UNIVERSIDADE FEDERAL DO OESTE DO PARÁ
PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO TECNOLÓGICA
PROGRAMA DE PÓS-GRADUAÇÃO EM RECURSOS NATURAIS DA AMAZÔNIA

KAREN LARISSA AUZIER GUIMARÃES

ESTUDOS DE TAXONOMIA INTEGRATIVA DO COMPLEXO *Hoplias malabaricus* (Bloch, 1794) NA BACIA AMAZÔNICA E DRENAGENS ADJACENTES

SANTARÉM-PA

2020

KAREN LARISSA AUZIER GUIMARÃES

ESTUDOS DE TAXONOMIA INTEGRATIVA DO COMPLEXO *Hoplias malabaricus* (Bloch, 1794) NA BACIA AMAZÔNICA E DRENAGENS ADJACENTES

Dissertação apresentada ao Programa de Pós-Graduação em Recursos Natuais da Amzônia para obtenção do título de Mestre em Ciências Ambientais; Universidade Federal do Oeste do Pará – UFOPA;
Área de concentração: Genética e Conservação da Biodiversidade.
Orientador: Dr. Luís Reginaldo Ribeiro Rodrigues

SANTARÉM-PA

2020

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema Integrado de Bibliotecas – SIBI/UFOPA

G963e Guimarães, Karen Larissa Auzier

Estudos de taxonomia integrativa do complexo *Hoplias malabaricus* (Bloch, 1794) na Bacia Amazônica e drenagens adjacentes. / Karen Larissa Auzier Guimarães. – Santarém, 2020.

62 p. : il.

Inclui bibliografias.

Orientador: Luís Reginaldo Ribeiro Rodrigues

Dissertação (Mestrado) – Universidade Federal do Oeste do Pará, Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação Tecnológica, Programa de Pós-Graduação em Recursos Naturais da Amazônia.

1. América do Sul. 2. Biogeografia. 3. Hoplias. I. Rodrigues, Luís Reginaldo Ribeiro, *orient*. II. Título.

CDD: 23 ed. 597.48098115

Bibliotecária - Documentalista: Renata Ferreira – CRB/2 1440

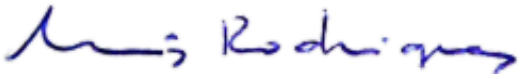
KAREN LARISSA AUZIER GUIMARÃES

ESTUDOS DE TAXONOMIA INTEGRATIVA DO COMPLEXO *Hoplias malabaricus* (Bloch, 1794) NA BACIA AMAZÔNICA E DRENAGENS ADJACENTES

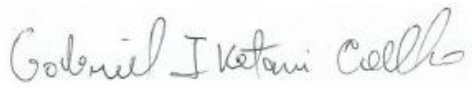
Dissertação apresentada ao Programa de Pós-Graduação em Recursos Natuais da Amzônia para obtenção do título de Mestre em Ciências Ambientais; Universidade Federal do Oeste do Pará – UFOPA;
Área de concentração: Genética e Conservação da Biodiversidade.

Conceito: 9,82

Data de Aprovação: 27/08/2020

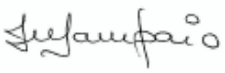


Dr. Luís Reginaldo Ribeiro Rodrigues – Orientador
UNIVERSIDADE FEDERAL DO OESTE PARÁ – UFOPA



Dr. Gabriel Iketani Coelho – Examinador 1
UNIVERSIDADE FEDERAL DO OESTE PARÁ – UFOPA

Dr. Luís Fernando Carvalho Costa – Examinador 2
UNIVERSIDADE FEDERAL DO MARANHÃO – UFMA



Dra. Maria Iracilda da Cunha Sampaio – Examinador 3
UNIVERSIDADE FEDERAL DO PARÁ – UFPA

AGRADECIMENTOS

Uma pesquisa é uma construção, resultado do esforço de inúmeros que contribuem direta e indiretamente... Utilizar de pequenas linhas para agradecer às pessoas que fizeram parte da construção deste trabalho é um desafio tão grande quanto realizar o trabalho.

Agradeço ao meu orientador Dr. Luís Reginaldo pela orientação não só desta dissertação, mas pelo apoio incansável, disposição, por me inspirar e me mostrar a nobre função da ciência e da pesquisa ao longo desses sete anos trabalhando juntos. Muito obrigada!

Agradeço aos membros do Laboratório de Genética e Biodiversidade da UFOPA (LGBio) (2018-2020) por todos os momentos compartilhados e por ajudas na bancada: Alline Rufino, Dyennef Pantoja, Heveline Campos, Luan Maciel, Mara Guimarães, Mendelsohn Fujiie, Thais Sena. Obrigada a todos!

Agradeço a todos os ictiólogos que me ajudaram na obtenção/doação de tecidos e sequências de *Hoplias*: André Netto-Ferreira, Arce Mariangeles, Caroline Arantes, Erick Guimarães, Jonas Oliveira, Mark Sabaj, Pâmella Brito.

Agradeço ao apoio de Dona Maria, Sr. Pena e Sr. Jameson durante a estadia no Creporizão e ajuda para obtenção de amostras.

Agradeço ao Grupo de Biotaxonomía Morfológica y Molecular de Peces da Universidad Nacional de Mar del Plata por toda a contribuição, troca de conhecimentos e recepção durante a minha estadia na Argentina: Dr. Juan José Rosso, Dr. Juan Martin Días de Astarloa e Dr. Mariano Gonzales-Castro.

Agradeço ao Dr. Diego José Santana (UFMS) pela paciencia e por compartilhar comigo seus conhecimentos em Bioinformática que muito me auxiliaram na análise de dados. Ao Dr. Elciomar de Oliveira por me acompanhar na viagem e sempre sanar minhas dúvidas quando necessitei. Agradeço também à Renata Fadel e Priscila por compartilhar vossa casa para minha estadia em Campo Grande.

Agradeço ao Dr. Gabriel Iketani por sempre ajudar nos sequenciamentos. Também ao apoio do Dr. Marcos Prado.

Agradeço à Tauanny Lima por toda a paciência e auxílio nas figuras e ilustrações que muito me ajudaram para a realização da parte morfológica deste trabalho! Agradeço também pelo companheirismo na ciência e na vida!

Agradeço aos alunos antigos do LGBio que iniciaram esta investigação: Diego Marques, Eugerlane Queiroz, Fabiola Araujo, Jennefer Alves, Marcos Prado, Vinícius Oliveira. Também agradeço à Beatriz Viana pela ajuda na edição de algunas imagens.

Agradeço à Rose (secretária do PPGRNA) e Dr. José Mauro pelo apoio prestado durante o curso.

Agradeço à PROPPIT (Dra Lenise Vargas) e ARNI (Dr. Rodrigo Silva) pelo apoio na realização da minha viagem para a Argentina.

Agradeço ao corpo docente do PPGRNA e turma (2018) pelo convívio e troca de conhecimentos durante o curso.

Aos meus velhos e novos amigos (sintam-se todos incluídos e abraçados) que de alguma maneira contribuíram para que essa caminhada tornasse mais leve! Obrigada a todos!

Ao meu primo Lucivaldo e minha tia Ivanete pelo apoio familiar concedido.

Esta dissertação teve apoio da Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES) através da bolsa concedida (processo 88882.457158/2019-01); CNPQ/FAPEAM-INCT ADAPTA II (Processos 465540/2014-7 e 062.1187/2017).

RESUMO

Hoplias malabaricus é um complexo de espécies difundido do norte ao sul da América do Sul. Este grupo pode constituir um bom modelo para investigações de eventos biogeográficos históricos e os padrões reais de estruturação genética que impulsionam a ictiofauna nos sistemas de água doce neotropicais. No entanto, é limitado devido à taxonomia confusa e das várias espécies crípticas escondidas no complexo. Neste trabalho, com base em evidências morfológicas e moleculares, estendemos a distribuição de *H. misionera*, conhecida apenas nas bacias do Uruguai, Paraná e Paraguai. Reconhecemos também uma verdadeira espécie de *H. malabaricus* e realizamos uma análise genética da população nessa linhagem. O DNA *barcoding* revelou três a oito espécies candidatas do complexo *H. malabaricus* que habitam a área de estudo. O maior clado recuperado (BIN ABZ3047) foi assumido como *H. malabaricus sensu stricto*. Esta espécie está estruturada em seis unidades populacionais: 1) Bacia do Rio Madeira (MRB), 2) Drenagens do Escudo das Guianas (GSD), 3) Bacia do Atlântico Nordeste Ocidental (WNAB), 4) Bacia do Rio Tapajós (TRB), 5) Baixo Rio Amazonas e confluências (LARC) e 6) bacia do rio São Francisco (SFRB). As populações TBR e SFRB foram as mais diferenciadas e apresentaram flutuações demográficas, onde as últimas mostraram evidências de declínio populacional.

Palavras-chave: América do Sul, Biogeografia, Diversificação, *Hoplias*, Rio Amazonas

ABSTRACT

Hoplias malabaricus is a species complex widespread from Northern to Southern South America continent. This group might constitute a good model for investigations of historical biogeographic events and the actual patterns of genetic structuring driving the ichthyofauna in the Neotropical freshwater systems. However, it is limited because of the confused taxonomy and the several cryptic species hidden in the complex. In this work, based on morphological and molecular evidence, we extend the distribution of *H. misionera*, which was only known from Uruguay, Paraná and Paraguay River basins. We also recognized the true *H. malabaricus* species and performed a population genetics analysis in this lineage. DNA barcoding revealed three to eight candidate species from the *H. malabaricus* complex inhabiting the study area. The largest clade recovered (BIN ABZ3047) was assumed as the true *H. malabaricus sensu stricto*. This species is structured in six population units: 1) Madeira River Basin (MRB), 2) Guiana Shields drainages (GSD), 3) Western Northeast Atlantic Basin (WNAB), 4) Tapajós River Basin (TRB), 5) Lower Amazonas River confluences (LARC) and 6) São Francisco River Basin (SFRB). The populations TBR and SFRB were most differentiated and showed demographic fluctuations, where the later showed evidence of declining.

Key words: South America, Biogeography, Diversification, *Hoplias*, Amazon River.

Sumário

ARTIGO 1* **6**
Introduction **9**
Materials and Methods **10**
Ethics statement **10**
Sampling and Study Area **10**
Morphological analysis **11**
Molecular analysis **12**
Cytogenetics analysis **12**
Results **13**
Morphological analysis **13**
Molecular analysis **17**
Cytogenetic analysis **19**
Geographic distribution **21**
Discussion **23**
Acknowledgements **27**
References **27**
ARTIGO 2* **36**
Introduction **37**
Materials and methods **39**
Ethics statements **39**
Sampling and Study area **39**
DNA extraction, PCR and Sequencing **40**
Molecular data analysis and species delimitation **41**
Population genetics and biogeographic analysis **42**
Results **43**
DNA barcoding and species delimitation **43**
Population genetics of *H. malabaricus sensu stricto* **46**
Demographic history of *H. malabaricus stricto sensu* **48**
Discussion **50**
Data availability **53**
References **53**
Acknowledgements **58**
Supplementary information **59**
Síntese Integradora **60**

ARTIGO 1*

First record of *Hoplias misionera* (Characiformes: Erythrinidae) in the Amazon River basin, Brazil: morphological, DNA barcoding and cytogenetic considerations

Guimarães et al. 2020

*O artigo apresentado foi redigido conforme as diretrizes da revista *Neotropical Ichthyology*. No entanto, visando facilitar a compreensão do manuscrito, imagens, tabelas e suas respectivas legendas estão inseridas no corpo do texto. Para a versão final a ser submetida à revista, imagens e tabelas serão enviadas separadamente.
As normas indicadas para redação de artigos pela revista estão disponíveis no link: https://www.scielo.br/revistas/ni/iinstruc.htm

Abstract

The *Hoplias malabaricus* group encompasses six valid species and is believed to hide cryptic species still not described. In this work, we analyze a population from the Amazon basin previously identified as *H. malabaricus*, but that was phylogenetically closer to *H. misionera*, a species described to La Plata basin. The DNA barcoding analysis revealed that the Amazon population nested together with *H. misionera* specimens from the La Plata basin (BIN AAB1732) in a single monophyletic clade. The intragroup distance (0.6%) was 10 times lower than the nearest neighbor (6.6%) distance. The morphometric analysis demonstrated slightly variation between Amazon and La Plata populations, being the former composed by larger specimens. Further morphological data supported the molecular evidence of *H. misionera* inhabiting Amazon basin. The karyotype of *H. misionera* in the Amazon population showed 2n=40 and karyotypic formulae 20m+20sm, that added to C-banding, Ag-NOR and 18S results, are suggestive of a cytotype C. This work reveals the first record of *H. misionera* outside of La Plata basin and expands the species distribution for 2.700 km northward until the Marajó Island, estuary of Amazonas River.

Keywords: Integrative taxonomy, Thraira, Fresh water, geographic distribution, La Plata basin

Resumo

O grupo *Hoplias malabaricus* compreende seis espécies válidas e estima-se haver espécies crípticas ainda não descritas. Neste trabalho, analisamos uma população da bacia amazônica previamente identificada como *H. malabaricus*, porém apresentou maior afinidade filogenética com *H. misionera*, uma espécie descrita na bacia La Plata. A análise molecular por DNA barcoding revelou que essa população amazônica forma um clado monofilético com espécimes de *H. misionera* provenientes da bacia La Plata (BIN AAB1732) cuja distancia genética intragrupo (0.6%) é 10 vezes menor do que para o vizinho mais próximo (6.6%). A comparação morfométrica demonstrou pequena variação entre as populações amazônica e La Plata, sendo os primeiros ligeiramente maiores. Entretanto, os dados morfológicos corroboram com evidencia molecular e confirmam a ocorrência de *H. misionera* na bacia amazônica. O cariótipo de *H. misionera* na população amazônica apresentou 2n=40 e formula cariotípica 20m+20sm, que aliada aos resultados de banda C, Ag-NOR e 18S, sugerem que se trata do citótipo C. Esse trabalho revela o primeiro registro de *H. misionera* fora da bacia La

Plata e estende a distribuição da espécie por 2.700 km ao Norte, até a Ilha do Marajó, estuário do rio Amazonas.

Palavras-chave: Taxonomia integrativa, Traíra, Água doce, Distribuição geográfica, Bacia La Plata

Introduction

The wolf fish, locally named as “thrairas” in most part of South America are classified in the family Erythrinidae, encompassing 19 valid species and 3 genera: *Hoplias, Erythrinus* and *Hoplerythrinus* (Fricke *et al*., 2020). These species have peculiar morphological features, such as, cylindrical body form, rounded caudal fin, absence of adipose fin, 8-15 dorsal-fin rays, 10-11 anal-fin rays, numerous teeth on the palate and 34-47 lateral-line scales (Oyakawa, 2003). This family is restricted to Neotropical region, from Costa Rica to southern Ecuador in the west, and to Argentina in the southeast, being widespread in the South America freshwaters systems (Oyakawa, 2003; Berra, 2007).

Hoplias Gill 1903 is the richest genus comprising 14 valid species (Fricke *et al*., 2020). Based on morphological features, the genus *Hoplias* is arranged in three groups: *H. aimara*, *H. lacerdae* and *H. malabaricus* (Oyakawa, 1990; Mattox *et al*., 2006; Oyakawa, Mattox, 2009). The taxonomy of this genus was recently revised and new species were described (Azpelicueta *et al*., 2015; Rosso *et al*., 2016, 2018) and re-described (Mattox *et al*., 2014) for the *H. malabaricus* species complex. Currently, this complex comprises six species: *Hoplias malabaricus* (Bloch, 1974), *Hoplias microlepis* (Günther, 1864), *Hoplias teres* (Valenciennes, 1847), *Hoplias mbigua* Azpelicueta, Benítez, Aichino & Mendez 2015, *Hoplias misionera* Rosso, Mabragaña, González-Castro, Delpiani, Avigliano, Schenone & Díaz de Astarloa 2016 and *Hoplias argentinensis* Rosso, González-Castro, Bogan, Cardoso, Mabragaña, Delpiani & Díaz de Astarloa 2018. The improvement in taxonomic discrimination, intimately agrees with historical cytogenetic studies demonstrating that *H. malabaricus* is a species complex that hinders cryptic diversity (Bertollo *et al*., 1997, 2000; Born, Bertollo, 2006; Cioffi *et al*., 2009; Blanco *et al*., 2010; da Rosa *et al*., 2014).

Just recently, molecular results revealed the existence of several fully supported lineages within the once considered to be the continentally distributed *H. malabaricus* (Cardoso *et al*., 2018; Jacobina *et al*., 2018). Many of these lineages were found in the Amazon basin, where Marques *et al*. (2013) earlier demonstrated a conspicuous genetic distinctiveness in a *Hoplias malabaricus* population (Haplogroup Gp2). This population was tentatively assigned to *H. misionera* (Cardoso *et al*., 2018) as it shares the same molecular identity (BIN AAB1732) of the type material of this species. Nevertheless, a taxonomic revision of the Amazon population is lacking. The eventual occurrence of *H. misionera* in the Amazon Basin would greatly expand the geographic distribution of this species, since it was

described from the Uruguay, Paraná and Paraguay River Basins in Argentina and southern Brazil (Rosso *et al*., 2016).

Hoplias misionera is distinguished from congeners by the presence of Y-shaped configuration in the medial margin of dentaries, total number of dorsal (14-16) and pectoral (12-14) fin rays, number of lateral-line scales (40-43), and series of the last vertical scales on caudal peduncle forming a marked curve (Rosso *et al.*, 2016). The evaluation of these characters as well as complementary molecular tools would certainly properly define the taxonomic status of the Amazon population postulated to be *H. misionera*. Indeed, modern integrative approaches combining morphological and molecular tools suggest that several divergent lineages may constitute fully independent species in Neotropical Teleosts (Pugedo *et al*., 2016; Rosso *et al*., 2018). Here, we investigate the taxonomic status of a *Hoplias* population from the Amazon basin by means of an integrative approach including morphological, DNA barcoding and cytogenetic considerations.

Materials and Methods

Ethics statement

The Brazilian government System of Authorization and Information in Biodiversity (SISBIO) provided the permits for fish collections (SISBIO N. 32653–3). The animals were anesthetized and euthanized through immersion in water containing Eugenol solution, following procedure approved by the Animal Use Ethics Committee (CEUA) of the Universidade Federal do Oeste do Pará (CEUA/UFOPA Nº 09003).

Sampling and Study Area

We analyzed 23 specimens collected during fieldwork in the Lower Amazonas and Trombetas Rivers, in Pará State (Tab. 1). Three specimens from Viçosa Island, in the Marajó archipelago, were purchased in a fish market in the Macapá city, Amapá State. These specimens were provisionally identified during fieldwork as *H. malabaricus*. The sampling localities are mapped in the Fig. 1.

