GENETICS, EVOLUTION, AND CONSERVATION OF NEOTROPICAL FISHES

EDITED BY: Rodrigo A. Torres and Roberto Ferreira Artoni
PUBLISHED IN: Frontiers in Genetics

frontiers Research Topics



frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88963-335-7
DOI 10.3389/978-2-88963-335-7

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

GENETICS, EVOLUTION, AND CONSERVATION OF NEOTROPICAL FISHES

Topic Editors:
Rodrigo A. Torres, Federal University of Pernambuco, Brazil
Roberto Ferreira Artoni, Universidade Estadual de Ponta Grossa, Brazil

Fish represent the most ancestral and specious group of vertebrates, and occupy more diverse aquatic environments around the world. Ichthyofauna is extremely diverse, especially in megadiverse countries occupying biogeographical regions such as the Neotropical Region, which covers an extensive area between North and South America. Much of this biodiversity will be extinct, even before science knows any aspect of its biology. Like this, Neotropical fish genetics started in the end of the 70's with papers studying the chromosomes of Hoplias malabaricus (Family Erythrinidae) and the karyotype variation among three genera of the family Anostomidae. The topic at that time was concentrated in two Institutions from the state of São Paulo, Southeastern Brazil. In the middle 80's, the first Symposium on Neotropical Fish Cytogenetics was organized. Nowadays, the field of Neotropical Fish Genetics is present in Brazil, Colombia, Argentina, Uruguai, Venezula, Chile, and Equador, as well as outside South America in Panama, Mexico, USA, Canada, Czech Republic, Germany, and Spain. The research developed in cytogenetics has focused mainly on karyotype evolution and cytotaxonomy, chromosome structure and, more recently, cytogenomics. In relation to the use of molecular markers, support has been sought for the management of populations for conservation or production in captivity. In addition, many studies have been carried out with the aim of establishing supra-specific phylogenetic relationships and clarifying species distribution scenarios by phylogeographic modeling. The genome and transcriptome of some model species begin to emerge as extremely promising and informative areas for neotropical fish.

In 2017, the Neotropical fish genetics research community celebrates the 30th anniversary of its main Meeting (today entitled Symposium on Neotropical Fish Genetics and Cytogenetics). This Research Topic is part of this celebration and aims at reporting the state of the art and its current advances in the frontier of knowledge in genetics, evolution, and conservation of neotropical fish, as well as to detect the challenges to be overcome in the next years.

Citation: Torres, R. A., Artoni, R. F., eds. (2020). Genetics, Evolution, and Conservation of Neotropical Fishes. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88963-335-7

Table of Contents

07 ***Editorial: Genetics, Evolution, and Conservation of Neotropical Fishes***
Rodrigo Augusto Torres and Roberto Ferreira Artoni

10 ***A Glimpse Into the Satellite DNA Library in Characidae Fish (Teleostei, Characiformes)***
Ricardo Utsunomia, Francisco J. Ruiz-Ruano, Duílio M. Z. A. Silva, Érica A. Serrano, Ivana F. Rosa, Patrícia E. S. Scudeler, Diogo T. Hashimoto, Claudio Oliveira, Juan Pedro M. Camacho and Fausto Foresti

21 **Triportheus albus *Cope, 1872 in the Blackwater, Clearwater, and Whitewater of the Amazon: A Case of Phenotypic Plasticity?***
José D. A. Araújo, Andrea Ghelfi and Adalberto L. Val

33 ***The* Bunocephalus coracoideus *Species Complex (Siluriformes, Aspredinidae). Signs of a Speciation Process Through Chromosomal, Genetic and Ecological Diversity***
Milena Ferreira, Caroline Garcia, Daniele A. Matoso, Isac S. de Jesus, Marcelo de B. Cioffi, Luiz A. C. Bertollo, Jansen Zuanon and Eliana Feldberg

45 ***A Flow Cytometry Protocol to Estimate DNA Content in the Yellowtail Tetra* Astyanax altiparanae**
Pedro L. P. Xavier, José A. Senhorini, Matheus Pereira-Santos, Takafumi Fujimoto, Eduardo Shimoda, Luciano A. Silva, Silvio A. dos Santos and George S. Yasui

53 ***Revealing Hidden Diversity of the Underestimated Neotropical Ichthyofauna: DNA Barcoding in the Recently Described Genus* Megaleporinus *(Characiformes: Anostomidae)***
Jorge L. Ramirez, Jose L. Birindelli, Daniel C. Carvalho, Paulo R. A. M. Affonso, Paulo C. Venere, Hernán Ortega, Mauricio Carrillo-Avila, José A. Rodríguez-Pulido and Pedro M. Galetti Jr.

64 ***Low Genetic Diversity and Structuring of the* Arapaima *(Osteoglossiformes, Arapaimidae) Population of the Araguaia-Tocantins Basin***
Carla A. Vitorino, Fabrícia Nogueira, Issakar L. Souza, Juliana Araripe and Paulo C. Venere

74 ***Hidden Diversity in the Populations of the Armored Catfish* Ancistrus *Kner, 1854 (Loricariidae, Hypostominae) From the Paraná River Basin Revealed by Molecular and Cytogenetic Data***
Ana C. Prizon, Daniel P. Bruschi, Luciana A. Borin-Carvalho, Andréa Cius, Ligia M. Barbosa, Henrique B. Ruiz, Claudio H. Zawadzki, Alberto S. Fenocchio and Ana L. de Brito Portela-Castro

87 ***Headwater Capture Evidenced by Paleo-Rivers Reconstruction and Population Genetic Structure of the Armored Catfish (*Pareiorhaphis garbei*) in the Serra do Mar Mountains of Southeastern Brazil***
Sergio M. Q. Lima, Waldir M. Berbel-Filho, Thais F. P. Araújo, Henrique Lazzarotto, Andrey Tatarenkov and John C. Avise

95 ***Genetic Diversity of an Imperiled Neotropical Catfish and Recommendations for its Restoration***
Fernando S. Fonseca, Rodrigo R. Domingues, Eric M. Hallerman and Alexandre W. S. Hilsdorf

107 *First Chromosomal Analysis in Hepsetidae (Actinopterygii, Characiformes): Insights Into Relationship Between African and Neotropical Fish Groups*
Pedro C. Carvalho, Ezequiel A. de Oliveira, Luiz A. C. Bertollo, Cassia F. Yano, Claudio Oliveira, Eva Decru, Oladele I. Jegede, Terumi Hatanaka, Thomas Liehr, Ahmed B. H. Al-Rikabi and Marcelo de B. Cioffi

119 *Population Genetic Structure of* **Cnesterodon decemmaculatus** *(Poeciliidae): A Freshwater Look at the Pampa Biome in Southern South America*
Aline M. C. Ramos-Fregonezi, Luiz R. Malabarba and Nelson J. R. Fagundes

129 *Genetic Pattern and Demographic History of* **Salminus brasiliensis***: Population Expansion in the Pantanal Region During the Pleistocene*
Lívia A. de Carvalho Mondin, Carolina B. Machado, Emiko K. de Resende, Debora K. S. Marques and Pedro M. Galetti Jr.

137 *Microsatellites Associated With Growth Performance and Analysis of Resistance to* **Aeromonas hydrophila** *in Tambaqui* **Colossoma macropomum**
Raquel B. Ariede, Milena V. Freitas, Milene E. Hata, Vito A. Mastrochirico-Filho, Fabiana Pilarski, Sergio R. Batlouni, Fábio Porto-Foresti and Diogo T. Hashimoto

145 *Extensive Karyotype Reorganization in the Fish* **Gymnotus arapaima** *(Gymnotiformes, Gymnotidae) Highlighted by Zoo-FISH Analysis*
Milla de Andrade Machado, Julio C. Pieczarka, Fernando H. R. Silva, Patricia C. M. O'Brien, Malcolm A. Ferguson-Smith and Cleusa Y. Nagamachi

155 *Genetic Diversity of the Endangered Neotropical Cichlid Fish (***Gymnogeophagus setequedas***) in Brazil*
Lenice Souza-Shibatta, Thais Kotelok-Diniz, Dhiego G. Ferreira, Oscar A. Shibatta, Silvia H. Sofia, Lucileine de Assumpção, Suelen F. R. Pini, Sergio Makrakis and Maristela C. Makrakis

165 *Identification of a New Mullet Species Complex Based on an Integrative Molecular and Cytogenetic Investigation of* **Mugil hospes** *(Mugilidae: Mugiliformes)*
Mauro Nirchio, Fabilene G. Paim, Valentina Milana, Anna R. Rossi and Claudio Oliveira

174 *Gene and Blood Analysis Reveal That Transfer From Brackish Water to Freshwater is More Stressful to the Silverside* **Odontesthes humensis**
Tony L. R. Silveira, Gabriel B. Martins, William B. Domingues, Mariana H. Remião, Bruna F. Barreto, Ingrid M. Lessa, Lucas Santos, Danillo Pinhal, Odir A. Dellagostin, Fabiana K. Seixas, Tiago Collares, Ricardo B. Robaldo and Vinicius F. Campos

184 *Molecular Evidences of a Hidden Complex Scenario in* **Leporinus** *cf.* **friderici**
Rosane Silva-Santos, Jorge L. Ramirez, Pedro M. Galetti Jr. and Patrícia D. Freitas

193 *Genetic Variation of the Endangered Neotropical Catfish* **Steindachneridion scriptum** *(Siluriformes: Pimelodidae)*
Rômulo V. Paixão, Josiane Ribolli and Evoy Zaniboni-Filho

203 ***Genetic Characterization of the Fish* Piaractus brachypomus *by Microsatellites Derived From Transcriptome Sequencing***
Paulo H. Jorge, Vito A. Mastrochirico-Filho, Milene E. Hata, Natália J. Mendes, Raquel B. Ariede, Milena Vieira de Freitas, Manuel Vera, Fábio Porto-Foresti and Diogo T. Hashimoto

215 ***Genetic and Morphological Analyses Demonstrate That* Schizolecis guntheri *(Siluriformes: Loricariidae) is Likely to be a Species Complex***
Camila S. Souza, Guilherme J. Costa-Silva, Fábio F. Roxo, Fausto Foresti and Claudio Oliveira

224 ***Sex Chromosome Evolution and Genomic Divergence in the Fish* Hoplias malabaricus *(Characiformes, Erythrinidae)***
Alexandr Sember, Luiz A. C. Bertollo, Petr Ráb, Cassia F. Yano, Terumi Hatanaka, Ezequiel A. de Oliveira and Marcelo de Bello Cioffi

236 ***High-Throughput Sequencing Strategy for Microsatellite Genotyping Using Neotropical Fish as a Model***
Juliana S. M. Pimentel, Anderson O. Carmo, Izinara C. Rosse, Ana P. V. Martins, Sandra Ludwig, Susanne Facchin, Adriana H. Pereira, Pedro F. P. Brandão-Dias, Nazaré L. Abreu and Evanguedes Kalapothakis

244 ***Evaluation of Reference Genes to Analyze Gene Expression in Silverside* Odontesthes humensis *Under Different Environmental Conditions***
Tony L. R. Silveira, William B. Domingues, Mariana H. Remião, Lucas Santos, Bruna Barreto, Ingrid M. Lessa, Antonio Sergio Varela Junior, Diego Martins Pires, Carine Corcini, Tiago Collares, Fabiana K. Seixas, Ricardo B. Robaldo and Vinicius F. Campos

257 ***Unrevealing Parasitic Trophic Interactions—A Molecular Approach for Fluid-Feeding Fishes***
Karine O. Bonato, Priscilla C. Silva and Luiz R. Malabarba

265 ***Fishing Into the MicroRNA Transcriptome***
Marcos E. Herkenhoff, Arthur C. Oliveira, Pedro G. Nachtigall, Juliana M. Costa, Vinicius F. Campos, Alexandre W. S. Hilsdorf and Danillo Pinhal

280 ***Little Divergence Among Mitochondrial Lineages of* Prochilodus *(Teleostei, Characiformes)***
Bruno F. Melo, Beatriz F. Dorini, Fausto Foresti and Claudio Oliveira

289 ***Trends in Karyotype Evolution in* Astyanax *(Teleostei, Characiformes, Characidae): Insights From Molecular Data***
Rubens Pazza, Jorge A. Dergam and Karine F. Kavalco

300 ***Molecular Identification of Shark Meat From Local Markets in Southern Brazil Based on DNA Barcoding: Evidence for Mislabeling and Trade of Endangered Species***
Fernanda Almerón-Souza, Christian Sperb, Carolina L. Castilho, Pedro I. C. C. Figueiredo, Leonardo T. Gonçalves, Rodrigo Machado, Larissa R. Oliveira, Victor H. Valiati and Nelson J. R. Fagundes

312 ***Hidden Diversity Hampers Conservation Efforts in a Highly Impacted Neotropical River System***
Naiara G. Sales, Stefano Mariani, Gilberto N. Salvador, Tiago C. Pessali and Daniel C. Carvalho

323 ***Remarkable Geographic Structuring of Rheophilic Fishes of the Lower Araguaia River***
Tomas Hrbek, Natasha V. Meliciano, Jansen Zuanon and Izeni P. Farias

335 ***A Multilocus Approach to Understanding Historical and Contemporary Demography of the Keystone Floodplain Species* Colossoma macropomum *(Teleostei: Characiformes)***
Maria da Conceição Freitas Santos, Tomas Hrbek and Izeni P. Farias

frontiers in Genetics

EDITORIAL
published: 04 November 2019
doi: 10.3389/fgene.2019.01124

Editorial: Genetics, Evolution, and Conservation of Neotropical Fishes

Rodrigo Augusto Torres[1] *and Roberto Ferreira Artoni*[2]*

[1] *Laboratório de Genômica Evolutiva e Ambiental, Departamento de Zoologia, Centro de Biociências, Universidade Federal de Pernambuco, Recife, Brazil,* [2] *Laboratório de Genética e Evolução, Departamento de Biologia Estrutural, Molecular e Genética, Universidade Estadual de Ponta Grossa, Ponta Grossa, Brazil*

Keywords: Chromosome, molecular marker, Genetic population, Evolutionary, fish

Editorial on the Research Topic

Genetics, Evolution, and Conservation of Neotropical Fishes

Fishes are a more basal and speciose vertebrate group and they occupy the most diverse aquatic environments on earth. Fish are extremely diverse, especially in biologically mega diverse countries such as those in the Neotropical region, encompassing a very large area in the Americas. The amazing diversity alone generates great interest in fish biology, ecology and evolution, but despite that interest, many challenges remain. Today, new genetic methods are available to study evolutionary and ecological processes that have influenced and given rise to their wide diversity and the many habitats that these fish occupy. Consequently, these genetic methods can also serve as tools for better conservation and management. Today, we must ask ourselves what are the classic questions not yet answered? What new information is available and useful to answer those questions? Where will this new science lead us into the future? To begin answering those questions, we asked these, and other, questions of the scientific community with the goal of gathering together answers to those questions for the Frontiers in Genetics, which celebrates the 30th anniversary of the first Symposium on Cytogenetics and Fish Genetics.

Traditionally, karyotpye analysis is among the genetic methods most often applied to Neotropical fish. Since the 1970s, karyotypes studies have been used to describe diversity and thereby provide support for evolutionary systematics, taxonomy and the evolution of karyotypes themselves. Banding and classical chromosomal markers, now used with the fluorescence *in-situ* hybridization (FISH) method and its derivatives, have advanced Neotropical fish cytogenetics which has allowed for more extensive comparative analyses with African fish (Carvalho et al.) and to investigate incipient speciation (Ferreira et al.) and evolutionary divergences of sex chromosomes (Sember et al.), as well as complex chromosomal rearrangements, that are a recurrent feature found in many neotropical fish families (Machado et al.). With integration of cytogenetics and molecular markers, a new cycle of methodological developments has come to pass, especially with the inclusion of dates that became available through massive sequencing of DNA and RNA. Trends in karyotype evolution have been refined in several species complexes, including the genera *Astyanax* (Pazza et al.) and *Mugil* (Nirchio et al.), as well as permitting the uncovering of surprising and previously unknown diversity in the armored catfish *Ancistrus* (Prizon et al.). While a good beginning, bioinformatics use in the selection of satellite DNA markers and their chromosome localization is still growing (Utsunomia et al.). This tremendous growth of information and the availability of more robust methodological tools, and finding undiscovered diversity, have resulted in a growing concern for understanding the evolutionary history of this diversity as well as its conservation (Moritz, 2002). Due to rapidly declining biodiversity of Neotropical fish, integrating conservation genetics with data describing systematics and ecological information including details of the adaptive landscape, together will become important for future research. Recently, the application of genetic evolutionary analysis

OPEN ACCESS

Edited and reviewed by:
Lior David,
Hebrew University of Jerusalem, Israel

***Correspondence:**
Roberto Ferreira Artoni
rfartoni@gmail.com

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

Received: *04 December 2018*
Accepted: *17 October 2019*
Published: *04 November 2019*

Citation:
Torres RA and Artoni RF (2019) Editorial: Genetics, Evolution, and Conservation of Neotropical Fishes. Front. Genet. 10:1124. doi: 10.3389/fgene.2019.01124

in fish conservation has shown its importance by demonstrating that commercially exploited species are losing genetic variability (Allendorf et al., 2014; Pinsky and Palumbi, 2014). Investigating taxonomic uncertainties and genetic diversity (evolutionary potential) of the many species of fish, diagnosing patterns of gene flow, and the processes that structure and fragment populations are among the most important topics for the conservation of fish in their natural environments. In this volume, we bring attention to analyses of genetic structure, population diversity, and research in evolutionary history. Conservation issues include problems that arise from hydroelectric dams that lead to the disappearance of rheophilic fish fauna that previously occupied the river being dammed (Hrbek et al.). This has resulted in reduced genetic diversity and changing population structure of geographical variants in the catfishes *Steindachneridion scriptum* (Paixão et al.) and *S. parahybae*, both of which are endangered (ICMBio/MMA, 2018) (Fonseca et al.). Additionally, dams are detrimental for the pirarucu, *Arapaima gigas*, cited by the European Union that set limits on the exploitation of vulnerable or potentially endangered species (CITES, 2017) (Vitorino et al.).

