River dolphin *Inia* sp. (Cetacea: Iniidae). *J. Hered.* 93, 312–322.
- Beerli, P., Felsenstein, J., 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. USA* 98, 4563–4568.
- Billuops, K., 2004. Low-down on a rhythmic high. *Nature* 427, 686–687.
- Campbell Jr., K.E., Heizler, M., Frailley, C.D., Romero Pittman, L., Prothero, D.R., 2001. Upper Cenozoic chronostratigraphy of the southwestern Amazon Basin. *Geology* 29, 595–598.
- Cantanhede, A.M., da Silva, V.M.F., Farias, I.P., Hrbek, T., Lazzarini, S.M., Alves-Gomes, J.A., 2005. Phylogeography and population genetics of the endangered Amazonian manatee, *Trichechus inunguis* Natterer, 1883 (Mammalia, Sirenia). *Mol. Ecol.* 14, 401–413.
- Crowley, T.J., Short, D.A., Mengel, J.G., North, G.R., 1986. Role of seasonality in the evolution of climate during the last 100 million years. *Science* 231, 579–584.
- Donaldson, K.A., Wilson Jr., R.R., 1999. Amphipanamic germinates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. *Mol. Phylogenet. Evol.* 13, 208–213.
- Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4, e88.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7, 214.
- Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1, 47–50.
- 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.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–472.
- Goulding, M., 1979. *Ecologia de Pesca do Rio Madeira*. INPA, Manaus, Brazil.
- Goulding, M., Barthem, R.B., Ferreira, E.J.G., 2003. *The Smithsonian Atlas of the Amazon*. Smithsonian Institution Press, Washington, DC.
- Goulding, M., Carvalho, M.L., 1982. Life history and management of the tambaqui (*Colossoma macropomum*, Characidae): an important amazonian food fish. *Rev. Bras. Zool.* 1, 107–133.
- Haffer, J., 1969. Speciation in Amazonian forest birds. *Science* 165, 131–166.
- Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.
- Hasegawa, M., Kishino, H., Yano, T.A., 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160–174.
- Hey, J., Nielsen, R., 2007. Integration with the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc. Natl. Acad. Sci. USA* 104, 2785–2790.
- Ho, S.Y.W., Phillips, M.J., Cooper, A., Drummond, A.J., 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* 22, 1561–1568.
- Hooghiemstra, H., Ran, E.T.H., 1994. Late Pliocene–Pleistocene high resolution pollen sequence of Colombia: an overview of climate change. *Quatern. Int.* 21, 63–80.
- Hoorn, C., Guerrero, J., Sarmiento, G.A., Lorente, M.A., 1995. Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. *Geology* 23, 237–240.
- Hubert, N., Duponchelle, F., Nuñez, J., Rivera, R., Bonhomme, F., Renno, J.-F., 2007. Isolation by distance and Pleistocene expansion of the lowland populations of the white piranha *Serrasalmus rhombeus*. *Mol. Ecol.* 16, 2488–2503.
- Isaac, V.J., Milstein, A., Ruffino, M.L., 1996. A pesca artesanal no baixo Amazonas: análise multivariada da captura por espécies. *Acta Amazon.* 26, 185–208.
- Isaac, V.J., Ruffino, M.L., 1996. Population dynamics of tambaqui, *Colossoma macropomum* Cuvier, in the lower Amazon. *Brazil. Fish. Manage. Ecol.* 1996, 315–333.
- Jobb, G., Haeseler, A.V., Strimmer, K., 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol. Biol.* 4, 1–9.
- Kass, R.E., Raftery, A.E., 1995. Bayes factors. *J. Am. Stat. Assoc.* 90, 773–795.
- Kopp, R.E., Simons, F.J., Maloof, A.C., Oppenheimer, M., 2009. Global and Local Sea Level During the Last Interglacial: A Probabilistic Assessment. [arXiv:0903.0752v1](http://arXiv.org/abs/0903.0752v1), pp. 1–23.
- Liu, Z., Herbert, T.D., 2004. High-latitude influence on the eastern equatorial Pacific climate in the early Pleistocene epoch. *Nature* 427, 720–723.
