



Mitochondrial DNA Part A

DNA Mapping, Sequencing, and Analysis

ISSN: 2470-1394 (Print) 2470-1408 (Online) Journal homepage: <https://www.tandfonline.com/loi/imdn21>

Molecular identification of the fish fauna from the pantanal flood plain area in Brazil

Cristiane Kioko Shimabukuro-Dias, Guilherme José da Costa Silva, Fernando Yuldi Ashikaga, Fausto Foresti & Claudio Oliveira

To cite this article: Cristiane Kioko Shimabukuro-Dias, Guilherme José da Costa Silva, Fernando Yuldi Ashikaga, Fausto Foresti & Claudio Oliveira (2017) Molecular identification of the fish fauna from the pantanal flood plain area in Brazil, Mitochondrial DNA Part A, 28:4, 588-592, DOI: [10.3109/24701394.2016.1149826](https://doi.org/10.3109/24701394.2016.1149826)

To link to this article: <https://doi.org/10.3109/24701394.2016.1149826>



Published online: 14 Mar 2016.



Submit your article to this journal [↗](#)



Article views: 55



View Crossmark data [↗](#)

SHORT COMMUNICATION

Molecular identification of the fish fauna from the pantanal flood plain area in Brazil

Cristiane Kioko Shimabukuro-Dias, Guilherme José da Costa Silva, Fernando Yuldi Ashikaga, Fausto Foresti and Claudio Oliveira

Departamento De Morfologia, Laboratório De Biologia E Genética De Peixes, Universidade Estadual Paulista, Botucatu, São Paulo, Brazil

ABSTRACT

The Pantanal is under the influence of the Paraguay River flood regime is considered to be one of the largest wetlands of the world, and has rich biodiversity, including fishes. Until now, the identification of fish species in this biome has only considered the morphological characteristics of individuals, and the present work aimed to identify the fish species of the Pantanal region through the DNA barcode methodology for investigating the biodiversity in this region. The genetic analysis of 638 samples via the GMYC approach identified 137 operational taxonomic units (OTUs) belonging to 127 species that have previously been described according to their morphological characteristics. Data suggest that 10 cases of morphospecies (*Eigenmannia trilineata*, *E. virescens*, *Pimelodella gracilis*, *Brachyhypopomus pinnicaudatus*, *Brachyhypopomus* sp., *Ancistrus* sp., *Hyphessobrycon eques*, *Jupiaba acanthogaster*, and *Serrapinnus calliurus*) represent complexes of cryptic species, and the number of species described in the Pantanal region has thus potentially been underestimated.

ARTICLE HISTORY

Received 24 June 2015
Revised 14 January 2016
Accepted 31 January 2016
Published online 9 March 2016

KEYWORDS

Cryptic species; cytochrome oxidase; DNA sequencing; molecular identification; sympatric

Introduction

The Pantanal is considered to be the largest wetland area in the world, covers areas in three countries (i.e. Brazil, Bolivia, and Paraguay) and is characterized by a rich diversity of species and the seasonal flooding caused by the Paraguay River (Alho et al. 1988; Harris et al. 2005; Britski et al. 2007).

This great biodiversity is highly dependent on the rainy seasons because among different species that compose the fauna, fishes depend most heavily on the flood periods (Alho & Vieira 1997; Bailly et al. 2008). However, this biome is currently suffering the effects of human population expansion (Agostinho et al. 2005; Abell et al. 2008).

Whereas species identification is an important step for the “*in situ*” conservation of biodiversity (Agostinho et al. 2005), with the advent of new technologies, molecular genetics has increased in importance in this area and brought new tools for determining taxonomies and elucidating many questions. Thus, Hebert et al. (2003) proposed the DNA barcoding methodology, which utilizes a short segment of 648 nucleotides of the Cytochrome Oxidase C subunit I (COI) gene from the mitochondrial DNA for the molecular identification of species.