The current distribution of neotropical fish genetic lineages may reflect ancient geological events such as paleo-drains, or more recent phenomena such as stream head capture.A special south Brazil hydrological system (Pampa) resulted from a long history of tectonism, climate, and sea-level changes directly affecting the diversification of fish evolutionary lineages in this biome (Ramos-Fregonezi et al.). Migratory behavior and gene flow in *Prochilodus* show little genetic divergence, however it is possible to delimit mitochondrial lineages suggestive of distinct species (Melo et al.). Natural populations of *Colossoma macropomum* comprise a single and large panmitic population in the main channel of the Solimões-Amazonas River basin, and structured populations in the headwaters of the tributaries with the greatest genetic differentiation in relation the Bolivian sub-basin (da Conceição Freitas Santos et al.). In paleo-drainage reconstructions, two putative paleo-rivers in southeastern Brazil include threatened and endemic species (Lima et al.). Pleistocene climate changes were major historical events that had a lasting impact on South American biodiversity for the endangered *Salminus brasiliensis* (ICMBio/MMA, 2018), which is found in the Pantanal basin, in which there was a period of significant population expansion (de Carvalho Mondin et al.). The use of microsatellite markers has expanded conservation tools for endemic species and populations (Souza-Shibatta et al.), and when used in conjunction with markers for phylogenetic and phylogeographic analyses, can uncover sometimes complex, often alarming, scenarios (Silva-Santos et al.). Using microsatellites, derived from transcriptome sequencing, can also be important for genetic management of natural or captive-bred commercial stocks (Jorge et al.; Ariede et al.). New strands of research constantly expand the set of evolutionary inferences about Neotropical fish.Forensic genetics has also benefitted from new tools, and using DNA barcoding has assisted in the identification of meats from different shark species (some endangered) in the domestic market in Brazil (Almerón-Souza et al.). Sequencing of cytochrome oxidase subunit 1 (COI) has helped unravel trophic relationships by the molecular identification of the stomach contents of parasitic catfishes (Bonato et al.). DNA barcoding has also revealed species complexes, such as *Megaleporinus* (Ramirez et al.) and *Schizolecis guntheri* (Souza et al.), and is a promising tool for future studies of the taxonomy of Neotropical fish. New methodologies and applications for the study of Neotropical fish genetics are constantly being developed and applied in new ways. The chromosome set manipulation by polyploidization has proved to be an important tool in obtaining sterile individuals. Recently, a protocol was developed for using flow cytometry in the confirmation of triploidy in *Astyanax altiparanae* (Xavier et al.). Beyond doubt, the coming years will be stimulated by the popularization of massive sequencing of DNA and RNA and the development of user-friendly software for analysing large data sets will follow. Improved experimental design and sequencing strategies will be important and one-such strategy for the use of next-gen sequencing technology to genotype many individuals in parallel is described for Neotropical fish (Pimentel et al.). The integration of Omics information (genome, transcriptome, and proteome) with environmental data in the analysis of biological functions and epigenetic responses is an exciting development for understanding the evolutionary biology of Neotropical fish. For example, the typical Amazonian condition, where rivers with blackwater, clearwater, and whitewater converge, transcriptomic analysis indicated phenotypic plasticity in *Triportheus albus* that allows it to survive in all these environments (Araújo et al.). In this volume, the expression of reference genes was determined by qRT-PCR to set environmental analysis parameters for study of *Odontesthes humensis* (Silveira et al.). Stable genes (i.e., good as control) were distinguished from genes that are tissue specific and unstable in that study. Another recent advance is combining physiological and gene expression analyses to better understand the ecology of Neotropical fishes. Salinity tolerances in the marine *Odonthestes* species demonstrated different physiological adaptations resulting from evolutionary processes that occurred in oceanic and estuarine environments (Silveira et al.). MicroRNA use is still incipient, but the role of these molecules in the regulation of gene expression is being investigated in fishes (Herkenhoff et al.).

Here, we certainly have not described the universe of evolutionary biology of, or the tools with which to study, Neotropical fish. Yet, we hope that those interested in this area of biological research now have an overview of the state of the art and that they will be encouraged to take additional steps to extend our current understanding of the evolutionary biology, ecology and conservation of Neotropical fish.

We greatly appreciate the invaluable contribution of all reviewers as well as the efforts of Chief Editor Dr. Samuel A. Cushman of Frontiers in Genetics: Evolutionary and Population Genetics with Research Topic and the experts and support staff at Frontiers.

AUTHOR CONTRIBUTIONS

RA and RT contributed equally to the preparation of this editorial manuscript of the special volume edited by these guest editors.

REFERENCES

Allendorf, F. W., Berry, O., and Ryman, N. (2014). So long to genetic diversity, and thanks for all the fish. *Mol. Ecol.* 23, 23–25. doi: 10.1111/mec.12574

CITES (2017). "Convention on International Trade in Endangered Species of Wild Fauna and Flora: Appendicies I, II and III," in *Convention on International Trade in Endangered Species of Wild Fauna*.

ICMBio/MMA (2018). *Livro Vermelho da Fauna Brasileira Ameaçada de Extinção: Volume VI – Peixes. 7*. Brasília: ICMBio.

Moritz, C. (2002). Strategies to protect biological diversity and the evolutionary processes that sustain it. *Syst. Biol.* 51, 238–254. doi: 10.1080/10635150252899752

Pinsky, M. L., and Palumbi, S. R. (2014). Meta-analysis reveals lower genetic diversity in overfished populations. *Mol. Ecol.* 23, 29–39. doi: 10.1111/mec.12509

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

Copyright © 2019 Torres and Artoni. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 14 August 2017
doi: 10.3389/fgene.2017.00103

A Glimpse into the Satellite DNA Library in Characidae Fish (Teleostei, Characiformes)

Ricardo Utsunomia[1], Francisco J. Ruiz-Ruano[2], Duílio M. Z. A. Silva[1], Érica A. Serrano[1], Ivana F. Rosa[1], Patrícia E. S. Scudeler[1], Diogo T. Hashimoto[3], Claudio Oliveira[1], Juan Pedro M. Camacho[2] and Fausto Foresti[1]*

[1] Department of Morphology, Institute of Biosciences, São Paulo State University, Botucatu, Brazil, [2] Departamento de Genética, Facultad de Ciencias, Universidad de Granada, Granada, Spain, [3] CAUNESP, São Paulo State University, Jaboticabal, Brazil

Satellite DNA (satDNA) is an abundant fraction of repetitive DNA in eukaryotic genomes and plays an important role in genome organization and evolution. In general, satDNA sequences follow a concerted evolutionary pattern through the intragenomic homogenization of different repeat units. In addition, the satDNA library hypothesis predicts that related species share a series of satDNA variants descended from a common ancestor species, with differential amplification of different satDNA variants. The finding of a same satDNA family in species belonging to different genera within Characidae fish provided the opportunity to test both concerted evolution and library hypotheses. For this purpose, we analyzed here sequence variation and abundance of this satDNA family in ten species, by a combination of next generation sequencing (NGS), PCR and Sanger sequencing, and fluorescence *in situ* hybridization (FISH). We found extensive between-species variation for the number and size of pericentromeric FISH signals. At genomic level, the analysis of 1000s of DNA sequences obtained by Illumina sequencing and PCR amplification allowed defining 150 haplotypes which were linked in a common minimum spanning tree, where different patterns of concerted evolution were apparent. This also provided a glimpse into the satDNA library of this group of species. In consistency with the library hypothesis, different variants for this satDNA showed high differences in abundance between species, from highly abundant to simply relictual variants.

Keywords: concerted evolution, repetitive DNA, *in situ* hybridization, satellite DNA, genome evolution

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Marcelo De Bello Cioffi, Federal University of São Carlos, Brazil
Paloma Morán, University of Vigo, Spain

****Correspondence:***
Ricardo Utsunomia
utricardo@ibb.unesp.br

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

Received: *05 July 2017*
Accepted: *26 July 2017*
Published: *14 August 2017*

Citation:
Utsunomia R, Ruiz-Ruano FJ, Silva DMZA, Serrano ÉA, Rosa IF, Scudeler PES, Hashimoto DT, Oliveira C, Camacho JPM and Foresti F (2017) A Glimpse into the Satellite DNA Library in Characidae Fish (Teleostei, Characiformes). Front. Genet. 8:103. doi: 10.3389/fgene.2017.00103

INTRODUCTION

Eukaryotic genomes are composed of huge amounts of highly dynamic repetitive DNA sequences that may be dispersed throughout the genomes, e.g., transposable elements, or tandemly repeated, such as multigene families or satellite DNA (satDNA; Charlesworth et al., 1994; Jurka et al., 2005). satDNA constitutes a non-coding fraction of the genome, consisting in long arrays of tandemly repeated sequences, preferentially located on the heterochromatin of pericentromeric and subtelomeric chromosome regions, although their presence in euchromatic regions has already been reported (López-Flores and Garrido-Ramos, 2012; Plohl et al., 2012; Garrido-Ramos, 2015; Ruiz-Ruano et al., 2016). In general, satDNA sequences

constitute different families that vary in localization, constitution, unit size and abundance (Garrido-Ramos, 2015). Since these sequences are highly dynamic genomic segments being susceptible to quick changes, these elements are generally species- or genus-specific (Vicari et al., 2010; Garrido-Ramos, 2015). According to the "library hypothesis" specific groups of related organisms share a common library of satDNAs that might be independently amplified in those distinct genomes (Fry and Salser, 1977). Such events might cause rapid changes in satDNA distribution and abundance profiles, even in closely related species (Plohl et al., 2012).

Up to now, most studies of satDNAs in fish genomes have focused on the development of chromosomal markers for evolutionary studies on B and sex chromosomes (Mestriner et al., 2000; Jesus et al., 2003; Vicari et al., 2010; Utsunomia et al., 2016). However, the evolutionary trends of satDNAs in closely related fish species have not yet been well evaluated, mainly if we consider that almost all discovered satDNA analyzed until now seemed to represent species- or genus-specific sequences (Garrido-Ramos et al., 1999; Leclerc et al., 1999; de la Herrán et al., 2001; Lanfredi et al., 2001; Robles et al., 2004; Martins et al., 2006).

Next generation sequencing (NGS) has been extensively used for several applications, including the in-depth characterization of satDNA sequences by similarity-based read clustering (Macas et al., 2011; Ruiz-Ruano et al., 2016). Such strategy has been frequently used for *de novo* characterization of repetitive DNA sequences in different organisms (Novák et al., 2010; Macas et al., 2011; Pagán et al., 2012; Camacho et al., 2015; García et al., 2015; Ruiz-Ruano et al., 2016; Utsunomia et al., 2016). In a recent study, Utsunomia et al. (2016) used graph-based clustering of sequence reads and isolated seven satDNAs (MS1-MS7) from the characid fish *Moenkhausia sanctaefilomenae*, two of which (MS3 and MS7) were fully characterized and mapped on chromosomes to unveil B chromosome origin in this species. More recently, it was evidenced that one of these satellites, MS1 satDNA (from now on referred to as MsaSat01-177, to follow the nomenclature rules suggested in Ruiz-Ruano et al., 2016), was found in the genomes of other characid fishes, such as *Astyanax paranae* and *A. mexicanus* (Silva et al., submitted), indicating its intergenera conservation and thus providing an interesting opportunity to investigate the evolutionary dynamics of this satellite in closely related species within Characidae.

Characidae is the largest family of freshwater fishes and comprises more than 1000 species (Eschmeyer and Fong, 2017). The phylogenetic relationships of this family are highly controversial and several species were considered *incertae sedis* by different authors (Javonillo et al., 2010; Oliveira et al., 2011; Thomaz et al., 2015). During the last few years, different studies using morphological and molecular evidence showed that Characidae is a well supported group which is subdivided into three different monophyletic clades (clades A, B, and C) (Weitzman and Malabarba, 1998; Javonillo et al., 2010; Oliveira et al., 2011; Thomaz et al., 2015). However, phylogenetic hypotheses within each of these clades are still scarce or unavailable and many genera are suspected to be non-monophyletic (Thomaz et al., 2015; Rossini et al., 2016).

Likewise, numerous cytogenetic studies were performed in representatives of this family during the last decades, which revealed extensive karyotype diversification at intra- and inter-species levels, including changes in diploid numbers, differential chromosomal location of multigene families and multiple origins of supernumerary chromosomes (Oliveira et al., 2009; Arai, 2011). However, the absence of satDNAs shared among species has impeded testing the main evolutionary hypotheses on this kind of repetitive DNA, such as concerted evolution and the library hypothesis (see above). Our main purpose here was to test these hypotheses on a satDNA shared between several Characidae species, using a combination of novel (Illumina sequencing) and traditional (PCR amplification, cloning, Sanger sequencing and FISH) approaches, in 10 species of Characidae fish belonging to A, B, and C clades. Therefore, our main objectives were: (i) delimiting the taxonomic spread of this satellite, (ii) comparing its chromosome abundance and localization between species, and (iii) investigating intra- and interespecific variation of MsaSat01-177 at nucleotide and chromosomal levels. All this information provided new insights on concerted evolution and the library hypothesis.

MATERIALS AND METHODS

Ethics Statement

Sampling was carried out on private lands and the owners gave permission to conduct this study. The animals were captured using nets, transported to the Laboratory, kept in a fish tank and were anesthetized before the analyses. The animals were collected in accordance with Brazilian environmental protection legislation (Collection Permission MMA/IBAMA/SISBIO—number 3245) and the procedures for sampling, maintenance and analysis of the fishes were performed in compliance with the Brazilian College of Animal Experimentation (COBEA) and was approved (protocols 405 and 504) by the BIOSCIENCE INSTITUTE//UNESP ETHICS COMMITTEE ON THE USE OF ANIMALS (CEUA).

Sampling, Chromosomal Preparations and DNA Extraction

In the present study, we analyzed ten allopatric Characidae species, namely *Astyanax paranae, A. bockmanni*, *A. altiparanae*, *A. fasciatus*, *A. jordani*, *M. sanctaefilomenae*, *Hasemania kalunga* and *Hyphessobrycon bifasciatus*, all of them belonging to clade C. In addition, *Bryconamericus stramineus* and *Serrapinus notomelas*, classified as clades A and B, respectively, were also analyzed (**Table 1**). The relationship between clades A, B, and C is represented in **Figure 1**. The available internal relationships among clade A species were not considered in this study, as several genera appear to be non-monophyletic. Cell suspensions from all species were already available in our laboratory from previous studies (Silva et al., 2013, 2014, 2016; Utsunomia et al., 2016), except for *H. bifasciatus*, *H. kalunga*, *B. stramineus* and *S. notomelas* whose karyotypes were analyzed here for the first time. Metaphase chromosomes were obtained from cell suspensions of the anterior kidney, according to Foresti et al.

TABLE 1 | Analyzed species in the present study and information regarding MsaSat01-177 distribution patterns.

Clade	Species	2n	PCR	Pattern	Sites
Clade C	*M. sanctaefilomenae*	50	+	c	36
Clade C	*A. paranae*	50 + B	+	c	10
Clade C	*A. fasciatus*	46 + B	+	nc	nc
Clade C	*A. jordani*	50	+	c	18
Clade C	*A. altiparanae*	50	+	c	10
Clade C	*A. bockmanni*	50	+	c	2
Clade C	*Hasemania kalunga*	50	+	c	2
Clade C	*Hyphessobrycon bifasciatus*	50	+	c	18
Clade A	*B. stramineus*	52	–	–	–
Clade B	*S. notomelas*	52	–	–	–

2n, Diploid chromosome number. Sites, number of chromosomes showing the satDNA. c, clustered; nc, non-clustered.