- Lovejoy, N.R., Collette, B.B., 2001. Phylogenetic relationships of New World needlefishes (Teleostei: Belontiidae) and the biogeography of transitions between marine and freshwater habitats. *Copeia* 2001, 324–338.
- Lundberg, J.G., Marshall, L.G., Guerrero, J., Horton, B., Malabarba, M.C.S.L., Wesselingh, F.P., 1998. The stage for Neotropical fish diversification: a history of tropical South American rivers. In: Malabarba, L.R., Reis, R.E., Vari, R.P., Lucena, Z.M.S., Lucena, C.A.S. (Eds.), *Phylogeny and Classification of Neotropical Fishes*. EDIPUCRS, Porto Alegre, Brazil, pp. 13–48.
- Marshall, L.G., Lundberg, J.G., 1996. Technical comment: miocene deposits in the Amazonian foreland basin. *Science* 273, 123–124.
- Nores, M., 1999. An alternative hypothesis for the origin of Amazonian bird diversity. *J. Biogeogr.* 26, 475–485.
- Pearse, D.E., Arndt, A.D., Valenzuela, N., Miller, B.A., Cantarelli, V.H., Sites Jr., J.W., 2006. Estimating population structure under nonequilibrium conditions in a conservation context: continent-wide population genetics of the giant Amazon River turtle, *Podocnemis expansa* (Chelonia; Podocnemididae). *Mol. Ecol.* 15, 985–1006.
- Pericchi, L.R., 2005. Model selection and hypothesis testing based on objective probabilities and Bayes Factors. In: Dey, D.K., Rao, C.R. (Eds.), *Bayesian Thinking: Modeling and Computation*. Elsevier, Amsterdam, The Netherlands, pp. 115–149.
- Posada, D., Crandall, K.A., 1998. MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Räsänen, M.E., Salo, J.S., Kalliola, R.J., 1987. Fluvial perturbation in the western Amazon basin: regulation by long-term sub-Andean tectonics. *Science* 238, 1398–1401.
- Reinert, T.R., Winter, K.A., 2002. Sustainability of harvested Pacú (*Colossoma macropomum*) populations in the northeastern Bolivian Amazon. *Conserv. Biol.* 16, 1344–1351.
- Rice, W.R., 1989. Analyzing tables of statistical tests. *Evolution* 43, 223–225.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular Cloning: A Laboratory Manual*. Cold Springs Harbor Laboratory Press, Cold Springs Harbor, NY.
- Sato, A., Takezaki, N., Tichy, H., Figueroa, F., Mayer, W.E., Klein, J., 2003. Origin and speciation of haplochromine fishes in East African crater lakes investigated by the analysis of their mtDNA, Mhc genes, and SINEs. *Mol. Biol. Evol.* 20, 1448–1462.
- Siddall, M., Rohling, E.J., Almogi-Labin, A., Hemleben, C., Meischner, D., Schmelzer, I., Smeed, D.A., 2003. Sea-level fluctuations during the last glacial cycle. *Nature* 423, 853–858.
- Sivasundar, A., Bermingham, E., Ortí, G., 2001. Population structure and biogeography of migratory freshwater fishes (*Prochilodus*: Characiformes) in major South American rivers. *Mol. Ecol.* 10, 407–417.
- Solomon, S.E., Bacci Jr., M., Martins Jr., J., Vinha, G.G., Mueller, U.G., 2008. Paleodistributions and comparative molecular phylogeography of leafcutter ants (*Atta* spp.) provide new insight into the origins of Amazonian diversity. *PLoS One* 3, e2738.
- Strasburg, J.L., Rieseberg, L.H., 2010. How robust are “Isolation with Migration” analyses to violations of the IM model? A simulation study. *Mol. Biol. Evol.* 27, 297–310.
- Suchard, M.A., Weiss, R.E., Sinsheimer, J.S., 2001. Bayesian selection of continuous-time Markov chain evolutionary models. *Mol. Biol. Evol.* 18, 1001–1013.
- Wesselingh, F.P., Räsänen, M.E., Irion, G., Vonhof, H.B., Kaandorp, R.J.G., Renema, W., Romero Pittman, L., Gingras, M., 2002. Lake Pebas: a palaeoecological reconstruction of a Miocene, long-lived lake complex in western Amazonia. *Cainozoic Res.* 1, 35–81.