Tab. 1. Updated geographic records of *H. misionera* in Amazon Basin

Amazon Basin Sector	Municipality	Collecting site	n	Lat	Lon	Reference
Lower Amazon River, Southern bank	Santarém	Maicá lake	2	-2.4745	-54.5357	Marques *et al.*, 2013
Lower Trombetas River, Northern bank	Óbidos	Pororoca-Mamauru streams	8	-1.9368	-55.5206	Marques *et al.*, 2013; This study
Lower Amazon River, Northern bank	Oriximiná	Sapucuá lake	2	-1.7751	-55.8699	Marques *et al.*, 2013; This study
Estuarine, Marajó archipelago	Chaves	Viçosa island	3	0.4494	-49.9999	This study
Lower Amazon River, Northern bank	Monte Alegre	Amazon River	5	-2.0105	-54.0713	This study
Lower Amazon River, Northern bank	Alenquer	Amazon River	1	-1.9592	-54.7411	This study
Lower Amazon River, Southern bank	Santarém	Aracampina	1	-2.4383	-54.3233	This study
Lower Amazon River, Northern bank	Alenquer	Centro do Arapiri	1	-2.0821	-54.9944	This study

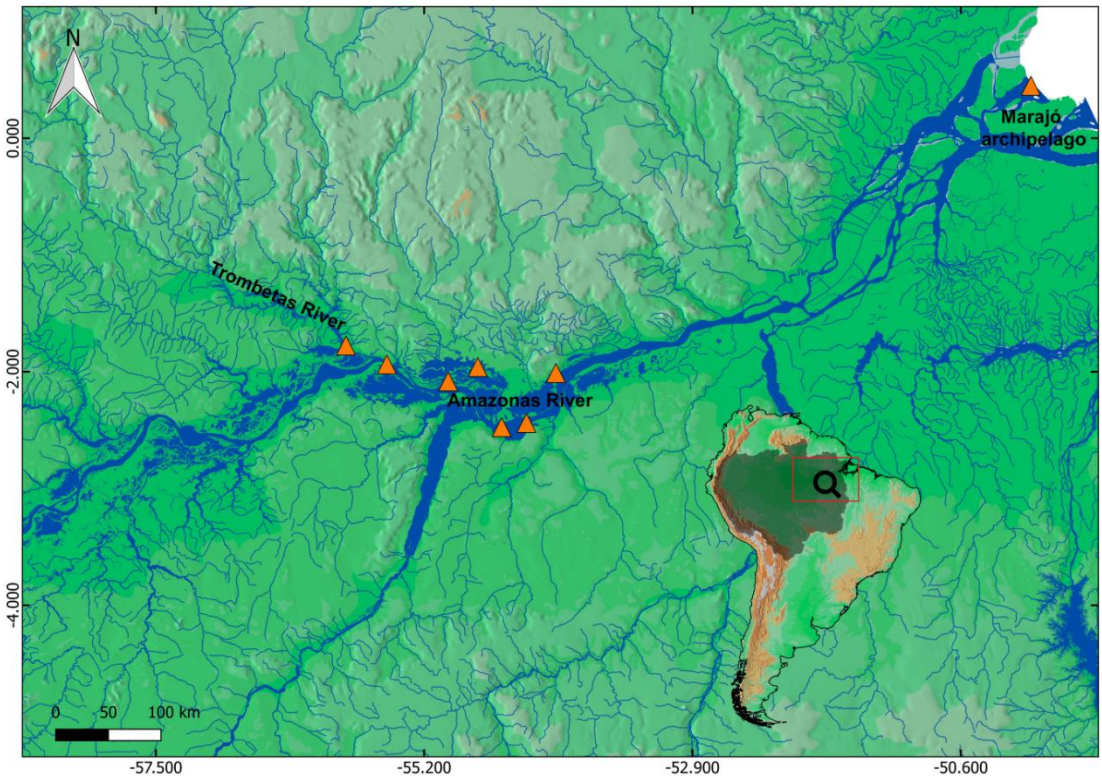


Fig. 1. Sampling localities of *H. misionera* in Amazon Basin.

Morphological analysis

The vouchers were fixed in 10% formalin during 72h, rinsed with tap water and preserved in 70% ethanol. Measurements and counts were made on the left side of the body following Fink and Weitzman (1974), Mattox *et al. (*2006) and Rosso *et al. (*2018). Counts were obtained by either visual or microscopic inspection. Linear body measurements were taken with a digital caliper to the nearest 0.1 mm. The examined vouchers (n=9) are deposited in the Ichthyological Collection of the Instituto de Ciências e Tecnologia das Águas, Universidade Federal do Oeste do Pará (UFOPA), Santarém, Pará, Brazil, and are listed in the material examined.

Molecular analysis

Before fixation, we collected muscle tissue fragments from each specimen and preserved them in absolute ethanol. Total genomic DNA was extracted following an adapted salting-out protocol (Aljanabi, Martinez, 1993; Vitorino *et al*., 2015). DNA barcoding sequences (COI mtDNA) were amplified by PCR using standard primers FishF1 and Fish R1 (Ward *et al*., 2005). Details of PCR profiles and sequencing reactions are given in Guimarães *et al.* (2018). In order to discriminate species we supplemented our COI data set with sequences downloaded from public repository Barcode of Life Database (www.boldsystems.org) (Ratnasingham, Hebert, 2007): *Hoplias misionera* (n=48) (Marques *et al*., 2013; Rosso *et al*., 2016; Cardoso *et al*., 2018), *Hoplias malabaricus* (n=5) (Cardoso *et al*., 2018), *Hoplias lacerdae* (n=2) (Cardoso *et al*., 2018). Detailed information on DNA barcoding sequences and specimen origin are listed in **S1**.

The sequences were aligned using the ClustalW Algorithm (Thompson *et al*., 1994) implemented in the software Bioedit (Hall, 1999). The Barcode Index Number (BIN) implemented in the BOLD System workbench was adopted to recognize Operational Taxonomic Units (OTUs) delimited as species (Ratnasingham, Hebert, 2013). For cluster visualization, we made a Neighbor-joining (NJ) tree based on Kimura-2-parameters (K2P) evolution model (Kimura, 1980) processed with the software MEGA X (Kumar *et al*., 2018) and edited with FigTree v.1.2.2 (http://tree.bio.ed.ac.uk/software/figtree/). We estimated K2P genetic distances using MEGA X (Kumar *et al*., 2018). All the new DNA barcoding sequences were deposited to the BOLD Systems database linked to the Project AMTRA: "Amazonian Trahiras".

Cytogenetics analysis

We analyzed the karyotype of ten specimens (AMTRA003-11, AMTRA022-11, AMTRA023-11, AMTRA024-11, AMTRA025-11, AMTRA028-11, AMTRA029-11, AMTRA030-11, AMTRA031-11, AMTRA037-11). Chromosome preparations were obtained from kidney cells after 24h of yeast mitosis stimulation (Lee, Elder, 1980) and exposed to 0.025% colchicine (0.01ml/g body mass) following Bertollo *et al.* (1978). The metaphases were examined through 5% Giemsa conventional staining. C-banding followed Sumner (1972). We detected Nucleolar Organizing regions by silver staining (Ag-NOR) following Howell, Black (1980). In situ hybridization (FISH) was applied to mapping ribosomal genes DNAr 18S. The probes were done by PCR using the primers Forward: 18Sf (5' CCG CTT

TGG TGA CTC TTG AT 3') and Reverse: 18Sr (5' CCG AGG ACC TCA CTA AAC CA 3') (Martins, Vicari, 2012). The FISH experiments followed procedures described in Pinkel *et al*. (1986) with minor adaptations as described in da Fonseca *et al*. (2018). The chromosomes in the karyotypes were visually classified as metacentrics (m) and submetacentrics (sm) based on arm ratio according to Levan *et al.* (1964) and arranged following Cioffi *et al*. (2009), Santos *et al*. (2009). In order to facilitate comparisons and cytotype discrimination we analyzed the pattern in size reduction of the first four largest pairs, a criterion previously adopted to distinguish cytotypes of similar diploid number (Bertollo *et al*., 1997).

Results

Morphological analysis

Taxonomy

Material examined (morphological data). *Hoplias misionera*: **Brazil: Lower Amazonas River basin**: UFOPA AMTRA126-130, 5, 214–238 mm SL • **Trombetas River: Sapucuá lake**: UFOPA AMTRA037, 1, Female, 188 mm SL • **Marajó archipelago: Viçosa Island:** UFOPA AMTRA122-125, 3, 228–262 mm SL).

Identification. Were examined nine specimens (n=9). We followed the morphological and meristic characters adopted in the species description (Rosso *et al*., 2016). The external morphology and ethanol-preserved coloration are showed in Fig. 2. Medial margins of contralateral dentaries converging to midline and then running parallel in a characteristic Y-shaped (n= 4) and V-shaped (n= 5) (Fig. 3a,b). A single premaxillary tooth row. First two premaxillary teeth large and caniniform, then four or five very small teeth followed by other two large canines. Maxilla with 35-44 teeth, first five increasing progressively in size. Dentary external series composed of 4 small teeth followed by two larger canines, then other series of 4-6 small teeth and 8-10 teeth arranged in a repetitive series of one large and one-two small conic teeth. Accessory ectopterygoids not fragmented, anteriorly expanded and bearing 12-14 conical teeth along their ventrolateral margins. Total dorsal-fin rays 14-15 (ii-12 n = 2; ii-13 n = 7). Total anal-fin rays 10-11 (i-9 n= 2; ii-9 n = 7). Total pectoral-fin rays 13 (i-12 n = 9). Tip of pectoral fin separated from pelvic-fin origin by 3-5 scales. Total pelvicfin rays 8 (i-7 n = 9). Tip of pelvic fin separated from vertical through anus by 2-3 scales. Total caudal-fin rays 17 (i-15-i n = 9). Predorsal scales (15-17) in an irregular series. Last vertical series of

scales on caudal peduncle forming a curve (Fig 4a). Lateral line complete with with one or two anterior scales without pores and 39-41 perforated scales. Longitudinal series of scales between dorsal fin origin and lateral line 5-5.5; between lateral line and pelvic fin origin 4-5. Longitudinal series of scales around caudal peduncle, invariable 20. First epibranchial with 9-12 plate-like denticulated gill rakers. One laminar gill raker on cartilage. First ceratobranchial with 4-6 more elongated rakers and 11-16 plate-like denticulated gill rakers. Laterosensory canal along ventral surface of dentary with four pores; a single laterosensory canal along infraorbitals invariable 11 pores. Laterosensory system of dorsal surface of head with 11-12 pores. Nasal bone: two pores, frontal bone: four-five pores, pterotic bone: two pores. One pore between parietal bones, on posterior end of symphysis. Supraopercle and extra-scapular bones with following combination of pores: 1:1 and 2:0.

Morphometric data of specimens of *H. misionera* from the different localities of the Amazon River Basin are summarized in Tab. 2.

Fig. 2. *Hoplias misionera*, UFOPA AMTRA129, 237 mm SL, Amazonas River, Alenquer, Pará, Brazil. Lateral view. Scale bar = 1 cm. Photo by L.R.R. Rodrigues.

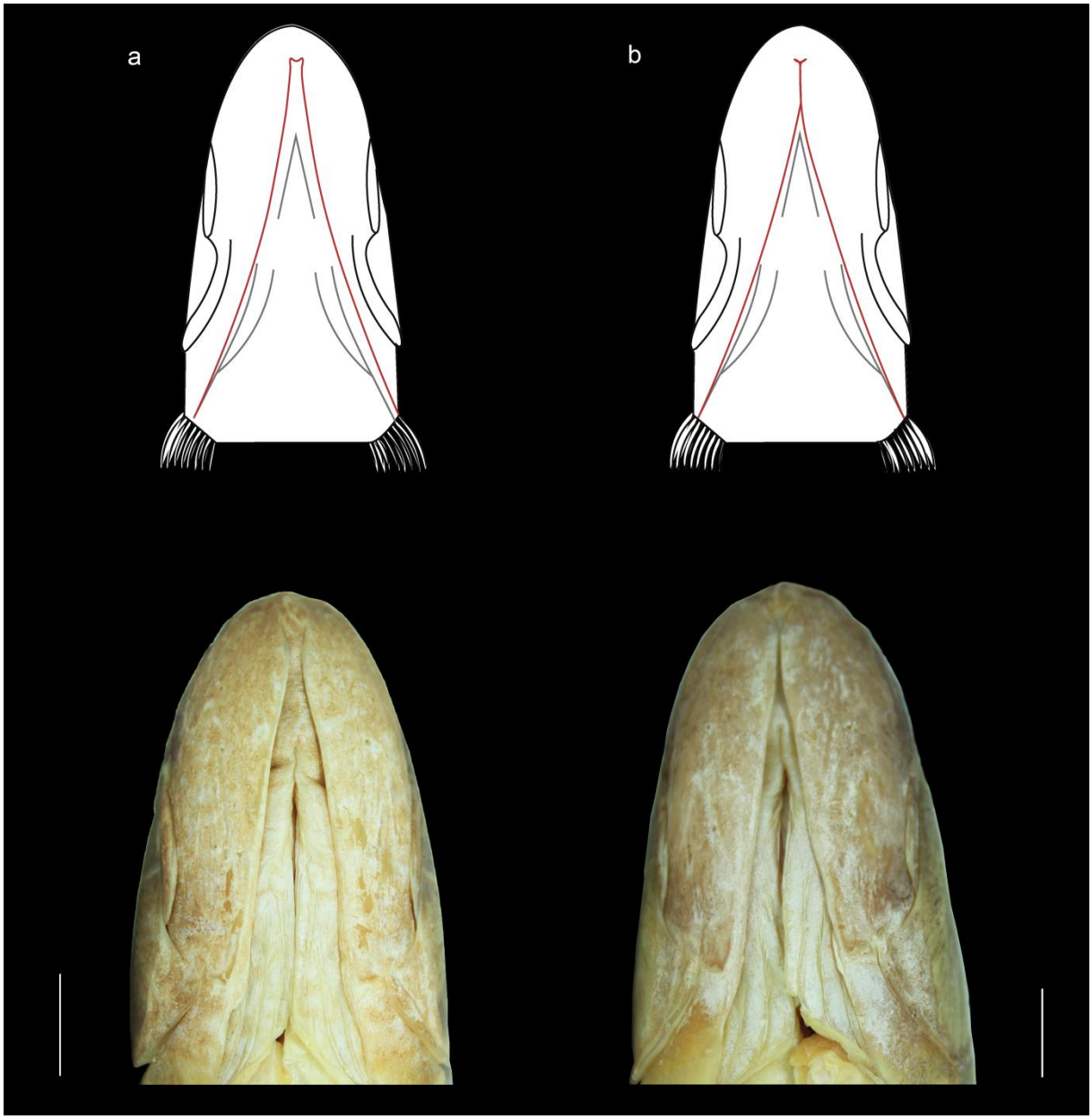


Fig. 3. Configuration of the medial margins of the dentary in *Hoplias misionera*. (a): V-shaped, UFOPA AMTRA127, 232 mm SL. (b): Y-shaped, UFOPA AMTRA126, 214 mm SL. Scale bars = 1 cm. Photos by L.R.R. Rodrigues. Illustration by T.M.A. Lima.

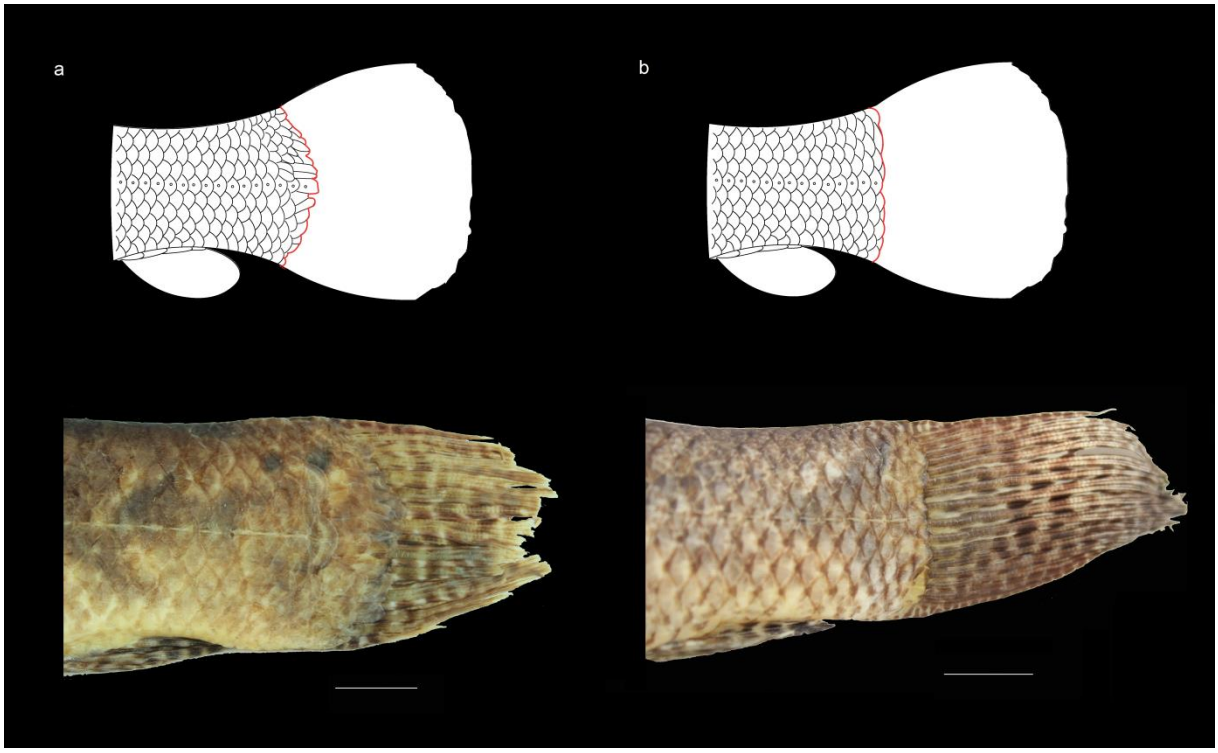


Fig. 4. Last vertical series of scales on the base of the caudal-fin rays. Comparison between *Hoplias misionera* (a), UFOPA AMTRA129, 237 mm SL and *Hoplias* cf. *malabaricus* (b), UFOPA AMTRA110, 201 mm SL and Scale bars = 1 cm. Photos by L.R.R. Rodrigues. Illustration by T.M.A. Lima.

Tab. 2. Morphometric data of *H. misionera* from the Amazon Basin; values 1-14 are percentages of the standard length and values 15-22 are percentages of head length.