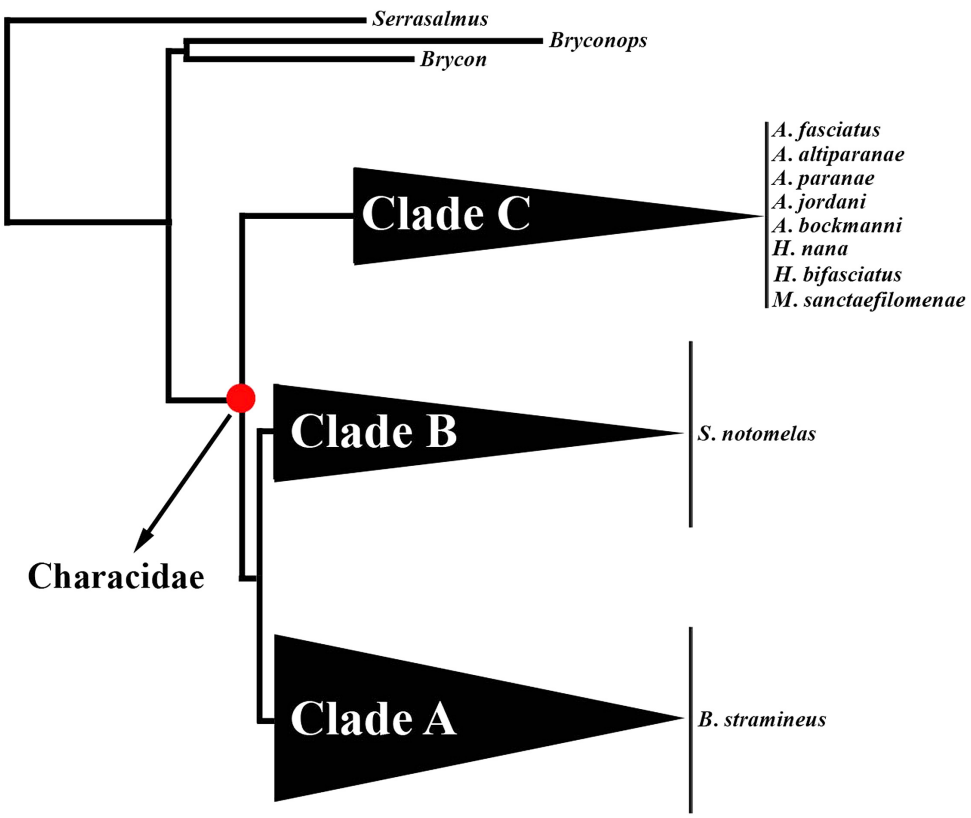


FIGURE 1 | Phylogeny showing the relationships between clades "A," "B," and "C" of Characidae adapted from Thomaz et al. (2015). The red circle on the internal branches indicates Characidae. Species analyzed here from the three clades are indicated on the right.

(1981). Genomic DNA was extracted from muscle or liver, using the Wizard Genomic DNA Purification Kit (Promega), following the manufacturer's instructions.

Whole-Genome Sequencing and Characterization of Monomers from Raw Reads

MsaSat01-177 was previously discovered in the *M. sanctaefilomenae* genome using RepeatExplorer (Utsunomia et al., 2016). Here, in order to perform a thorough search for MsaSat01-177 monomers in different genomic libraries, we used gDNA Illumina HiSeq2000 reads (2x101bp) from *M. sanctaefilomenae* and *A. paranae* stored in SRA (accession numbers SRR5839692 and SRR5461470, respectively). In addition, two individuals of *A. fasciatus* were sequenced on the Illumina MiSeq, yielding 2 × 250 bp paired-end reads. Firstly, in these three species, we performed a random sequence subsampling step of 5.000.000 paired-end reads per species. Detailed information about the used Illumina libraries is shown in **Table 2**. In addition, we used other

TABLE 2 | Genetic variation found in the monomers of MsaSat01-177 extracted from Illumina reads of three different species and PCR-amplified from other five different characid species.

Species	Library	Seq	*N*	Size	Hap	Hd	π
M. sanctaefilomenae	2 × 101 bp	10^6	470	172–180bp	84	0.884	0.013
A. paranae	2 × 101 bp	10^6	201	170–180bp	68	0.9844	0.055
A. fasciatus	2 × 250 bp	10^6	8	175–178bp	8	1	0.18

Seq, Number of sequences in the primary fastq library. N, number of isolated monomers. Hap, number of haplotypes. Hd, haplotype diversity. π, nucleotide diversity.

gDNA Illumina MiSeq reads (2 × 250 bp) recently sequenced in our laboratory (data not shown) from several characiform fishes, including Anostomidae (*Megaleporinus macrocephalus* and *Leporinus friderici*), Crenuchidae (*Characidium gomesi*) and Serrasalmidae (*Piaractus mesopotamicus*), all belonging to the Characiformes order (Oliveira et al., 2011).

To obtain a detailed and reliable score of haplotype abundance for MsaSat01-177 sequences from the genomic libraries of *A. paranae*, *A. fasciatus* and *M. sanctaefilomenae*, we extracted complete monomers directly from the Illumina raw reads, as this is expected to provide accurate estimates of haplotype abundance without the bias of PCR amplification. For this purpose, we performed a series of bioinformatic workflows that included joining the paired-end reads, aligning them against the MsaSat01-177 sequence and trimming the ends to get full monomers, as described in Utsunomia et al. (2016). Importantly, singletons (e.g., sequence variants found only once) were discarded at this stage of the analysis in order to minimize the impact of possible sequencing errors. Collected monomers from Illumina reads in the three species were aligned separately using the Muscle algorithm (Edgar, 2004), under default parameters, to be displayed as sequence logos using the WebLogo 3.3 software (Crooks et al., 2004). The obtained monomers were used for all downstream analyses in this study, except for RepeatExplorer (described below).

In order to investigate possible structural variation of MsaSat01-177 in these three species and to search for possible associations with other repetitive elements, we selected pairs of reads showing homology with this satDNA in each gDNA library separately, by using BLAT (Kent, 2002). This step is implemented in a custom script[1]. We then used the selected read pairs from each library to run RepeatExplorer clustering (Novák et al., 2013) with at least 2 × 2500 reads.

satDNA Amplification, Cloning and Sequencing

After complete characterization of the repetition unit of MsaSat01-177, different sets of divergent primer pairs were designed: MsaSat01F1 (5′-TTTTGACCATTCATGAAACCTTG-3′) and MsaSat01R1 (5′-ACCAGAATCACATACCGCGG-3′); MsaSat01F2 (5′-TGCCCATGCATTTTCCCACT-3′) and MsaSat01R2 (5′-GAARGATTTCATGAAATTTYGC-3′). PCR reactions were performed in 1x PCR buffer, 1.5 mM $MgCl_2$, 200 μM each dNTP, 0.1 μM each primer, 2 pg–10 ng of DNA and 0.5 U of *Taq* polymerase (Invitrogen). The cycling program for amplification consisted of an initial denaturation at 95°C for 5 min, followed by 30 cycles at 95°C for 20 s, 63°C for 30 s, 72°C for 20 s and a final extension at 72°C for 15 min. The PCR products were visualized in 2% agarose gels, and the fragment obtained from each sample was extracted from the gel and cloned into the pGEM-T Easy Vector (Promega, Madison, WI, United States). DNA sequencing was performed with the Big Dye TM Terminator v3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems) following the manufacturer's instructions. Consensus sequences from forward and reverse strands of the sequenced clones were obtained using Geneious Pro v.8.04.

DNA Probes and FISH

DNA probes for MsaSat01-177 were obtained by PCR amplification on genomic DNA from all species, except *B. stramineus* and *S. notomelas*, using the same conditions described above and labeling DNA with digoxigenin-11-dUTP or biotin-16-dUTP. Complementarily, probes were also obtained directly from single cloned sequences to compare the results. Thus, for every species, FISH was performed using probes obtained from their own genomes.

Fluorescence *in situ* hybridization was performed under high-stringency conditions using the method described by Pinkel et al. (1986). Pre-hybridization conditions included a 1-h incubation with RNAse (50 μg/ml) followed by chromosomal DNA denaturation in 70% formamide/2x SSC for 5 min at 70°C. For each slide, 300 μl of hybridization solution (containing 200 ng of labeled probe, 50% formamide, 2x SSC and 10% dextran sulfate) was denatured for 10 min at 95°C, then dropped onto the slides and allowed to hybridize overnight at 37°C in a moist chamber containing 2x SSC. Post-hybridization, all slides were washed in 0.2x SSC/15% formamide for 20 min at 42°C, followed by a second wash in 0.1x SSC for 15 min at 60°C and a final wash at room temperature in 4x SSC, 0.5% Tween for 10 min. Probe detection was carried out with avidin-FITC (Sigma) or anti-digoxigenin-rhodamine (Roche), and the chromosomes were counterstained with DAPI (4′,6-diamidino-2-phenylindole, Vector Laboratories) and analyzed under an optical photomicroscope (Olympus BX61). Images were captured with an Olympus DP70 digital camera and with the Image Pro plus 6.0software (Media Cybernetics). From each individual, a minimum of five cells was analyzed for FISH.

Nucleotide Analyses

A global alignment from both Illumina-derived and PCR-derived sequences was generated using the Muscle algorithm (Edgar, 2004) under default parameters. DNA diversity analyses,

[1]https://github.com/fjruizruano/ngs-protocols/blob/master/mapping_blat_gs.py

considering indels and all haplotypes, were performed with DnaSP v5.05 (Librado and Rozas, 2009). In order to get fewer haplotypes in the Minimum spanning tree (MST), we performed a clustering analysis with CD-HIT-EST (Li and Godzik, 2006) selecting a sequence identity level of 99%. The MST was built on the basis of pairwise differences using ARLEQUIN v3.5.1.3 (Excoffier and Lischer, 2010) and was visualized with HAPSTAR (Teacher and Griffiths, 2011).

RESULTS

Chromosomal Analysis

Cytogenetic analyses evidenced different diploid chromosome numbers for the analyzed species (**Table 1**). PCR amplification of MsaSat01-177 yielded a ladder pattern in 2% agarose gels for all species within clade C, while no visible banding patterns were detected for species from clade A (*B. stramineus*) and B (*S. notomelas*), suggesting that these sequences are not present in these species or were not amplified with the designed primers due to high sequence divergence. Also, FISH with inter-specific probes did not returned any visible signal on the chromosomes of these two species (data not shown).

FISH evidenced that MsaSat01-177 shows a non-clustered organization in *A. fasciatus*, but a clustered distribution in the other C-clade species. Remarkably, all clusters for this satDNA were located pericentromerically (**Figure 2**), but showing extensive variation among species concerning the number of chromosomes carrying it, namely two in *H. kalunga* and *A. bockmanni*, 10 in *A. altiparanae* and *A. paranae*, 18 in *H. bifasciatus* and *A. jordani*, and 36 in *M. sanctaefilomenae*.

Bioinformatic and Molecular Analyses

Selection of Illumina reads showing homology with MsaSat01-177 resulted in 298.622, 3.160 and 24 reads in *M. sanctaefilomenae*, *A. paranae* and *A. fasciatus*, respectively. Those found in the latter species were insufficient for RepeatExplorer analysis and, in *A. paranae*, we had to use two copies of each read in order to meet the requirement of 5.000 reads minimum. Finally, for *M. sanctaefilomenae*, we subsampled the reads from 298.622 to 30.000 reads to optimize RepeatExplorer calculations. Output data evidenced spherical graphs for MsaSat01-177 in both species (**Supplementary Figure S1**), as expected for satDNAs. Although these results do not exclude the possibility of association with other repetitive sequences, they indicate that this satDNA is not primarily associated with other repetitive elements.

We successfully extracted monomers directly from sequencing reads of *A. paranae*, *A. fasciatus* and *M. sanctaefilomenae* and the detailed information is summarized in **Table 2**. Conversely, searches for MsaSat01-177 in distinct Characiformes genomes, except Characidae, did not yield any result, suggesting that MsaSat01-177 is not present in other families than Characidae within this order. In this context, we restricted our high-throughput analyses to the three Characidae fishes available. The extraction of MsaSat01-177 monomers from read pairs showing overlapping, resulted in a total of 470, 201 and 8 monomers in *M. sanctaefilomenae*, *A. paranae* and *A. fasciatus*, respectively. The eight sequences in the latter species showed the highest nucleotide diversity (π), whereas those in *A. paranae* showed higher nucleotide diversity than those in *M. sanctaefilomenae* (**Table 2**). Sequence logos corroborated this result and exhibited different levels of sequence conservation between the analyzed species for MsaSat01-177 monomers, with those in *M. sanctaefilomenae* showing higher conservation than those in *A. paranae* and *A. fasciatus* (**Figure 3**).

PCR amplification in the C-clade species, and subsequent cloning and sequencing, yielded several sequences per species (**Table 3**). Notably, most of the few sequenced clones in *A. paranae*, *A. fasciatus* and *M. sanctaefilomenae* were also found among the Illumina reads. In general, the number of haplotypes was almost equal to the number of sequenced clones for all species, while nucleotide diversity (π) values were variable, with those in *A. altiparanae* showing the highest values.

In order to obtain a global alignment and generate a MST, we firstly performed a clustering step with CD-HIT-EST to the Illumina-derived monomers to reduce the numbers of haplotypes. Thus, a total of 470, 201 and 8 monomers were reduced to a final matrix with 55, 51 and 8 clusters, from *M. sanctaefilomenae*, *A. paranae* and *A. fasciatus*. After that, a final alignment matrix was composed of 150 haplotypes, 114 of which were obtained from Illumina reads and 36 from PCR clones. Considering this whole alignment, we built a MST, considering haplotype relative abundance, which evidenced overall species-specific groups of haplotypes, the main exception being *A. paranae* which showed several groups linked with those in most remaining species (**Figure 2**). The main steps performed in this study to obtain the described results are represented in **Figure 4**.

DISCUSSION

It is generally assumed that satDNA sequences evolve following a pattern of "concerted evolution," as a consequence of intraspecific sequence homogenization and fixation (Dover, 1982, 1986). Notably, the homogenization process is driven by molecular mechanisms such as unequal crossing-over or gene conversion (Smith, 1976; Dover, 1986), which usually lead to a quite low sequence divergence among monomers within satDNA arrays (Plohl et al., 2012). Our sequence analysis of MsaSat01-177 in eight Characidae fish species has revealed some interesting features. First, we could indeed observe a higher homogenization within species since most haplotypes in **Figure 2**, coming from a same species, tended to group together, with the exception of those in *A. paranae*, which were distributed into three different groups. This is the expected pattern for concerted evolution of satDNA, but the *A. paranae* case demands additional explanations (see below).

Second, a comparison between the number of clusters observed by FISH and nucleotide diversity, in the three species analyzed by Illumina sequencing, revealed an interesting pattern, since the species showing the highest number of clusters (*M. sanctafilomenae*) showed the lowest nucleotide

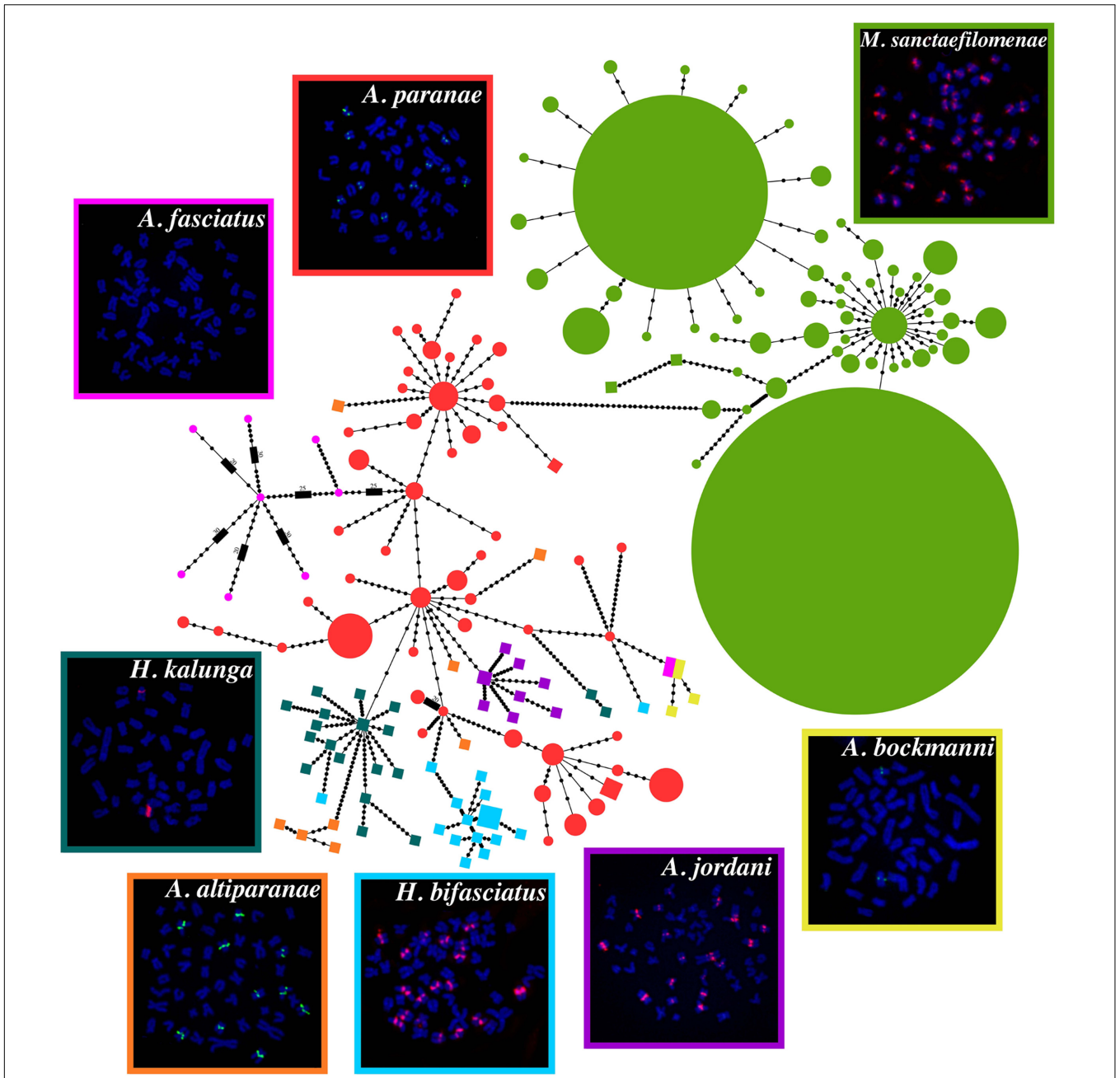


FIGURE 2 | Minimum spanning tree (MST) showing the relationships between the different haplotypes of MsaSat01-177 obtained from distinct species. Colored circles represent haplotypes retrieved from Illumina reads, and the diameter of the circles is proportional to their abundance, whereas PCR-amplified haplotypes are represented by colored squares. Each black dot represents a mutational step. Metaphase plates after FISH with MsaSat01-177 probe are also shown and the colors of the borders correspond to the colors of the circles/squares.

diversity, whereas the species that failed to show clusters (*A. fasciatus*) showed the highest diversity, with *A. paranae* showing intermediate values for both parameters. Population demographical events (e.g., bottlenecks) might have contributed to yield a pattern like this (Ardern et al., 1997; Pons et al., 2002), but between species differences in the homogenization/mutation balance could also provide an explanation (reviewed in Ugarković and Plohl, 2002). For example, a recent amplification event of the MsaSat01-177 in *M. sanctaefilomenae* could explain its low nucleotide diversity. On the other hand, differential amplification between satellite subfamilies could explain the high diversity in *A. paranae* (Willard and Waye, 1987).

