Molecular analysis reveals high diversity in the *Hoplias malabaricus* (Characiformes, Erythrinidae) species complex from different Amazonian localities

Alex M. V. Ferreira¹*, Leila Braga Ribeiro², Eliana Feldberg¹

¹Instituto Nacional de Pesquisas da Amazônia, Programa de Pós-graduação em Genética, Conservação e Biologia Evolutiva, Manaus, Amazonas, Brazil
²Universidade Federal de Roraima, Centro de Ciências da Saúde, Av. Capitão Ene Garecz 2413, CEP 69310-000, Boa Vista, Roraima, Brazil
* Corresponding author: alex.matheus.viana@gmail.com; https://orcid.org/0000-0003-3969-177X

ABSTRACT

DNA barcoding proposes that a fragment of DNA can be used to identify species. In fish, a fragment of cytochrome oxidase subunit I (COI) has been effective in many studies with different foci. Here we use this molecular tool to provide new insights into the cryptic diversity found in the *Hoplias malabaricus* species complex. Popularly known as trahira, *H. malabaricus* is widely distributed in South America. The clade shows molecular and cytogenetic diversity, and several studies have supported the occurrence of a complex of species. We performed molecular and karyotypic analysis of *H. malabaricus* individuals from eight Amazonian localities to assess the diversity present in the nominal taxon, and to clarify relationships within this group. We used 12 samples in cytogenetic analyses and found two karyomorphs: 2n = 40 (20m + 20sm) (karyomorph C) and 2n = 42 (22m + 20sm) (karyomorph A). We used 19 samples in molecular analyses with COI as a molecular marker, maximum likelihood analyses, and the Kimura-2-parameter evolutionary model with bootstrap support. We found karyomorph-related differentiation with bootstrap of 100%. However, we found high molecular diversity within karyomorph C. The observed pattern allowed us to infer the presence of cryptic diversity, reinforcing the existence of a species complex.

KEYWORDS: trahira, DNA barcoding, karyotype, karyomorph

INTRODUCTION

The family Erythrinidae is composed of three genera widely distributed in South and Central-America: *Hoploerythrinus* (Gill 1896), *Erythrinus* (Scopoli 1777) and *Hoplias* (Gill 1903) (Nelson et al. 2016; Betancur et al. 2017). *Hoplias* is divided in three groups: *Hoplias macrophthalmus* (Pellegrin 1907), a group comprised of a single valid species; *Hoplias lacerdae* (Miranda Ribeiro 1908), a group currently including five valid species; and *Hoplias malabaricus* (Bloch 1794), a group still lacking a proper taxonomic approach (Reis et al. 2003; Oyakawa and Mattox 2009). *Hoplias malabaricus* is frequently
found in temporary lakes due to its capacity to survive at low oxygen concentrations and high water temperatures (Reis et al. 2003). It also represents one of the most complex problems in Neotropical fish taxonomy (Reis et al. 2003; Mattox et al. 2014). This species shows extensive karyotypic variation, with seven karyomorphs described so far (A-G), occurring in either allopatric and sympatric conditions. The karyomorphs differ in diploid number, karyotypic formula, and some exhibit sexual chromosomes (Bertollo et al. 2000). Such diversity in this species might be related to low gene flow among populations, since *H. malabaricus* has sedentary habits (Dergam et al. 2002; Blanco et al. 2010).

In recent decades, molecular and cytogenetic analyses have formed the basis of useful new interpretations of Neotropical fish taxonomy and phylogeny (Albert and Reis 2011; Pereira et al. 2013). Among the molecular markers used for species differentiation, DNA barcoding is one of the most popular and is widely used in species classification (Ardura et al. 2010). This approach uses a DNA fragment from cytochrome oxidase subunit I (COI), which codes part of an enzyme from the electron transfer chain, as a molecular tool for either identification of species already described or to provide supportive evidence for new species (Hebert et al. 2003). This genetic tool can be used for a variety of purposes, from forensic analyses and studies for species conservation (Ardura et al. 2010; Carvalho et al. 2015; Shen et al. 2016; Barman et al. 2018) to taxonomic studies (de Borba et al. 2019). Several authors have also shown the efficiency of DNA barcoding for fish species identification (Benzaquem et al. 2015; Ferreira et al. 2017; Barendse et al. 2019).