	Count	Mean	Minimum	Maximum	standard deviation
Standard length (mm)	9	230.4	188	262	-
1. Body depth	6	19.2	18.1	20.3	0.6
2. Head length	9	29.2	28.0	30.9	1.0
3. Pectoral fin length	9	17.3	15.0	19.0	1.1
4. Pelvic fin length	9	18.0	16.1	20.0	1.2
5. Anal fin length	9	17.1	15.0	18.9	1.1
6. Dorsal fin length	9	32.5	30.4	35.9	1.8
7. Dorsal fin base length	9	18.9	17.1	20.9	1.0
8. Anal fin base length	9	8.4	7.2	9.4	0.7
9. Prepectoral distance	9	28.8	25.6	31.0	1.6
10. Prepelvic distance	9	52.4	48.5	55.0	2.1
11. Predorsal distance	9	44.1	42.4	46.2	1.2
12. Preanal distance	9	78.4	72.9	83.5	3.5
13. Caudal peduncle depth	9	13.2	12.3	14.5	0.7
14. Caudal peduncle length	9	12.8	11.9	13.5	0.5
15. Head depth	8	49.4	44.9	55.4	3.1
16. Snout length	9	25.7	24.0	28.5	1.4
17. Snout width	9	27.7	22.8	29.7	1.9
18. Snout depth	9	19.7	17.4	22.6	1.7
19. Pre nasal distance	9	15.8	13.7	18.7	1.4
20. Orbital diameter	9	16.2	13.8	18.0	1.2
21. Interorbital width	9	30.8	25.8	35.6	2.5
22. Upper jaw length	9	55.3	48.3	61.3	3.6

Molecular analysis

A data set with 68 DNA barcoding sequences (COI gene) of three species of *Hoplias* was assembled with 652bp long each sequence and revealed a base composition of 17.48% G, 27.30% C, 24.08% A and 31.13% T. The sequences did not show indels or stop codons. A genealogical tree based on distances (NJ) revealed three clusters that were congruent with species delimitation by BIN - *Hoplias misionera* (BIN AAB1732), *H. malabaricus* (BIN ABZ3047) and *H. lacerdae* (BIN ABW2258) (Fig. 5). All the specimens sampled to this work nested to *H. misionera* clade. Pairwise genetic distances indicated deep divergences (6.6-14%) between species; in contrast *H. misionera* populations from Amazon Basin and La Plata Basin diverged by just 0.6% (Tab. 3).

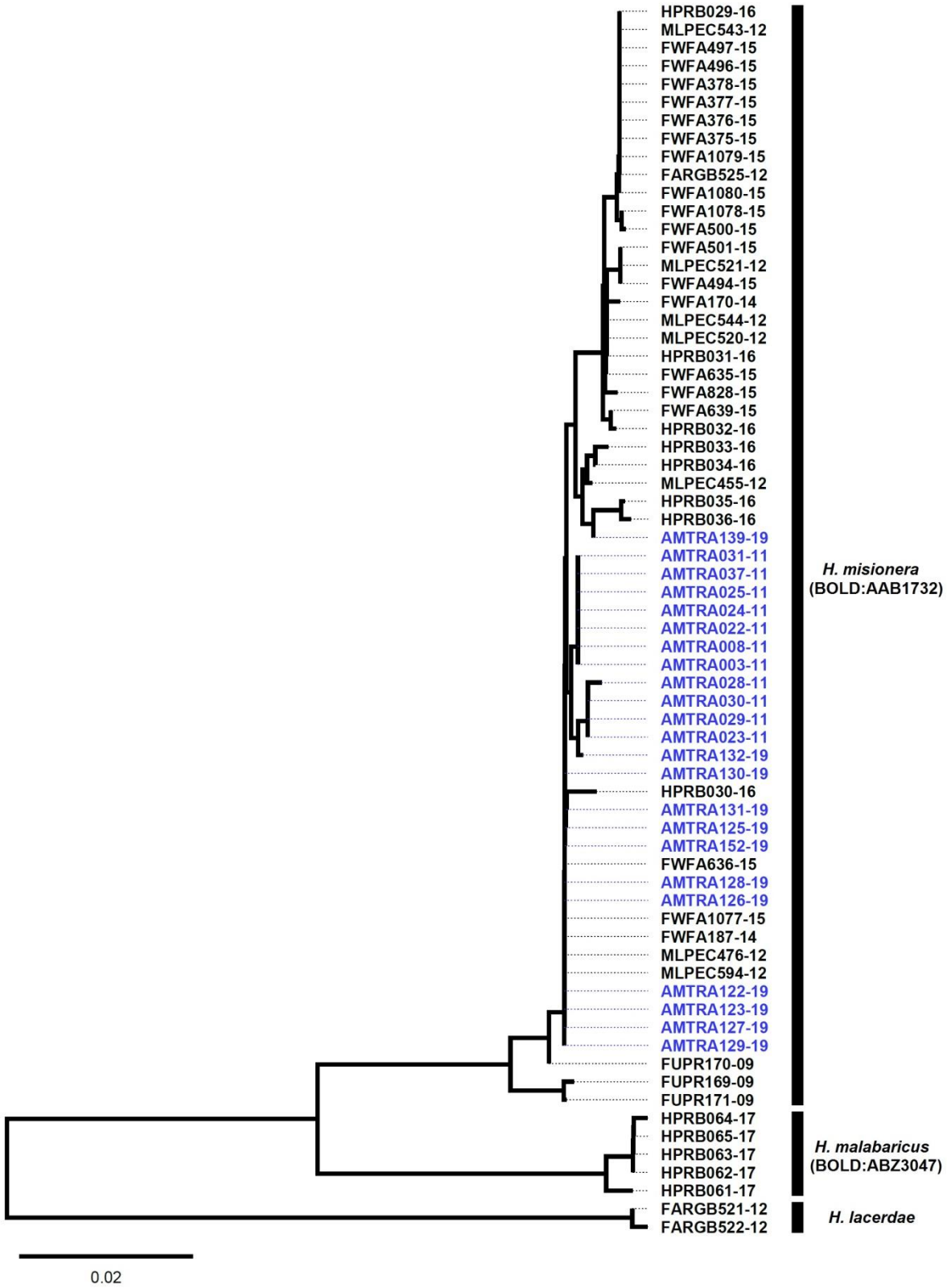


Fig. 5. Neighbor-joining tree of COI sequences of *Hoplias* congeners, based on K2P evolution model. The lateral bar indicate the partitions of species delimitation with BIN. The clade *Hoplias misionera* nested individuals from La Plata Basin population and Amazon Basin population (blue tips).

Tab. 3. Mean genetic distances (K2P) between *Hoplias* clusters. Bold values indicates the mean intraspecific distance.

Species/BIN(Population)	*H. misionera*		ABZ3047	ABW2258
	Amazon	La Plata		
H. misionera AAB1732 (Amazon)	**0.002**			
H. misionera AAB1732 (La Plata)	0.006	**0.006**		
H. malabaricus ABZ3047	0.066	0.069		
H. lacerdae ABW2258	0.139	0.139	0.145	

Cytogenetic analysis

The karyotype of *H. misionera* from Amazon basin presented 2n=40 chromosomes, where 20 were metacentrics and 20 submetacentrics (Fig. 6a). C-banding showed heterochromatic regions in the centromeres of all chromosomes and variable amount in the distal region in most of the metacentrics (pairs 4-10) and few submetacentrics (pairs 17-20). A conspicuous heterochromatic block was observed in the pericentromeric region of pair 13 (Fig. 6b). The Nucleolar organizing regions (Ag-NORs) were detected in the telomeric region of two submetacentric pairs (Fig. 7a). This Ag-NOR positions were coincident with the 18S rDNA hybridization, which revealed an additional fluorescent mark in the centromeric region of one metacentric pair (Fig. 7b). The size and heterochromatic band comparison between the largest chromosome pairs (1-2, 11-12) are showed in Figure 8.

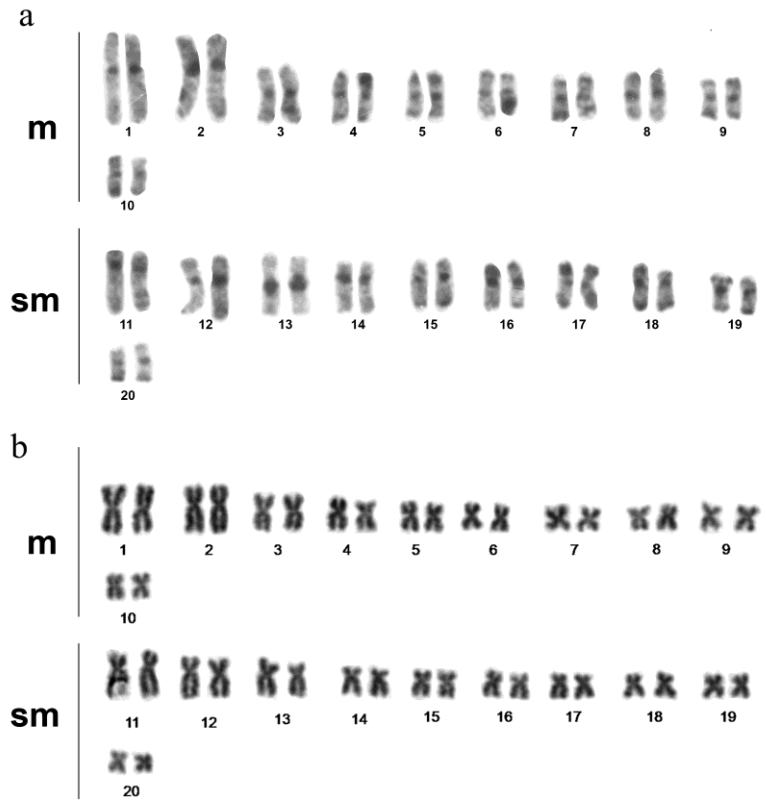


Fig. 6. Giemsa stained karyotype of *Hoplias misionera* from Amazon basin (2n=40 chromosomes) (**a**). C-banded stained karyotypes of *Hoplias misionera* from Amazon basin. (**b**). m=metacentric; sm=submetacentric.

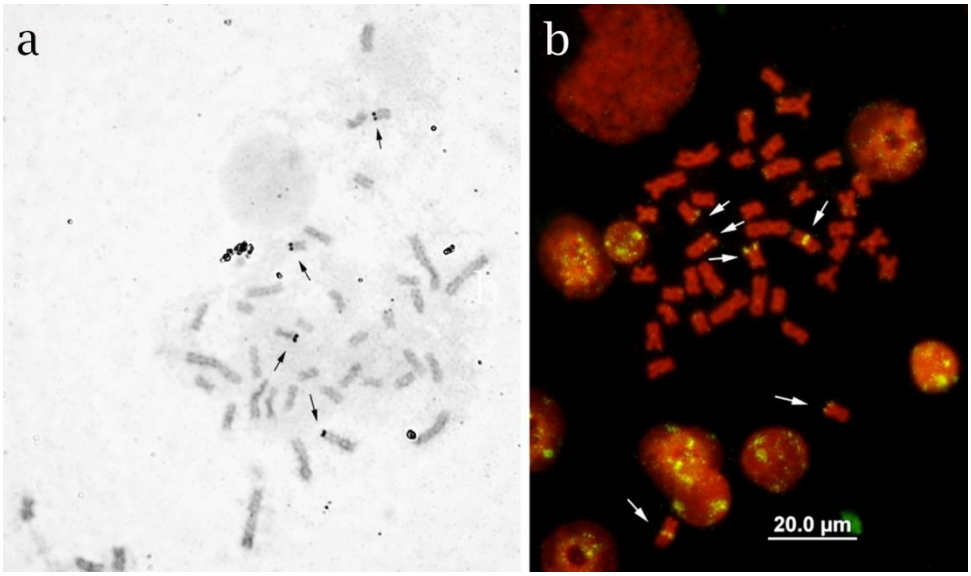


Fig. 7. The Ag-NOR bearing chromosomes (**a**); FISH of 18S rDNA probe in *Hoplias misionera* from Amazon basin. The chromosomes were stained with propidium iodide (red) and the probes were marked with Biotin-FITC (green) (**b**).

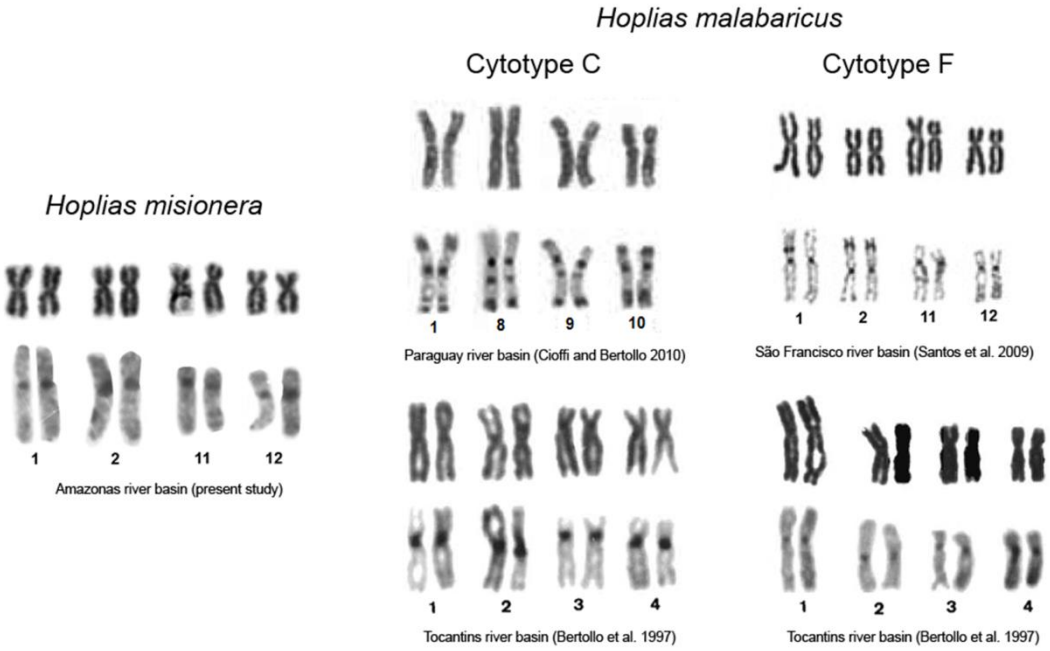


Fig. 8. Partial karyotype of *Hoplias misionera* and *H. malabaricus* (cytotypes C and F) that shows Giemsa and C-banding of the first four largest pairs.

Geographic distribution

The easternmost (and northernmost) collecting point in the Amazon basin, Viçosa Island, was situated in the Marajó Archipelago, a large island that lies at the mouth of the Amazon River. An updated distribution map of *H. misionera* is provided in the Fig. 9.

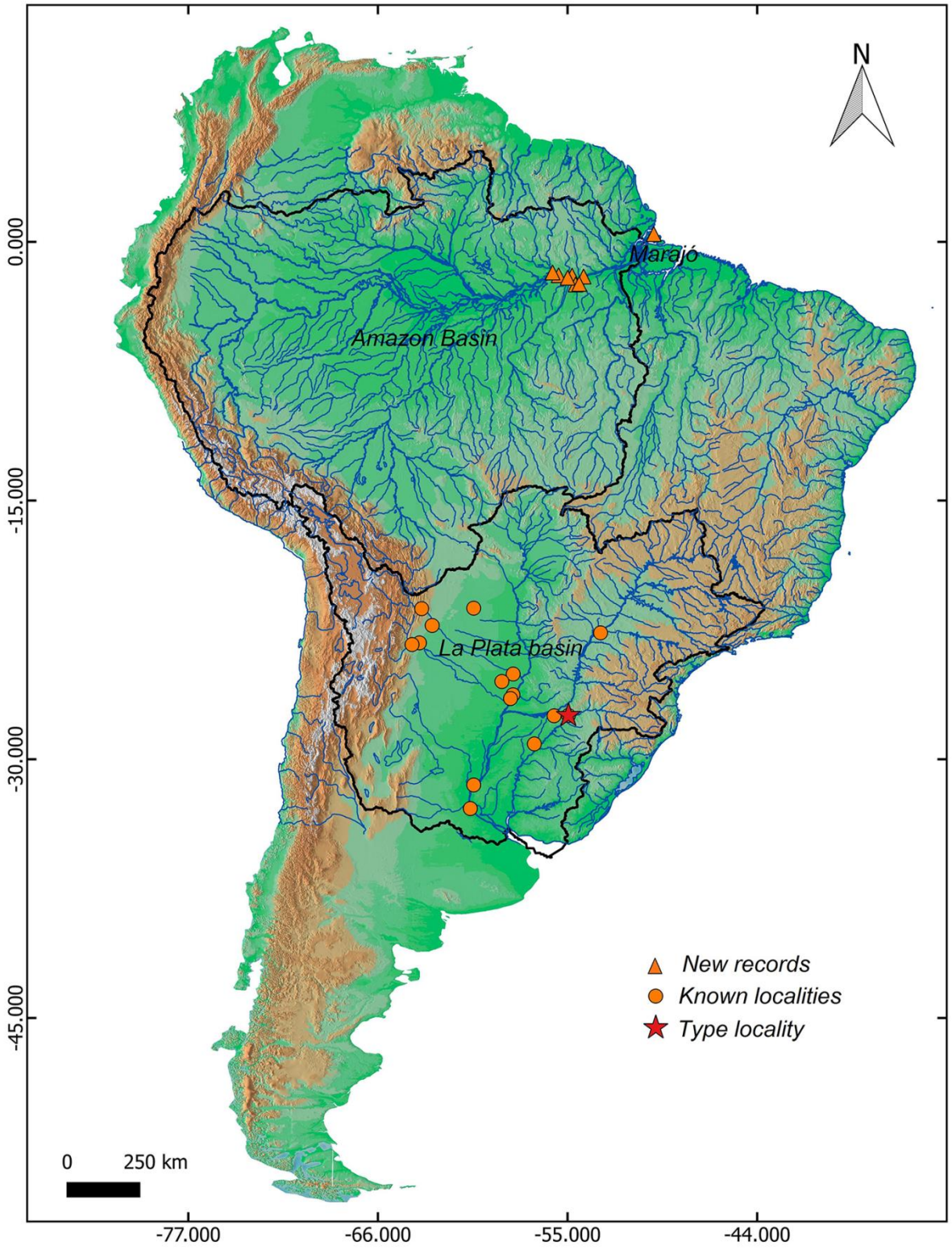


Fig. 9. Updated distribution map of *Hoplias misionera* showing former known localities in Argentina and southern Brazil and the new records from Amazon Basin (triangles). Star: type locality.

Discussion

Our results combined morphology, DNA and cytogenetics to characterize a population of *H. misionera* from the Amazon Basin and confirm the first record of this species outside the La Plata River Basin. The set of available diagnostic characters examined, fits the diagnosis of *H. misionera* (Rosso *et al*., 2016), with remarks on some characters. Overall, when compared with type specimens from La Plata River basin, the population of *H. misionera* from the Amazon basin can be characterized by having a slightly higher number of total anal-fin rays (10-11 vs. 10) and scales separating pelvic fin from anus (3-5 vs. 2). They also presented lower number of teeth (8-10 vs. 10-16) in the posterior repetitive series of the dentary and scales in the lateral line (39-41 vs. 40-43). A slightly wider range was observed in counts of gill rakers in epibranchial (9-12 vs. 10-11) and ceratobranchial (15-22 vs, 17-21) bones.