Thus, the aim of this work was to characterize the biodiversity of the fishes in the Pantanal (Brazil) through the DNA barcode methodology using the General Mixed Yule

Coalescent (GMYC) approach as a means of improving the estimate of the number of different operational taxonomic units (OTUs) in this area. We also sought to identify the cases in which discontinuities of genetic patterns were present among morphologically indistinguishable individuals captured in sympatry to enable a better estimate of the specific abundance in this region.

Material and methods

Sampling

For this study, 638 specimens belonging to 127 distinct morphospecies were analyzed (Table 1) and all voucher are deposited in the *Laboratório de Biologia de Peixes (LBP)*, from *Universidade Estadual Paulista “Julio de Mesquita Filho” – UNESP – Botucatu – Brazil*. The morphological identifications to the species level were performed by experts using the morphological keys described by Britski et al. (2007), but in some cases, it was only possible to extend the identifications to the genus level. A portion of the vouchers analyzed here (75%) is coincident to that sampled on the Pantanal National Park published by Polaz et al. (2014) (Figure 1). Tissue samples were extracted, fixed in ethanol 96%, and then stored at -20°C . The specimens were fixed in formalin solution (10%) and preserved in ethanol (70%).

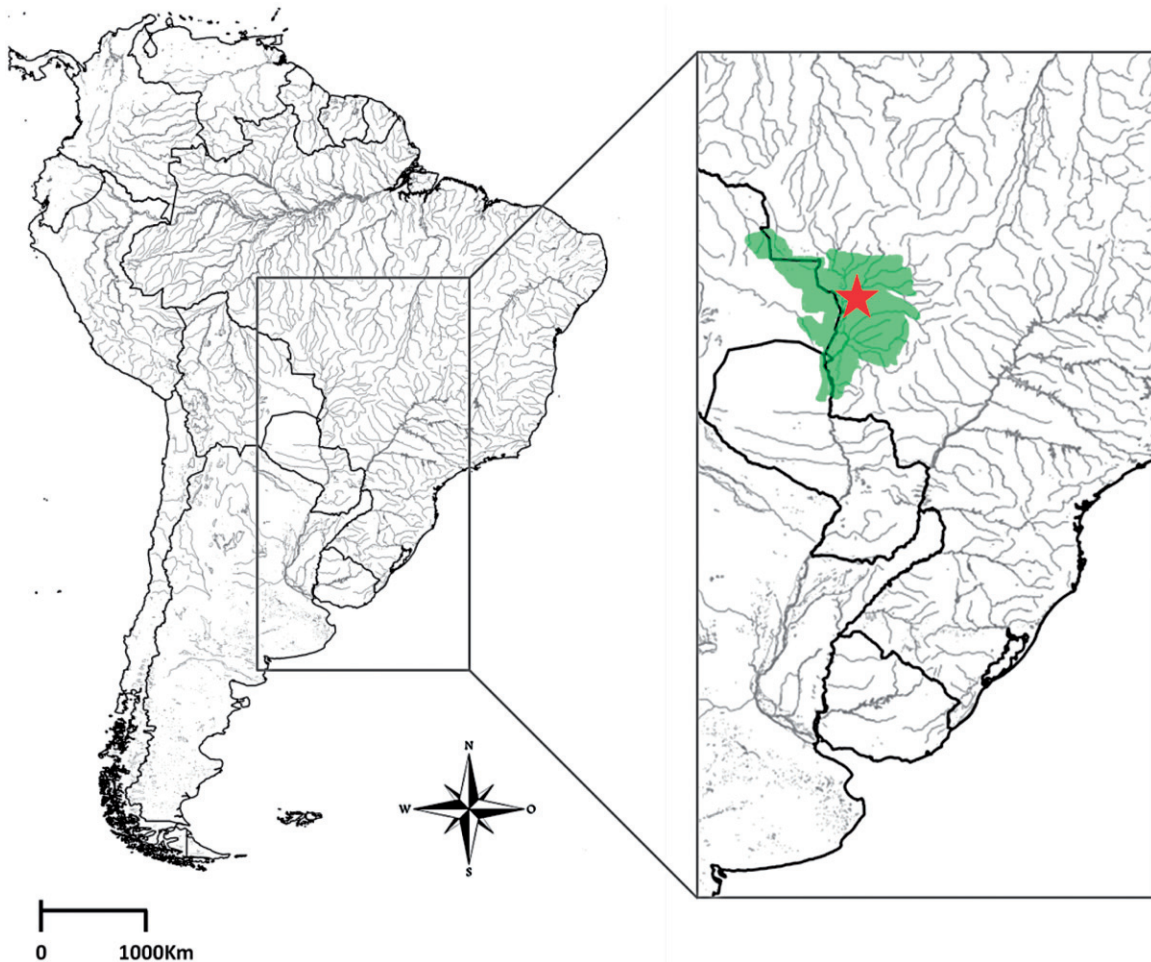


Figure 1. Coverage area of the Pantanal biome (in green), drained by the Paraguay River Basin. In detail (red star) is the location of the Parque Nacional do Pantanal Matogrossense – Brazil, the main point of sampling.

DNA extraction and sequencing

The genomic DNA was extracted using DNeasy Tissue Kits (Qiagen, Valencia, CA). The primers used for amplification and sequencing were the *Fish F1* and *Fish R1* (Ward et al. 2005). Thus, the PCR products were sequenced in both directions using the Big DyeTM Terminator v 3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Waltham, MA), purified again by ethanol precipitation and loaded on an automatic sequencer 3130-Genetic Analyzer (Applied Biosystems, Waltham, MA).

Sequence analysis

