



# A Multilocus Approach to Understanding Historical and Contemporary Demography of the Keystone Floodplain Species *Colossoma macropomum* (Teleostei: Characiformes)

## OPEN ACCESS

Maria da Conceição Freitas Santos<sup>1</sup>, Tomas Hrbek<sup>2</sup> and Izeni P. Farias<sup>2\*</sup>

### Edited by:

Rodrigo A. Torres,  
Universidade Federal de Pernambuco,  
Brazil

### Reviewed by:

Yessica Rico,  
Instituto de Ecología (INECOL), Mexico  
Fernanda Dotti do Prado,  
Universidade Estadual Paulista Júlio  
de Mesquita Filho (UNESP), Brazil  
Fabio Porto-Foresti,  
Universidade Estadual Paulista Júlio  
de Mesquita Filho (UNESP), Brazil

### \*Correspondence:

Izeni P. Farias  
izeni@evoamazon.net

### Specialty section:

This article was submitted to  
Evolutionary and Population Genetics,  
a section of the journal  
Frontiers in Genetics

**Received:** 08 November 2017

**Accepted:** 28 June 2018

**Published:** 14 August 2018

### Citation:

Santos MCF, Hrbek T and Farias IP  
(2018) A Multilocus Approach to  
Understanding Historical and  
Contemporary Demography of the  
Keystone Floodplain Species  
*Colossoma macropomum* (Teleostei:  
Characiformes). *Front. Genet.* 9:263.  
doi: 10.3389/fgene.2018.00263

<sup>1</sup> Departamento de Biologia, Universidade do Estado do Amazonas, Manaus, Brazil, <sup>2</sup> Laboratório de Evolução e Genética Animal, Departamento de Genética, Universidade Federal do Amazonas, Manaus, Brazil

We studied the natural populations of a flagship fish species of the Amazon, *Colossoma macropomum* which in recent years has been suffering from severe exploitation. Our aim was to investigate the existence or not of genetic differentiation across the wide area of its distribution and to investigate changes in its effective population size throughout its evolutionary history. We sampled individuals from 21 locations distributed throughout the Amazon basin. We analyzed 539 individuals for mitochondrial genes (control region and ATPase gene 6/8), generating 1,561 base pairs, and genotyped 604 individuals for 13 microsatellite loci obtaining, on average, 21.4 alleles per locus. Mean  $H_E$  was 0.78 suggesting moderate levels of genetic variability. AMOVA and other tests used to detect the population structure based on both markers indicate that *C. macropomum* comprises a single and large panmictic population in the main channel of the Solimões-Amazonas River basin, on the other hand localities in the headwaters of the tributaries Juruá, Purus, Madeira, Tapajós, and localities of black water, showed genetic structure. The greatest genetic differentiation was observed between the Brazilian Amazon basin and the Bolivian sub-basin with restricted genetic flow between the two basins. Demographic analyzes of mitochondrial genes indicated population expansion in the Brazilian and Bolivian Amazon basins during the Pleistocene, and microsatellite data indicated a population reduction during the Holocene. This shows that the historical demography of *C. macropomum* is highly dynamic. Conservation and management strategies should be designed to respect the existing population structure and minimize the effects of overfishing by limiting fisheries *C. macropomum* populations.

**Keywords:** tambaqui, microsatellites, mitochondrial DNA, genetic variability, gene flow, genetic structure, Amazon basin

## INTRODUCTION

The Amazon basin holds the largest diversity of fishes in the world. It is estimated that approximately 2,411 fish species occur there (Reis et al., 2016), with 1,089 species being endemic. Aquatic biodiversity of the Amazon basin is thought to be the consequence of diversification of modern fauna that occurred mainly during the Miocene (Lovejoy et al., 2010), driven, to a large extent by the establishment of the current hydroscape. Amazonian rivers also drain three principal geological formations, the Andes and the Guyana and Brazilian Shields, with consequences for the physicochemical properties of the waters draining these geological formations. Thus some of these rivers present physical barriers which limit gene flow between different sections of the river, further acting as agents of divergence (Hoorn et al., 2010). Naturally, all of these forces interact, producing an amazingly diverse ichthyofauna. Part of this ichthyofauna is also exploited as a fisheries resource that represent the production base of an economic sector that contributes more than US\$ 200 million per year to the economy of the Brazilian Amazon basin (Barthem and Fabr e, 2003). *Colossoma macropomum* (tambaqui) is on the top of the list of most important commercial species. This species also has an important ecological role, as it is an important disperser of seeds of trees and shrubs of the Amazonian floodplain (Ara ujo-Lima and Goulding, 1998). For all its commercial importance, this species has suffered over-exploitation of its natural stocks over the last years and today, juveniles account for most of the catch (Barthem and Goulding, 2007). The average size of the fish landed and sold in the main markets in the Amazon suggests that many individuals are captured before reaching sexual maturation, which occurs in females between 50 and 55 cm in length, at an estimated mean age of 3 years (Goulding and Carvalho, 1982; Isaac et al., 1996) based on the length/age relationship estimated from Bertalanffy's model by Isaac and Ruffino (1996).

*Colossoma macropomum* is found throughout almost the entire length of the Amazon River and most of its affluents, as well as in the Orinoco basin (Ara ujo-Lima and Goulding, 1998). Thus, the species is found in the three main Amazonian water types (white, clear, and black), as well as upstream and downstream of geographic barriers such as rapids (Goulding et al., 2003). *Colossoma macropomum* is thus an idea candidate for the study of and the understanding of the structuring patterns in the Amazon basin, which are important for the implementation of science-driven conservation measures.

Previous studies have found a high degree of genetic variability of populations of *C. macropomum* (Santos et al., 2007; Farias et al., 2010), indicating that overfishing has not yet affected genetic diversity of wild populations, nor were signals of population reduction detectable. The authors suggested that the absence of a genetic sign of population reduction was likely related to the large effective population size of the species. Moreover, *C. macropomum* has migratory behavior and moves through the rivers of the Amazon seasonally for the purposes of feeding and breeding (Ara ujo-Lima and Ruffino, 2004). The behavior of *C. macropomum* and the hydrological

dynamic of the floodplain habitat it predominantly occupies may partially explain the panmixia reported for this species. This apparent lack of population structuring is found throughout the mainstream of the Amazon basin, with the exception of individuals found upstream of the series of rapids delimiting the Bolivian sub-basin from the Amazon basin (Farias et al., 2010). The authors also suggested that populations of *C. macropomum* from Bolivian sub-basin were largely demographically stable, while the Brazilian Amazon basin populations evidenced a historical population growth from the Pleistocene onward.

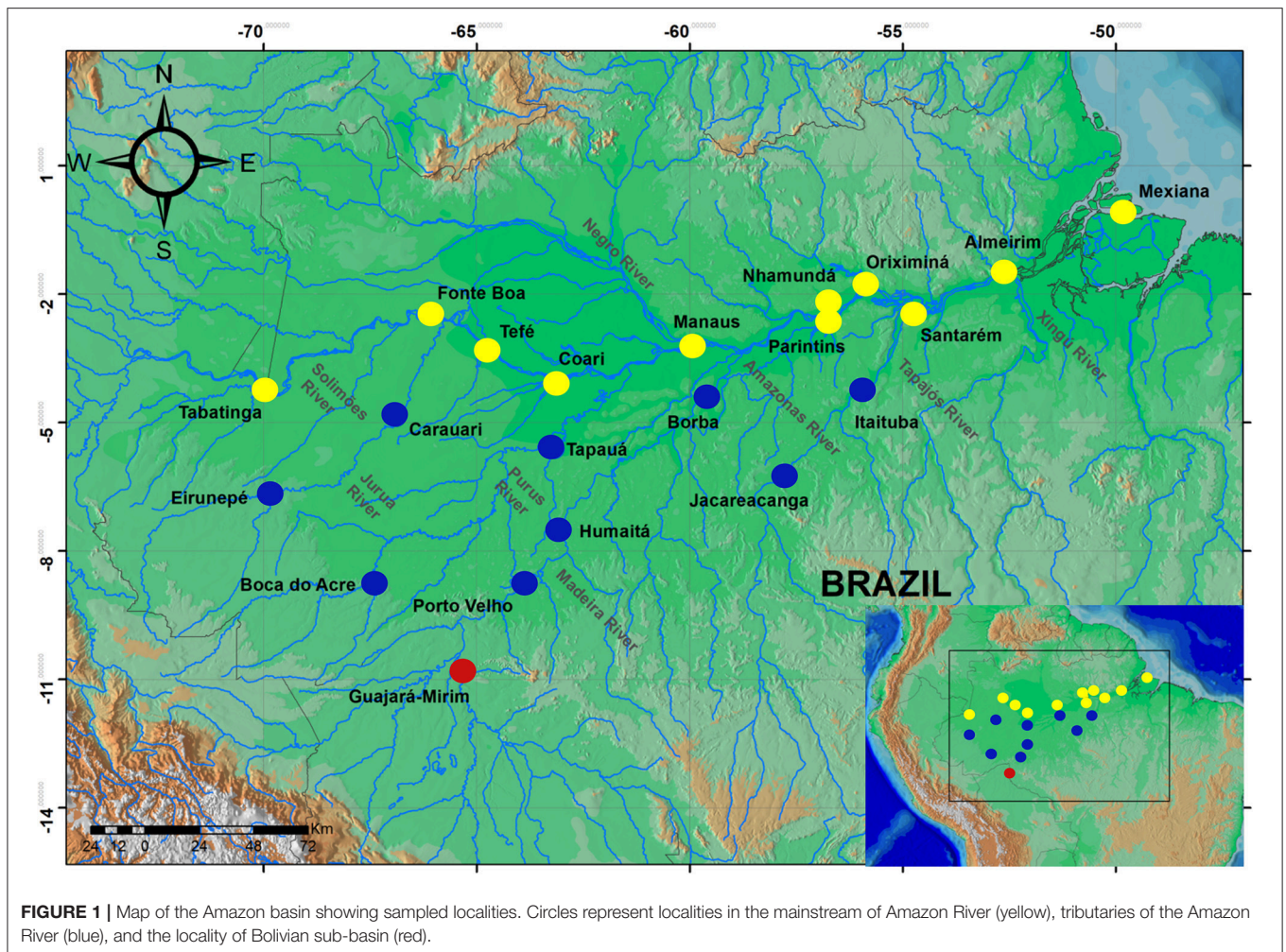
Knowledge of changes of effective population sizes of *C. macropomum* is important for understanding the demography of the species. In addition, robust estimates of population differentiation, are important for implementing conservation and management strategies. Therefore, the aim of the present study was to test the two hypotheses raised in previous studies of *C. macropomum* using samples from the entire area of distribution of *C. macropomum* in the Amazon basin, and using both nuclear-encoded microsatellites and mtDNA genes sequences. As the first hypothesis we test if *C. macropomum* populations are differentiated, considering: (i) samples of the mainstream of the Amazon River, as well as eight of its main tributaries; (ii) samples of all three major water types of the Amazon (white, clear, and black water) based on the classifications of Sioli (1984) and Venticinque et al. (2016); (iii) samples of water upstream and downstream of rapids in the Madeira and Tapaj s rivers. In the second hypothesis, historical and contemporary demographic approaches were used to test if *C. macropomum* underwent changes in the effective population size throughout its evolutionary history in the Amazon basin.

## MATERIALS AND METHODS

### Samples and Data Collection

A total of 637 samples of *Colossoma macropomum* were collected directly from artisanal fishers at 21 localities within the Brazilian Amazon basin and one locality in the Bolivian sub-basin (Figure 1). The adipose fin or fragment of muscle tissue was removed from between 20 (localities within the Brazilian Amazon Basin) and 69 individuals (within the Bolivian sub-basin) and then preserved in 100% ethanol for subsequent laboratory analyzes.

Total genomic DNA was extracted using Proteinase K/Phenol-chloroform/isoamyl alcohol protocol and precipitated with 70% ethanol (Sambrook et al., 1989). Approximately 50 to 100 ng of genomic DNA was used as a template for PCR reactions. We amplified the mitochondrial DNA control region (mtDNA control region) and the ATPase subunits 6 and 8, using the primers Chara\_LDloop and Chara\_RDloop; CMF2 and CMR2 (control region) and ATP 8.2\_L8331 and CO3.2\_H9236 (ATPase genes) listed in **Supplementary Table S1**. The PCR reactions for the two regions were performed in a final volume of 15  $\mu$ L containing 1.5  $\mu$ l of the forward and reverse primer (2 mM), 1.5  $\mu$ l of buffer (Tris-KCL 200 mM, pH 8.5), 1.5  $\mu$ L of 25 mM MgCl, 1.5  $\mu$ L of 25 mM dNTP, 0.3  $\mu$ L of 5 U/ $\mu$ L Taq polymerase and 6.2  $\mu$ L ddH<sub>2</sub>O. PCR conditions (for control region and ATPase gene) were as follows: denaturation at 94°C for 60 s, primer annealing



at 50°C for 30 s, primer extension at 68°C for 90 s, followed by a final extension at 68°C for 5 min. The first three steps were repeated 35 times.