Morphometric analysis revealed some trends of variation between the Amazon and La Plata populations of *H. misionera*. The population of the Amazon basin was composed by larger specimens (188-262 mm vs. 39.22-174 mm of standard length) and showed different values in the following measures: body depth (18.1-20.3% vs 20.6-25.46%), head length (28.0-30.9% vs 30.61-34.57%), predorsal distance (42.4-46.2% vs 46.88-51.83%) and snout length (24.0-28.5% vs 20.47-24.72%).

Morphological differences between natural populations of geographically isolated fish that make up different hydrographic systems have been frequently reported (*e.g*. Neves, Monteiro, 2003; Shibatta, Hoffmann, 2005; Silva *et al*., 2009). In particular, these studies proposed that this external morphological distinction could be, in part, a response to selective pressure in different environmental conditions. *H. misionera* is a species that inhabits different environments, such as rivers, streams, lakes and dams, where these morphometric differences may be the result of different evolutionary patterns due to environmental conditions, emphasizing the singularity of each basin. Some of the observed variable morphometric characters could also be the result of ontogenetic variation, since there was a complete non-overlapping range of standard length between both populations. The entire range of lateral-line scales (39-43) combining Amazon and La Plata populations, was observed by Ota *et al*. (2018) in the upper Paraná River basin. Reia *et al*. (2020) further expanded the lower limit to 38 when revising four specimens in the Sucurí River basin, upper Paraná River.

Some specimens of *H. misionera* from Amazon basin displayed a Y-shaped arrangement on the medial margins of dentaries, a feature firstly proposed to distinguish *H. misionera* from all remainder species of *Hoplias* (Rosso *et al*., 2016). However, we also observed the V-shaped configuration in some specimens, as it was also observed lately for other populations of *H. misionera* elsewhere (Rosso, unpublished data). Despite the variable state of this character, the remaining morphological and meristic characters analyzed in the specimens collected in the Amazon River basin were largely congruent with those provided by Rosso *et al.* (2016).

The molecular evidence also confirms that tissue samples analyzed represent *H. misionera*. The speciation process within *Hoplias* has been linked to deep divergence in COI sequences. For instance, *H. misionera* and the recently described *H. argentinensis* diverged in 5.6 and 9.0% from the nearest neighbor (Rosso *et al*., 2016, 2018) respectively. In the Amazon basin, only one species of *Hoplias malabaricus* group (*H. malabaricus*) has been reported (see Oyakawa, 2003). Our results showed that *Hoplias misionera* of the Amazon Basin presented 6.6% of genetic distance to this species. Generally, a 2% divergence threshold is commonly observed to discriminate most Neotropical fish species (Pereira *et al*., 2013). The molecular data also demonstrated low divergence between the two populations of *H. misionera* (Amazon x La Plata) diverging by just 0.6%, further supporting the specific status of the Amazon specimens. Indeed, phylogenetic analyses using COI gene, showed that several taxonomically recognized species in the genus *Hoplias* form monophyletic groups (Cardoso *et al*., 2018). These results strongly suggest that barcode methodology should be considered as an additional diagnostic tool for confirmation of future new records for the genus *Hoplias*.

H. misionera from the Amazonas River showed karyotype 2n=40 and its macrostructure resembles to the cytotypes C and F (Bertollo *et al*., 1997, 2000; Cioffi *et al*., 2009; Santos *et al*., 2009). Besides the conservative diploid number both cytotypes are clearly distinguished based on the relative chromosomal size between the first four largest pairs (Bertollo *et al*., 1997) and some minor differences in the amount of constitutive heterochromatin that is slightly increased in the cytotype C (see Cioffi, Bertollo, 2010; Santos *et al*., 2009, 2016). These attributes are assumed to validate both cytotypes as distinct entities. Indeed, Bertollo *et al*. (1997) observed sympatry of cytotypes C and F in the Tocantins River population without evidence of hybridization.

Our specimens showed karyotypic formulae (20m+20sm) similar to that observed in *H. malabaricus* cytotype F from São Francisco river (Santos *et al*., 2009) and cytotype C from

Amazon basin (Marques *et al*., 2013; Santos *et al*., 2016). However, it diverges from cytotype C population from Bento Gomes River, a tributary of Paraguay River basin (Cioffi *et al*., 2009; Cioffi, Bertollo, 2010). In addition, *H. misionera* from Amazon basin shared identical pattern of Ag-NOR bearing chromosomes with *H. malabaricus* cytotype F and C from Amazon basin, but diverges from the Paraguay River basin population (see Cioffi *et al*., 2009; Cioffi, Bertollo, 2010).

It is noteworthy that cytotype C is widespread through South America (Bertollo *et al*., 2000) and shows variation in some cytogenetic markers, such as karyotypic formulae, Ag-NOR and 18S FISH marks (Cioffi *et al*., 2009; Cioffi, Bertollo, 2010; Santos *et al*., 2016; Guimarães *et al*., 2017). Variation of karyotypic formulae is frequently explained by the occurrence of chromosomal rearrangements but sometimes it could be an artifact resulted from misinterpretation of chromosome morphology in poor metaphases plates. Given the good quality of *Hoplias* chromosome preparations is plausible that populations from distinct hydrographic basins, showing cytotype C variants, can diverge by chromosomal rearrangements type pericentric inversions, which is a good explanation for the transformation of 20m+20sm to 14m+26sm such as observed between *H. malabaricus* (cytotype C) from Amazon and Paraguay river basins. It has been frequently demonstrated that *H. malabaricus* display multiple and variable Ag-NORs sites among distinct populations (Bertollo, 1996; Born, Bertollo, 2000; Vicari *et al*., 2005; Santos *et al*., 2016). *H. misionera* conserved this cytogenomic feature (multiple Ag-NORs) and because this species shared a similar Ag-NOR pattern with cytotypes C and F, is reasonable to conclude that based on this trait we cannot resolve its karyotype classification.

The relative size of the first four largest pairs has been considered a reliable trait to separate both cytotypes C and F (Bertollo *et al*., 1997). This last analysis failed to demonstrate the marked size reduction from the first to second pair, which is the main cytogenetic signature of cytotype F. In contrast, we observed a gradual size reduction congruent with the cytotype C. Distinct populations of cytotype C share these same characters and additionally show high amounts of heterochromatin (Bertollo *et al*., 1997; Santos *et al*., 2016; Cioffi *et al*., 2010), a feature that was also clearly observed in studied specimens of *H. misionera*. Additionally, the cytotype C is characterized by a nascent XX/XY sex system that leads to heterochromatin accumulation in the centromere of pair 11 (Cioffi, Bertollo, 2010), homologue to pair 14 (Santos *et al*., 2016) and that we postulate homologue to the *H. misionera* pair 13. In contrast, Santos *et al*. (2009) also recognized a probable XX/XY sex

system in the cytotype F but in this case, the chromosomal pair involved is the largest metacentric pair 1.

Based on the cytogenetic markers analyzed herein the karyotype of *H. misionera* from the Amazonas River must be classified as cytotype C. This cytotype has been recorded in the eastern and central portions of Amazon Basin (Bertollo *et al*., 2000; Santos *et al*., 2016) and distributes southward to Paraná and Paraguay Basins reaching the northeast Argentina in the region of Misiones Province (Lopes, Fennochio, 1994) where the type locality of *H. misionera* is situated (Rosso *et al*., 2016). Therefore, there is a possibility that the cytotype C remains conserved throughout the species distribution range. However, we recommend treating this assumption cautiously because karyotypic macrostructure can be a homoplastic character and could lead to mistaken inferences. Indeed, our results do not support the hypothesis that *H. misionera* populations from Argentina must be characterized as cytotype A, as proposed by Jacobina *et al*. (2018) from a geographic distribution interpretation. The sympatry of cytotypes A and C in the northeast Argentina has been already reported (Lopes, Fennochio, 1994). All these aspects highlight the need for conducting cytogenetic studies only for well-defined taxonomic species, if we wish to improve our knowledge about the relationships between taxonomic and karyotype diversity. Further investigations are also needed to understand other aspects of the cytogenomic patterns of *H. misionera* populations from the Amazon and La Plata basins.

The geographic range of *H. misionera* is widely expanded northerly from the original localities included in the species description. New occurrences reported for the Amazon Basin are situated 2700 km northwards from the northernmost location previously known for *H. misionera* and 3180 km from the type locality of the species (Rosso *et al*., 2016). The actual disjunct distribution of *H. misionera* in the Amazon-Paraguay systems confirms a biogeographic condition formerly suggested by Cardoso *et al*. (2018) grounded only in molecular data. Fish fauna shared between the Amazon and Paraguay rivers has been explained as the result of biotic dispersal events across wetlands connecting the headwaters of neighboring drainages (see Lundberg *et al*., 1998; Carvalho, Albert, 2011; Ota *et al*., 2014). The Paraguay Basin has approximately 307 species (Koerber *et al*., 2017), with about one-third shared with the Amazon basin (see Carvalho, Albert, 2011; Dagosta, de Pinna, 2019). Clearly, given the vastness of the Amazon basin, this region might still harbor large extensions of underexplored river systems hindering species diversity (Jezequel *et al*., 2020). In this scenario, the occurrence of other populations of *H. misionera* should not be ruled out. Further studies focusing on biogeographic and integrative scopes may fill these sampling gaps

and assess the morphological and genetic trait variation of *H. misionera* populations throughout the entire species distribution range.

Acknowledgements

We thank to the local fishermen who helped with the collection of specimens in the field. We are grateful to Tauanny Lima for provided the illustrations to Figs 3-4 and B. Viana for helped with the Figs 2- 4. KLAG thanks the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) by the master scholarship (88882.457158/2019-01) and Universidade Federal do Oeste do Pará for providing the finance support for trip to Argentina (process 23204.013496/2019-17). LRRR was granted by CNPq and FAPEAM by the project INCT-Adapta II.

References

Aljanabi SM, Martinez I. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Research. 1997; 25 (22):4692–4693.

Azpelicueta MM, Benítez M, Aichino D, Mendez CMD. A new species of the genus *Hoplias* (Characiformes, Erythrinidae), a tararira from the lower Paraná River, in Misiones, Argentina. Acta Zoologica Lilloana. 2015; 59 (1-2): 71–82.

Blanco DR, Lui RL, Bertollo LAC, Diniz D, Filho OM. Characterization of invasive fish species in a river transposition region: evolutionary chromosome studies in the genus *Hoplias* (Characiformes, Erythrinidae). Reviews in Fish Biology and Fisheries. 2010; 20(1), 1–8. https://doi.org/10.1007/s11160-009-9116-3

Berra TM. Freshwater fish distributions. The University of Chicago Press, Chicago. 2007; 645p.

Bertollo LAC, Takahashi KS, Moreira-Filho O. Cytotaxonomic considerations on *Hoplias lacerdae* (Pisces-Erythrinidae). Revista Brasileira de Genética. 1978; 2: 103–120.

Bertollo LAC. The Nucleolar Organizer Regions of Erythrinidae Fish. An uncommon situation in the genus *Hoplias*. Cytologia. 1996; (61): 75–81.

Bertollo LAC, Moreira-Filho O, Fontes MS. Karyotypic diversity and distribution in *Hoplias malabaricus* (Pisces, Erythrinidae): Cytotypes with 2n=40 chromosomes. Brazilian Journal of Genetics. 1997; (20): 237–242.

Bertollo LAC, Born GG, Dergam JA, Fenocchio AS, Moreira-Filho O. A biodiversity approach in the Neotropical Erythrinidae fish, *Hoplias malabaricus*. Karyotypic survey, geographic distribution of karyomorphs and cytotaxonomic considerations. Chromosome Research. 2000; 8(7): 603–613.

Born GG, Bertollo LAC. An XX/XY sex chromosome system in a fish species, *Hoplias malabaricus*, with a polymorphic NOR-bearing X chromosome. Chromosome Research. 2000; (8):111–118.

Cardoso YP, Rosso JJ, Mabragaña E, González-Castro M, Delpiani M, Avigliano E, Bogan S, Covain R, Schenone NF, Díaz de Astarloa JM. A continental-wide molecular approach unraveling mtDNA diversity and geographic distribution of the Neotropical genus *Hoplias*. PLoS ONE. 2018; 13(8):e0202024. https://doi.org/10.1371/journal.pone.0202024

Carvalho T, Albert J. The Amazon-Paraguay divide. In: Albert J, Reis RE, editors. Historical Biogeography of Neotropical Freshwater Fishes. London: University of California Press; 2011; p.193–202.

Cioffi M, Martins C, Centofante L, Jacobina U, Bertollo L. Chromosomal variability among allopatric populations of erythrinidae fish *Hoplias malabaricus*: mapping of three classes of repetitive DNAs. Cytogenetic and Genome Research. 2009; 125: 132–141. https://doi.org/10.1159/000227838

Cioffi MB, Bertollo LAC. Initial steps in XY chromosome differentiation in *Hoplias malabaricus* and the origin of an X1X2Y sex chromosome system in this fish group. Heredity. 2010; (105):554–561. https://doi.org/10.1038/hdy.2010.18

Dagosta FCP, De Pinna M. The fishes of the Amazon: distribution and biogeographical patterns, with a comprehensive list of species. Bulletin of the American Museum of Natural History. 2019; 431:1–163.

Da Rosa R, Vicari MR, Dias AL, Giuliano-Caetano L. New Insights into the Biogeographic and Karyotypic Evolution of *Hoplias malabaricus*. Zebrafish. 2014; 11(3), 198–206; https://doi.org/10.1089/zeb.2013.0953

Fink W, Weitzman S. The so called Cheirodontin fishes of Central America with descriptions of two new species (Pisces: Characidae). Smithsonian Contributions to Zoology. 1974; 172: 1–45. https://doi.org/10.5479/si.00810282.172

Fonseca IC, Maciel LAM, Ribeiro FRV, Rodrigues LRR. Karyotypic variation in the long-whiskered catfish *Pimelodus blochii* Valenciennes, 1840 (Siluriformes, Pimelodidae) from the lower Tapajós, Amazonas and Trombetas Rivers. Comparative Cytogenetics. 2018; 12(3):285–298. https://doi.org/10.3897/CompCytogen.v12i3.22590

Fricke R, Eschmeyer WN, Van der Laan R. Eschmeyer's catalog of fishes: genera, species, references [Internet]. San Francisco: California Academy of Science; 2020. Available from: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp

Guimarães EMC, Carvalho NDM, Schneider CH, Feldberg E, Gross MC. Karyotypic Comparison of *Hoplias malabaricus* (Bloch, 1794) (Characiformes, Erythrinidae) in Central Amazon. Zebrafish. 2017; 14(1):2017. https://doi.org/10.1089/zeb.2016.1283

Guimarães KLA, de Sousa MPA, Ribeiro FRV, Porto JIR, Rodrigues LRR. DNA barcoding of fish fauna from low order streams of Tapajós River basin. PLoS ONE. 2018; 13(12): e0209430. https://doi.org/10.1371/journal.pone.0209430

Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium. 1999; 41:95–98.

Howell WM, Black DA. Controlled silver staining nucleolus organizer regionswith protective colloidal developer: a 1-step method. Experientia. 1980; 36: 1014–1015.

Jacobina UP, Lima SMQ, Maia DG, Souza G, Batalha-Filho H, Torres RA. DNA barcode sheds light on systematics and evolution of neotropical freshwater trahiras. Genetica. 2018; (146):505. https://doi.org/10.1007/s10709-018-0043-x

Jézéquel C, Tedesco PA, Bigorne R, Maldonado-Ocampo JA, Ortega H, Hidalgo M. *et al.* A database of freshwater fish species of the Amazon Basin. Scientific Data. 2020; 7:1–96. https://doi.org/10.1038/s41597-020-0436-4

Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution. 1980; 16: 111–120.

Koerber S, Litz TO, Mirande JM. Checklist of the freshwater fishes of Argentina (CLOFFAR). Ichthyological Contributions of Peces Criollos. 2017; 55: 1-11.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution. 2018; 35:1547–1549. https://doi.org/10.1093/molbev/msy096

Lee MR, Elder FFB. Yeast stimulation of bone marrow mitosis for cytogenetic investigtions. Cytogenetics and Cell Genetics. 1980; 26:36–40.

Levan A, Fredga K, Sandberg AA. Nomenclature for centromeric position on chromosomes. Hereditas. 1964; 52:201–220.

Lopes PA, Fenocchio AS. Confirmation of two different cytotypes for the neotropical fish *Hoplias malabaricus* Gill 1903 (Characiformes). Cytobios. 1994; 80:217–221.

Lundberg JG, Marshall LG, Guerrero J, Horton B, Malabarba MCSL, Wesselingh F. In: Malabarba LR, Reis RE, Vari RP, Lucena ZM, Lucena CAS, editors. The stage for Neotropical fish diversification: a history of tropical South American rivers. Phylogeny and classification of Neotropical fishes. Porto Alegre: Edipucrs. 1998; p.13–48.

Marques DF, Santos FA, da Silva SS, Sampaio I, Rodrigues LRR. Cytogenetic and DNA barcoding reveals high divergence within the trahira, *Hoplias malabaricus* (Characiformes:

Erythrinidae) from the lower Amazon River. Neotropical Ichthyology. 2013; 11(2):459–466. https://doi.org/10.1590/S1679-62252013000200015

Martins C, Vicari MR. Hibridização in situ em cromossomos de peixes. In: Guerra M, editor. Citogenética Molecular: Protocolos comentados. Ed. SBG, Ribeirão Preto. 2012; p.59–87.

Mattox GMT, Toledo-Piza M, Oyakawa OT. Taxonomic Study of *Hoplias aimara* (Valenciennes, 1846) and *Hoplias macrophthalmus* (Pellegrin, 1907) (Ostariophysi, Characiformes, Erythrinidae). Copeia. 2006; 2006: 516–528. https://doi.org/10.1643/0045-8511(2006)2006[516:TSOHAV]2.0.CO;2

Mattox GMT, Bifi AG, Oyakawa OT. Taxonomic study of *Hoplias microlepis* (Gunther, 1864), a trans- Andean species of trahiras (Ostariophysi: Characiformes: Erythrinidae). Neotropical Ichthyology. 2014; https://doi.org/10.1590/1982-0224-20130174

Neves FM, Monteiro LR. Body shape and size divergence among populations of *Poecilia vivipara* in coastal lagoons of south-eastern Brazil. Journal of Fish Biology. 2003; 63(4):928–941. https://doi.org/10.1046/j.1095-8649.2003.00199.x

Ota RP, Lima F, Pavanelli C. A new species of *Hemigrammus* Gill, 1858 (Characiformes: Characidae) from the Rio Madeira and Rio Paraguai basins, with a redescription of *H. lunatus*. Neotropical Ichthyology. 2014; 12: 265–279.