We performed cytogenetic and molecular analyses on samples of *H. malabaricus* from eight Amazonian localities to contribute to the understanding of the relationships within this taxon, and test the efficiency of DNA barcoding for identification of this group.

**MATERIAL AND METHODS**

We collected 19 individuals of *H. malabaricus* in eight localities (Figure 1; Table 1). We also collected two individuals of *Erythrinus erythrinus* (Bloch & Schneider, 1801) and one individual of *Hoplerythrinus unitaeniatus* (Spix & Agassiz, 1829) for use as outgroup in the analyses (Table 1). Individuals were euthanized, and had their kidneys and muscle samples removed. Collection and animal handling occurred between 1995 and 2004 under SISBIO permanent license no. 28095-1 issued by Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis – IBAMA. Tissue samples were stored at the collection of genetic resources – INPA and the specimens were deposited in the Ichthyological Collection of Instituto Nacional de Pesquisas da Amazônia (INPA-ICT 059630, INPA-ICT 059631, INPA-ICT 059632).

**Cytogenetic analysis**

Kidney samples were used for chromosomal preparations, according to the protocol of Bertollo et al. (1978). We classified chromosomes as metacentric, submetacentric or acrocentric, following Levan et al. (1964). Karyomorphs were separated following Bertollo et al. (2000).
For molecular analyses, we used 19 individuals of *H. malabaricus*, and two *E. erythrinus* and one *H. unitaeniatus* as outgroup. Total DNA was extracted using the Wizard Extraction Kit (Promega) according to the manufacturer’s recommendations, and quantified on a NanoVue Plus spectrophotometer (GE Healthcare). COI amplification was performed in a thermal cycler (T100™ Thermal Cycler), using the following primers: VR1 5’ CAG GAA ACA GCT ATG ACT AGA CTT CTG GGT GGC CAA AGA ATC A 3’ and VF1 5’ TGT AAA ACG CCG CAT TGG CT CAA CCA CAA CAT TGG 3’ (Ivanova et al. 2007). Amplification was conducted using 35 cycles with the following profile: 92 °C for 30 sec, 56 °C for 30 sec and 72 °C for 1 min 30 sec.

DNA sequencing

Polymerase Chain Reaction (PCR) products were visualized on a 1.7% agarose gel, and purified with 20% polyethylene glycol (PEG) (Lis 1980). For sequencing, a Big Dye Sequence Terminating kit (Applied Biosystems) was used, according to manufacturer instructions. Amplification conditions comprised 25 cycles at 96 °C for 30 sec; 15 sec at 50 °C; and 4 min at 60 °C. Subsequently, reaction products were precipitated and sequenced (sequencer model ABI PRISM 3100 Genetic Analyzer from Applied Biosystems/made by HITACHI).

Barcoding analysis

We used 22 COI sequences with 600 bp to perform the molecular analyses, which were aligned using the GeneiousR 10.1.3 software. Kimura-2 distance model parameters (Kimura 1980) were used to build a Maximum Likelihood (ML) tree, created to provide a graphic representation of the relationships between individuals and clusters using GeneiousR 10.1.3 software, and a bootstrap analysis was performed (Felsenstein 1985) with 1,000 replicates. We used the other two Erythrinidae genera (*Erythrinus* and *Hoploerythrinus*) as an outgroup. All sequences obtained in this study are available on GenBank (https://www.ncbi.nlm.nih.gov/genbank/), accession numbers MT076970-MT076984. In addition, we used sequences from the BOLD Systems (https://www.boldsystems.org/) to build a ML tree to produce a more robust interpretation of our results (accession numbers BFFDF148-19; BSB280-10 - BSB492-10; CYTC4372-12; DSFRE175-08 - DSFRE178-08; FUPR169-09 - FUPR171-09; GBMNB1689-20 - GBMNB12169-20; HPRB006-16 - HPRB024-16; ITAPE092-15 - ITAPE131-15; JEQUI041-12 - JEQUI210-13; LARI001-12 - LARI212-12; MUCU030-13 - MUCU086-13; RDOCE022-13 - RDOCE099-13; RENA032-16 - RENA039-16; UDEA133-18).