Our present results have shown the presence of the MsaSat01-177 satDNA, previously described in the characid

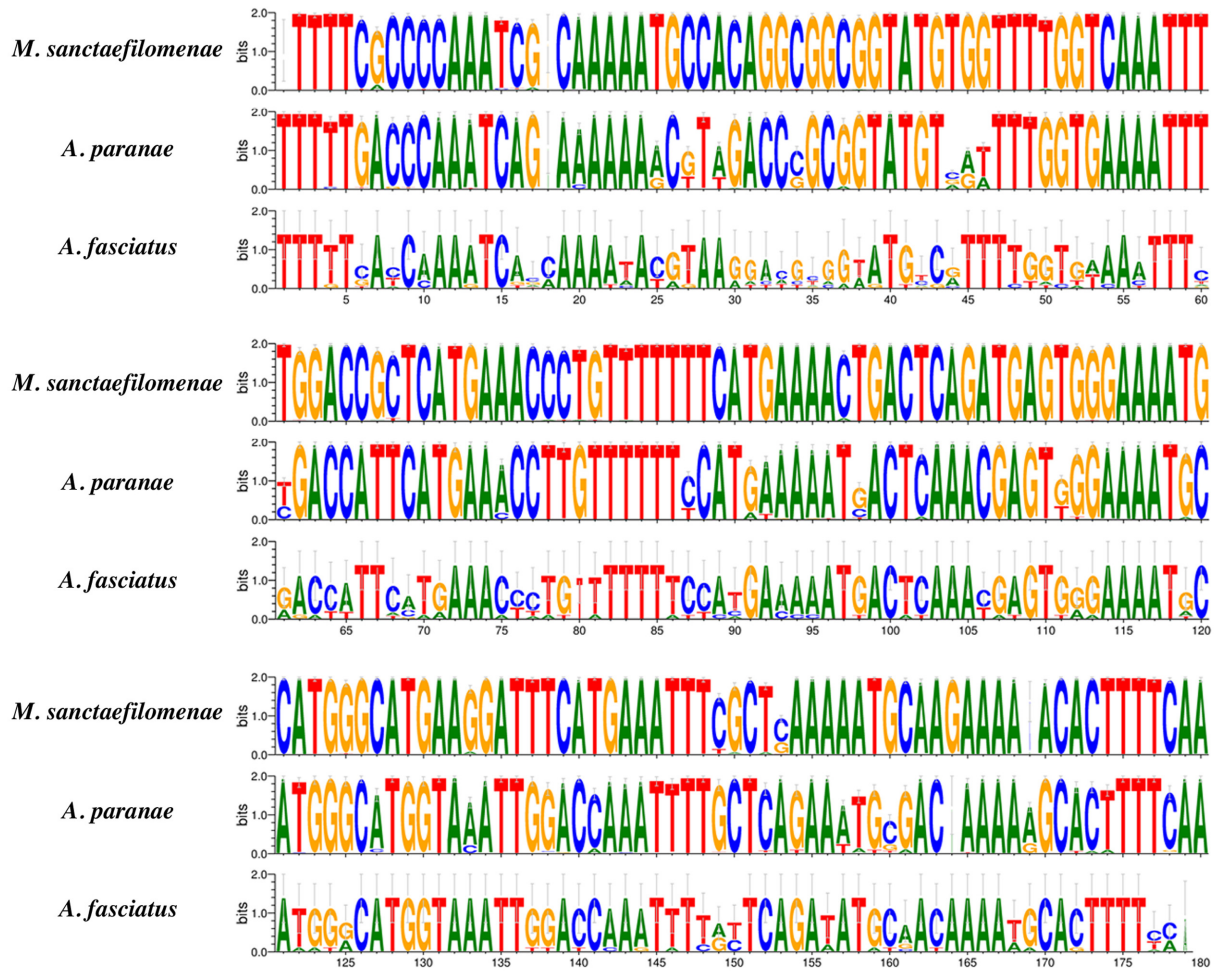


FIGURE 3 | MsaSat01-177 sequence logos, where differences between consensus sequences obtained from *M. sanctaefilomenae*, *A. paranae* and *A. fasciatus* might be observed.

fish *M. sanctaefilomenae* (Utsunomia et al., 2016), in eight species belonging to four different Characidae genera, belonging to Clade C (Oliveira et al., 2011). Notably, we could neither amplify it by PCR in species belonging to clades A or B, nor find any trace of this sequence in genomic libraries of other Characiform families, suggesting that the MsaSat01-177 satDNA might be restricted to the C-clade species. This suggests the conservation of this satDNA in this fish group and allows testing some features of the satDNA library hypothesis (Fry and Salser, 1977) in these Characidae fish.

The MST shown in **Figure 2** provides a glimpse into a small part of the satDNA library of the Characid C-clade species, as it only shows some library volumes for the MsaSat01-177 satDNA in only eight species. Clearly, the complete library should include all haplotypes found in all species for the whole satellitome catalog, with MSTs for each satDNA family and some families connected by common branches if belong to a same superfamily. Of course, this appears to be an impossible task, but looking at a small corner of the library is also very illustrative. Firstly, **Figure 2** shows that species from four different genera share the MsaSat01-177 satellite, with *Astyanax paranae* showing connections with all remaining species, but with higher number of differences with *M. sanctafilomenae*. The central position of *A. paranae* among the five *Astyanax* species might actually be an artifact due to the higher number of sequences obtained from sequence reads in this species. However, in the case of *A. fasciatus*, we also employed this approach but we found only 24 reads showing homology with the MsaSat01-177 satellite, indicating that its presence in this species is just a relic, with only a few small arrays for eight highly divergent haplotypes scattered through the genome, since it was not apparent by FISH (see **Figure 2**). Ruiz-Ruano et al. (2016) suggested that satDNA follows a three-step evolutionary pathway: birth, dissemination and clustering. It is thus conceivable that, in *A. fasciatus*, this satDNA has not reached the third stage, and the extremely high divergence shown by the few units found ($\pi = 0.18$) suggests that they are not subjected

TABLE 3 | Genetic variation found in the monomers of PCR-amplified MsaSat01-177 from eight different characid species.

Species	*N*	Size	Hap	Hd	π
M. sanctaefilomenae	2	179	2	1	0.06
A. paranae	3	177–178	2	0.66	0.07
A. fasciatus	2	167	1	0	0
A. altiparanae	8	177–186bp	8	1	0.086
A. bockmanni	5	167bp	4	0.9	0.01
A. jordani	9	177–182bp	9	1	0.03
H. bifasciatus	19	166–180bp	17	0.98	0.02
H. kalunga	19	177–178bp	18	0.99	0.05

N, number of sequences. Hap, number of haplotypes. Hd, haplotype diversity. π, nucleotide diversity.

to concerted evolution thus probably being disseminated across the genome. In high contrast, in *M. sanctafilomenae*, we found two extremely abundant haplotypes along with many other less abundant ones at few mutational steps, suggesting that sequence homogenization works very efficiently in this species, as also indicated by its low nucleotide diversity ($\pi = 0.013$). This indicates that this species has lost much of the satDNA variants, which were originally present in the common ancestor of the eight species here analyzed. The case of *A. paranae* is intermediate ($\pi = 0.055$), suggesting that it has preserved a higher proportion of the original satDNA variation, presumably because satDNA homogenization works poorer than in *M. sanctafilomenae*. This might explain the central position of *A. paranae* in the tree. As a whole, the former observations are consistent with the independent amplification of satDNA variants in different genomes, suggested by the library hypothesis (Fry and Salser, 1977).

The MST in **Figure 2** also suggests that the Illumina approach is much more informative than the PCR one and that the traditional conclusions on concerted evolution inferred from the latter method could be biased by the unavoidable filtering inherent to the PCR reaction, with products enriched in those sequences to which the primers are able to anchor. Illumina sequencing, however, provides a random sample of sequences thus giving more realistic information. The PCR bias might explain why in **Figure 2** the immense majority of haplotypes obtained by PCR were grouped per species. However, on the basis of the multiple connections shown by the Illumina haplotypes in *A. paranae*, we can imagine a much more intricate haplotype tree with connections between most species. Therefore, the analysis of satDNA variation through Illumina sequencing can open, in the next future, the satDNA library doors wide.

At genus level, it appears that the *Astyanax* satDNA library keeps more variation in common with those in *Hasemania* and *Hyphessobrycon* than with that in *Moenkhausia*, suggesting closer relationship between the three former genera. However, this might be a false impression due to the efficient homogenization in *M. sanctafilomenae*, which has erased, in the satDNA library, many signs of their common descent.

In general, satDNA accumulation in heterochromatic areas is an overall trend, although recent analyses have revealed that euchromatic areas might also be occupied by this kind of repetitive sequences (Kuhn et al., 2012; Ruiz-Ruano et al., 2016). Here, we found that all clusters of MsaSat01-177 found on the chromosomes of seven species were exclusively located on heterochromatic pericentromeric regions. The high differences between species for the number and size of MsaSat01-177 clusters, in the seven species where this satellite was visualized by FISH, from two in *A. bockmanni* and *H. kalunga* to 36 in *M. sanctafilomenae* (see **Table 1**), indicates that satDNA clustering has followed different evolutionary pathways in most species, although it is also conceivable that some clusters residing in chromosomes showing synteny among species might have descended from a common ancestor. According to the three-step hypothesis (Ruiz-Ruano et al., 2016), the former results suggest that satDNA evolution may follow different pathways in different species by reaching variable degrees of interchromosomal spread. Notably, *A. paranae* is phylogenetically more related to *A. bockmanni* than to *A. altiparanae* (Rossini et al., 2016) consistent with our MST. However, the number of sites per genome evidenced by FISH (10, 2 and 10, respectively) would not indicate that. In this context, as other repetitive DNA sequences, the number of sites and satDNA-bearing chromosomes do not appear to completely reflect phylogenetic relationships, and thus probably reflect historical contingency. Unfortunately, as mentioned before, a complete phylogeny considering the taxa sampled in our study is not available (Oliveira et al., 2011; Thomaz et al., 2015; Rossini et al., 2016).

As components of the repetitive fraction of genomes, satDNA is highly dynamic and its abundance might rapidly change due to expansion and/or decrease of these sequence arrays (Plohl et al., 2008; Garrido-Ramos, 2015). Therefore, different mechanisms may have led MsaSat01-177 to be highly abundant and homogenized in *M. sanctaefilomenae*, presumably due to recent amplification on 72% of its chromosomes (50 chromosomes – 36 FISH signals) (**Supplementary Figure S2**). In contrast, this satellite is relictual in *A. fasciatus*, it shows a cluster on a single chromosome pair in *A. bockmanni* and *H. kalunga*, or several chromosome pairs in the remaining species. Such dynamics, at the chromosomal level, has frequently been reported for several satDNA sequences in a wide range of organisms, at the intra- and interespecific levels (Plohl et al., 2012; Garrido-Ramos, 2015). Although multiple mechanisms have been put

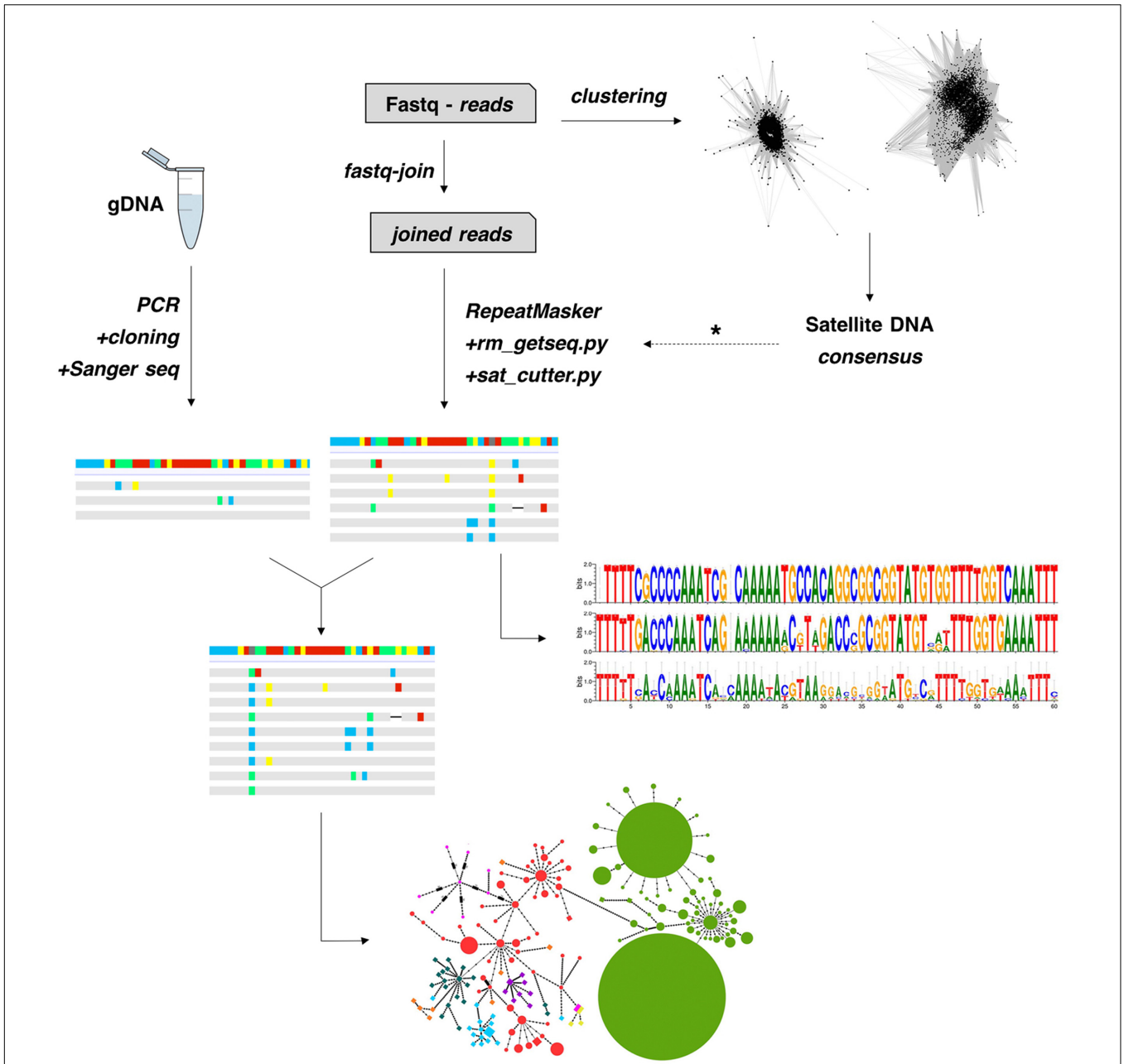


FIGURE 4 | Pipeline used for satDNA analysis in this study. A typical clustering of raw reads with RepeatExplorer is performed to search for a consensus satDNA sequence. After that, this consensus sequence is used as reference (dotted arrow) to collect monomers directly from the reads using a series of custom python scripts. The isolated monomers are then analyzed for different features, such as intragenomic diversity and nucleotide analyses. Also, a strategy combining PCR, cloning and Sanger sequencing might be performed to isolate monomers if NGS reads are not available. Finally, a global alignment is then performed to construct minimum-spanning-trees.

forward to explain this variation, such as unequal crossing-over, ectopic recombination, replication slippage, association with transposable elements and extrachromosomal circular DNA (Dover, 1993; McMurray, 1995; Hancock, 1996; Cohen et al., 2010; Milani and Cabral-de-Mello, 2014; Ruiz-Ruano et al., 2015, 2016), the ultimate explanation has not yet been figured out.