Purification of the PCR products was performed using ExoSAP (Exonuclease Enzymes and Shrimp Phosphatase Alkaline). The samples were sequenced using the BigDye terminator v3 kit (ThermoFisher), following the manufacturer's protocol. Due to the size of the control region of *C. macropomum* (approximately 1,100 bp), each sample was sequenced in two steps, using the CMF2 (forward) and CMR2 (reverse) internal primers (**Supplementary Table S1**). For the ATPase gene only the primer ATP.8.2 (forward) was used. The precipitated product was resolved in the ABI 3130xl DNA Analysis System sequencer (ThermoFisher), according to the manufacturer's standard protocol.

Microsatellite genotypes were generated using a multiplex design (**Supplementary Table S2**) using 13 pairs of primers developed by Santos et al. (2009) for *C. macropomum*. The amplification conditions for each multiplex were: For three pairs of primers: 1.5 µl MgCl<sub>2</sub> (25 mM), 1.5 µl dNTPs (10 mM), 1.5 µl 10x buffer (100 mM Tris-HCl, 500 mM KCl), 1.0 µL of each

forward primer containing one of the two M13 tails (2 µM), 1.5 µL of each reverse primer, 1.5 µL of fluorescence-labeled M13f (FAM) primer, 0.7 µL of primer fluorescence-labeled M13r (HEX), 0.8 µl Taq DNA Polymerase (5 U/µl) and 1 µl DNA (50–100 ng), with a final volume of 14.5 µl. For two pairs of primers: 3.0 µl ultra pure water, 1.5 µl MgCl<sub>2</sub> (25 mM), 1.5 µl dNTPs (10 mM), 1.5 µl 10x buffer (100 mM Tris-HCl, 500 mM KCl), 1.0 µL of each forward primer containing one of the M13 tails (2 µM), 1.5 µL of each reverse primer, 1.5 µL of the fluorescently labeled M13f primer (FAM), or 0.7 µL of fluorescently labeled M13r primer (HEX), 0.6 µl of Taq DNA Polymerase (5 U/µl), and 1 µl of the DNA (50–100 ng), with a final volume of 14 µl.

PCR conditions were as follows: denaturation at 94°C for 20 s, primer annealing at 60–65°C (depending on the primer combination) for 20 s, and extension at 68°C for 30 s, repeated for 30 times, followed by another cycle for annealing the M13 primers with the following conditions: denaturation at 94°C for 20 s, annealing of the M13 fluorescence-labeled primer at 50°C for 20 s, and extension at 68°C for 30 s, repeated for 20 times, with final extension of 30 min at 68°C.

For the genotyping reaction the PCR products were diluted to between 10–50  $\mu\text{L}$  with ultra-pure water depending on the intensity of PCR products on an agarose gel. For each 1  $\mu\text{L}$  of diluted product, 8.0  $\mu\text{L}$  of Hi-Di formamide (ThermoFisher, Inc.), and 1.0  $\mu\text{L}$  6-carboxy-X-rhodamine (ROX) size standard from DeWoody et al. (2004) were added. The samples were genotyped in ABI 3130xl automatic sequencer (ThermoFisher, Inc.) and allele sizes (in base pairs) were estimated in *GeneMapper*<sup>TM</sup> software version 4.0 (ThermoFisher, Inc.). Matrix of genotypes is available at <https://github.com/legalLab/datasets>.

The sequences of the control region and subunits 6 and 8 of the ATPase gene were verified, edited and aligned in the program BIOEDIT v7.0.5 (Hall, 1999). The ATPase genes were translated into hypothetical amino acids in the program MEGA 6.0 (Tamura et al., 2013) to verify the presence of any unexpected stop codons. Sequences were deposited in the GenBank under accession numbers MH514288–MH514827 for control region and MH520124–MH520663 for ATPase genes.

## Mitochondrial DNA Analyses

The existence of population structure was tested for using the Analysis Molecular Variance (AMOVA) implemented in the program ARLEQUIN v3.5 (Excoffier and Lischer, 2010). We analyzed three datasets: (1) all 21 locations were analyzed as a single hierarchical level; (2) the Guajar -Mirim (Bolivian basin) locality was removed from the data matrix; (3) tributaries vs. locations of the main channel, of Amazon River. Pairwise  $\Phi_{ST}$  were also estimated, and statistical significance was corrected for multiple comparisons (Rice, 1989). We also tested for population structuring using the Spatial Analysis of Molecular Variance (SAMOVA) (Dupanloup et al., 2002), and the Bayesian Analysis of Genetic Population Structure (BAPS) (Mantel, 1967; Corander and Tang, 2007). Spatial structuring was tested using the Mantel test as implemented in ARLEQUIN v3.51 (Excoffier and Lischer, 2010).

In order to investigate patterns of change in *C. macropomum* historical effective population sizes, we carried out a Bayesian Skyline plot analyses in the program BEAST v.1.8.4 (Drummond and Rambaut, 2007). We collected 50,000,000 Monte Carlo Markov Chain steps (MCMC), discarded the first 5,000,000 steps as burnin, and subsequently sampled every 1,000th step, retaining 45,000 topologies. The HKY85 (Hasegawa et al., 1985) model of molecular evolution was selected as the best fitting model in the program Modeltest. We estimated a genetic network of mtDNA haplotypes from all samples using Network (<http://www.fluxus-engineering.com/>) using the median-joining algorithm.

To convert the results of the coalescent analyzes into years and effective number of individuals, we assumed a three-year generation time (Goulding and Carvalho, 1982; Isaac and Ruffino, 1996), and a rate of molecular evolution of  $2.0 \times 10^{-8}$  mutations per site and per year (Farias et al., 2010).

The Tajima's *D* (Tajima, 1989) and the Fu's *F<sub>s</sub>* (Fu, 1997) tests were used to examine whether populations are at a mutation-drift equilibrium assuming no selective differences among haplotypes. Both tests were performed using the program ARLEQUIN v3.5 (Excoffier and Lischer, 2010). Demographic history may also

be inferred from frequency distribution of pairwise haplotype differences. In populations that are at a demographic equilibrium, the distribution of differences are generally multimodal, while populations that have undergone recent expansion or reduction typically have a unimodal distribution (Slatkin and Hudson, 1991). In order to distinguish population reduction and expansion, we used the results of two tests. The first test evaluates the distribution of the sum of the squares of differences (SSD) between the mismatch distribution observed for each locality and the expected distribution for a null expansion model, where significant values for SSD indicate deviations from the population expansion model (Schneider and Excoffier, 1999). The other test is based on the Harpending inequality index ( $H_{ri} = r$ ) (Harpending, 1994), which quantifies the variance of the mismatch distribution, assuming that the mismatch distribution is unimodal. These analyzes were performed in the programs DNASP v5.0 (Librado and Rozas, 2009) and ARLEQUIN v3.5 (Excoffier and Lischer, 2010); significance was tested via 10,000 permutations with a  $P = 0.05$  cut-off.

## Microsatellite DNA Analyses

The data matrix with allele sizes was verified for the occurrence of null alleles, allelic stutter, and large allele dropout in the program MICRO-CHECKER (Van Oosterhout et al., 2004). The number of alleles (*A*), observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosities, gene diversity (*h*), nucleotide diversity ( $\pi$ ), linkage disequilibrium (LD) between pairs of loci and the Hardy-Weinberg equilibrium (HWE) were calculated. All these parameters were estimated using ARLEQUIN v3.5 (Excoffier and Lischer, 2010). Considering that some of these estimates suffer influence of sample size (Leberg, 2002), we implemented a rarefaction analysis and calculated allelic richness (AR) and private allelic richness (PAR) in the program HP-Rare (Kalinowski, 2005), so that the number of alleles and allelic richness estimates could be compared between localities. Additionally, we estimated the inbreeding coefficient ( $F_{IS}$ ) for each sample. The effective population size ( $N_e$ ) for each population was estimated using the LD method (Waples and Do, 2008) as implemented in NeEstimator 2.0 (Do et al., 2014). The  $N_e$  estimates are equivalent to the effective number of breeders that produced offspring during a certain period of time and assuming that sample sizes are not representative of the entire generation (Palstra and Fraser, 2012). In all instances, significance levels for tests involving multiple comparisons were adjusted using the sequential Bonferroni correction (Rice, 1989).

The overall genetic structuring was estimated using the analysis of molecular variance (AMOVA—Excoffier et al., 1992) performed in the program Arlequin v3.5 (Excoffier and Lischer, 2010). We also analyzed three datasets: (1) all 21 locations were analyzed as a single hierarchical level; (2) the Guajar -Mirim (Bolivian basin) locality was removed from the data matrix; (3) tributaries vs. locations of the main channel of Amazon River. Genetic differentiation between pairs of populations was estimated using  $F_{ST}$ . Additionally, pairwise genetic differentiation between populations was estimated using Hedrick's  $G_{ST}$  (Hedrick, 2005) based on the empirical Bayes (EB)  $G_{ST}$  estimator (Kitada et al., 2007) suitable for high gene flow

species (Kitada et al., 2017), using the FinePop 1.3.0 package (Kitada et al., 2017) implemented in the R statistical language (R Development Core Team, 2011).

We used SAMOVA (Dupanloup et al., 2002) to infer spatial population structure and STRUCTURE v2.3.4 (Pritchard et al., 2000) to identify biological populations. For STRUCTURE analyses we used the admixture and correlated allelic frequencies models with and without location prior, and we tested one to 20 groups ( $K = 1-20$ ). The analysis was run with 1,000,000 MCMC step, discarding the first 100,000 steps. The Isolation by Distance (IBD) was tested via a correlation between genetic and geographical distance using the Mantel test implemented in the Software Arlequin (Excoffier and Lischer, 2010). In addition, we also used a multivariate approaches implementing the Discriminant Analysis of Principal Components (DAPC; Jombart et al., 2010) to cluster genotypes using the R package Adegenet (Jombart, 2008).

Coalescent analyses implemented in the program IMA2 (Hey and Nielsen, 2007) were used to partition allele sharing between populations due to ongoing geneflow and ancestral haplotype sharing. We estimated the parameters  $t$  (splitting time),  $m$  (migration rates), and theta ( $\theta$ ) where  $\theta = 2Ne\mu$ . We sampled 20,000,000 Monte Carlo Chain Markov Chain Monte Carlo (MCMC) generations after discarding the first 1,000,000 generations as burn-in. Two independent runs were carried out with different starting points, in order to verify convergence. The two independent runs converged and thus were combined and the parameters  $\theta$ ,  $m$ , and  $t$  were estimated. Then, these were converted into demographic parameters: contemporary effective population size, number of migrants per generation, and time of divergence of the populations in generations. The analyses using IMA2 were performed with pairs of sampling sites located at the geographic extremes along the main channel of the Amazon River and between upstream and downstream localities of principal tributaries. We also estimated geneflow using the program MIGRATE 3.1.6 (Beerli and Felsenstein, 2001), where for diploid data  $\theta = 4Ne\mu$  and  $M = m/\mu$  migration rate ratio and mutation rate. For the Bayesian analysis, we ran ten short chains, sampling each chain 10,000 times. Then we ran six long chains of 2,000,000 steps, sampling each chain 200,000 times and discarding the first 2,000 samples. The runs were replicated, and the convergence between the chains was evaluated using the Gelman-Rubin statistic implemented in the program. We estimated the historical migration rates ( $M$ ) between the localities and the relative number of migrants per generation  $Nm = M\theta/2$ . To convert the results into biological information, we assumed a 3-year generation time for *C. macropomum* (previously justified with mtDNA data), and a mutation rate  $\mu = 5 \times 10^{-4}$  (mean rate of evolution of microsatellites, Di Rienzo et al., 1994).

To detect, quantify, and date the historical and contemporaneous demographic changes in *C. macropomum* populations we implemented the coalescent sampler implemented in the program MSVar v1.3 (Beaumont, 1999; Storz et al., 2002). We ran 11 independent parallel chains sampling every 1,000th proposal to collect 20,000 proposals in the MCMC chain in each parallel run. Priors for current and historical population size means and variances were equal,

and variances encompassed three orders of magnitude. Prior for mean time of population size change was set at 1,000 with variance encompassing time range from 1,000,000 to 0 generations ago. The runs were evaluated for convergence and were pooled to provide an estimate of current and historical effective population size. Convergence was assessed using the Gelman-Rubin criterion (Gelman and Rubin, 1992) and the test of alternative hypotheses (population decline vs. stable population size) as suggested by Beaumont (1999) was tested using Bayes factors. Calculations and plots were performed in the R statistical programming language (R Development Core Team, 2011) using the packages CODA and ggplot2.