Ota RR, Deprá GC, Graça WJ, Pavanelli CS. Peixes da planície de inundação do alto rio Paraná e áreas adjacentes: revised, annotated and updated. Neotropical Ichthyology. 2018; 16 (2): 1–111. https://doi.org/10.1590/1982-0224-20170094

Oyakawa OT. Revisão sistemática das espécies do gênero *Hoplias* (grupo lacerdae) da Amazônia brasileira e região leste do Brasil (Teleostei: Erythrinidae). MS.c. Thesis, Universidade de São Paulo, São Paulo. 1990; 114 p.

Oyakawa OT. Family Erythrinidae (Trahiras). In: Reis RE, Kullander SO, Ferraris CJr. (Orgs.). Check list of the freshwater fishes of South and Central America. Porto Alegre: Edipucrs. 2003; p.238–240.

Oyakawa OT, Mattox GMT. Revision of the Neotropical trahiras of the *Hoplias lacerdae* species group (Ostariophysi: Characiformes: Erythrinidae) with descriptions of two new species. Neotropical Ichthyology. 2009; 7(2):117–140.

Pereira LH, Hanner R, Foresti F, Oliveira C. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? BMC Genetics. 2013; 14: 20. https://doi.org/10.1186/1471-2156-14-20

Pinkel D, Straume T, Gray J. Cytogenetic analysis using quantitative, high sensitivity, fluorescence hybridization. Proceedings of the National Academy Sciences. 1986; 83: 2934–2938.

Pugedo ML, de Andrade Neto FR, Pessali TC, Birindelli JLO, Carvalho DC. Integrative taxonomy supports new candidate fish species in a poorly studied neotropical region: the Jequitinhonha River Basin. Genetica. 2016; 144(3):341–349.

Ratnasingham S, Hebert PDN. BARCODING, BOLD: The Barcode of Life Data System (www.barcodinglife.org). Molecular Ecology Notes. 2007; 7: 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Ratnasingham S, Hebert PDN. DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. PLoS ONE. 2013; 8(7): e66213. https://doi.org/10.1371/journal.pone.0066213

Reia L, Silva GSC, Garcia-Ayala JR, Vicensotto AMPF, Benine RC. Ichthyofauna of the ribeirão Sucuri, a tributary of the rio Tietê, upper rio Paraná basin, southeastern Brazil. Check List. 2020; 16 (3):711–728. https://doi.org/10.15560/16.3.711

Rosso JJ, Mabragaña E, González-Castro M, Delpiani MS, Avigliano E, Schenone N, Díaz de Astarloa JM. A new species of the *Hoplias malabaricus* species complex (Characiformes: Erythrinidae) from the La Plata Riverbasin. Cybium. 2016; 40(3): 199–208.

Rosso JJ, Gonzalez-Castro M, Bogan S, Cardoso Y, Mabragaña E, Delpiani M, Díaz de Astarloa JM. Integrative taxonomy reveals a new species of the *Hoplias malabaricus* species complex (Teleostei: Erythrinidae). Ichthyological Exploration of Freshwaters. 2018; 1–18. https://doi.org/10.23788/IEF-1076

Santos U, Völcker CM, Belei FA, Cioffi MB, Bertollo LAC, Paiva SR, Dergam JA. Molecular and karyotypic phylogeography in the Neotropical *Hoplias malabaricus* (Erythrinidae) fish in eastern Brazil. Journal of Fish Biology. 2009; 75(9), 2326–2343. https://doi.org/10.1111/j.1095-8649.2009.02489.x

Santos FA, Marques DF, Terencio ML, Feldberg E, Rodrigues LRR. Cytogenetic variation of repetitive DNA elements in *Hoplias malabaricus* (Characiformes-Erythrinidae) from white, black and clear water rivers of the Amazon basin. Genetics and Molecular Biology. 2016; 39(1):40–48. https://doi.org/10.1590/1678-4685-gmb-2015-0099

Silva EL, Centofante L, Miyazawa CS. Análise morfométrica em *Thoracocharax stellatus* (Kner, 1858) (Characiformes, Gasteropelecidae) proveniente de diferentes bacias hidrográficas Sul-americanas. Biota Neotropica, 2009; 9(2):71–76. https://doi.org/10.1590/S1676-06032009000200006

Shibatta OA, Hoffmann AC. Variação geográfica em *Corydoras paleatus* (Jenyns) (Siluriformes, Callichthyidae) do sul do Brasil. Revista Brasileira de Zoologia. 2005; 22(2):366–371. https://doi.org/10.1590/S0101-81752005000200010

Sumner AT. A simple technique for demonstrating centromeric heterochromatin. Experimental Cell Research. 1972; (75):304–306.

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research. 1994; 22(22):4673–80.

Vicari MR, Artoni RF and Bertollo LAC. Comparative cytogenetics of *Hoplias malabaricus* (Pisces, Erythrinidae). A population analysis in adjacent hydrographic basins. Genetics and Molecular Biology. 2005; (28):103–110. https://doi.org/10.1590/S1415-47572005000100018

Vitorino CA, Oliveira RCC, Margarido VP, Venere PC. Genetic diversity of *Arapaima gigas* (Schinz, 1822) (Osteoglossiformes: Arapaimidae) in the Araguaia-Tocantins basin estimated by ISSR marker. Neotropical Ichthyology. 2015; (13):557–568. https://doi.org/10.1590/1982-0224-20150037

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. DNA barcoding Australia's fish species. Philosophical transactions of The Royal Society B. 2005; 359:1847–1857. https://doi.org/10.1098/srtb.2005.1716

Supplementary Files

S1. Detailed information on DNA barcoding sequences and specimen origin.

Figure legends

Fig. 1. Sampling localities of *Hoplias misionera* in the Amazon Basin.

Fig. 2. *Hoplias misionera*, UFOPA AMTRA131-19, 237 mm SL, Amazonas River, Alenquer, Pará, Brazil. Lateral view. Scale bar = 1 cm. Photo by L.R.R. Rodrigues.

Fig. 3. Configuration of the medial margins of the dentary in *Hoplias misionera*. Y-shaped, UFOPA AMTRA126-19, 214 mm SL (**a**); V-shaped, UFOPA AMTRA127-19, 232 mm SL (**b**). (Scale bars = 1 cm). Illustration by T.M.A. Lima. Photos by L.R.R. Rodrigues.

Fig. 4. Last vertical series of scales on the base of the caudal-fin rays. Comparison between *Hoplias misionera* (**a**), UFOPA AMTRA131-19, 237 mm SL and *Hoplias* cf. *malabaricus* (**b**), UFOPA AMTRA110-18, 201 mm SL (Scale bars = 1 cm). Illustration by T.M.A. Lima. Photos by L.R.R. Rodrigues.

Fig. 5. Neighbor-joining tree of COI sequences of *Hoplias* species, based on K2P evolution model. The lateral bar indicates the partitions of species delimitation by means of the BIN. The clade *Hoplias misionera* nested individuals from La Plata Basin population and Amazon Basin population (blue tips).

Fig. 6. Giemsa stained karyotype of *Hoplias misionera* from Amazon basin (2n=40 chromosomes) (**a**). C-banded stained karyotypes of *Hoplias misionera* from Amazon basin. (**b**). m=metacentric; sm=submetacentric.

Fig. 7. The Ag-NOR bearing chromosomes (**a**); FISH of 18S rDNA probe in *Hoplias misionera* from Amazon basin. The chromosomes were stained with propidium iodide (red) and the probes were marked with Biotin-FITC (green) (**b**).

Fig. 8. Partial karyotype of *Hoplias misionera* and *H. malabaricus* (cytotypes C and F) that shows Giemsa and C-banding of the first four largest pairs.

Fig. 9. Updated distribution map of *Hoplias misionera* showing former known localities in Argentina and southern Brazil and the new records from Amazon Basin (triangles). Star: type locality.

Tables

Tab. 1. Detailed geographic information of new and former records of *Hoplias misionera* in Amazon Basin.

Tab. 2. Morphometric data of *Hoplias misionera* from the Amazon Basin; values 1-14 are percentages of the standard length and values 15-22 are percentages of head length.

Tab. 3. Mean genetic distances (K2P) between *Hoplias* clusters. Bold values indicate the mean intraspecific distance.

ARTIGO 2*

Molecular taxonomy of the species *Hoplias malabaricus* (Bloch, 1794) complex (Characiformes, Erythrinidae) and population genetics of a hidden cryptic species

Guimarães et al. 2020

*O artigo apresentado foi redigido conforme as diretrizes da revista *Scientific Reports*. No entanto, visando facilitar a compreensão do manuscrito, imagens, tabelas e suas respectivas legendas estão inseridas no corpo do texto. Para a versão final a ser submetida à revista, imagens serão enviadas separadamente e tabelas serão localizadas no final do manuscrito. As normas indicadas para redação de artigos pela revista estão disponíveis no link: https://www.nature.com/srep/author-instructions

Abstract

Hoplias malabaricus (Characiformes – Erythrinidae) is a species complex widespread from Northern to Southern South America continent. This group might constitute a good model for investigations of historical biogeographic events and the actual patterns of genetic structuring driving the ichthyofauna in the Neotropical freshwater systems. However, it is limited because of the confused taxonomy and the several cryptic species hidden in the complex. In this paper, we used DNA barcoding to delimit species from the Amazon Basin and adjacent drainages. We recognized the true *H. malabaricus* species and performed a population genetics analysis in this lineage. DNA barcoding revealed three to eight candidate species from the *H. malabaricus* complex inhabiting the study area. The largest clade recovered (BIN ABZ3047) was assumed as the true *H. malabaricus sensu stricto*. This species is structured in six population units: 1) Madeira River Basin (MRB), 2) Guiana Shields drainages (GSD), 3) Western-Northeast Atlantic Basin (WNAB), 4) Tapajós River Basin (TRB), 5) Lower Amazonas River confluences (LARC) and 6) São Francisco River Basin (SFRB). The populations TBR and SFRB were most differentiated and showed demographic fluctuations, where the later showed evidence of declining. The present distributional pattern is largely explained through a scenario from the last maximum glacial.

Keywords: Amazon Basin; COI; DNA barcoding; species delimitation; trahira.

Introduction

The river systems of Amazon basin covers a large portion of South America continent and harbor the major freshwater fish diversity of the world [1], [2]. For centuries, generations of naturalists explored the fish fauna in the Amazonas River and its larger tributaries revealing high levels of local richness and complex phylogenetic and evolutionary histories [3]. On the other hand, our knowledge about the diversification of Amazonian ichthyofauna and its biogeography and evolutionary history remains largely limited [3], [4]. Few studies have focused on the evolutionary history with a spatial/temporal scale of the Amazonian fish groups [5], [6].

The recent progress in molecular systematics approaches (e.g. DNA barcoding) has contributed to the knowledge on the hyperdiverse Amazonian ichthyofauna by improving the species discovery and taxonomy resolution, e.g. [7-9]. DNA barcoding is a high throughput

method to delimit animal species [10], including Neotropical fishes, e.g. [7], [8], [11]. Firstly, a cut-off threshold of 2% in the pairwise genetic distances between COI sequences was advocated as a good evidence to detect species boundaries [10], but nowadays the molecular species delimitation methods evolved to sophisticated statistical models (e.g. BIN, ABGD, GMYC) and integrative taxonomy approaches that combine total evidence from morphology, DNA, chromosomes, etc. [12], [13].

Understanding population divergences in a regional scale with an appropriate biogeographic unit are invaluable to disentangle intricate questions on the Amazonian ichthyofauna evolution. Then, fishes with large distributional ranges are considered suitable models for biogeographic studies, since dispersal is a consequence of temporary connections and displacement of river limits [14-16]. The thraira, *Hoplias malabaricus* (Bloch, 1974) is widely distributed throughout the South America continent, occurring from Northern watersheds (e.g. Orinoco and Amazon Basins) to Northeastern and Southeastern Brazilian drainages (e.g. São Francisco River Basin, Atlantic Coastal Basins) reaching some parts of the Northern Argentina [17-19]. This species might be a good model for biogeographic studies; however, *H. malabaricus* is a species complex, how demonstrated by cytogenetic and molecular studies [17, 20-22]. This nominal taxon is believed to hide several independent lineages that possibly are full species [23].

Hidden diversity within *H. malabaricus* complex is surely high. Cardoso and colleagues [23] recognized 16 clades of *H. malabaricus* complex delimited with DNA barcoding. Indeed, three of these clades (BINs: ACO5223, AAZ3734, AAB1732) were assigned to new species of this complex recently described from material collected in Argentina, *H. mbigua* Azpelicueta, Benítez, Aichino & Mendez 2015, *H. misionera* Rosso, Mabragaña, González-Castro, Delpiani, Avigliano, Schenone & Díaz de Astarloa 2016 and *H. argentinensis* Rosso, González-Castro, Bogan, Cardoso, Mabragaña, Delpiani & Díaz de Astarloa 2018 [24-26]. The remaining 12 putative species candidates were assigned to *H. malabaricus* and six of these lineages were recorded in Amazon basin (BINs: ABZ3046, ABZ3047, AAB1731, ACF3787, ACK2158 and ADG3393). DNA barcoding sequences revealed high level of genetic divergences between *H. malabaricus* lineages, even when they shared identical karyomorphs [22], [27]. In the Amazon Basin, five karyomorphs were recorded (A, C, E, F and G) but only karyomorphs E and G are restricted to this province [17].

Therefore, *H. malabaricus* complex remain as one of the most intriguing problem in the Neotropical ichthyofauna and the taxonomic confusion in this group is a serious constraint

for its use in studies of micro evolutionary processes [24]. In this study, we investigated the taxonomic status of *Hoplias malabaricus* species complex from the Amazon basin and adjacent drainages aiming to delimit species and explore microevolution patterns in *H. malabaricus sensu stricto* (BIN ABZ3047) that was supposed to be the true *H. malabaricus* species [23].

Materials and methods

Ethics statements

Fish specimens were collected under a SISBIO permission (N. 24384-1). For tissue sample and voucher preservation the specimens were euthanized with Eugenol following procedure approved by CEUA/UFOPA (Ethical Committee for Animal Research) Protocol Nº 09003/2016.

Sampling and Study area

We sampled 153 *H. malabaricus* specimens from 38 localities in Amazon Basin, Orinoco Basin and Guiana Shield drainages, (Supplemental S1, Fig. 1). The fish were captured using seine nets, casting nets and fish hook. Samples of epaxial muscle were preserved in absolute Ethanol and stored at -20°C. Vouchers specimens were photographed and measured for standard length (mm) and weight (g). The specimens were fixed with 10% formalin during 48h, washed and preserved in 70% Ethanol for deposit in the Fish Collection of the Institute of Water Science and Technology, Federal University of Western Pará (Brazil).

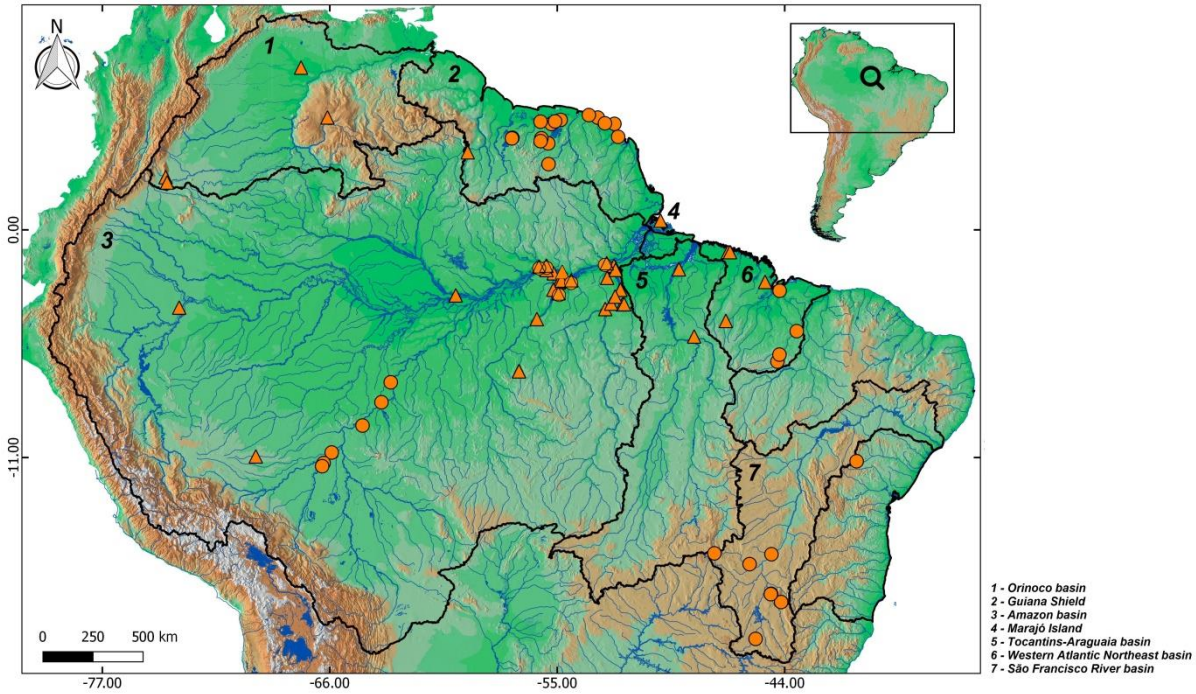


Figure 1. Collection sites of *H. malabaricus* in the Amazon basin and adjacent drainages. Localities sampled in this study (triangles), localities of sequences mined from internet (circles).

DNA extraction, PCR and Sequencing

DNA extraction followed a Salting-out protocol and the amounts were evaluated in a 1% agarose gel stained with Gelred (Biotium) [28], [29]. DNA fragments of Cytochrome c Oxidase subunit I (COI) mitochondrial gene was amplified by PCR using the standard DNA Barcoding primers Fish F1 and Fish R1 [30]. The reactions were assembled in 25 µL, containing 15 µL sterile H_2O, 2.8 µL dNTP mix (1.25 mM), 2.5 µL buffer 10X (200 mM Tris-HCl (pH = 8,4) + 500 mM KCl), 2.5 µL $MgCl_2$ (50 mM), 0.5 µL of each primer (5µM), 0.2 µL Taq DNA polymerase (5U/µL) and 1 µL of genomic DNA (around 100ng). The cycling profile followed as 95°C/2min, 35 cycles of 94°C/30sec, 54°C/30sec and 72°C/1min, and a final step of 72°C/10min. The PCR were processed with a Pxe 0.2 thermocycler (Thermo Scientific) and the amplified products were evaluated in a 1% agarose gel stained with Gelred. PCR products were cleaned with an adapted PEG8000 protocol [31]. COI sequences were obtained by Sanger method using the ABI PRISM Big Dye Terminator V.3 Cycle Sequencing kit (Applied Biosystems, Waltham, Massachusetts, USA). Sequencing reactions were made in 96-well plates with final volume of 10µL, containing 5 µL of sterile H_2O, 1.5 µL of sequencing buffer 5X, 0.5 µL of primer (10 µM), 1 µL of Big Dye mixture and 2 µL of cleaned PCR. The dye incorporation reactions followed 96°C/1 min; 35 cycles of 96°C/15

sec, 50°C/15 sec, and 60°C/4 min. The plates were precipitated in ethanol/EDTA, eluted with 10 μL Formamida Hi-Di and detected with an ABI 3500 genetic analyzer (Applied Biosystems), following the manufacturer instructions.