Taken together, our present results have provided evidence for the presence of a shared satDNA among several species within Characidae, which probably arose after the split of Clade C. The chromosomal distribution of MsaSat01-177 was highly variable and several spreading mechanisms might be acting in this case. As expected, monomers from all species are subjected to concerted evolution, except those in *A. fasciatus* where

short tandem arrays of MsaSat01-177 are probably scattered across the genome. In addition, sequence homogenization levels were also different among species, and our results have also shown the differential amplification of some variants for this satellite in different species. This is in high consistency with the library hypothesis (Fry and Salser, 1977), that a same satellite family can follow different evolutionary pathways in different species, including not only for amplification levels but also for chromosome distribution.

AUTHOR CONTRIBUTIONS

RU, DS, PS, ES, and IR collected the samples, performed the cytogenetic analyses, the production of DNA probes and the cloning and FISH experiments. RU, DS, and FR-R performed the bioinformatics analyses. RU, DS, PS, ES, IR, and FR-R drafted the text and designed the figures. DH, JC, CO, and FF critically revised the manuscript and approved the final version.

ACKNOWLEDGMENTS

Authors thank to Renato Devide for help with obtaining samples. This study was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo – (grant numbers 2010/17009-2 and 2014/26508-3 to CO) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (grant number 306054/2006-0 to CO and 403066/2015-8 to RU).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fgene.2017.00103/full#supplementary-material

FIGURE S1 | Cluster-graphs obtained for MsaSat01-177 from *A. paranae* and *M. sanctaefilomenae* after analysis with RepeatExplorer.

FIGURE S2 | Metaphases from different Characidae species after FISH with MsaSat01-177 probes.

REFERENCES

Arai, R. (2011). *Fish Karyotypes: A Check List*. Berlin: Springer Science & Business Media. doi: 10.1007/978-4-431-53877-6

Ardern, S. L., Lambert, D. M., Rodrigo, A. G., and McLean, I. G. (1997). The effects of population bottlenecks on multilocus DNA variation in robins. *J. Hered.* 88, 179–186. doi: 10.1093/oxfordjournals.jhered.a023086

Camacho, J. P. M., Ruiz-Ruano, F. J., Martín-Blázquez, R., López-León, M. D., Cabrero, J., Lorite, P., et al. (2015). A step to the gigantic genome of the desert locust: chromosome sizes and repeated DNAs. *Chromosoma* 124, 263–275. doi: 10.1007/s00412-014-0499-0

Charlesworth, B., Sniegowski, P., and Stephan, L. W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371, 215–220. doi: 10.1038/371215a0

Cohen, S., Agmon, N., Sobol, O., and Segal, D. (2010). Extrachromosomal circles of satellite repeats and 5S ribosomal DNA in human cells. *Mob. DNA* 1:11. doi: 10.1186/1759-8753-1-11

Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

de la Herrán, R., Fontana, F., Lanfredi, M., Congiu, L., Leis, M., Rossi, R., et al. (2001). Slow rates of evolution and sequence homogenization in an ancient satellite DNA family of sturgeons. *Mol. Biol. Evol.* 18, 432–436. doi: 10.1093/oxfordjournals.molbev.a003820

Dover, G. A. (1982). Molecular drive: a cohesive mode of species evolution. *Nature* 299, 111–117. doi: 10.1038/299111a0

Dover, G. A. (1986). Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends Genet.* 2, 159–165. doi: 10.1016/0168-9525(86)90211-8

Dover, G. A. (1993). Evolution of genetic redundancy for advanced players. *Curr. Opin. Genet. Dev.* 3, 902–910. doi: 10.1016/0959-437X(93)90012-E

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Eschmeyer, W. N., and Fong, J. (2017). *Species by Family/Subfamily in the Catalog of Fishes, Electronic Version*. San Francisco, CA: California Academy of Sciences.

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Foresti, F., Almeida-Toledo, L. F., and Toledo-Filho, S. A. (1981). Polymorphic nature of nucleolus organizer regions in fishes. *Cytogenet. Genome Res.* 31, 137–144. doi: 10.1159/000131639

Fry, K., and Salser, W. (1977). Nucleotide sequences of HS-α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* 12, 1069–1084. doi: 10.1016/0092-8674(77)90170-2

García, G., Ríos, N., and Gutiérrez, V. (2015). Next-generation sequencing detects repetitive elements expansion in giant genomes of annual killifish genus *Austrolebias* (Cyprinodontiformes, Rivulidae). *Genetica* 143, 353–360. doi: 10.1007/s10709-015-9834-5

Garrido-Ramos, M. A. (2015). Satellite DNA in plants: more than just rubbish. *Cytogenet. Genome Res.* 146, 153–170. doi: 10.1159/000437008

Garrido-Ramos, M. A., de la Herrán, R., Jamilena, M., Lozano, R., Ruiz Rejón, C., and Ruiz Rejón, M. (1999). Evolution of centromeric satellite DNA and its use in phylogenetic studies of the Sparidae family (Pisces, Perciformes). *Mol. Phylogenet. Evol.* 12, 200–204. doi: 10.1006/mpev.1998.0609

Hancock, J. M. (1996). Simple sequences and the expanding genome. *Bioessays* 18, 421–425. doi: 10.1002/bies.950180512

Javonillo, R., Malabarba, L. R., Weitzman, S. H., and Burns, J. R. (2010). Relationships among major lineages of characid fishes (Teleostei: Ostariophysi: Characiformes), based on molecular sequence data. *Mol. Phylogenet. Evol.* 54, 498–511. doi: 10.1016/j.ympev.2009.08.026

Jesus, C. M., Galetti, P. M. Jr., Valentini, S. R., and Moreira-Filho, O. (2003). Molecular characterization and chromosomal localization of two families of satellite DNA in *Prochilodus lineatus* (Pisces, Prochilodontidae), a species with B chromosomes. *Genetica* 118, 25–32. doi: 10.1023/A:1022986816648

Jurka, J., Kapitonov, V. V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. (2005). Repbase update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* 110, 462–467. doi: 10.1159/000084979

Kent, W. J. (2002). BLAT—the BLAST-like alignment tool. *Genome Res.* 12, 656–664. doi: 10.1101/gr.229202

Kuhn, G. C. S., Küttler, H., Moreira-Filho, O., and Heslop-Harrison, J. S. (2012). The 1.688 repetitive DNA of *Drosophila*: concerted evolution at different genomic scales and association with genes. *Mol. Biol. Evol.* 29, 7–11. doi: 10.1093/molbev/msr173

Lanfredi, M., Congiu, L., Garrido-Ramos, M. A., Leis, M., Chicca, M., Rossi, R., et al. (2001). Chromosomal location and evolution of a satellite DNA family in seven sturgeon species. *Chromosome Res.* 9, 47–52. doi: 10.1023/A:1026739616749

Leclerc, G. M., Han, K., Leclerc, G. J., and Ely, B. (1999). Characterization of a highly repetitive sequence conserved among the North American Morone species. *Mar. Biotechnol.* 1, 122–130. doi: 10.1007/PL00011759

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187

López-Flores, I., and Garrido-Ramos, M. A. (2012). "The repetitive DNA content of eukaryotic genomes," in *Repetitive DNA*, ed. M. A. Garrido-Ramos (Basel: Karger Publishers), 1–28. doi: 10.1159/000337118

Macas, J., Kejnovský, E., Neumann, P., Novák, P., Koblížková, A., and Vyskot, B. (2011). Next generation sequencing-based analysis of repetitive DNA in the model dioceous plant *Silene latifolia*. *PLoS ONE* 6:e27335. doi: 10.1371/journal.pone.0027335

Martins, C., Ferreira, I. A., Oliveira, C., Foresti, F., and Galetti, P. M. Jr., (2006). A tandemly repetitive centromeric DNA sequence of the fish *Hoplias malabaricus* (Characiformes: Erythrinidae) is derived from 5S rDNA. *Genetica* 127, 133–141. doi: 10.1007/s10709-005-2674-y

McMurray, C. T. (1995). Mechanisms of DNA expansion. *Chromosoma* 104, 2–13. doi: 10.1007/BF00352220

Mestriner, C. A., Galetti, P. M. Jr., Valentini, S. R., Ruiz, I. R., Abel, L. D., Moreira-Filho, O., et al. (2000). Structural and functional evidence that a B chromosome in the characid fish *Astyanax scabripinnis* is an isochromosome. *Heredity* 85, 1–9. doi: 10.1046/j.1365-2540.2000.00702.x

Milani, D., and Cabral-de-Mello, D. C. (2014). Microsatellite organization in the grasshopper *Abracris flavolineata* (Orthoptera: Acrididae) revealed by FISH mapping: remarkable spreading in the A and B chromosomes. *PLoS ONE* 9:e97956. doi: 10.1371/journal.pone.0097956

Novák, P., Neumann, P., and Macas, J. (2010). Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC Bioinformatics* 11:378. doi: 10.1186/1471-2105-11-378

Novák, P., Neumann, P., Pech, J., Steinhaisl, J., and Macas, J. (2013). RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 29, 792–793. doi: 10.1093/bioinformatics/btt054

Oliveira, C., Avelino, G. S., Abe, K. T., Mariguela, T. C., Benine, R. C., Ortí, G., et al. (2011). Phylogenetic relationships within the speciose family Characidae (Teleostei: Ostariophysi: Characiformes) based on multilocus analysis and extensive ingroup sampling. *BMC Evol. Biol.* 11:275. doi: 10.1186/1471-2148-11-275

Oliveira, C., Foresti, F., and Hilsdorf, A. W. S. (2009). Genetics of neotropical fish: from chromosomes to populations. *Fish Physiol. Biochem.* 35, 81–100. doi: 10.1007/s10695-008-9250-1

Pagán, H. J. T., Macas, J., Novák, P., McCulloch, E. S., Stevens, R. D., and Ray, D. A. (2012). Survey sequencing reveals elevated DNA transposon activity, novel elements, and variation in repetitive landscapes among vesper bats. *Genome Biol. Evol.* 4, 575–585. doi: 10.1093/gbe/evs038

Pinkel, D., Straume, T., and Gray, J. W. (1986). Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc. Natl. Acad. Sci. U.S.A.* 83, 2934–2938. doi: 10.1073/pnas.83.9.2934

Plohl, M., Luchetti, A., Meštrović, N., and Mantovani, B. (2008). Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero) chromatin. *Gene* 409, 72–82. doi: 10.1016/j.gene.2007.11.013

Plohl, M., Meštrović, N., and Mravinac, B. (2012). "Satellite DNA evolution," in *Repetitive DNA*, ed. M. A. Garrido-Ramos (Basel: Karger Publishers), 126–152. doi: 10.1159/000337122

Pons, J., Juan, C., and Petitpierre, E. (2002). Higher-order organization and compartmentalization of satellite DNA PIM357 in species of the coleopteran genus *Pimelia*. *Chromosome Res.* 10, 597–606. doi: 10.1023/A:1020918803675

Robles, F., de la Herrán, R., Ludwig, A., Ruiz Rejón, C., Ruiz Rejón, M., and Garrido-Ramos, M. A. (2004). Evolution of ancient satellite DNAs in sturgeon genomes. *Gene* 338, 133–142. doi: 10.1016/j.gene.2004.06.001

Rossini, B. C., Oliveira, C. A. M., de Melo, F. A. G., de Araújo Bertaco, V., de Astarloa, J. M. D., Rosso, J. J., et al. (2016). Highlighting Astyanax species diversity through DNA barcoding. *PLoS ONE* 11:e0167203. doi: 10.1371/journal.pone.0167203

Ruiz-Ruano, F. J., Cuadrado, Á., Montiel, E. E., Camacho, J. P. M., and López-León, M. D. (2015). Next generation sequencing and FISH reveal uneven and nonrandom microsatellite distribution in two grasshopper genomes. *Chromosoma* 124, 221–234. doi: 10.1007/s00412-014-0492-7

Ruiz-Ruano, F. J., López-León, M. D., Cabrero, J., and Camacho, J. P. M. (2016). High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Sci. Rep.* 6:28333. doi: 10.1038/srep28333

Silva, D. M. Z. A., Daniel, S. N., Camacho, J. P. M., Utsunomia, R., Ruiz-Ruano, F. J., Penitente, M., et al. (2016). Origin of B chromosomes in the genus *Astyanax* (Characiformes, Characidae) and the limits of chromosome painting. *Mol. Genet. Genomics* 291, 1407–1418. doi: 10.1007/s00438-016-1195-y

Silva, D. M. Z. A., Pansonato-Alves, J. C., Utsunomia, R., Araya-Jaime, C., Ruiz-Ruano, F. J., Daniel, S. N., et al. (2014). Delimiting the origin of a B chromosome by FISH mapping, chromosome painting and DNA sequence analysis in *Astyanax paranae* (Teleostei, Characiformes). *PLoS ONE* 9:e94896. doi: 10.1371/journal.pone.0094896

Silva, D. M. Z. A., Pansonato-Alves, J. C., Utsunomia, R., Daniel, S. N., Hashimoto, D. T., Oliveira, C., et al. (2013). Chromosomal organization of repetitive DNA sequences in *Astyanax bockmanni* (Teleostei, Characiformes): dispersive location, association and co-localization in the genome. *Genetica* 141, 329–336. doi: 10.1007/s10709-013-9732-7

Smith, G. P. (1976). Evolution of repeated DNA sequences by unequal crossover. *Science* 191, 528–535. doi: 10.1126/science.1251186

Teacher, A. G. F., and Griffiths, D. J. (2011). HapStar: automated haplotype network layout and visualization. *Mol. Ecol. Resour.* 11, 151–153. doi: 10.1111/j.1755-0998.2010.02890.x

Thomaz, A. T., Arcila, D., and Ortí, G. (2015). Molecular phylogeny of the subfamily Stevardiinae Gill, 1858 (Characiformes: Characidae): classification and the evolution of reproductive traits. *BMC Evol. Biol.* 15:269. doi: 10.1186/s12862-015-0489-8

Ugarković, Đ., and Plohl, M. (2002). Variation in satellite DNA profiles—causes and effects. *EMBO J.* 21, 5955–5959. doi: 10.1093/emboj/cdf612

Utsunomia, R., Silva, D. M. Z. A., Ruiz-Ruano, F. J., Araya-Jaime, C., Pansonato-Alves, J. C., Scacchetti, P. C., et al. (2016). Uncovering the ancestry of B chromosomes in *Moenkhausia sanctaefilomenae* (Teleostei, Characidae). *PLoS ONE* 11:e0150573. doi: 10.1371/journal.pone.0150573

Vicari, M. R., Nogaroto, V., Noleto, R. B., Cestari, M. M., Cioffi, M. B., Almeida, M. C., et al. (2010). Satellite DNA and chromosomes in Neotropical fishes: methods, applications and perspectives. *J. Fish Biol.* 76, 1094–1116. doi: 10.1111/j.1095-8649.2010.02564.x

Weitzman, S. H., and Malabarba, L. R. (1998). "Perspectives about the phylogeny and classification of the Characidae (Teleostei: Characiformes)," in *Phylogeny and Classification of Neotropical Fishes*, eds L. R. Malabarba, E. R. Reis, R. P. Vari, Z. M. S. Lucena, and C. A. Lucena (Porto Alegre: Museu de Ciências e Technologia), 161–170.

Willard, H. F., and Waye, J. S. (1987). Hierarchical order in chromosome-specific human alpha satellite DNA. *Trends Genet.* 3, 192–198. doi: 10.1016/0168-9525(87)90232-0

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

Copyright © 2017 Utsunomia, Ruiz-Ruano, Silva, Serrano, Rosa, Scudeler, Hashimoto, Oliveira, Camacho and Foresti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 31 August 2017
doi: 10.3389/fgene.2017.00114

Triportheus albus Cope, 1872 in the Blackwater, Clearwater, and Whitewater of the Amazon: A Case of Phenotypic Plasticity?