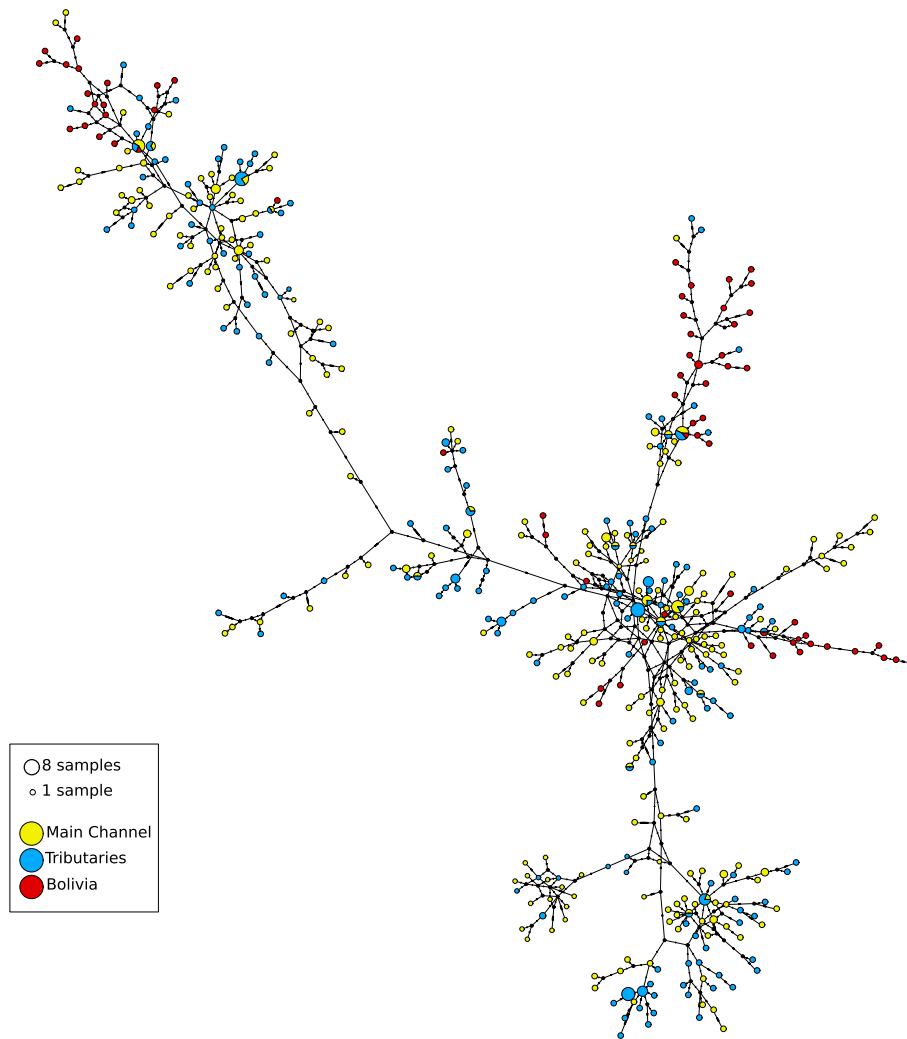
In order to verify reduction in effective population size ( $Ne$ ) or bottleneck effect, we tested for heterozygosity excess in the program BOTTLENECK (Piry et al., 1999) using three different mutation models: the stepwise mutation model, SMM (Ohta and Kimura, 1973); the two-phase model, TPM (Di Rienzo et al., 1994); and the infinite alleles model, IAM (Estoup et al., 1995). Genetic bottlenecks can also leave a signature in the ratio of number alleles to the allele size range (the  $M$ -ratio), where a bottleneck depletes the number of alleles faster than reducing allelic size range of the microsatellite (Garza and Williamson, 2001). We calculated the  $M$ -ratio using ARLEQUIN, and considered a reduction in the number of alleles to occur when  $M < 0.68$ , as suggested by Garza and Williamson (2001).

## RESULTS

### Genetic Diversity of *C. macropomum*

Eight hundred and thirty-nine base pairs from the control region and 732 base pairs from the ATPase6/8 gene were obtained from 539 individuals. Approximately 5% of the samples were resequenced to confirm the sequences obtained. The sequences of the mitochondrial gene fragments were concatenated, resulting in a total of 1,561 base pairs. A total of 444 haplotypes were found, 400 of which were unique. The haplotype network showed numerous reticulations between haplotypes (Figure 2). There was very little clustering among the haplotypes found at most localities, implying in a high degree of gene exchange. High and relatively homogeneous values of haplotype diversity was found, ranging from  $h = 0.895$  in Porto Velho to  $h = 1.000$  at nine of the 21 sampled localities (Table 1).

A total of 604 individuals were genotyped for 13 microsatellite loci. The data revealed no evidence of allelic stutters or large allele dropouts (genotyping errors) and neither linkage disequilibrium (LD). However, deviation from the Hardy-Weinberg equilibrium (EHW) was observed at the Cm1E3 locus in 17 of the 21 sampled localities and the locus was therefore removed from the population analyses. Genetic variability parameters were quite homogeneous among the individuals from different localities (See Table 1). The expected heterozygosity ( $H_E$ ) ranged from 0.714 in Guajará-Mirim to 0.797 in Eirunepé (Juruá River). Mean heterozygosity was  $0.777 \pm 0.395$  for all loci and all locations (Supplementary Table S3). Allelic richness varies from 5.32 alleles in Guajará-Mirim (Guaporé River) to 6.53 in Carauari (Juruá River). The endogamy coefficient ( $F_{IS}$ ) ranged from 0.023 to 0.144 and was significant for all localities.



**FIGURE 2 |** Haplotype network of *Colossoma macropomum* haplotypes estimated using Network. Each line represents a single mutation. Circle size correspond to the number of observations, and missing haplotypes remain unfilled. Colors corresponds to observation of a haplotype in one of three main regions (yellow = mainstream of Amazon River, blue = tributaries of the Amazon River, and red = Bolivian sub-basin).

## Population Structure

AMOVA of the mtDNA data demonstrated that more than 90% of genetic variance was within sampling sites. When AMOVA was performed without Guajar-Mirim (Bolivian sub-basin),  $\Phi_{ST}$  was 0.032, which is lower than the  $\Phi_{ST} = 0.062$  found in the analysis including all sampling sites. Considering the tributaries vs. locations of the main channel,  $\Phi_{ST}$  was 0.052. Nonetheless, AMOVA was significant for all three datasets analyzed. The result of the Mantel test was non-significant ( $r = 0.1587$ ,  $P = 0.157$ ), demonstrating no correlation between the genetic distances of the sampling sites and their respective geographic distances.

Global AMOVA of microsatellite data resulted in partitioning more than 98% variance within sites ( $F_{ST} = 0.0111$ ,  $P = 0.00124$ ). Based in this result, additional AMOVA tests were implemented assuming two main groups: Amazon basin vs. Bolivia basin and, within the Amazon basin, tributaries vs. main channel. The

AMOVA results were significant for both analyses ( $F_{ST} = 0.0192$ ,  $P = 0.0478$ ;  $F_{ST} = 0.0086$ ,  $P = 0.9277$ ;  $F_{ST} = 0.0026$ ,  $P = 0.0004$ ), respectively.

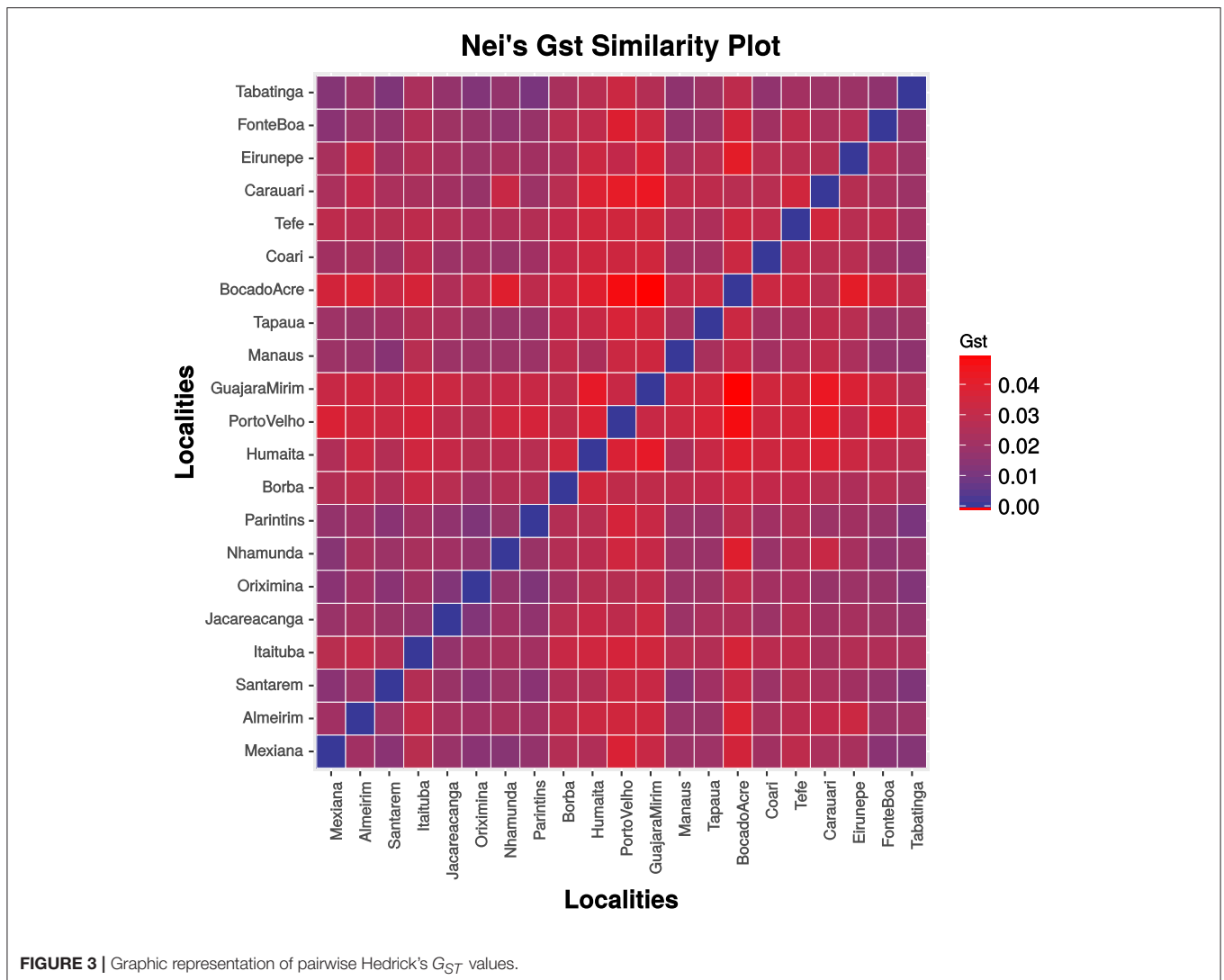
The matrix of pairwise  $G_{ST}$  (Figure 3, Supplementary Table S4) and  $F_{ST}$  values (Supplementary Table S5) were congruent and indicated population structuring. Significant values were observed for almost all comparisons involving the Bolivian sub-basin (Guapore River), and the Madeira (Porto Velho, Humaita, Borba), upper Jurua (Eirunep), upper Purus (Boca do Acre), and upper Tapajs (Itaituba, Jacareacanga) rivers. The Mantel test was significant ( $r = 0.34260$ ,  $P < 0.05$ ) only when the Bolivian sub-basin was included in the analysis.

SAMOVA analyses indicated the existence of two geographic groups, one group comprised of Guajar-Mirim and another group comprising all remaining localities. At  $K = 2$  (Group 1: Guajar-Mirim; Group 2: other sampling sites in the Brazilian

**TABLE 1** | Main genetic pattern estimates from mtDNA and microsatellites of *C. macropomum* individuals from the Amazon drainage and Bolivian sub-basin.