The consensus sequences were obtained from the forward and reverse strands using GENEIOUS PRO 5.4.2 (Drummond et al. 2010). The alignments were generated in MUSCLE (Edgar 2004). The presence of stop codons was determined using GENEIOUS PRO 5.4.2.

MEGA 6.0 (Tamura et al. 2013) was used for the detection of nucleotide variations, substitution patterns, and genetic distances. To evaluate the occurrence of substitution saturation, we estimated the *I*_{ss} index in DAMBE 5.2.31 (Xia 2001). This software was also used to evaluate the rate of transitions/transversions. Finally, MODELTEST 3.7 (Posada & Crandall 1998)

was used to evaluate the best nucleotide evolution models for the *COI* genes.

Molecular species definition

We used the general mixed Yule coalescent (GMYC) model (Pons et al. 2006) for molecular species identification. We generated the ultrametric on Beast 1.5.1 (Rambaut et al. 2014) with the lognormal relaxed clock not calibrated and birth and death model as the tree prior and the general time reversible (GTR) nucleotide evolution model.

The Bayesian topology reconstruction was performed with a Monte Carlo Markov chain (MCMC) procedure with 10⁷ generations in which one tree per 1000 generations was sampled, and we used a UPGMA tree to start the search. In the program TRACER (Rambaut et al. 2014), we checked the stationary phase in the log-likelihood score distribution to calculate the best point for burnin. After burn-in, the tree with the best credibility nodes (ultrametric tree) was found via TreeAnnotator (Rambaut & Drummond 2012) analysis.

The ultrametric tree was examined in R 3.0.0 (R Development Core Team 2013) with the “splits” (species limits by threshold statistics) package and the single threshold method using standard parameters.

Results

All sequences were deposited to the BOLD database, belonging to the project entitled: Barcode Fishes From Paraguay River (BFPR). No pseudogenes, stop codons, insertions, and deletions were detected. After alignment and editing of the consensus sequences, we obtained a matrix with 549 base pairs with the following nucleotide composition: 30.5% thymine, 27.6% cytosine, 24.8% adenine, and 17.2% guanine. Among all analyzed bases, 263 were conserved and 286 were variable, and there were 12 singletons. The overall distance was 24.4% (i.e. the K2P distance).