José D. A. Araújo[1,2], Andrea Ghelfi[3] and Adalberto L. Val[1]*

[1] Laboratory of Ecophysiology and Molecular Evolution, National Institute of Amazonian Research, Manaus, Brazil, [2] Federal University of Amazonas, Manaus, Brazil, [3] Kazusa DNA Research Institute, Kisarazu, Japan

The Amazon basin includes 1000s of bodies of water, that are sorted according to their color in three types: blackwater, clearwater, and whitewater, which significantly differ in terms of their physicochemical parameters. More than 3,000 species of fish live in the rivers of the Amazon, among them, the sardine, *Triportheus albus*, which is one of the few species that inhabit all three types of water. The purpose of our study was to analyze if the gene expression of *T. albus* is determined by the different types of water, that is, if the species presents phenotypic plasticity to live in blackwater, clearwater, and whitewater. Gills of *T. albus* were collected at well-characterized sites for each type of water. Nine cDNA libraries were constructed, three biological replicates of each condition and the RNA was sequenced (RNA-Seq) on the MiSeq® Platform (Illumina®). A total of 51.6 million of paired-end reads, and 285,456 transcripts were assembled. Considering the FDR $\leq$ 0.05 and fold change $\geq$ 2, 13,754 differentially expressed genes were detected in the three water types. Two mechanisms related to homeostasis were detected in *T. albus* that live in blackwater, when compared to the ones in clearwater and whitewater. The acidic blackwater is a challenging environment for many types of aquatic organisms. The first mechanism is related to the decrease in cellular permeability, highlighting the genes coding for claudin proteins, actn4, itgb3b, DSP, Gap junction protein, and Ca^{2+}-ATPase. The second with ionic and acid-base regulation [rhcg1, slc9a6a (NHE), ATP6V0A2, Na^{+}/K^{+}-ATPase, slc26a4 (pedrin) and slc4a4b]. We suggest *T. albus* is a good species of fish for future studies involving the ionic and acid-base regulation of Amazonian species. We also concluded that, *T. albus*, shows well defined phenotypic plasticity for each water type in the Amazon basin.

Keywords: Rio Negro, Tapajós River, Solimões River, differential expression, RNA-Seq, acidic pH, ionic regulation

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Luciana Rodrigues De Souza-Bastos, Institutos LACTEC, Brazil
Rafael Mendonça Duarte, Universidade Estadual Paulista Júlio de Mesquita Filho, Brazil
Rui Rosa, Universidade de Lisboa, Portugal

****Correspondence:***
José D. A. Araújo
deneyaraujo@gmail.com

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

Received: *30 June 2017*
Accepted: *17 August 2017*
Published: *31 August 2017*

Citation:
Araújo JDA, Ghelfi A and Val AL (2017) Triportheus albus Cope, 1872 in the Blackwater, Clearwater, and Whitewater of the Amazon: A Case of Phenotypic Plasticity? Front. Genet. 8:114. doi: 10.3389/fgene.2017.00114

INTRODUCTION

The rivers of the Amazon are interconnected to the central channel, the Amazon River, sheltering a rich ichthyofauna and allowing the entire Amazon basin to be linked through its waters. The river connection make possible for species to migrate among the rivers of the region (Val and Almeida-Val, 1995). However, these rivers contain different water types, due to the geographical location of each river and the materials that are deposited on their beds (Sioli, 1984; Konhauser et al., 1994).

In many cases, the physicochemical parameters of the waters govern the selection of species which inhabit them (Val and Almeida-Val, 1995). This selection depends on the capacity that species possess to get adapted to the environmental conditions to which they are exposed (Evans et al., 2005; Wood et al., 2007). In this context, phenotypic plasticity has recently gained significance in transcriptome analyses and can be defined as the ability of a genotype to produce variable phenotypes under different environmental conditions. Thus, the phenotypic plasticity supporting physiological adjustments necessary for the transition between environments is valuable for many species (Gibbons et al., 2017).

The waters of the Amazon constitute a singular place to analyze that issue, as water bodies containing blackwater, clearwater, and whitewater (Sioli, 1984) are interconnected.

Indeed, physicochemical parameters are closely related to their characteristic colors. Three major rivers have these water profiles: the Negro River (blackwater), considered one of the most challenging aquatic environments for aquatic species, due to its natural acidic water and low level of ions. The presence of significant concentrations of dissolved organic matter (DOM), produced by plant decomposition because of the seasonal flooding of part of the forest (Sioli, 1984; Ertel et al., 1986), is responsible for releasing humic and fulvic acids in the water, which in turn, have carboxylic groups (−COOH) and hydroxylic (−OH) groups in their structures. These dissociate themselves and release H^+ ions into the water, thus reducing the pH in the environment (Queiroz et al., 2009). Several studies have been developed in an attempt to understand how a river with such an enormous environmental challenge houses a significant diversity of fish, estimated at, approximately, 1,000 species (Val and Almeida-Val, 1995; Gonzalez et al., 2002; Wood et al., 2014; Duarte et al., 2016).

The Solimões/Amazonas river, the primary representative whitewater, is the largest freshwater river on the planet. It has its distinctive color due to the high quantity of material in suspension derived from sediments from the Andes. These sediments are transported in large volumes along the whole river, being deposited on the banks, forming sandbanks with the clayey material known in the region of the várzea (the area on the riverbanks that gets flooded during the flood season) (Furch, 1984). The Solimões River is richer than the Rio Negro in dissolved ions and its electrical conductivity is the highest one when compared to clearwater and blackwater, with an average of 98.8 μS/cm (Küchler et al., 2000). The predominant ions in its waters are Na^+, K^+, Mg^{2+}, and Ca^{2+}. The pH of whitewater is close to neutrality (6.6) (Sioli, 1984; Gaillardet et al., 1997).

The Tapajós River is known for its crystalline, slightly greenish waters. Its waters contain little clayey sediment, due to the drainage that the river makes on the soil of pre-Cambrian origin (Sioli, 1984). Just as the Negro River, it has areas of igapós (flooded forests) on its banks, depending on the seasonal cycle. The pH is around 6.5 and it has low conductivity (14.4 μS/cm), with a reduced amount of suspended material. Whereas the physicochemical parameters of its waters are referenced to as an intermediary between blackwater and whitewater rivers (Duncan and Fernandes, 2010), the clearwater is the one with the highest annual variation on its physicochemical parameters in the Amazon basin, mainly as regards to the water pH (Junk and Furch, 1980). Therefore, both the Tapajós River and the Solimões River have a pH close to neutrality, while the Negro River has acidic waters (Sioli, 1984).

Depending on the annual water fluctuation (seasonal ebb or river flooding), the physicochemical parameters can be altered. Observing the environmental conditions of different types of waters makes the challenges that these environments impose on ichthyofauna evident (Furch, 1982).

These environments are home to the most diverse ichthyofauna in the world. Many species simultaneously inhabit two of these environments (reviewed by Val and Almeida-Val, 1995). Few, however, have developed biological mechanisms to live simultaneously in the three types of water in the Amazon. Among these, we highlight the species *Triportheus albus*, popularly known in the region as sardine. This species is often found in all three types of water in the Amazon basin (INCT/ADAPTA Project Report 2012–2013). The understanding of how this species responds to the different environmental conditions prevailing in the three types of Amazonian environments was the primary factor that challenged us in the present study.

Considering the environmental and physicochemical characteristics of the waters in which the presence of the *T. albus* occurs, the purpose of our study was to analyze if the gene expression of *T. albus* is determined by the different types of water. Our hypothesis is that *T. albus* presents phenotypic plasticity to live in blackwater, clearwater, and whitewater.

MATERIALS AND METHODS

The experimental procedures were approved by the Animal Use Ethics Committee of the Brazilian National Institute for Research of the Amazon (CEUA-INPA) (Protocol 026/2015). The permit for the collection of the biological material to carry out the research was authorized by the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA/SISBio), number 29837-8/2015.

Collection of Samples

Specimens of *T. albus* were captured in their natural environments, covering the different water types of the Amazon basin in expeditions carried out in July and August 2015. See Supplementary Table S1 for length and mass of the analyzed fish. Blackwater specimens were captured on the banks of the Negro River, in the Anavilhanas Archipelago (2°43′11.8″S, 60°45′19.1″W). Fish collection in whitewater was carried out on the banks of the Solimões River, near the island of Marchantaria (3°14′47.0″S, 59°57′23.3″W). The collection of specimens in clearwater was carried out on the banks of the Tapajós River (2°48′46.3″S, 55°02′21.2″), for information on fish in Supplementary Table S1.

Gills were the analyzed tissue because they are the primary site for osmotic regulation, excretion of ammonia, as well as an important site for gas exchange and regulation of body fluid pH.

Thus, nine *T. albus* specimens were captured using fishing line and hook, three individuals for each environmental condition. After captured, they were sacrificed by severing their spinal cord, the gills were removed and immediately stored in RNALater® (Ambion®), and kept at 4°C, until arrival at the Laboratory of Ecophysiology and Molecular Evolution (LEEM), at INPA's facilities in Manaus, Amazonas, Brazil. Water physicochemical parameters were measured at the same collection site of the organisms between 5 and 7 pm. During the period of collection, the seasonal "river flooding" period prevailed. The pH values were measured using a pHmeter (Micronal model B374) and dissolved oxygen and conductivity was measured with an oxygen meter (YSI, model 55/12 FT).

RNA Extraction and Library Construction

Total RNA extraction was performed individually for each of the nine *T. albus* gill tissue samples using the TRIzol® reagent protocol (Invitrogen™). After that, 30 μL of ultrapure water (Nuclease-Free) was added to each bullet tube containing the RNA extract, which was stored at −80°C until the time of analyses. The quantification of extracted RNA was determined using a BioAnalyzer Agilent 2100 (Agilent Technologie®). The procedures for building the libraries were performed according to the TruSeq RNA Sample Prep v2 LS protocol (Illumina®), observing the manufacturer's recommendations.

The mRNA was isolated from the total RNA using magnetic oligo (dT) beads that were bound to the poly (A) tail of the mRNA. Then, the first cDNA strand was synthesized using reverse transcriptase and random primers. The second cDNA chain was immediately synthesized, using the enzymes DNA Polymerase I and RNAse H. It was then added to the end 3′ of the fragments of a single A nucleotide (adenine). The adapters were bound to these fragments, the PCR was then performed for the enrichment of the libraries.

Finally, the quality and quantification of the constructed libraries were analyzed using the BioAnalyzer Agilent 2100 (Agilent Technologie®) and Real-Time PCR 7500 (Applied Biosystems®). Three biological replicates of samples from each environmental condition were sequenced on the MiSeq® platform (Illumina®) in three sequencing runs – a paired-end run (2 × 150 cycles) and two paired-end runs (2 × 250 cycles). The data generated in this study are stored in the National Center for Biotechnology Information/Sequence Read Archive (NCBI-SRA), project number PRJNA391967.

Quality Control and Reassembly

The bioinformatics analyses were performed at the Bioinformatics Laboratory of LEEM in the facilities of INPA. The FastQC program [v.0.11.3] (Andrews, 2010) was used to analyze the quality of sequenced reads. Treatment of the 5′ and 3′ ends of the low quality reads (Q-score ≤ 20) and filtering of reads of less than 50 bp (base pairs) (≤50) were performed using the program Trimmomatic [v. 0.33] (Bolger et al., 2014).

The reassembly of the transcriptome was performed with the Trinity program [version 2.1.1] (Grabherr et al., 2011). In addition, programs that assisted Trinity to assemble the transcriptome and to calculate the abundance of transcripts were used, among them, the Bowtie2 [v. 2.2.6] (Langmead and Salzberg, 2012), RSEM [v. 1.2.19] (Li and Dewey, 2011), and R/Bioconductor packages [v. 3.1] (Gentleman et al., 2004).

TABLE 1 | Physicochemical parameters of the waters of the Negro River (blackwater), Tapajós River (clearwater) and Solimões River (whitewater).

Characteristics	Types of water		
	Negro River	Tapajós River	Solimões River
Color	Black	Clear	White
Conductivity	10.5 ± 0.5 μS/cm^a	17.3 ± 0.3 μS/cm^b	74.2 ± 0.1 μS/cm^c
Oxygen	3.1 ± 0.5 mg/L^a	6.6 ± 0.1 mg/L^b	2.4 ± 0.2 mg/L^c
Temperature	29.5 ± 0.1°C^a	31.2 ± 0.4°C^b	28.3 ± 0.1°C^c
pH	4.6 ± 0.1^a	6.0 ± 0.4^b	6.3 ± 0.1^c
DOC*	8.75 mg/L^a	<0.4 mg/L^b	3.86 mg/L^c

*Values expressed as mean ± SD; N = 5 in each condition. Different letters indicate significant differences (P ≤ 0.05), using one-way ANOVA and Tukey test. *DOC values from Holland et al. (2017).*

Differential Expression Analysis and Gene Annotation

The analysis of differentially expressed genes (DEGs) was performed with the R/Bioconductor package, edgeR [v. 3.8.6] (Robinson et al., 2010), with a False Discovery Rate (FDR) ≤ 0.05. After the transcript quantification, the data generated by the RSEM, served as the input to the edgeR, when fold change was calculated. The DEGs were annotated with the BLASTx program [v. 2.3.1] (Altschul et al., 1997), through consultation of the database of Uniprot/TrEMBL proteins (class Actinopterygii)[1], with *E*-value 1.0E-5. After functional annotation, the genes were classified according to their gene ontologies (GO), using a script developed at the Multidisciplinary Support Center of the Federal University of Amazonas (CAM/UFAM).

RESULTS

Physicochemical Parameters of the Waters

The physicochemical parameters are quite different, with a specific water pattern for each environmental condition (**Table 1**). Conductivity is one of the reflexes of the differences between the waters, being directly connected to the quantity of dissolved ions. It presents low conductivity in waters that are poor in ions such as the Negro River (10.5 μS/cm), with a moderate increase in the Tapajós River (17.3 μS/cm) and very high in the Solimões River (74.2 μS/cm).

Data recently published by Holland et al. (2017), who, for their part, carried out water collections in the same period of our collections of the specimens (July, 2015), in locations close to our collection points, corroborate with additional information on some parameters Physical-chemical properties of water. Among them, we highlight the concentration of dissolved organic

[1]www.uniprot.org

carbon (DOC) for the Rio Negro of 8.75 mg/L, Solimões River (3.86 mg/L) and Tapajós River (<0.4 mg/L). We also highlight the difference concerning the pH of the respective rivers: acidic water from the Negro River with a pH of 4.6, being as low as a pH 3, the clearwater of the Tapajós River with a pH of around 6.0 and the whitewater of the Solimões River with a pH around 6.3, both close to neutrality. The blackwater is low in ions, while it has a high concentration of H^+ (2.5×10^{-5} mol/L), derived from the humic and fulvic acids of the plant decomposition that occurs on the Rio Negro bed (reviewed by Val and Almeida-Val, 1995). On the other hand, the Tapajós River and the Solimões River have a low concentration of DOC, with a concentration of H^+ equal to 1.0×10^{-6} mol/L and 5.0×10^{-7} mol/L, respectively. This represents a concentration of H^+ in the Negro River of approximately 25 times greater than that observed in the Tapajós River and 50 times greater than that present in the Solimões River.

Sequencing and Quality Control

Nine *T. albus* cDNA libraries were constructed, three biological replicates for each environmental condition: blackwater, clearwater, and whitewater. The RNA-Seq Sequencing in the MiSeq® Platform (Illumina®) produced about 51.6 million paired-end reads. In the quality control and filtering of the raw data, the bases were removed from the ends of the reads with Q-Score ≤ 20 and the reads with a size of less than 50 pb were excluded, resulting in a total of 45.8 million paired-end reads that were saved. Only 11.17% of the total reads sequenced were discarded (Supplementary Table S2).

Assembly *De Novo* and Differential Expression Analysis

The reads resulting from the pre-processing of the data were grouped and the transcriptome assembly *de novo* was performed. A total of 285,456 transcripts were assembled, using the Bruijn graph analysis (Grabherr et al., 2011), from which the contigs with average lengths of 584.93 pb were derived. The value of N50 was 751 pb, totaling 166,972,252 bases assembled. The contigs were aligned and the abundance of transcripts quantified in each environmental condition. This was done for all biological replicates. The analysis of the differential expression of *T. albus* among environmental conditions was performed using FDR ≤ 0.05. In addition to FDR, we considered fold change ≥ 2. A total of 13,754 genes were differentially expressed in the three environmental conditions. In blackwater conditions versus whitewater conditions, 3,956 DEGs were found, 265 up-regulated ones (6.7%), 3,691 down-regulated ones (93.3%). In clearwater versus whitewater conditions, only 30 differentially expressed transcripts, 2 up-regulated ones (6.7%) and 28 down-regulated ones (93.3%) were found. When it comes to blackwater versus whitewater conditions, 9,768 DEGs, 4,318 up-regulated ones (43.2%) and 5,550 down-regulated ones (56.8%) were found (**Figure 1**).

Transcriptome Annotation of *T. albus*

Using BLASTx, with E-value of 1.0E-5, through consultation to the database Uniprot/TrEMBL (Actinopterygii class), 33,739 genes were identified. The top hits returned by a BLAST search were associated with the following species of fish: *Astyanax mexicanus* (43%), *Danio rerio* (14%), *Oncorhynchus mykiss* (7%), *Poeciliopsis prolifica* (6%), *Ictalurus punctatus* (5%) and other species (25%).