Basin	Locality	River	Water type*	mtDNA										Microsatellites				
				N	H	S	h	$\pi$	N	A	Ana	AR	PAR	AGD	Ho-He	FIS	Ne (Cis)	p_SMM
	Mexiana	Amazon	White	15	14	1.000 ± 0.024	0.009 ± 0.005	31	139	11.5	6.23	0.12	0.791 ± 0.407	0.779-0.791	<b>0.049</b>	4183.1 (140.7-inf)	<b>0.040</b>	0.71
	Almeirim	Amazon	White	21	20	1.000 ± 0.016	0.009 ± 0.005	36	150	12.5	6.31	0.13	0.781 ± 0.402	0.754-0.781	<b>0.070</b>	-234.4 (1094.3-inf)	<b>0.008</b>	0.80
	Santarém	Lower Tapajós	Clear	26	26	1.000 ± 0.010	0.012 ± 0.006	36	143	11.9	6.29	0.13	0.771 ± 0.397	0.680-0.771	<b>0.144</b>	-266.3 (845.1-inf)	0.243	0.78
	Itaituba	Tapajós	Clear	32	24	1.000 ± 0.014	0.011 ± 0.006	32	138	11.5	6.01	0.20	0.762 ± 0.393	0.734-0.762	<b>0.102</b>	476.9 (102.6-inf)	<b>0.011</b>	0.68
	Jacareacanga	Tapajós	Clear	29	19	0.958 ± 0.020	0.009 ± 0.005	30	134	11.1	6.04	0.13	0.757 ± 0.391	0.730-0.757	<b>0.078</b>	-2610.5 (169.4-inf)	0.277	0.79
	Oriximá	Lower Trombetas	Clear	34	34	1.000 ± 0.007	0.012 ± 0.006	32	139	11.5	6.04	0.12	0.773 ± 0.399	0.715-0.773	<b>0.102</b>	275.7 (89.5-inf)	0.118	0.79
	Nhamundá	Lower Nhamundá	Black	22	21	1.000 ± 0.015	0.011 ± 0.005	22	127	10.5	6.25	0.19	0.766 ± 0.398	0.727-0.766	<b>0.078</b>	100.3 (45.1-inf)	<b>0.034</b>	0.74
	Parintins	Amazon	White	20	20	1.000 ± 0.015	0.012 ± 0.006	35	148	12.3	6.13	0.17	0.768 ± 0.396	0.719-0.768	<b>0.081</b>	96.2 (57.9-233)	<b>0.012</b>	0.76
Amazon	Borba	Madeira	White	19	19	1.000 ± 0.017	0.013 ± 0.007	23	131	10.9	6.13	0.21	0.776 ± 0.403	0.775-0.776	<b>0.023</b>	22.3 (16.4-32.3)	<b>0.037</b>	0.74
	Humaitá	Madeira	White	22	20	1.000 ± 0.016	0.011 ± 0.005	22	110	9.10	5.66	0.10	0.780 ± 0.405	0.746-0.780	<b>0.077</b>	11.6 (8.8-15.6)	<b>0.042</b>	0.75
	Porto Velho	Madeira	White	21	14	0.895 ± 0.061	0.011 ± 0.006	21	103	8.5	5.59	0.15	0.737 ± 0.385	0.674-0.737	<b>0.143</b>	39.7 (24.3-86.4)	0.249	0.64
	Manaus	lower Negro	Black	30	29	0.997 ± 0.009	0.012 ± 0.006	35	152	12.6	6.38	0.14	0.785 ± 0.404	0.738-0.785	<b>0.090</b>	-437.1 (333.5-inf)	<b>0.002</b>	0.77
	Tapauá	Purus	White	28	28	1.000 ± 0.009	0.014 ± 0.007	28	144	12.0	6.37	0.18	0.766 ± 0.396	0.755-0.766	<b>0.056</b>	-168.7 (73.5-inf)	<b>0.037</b>	0.73
	Boca do Acre	Purus	White	21	21	1.000 ± 0.014	0.013 ± 0.007	21	130	10.8	6.35	0.27	0.771 ± 0.401	0.742-0.771	<b>0.076</b>	240.3 (62.6-inf)	0.112	0.73
	Coari	Lake Coari	Black	20	20	1.000 ± 0.015	0.012 ± 0.006	29	141	11.7	6.20	0.16	0.762 ± 0.394	0.683-0.762	<b>0.123</b>	924.8 (119.3-inf)	<b>0.002</b>	0.78
	Tefé	Lake Tefé	Black	18	16	0.986 ± 0.022	0.012 ± 0.006	20	117	9.70	6.04	0.16	0.751 ± 0.392	0.733-0.751	<b>0.058</b>	652.6 (86.4-inf)	0.439	0.64
	Carauari	Juruá	White	25	25	1.000 ± 0.011	0.013 ± 0.007	25	141	11.6	6.53	0.18	0.790 ± 0.409	0.760-0.790	<b>0.067</b>	-448.2 (174.2-inf)	0.244	0.71
	Eirunepé	Juruá	White	16	15	0.991 ± 0.025	0.010 ± 0.005	20	113	9.40	6.00	0.08	0.797 ± 0.414	0.754-0.797	<b>0.066</b>	47.9 (30.1-100.3)	0.248	0.73
	Fonte Boa	Amazon	White	20	18	0.989 ± 0.019	0.011 ± 0.005	31	142	11.8	6.18	0.16	0.773 ± 0.399	0.736-0.773	<b>0.050</b>	-309.0 (250.3-inf)	<b>0.036</b>	0.73
	Tabatinga	Amazon	White	31	30	0.997 ± 0.008	0.013 ± 0.007	37	146	12.1	6.15	0.15	0.779 ± 0.401	0.734-0.779	<b>0.077</b>	40.0 (77.0-527.4)	0.290	0.76
Bolivian	Guajará-Mirim	Guaporé	White	69	68	0.999 ± 0.002	0.012 ± 0.006	38	123	10.2	5.32	0.09	0.714 ± 0.369	0.677-0.714	<b>0.087</b>	-329.8 (443.5-inf)	<b>0.038</b>	0.75

*N*, number of individuals; *H*, number of haplotypes; *S*, singletons; *h*, gene diversity;  $\pi$ , nucleotide diversity; *A*, number of alleles; *A<sub>95%</sub>*, average number of alleles per locus; *AR*, allelic richness; *PAR*, private allelic richness; *AGD*, average genetic diversity; *Ho*, observed heterozygosity; *He*, expected heterozygosity; *F<sub>is</sub>*, endogamy index; *Ne [C]*, effective population size [95% confidence intervals]; *p\_SMM*= Stepwise Mutation Model from Bottleneck program, *Mvalue*= *M*-ratio using ARLEQUIN (Garza and Williamson index). Values in bold are significant at 0.05. \*Classification of water types based in Slioi (1984) and Venticinque et al. (2016).



Amazon basin),  $F_{CT}$  was maximized for both the mtDNA and microsatellite datasets, but with significant support only for the mtDNA data. Bayesian analyses implemented in STRUCTURE v2.3.4 identified three biological groups. The highest posterior probability was  $\text{LnP}(K = 3) = -31766.0000$ . The three populations comprised individuals from the Bolivian sub-basin and the Brazilian Amazon basin. Individuals from the three localities in the Madeira River showed a linear gradient of admixture between these two populations, and a contribution of an additional biological group principally within the Humaitá locality (Figure 4). Results based on DAPC analysis displayed a general pattern of low genetic differentiation (Figure 5), however, as observed in the previous results, some individuals from upper Madeira River and Guaporé drainage are partially differentiated from the other localities.

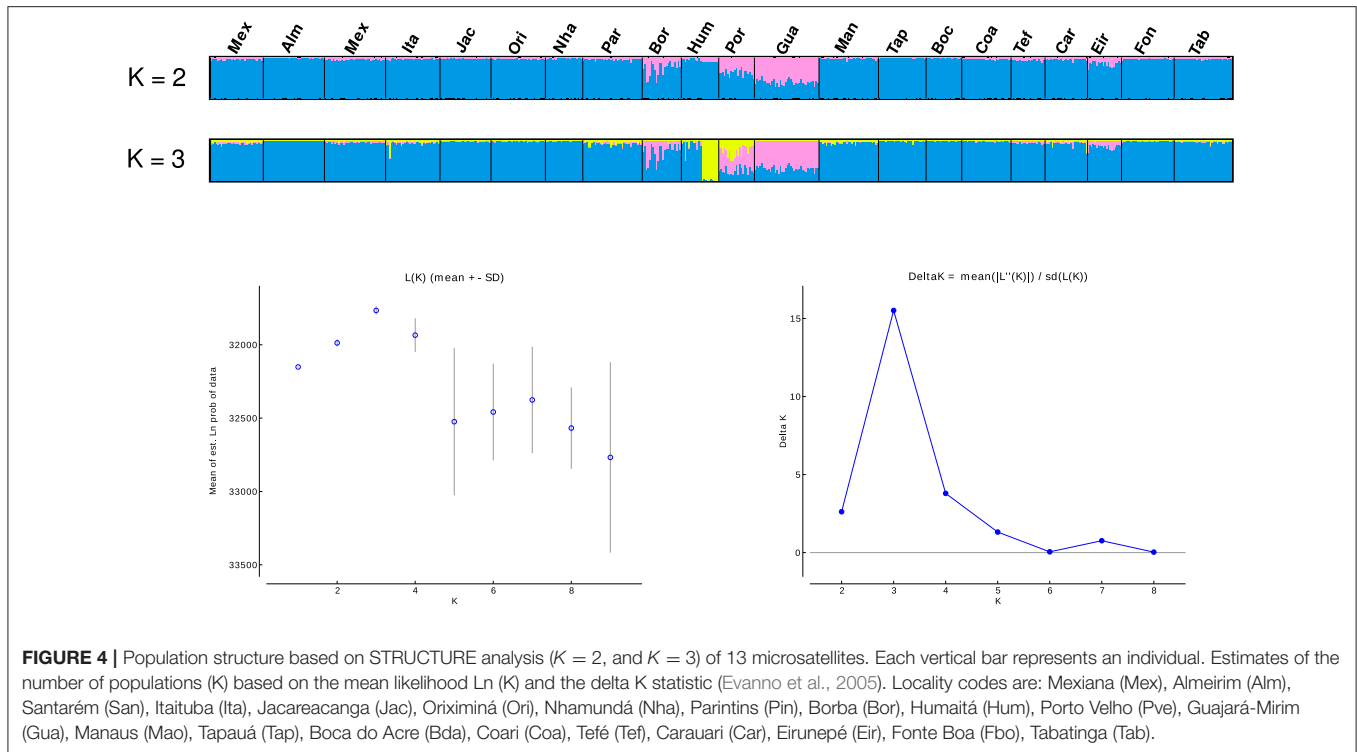
## Gene Flow

Results of the isolation-with-migration analyses using microsatellite data are in **Supplementary Table S6**. Thus,

Mexiana and Tabatinga (on the Amazon River) were paired, and a group denominated the main channel was formed by randomly sampling 30 individuals from among the sampling sites of the main Amazon River channel, which was then analyzed with upper-most tributary localities: Jacareacanga (Tapajós River), Guajará-Mirim (Bolivian sub-basin), Boca do Acre (Purus River), and Eirunepé (Juruá River) (Table 2). The result indicated bidirectional gene flow between all localities. In all cases, the direction of migration from upstream areas of tributaries to the central Amazon basin predominated except in the case of the Jacareacanga locality.

The results of MIGRATE analyses supported substantial levels of geneflow between sites in the main stream of the Amazon, but reduced gene flow levels between localities at tributary headwaters, and of the Madeira River (Table 3). The genetic parameters estimated for *C. macropomum* in MIGRATE version 3.1.3 inferred from microsatellites data is reported in **Supplementary Table S7**.





## Population Demography

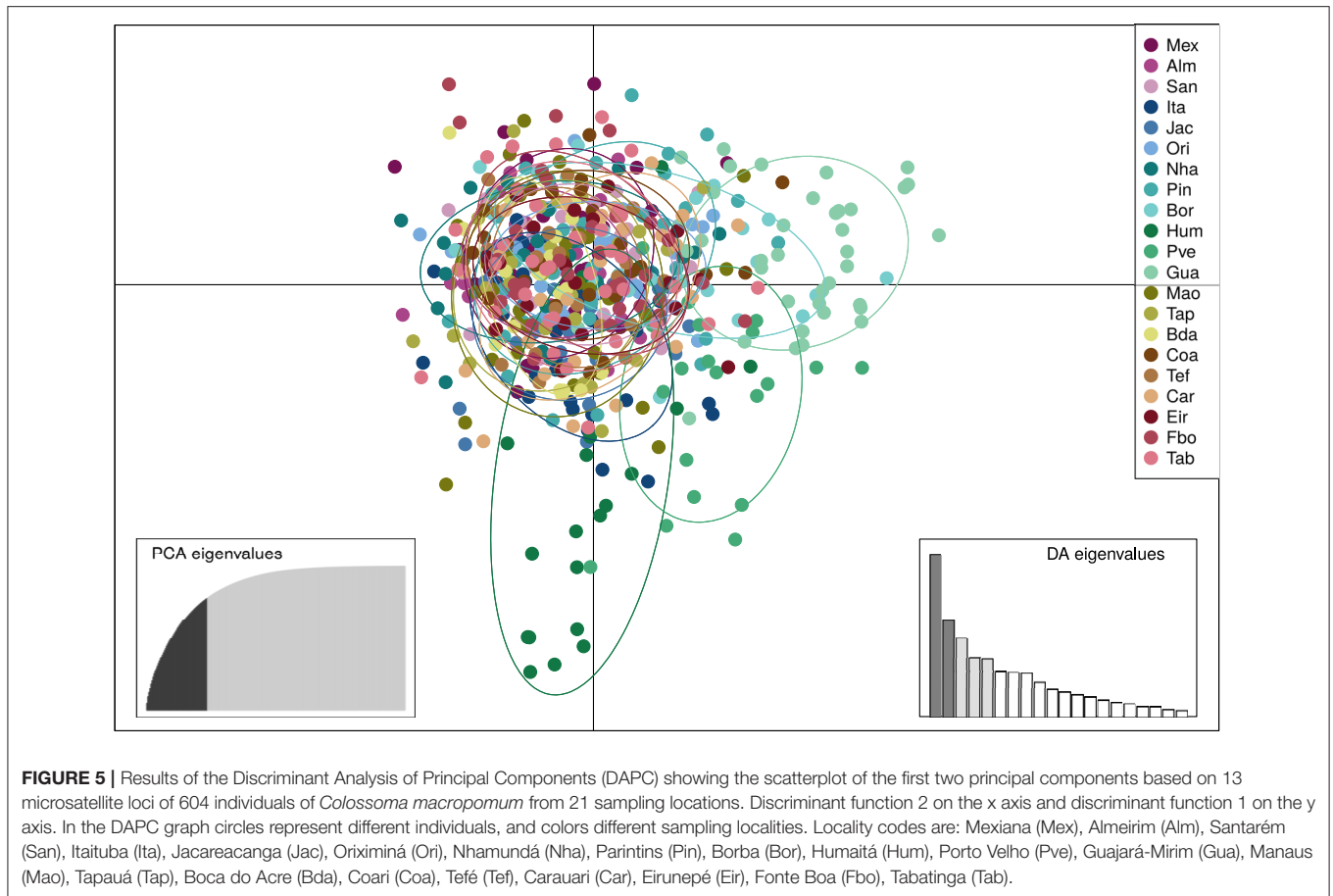
Using the genetic parameters from IMA2 analyses (Table 2), the coalescent effective population size did not differ substantially for almost all pair of localities examined. As a whole, effective population sizes were of thousands of individuals, with the exception of Guajará-Mirim.