Molecular data analysis and species delimitation

In order to complement our dataset with reference material (*H. malabaricus stricto sensu*, BIN ABZ:3047), we downloaded 93 sequences of populations from Amazon, Surinam, French Guiana, Madeira River in Brazil and Bolivia, Itapecuru River and São Francisco River Basin, see [8], [22], [23], [27] (Supplemental S1).

The consensus sequences were assembled with Geneious R7 software (Biomatters, New Zeland) and aligned with Clustal W v1.4 [32]. We used GBLOCKS v0.91b [33] to inspect the alignment and trim the sequence tips and poor aligned regions. The new sequences generated in this work were deposited in a DNA barcoding repository (http://www.boldsystems.org) linked to the Project "Amazonian Trahiras (AMTRA)" (Supplemental S2).

H. malabaricus is a species complex. Then, in order to ensure that our data set for population genetic investigation had a single species, we performed a molecular species delimitation by: 1) Barcode Index Number (BIN) [34], 2) Generalized Mixed Yule Coalescent – GMYC [35], [36] and 3) Automatic Barcode Gap Discovery – ABGD [37]. The species delimitation through BIN analysis is an automated process implemented in the platform www.boldsystems.org. To GMYC analysis, we removed haplotype duplicates with ElimDupes (https://www.hiv.lanl.gov/content/sequence/elimdupesv2/elimdupes.html) and made an ultrametric tree using BEAST v1.8.0 [38], following these parameters: evolutionary model GTR+G chosen with jModelTest [39], molecular clock lognormal relaxed, Yule speciation process. The congener *Hoplias lacerdae* (BIN ABW2258) was adopted as outgroup. This Bayesian reconstruction done 80 million MCMC iterations, sampled each 1000 iterations with burn-in of 10%. The tree convergence and stability were checked with software Tracer v.1.7.1 [38] retaining the effective sample size (ESS) >200. The trees were combined with TreeAnotator v1.8.0 [38] and the output file saved in Newick tree format to be used for GMYC delimitation. The analysis of coalescence/speciation (GMYC) was processed following the model single threshold, in the environment R 3.4.3 [40] supplemented with libraries Splits (Species Limits by Threshold Statistics) [41] and Ape (Analyses of Phylogenetics and Evolution in R language) [42]. ABGD was processed in the platform

www.bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html using the alignment data set (fasta file) as input file. We set the parameters: model K80, Pmin. 0.001, Pmax 0.01 and barcoding gap width X=0.2.

In order to integrate the phylogenetic information, species delimitation and divergence time we processed a second Bayesian reconstruction following the procedures mentioned above with minor modifications: the best model was HKY+I+G, 200 million MCMC iterations sorted at each 1000 and 10% burn-in, and strict clock model. Divergence times were calibrated with a mutation rate of 1% per million years (Myr), which is conservative for fish mtDNA, e.g. [43], [44]. The resulting trees were assembled with TreeAnotator and the topology visualized/edited with FigTree v1.2.2 (http://tree.bio.ed.ac.uk/software/figtree/). Pairwise genetic distances between delimited species were measured following K80 model [45] using the software MEGA X [46].

Population genetics and biogeographic analysis

The individuals assigned to species *Hoplias malabaricus stricto sensu* following designation proposed in Cardoso and colleagues [23] were investigated for intraspecific genetic diversity and population structure. Parameters of the population genetics (e.g. haplotypes, nucleotide diversity, polymorphic sites) were analyzed with DNAsp v.5 [47]. A haplotype network was constructed based on Median Joining algorithm [48] with assistance of PopART software [49]. We adopted the software Geneland [50] to investigate population subdivisions and find the geographic population units, based on Bayesian statistics implemented with R package v3.4.0 [40]. The population genetic structure was evaluated through F_{ST} statistics and molecular variance analysis (AMOVA) implemented with Arlequin v.3.1 [51]. We assumed populations as the clusters of individuals such as revealed by Geneland analysis. For F_{ST} divergence, we follow Wright and colleagues [52] categories: low (0.00 – 0.05), moderate (0.05 – 0.15), high (0.15 – 0.25) and elevated (> 0.25).

To explore the demographic history we applied neutrality tests Tajima's D [53] and Fu's Fs [54], implemented with Arlequin v.3.1 [51]. Additionally, to detect population size variations we investigated the mismatch distributions and Bayesian skyline plot (BSP). These analyses were implemented with DNAsp v.5 [47] and BEAST v.1.8.0 [38]. BSP analysis adopted HKY+I+G model and 100 million MCMC sorted each 1000 iterations.

We constructed an ecological niche model with a maximum entropy algorithm MAXENT version3.3.3k [55-57] based on 66 georeferenced occurrence records (Fig 6a) and

19 bioclimatic variables from WorldClim (https://www.worldclim.org/data/bioclim.html). Such variables were correlated with 2.5 arc-minute spatial scale [58] and the distributional limits were assessed from the median occurrence with 50 bootstrap pseudo replicates. The theoretical distributional patterns were visualized with QGIS (Quantum GIS Development Team, www.qgis.org). We used jackknife permutations to evaluate the model performance gain, to identify and retain the most relevant explanatory variables.

Results

DNA barcoding and species delimitation

We analyzed 246 COI sequences of *H. malabaricus* complex from the Amazon basin and adjacent drainages. The sequences were 600bp long without stop codons neither indels. The data set showed a base composition of 29.4% (T), 29.5% (C), 23.3% (A), 17.7% (G).

Our species delimitation procedures revealed multiple independent evolutionary units that we assumed as putative candidate species. We found three (GMYC), four (ABGD) and eight putative species (BIN). The Bayesian inference showed two large clusters with several groups of individuals nested in sub clades that in part are restricted to some localities/eco regions, for instance: São Francisco River and Western Northeastern Atlantic Basin, Lower Tapajós River, Guiana Shields Drainages and Crepori River Basin. On the other hand, we observed sub clades nesting individuals from several localities scattered throughout the Amazon Basin (Fig. 2).

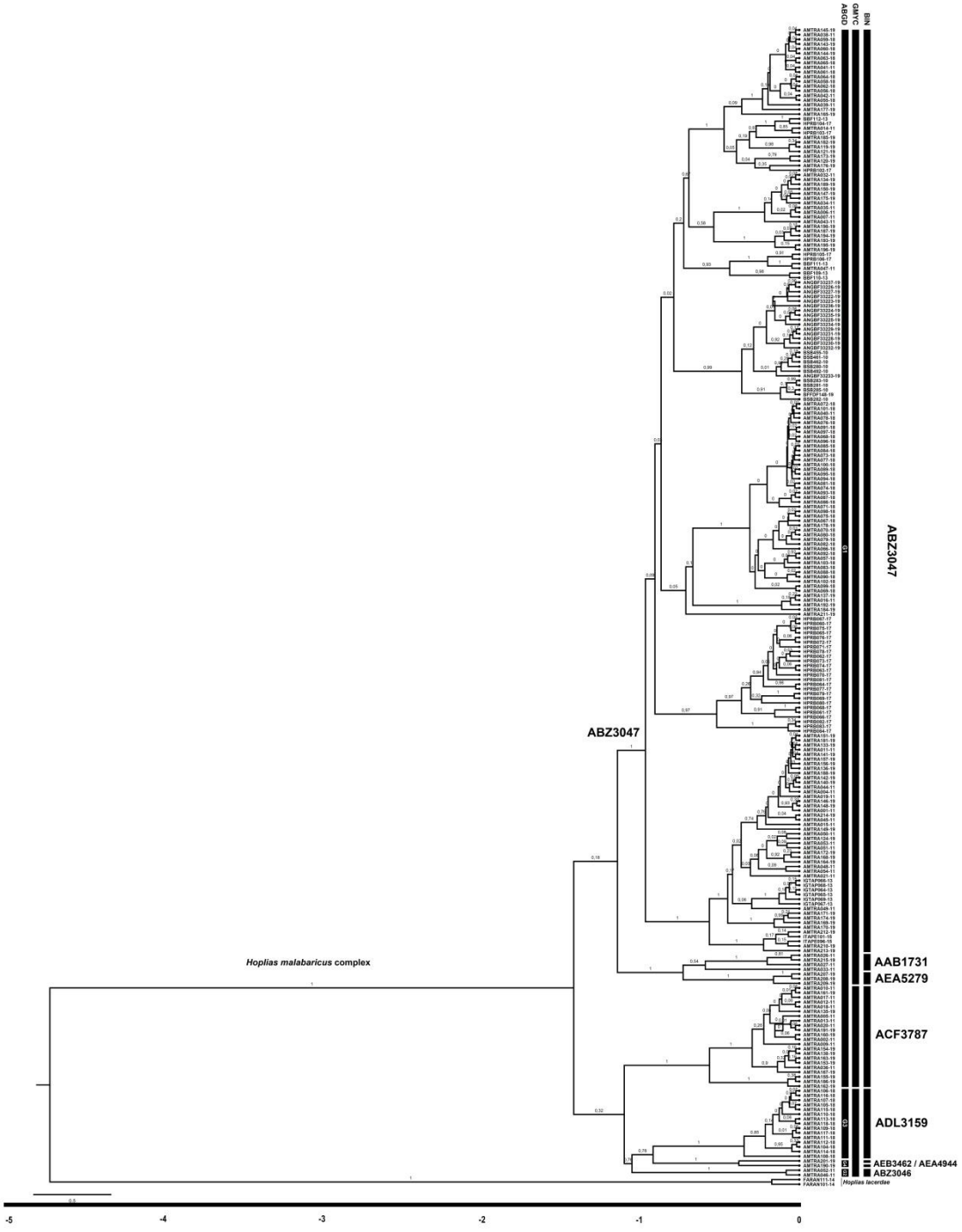


Figure 2. Bayesian phylogeny reconstructed from mtDNA lineages (COI) of *H. malabaricus* complex from Amazon Basin and adjacent drainages. Numbers in the nodes indicates statistical support of posterior probability. Black bars in the right show the partitions inferred as candidate species.

The BIN analysis revealed eight mitochondrial lineages (MOTUs): BINs–ABZ3047, AAB1731, AEA5279, ACF3787, ADL3159, AEB3462, AEA4944 and ABZ3046, that received strong support from the Bayesian genealogic inference (100% posterior probability -

PP). An exception was the BIN AAB1731 supported by 54%PP (Fig. 2). The pairwise genetic distances between groups (BINs) ranged from 1.5 to 4.7%, while the intra-BIN distances ranged from 0 to 1.7% (Table 1). The largest distances (4.6 - 4.7%) was detected for groups from the Orinoco drainage (BINs: AEB3462, AEA4944 and AEA5279).

Table 1. Pairwise genetic distances (K2P) between and within putative species of *H. malabaricus* complex based on COI sequences. Bold values indicate the intra-BIN distance. Number of individuals are showed in brackets.

Groups (Barcode Index Number) [n]	1	2	3	4	5	6	7	8
1 (ABZ3047) [198]	**0.017**							
2 (AAB1731) [4]	0.026	**0.009**						
3 (AEA5279) [3]	0.034	0.022	**0.003**					
4 (ACF3787) [22]	0.028	0.020	0.030	**0.004**				
5 (ADL3159) [15]	0.030	0.030	0.036	0.028	**0.001**			
6 (AEB3462) [1]	0.044	0.043	0.046	0.041	0.032	-		
7 (AEA4944) [1]	0.040	0.038	0.047	0.036	0.030	0.015	-	
8 (ABZ3046) [2]	0.035	0.030	0.036	0.024	0.029	0.041	0.035	**0.003**

The largest clade comprises all the individuals assembled to BIN ABZ3047. Based on coalescence/speciation analysis (GMYC), this group was delimited as an independent lineage with the addition of two small clades (BINs AAB1731 and AEA5279). Individuals assembled to BIN ABZ3047 are widely distributed throughout the Amazon Basin and adjacent drainages, including a sub clade restrict to the Guyana shields drainages. The second largest clade encompasses five mitochondrial lineages, BINs: ACF3787, ADL3159, AEA4944, AEB3462 and ABZ3046. In this clade, GMYC delimited two species that were congruent with BINs ACF3787 and ADL3159.

From barcoding gap assumption (ABGD), we found that putative species “Group 1” (G1) comprises four BINs (ABZ3047, AAB1731, AEA5279 and ACF3787). This arrangement results paraphyletic in the Bayesian reconstruction (Fig. 2). Other delimited species, Groups 2 and 3 (G2, G3) were fully congruent with BINs ABZ3046 and ADL3159. In turn, the Group 4 comprised two singleton BINs AEB3462 and AEA4944.

The group ADL3159 was restricted to Crepori River, a tributary from Tapajós drainage distantly more than 500km from the confluence zone between Tapajós and Amazonas Rivers. The group ABZ3046 includes only two individuals collected from the Urumari stream, a highly disturbed riverbed situated in the Santarém town, in the confluence zone of Tapajós and Amazonas Rivers. Despite minor discrepancy, we found clear evidences supporting a specific status to the largest clade comprising the BIN ABZ3047. Therefore, we

assumed herein that lineage BIN ABZ3047 is representative of a single species and proceeded analyzing its micro evolutionary history. We followed Cardoso and colleagues [23] designating ABZ3047 as *H. malabaricus sensu stricto* because this group includes individual from Suriname, type locality of true *H. malabaricus* species.

Population genetics of *H. malabaricus sensu stricto*

We filtered 198 individuals assigned to BIN ABZ3047 that we designated as *H. malabaricus sensu stricto*. We observed 71 polymorphic sites and 55 haplotypes (Supplemental S3). Based on genetic and spatial data processed through Geneland analysis, we depicted six subpopulation units distributed in the following regions: 1) Madeira River Basin (MRB), 2) Guiana Shields drainages (GSD), 3) Western-Northeast Atlantic Basin (WNAB), 4) Tapajós River Basin (TRB), 5) Lower Amazonas River confluences (LARC) and 6) São Francisco River Basin (SFRB), (Fig. 3).

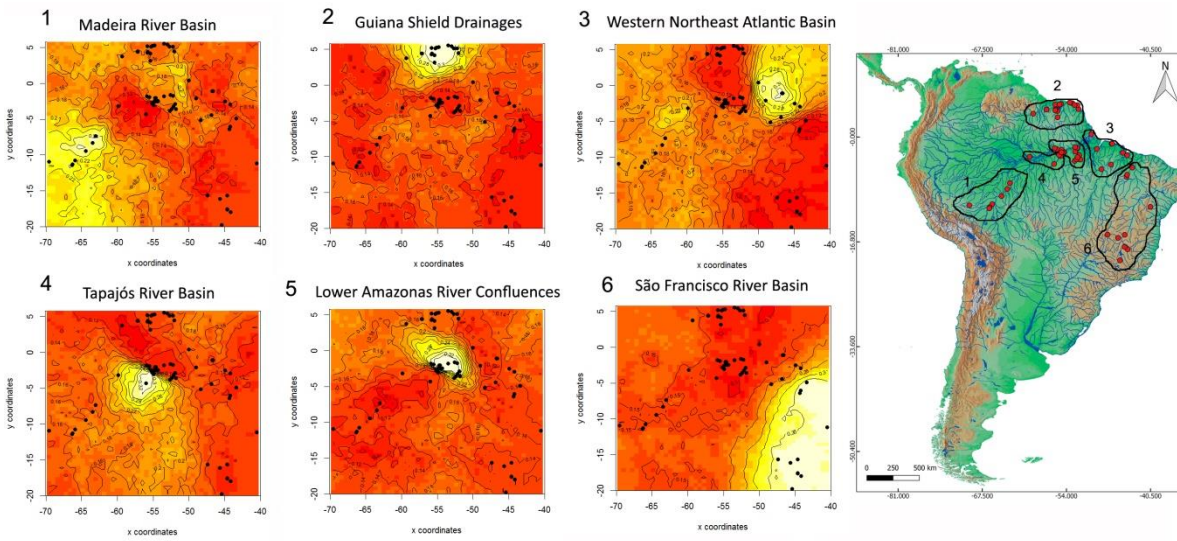


Figure 3. Subpopulations of *H. malabaricus* (BIN BOLD:ABZ3047) from Amazon Basin and adjacent drainages. The polygons shaded white indicates largest probabilities to associate haplotypes (black dots) to populations.

The subpopulations of *H. malabaricus sensu stricto* presented high haplotypic diversity (h) with 4 to 21 haplotypes. Exceptionally, the subpopulation 4 (TRB) showed the lowest values of diversity estimators: h=0.197 and nucleotide diversity (π)=0.0004. This subpopulation resulted negative values to Fu's Fs and Tajima's D statistics that is suggestive of neutrality deviation by purifying selection or population expansion (Table 2). The haplotype network revealed evidence for genetic structuring within *H. malabaricus sensu stricto* subpopulations (Fig. 4).

Table 2. Genetic diversity and values of neutrality tests of *H. malabaricus* subpopulations from the Amazon basin and adjacent drainages, based on mtDNA (COI gene). **N**= individuals, **Ha**= haplotypes, **S**= polymorphic sites, **h**= haplotypic diversity, **π**= nucleotide diversity. Bold values are statistically significant ($p<0.05$). Subpopulations: 1) Madeira River Basin (MRB), 2) Guiana Shields drainages (GSD), 3) Western Northeast Atlantic Basin (WNAB), 4) Tapajós River Basin (TRB), 5) Lower Amazonas River confluences (LARC) and 6) São Francisco River Basin (SFRB).