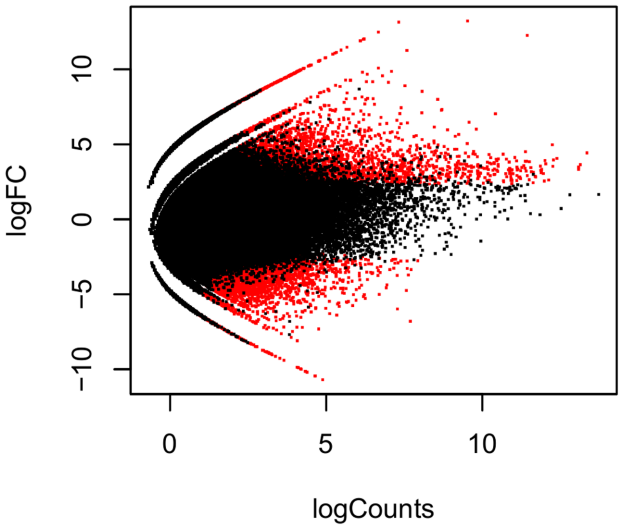


FIGURE 1 | Differentially expressed genes of *Triportheus albus* in blackwater versus clearwater conditions. The data are shown in logarithmic scale (base 2), considering the fold change of the expression versus the mean of the level of expression between the conditions analyzed. The red points above zero on the y-axis represent the up-regulated transcripts, while the ones which are below zero, represent down-regulated transcripts.

Differentially expressed genes were grouped according to their GO. In blackwater versus whitewater conditions, 3,206 terms were annotated. Out of these, 159 up-regulated ones (Biological Process – PB: 54, Molecular Function – FM: 50 and Cell Component – CC:55) (**Figure 2A**) and 3,047 down-regulated ones (PB: 1,100, FM: 1,064 and CC: 883) (**Figure 2B**).

In blackwater versus clearwater conditions, 9,566 terms were annotated – 5,938 up-regulated ones (PB: 2,476, FM: 2,077 and CC: 1,385) (**Figure 3A**) and 3,628 down-regulated ones (PB: 2,476, FM: 2,077 and CC: 1,385) (**Figure 3B**). When it comes to clearwater versus whitewater conditions, only 21 terms were annotated (9 in PB, 9 in FM and 3 in CC, all of them were up-regulated).

Common Terms in "Blackwater versus Clearwater" and "Blackwater versus Whitewater"

In the DEGs analysis, a greater number of transcripts expressed in the conditions of blackwater versus clearwater and blackwater versus whitewater were observed. We then verified whether such conditions share the same gene ontology terms. Thus, we located 1,551 up-regulated terms in common under both conditions. Among these, we highlight some terms such as integral component of membrane, calcium ion binding, plasma membrane, cell adhesion, cytoskeleton, myosin

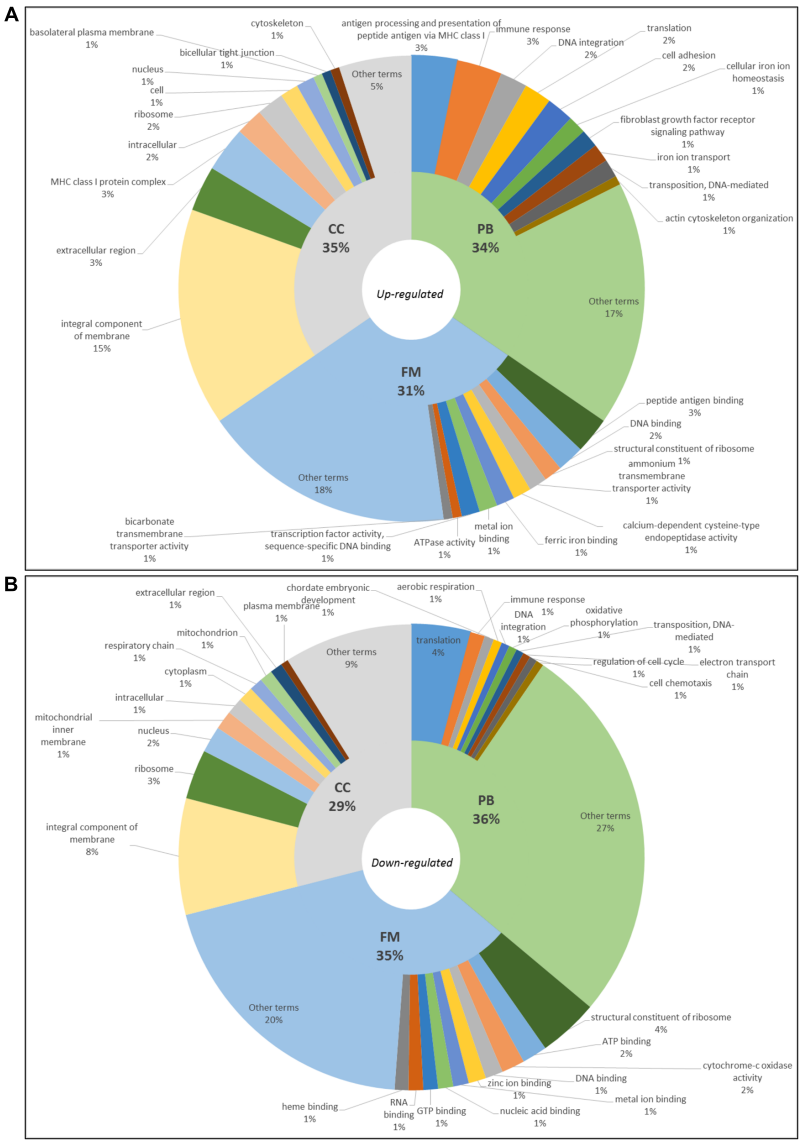


FIGURE 2 | Gene ontology of top10 up-regulated **(A)** and top10 down-regulated **(B)** terms in blackwater versus whitewater conditions for *Triportheus albus*.

complex, transporter activity, ammonium transmembrane transporter activity, bicellular tight junction and ATPase activity (Supplementary Figure S1).

After that, we analyzed the genes linked to the annotated terms. We observed that the genes identified were related to a fundamental role in cellular permeability, ionic regulation and acid-base (**Table 2**). Among these, we found the claudin genes, actn4 (actinin, alpha 4), itgb3b (integrin beta) (paracellular junction and cell adhesion), rhcg1 (ammonium transporter Rh type C 1), slc9a6a (sodium/hydrogen exchanger) (NHE), ATP6V0A2 (V-type proton ATPase subunit a), Na^+/K^+-ATPase (sodium/potassium-transporting ATPase subunit alpha) (ionic regulation and acid-base) nr3c1 (glucocorticoid receptor), prlra (prolactin receptor a) (endocrine response) (**Table 2**).

DISCUSSION

The diversity of rivers of the Amazon basin, besides containing the largest number of species of freshwater fish, represents geographical barriers by means of its restricted water profile within each river. This requires biological adjustments of species living in these different environments (De Pinna, 2006). Many fish species live in an environment close to neutrality (Val and Almeida-Val, 1995). However, many fish species live in waters with low pH may that influence several physiological processes (Wood et al., 1998; Matsuo and Val, 2007). The species *T. albus* is one of the few species that differs from the innumerable other species of fish in the region, since it occurs in all three types of waters (blackwater, clearwater, and whitewater) in the Amazon Basin.

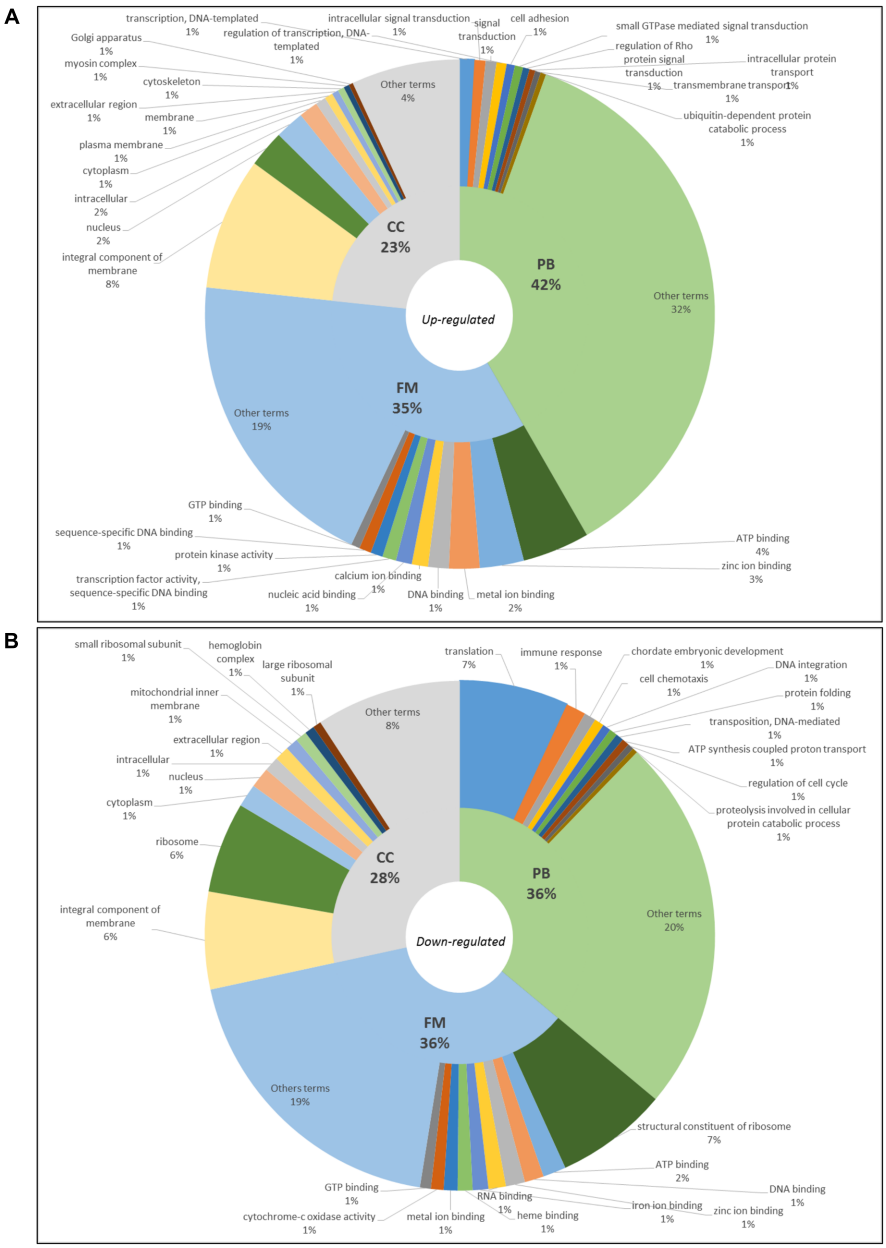


FIGURE 3 | Gene ontology of top10 up-regulated **(A)** and top10 down-regulated **(B)** terms in blackwater versus clearwater conditions for *Triportheus albus*.

The work of Gibbons et al. (2017), with the fish species *Gasterosteus aculeatus*, which survives in freshwater and saltwater environments, shows an excellent approach to how the gene expression of fish responds to the most diversified environments. Although living in environments that are different from those of the species studied here, the authors observed that the environment in which the organism was being exposed directly influenced the gene expression pattern, a situation corroborated by the present study. They have also described the occurrence of phenotypic plasticity to respond to the environment to which they are exposed and to maintain the homeostasis under the prevailing conditions, a very similar situation to the one found here for *T. albus*.

To date, there is no information on the transcriptome profile of another species that is able to survive in all three types of water in the Amazon. In this context, our study is unique since we present the analysis of the transcriptome of *T. albus* living in all three types of water in the Amazon.

We observed 13,754 differentially expressed transcripts in all three environmental conditions. The DEGs were grouped by their functional categories according to the GOs (Ashburner et al.,

TABLE 2 | Common genes of *Triportheus albus* in blackwater versus clearwater conditions and blackwater versus whitewater conditions, candidates involved in the response to low pH.

Contig ID	*E*-value	LogFC	Gene/Proteína	Annotation (GO)
TR35387| c0_g1_i1	5.00E-157	6.65	Ca^{2+}-ATPase	Calcium ion transport (PB) and calcium-transporting ATPase activity (FM)
TR31771| c0_g1_i1	2.00E-11	6.33	Claudin	Bicellular tight junction (CC) and cell junction (CC)
TR44997| c0_g2_i1	0.00E-00	6.18	DSP	cell adhesion molecule binding (FM) and cell junction (CC)
TR53695| c1_g2_i1	0.00E-00	4.56	actn4	Bicellular tight junction assembly (PB) and protein localization to bicellular tight junction (PB)
TR49051| c2_g2_i1	0.00E-00	4.55	itgb3b	Cell-substrate junction assembly (PB) and cell adhesion (PB)
TR58189| c5_g2_i3	8.00E-158	4.11	Gap junction protein	Gap junction (CC) and cell junction (CC)
TR57503| c3_g54_i1	8.00E-48	5.86	slc26a4	Bicarbonate transport (PB) and chloride transport (PB)
TR39575| c0_g1_i2	0.00E-00	4.71	slc9a6a (NHE)	Sodium ion transport (PB) and hydrogen ion transmembrane transport (PB)
TR53402| c0_g3_i1	0.00E-00	4.04	slc4a4b	Bicarbonate transport (PB) and anion transmembrane transporter activity (FM)
TR42005| c0_g1_i1	8.00E-173	4.34	Na^{+}/K^{+}-ATPase	Sodium ion transport (PB) and sodium:potassium-exchanging ATPase activity (FM)
TR54587| c4_g2_i6	0.00E-00	3.82	rhcg1	Ammonium transport (PB) and ammonium transmembrane transporter activity (FM)
TR42033| c0_g1_i1	0.00E-00	3.16	ATP6V0A2	Ion transport (PB) and hydrogen ion transmembrane transporter activity (FM)
TR46875| c0_g2_i1	0.00E-00	4.73	Prlra	Prolactin signaling pathway (PB) and prolactin receptor activity (CC)
TR33864| c0_g2_i2	0.00E-00	3.45	nr3c1	Sodium ion homeostasis (PB) and regulation of ion homeostasis (PB)

2000). Thus, we selected the top10 terms, both up-regulated and down-regulated, from each functional category (PB, FM and CC).

In blackwater versus whitewater conditions and blackwater versus clearwater conditions, GO enriched up-regulated terms are involved mainly with membrane components, active transport of ions across the membrane, cytoskeletal/cell adhesion change, and synthesis of proteins (**Figures 2A**, **3A**). These functional categories show responses to exposure to the acidic and low ion environment, as the blackwater, contrasting with whitewater and clearwater that have a pH close to neutrality (6.0 and 6.4 respectively) and are relatively richer in ions. These responses trigger cellular processes aimed at maintaining the body's homeostasis in relation to the environment to which they are exposed. In general, the response is due to the increase in tightness of the paracellular junctions and active transport through membrane proteins (Bonga et al., 1990; Evans et al., 2005; Kumai and Perry, 2011).

In contrast, the down-regulated terms are related to reduced protein synthesis, embryonic development, cell cycle regulation, membrane components, mitochondria, and the respiratory chain (**Figures 2B**, **3B**). These repressed categories may indicate a metabolic readjustment, because in addition to the essential mechanisms of the organism, it has to maintain the homeostasis under the environmental challenges to which it is exposed (Li et al., 2016). In the gene expression pattern, both in blackwater versus whitewater conditions, as in blackwater versus clearwater conditions, up-regulated genes were mainly involved with mechanisms to maintain the body's osmotic and ionic homeostasis (**Table 2**).

Thus, some cellular mechanisms may have been deactivated/reduced to reduce energy expenditure (Li et al., 2016). Vidal-Dupiol et al. (2013) and Logan and Buckley (2015) also observed that in order to maintain the vital functions of the organism, some genes can be readjusted until the establishment of the homeostatic balance. However, future studies should analyze the reaction of repressed genes when they are exposed to the environment that is predominant in the Negro River.

In clearwater versus whitewater conditions, there was no significant difference. Only 30 differentially expressed transcripts were found and 21 terms were annotated. Such terms are related to the immune response, hemoglobin complex and ribosomal RNA. Given the small difference, considering the whole transcriptome, we concentrated on the analyses for contrasts of blackwater versus whitewater, and blackwater versus clearwater, which presented greater difference with regards to gene expression in this study.

According to the terms annotated in common, both blackwater versus whitewater conditions and blackwater versus clearwater conditions (Supplementary Figure S1) the genes of the referred terms were located, and it was found that these genes could be linked to the response to the acidic environment. Thus, we selected 14 genes involved in such a response (**Table 2**). We have grouped these genes of *T. albus* into two main response mechanisms to low pH.