The Bayesian skyline plot for *C. macropomum* from the Brazilian Amazon demonstrated a strong sign of population expansion, which began slowly approximately 3,000,000 years ago. Demographic growth accelerated considerably approximately 450,000 years ago, with a weak signature of a recent population decline. From the beginning of the initial growth phase, the population size of this species have increased two orders of magnitude from approximately little more than 750 thousand to 75 million individuals in the coalescent history of the populations sampled (Figure 6). Population from Bolivian sub-basin shows demographic growth beginning at 500,000 years ago, with a signature of recent population stability.

Population expansion was also supported by Harpending's raggedness index, which was significantly small ( $r = 0.0046$ ,  $P = 1.0000$ ) considering all samples as well as when considering the two basins separately (Amazon basin:  $r = 0.00060$ ,  $P = 1.0000$ ; Bolivian basin:  $r = 0.0018$ ,  $P = 0.9990$ ) (Table 4). These indexes statistically support the inference of population expansion based on the observation of the distribution of mismatch distribution (Harpending, 1994) for all samples of *C. macropomum*. However, when mismatch distribution was investigated for the basins separately, unimodal distribution was found only within the Amazon basin, whereas multimodal distribution was found for the Bolivian basin (results not shown),

which fits a pattern expected under stable population size, although this was not supported by Harpending's raggedness index. The sum of squared deviations (Schneider and Excoffier, 1999) was non-significant for the overall sample ( $SSD = 0.0020$ ,  $P = 0.6260$ ) as well as the inference performed for the two basins separately (Amazon basin:  $SSD = 0.0021$ ,  $P = 0.6430$ ; Bolivian basin:  $SSD = 0.0047$ ,  $P = 0.8420$ ). Thus, these values neither support nor reject the null hypothesis of a demographic population expansion for *C. macropomum* (Table 4). The Tajima's  $D$  was non-significant for all sampling localities. The same was found for Fu's  $F_s$ . When considering the basins separately, Fu's  $F_s$  was significantly negative for the Bolivian basin, suggesting a population expansion.

Results based in MSVar analyses show that historically *C. macropomum* has undergone a pronounced population decline in both the Amazon and Bolivian basins (Figure 7). The mean estimated ancestral effective population size were at approximately 100,000 individuals, declining recently to approximately 5,000 individuals, an approximately 1.5 orders of magnitude decrease. Demographic decrease was strongly supported ( $BF = 1206$ ) and occurred with 0.9992 probability. Population decline started at approximately 10,000 (Amazon) and 2,500 (Bolivia) years ago. Signs of population reduction in the Bottleneck program were significant for 18 locations under the SMM model, however  $M$ value showed no signal of population reduction, with exception of individuals from Tefé (Table 1). Effective population sizes ( $N_e$ ) were low for majority of the localities. However, the confidence intervals were also "infinite" for all but the Madeira River, Parintins, Eirunepé and Tabatinga localities.



**TABLE 2 |** Demographic parameters estimated in IMa2 program for microsatellite data.

Localities	$Ne_1$	$Ne_2$	$Ne_A$	T	Nm1 (receiving)	Nm2 (receiving)
1-Mexiana and 2-Tabatinga	2,209 (637–26,757)	2,476 (311–21,895)	15,412 (8,287–24,029)	2.5 (0.0–2,000)	11.37 (4.37–1,365)	0.87 (0.87–1,001)
1-Jacareacanga and 2- Main channel	5,068 (1,705–11,114)	2,568 (523–5,932)	16,569 (10,796–25,751)	1,212 (121–6,316)	0.36 (0.36–24.48)	8.65 (1.06–18.71)
1- Guajará-Mirim and 2-Main channel	892 (206–945)	3,959 (892–7,620)	18,695 (12,654–29,405)	725.7 (43–4,892)	4.87 (1.64–8.35)	0.23 (0.13–10.83)
1- Boca do Acre and 2-Main channel	8,630 (3,548–39,410)	2,457 (357–6,152)	17,910 (12,157–29,500)	320.3 (18–6,202)	27.11 (7.87–396.2)	6.33 (0.36–20.72)
1-Eirunepé and 2-Main channel	3,317 (1,283–7,045)	4,237 (1,767–7,917)	19,732 (12,904–33,725)	1,855 (436–7,541)	5.27 (0.47–14.56)	1.11 (0.26–13.12)

Effective population size  $Ne = 4 Ne_{\mu}/4 \mu \times 3$  ( $Ne_1$  of the locality1,  $Ne_2$  of the locality 2 and  $Ne_A$  of the ancestor of the two localities); current time in years  $T = (t \times \mu) \times 3$ ; number of migrants per generation  $Nm = Mt/2$ . Values in parentheses = minimum and maximum of confidence limits. Three year generation time is assumed.

## DISCUSSION

### The Role of Rapids in *C. macropomum* Gene Flow

The use of more variable genetic markers, such as microsatellites, has confirmed some of our earlier findings. *Colossoma macropomum* is not panmictic throughout its distribution area. Considering the entire sample, AMOVA, SAMOVA, STRUCTURE,  $\Phi_{ST}$  (DNAmt), and  $F_{ST}/G_{ST}$  (microsatellites)

analyses, suggested genetic differentiation of the Bolivian sub-basin and Brazilian Amazon basin localities.

Within a given river system, freshwater fishes can either form a large panmictic population or be divided into genetically differentiated groups with sufficient gene flow between groups to maintain the integrity of the meta-population. Gene flow measured indirectly by the number of effective migrants per generation ( $Nm$ ) for DNAmt and microsatellite data evidenced restricted gene flow between the two basins, but enough to

**TABLE 3** | Demographic parameters estimated in the program Migrate v3.1.3 for microsatellite data.

	Mex	Alm	San	Ita	Jac	Ori	Nha	Pin	Bor	Hum	Pve	Gua	Mao	Tap	Bda	Coa	Tef	Car	Eir	Fbo	Tab
Mex	—	4.2	3.2	1.2	0.5	2.2	3.9	4	1.5	0.5	0.2	0.8	0.3	4.5	1.1	2.7	1.8	1.6	0.9	4.2	13.6
Alm	15.5	—	2.5	0.6	0.5	2.8	5.2	5.3	1.1	0.6	0.3	0.4	1.8	2.7	1.1	3.0	0.8	3.9	0.3	2.6	3.1
San	11.3	2.9	—	1.1	0.3	2.8	3.0	1.5	1.8	0.5	0.1	0.2	1.0	3.1	1.2	4.1	1.2	2.2	0.7	7.0	2.8
Ita	15.9	1.1	2.8	—	0.3	6	1.2	3.6	2.3	0.5	0.7	0.4	0.9	1.6	3.4	4.1	5.7	5.7	0.5	5.9	3.0
Jac	16.4	3.3	4.3	0.8	—	4.3	9.9	2.8	0.3	0.5	0.6	0.2	1.3	10.5	0.4	2.5	5.9	4.0	0.3	2.3	1.9
Ori	10.3	3.9	4	0.7	0.1	—	5.2	2.0	3.6	0.6	0.7	0.4	1.3	3.8	1.1	2.1	1.2	3.0	0.2	4.7	3.3
Nha	23.3	4.9	3.4	0.5	0.4	4.3	—	4.2	3.0	0.3	0.3	0.5	0.8	4.1	0.7	1.7	2.4	2.2	0.1	3.9	1.6
Pin	11.7	2.9	5	0.7	0.1	3.7	4.4	—	4.2	0.7	0.8	0.8	1.0	5.0	1.3	1.8	1.8	5.8	0.3	2.6	1.8
Bor	14.1	1.8	1.9	0.8	0.6	1.2	5.3	3.3	—	1	0.6	0.4	1.6	2.2	1.3	1.8	1.7	9.0	0.2	3.4	1.4
Hum	15.5	8.3	2.4	4.2	0.4	3.5	7.4	1.8	4.0	—	0.1	0.3	2.0	2.2	0.9	4.5	1.8	1.6	0.3	5.9	3.0
Pve	17.3	5.5	0.5	2.1	0.3	2.2	2.0	2.3	2.1	0.6	—	0.4	1.0	5.0	0.5	3.5	3.3	3.7	0.3	6.0	1.9
Gua	16.9	4.5	4	1.6	0.4	3.6	16	2.0	1.7	0.8	0.5	—	1.3	2.7	0.7	4.0	1.5	7.0	0.7	5.7	1.6
Mao	8.5	7.5	2.9	3.5	0.4	3.6	1.2	4.2	6.0	0.5	0.3	0.5	—	4.5	0.5	0.5	3.0	5.5	0.4	3.9	2.2
Tap	6.2	2.9	0.7	0.3	0.1	2.2	3.0	2.8	0.6	0.4	0.6	0.2	1.5	—	2.1	2.0	0.9	5.5	0.5	3.4	2.7
Bda	13.2	2.9	2.6	0.3	0.2	1.8	2.7	2.0	1.6	0.2	0.5	0.2	2.3	3.8	—	1.7	1.2	1.3	0.3	1.8	1.5
Coa	9.5	3.4	5.1	0.7	0.3	5.9	0.7	4.5	1.0	0.4	0.6	0.6	2.0	7.0	0.6	—	2.6	2.2	1.0	5.2	2.2
Tef	14.6	0.8	2.1	0.3	0.2	1.8	8.8	1.1	1.2	0.6	0.3	0.5	0.9	2.2	0.1	1.7	—	4.7	0.3	7.6	2.8
Car	3.0	3.5	5.5	3.0	0.2	3.6	3.9	0.1	1.8	0.3	0.4	0.5	2.6	5.0	0.8	4.8	2.1	—	0.3	4.2	2.0
Eir	12.2	2.1	1.1	0.8	0.3	2.8	7.4	1.5	0.3	1.7	0.6	0.8	0.3	8.9	0.6	1.5	1.8	3.7	—	5.9	0.8
Fbo	27.5	5	4.5	2.0	0.4	3.5	2.7	1.0	1.4	0.2	0.6	0.2	1.6	7.1	1.1	2.9	0.9	4.0	0.3	—	2.2
Tab	5.5	9.4	4.6	0.7	0.1	8.1	5.8	4.5	3.4	0.5	0.6	0.5	3.0	3.4	1.3	2.87	3.2	0.7	0.3	3.6	—

Number of migrants per generation  $Nm = M\theta/2$ . Row localities are sending individuals, while column localities are receiving individuals. Locality codes are: Mexiana (Mex), Almeirim (Alm), Santarém (San), Itaituba (Ita), Jacareacanga (Jac), Oriximiná (Ori), Nhamundá (Nha), Parintins (Pin), Borba (Bor), Humaitá (Hum), Porto Velho (Pve), Guajará-Mirim (Gua), Manaus (Mao), Tapauá (Tap), Boca do Acre (Bda), Coari (Coa), Tefé (Tef), Carauari (Car), Eirunepé (Eir), Fonte Boa (Fbo), Tabatinga (Tab).

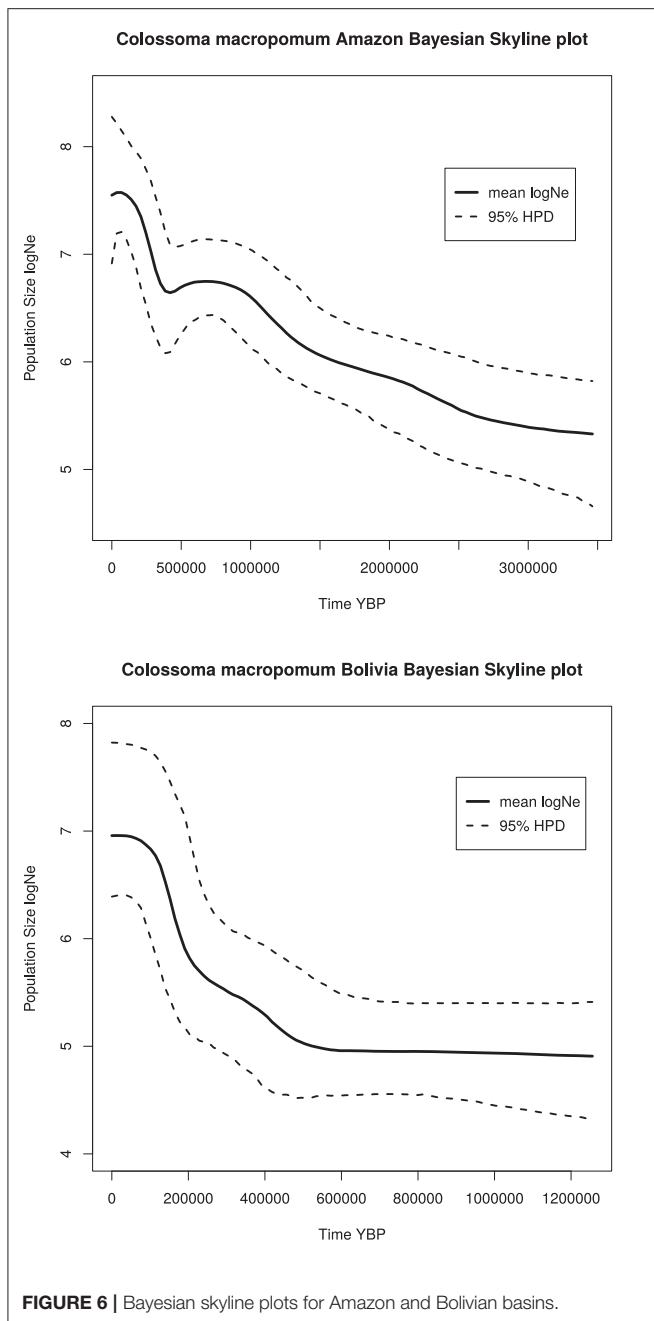
maintain the exchange of genes, thereby minimizing effects of genetic drift. The microsatellite data demonstrate that migration between the two basins is bidirectional. The results of both the IMA and MIGRATE analyses show that gene flow is greater from Bolivia to the Brazilian Amazon, which is in agreement with the results described by Farias et al. (2010). The Brazilian part of the Amazon basin receives more migrants, probably through the passive downstream transport of larvae and juveniles.