The chosen tree was used in the GMYC analysis with a threshold time = -0.005705752 (or 0.2% of genetic divergence

following the topology of the ultrametric tree), which revealed the presence of 137 independent evolutionary units (confidence interval 136–142; Figure 2). The genetic divergences were estimated at three hierarchical levels of analysis as follows: (1) the average genetic divergence between the individuals within the same OTU was 0.3% with a range of 0–0.9%; (2) the average genetic divergence among individuals within a morphospecies was 0.6% with a range of 0–11.9%; and (3) the minimum genetic divergence found between different morphospecies was 1.8%.

Among the 127 analyzed morphospecies, the identities of 118 were confirmed by DNA barcode analysis, and the variation in the number of OTU was probably due to the haplotype differentiations between individuals within the same group.



Figure 2. Ultrametric tree of the fishes from Pantanal flood plain obtained with the COI dataset. The name of the taxa was defined by traditional morphological identification. In red are shown the morphospecies with more than one Operational Taxonomic Units (OTUs) in sympatry.

This situation was observed in morphospecies (i.e. *Eigenmannia trilineata*, *E. virescens*, *Pimelodella gracilis*, *Brachyhypopomus pinnicaudatus*, *Brachyhypopomus* sp., *Ancistrus* sp., *Hyphessobrycon eques*, *Jupiaba acanthogaster*, and *Serrapinnus calliurus*) that exhibited high levels of haplotypic differentiation among individuals within each group (exceeding 1.8%). Two genetically divergent groups occurred in the same location but had been identified as a single entity according to morphological characters, as frequently occurs in cases of cryptic species in sympatry (Kekkonen & Hebert 2014).

Discussion

Based on comparisons of the sequences of the fishes deposited in the Barcode of Life Data Systems (BOLD), Ward (2009) suggested that entities with genetic divergence greater than 2% are probably different species. However, this estimate may be overestimated, since samples were used from different parts of the world and not necessarily of phylogenetically related groups. In the present study, the GMYC analysis revealed values ranging between 0.2% and 1.8%, which corresponded to the lower value of divergence between the different OTUs and the greater value of divergence between the entities within the same OTU, respectively. These values may have been approximately ten times smaller than the values suggested by Ward (2009), but these results are similar to those found by Pereira et al. (2013) in a study in which many of the species were separated from the nearest congeners by less than 2% genetic divergence. Although one of the putative factors that influences the high number of species of neotropical regions is the high speciation rate (Theory of "Cradle"; Albert & Reis 2011), the lower interspecific genetic divergence observed among neotropical fishes could be related to recent high levels of diversification of the species of these regions.

Among the fishes included in this study, 127 species were morphologically identified; however, using the genetic distance threshold of 2% suggested by Ward (Ward 2009), species such as *Pimelodus argenteus* and *P. maculatus* were recognized as single entities. The GMYC analysis identified 137 OTUs (Figure 2), and nine morphospecies were formed by at least two OTUs. These data emphasize the differences in results that depend on the model used and revealed that the number of species present in the Pantanal region may be underestimated.

The occurrence of genetic differentiation among groups inside a morphospecies is not a safe argument for defining two species. Such genetic differentiation can be only due to spatial separation between groups, and this does not mean that the entities are incompatible and represent distinct species according to the biological species concept (Mayr 1940). In contrast, Kekkonen and Hebert (2014) argued that when two groups are in sympatry, the better explanation for deep genetic differentiation is reproductive isolation. Thus, findings of different OTUs in the absence of morphological distinctions in conditions of sympatry likely indicate that we are dealing with a species complex.