RESULTS

The COI fragment contained 600 bp, and the Blast analysis indicated that most samples were correctly identified as *H. malabaricus*. All individuals belonging to *H. malabaricus* formed a group with 100% bootstrap support, but we also found four different linages, which were classified as Evolutionary Significant Units (ESUs). ESU-A: samples from Barcelos (Jaradi Stream) + Arraia Lake (corresponding to karyomorph B); ESU-B: samples from Barrentinho Stream + Negro River + Capivara Lake and Catalão Lake; ESU-C: Capivara lake + Catalão Lake and Anavilhanas; and ESU-D: samples from the Amazonas River. Noticeably, we found two different ESUs (B and C) occurring in the same localities (Capivara Lake and Catalão Lake). In addition, the other species of Erythrinidae demonstrated robust bootstrap values. Our two *E. erythrinus* outgroup samples formed a clade with 100% bootstrap support, and the analysis supported a clade formed by *E. erythrinus* + *H. unitaeniatus* with 100% support (Figure 2).

The complementary analysis with BOLD sequences showed that our samples from Jaradi Stream and Arraia Lake (ESU A) shared a cluster with samples from Negro River, Barrentinho Stream (ESU B), and sequences from Caracas (Venezuela). Samples from ESU C shared a cluster with sequences
from Minas Gerais (Brazil), and Buenos Aires (Argentina). Additionally, this cluster was close to sequences from Maranhão (Brazil). ESU D shared a cluster with sequences from Paraná (Brazil) (Supplementary Material, Figure S1).

We only obtained cytogenetic results for 12 of the 19 collected individuals of *Hoplias malabaricus*. Cytogenetically analyzed individuals showed two karyotypes, karyomorph A (2n = 42, 22m + 20sm) and karyomorph C (2n = 40, 20m + 20sm) (Figure 3), both of which have already been described for the species. Karyomorph A was found in samples from Jaradi stream and Arraia Lake (ESU A), while karyomorph C was found in samples from Catalão Lake, Barrentinho Stream, Anavilhanas, and Amazon River (ESU B, C and D). We did not record any sex-related differentiation at the karyotypic level.

**Figure 2.** Phylogenetic relationships of COI haplotypes derived from *Hoplias malabaricus* karyomorphs and its outgroups. The topology was obtained with maximum likelihood analysis. Bootstrap values equal 100% for all clusters.

**Figure 3.** Karyotypes of *Hoplias malabaricus* using Giemsa stain showing karyomorph A (A), and karyomorph C (B).
DISCUSSION

Despite having its efficiency questioned (Blaxter 2004), DNA barcoding has demonstrated its efficiency in resolving a variety of problems related to identification of species and species complexes (Hubert et al. 2016; Ferreira et al. 2017). We confirmed the efficiency of this molecular tool, which revealed identification mistakes made by the use of morphological characteristics alone, showing the potential of DNA barcoding to confirm field identifications, detect misidentifications and uncover cryptic diversity. Furthermore, we observed a division between populations and karyomorphs (2n = 40 and 2n = 42), suggesting that karyotypic variation in *Hoplias malabaricus* is a consequence of different chromosomal rearrangements occurring independently in different areas of the distribution range of the species. Guimarães et al. (2017) compared two populations of karyomorph A, and found differences related to centromere position, number and location of nucleolar organizer regions (NORs), and organization patterns of heterochromatic regions, demonstrating that evolutionary divergence has occurred within these groups to some degree. This karyotypic pattern is supported by the sedentary habit of *H. malabaricus*, which may result in restricted gene flow (Dergam et al. 2002; Blanco et al. 2010).