The genetic differentiation evidenced between the two basins (Brazilian and Bolivian) is associated with the upper Madeira River rapids, which serve as a natural barrier that restricts, but does not prevent, geneflow between the populations of *C. macropomum* of the two basins. The origin of the Bolivian sub-basin is related to the elevation of the Fitzcarrald arch at the beginning of the middle Pliocene (4 to 3 Ma), which gradually isolated the Bolivian basin, resulting in considerable changes in the drainage pattern, in which the main rivers north of Bolivia drain into the Amazon River through the Madeira River (Hoorn et al., 1995; Campbell et al., 2001), which is the largest tributary of the southern margin of the Solimões-Amazonas basin (Lundberg et al., 1998). The Bolivian sub-basin includes the main Beni and Mamoré rivers, as well as approximately 60% of the entire drainage area of the Madeira River (upper part of Madeira) and is separated from the Amazon basin by a set of 18 rapids and cataracts located between Guajará-Mirim and Porto Velho (Cella-Ribeiro et al., 2013). The largest of these cataracts, the Teotônio cataract, constituted the greatest barrier to navigation

on this river, as well as to the movement of many species of fish (Goulding et al., 2003). The rapids of the upper Madeira River play an important role in the structuring of populations of other Amazonian aquatic species, such as the river turtle *Podocnemis expansa* (Pearse et al., 2006), river dolphins (Gravena et al., 2014, 2015), the black Amazonian flannelmouth characin *Prochilodus nigricans* (Machado et al., 2017), and the catfish *Brachyplatystoma rousseauxii* (Batista, 2010; Carvajal-Vallejos et al., 2014). The Teotônio cataract, as well as the Jirau cataract, the second largest of the Madeira River, have been submerged by hydroelectric reservoirs.

## Population Genetic Structure in the Amazonas River

The lack of genetic differentiation of *C. macropomum* in the main channel of Amazonas River in the Brazilian Amazon basin was supported by all population structure analyses based on mitochondrial genes and microsatellites loci. These findings confirm the pattern reported by Santos et al. (2007) and Farias et al. (2010) who used mtDNA only (supported by  $\Phi_{ST}$ ), and worked at much smaller geographic scales. However,  $\Phi_{ST}$  comparisons involving the tributaries were significant even after Bonferroni corrections. Fisher's exact test and Hedrick's  $G_{ST}$  analysis with microsatellites markers show a weak population differentiation between localities, and stronger differentiation involving comparisons with tributaries, and also black water sites. Geneflow between the main stream localities



and tributary headwaters and black water sites generally is smaller than one effective migrant per generation which also confirms the migration patterns observed in IMA and MIGRATE analyses. Contrary to this pattern, STRUCTURE analysis shows population differentiation only for upper Madeira (localities below the rapids). STRUCTURE program uses a Bayesian approach to investigate the number of biological groups in the dataset. The discordance, between these analyses, could be due to the fact that  $G_{ST}$  analyses are based on variance in allelic frequencies between/among groups and Fisher's exact test uses contingency tables to test null hypothesis that the

alleles are drawn from the same distribution in all populations. These two analyses are more sensitive to detecting smaller and finer levels of genetic differentiation, while STRUCTURE tests for shared system of mating among individuals of the same group. The algorithm of STRUCTURE will not necessarily detect weak structure (Evanno et al., 2005). The population structure observed in *C. macropomum* falls within the category of weak to moderate level of population differentiation (according to Wright, 1965), which could be limiting the sensitivity of the analysis to find more refined substructuring.

Putman and Carbone (2014) emphasize that analyzes used to infer population differentiation have limitations in detecting or not the population structure. Therefore, being conservative for purposes of management and conservation of this species, we consider that the populations of the tributaries (Juruá, Purus, Madeira, Tapajós rivers, and blackwater localities) are different management units until proven otherwise. In this case, the management of fisheries, and seasonal fishing closures during reproductive period, must be effectively respected to preserve the evolutionary potential for the species sustainability. Encouraging aquaculture of the species could also minimize the impact of harvesting natural stocks, which is in fact already occurring.

On the other hand, in the great corridor of the main channel of the Amazon River, from Mexiana (1) to Tabatinga (20) there seem to be not even a signal of isolation-by-distance between the two localities separated by approximately 2,500 Km. The observed lack of genetic structuring of *C. macropomum* in the main channel of the Amazon River basin is probably the result of living in a floodplain environment. The life cycle of this species is tied directly to the seasonal flood cycle of the Amazon; during the flood *C. macropomum* disperses to reproduce and feed in the floodplains and in the flooded forest, while in the dry period fishes become concentrated in lakes and rivers (Araújo-Lima and Goulding, 1998). During the reproduction season, the eggs and larvae are passively transported by the millions to the floodplains, as this species is highly fecund (Araújo-Lima and Goulding, 1998; Araújo-Lima and Ruffino, 2004). This dynamic together with the interlinking of river channels during from flood pulses thus is the primary factor in homogenizing differences among populations of *C. macropomum* (Junk, 1997) and this is observed in molecular data as well. The lack of population structuring is not a unique feature of *C. macropomum*; numerous other species occupying the Amazonian floodplain [e.g., *Prochilodus nigricans* (Machado et al., 2017), *Brycon amazonicus* (Oliveira et al., 2018), and *Paratrygon aiereba* (Frederico et al., 2012)] show this pattern as well. Thus, the pattern found in the present study likely stems from events acting at a macro time scale that affected the region, which, together with current water cycles and the migratory movements of *C. macropomum*, may maintain intra-population homogeneity over generations.

### Genetic Diversity: Large or Small?

Considering that  $H_O$  suffers from sampling effects, we used  $H_E$  as a minimally biased estimate of diversity (Frankham et al., 2002). The  $H_E$  for *C. macropomum* for the microsatellite data were moderate and uniform across the sampling localities, ranging

**TABLE 4** | Demographic parameters estimated for *Colossoma macropomum*, inferred from mitochondrial DNA data.

Basin	Localities	N	Mismatch distribution			Neutrality test	
			Hri (DnaSP)	Hri (Arlequin)	SSD (Arlequin)	Tajima's D	Fu's Fs
Brazilian Amazon basin	Mexiana	15	0.0921 ( $P = 0.0000$ )	0.0183 ( $P = 0.6970$ )	0.0099 ( $P = 0.6630$ )	-0.3633	-5.6331
	Almeirim	21	0.0908 ( $P = 0.0000$ )	0.0190 ( $P = 0.3920$ )	0.0124 ( $P = 0.4090$ )	-0.4285	-7.4138
	Santarém	26	0.0906 ( $P = 0.0000$ )	0.0063 ( $P = 0.8370$ )	0.0041 ( $P = 0.5670$ )	0.6759	-12.6256*
	Itaituba	32	0.0892 ( $P = 0.0000$ )	0.0119 ( $P = 0.2240$ )	0.0043 ( $P = 0.5430$ )	-0.3504	-3.6772
	Jacareacanga	29	0.0882 ( $P = 0.0000$ )	0.0163 ( $P = 0.4080$ )	0.0108 ( $P = 0.5590$ )	-0.0637	-1.2647
	Oriximiná	34	0.0896 ( $P = 0.0000$ )	0.0021 ( $P = 0.9990$ )•	0.0031 ( $P = 0.7270$ )	-0.6243	-19.8803*
	Nhamundá	22	0.0885 ( $P = 0.0000$ )	0.0049 ( $P = 0.9920$ )•	0.0043 ( $P = 0.8760$ )	-0.6959	-7.4689
	Parintins	20	0.0921 ( $P = 0.0000$ )	0.0060 ( $P = 0.9780$ )•	0.0070 ( $P = 0.6550$ )	0.0280	-7.7704
	Borba	19	0.0890 ( $P = 0.0000$ )	0.0105 ( $P = 0.8350$ )	0.0074 ( $P = 0.7770$ )	-0.5002	-6.6767
	Humaitá	22	0.0897 ( $P = 0.0000$ )	0.0061 ( $P = 0.9120$ )	0.0044 ( $P = 0.8820$ )	0.1044	-5.4790
	Porto Velho	21	0.0898 ( $P = 0.0000$ )	0.0251 ( $P = 0.3850$ )	0.0233 ( $P = 0.3610$ )	-0.0231	0.3008
	Manaus	30	0.0937 ( $P = 0.0000$ )	0.0045 ( $P = 0.9690$ )•	0.0071 ( $P = 0.5700$ )	-0.0405	-12.9266*
	Tapauá	28	0.0881 ( $P = 0.0000$ )	0.0067 ( $P = 0.7370$ )	0.0045 ( $P = 0.6940$ )	-0.4853	-13.0369*
	Boca do Acre	21	0.0850 ( $P = 0.0000$ )	0.0060 ( $P = 0.9340$ )	0.0075 ( $P = 0.4240$ )	0.0123	-7.9591
	Coari	20	0.0896 ( $P = 0.0000$ )	0.0120 ( $P = 0.6820$ )	0.0665 ( $P = 0.6180$ )	-0.3213	-7.9575
	Tefé	18	0.0952 ( $P = 0.0000$ )	0.0298 ( $P = 0.2560$ )	0.0177 ( $P = 0.2690$ )	0.0482	-2.7381
	Carauari	25	0.0902 ( $P = 0.0000$ )	0.0063 ( $P = 0.8510$ )	0.0034 ( $P = 0.8900$ )	-0.2137	-10.6621*
	Eirunepé	16	0.0921 ( $P = 0.0000$ )	0.0232 ( $P = 0.6630$ )	0.0150 ( $P = 0.6460$ )	-0.3480	-3.6541
Fonte Boa	20	0.0943 ( $P = 0.0000$ )	0.0244 ( $P = 0.3590$ )	0.0264 ( $P = 0.0420$ )	0.7627	-4.2779	
Tabatinga	31	0.0912 ( $P = 0.0000$ )	0.0031 ( $P = 0.9690$ )•	0.0025 ( $P = 0.8240$ )	-0.2506	-12.8994*	
Bolivian basin	Guajará-Mirim	69	0.0898 ( $P = 0.0000$ )	0.0018 ( $P = 0.9990$ )•	0.0047 ( $P = 0.8420$ )	0.9462	-24.0981*
	All localities	539	0.0912 ( $P = 0.0000$ )	0.0046 ( $P = 1.0000$ )•	0.0020 ( $P = 0.6260$ )	-1.2345	-23.3960

N, number of individuals; Hri, Harpending raggedness index; SSD, sum of square deviations. Significant values \* $P < 0.05$ , • $P > 0.95$ .

from 0.71 to 0.79 with an average of 0.78. For all microsatellite data, variability measures, such as expected heterozygosity and allelic diversity, were similar to those observed in other exploited migratory Amazonian fishes (Carvajal-Vallejos et al., 2014; Ochoa et al., 2015; Oliveira et al., 2018). A compilation of average  $H_E$  values for microsatellite loci obtained from the literature for exploited Amazonian migratory fish shows average  $H_E$  values ranging from 0.50 of *Brachyplatystoma platynemum* (Ochoa et al., 2015), to 0.87 of *Semaprochilodus insignis* (Passos et al., 2010). Between this minimum and maximum, one can observe  $H_E = 0.61$  of *Brachyplatystoma rousseauxii* (Batista, 2010),  $H_E = 0.75$  of *Brachyplatystoma vaillantii* (Rodrigues et al., 2009), and  $H_E = 0.83$  of *Brycon amazonicus* (Oliveira et al., 2018). An average of these values (mean  $H_E = 0.72$ ) is not in agreement with DeWoody and Avise (2000), who report an average  $H_E = 0.54$  for freshwater fishes. The  $H_E = 0.78$  of *C. macropomum* and most of the  $H_E$  reported for exploited Amazonian migratory fishes are more similar to the  $H_E$  of marine fishes reported by DeWoody and Avise (2000) with a mean  $H_E = 0.77$ . In the review of DeWoody and Avise (2000), the authors summarized microsatellite data from North American and European freshwater species, that in comparison to the Amazon basin are geographically very restricted. At 5.5 million km<sup>2</sup>, the Amazon basin is by far the largest hydrographic basin on the planet, and the size of area

occupied by many fish species is comparable to that for marine fishes. The Amazon is in a sense a “sea” that provides an expansive and effectively continuous environment for migratory freshwater fishes such as *C. macropomum* and the other cited species. Migratory fishes, whether freshwater or marine, are usually  $r$  strategists, that is, they have high dispersal capacity, produce a lot of offspring, and in general have a large effective population size, which is turn is reflected in high heterozygosity levels.