In the present data, nine cases involving morphologically indistinguishable groups exhibited two or more OTUs that occurred in sympatry, which indicates that approximately 7% of the species corresponded to species complexes. If we

consider that in the region of the Pantanal National Park (i.e. the main sampling area of the present study), 182 morphospecies species have been reported (Polaz et al. 2014), the real biodiversity is likely greater, even more so if we consider the diversity of the 269 species of fishes that have been described by for the Pantanal area (Britski et al. 2007).

Acknowledgements

The authors are grateful to Renato Devidé, Bruno F. Melo, and Ricardo Britski for their help with sampling and specimens identification.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding information

This research was supported by the Brazilian agencies FAPESP (Fundação de Amparo à pesquisa do Estado de São Paulo) e CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas) e CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

References

- Abell R, Thieme ML, Revenga C, Bryer M, Kottelat M, Bogutskaya N, Coad B, Mandrak N, Balderas SC, Bussing W, et al. 2008. Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. *BioScience*. 58:403–414.
- Agostinho AA, Thomaz SM, Gomes LC. 2005. Conservation of the biodiversity of Brazil's Inland waters. *Conserv Biol*. 19:646–652.
- Albert JS, Reis RE. 2011. Historical biogeography of neotropical freshwater fishes. Historical biogeography of neotropical freshwater fishes. Los Angeles: University of California Press.
- Alho C, Lacher TJ, Gonçalves H. 1988. Environmental degradation in the Pantanal ecosystem. *Bioscience*. 38:164–171.
- Alho CJR, Vieira LM. 1997. Fish and wildlife resources in the Pantanal wetlands of Brazil and potential disturbances from the release of environmental contaminants. *Environ Toxicol Chem*. 16:71–74.
- Bailly D, Agostinho AA, Suzuki HI. 2008. Influence of the flood regime on the reproduction of fish species with different reproductive strategies in the Cuiabá River, Upper Pantanal, Brazil. *River Res Appl*. 24:1218–1229.
- Britski HA, Silimon K, Lopes B. 2007. Peixes do Pantanal: manual de identificação. Brasília: Embrapa.
- Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C., Field M, Heled J, Kearse M, Markowitz S, et al. 2010. Geneious v5.5.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32:1792–1797.
- Hebert PDN, Cywinska A, Ball SL, de Waard JR. 2003. Biological identifications through DNA barcodes. *Proc Biol Sci*. 270:313–321.
- Herris MB, Tomas W, Mourao G, Da Silva CJ, Guimaraes E, Sonoda F, et al. 2005. Safeguarding the Pantanal wetlands: threats and conservation initiatives. *Conserv Biol*. 19:714–720.
- Kekkonen M, Hebert PDN. 2014. DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. *Mol Ecol Resour*. 14:706–715.
- Mayr E. 1940. Speciation phenomena in birds. *Am Nat*. 74:249–278.
- Pereira LHG, Hanner R, Foresti F, Oliveira C. 2013. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genet*. 14:20.
- Polaz CNM, Melo BF, Britzke R, Resende EK, Machado FA, Lima JAF, Petrere Jr M. 2014. Fishes from the Parque Nacional do Pantanal Matogrossense, upper Paraguai River basin, Brazil. *Check List*. 10:122–130.

- Pons J, Barraclough T, Gomez-Zurita J, Cardoso A, Duran D, Hazell S, Kamoun S, Sumlin WD, Vogler AP. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst Biol.* 55:595–609.
- Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics.* 14:817–818.
- R Development Core Team. 2013. R Development Core Team. *R: A Language and Environment for Statistical Computing.*
- Rambaut A, Drummond AJ. 2012. Tree Annotator, Version 1.7.4.
- Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v1.6.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol.* 30:2725–2729.
- Ward RD. 2009. DNA barcode divergence among species and genera of birds and fishes. *Mol Ecol Resour.* 9:1077–1085.
- Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish species. *Philos Trans R Soc Lond B: Biol Sci.* 360:1847–1857.
- Xia X. 2001. DAMBE: software package for data analysis in molecular biology and evolution. *J Hered.* 92:371–373.