Our results show low molecular differentiation between ESU A, ESU B and ESU C, which indicates that *H. malabaricus* karyomorphs A and C may hybridize, although no intermediate karyomorph has ever been found (e.g. 2n = 41), even in the areas where both A and C are found in sympathy (Bertollo et al. 2000). However, we did not find karyomorph F, which is similar to karyomorph C, except for the presence of a metacentric pair, the first in the karyotype (Bertollo et al. 2000). Furthermore, variations can be found in the same karyomorph, as reported for karyomorph A from the upper Paraná River basin (Blanco et al. 2010), and karyomorphs A and C from the Central Amazon (Guimarães et al. 2017). Taken together, all previous evidence indicates that this species group is characterized by very high levels of cytogenetic and molecular variation.

Such diversity has also been shown at the molecular level. Recently, using the COI gene, Cardoso et al. (2018) revealed the existence of high levels of diversity within members of the genus *Hoplias* collected in South America, with 15 fully supported mitochondrial lineages within what was once considered to be a single species, *H. malabaricus*, with a continental-wide distribution. Only four of these lineages are currently described as valid species, leaving 11 mitochondrial lineages still hidden within the species complex. The same study highlighted the existence of different lineages within the *H. lacerdae* group, hinting at the complexity within *Hoplias*. Also based on COI analyses, Marques et al. (2013) found six haplogroups for populations of *H. malabaricus* from seven locations in Pará state (Brazil), yet all individuals had the karyomorph C (2n = 40). In a study with *H. malabaricus* from populations in eastern Brazil, Santos et al. (2009) found genetic differentiation using ATPase6 as molecular marker, two karyomorphs [A (2n = 42) and F (2n = 40)], and six haplogroups.

Barcode DNA has been used to reveal cryptic diversity in other fish groups. For example, Benzaquem et al. (2015) found high levels of divergence in the genus *Nannostomus*, which suggests the presence of cryptic species. Likewise, barcoding and cytogenetic data indicated great divergence among Ancistrus species from the Amazonas and Paraguay basins, and the molecular tool efficiently discriminated Ancistrus lineages, indicating the occurrence of cryptic species in these regions (de Borba et al. 2019). In the current study, we found great differentiation between samples from the Amazonas River and the other sampled localities, as well as the segregation between karyomorphs A and C, suggesting that different processes are acting on the populations of Amazonian trahira, which can lead to full speciation. In view of this variation within karyomorphs, the karyotype classification previously proposed by Betorllo et al. (2000) is insufficient for describing phylogenetic relationships within this species complex.

CONCLUSIONS

The pattern in our cytogenetics and molecular data indicates the presence of distinct lineages of *Hoplias malabaricus* in the Amazonas basin and support the existence of a species complex. The use of nuclear markers would be needed to clarify the processes related to *H. malabaricus* population structure and dynamics. In addition, comparative analyses with of other regions would also be necessary to understand the evolutionary history and biogeographic patterns of this group.-

ACKNOWLEDGMENTS

The authors are grateful to the Brazilian institutions, Instituto Nacional de Pesquisas da Amazônia (INPA), Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). AMVF was a student of Programa de Apoio à Iniciação Científica do Amazonas (PAIC) and received a fellowship from FAPEAM. The authors also thank Dr. Jorge Ivan Rebelo Porto for collecting the fish and discussing the data and special thanks to Arlindo Batista for all the support provided in the laboratory.

REFERENCES


FERREIRA et al. Molecular diversity in the *H. malabaricus* complex

SUPPLEMENTARY MATERIAL (only available in the electronic version)

Ferreira et al. Molecular analysis reveals high diversity in the *Hoplias malabaricus* (Characiformes, Erythrinidae) species complex from different Amazonian localities.

*Figure S1.* Phylogenetic relationships of COI haplotypes derived from *Hoplias malabaricus* karyomorphs from the Amazon region, its outgroups and BOLD sequences. The topology was obtained with maximum likelihood analysis. Bootstrap values are indicated at the nodes.