In this respect, when compared to freshwater fishes of Europe and North America, the heterozygosities of Amazonian fishes may appear to be high, but this is an illusion. The expected heterozygosities are *on par* with those expected for fishes occupying large areas and having large census numbers. Within the Amazonian species analyzed, the large predatory catfishes of the genus *Brachyplatystoma* have lower  $H_E$  as would be expected by their smaller census sizes. At the opposite end of the spectrum are the relatively small, detritivorous and frugivorous migratory characids (*Semaprochilodus*, *Prochilodus*, and *Brycon*) which have higher  $H_E$  as would be expected by their much larger census sizes. *Colossoma macropomum* has an intermediate  $H_E$ , again a reflection of its frugivorous lifestyle combined with large body size, and thus smaller census size than the other migratory characids but larger census size than the predatory catfishes.

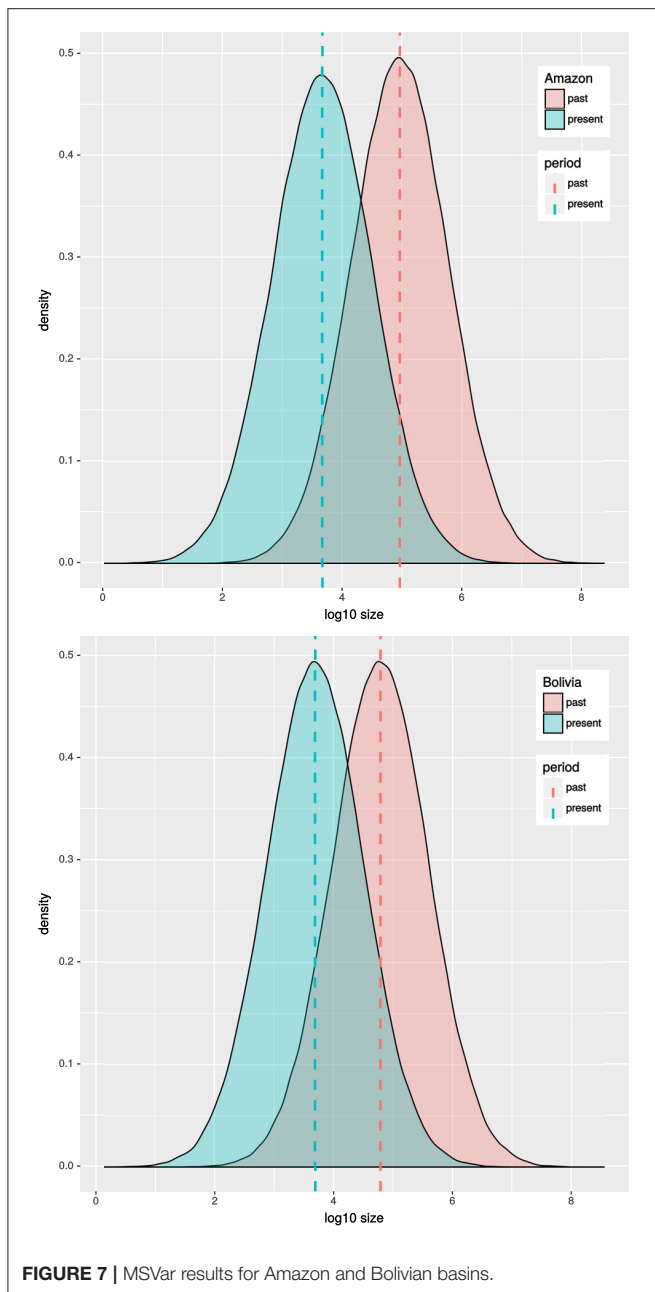


FIGURE 7 | MSVar results for Amazon and Bolivian basins.

## Population Genetic Demography

An analysis of historical demography of *C. macropomum* suggested population expansion in the Amazon basin, which is the same scenario suggested by Farias et al. (2010). However, the result of the current study suggest a population size reduction in the Holocene (Figure 5). This result is confirmed by the very recent population decline observed in the Skyline plot analyses (Figure 4). Similar pattern are observed for the Bolivian basin, however, very recent population decline is not evidenced in the Skyline plot.

Most studies conducted in the Amazon involving fish species such as *Prochilodus nigricans* (Machado et al., 2017), *Brachyplatystoma rousseauxii* (Batista and Alves-Gomes, 2006;

Carvajal-Vallejos et al., 2014), and *Brachyplatystoma platynemum* (Ochoa et al., 2015) indicate recent population expansion, at least as indicated by Fu's  $F_s$  test. The difference in effective population size of *C. macropomum* prior estimated for the Bolivian and Amazon basin broadly corresponded to the relative proportion of potential habitat in the Bolivian basin. This basin accounts for approximately 20% of the total Amazon basin, suggesting that the Bolivian basin had a *C. macropomum* population approximately 20% smaller than the rest of the Amazon basin during the Holocene.

Glacial and inter-glacial periods of the Pleistocene exerted considerable impact on the climate, which consequently affected the vegetation in South America (Ledru et al., 1996) and also had impact on aquatic and terrestrial fauna. It is during the Late Pleistocene that *C. macropomum* in the Amazon basin began expanding (Figure 4) associated with the expansion of the várzea-like habitat (Irion and Kalliola, 2010). Therefore, the population expansion of *C. macropomum* in the Amazon basin likely occurred due to the increase in the availability of habitat for this species starting in the later half of the Pleistocene; however, population growth is no longer observed in the Holocene.

Corroborating the observation for the Holocene, analyses of microsatellite data in MSVar indicate that *C. macropomum* has undergone a pronounced population decline in both drainages during the Holocene (10,000 years ago—Amazon and 2,500 years ago Bolivia), probably due to climate change related to Last Glacial Maximum and during the mid-Holocene epoch (Wang et al., 2017). Demographic decrease was strongly supported ( $BF = 1206$ ) and occurred with 0.9992 probability.

The demographic decrease was from approximately 100,000 effective individuals to approximately 5,000 individuals, an approximate 1.5 orders of magnitude decrease. Similar values for current effective population size were inferred using IMA2 (Table 2). By any measure, the effective population size is small, and much smaller than in the last several thousand years.

DeWoody and Avise (2000) estimate that at equilibrium  $H_E = 0.79$  represents 25,000 effective individuals assuming a substitution rate of  $10^{-4}$ . Our estimate of substitution rate from MSVar was  $10^{-3.81}$ , or just about 17,000 effective individuals are expected at equilibrium. In this sense, the  $N_e$  values of *C. macropomum* are below an equilibrium expectation, which also suggests a current reduction of  $N_e$ . In fact, the results of the Bottleneck program indicate decrease in population size in the majority of sampling localities, which is also corroborated by the  $N_e$  values (Table 1). The only major event that may have contributed to population decrease of *C. macropomum* in the nowadays, within a time window of decades, is the over-exploitation of the species and the destruction of its floodplain habitat. Natural stocks of the *C. macropomum* suffer from overfishing and juveniles currently account for the largest part of the catch (Barthem and Goulding, 2007). Although aquaculture of *C. macropomum* has grown in recent years, there is strong evidence that the natural population of this species are still depressed because of over-fishing. This can be evidenced in the continuous reduction of the tonnage landed in the port of Manaus and other major Amazonian ports. Araújo-Lima (2002) report that in 1976 *C. macropomum* reached 16,000 tons/year landed in the port of Manaus, while data from late

1990's indicated less than 4 thousand tons. During this time, the population of Manaus more than doubled. Furthermore, another worrying factor is the mean size of the fish landed, with juveniles representing the majority of the catch (Barthem and Goulding, 2007). The average size of the fish landed in the main markets in the Amazon suggests that most individuals are fished before reaching sexual maturation, which in the case of females occurs between 50 and 55 cm (Isaac and Ruffino, 2000).

In conclusion, naturally exogamous species with large census sizes have considerable genetic diversity and large effective population sizes (Frankham et al., 2002). In this context, *C. macropomum* has levels of genetic diversity that are *on par* with expectations for species of similar lifestyle and body size. However, with the historical decline of *C. macropomum* populations, it is evident that part of the genetic diversity that existed in the past has been lost. Still the remaining diversity is representative of this species's historical genetic diversity and it is this genetic diversity that can secure the recovery and long-term persistence of natural populations of *C. macropomum* in the Amazon basin.

## ETHICS STATEMENT

All field collections were authorized by IBAMA/SISBIO 11325-1, and access to genetic resources was authorized by permit No. 034/2005/IBAMA. Field collection permits are conditional that collection of organisms be undertaken in accordance with the ethical recommendations of the Conselho Federal de Biologia (CFBio; Federal Council of Biologists), Resolution 301 (December 8, 2012).

## REFERENCES

- Araújo-Lima, C. A. R. M. (2002). "Piscicultura extensiva de tambaqui na floresta de várzea," in *Livro de Resultados dos Projetos de Pesquisa Dirigida (PPDS)-PPG7* (Brasília: MCT), 131–135.
- Araújo-Lima, C. A. R. M., and Goulding, M. (1998). *Os Frutos do Tambaqui. Ecologia, Conservação e Cultivo na Amazônia*. Belem: Sociedade Civil Mamirauá-MCT-CNPq.
- Araújo-Lima, C. A. R. M., and Ruffino, M. L. (2004). "Migratory fishes of the Brazilian Amazon," in *Migratory Fishes of South America. Biology, Fisheries, and Conservation Status*, eds J. Carolsfield, B. Harvey, C. Ross, and A. Baer (Victoria, BC: Co-published by World Fisheries Trust/World Bank/International Development Research Center), 233–302.
- Barthem, R. B., and Fabr e, N. N. (2003). "Biologia e diversidade dos recursos pesqueiros da Amaz nia," in *A Pesca e os Recursos Pesqueiros na Amaz nia Brasileira*, ed M. L. Ruffino (Manaus, AM: ProVarzea), 11–55.
- Barthem, R. B., and Goulding, M. (2007). *An Unexpected Ecosystem: The Amazon as Revealed by Fisheries*. St. Louis, MO: Amazon Conservation Association and Missouri Botanical Garden Press.
- Batista, J. S. (2010). *Caracteriza o Gen tica da Dourada-Brachyplatystoma Rousseauxii, Castelnau, 1855 (Siluriformes - Pimelodidae) na Amaz nia por Meio de Marcadores Moleculares Mitocondriais e Microsat lites: Subs dios Para Conserva o e Manejo*. Tese de Mestrado do Programa de P s-gradua o de Gen tica, Conserva o e Biologia Evolutiva, Instituto Nacional de Pesquisas da Amaz nia, Manaus, Brasil.
- Batista, J. S., and Alves-Gomes, J. A. (2006). Phylogeography of *Brachyplatystoma rousseauxii* (Siluriformes - Pimelodidae) in the Amazon Basin offers preliminary evidence for the first case of "homing" for an Amazonian migratory catfish. *Gen. Mol. Res.* 5, 723–740.

## AUTHOR CONTRIBUTIONS

IF and TH conceived the experiment and obtained funding. MS, IF, and TH conducted fieldwork, collected specimens, analyzed the results, and wrote the manuscript. MS collected molecular data. All authors contributed to and reviewed the manuscript.

## FUNDING

This research was supported by the MCT/CNPq/PPG7 557090/2005-9, CNPq/CT-Amaz nia 554057/2006-9, CNPq/CT-Amaz nia 575603/2008-9, and FINEP/DARPA (Conv nio No. 01.09.0472.00) to IF. Brazilian permits for field collection and molecular analyses were given by IBAMA/CGEN 11325-1, and IBAMA/MMA-N  086/2006 de 08/09/2006. TH and IF were supported by a Bolsa de Pesquisa scholarship from CNPq during the study and MS by a FAPEAM fellowship.