Subpopulation	N	Ha	S	*H*	π	Fu's FS	p-value	Tajima's D	p-value
1 MRB	10	8	16	0,956	0,0113	-1.69595	0.12600	-0.06384	0.52100
2 GSD	27	10	18	0,792	0,0062	-2.08809	0.16700	-1.44484	0.06000
3 WNAB	20	11	21	0,947	0,0156	-0.99411	0.33300	0.13010	0.57900
4 TRB	39	5	4	0,197	0,0004	**-4.97871**	**0.00000**	**-1.88116**	**0.00300**
5 LARC	75	21	42	0,876	0,0182	-0.95813	0.44000	-0.37985	0.41000
6 SFRB	26	4	3	0,769	0,0022	-1.12306	0.10000	-0.74439	0.25000

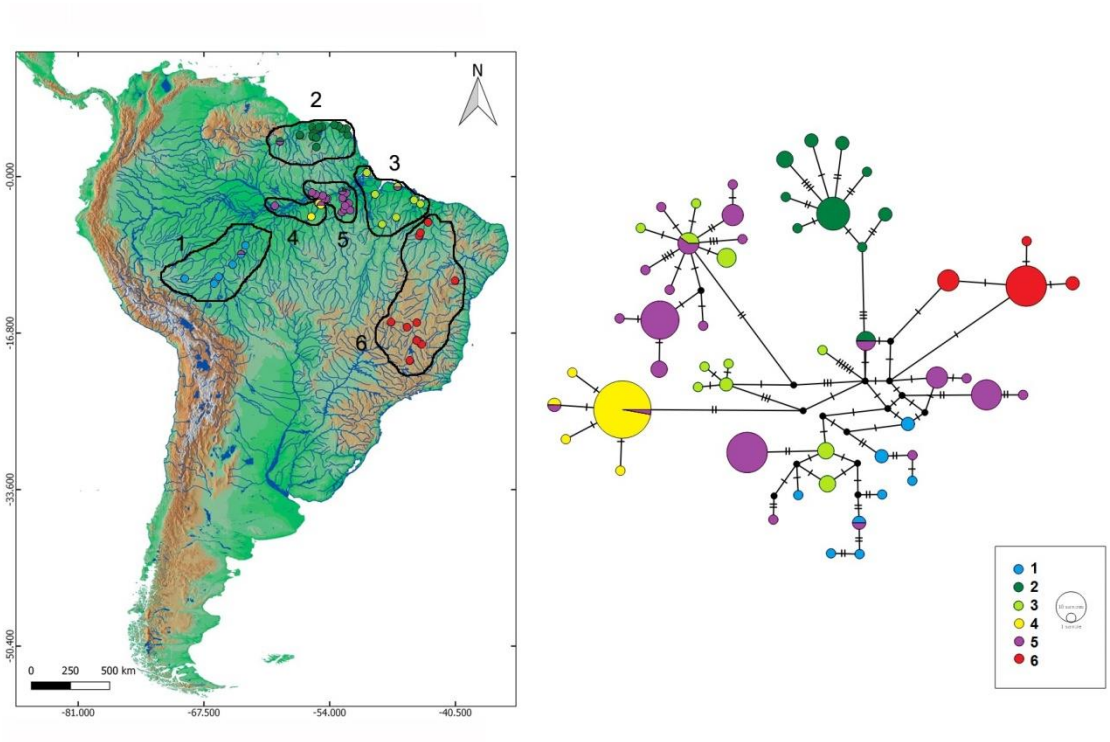


Figure 4. Haplotype network of *H. malabaricus stricto sensu* based on mitochondrial gene COI. The size of circles is proportional to haplotype frequency. Black circles indicate not sampled or possibly extinct haplotypes. Perpendicular bars show the mutational steps. Subpopulations: 1) Madeira River Basin (MRB), 2) Guiana Shields drainages (GSD), 3) Western Northeast Atlantic Basin (WNAB), 4) Tapajós River Basin (TRB), 5) Lower Amazonas River confluences (LARC) and 6) São Francisco River Basin (SFRB).

Subpopulations from the Amazon Basin (MRB, LARC and TRB) are geographically interconnected and share at least one haplotype with the adjacent basins (GSD and WNAB). In contrast, SFRB subpopulations showed only private haplotypes demonstrating higher

genetic differentiation. A pairwise Fst comparison demonstrate that subpopulation TRB is deeply differentiated from the all (Table 3). Its lowest Fst value (0.397) was recorded to TRBxLARC comparison and the highest differentiation was detected to TRBxSFRB (0.637). The subpopulation SFRB also presented high Fst values, ranging from 0.287 to 0.637 (Table 3). AMOVA results demonstrate that partitions of genetic variation support population genetic structure of *H. malabaricus sensu stricto*, because 48.49% of variation was detected among populations (p=0.000) (Table 4).

Table 3: Pairwise F_{ST} values between subpopulations of *Hoplias malabaricus stricto sensu* (BOLD:ABZ3047) based on mtDNA (COI gene) haplotypes. The numbers in brackets are p-values at significance level of 0.05. Subpopulations: 1) Madeira River Basin (MRB), 2) Guiana Shields drainages (GUSD), 3) Western Northeast Atlantic Basin (WNAB), 4) Tapajós River Basin (TRB), 5) Lower Amazonas River confluences and 6) São Francisco River Basin (SFRB).

Hierarchical level	F_{ST} Matrix					
	1	2	3	4	5	6
1 - MRB	0.00000					
2 - GSD	0.13619 (0.00000)	0.00000				
3 - WNAB	0.06009 (0.01465)	0.14309 (0.00000)	0.00000			
4 - TRB	0.51510 (0.00000)	0.50165 (0.00000)	0.46406 (0.00000)	0.00000		
5 - LARC	0.09188 (0.00391)	0.16307 (0.00000)	0.10016 (0.00000)	0.39706 (0.00000)	0.00000	
6 - SFRB	0.31900 (0.00000)	0.35512 (0.00000)	0.29974 (0.00000)	0.63698 (0.00000)	0.28766 (0.00000)	0.00000

Table 4. AMOVA results of *Hoplias malabaricus stricto sensu* from Amazon basin and adjacent drainages.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	P value
Among populations	5	316.925	2.01385	48.49	0.000
Within populations	191	408.654	2.13955	51.51	0.000
Total	196	725.579	415.339		

Demographic history of *H. malabaricus stricto sensu*

We found long term of demographic stability followed by a recent population expansion around 25.000 ybp, as demonstrated through BSP plots and mismatch distributions (Fig. 5). In particular, population from SFRB showed a tiny evidence for declining around 10.000 mya and the GSD population at 100.000 ybp. On the other hand, LARC population started an expansion around 50.000 ybp after a declining period of almost 100.000 years. The WNAB population expanded since 150.000 ybp and is stable at present. The unimodal distributions observed in the populations SFRB and TRB are suggestive of demographic fluctuations (Fig. 5).

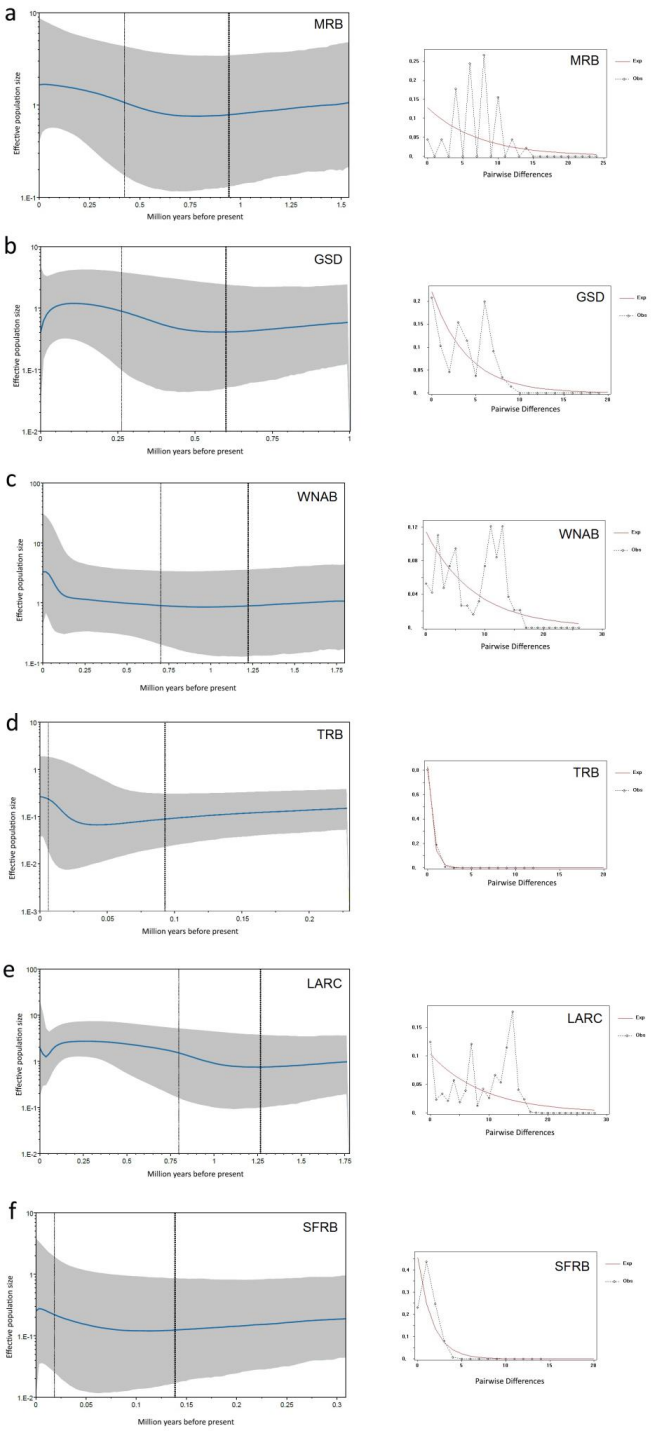


Figure 5. Mismatch distribution (rigth) and Bayesian skyline plot (BSP) (left) for the six population subdivisions of *H. malabaricus* (BIN BOLD:ABZ3047): a) Madeira River Basin (MRB), b) Guiana Shields drainages (GSD), c) Western Northeast Atlantic Basin (WNAB), d) Tapajós River Basin (TRB), e) Lower Amazonas River confluences (LARC) and f) São Francisco River Basin (SFRB). BSPs show changes in the effective population size. The thick solid line represents the median estimate and the margins of the surrounding area represent the largest posterior density ranges of 95%.

A paleogeographic model reconstructed with five bioclimatic variables (2, 9, 11, 15 and 16) showed that present distributional pattern is largely explained through a scenario from the last maximum glacial (21.000 ybp) and Anthropocene climate. However, this model indicates local extinctions of populations from SFRB through this geologic period. On the other hand, the areas from eastern Amazon basin seems to be the most propitious environment to the occurrence of this species for a long time (Fig. 6).

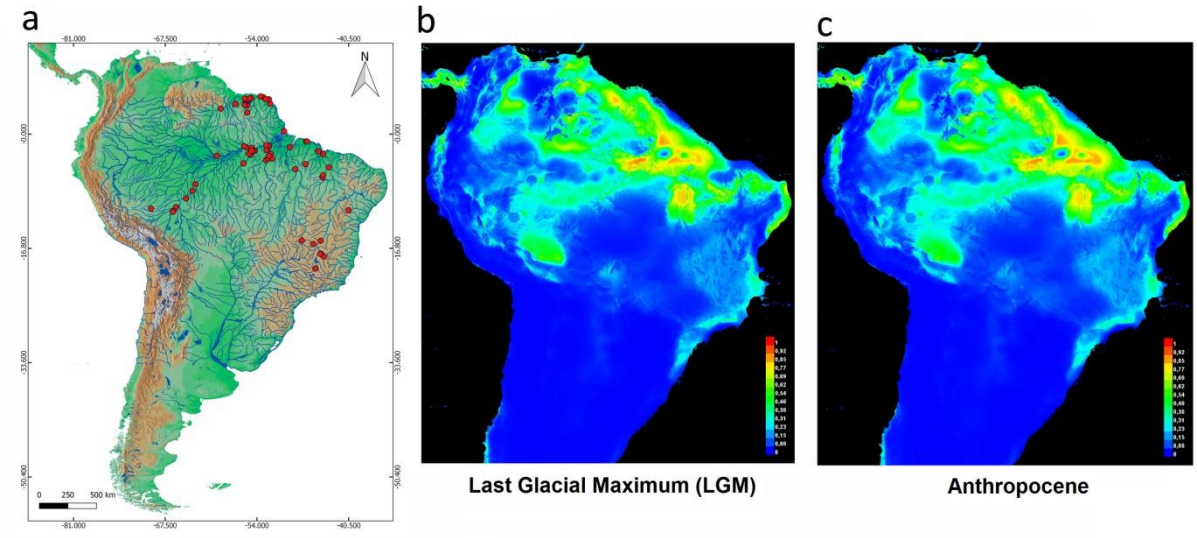


Figure 6: Present distribution of *H. malabaricus stricto sensu* (a). Bioclimatic model of the most suitable geographic areas for the occurrence of this species in the last glacial maximum (b) and Anthropocene (c). The warm colors (red, yellow) indicates high probability.

Discussion

The taxonomy of *Hoplias malabaricus* complex is still regarded as one of the great problems on the Neotropical ichthyology field [24]. Despite its conspicuous karyotypic variation, cytogenetic markers showed poor resolution to reconstruct phylogeny [17], [22], [27]. On the other hand, integrative approaches using morphological traits and DNA barcoding start to delineate a more detailed picture on the evolutionary history of this group [24], [26].

In this study, we observed deep divergence (max. 4.7%) between mitochondrial lineages (COI sequences) of *H. malabaricus* from the Amazon basin and adjacent drainages. This magnitude of genetic divergence was interpreted as a signature of cryptic speciation in the *H. malabaricus* complex, which was clearly demonstrated with species delimitation analysis based on distinct assumptions (ABGD, GMYC and BIN). Therefore, our results corroborate the existence of *H. malabaricus* species complex inhabiting Amazon Basin and

adjacent drainages, as previously postulated [17], [22], [23], [27]. Some of these lineages "putative species" (ABZ3047, ABZ3046, AAB1731 and ACF3787) were revealed in previous studies, see [23], [27]. However, we illuminated four new candidate species from the *H. malabaricus* complex (AEA5279, ADL3159, AEA4944 and AEB3462).

The BIN AEA5279, composed with three individuals from Guaviare-Orinoco drainage, was not supported with GMYC and ABGD. This group is closer to AAB1731, which is representative of Trombetas river population. Such phylogenetic affinity is suggestive of a population differentiation in the Northern drainages from Amazon basin. This portion of the Amazon territory is still poorly explored and the riverbed is mostly inaccessible because of many rapids and waterfalls in the middle to upper courses [59]. Further investigations are needed to give resolution on the taxonomy of this group. The group ADL3159 enclosed individuals from the Crepori River, a tributary of the middle Tapajos drainage. This group was supported by distinct methods and we believe it is a new undescribed species in the *H. malabaricus* complex. Currently, an integrative analysis with DNA barcoding, morphology and cytogenetics is ongoing in our lab, aiming to describe formally this taxon.

The BINs AEA4944 and AEB3462 are singletons, collected from the Rio Manapiare (Venezuela) in the Ventuari-Orinoco system. Both individuals were delimited as a single species with ABGD (Group 4). Herein, we consider the taxonomic status of these lineages most imprecise, and its delimitations as putative species must be regard cautiously. Putative species delimited from small number of sequences (< 5 individuals) are susceptible to bias [37]. On the other hand, we noted a phylogenetic link between population from Orinoco drainage with populations from the Amazon basin, in Trombetas River (Northern Amazonas drainage) and Crepori River (Southern Tapajós-Amazonas drainage). This genetic connection is tentatively explained by lineage sorting taking to maintenance of ancestral haplotypes or could be an artifact resulted from insufficient sampling. Although, Amazonas and Orinoco systems stay currently isolated, past contacts between them are well documented and such gene flow exchange had been observed in needlefish genus *Potamorrhaphis* [60].

Deep divergence of COI sequences is a known phenomenon of the lineage differentiation within *H. malabaricus* complex. Such phenomenon had been observed between populations from the same and distinct hydrographical basins [22] and between populations that share identical cytotypes [27]. COI deep divergence, 5.6 and 9.0% marked the speciation between *H. misionera* and *H. argentinensis* from their nearest neighbor [24], [26]. Jacobina and colleagues [22] also recorded deep divergence (7 to 7.3%) delimiting

putative species from *H. malabaricus* lineages in distinct Brazilian hydrographic basins. However, Cardoso and colleagues [23] demonstrated that speciation in *Hoplias* shows a large range of genetic distances (from around 1% to 20%). Indeed, *H. mbigua* diverged from its nearest neighbor (*H. intermedius*) by only 1.13% [23].

Because of the incomplete taxonomic resolution of the genus *Hoplias,* particularly in the *H. malabaricus* species group, few studies have focused on population genetics. Due to the high genetic diversity (cytotypes, DNA), morphological similarity and poor taxonomic knowledge, the species discrimination in field conditions is a challenge that certainly limits the investigation of micro evolutionary processes in *Hoplias*. Herein, we successfully identified *H. malabaricus* from cryptic congeners based on DNA barcoding sequences. In spite of COI gene has been widely used for species delimitation [10], [61], it can be suitable to explore population genetic structuring [9], [62]. The populations of *H. malabaricus stricto sensu* presented a clear genetic structure pattern, which was evidenced with Bayesian topology that shows multiple sub clades and corroborated with high Fst differentiation values and almost 50% of the genetic variation among populations. For comparison, Fst values of Tapajos River Basin population (TRB), the most differentiated, ranged from 46 to 63%. Aguirre and colleagues [63] measured a maximum F_{ST} divergence of 20% between *H. microlepis* populations from rivers and artificial impoundments in Ecuador. Demonstrations of genetic structure in non-migratory fishes from Amazon basin dispersed to adjacent drainages are found in Arapaimidae species, *Osteoglossum bicirrhosum* [64] and *Arapaima gigas* [65].

The modern Amazonian biodiversity dimensions was achieved during the Neogene (23-2.6 Mya), but the most ancient lineages probably stay there since Paleogene (66-23Mya) and Late Cretaceous (100-66Mya). During the middle to late Cenozoic, the Western Amazon basin was a lacustrine habitat while the eastern and central portions were repeatedly invaded by marine incursions resulting the isolation of Guiana and Brazilian Shield tributaries [5]. The river capture dynamics during the Neogene has been proposed as the main force driving the last diversification of aquatic and terrestrial Amazonian taxa that are ecologically restricted to water bodies and riparian forests [6], [66].

The *H. malabaricus* populations showed demographic evidences of expansion-decline cycles during Pleistocene, immediately before the last glacial maximum (LGM). We hypothesize that such demographic fluctuation in *H. malabaricus* may be synchronized with the recent glaciation and the geologic events linked to the formation of modern Amazonas River system [66], [67]. The population TRB presented demographic instability and neutrality

deviation that was interpreted as signature of recent expansion. The biological drives in such population may be adjusted with Pleistocenic phenomena. The Pleistocene glaciations had periods with high dry weather and marine regression with sea level 100 to 120m below the present. The highlands in Guiana shields and Andes Mountains accumulated thick ice caps and Tapajós region experienced long periods of erosion [68], [69].

In the present, *H. malabaricus sensu stricto* is dispersed to Guiana drainages, Amazon basin, Western Atlantic Northeast basin and São Francisco River basin [23]. These systems are supposed to experienced connection-isolation cycles during the Pleistocene period driven by climatic fluctuations and geomorphologic forces [14], [23], [70]. Our paleogeographic model revealed that hypothetical distribution in Pleistocene (Last Glacial Maximum, Anthropocene) is mostly congruent with the present day distribution. However, some discontinuities can be explored. We suppose that Pleistocene constraints affected mainly the populations from the São Francisco River Basin. The drainages from the Eastern portion of Amazon basin and Coastal drainages in the Western Northeast Atlantic basin seems to be important vectors for population dispersal.

Geomorphologic forces, mainly plate tectonics, marine incursions/regressions and climate fluctuations, recurrently shaped the amazon landscapes driving important processes on the aquatic systems, leading to intensive changes in the river courses and river captures. These phenomena may be involved with the taxonomic radiation and geographic dispersion of the fish populations throughout the Amazon basin [66], [71-76].

Data availability

All data generated or analysed during this study are included in this published article (and its Supplementary Information files).

References

1. Junk, W. J., Soares, M. G. M. & Baylet, P. B. Freshwater fishes of the Amazon River basin: their biodiversity, fisheries, and habitats. *Aquatic Ecosystem Health & Management* **10**(2), 153-173; 10.1080/14634980701351023 (2007).
2. Reis, R. E. *et al*. Fish biodiversity and conservation in South America. *Journal of Fish Biology* **89**(1), 12-47; 10.1111/jfb.13016 (2016).
3. Dagosta, F. C. P. & Pinna, M. C. C. de. A history of the biogeography of Amazonian fishes. *Neotropical Ichthyology* **16**(3); 10.1590/1982-0224-20180023 (2018).

4. Dagosta, F. C. P. & Pinna, M. C. C. de. Biogeography of Amazonian fishes: deconstructing river basins as biogeographic units. *Neotropical Ichthyology* **15**(3), 1-24; 10.1590/1982-0224-20170034 (2017).
5. Farias I. P. & Hrbek, T. Patterns of diversification in the discus fishes (Symphysodon spp. Cichlidae) of the Amazon basin. *Molecular Phylogenetics and Evolution* **49**, 32-43; doi:10.1016/j.ympev.2008.05.033 (2008).
6. Tagliacollo, V. A., Bernt, M. J., Craig, J. M., Oliveira, C. & Albert, J. S. Model-based total evidence phylogeny of Neotropical electric knifefishes (Teleostei, Gymnotiformes). *Molecular Phylogenetics and Evolution* **95**, 20-33; 10.1016/j.ympev.2015.11.007 (2015).
7. Paz, F. P. C., Batista, J. S. & Porto, J. I. R. DNA Barcodes of Rosy Tetras and Allied Species (Characiformes: Characidae: Hyphessobrycon) from the Brazilian Amazon Basin. *PLoS ONE* **9**(5), e98603; 10.1371/journal.pone.0098603 (2014).
8. Guimarães, K. L. A., de Sousa, M. P. A., Ribeiro, F. R. V., Porto, J. I. R. & Rodrigues, L. R. R. DNA barcoding of fish fauna from low order streams of Tapajós River basin. *PLoS ONE* **13**(12): e0209430; 10.1371/journal.pone.0209430 (2018).
9. Machado, V. N. *et al*. One thousand DNA barcodes of piranhas and pacus reveal geographic structure and unrecognised diversity in the Amazon. *Scientific Reports* **8**, 8387; 10.1038/s41598-018-26550-x (2018).
10. Hebert, P. D. N., Cywinska, A., Ball, S. L. & Dewaard, J. R. Biological identifications through DNA barcodes. *Philosophical transactions of The Royal Society B* **270**(1512), 313-321; 10.1098/rspb.2002.2218 (2003).
11. Pugedo, M. L., de Andrade Neto, F. R., Pessali, T. C., Birindelli, J. L. O. & Carvalho, D. C. Integrative taxonomy supports new candidate fish species in a poorly studied neotropical region: the Jequitinhonha River Basin. *Genetica* **144** (3):1-9; 10.1007/s10709-016-9903-4 (2016).
12. Dayrat, B. Towards integrative taxonomy. *Biological Journal of the Linnean Society* **85**(3), 407–415; 10.1111/j.1095-8312.2005.00503.x (2005).
13. Padial, J. M., Miralles, A., De la Riva, I. & Vences, M. The integrative future of taxonomy. *Frontiers in Zoology* **7**(1), 16; 10.1186/1742-9994-7-16 (2010).
14. Cardoso, Y. P. & Montoya-Burgos, J. I. Unexpected diversity in the catfish Pseudancistrus brevispinis reveals dispersal routes in a Neotropical center of endemism: The Guyanas Region. *Molecular Ecology* **18**(5), 947-64; 10.1111/j.1365-294X.2008.04068.x (2009).
15. Hoorn, C., Wesselingh, F. P., Hovikoski, J. & Guerrero, J. The development of the Amazonian mega-wetland (Miocene; Brazil, Colombia, Peru, Bolivia). *Amazonia, Landscape and Species Evolution* 123-142; 10.1002/9781444306408.ch8 (2010).
16. Albert, J. S. & Reis, R. E. Introduction to Neotropical Freshwaters. Historical Biogeography of Neotropical Freshwater Fishes (eds. Albert, J. S & Reis, R. E.). *University of California Press* 3-19 (Berkeley, 2011)
17. Bertollo, L. A. C., Born, G. G., Dergam, J. A, Fenocchio, A. S. & Moreira-Filho, O. A biodiversity approach in the Neotropical Erythrinidae fish, Hoplias malabaricus. Karyotypic survey, geographic distribution of karyomorphs and cytotaxonomic considerations. *Chromosome Research* **8**(7):603-613 (2000)
18. Oyakawa, O. T. Family Erythrinidae (Trahiras). Check list of the freshwater fishes of South and Central America (Reis, R. E., Kullander, S. O. & Ferraris, C. J. Jr. Orgs.). *Edipucrs* 238-240 (Porto Alegre 2003).

19. Dagosta, F. C. P. & Pinna, M. C. C. de. The fishes of the Amazon: distribution and biogeographical patterns, with a comprehensive list of species. *Bulletin of the American Museum of Natural History* **431**, 1-163 (2019).
20. Da Rosa, R., Vicari, M. R., Dias, A. L. & Giuliano-Caetano, L. New Insights into the Biogeographic and Karyotypic Evolution of Hoplias Malabaricus. *Zebrafish* **11**(3), 198-206; 10.1089/zeb.2013.0953 (2014).
21. Blanco, D. R., Lui, R. L., Bertollo, L. A. C., Diniz, D. & Filho, O. M. Characterization of invasive fish species in a river transposition region: evolutionary chromosome studies in the genus Hoplias (Characiformes, Erythrinidae). *Reviews in Fish Biology and Fisheries* **20**(1), 1-8; 10.1007/s11160-009-9116-3 (2010).
22. Jacobina, U. P. *et al*. DNA barcode sheds light on systematics and evolution of neotropical freshwater trahiras. *Genetica* **146,** 505; 10.1007/s10709-018-0043-x (2018)
23. Cardoso, Y. P. *et al*. A continental-wide molecular approach unraveling mtDNA diversity and geographic distribution of the Neotropical genus Hoplias. *PLoS ONE* **13**(8), e0202024; 10.1371/journal.pone.0202024 (2018).
24. Rosso, J. J. *et al*. Integrative taxonomy reveals a new species of the Hoplias malabaricus species complex (Teleostei: Erythrinidae). *Ichthyological Exploration of Freshwaters* 1-18; 10.23788/IEF-1076 (2018).
25. Azpelicueta, M. M., Benítez, M., Aichino, D. & Mendez, C. M. D. A new species of the genus Hoplias (Characiformes, Erythrinidae), a tararira from the lower Paraná River, in Missiones, Argentina. *Acta Zoologica Lilloana* **59** (1-2), 71-82 (2015).
26. Rosso, J. J. *et al*. A new species of the Hoplias malabaricus species complex (Characiformes: Erythrinidae) from the La Plata River basin. *Cybium* **40**(3), 199-208 (2016).
27. Marques, D. F., Santos, F. A., da Silva, S. S., Sampaio, I. & Rodrigues, L. R. R. Cytogenetic and DNA barcoding reveals high divergence within the trahira, Hoplias malabaricus (Characiformes: Erythrinidae) from the lower Amazon River. *Neotropical Ichthyology* **11**(2), 459-466; 10.1590/S1679-62252013000200015 (2013).
28. Aljanabi, S. M., Martinez, I. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* **25** (22), 4692-4693 (1997).
29. Vitorino, C. A., Oliveira, R. C. C., Margarido, V. P. & Venere, P. C. Genetic diversity of Arapaima gigas (Schinz, 1822) (Osteoglossiformes: Arapaimidae) in the Araguaia-Tocantins basin estimated by ISSR marker. *Neotropical Ichthyology* **13**, 557-568; 10.1590/1982-0224-20150037 (2015).
30. Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. & Hebert, P. D. N. DNA barcoding Australia's fish species. *Philosophical transactions of The Royal Society B* **359**:1847-1857; 10.1098/srtb.2005.1716 (2005).
31. Dunn, I. S. & Blattner, F. R. Sharons 36 to 40: Multi-enzyme, high capacity, recombination deficient replacement vectors with polylinkers and polystuffers. *Nucleic Acids Research* **15**, 2677-2698 (1987).
32. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**(22), 4673-80 (1994).

33. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**, 540-552; 10.1093/oxfordjournals.molbev.a026334 (2000).
34. Ratnasingham, S. & Hebert, P. D. N. DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *PLoS ONE* **8**(7), e66213; 10.1371/journal.pone.0066213(2013).
35. Pons, J. *et al*. Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. *Systematic Biology* **55**(4), 595-609; 10.1080/10635150600852011(2006).
36. Fujisawa, T. & Barraclough, T. G. Delimiting Species Using Single-Locus Data and the Generalized Mixed Yule Coalescent Approach: A Revised Method and Evaluation on Simulated Data Sets. *Systematic Biology* **62**(5), 707-724. 10.1093/sysbio/syt033 (2013).
37. Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* **21**(8), 1864-77; 10.1111/j.1365-294X.2011.05239.x (2012).
38. Drummond, A. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**, 214; 10.1186/1471-2148-7-214 (2007).
39. Posada, D. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* **25**, 1253-1256; 10.1093/molbev/msn083 (2008).
40. R Core Team. R: A Language and Environment for Statistical Computing. https://www.R-project.org/ (2017).
41. Ezard, T., Fujisawa, T. & Barraclough, T. splits: SPecies' LImits by Threshold Statistics. *R package version 1.0-19/r52* https://R-Forge.R-project.org/projects/splits/ (2017).
42. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526-528 (2018).
43. Bermingham, E., McCafferty, S. S., Martin, A. P. Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. Molecular systematics of fishes (eds. Kocher, T. D., Stepien, C. A.). *Academic Press* 113-128 (San Diego, 1997).
44. Thomaz, A. T, Malabarba, L. R., Bonatto, S. L. & Knowles, L. L. Testing the effect of palaeodrainages versus habitat stability on genetic divergence in riverine systems: study of a Neotropical fish of the Brazilian coastal Atlantic Forest. *Journal of Biogeography* **42**, 2389-2401; 10.1111/jbi.12597 (2015).
45. Kimura, M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111-120 (1980).
46. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* **35**, 1547-1549; 10.1093/molbev/msy096 (2018).
47. Librado, P. & Rozas, J. DNAsp5. A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452; 10.1093/bioinformatics/btp187 (2009).
48. Bandelt, H. J., Forster, P. & Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**(1), 37-48 (1999).
49. Leigh, J. W. & Bryant, D. POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6**, 1110-1116; 10.1111/2041-210X.12410 (2015)

50. Guillot, G., Renaud, S., Ledevin, R., Michaux, J. & Claude, J. A Unifying Model for the Analysis of Phenotypic, Genetic and Geographic Data. *Systematic Biology* **61**(6), 897-911; 10.1093/sysbio/sys038 (2012).
51. Excoffier, L., Laval, G. & Schneider, S. Arlequin: a software for population data analysis. Version 3.1 http://cmpg.unibe.ch/software/arlequin3 (2007).
52. Wright, S. Evolution and the genetics of populations: variability within and among natural populations. *The University of Chicago* **4**, 580 (1978).
53. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-95 (1989).
54. Fu, Y. X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915-925 (1997).
55. Austin, M. P. Continuum concept, ordination methods, and niche theory. *Annual review of ecology and systematics* **16**(1), 39-61; 10.1146/annurev.es.16.110185.000351 (1985).
56. Graham, A., Atkinson, P. & Danson, F. Spatial analysis for epidemiology. *Acta Tropica* **91** (3), 219-225; 10.1016/j.actatropica.2004.05.001 (2004).
57. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**(3-4), 231-259; 10.1016/j.ecolmodel.2005.03.026 (2006).
58. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**(15), 1965-1978; 10.1002/joc.1276 (2005).
59. Jézéquel, C. *et al.* A database of freshwater fish species of the Amazon Basin. *Scientific Data* **7**, 96; 10.1038/s41597-020-0436-4 (2020).
60. Lovejoy, N. R. & Araujo, M. L. G. Molecular systematics, biogeography and population structure of Neotropical freshwater needlefishes of the genus Potamorrhaphis. *Molecular Ecology* **9**(3), 259-268; 10.1046/j.1365-294x.2000.00845.x (2000)
61. Blaxter, M. L. The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences* **359**(1444), 669-679; 10.1098/rstb.2003.1447 (2004).
62. Nwani, C. D. *et al*. DNA barcoding discriminates freshwater fishes from southeastern Nigeria and provides river system-level phylogeographic resolution within some species. *Mitochondrial DNA* **22**(1), 43-51; 10.3109/19401736.2010.536537 (2011).
63. Aguirre, W. E., Shervette, V. R., Navarrete, R., Calle, P. & Agorastos, S. Morphological and Genetic Divergence of Hoplias microlepis (Characiformes: Erythrinidae) in Rivers and Artificial Impoundments of Western Ecuador. *Copeia* **2013**(2), 312-323; 10.1643/ci-12-083 (2013).
64. Souza, F. H. S. *et al*. Interspecific Genetic Differences and Historical Demography in South American Arowanas (Osteoglossiformes, Osteoglossidae, Osteoglossum). *Genes* **10**(9), 693; 10.3390/genes10090693 (2019).
65. Torati, L. S. *et al*. Genetic diversity and structure in Arapaima gigas populations from Amazon and Araguaia-Tocantins river basins. *BMC Genetics* **20**(1); 10.1186/s12863-018-0711-y (2019)
66. Albert, J. S., Val, P. & Hoorn, C. The changing course of the Amazon River in the Neogene: center stage for Neotropical diversification. *Neotropical Ichthyology* **16**(3), e180033; 10.1590/1982-0224-20180033 (2018).
67. Silva, W. C., Marceniuk, A. P., Sales, J. B. L., & Araripe, J. Early Pleistocene lineages of Bagre bagre (Linnaeus, 1766) (Siluriformes: Ariidae), from the Atlantic coast of

South America, with insights into the demography and biogeography of the species. *Neotropical Ichthyology* **14**(2), 10.1590/1982-0224-20150184 (2016).
68. Haffer, J. & Prance, G. T. "Impulsos climáticos da evolução na Amazônia durante o Cenozóico: sobre a teoria dos Refúgios da diferenciação biótica". *Estudos Avançados USP* **46**, 175-208; 10.1590/S0103-40142002000300014 (2002).
69. Riker, S. R. L., Lima, F. J. C., Motta, M. B. Evidências de glaciação Pleistocênica na Amazônia Brasileira. *Anais do 14° Simpósio de Geologia da Amazônia*, *Sociedade Brasileira de Geologia* 15-18. (Marabá 2015).
70. Lemopoulos, A. & Covain, R. Biogeography of the freshwater fishes of the Guianas using a partitioned parsimony analysis of endemicity with reappraisal of ecoregional boundaries. *Cladistics* **35**(2019), 106-124; 10.1111/cla.12341 (2018).
71. Hoorn, C. Marine incursions and the influence of Andean tectonics on the Miocene depositional history of northwestern Amazonia: results of a palynostratigraphic study. *Palaeogeography, Palaeoclimatology and Palaeoecology* **105**, 267-309; 10.1016/0031-0182(93)90087-Y (1993).
72. Hoorn, C., Guerreiro, J. & Sarmiento, G. Andean tectonics as a cause for changing drainage patterns in Miocene Northern South America. *Geology* **23**(3), 237-240; 10.1130/0091-7613(1995)023<0237:ATAACF>2.3.CO;2 (1995).
73. Lundberg, J. G. *et al*. The stage for Neotropical fish diversification: a history of tropical South American rivers. (eds. Malabarba, L. R., Reis, R. E., Vari, R. P., Lucena, Z. M., Lucena, C. A. S. Phylogeny and classification of Neotropical fishes). *Edipucrs* 13-48 (Porto Alegre 1998).
74. Hubert, N., Renno, J. F. Historical biogeography of South American freshwater fishes. *Journal of Biogeography* **33**(8), 1414-1436; 10.1111/j.1365-2699.2006.01518.x (2006).
75. Ribeiro, A. C. Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: an example of faunal evolution associated with a divergente continental margin. *Neotropical Ichthyology* **4**(2), 225-46; 10.1590/S1679-62252006000200009 (2006).
76. Lovejoy, N. R., Albert, J. S. & Crampton, W. G. R. Miocene marine incursions and marine/freshwater transitions: evidence from Neotropical fishes. *Journal of South American Earth Sciences* **21**(1-2), 5-13; 10.1016/j.jsames.2005.07.009 (2006).

Acknowledgements

We tank to Dr. Gabriel Iketani for his help in sequencing and to everyone who helped with field samples and laboratory experiments. To Dr. C.C. Arantes for the samples of the lower Amazon floodplain. We thank the iXINGU Project (NSF DEB-1257813). We are grateful to T.M.A. Lima for help in figures prepared. KLAG received Mastership from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES grant # 88882.457158/2019-01). DJS thanks Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq research fellowship #311492/2017-7). LRRR was granted by CNPq and FAPEAM by the project INCT-Adapta II.

Supplementary information

Supplemental Information S1.
Supplemental Information S2.
Supplemental Information S3.

Síntese Integradora

Durante séculos, explorar a ictiofauna Neotropical e os padrões de distribuição geográfica de peixes tem sido alvo de estudos de muitos cientistas. Apesar disso, muitas bacias hidrográficas continuam produzindo novas descobertas e, certamente, a bacia Amazônica se destaca como um dos maiores desafios a serem enfrentados, pois estamos longe de catalogar completamente a surpreendente diversidade de peixes presente nesta região.

A compreensão sobre a taxonomia e biogeografia de peixes da Amazônia evoluiu ao longo dos anos e novas ferramentas e abordagens podem ser adotadas. A utilização da taxonomia integrativa na descoberta de novas espécies tem tido uma importante contribuição no conhecimento sobre a real identidade taxonômica de espécies.

Este estudo revelou novas espécies, novos registros, distribuições disjuntas, presença de espécies crípticas, preencheu lacunas geográficas e apresentou novos *insights* sobre a microevolução de *Hoplias* gr. *malabaricus*. Esta contribuição agregada à grande quantidade de dados já disponíveis sobre *Hoplias* certamente será de grande valor para outros ictiólogos e futuras investigações. Além disso, fornece um bom exemplo do poder da taxonomia integrativa na aceleração do conhecimento ictiológico.

Considerando que a bacia Amazônica possui a mais diversificada ictiofauna do mundo e que os vários sistemas hídricos que a compõem estão ameaçados por diversos estressores antropogênicos que colocam em risco a viabilidade de muitas espécies, a continuidade de estudos desta natureza poderá melhorar nossa capacidade de compreender e conservar a biota mais diversa do mundo.