## ACKNOWLEDGMENTS

We thank M rio Nunes, Pedro Bittencourt and Rommel Rojas for technical support. This study is part of MS's Ph.D. thesis in the Biotechnology graduate program of UFAM.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2018.00263/full#supplementary-material>

- Beaumont, M. A. (1999). Detecting population expansion and decline using microsatellites. *Genetics* 153, 2013–2029.
- Berli, P., and 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. U.S.A.* 98, 4563–4568. doi: 10.1073/pnas.081068098
- Campbell, K. E. Jr., Heizler, M., Frailey, C. D., Romero Pittman, L., and Prothero, D. R. (2001). Upper Cenozoic chronostratigraphy of the southwestern Amazon Basin. *Geology* 29, 595–598. doi: 10.1130/0091-7613(2001)029<0595:UCCOTS>2.0.CO;2
- Carvajal-Vallejos, F. M., Duponchelle, F., Desmarais, E., Cerqueira, F., Querouil, S., Nu ez, J., et al. (2014). Genetic structure in the Amazonian catfish *Brachyplatystoma rousseauxii*: influence of life history strategies. *Genetica* 142, 323–336. doi: 10.1007/s10709-014-9777-2
- Cella-Ribeiro, A., Torrente-Vilara, G., Hungria, D. B., and Oliveira, M., de (2013). "As corredeiras do Rio Madeira," in *Peixes do Rio Madeira*, eds L. J. de Queiroz, G. Torrente-Vilara, W. M. Ohara, T. H. da S. Pires, J. Zuanon, and C. R. C. Doria (Santo Antonio Energia), 47–53.
- Corander, J., and Tang, J. (2007). Bayesian analysis of population structure based on linked molecular information. *Math. Biosci.* 205, 19–31. doi: 10.1016/j.mbs.2006.09.015
- DeWoody, J. A., and Avise, J. C. (2000). Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *J. Fish Biol.* 56, 461–473. doi: 10.1006/jfbi.1999.1210
- DeWoody, J. A., Schupp, J., Kenefic, L., Busch, J., Murfitt, L., and Keim, P. (2004). Universal method for producing ROX-labeled size standards suitable for automated genotyping. *BioTechniques* 37, 348–352. doi: 10.2144/04373BM02
- Di Rienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M., and Freimer, N. B. (1994). Mutational processes of simple-sequence repeat

- loci in human populations. *Proc. Natl. Acad. Sci. U.S.A.* 91, 3166–3170. doi: 10.1073/pnas.91.8.3166
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., and Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size ( $N_e$ ) from genetic data. *Mol. Ecol. Res.* 14, 209–214. doi: 10.1111/1755-0998.12157
- Drummond, A. J., and Rambaut, A. (2007). BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214. doi: 10.1186/1471-2148-7-214
- Dupanloup, I., Schneider, S., and Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* 11, 2571–2581. doi: 10.1046/j.1365-294X.2002.01650.x
- Estoup, A., Tailliez, C., Cornuet, J.-M., and Solignac, M. (1995). Size homoplasy and mutational processes of interrupted microsatellite in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Mol. Biol. Evol.* 12, 1074–1084.
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x
- Excoffier, L., Smouse, P. E., and 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.
- Farias, I. P., Torrico, J. P., García-Dávila, C., da Santos, M. C. F., Hrbek, T., and Renno, J.-F. (2010). Are rapids a barrier for floodplain fishes of the Amazon basin? A demographic study of the keystone floodplain species *Colossoma macropomum* (Teleostei: Characiformes). *Mol. Phylogenet. Evol.* 56, 1129–1135. doi: 10.1016/j.ympev.2010.03.028
- Frankham, R., Ballou, J. D., and Briscoe, D. A. (2002). *Introduction to Conservation Genetics*. Cambridge: Cambridge University Press.
- Frederico, R. G., Farias, I. P., Araújo, M. L. G., Charvet-Almeida, P., and Alves-Gomes, J. A. (2012). Phylogeography and conservation genetics of the Amazonian freshwater stingray *Paratrygon aiereba* Müller & Henle, 1841 (Chondrichthyes: Potamotrygonidae). *Neotrop. Ichthyol.* 10, 71–80. doi: 10.1590/S1679-62252012000100007
- Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.
- Garza, J. C., and Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. *Mol. Ecol.* 10, 305–318. doi: 10.1046/j.1365-294X.2001.01190.x
- Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–472. doi: 10.1214/ss/1177011136
- Goulding, M., Barthem, R. B., and Ferreira, E. J. G. (2003). *The Smithsonian Atlas of the Amazon*. Washington, DC: Smithsonian Institution Press.
- Goulding, M., and 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. doi: 10.1590/S0101-81751982000200001
- Gravena, W., da Silva, V. M. F., da Silva, M. N. F., Farias, I. P., and Hrbek, T. (2015). Living between rapids: genetic structure and hybridization in botos (Cetacea: Iniidae: *Inia* spp.) of the Madeira River, Brazil. *Biol. J. Linn. Soc.* 114, 764–777. doi: 10.1111/bij.12463
- Gravena, W., Farias, I. P., da Silva, M. N. F., da Silva, V. M. F., and Hrbek, T. (2014). Looking to the past and the future: were the Madeira River rapids a geographic barrier to the boto (Cetacea: Iniidae)? *Conserv. Genet.* 15, 619–629. doi: 10.1007/s10592-014-0565-4
- 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. Available online at: <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>
- Harpending, H. C. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum. Biol.* 66, 591–600.
- Hasegawa, M., Kishino, H., and Yano, T. A. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160–174. doi: 10.1007/BF02101694
- Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution* 59, 1633–1638. doi: 10.1111/j.0014-3820.2005.tb01814.x
- Hey, J., and Nielsen, R. (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2785–2790. doi: 10.1073/pnas.0611164104
- Hoorn, C., Guerrero, J., Sarmiento, G. A., and Lorente, M. A. (1995). Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. *Geology* 23, 237–240. doi: 10.1130/0091-7613(1995)023<0237:ATAACF>2.3.CO;2
- Hoorn, C., Wesselingh, F. P., ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., et al. (2010). Change, landscape evolution, and biodiversity Amazonia through time: Andean uplift, climate. *Science* 330, 927–931. doi: 10.1126/science.1194585
- Irion, G., and Kalliolu, R. J. (2010). “Long-term landscape development processes in Amazonia,” in *Amazonia: Landscape and Species Evolution: A Look into the Past*, eds C. Hoorn and F. P. Wesselingh (Chichester: Wiley-Blackwell), 185–197.
- 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. doi: 10.1590/1809-43921996263208
- Isaac, V. J., and Ruffino, M. L. (1996). Population dynamics of tambaqui, *Colossoma macropomum* Cuvier, in the lower Amazon, Brazil. *Fish. Manag. Ecol.* 1996, 315–333. doi: 10.1046/j.1365-2400.1996.d01-154.x
- Isaac, V. J., and Ruffino, M. L. (2000). “Biologia pesqueira do tambaqui, *Colossoma macropomum*, no Baixo Amazonas,” in *Recursos Pesqueiros do Médio Amazonas: Biologia e Estatística Pesqueira* (Brasília: Edições IBAMA; Coleção meio ambiente; Série Estudos Pesca), 65–88.
- Jombart, T. (2008). ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129
- Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11:94. doi: 10.1186/1471-2156-11-94
- Junk, W. J. (1997). *The Central Amazon System: Ecology of a Pulsing System*. Berlin: Springer Verlag.
- Kalinowski, S. T. (2005). HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol. Ecol. Notes* 5, 187–189. doi: 10.1111/j.1471-8286.2004.00845.x
- Kitada, S., Kitakado, T., and Kishino, H. (2007). Empirical bayes inference of pairwise  $F_{ST}$  and its distribution in the genome. *Genetics* 177, 861–873. doi: 10.1534/genetics.107.077263
- Kitada, S., Nakamichi, R., and Kishino, H. (2017). The empirical Bayes estimators of fine-scale population structure in high gene flow species. *Mol. Ecol. Resour.* 17, 1210–1222. doi: 10.1111/1755-0998.12663
- Leberg, P. L. (2002). Estimating allelic richness: effects of sample size and bottlenecks. *Mol. Ecol.* 11, 2445–2449. doi: 10.1046/j.1365-294X.2002.01612.x
- Ledru, M.-P., Braga, P. I. S., Soubiès, F., Fournier, M., Martin, L., Suguio, K., et al. (1996). The last 50,000 years in the Neotropics (Southern Brazil): evolution of vegetation and climate. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 123, 239–257. Available online at: <http://www.sciencedirect.com/science/article/B6V6R-3VVT2P4-D/2/b91b6aa40adb8b4b114c150b06404376>. doi: 10.1016/0031-0182(96)00105-8
- Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187
- Lovejoy, N. R., Willis, S. C., and Albert, J. S. (2010). “Molecular signatures of Neogene biogeographical events in the Amazon fish fauna,” in *Amazonia: Landscape and Species Evolution: A Look into the Past*, eds C. Hoorn and F. P. Wesselingh (Oxford: Wiley-Blackwell), 405–417.
- Lundberg, J. G., Marshall, L. G., Guerrero, J., Horton, B., Malabarba, M. C. S. L., and Wesselingh, F. P. (1998). “The stage for Neotropical fish diversification: a history of tropical South American rivers,” in *Phylogeny and Classification of Neotropical Fishes*, eds L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M. S. Lucena, and C. A. S. Lucena (Porto Alegre: EDIPUCRS), 13–48.
- Machado, V. N., Willis, S. C., Hrbek, T., and Farias, I. P. (2017). Population genetic structure of the Amazonian black flannelmouth characin (Characiformes, Prochilodontidae: *Prochilodus nigricans* Spix and Agassiz, 1829): contemporary and historical gene flow of a migratory and abundant fishery species. *Environ. Biol. Fishes* 100, 1–16. doi: 10.1007/s10641-016-0547-0
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27, 209–220.
- Ochoa, L. E., Pereira, L. H. G., Costa-Silva, G. J., Roxo, F. F., da Batista, J. S., Formiga, K., et al. (2015). Genetic structure and historical diversification



- of catfish *Brachyplatystoma platynemum* (Siluriformes: Pimelodidae) in the Amazon basin with implications for its conservation. *Ecol. Evol.* 5, 2005–2020. doi: 10.1002/ece3.1486
- Ohta, T., and Kimura, M. (1973). A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Gen. Res.* 22, 201–204. doi: 10.1017/S0016672300012994
- Oliveira, R. C., Santos, M. C. F., Bernardino, G., Hrbek, T., and Farias, I. P. (2018). From river to farm: an evaluation of genetic diversity in wild and aquaculture stocks of Brycon amazonicus (Spix & Agassiz, 1829), Characidae, Bryconinae. *Hydrobiologia* 805, 75–88. doi: 10.1007/s10750-017-3278-0
- Palstra, F. P., and Fraser, D. J. (2012). Effective/census population size ratio estimation: a compendium and appraisal. *Ecol. Evol.* 2, 2357–2365. doi: 10.1002/ece3.329
- Passos, K. B., Leão, A. S. A., Oliveira, D. P., Farias, I. P., and Hrbek, T. (2010). Polymorphic microsatellite markers for the overexploited Amazonian fish, *Semaprochilodus insignis* (Jardine and Schomburgk 1841). *Conserv. Gen. Res.* 2, 231–234. doi: 10.1007/s12686-010-9245-y
- Pearse, D. E., Arndt, A. D., Valenzuela, N., Miller, B. A., Cantarelli, V. H., and Sites, 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. doi: 10.1111/j.1365-294X.2006.02869.x
- Piry, S., Luikart, G., and Cornuet, J. (1999). BOTTLENECK: a computer program for detecting recent reduction in the effective size using allele frequency data. *J. Hered.* 90, 502–503. doi: 10.1093/jhered/90.4.502
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Putman, A. I., and Carbone, I. (2014). Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecol. Evol.* 4, 4399–4428. doi: 10.1002/ece3.1305
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing Available online at: [www.R-project.org](http://www.R-project.org)
- Reis, R. E., Albert, J. S., Di Dario, F., Mincarone, M. M., Petry, P., and Rocha, L. A. (2016). Fish biodiversity and conservation in South America. *J. Fish Biol.* 89, 12–47. doi: 10.1111/jfb.13016
- Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* 43, 223–225.
- Rodrigues, F. C., Farias, I. P., Batista, J. S., and Alves-Gomes, J. (2009). Isolation and characterization of microsatellites loci for “piramutaba” (*Brachyplatystoma vaillantii*, Siluriformes: Pimelodidae), one of the commercially most important migratory catfishes in the Amazon Basin. *Conserv. Genet. Res.* 1:365. doi: 10.1007/s12686-009-9084-x
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Springs Harbor, NY: Cold Springs Harbor Laboratory Press.
- Santos, M. C. F., Ruffino, M. L., and Farias, I. P. (2007). High levels of genetic variability and panmixia of the tambaqui *Colossoma macropomum* (Cuvier, 1818) in the main channel of the Amazon River. *J. Fish Biol.* 71A, 33–44. doi: 10.1111/j.1095-8649.2007.01514.x
- Santos, M. D., Hrbek, T., and Farias, I. P. (2009). Microsatellite markers for the tambaqui (*Colossoma macropomum*, Serrasalminae, Characiformes), an economically important keystone species of the Amazon River floodplain. *Mol. Ecol. Resour.* 9, 874–876. doi: 10.1111/j.1755-0998.2008.02331.x
- Schneider, S., and Excoffier, L. (1999). Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 152, 1079–1089. Available online at: <http://www.genetics.org/cgi/content/abstract/152/3/1079>
- 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 H. Sioli (New York, NY: Springer Verlag), 127–165.
- Slatkin, M., and Hudson, R. R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129, 555–562.
- Storz, J. F., Beaumont, M. A., and Alberts, S. C. (2002). Genetic evidence for long-term population decline in a savannah-dwelling primate: inferences from a hierarchical bayesian model. *Mol. Biol. Evol.* 19, 1981–1990. doi: 10.1093/oxfordjournals.molbev.a004022
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.
- Tamura, K., Stecher, G., Peterson, D. G., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4, 535–538. doi: 10.1111/j.1471-8286.2004.00684.x
- Venticinque, E., Forsberg, B., Barthen, B. R., Petry, P., Hess, L., Mercado, A., et al. (2016). An explicit GIS-based river basin framework for aquatic ecosystem conservation in the Amazon. *Earth Syst. Sci. Data* 8, 651–661. doi: 10.5194/essd-2016-17
- Wang, X., Lawrence Edwards, R., Auler, A. S., Cheng, H., Kong, K., Wang, Y., et al. (2017). Hydroclimate changes across the Amazon lowlands over the past 45,000 years. *Nature* 541, 204–207. doi: 10.1038/nature20787
- Waples, R. S., and Do, C. (2008). Ldne: a program for estimating effective population size from data on linkage disequilibrium. *Mol. Ecol. Res.* 8, 753–756. doi: 10.1111/j.1755-0998.2007.02061.x
- Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19, 395–420.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Santos, Hrbek and Farias. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.