The first mechanism is related to the modulation of the branchial epithelium (paracellular junctions - JPs), highlighting the genes coding claudin proteins, actn4 (actinin, alpha 4), itgb3b (integrin beta), DSP (desmoplakin), Gap junction protein, and Ca^{2+}-ATPase (calcium-transporting ATPase) (**Table 2**). Several studies have shown differentiated responses in some Amazon species exposed to low pH (Gonzalez and Wilson, 2001; Matsuo and Val, 2007; Wood et al., 2014). One of the fundamental characteristics to maintain the homeostasis in the acidified waters of the Negro River is the increase of the stiffness of JPs, avoiding excessive losses of Na^+ and Cl^- to the environment (Wood et al., 1998; Gonzalez et al., 2002; Matsuo and Val, 2007). Paracellular junctions represent one mode of cell-to-cell adhesion in epithelial or endothelial cell sheets, forming continuous seals around cells. The up-regulated gene claudin, together with others found in this study as DSP, itgb3b and Gap junction, encode for the primary sealing agents of JPs (Kumai and Perry, 2011; Kwong and Perry, 2013). This mechanism, supported by many genes had not been previously reported, as this mechanism is found in Amazonian cichlids, not in characids (Gonzalez and Wilson, 2001; Duarte et al., 2013).

It is worth noting that blackwater is rich in H^+, and this could in some way affect the JPs negatively, increasing permeability, due to the important role of Ca^{2+} in cell adhesion (McDonald, 1983; Freda et al., 1991). However, blackwater have a unique characteristic that differs them from the others. It is rich in DOM with a high concentration of DOC (Walker, 1995; Johannsson et al., 2016). An important function of DOC in acidic water is the buffering capacity against damaging effects caused by low pH (Holland et al., 2012).

Campbell et al. (1997) showed that DOC can bind directly to the biological membrane of gill cells and alter the permeability of the cell membrane. In addition to that, Wood et al. (2011) pointed out an important role of DOC on JPs, stating that DOC could act with similar functions to Ca^{2+} in JPs, reducing paracellular permeability. This statement about the protective role of DOC has recently been confirmed by Duarte et al. (2016). Duarte and collaborators observed that in Ca^{2+} poor waters such as the ones in the Negro River, DOC can act with functions similar to Ca^{2+} functions to protect the integrity of JPs. Thus, our data corroborate the information of Duarte et al. (2016), showing that even in Ca^{2+} poor waters, the genes continued to be expressed so as to maintain the integrity of JPs and, consequently, to maintain the ion balance allowing survival in waters that are challenging for many aquatic species.

The second mechanism related to the up-regulated genes encompasses ion and acid-base regulation [rhcg1 (ammonium transporter Rh type C 1), slc9a6a (sodium/hydrogen exchanger) (NHE), ATP6V0A2 (V-type proton ATPase subunit a) Na^+/K^+-ATPase (sodium/potassium-transporting ATPase subunit alpha), slc26a4 (pedrin), slc4a4b (anion exchange protein) (**Table 2**)]. One of the essential functions to maintain homeostasis in fish is ammonia excretion (NH_3^+), mainly through the gills (Wood et al., 2007). Several studies have shown that the organism exposed to an acidic environment can use the NH_3^+ excretion function to capture Na^+ from the environment (Kumai and Perry, 2011; Wright and Wood, 2012; Wood et al., 2014). This information was questioned taking thermodynamics into account (Parks et al., 2008). However, Kumai and Perry (2011) observed that the excretion of ammonia through the gills increased Na^+ uptake through the NHE complex. This interaction was associated with the presence of rhcg1 glycoprotein responsible for ammonium dissociation (NH_4^+) in NH_3^+ and H^+ creating a favorable microenvironment for the transport of NH_3^+ and H^+, through rhcg1 and NHE, respectively (Wright and Wood, 2012; Ito et al., 2013). These genes were up-regulated in the analyzed speciemens *T. albus* from Rio Negro.

In addition to the rhcg1 and slc9a6a (NHE), we also observed the expression of ATP6V0A2 (H^+-ATPase). Like the rhcg1 and NHE proteins, H^+-ATPase has been the focus of past studies, since it is responsible for providing an intracellular environment that favors Na^+ uptake during low pH exposure through the active H^+ extrusion from the cell (Lin et al., 2006; Chang et al., 2009).

A favorable electrochemical environment is required for a continuous ion regulation. In this study, we also observed an up-regulation of the gene encoding Na^+/K^+-ATPase. This was already expected, since organisms exposed to acidic environments perform Na^+ uptake through the mechanism described above. Thus, for the Na^+ uptake through the NHE exchanger, Na^+/K^+-ATPase is vital, since intracellular Na^+ in excess could disturb ion regulation (Marshall, 2002; Evans et al., 2005; Wood et al., 2014). In addition to the control of cellular permeability and ion regulation, it is necessary to maintain the acid-base, in view of the natural tendency of Cl^- loss to the acidic environment (Evans et al., 2005). We observed the up-regulated expression of the slc26a4 and slc4a4b genes. These genes are involved in the control of intracellular pH. This information corroborates Bayaa et al. (2009) and Perry et al. (2009) proposing that the uptake of Cl^- from the environment would occur through families of exchange proteins, such as slc26.

However, for this exchange to be possible, an intracellular HCO_3^- molecule is required for the coupling of the Cl^-/HCO_3^- exchanger (Evans et al., 2005). Studies have shown that carbonic anhydrase is responsible for the supply of internal HCO_3^- through the hydration of CO_2 (Lin et al., 2008; Gilmour and Perry, 2009). In our data, we did not find the differential expression of carbonic anhydrase. However, we can verify its performance through the by-product (HCO_3^-), which is being used in the acid-base balance, through the gene expression for the protein slc26a4 that acts on the apical membrane of the cell by exchanging Cl^-/HCO_3^-,

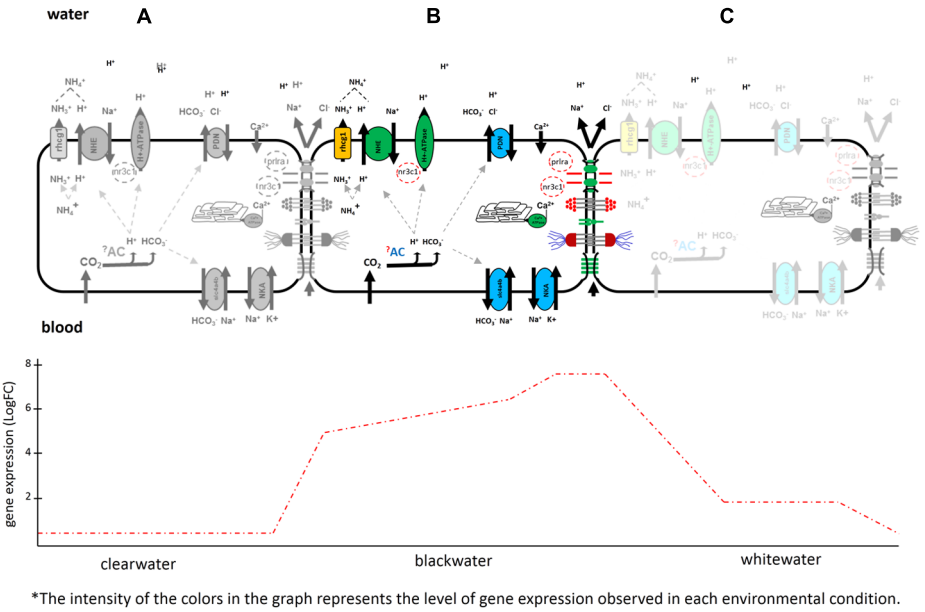


FIGURE 4 | Schematic representation of phenotypic plasticity observed for *Triportheus albus*. In clearwater there was almost no expression **(A)**. Following the trajectory of the graph, there was a high differential expression of the genes found in animals exposed to blackwater **(B)**. In whitewater, the genes related to ionic regulation was reduced and the modulation of the branchial epithelium (paracellular junctions) completely disconnected **(C)**.

and also for slc4a4b, a protein located in the basolateral membrane, which uses HCO_3^- and Na^+ to transport HCO_3^- to the blood, keeping the internal pH balanced (Evans et al., 2005).

The mechanisms described above are essential to maintain the homeostasis of the organism when exposed to acidic environments, such as the waters of the Negro River. For these mechanisms to be triggered, many authors have reported the importance of endocrine responses in fish (as reviewed by Breves et al., 2014; Kwong et al., 2014). We highlighted the genes nr3c1 (glucocorticoid receptor) and prlra (prolactin receptor a) (**Table 2**), both involved in hormonal responses, cortisol and prolactin, respectively. According to Kwong and Perry (2013), these hormones contribute to reduce epithelial permeability, avoiding excessive losses of Na^+ and Cl^- to the environment. They can also promote the reabsorption of these ions (Flik and Perry, 1989; Kelly and Wood, 2001; Kumai et al., 2012), and increase H^+-ATPase activity (Lin and Randall, 1993).

The ability of *T. albus* to modulate such mechanisms demonstrates a well-developed phenotypic plasticity system for this species. The potential mechanisms described here, based on the gene expression, demonstrate the potential of *T. albus* to survive the challenges presented by the different types of water in the Amazon basin. Depending on the environmental condition to which it is exposed, the genotype of this species can be differentially transcribed (**Figure 4**). The review by Schneider and Meyer (2017) highlights the mechanisms of differential expression. Analyzing **Figure 4**, in clearwaters (A) the activation level of ion and acid-base response mechanisms (genes coding for respective proteins involved) are almost null when compared to animals collected in blackwaters (B). In blackwater animals, every expressed mechanism for the maintenance of ion homeostasis is well characterized. In whitewater animals, these mechanisms are being expressed in an intermediate form (C), when compared to clearwater and blackwater conditions.

Physiological responses of Amazonian fish have been better understood over the years (Matsuo and Val, 2007; Duarte et al., 2013; Wood et al., 2014). The use of new techniques of molecular biology and bioinformatics resources, has increased the knowledge of how some species of fish respond to adverse environmental conditions (Lemgruber et al., 2013; Jesus et al., 2016; Prado-Lima and Val, 2016). In this context, we could observe in the present study that the *T. albus* species responds differently according to the exposed environment (**Figure 5**). When analyzing the heatmap graph, we verified two clusters of differential gene responses, related to the two extremes of aquatic environments.

Although literature emphasizes that clearwater is classified as intermediate between blackwater and whitewater (Sioli, 1984; Duncan and Fernandes, 2010), we can observe that *T. albus* presents a larger set of adjustments in gene expression in the Tapajós River (clearwater) and in the Negro River (blackwater). Thus, through the gene response observed in this species, we can conclude that clearwater and blackwater are two distinct extremes, and they present a greater environmental challenge for the survival of the organism. On the other hand, we verified that the whitewater was the one that presented an intermediate level, requiring less quantitative adjustment of the gene response to

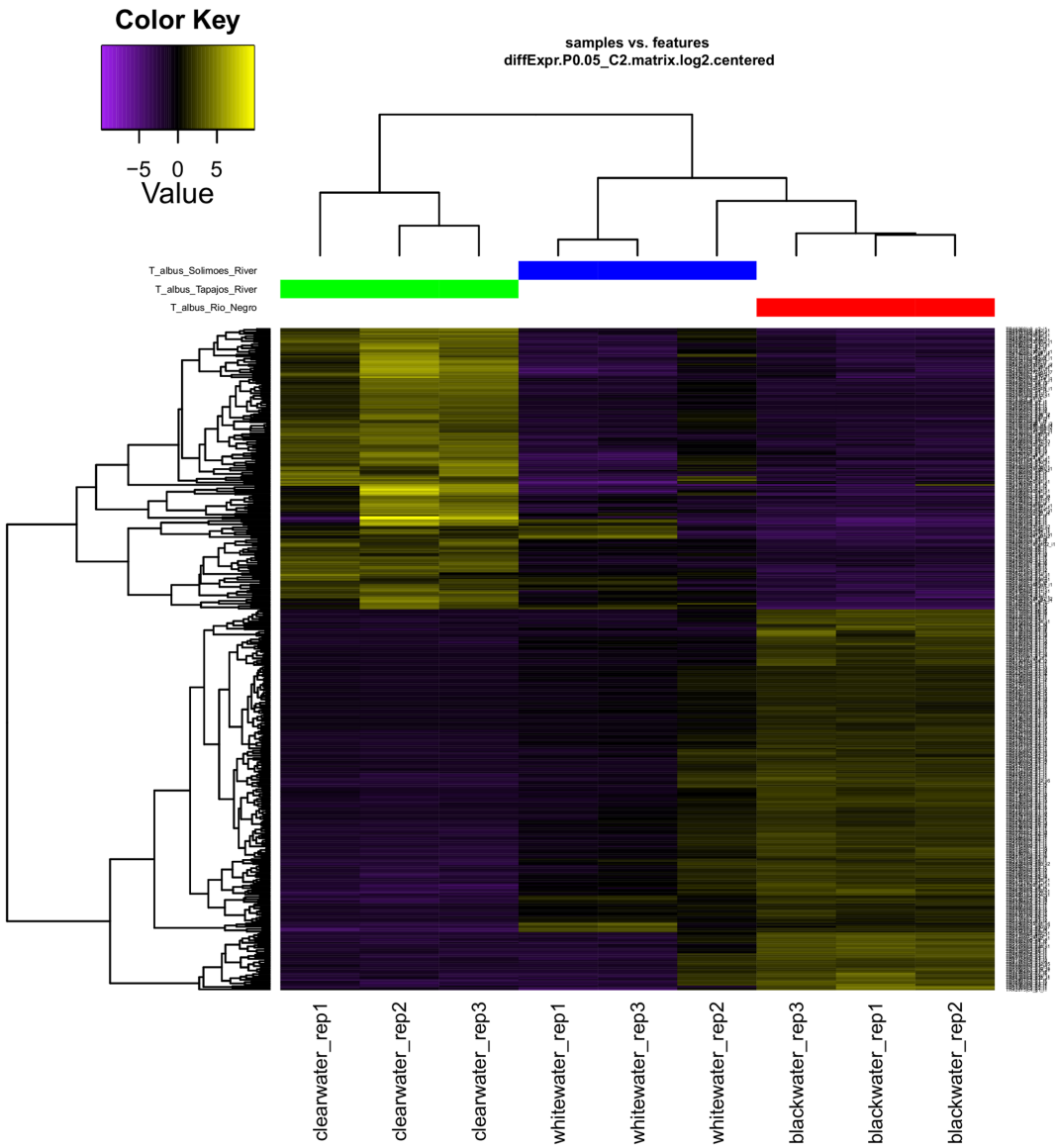


FIGURE 5 | Expression patterns (heatmap) and hierarchical clusters of the genes of *Triportheus albus* specimens differentially expressed in different habitats (blackwater, whitewater and clearwater). Dendrogram transcription patterns were estimated only for differentially expressed genes. Bar colors reflect expression levels of the gene, black (low), purple (down-regulated), and yellow (up-regulated).

the environment to which it was exposed. Therefore, the gene expression pattern observed for *T. albus* suggests that this species presents phenotypic plasticity to live in the three main types of water in the Amazon.

CONCLUSION

The Negro River is the most critical environment for the survival of many aquatic species in the Amazon basin, due to its high acidity and low ion concentration. As this river system harbors more 1,000 species, it is possible that more species present similar phenotypic plasticity as *T. albus* that showed two main mechanisms that allow survival in Amazonian aquatic environments, including those with low pH. The first mechanism is the control of genes in paracellular junctions, such as claudin, actn4, itgb3b, DSP that are involved in the process of maintaining the paracellular permeability control and, consequently, the loss of Na^+ and Cl^- ions to the environment. This characteristic that until then had been observed only in Amazon cichlids, was well developed in *T. albus*, a species of characid.

The second mechanism was attributed to the ability of ionic and acid-base regulation developed by this species. We observed high expression of the genes involved in Na^+ uptake, where excretion of NH_3^- through the rhcg1 protein somehow favors Na^+ uptake through the NHE exchanger, in addition to the

H^+-ATPase and the Na^+/K^+-ATPase pump. We also find the prlra and nr3c1 genes, responsible for triggering the two mechanisms described above. Therefore, we could verify that the species *T. albus* presents phenotypic plasticity with mechanisms that confer abilities to survive in environments considered critical for many species. We suggest that the species *T. albus* is a good candidate for future studies involving ion and acid-base regulation processes, as well as to analyze the activities of the respective enzymes involved in these processes.

AUTHOR CONTRIBUTIONS

JA, AG, and AV designed the work. JA and AG analyzed and interpretation of the data. JA and AV drafted the work. All the authors approve the final version.

ACKNOWLEDGMENTS

We thank MSc. Maria de Nazaré Paula da Silva, for the support in logistics in order to collect the samples. The MSc. Luciana Mara Fé Gonçalves and MSc. Erica Martinha Silva de Souza for the support in the sequencing of the data. This study is part of INCT-ADAPTA (CNPq/FAPEAM). AV is a recipient of a research fellowship from CNPq. JA was a recipient of a MSc. fellowship from FAPEAM and CNPq.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fgene.2017.00114/full#supplementary-material

REFERENCES

Altschul, S., Madden, T., Schäffer, A., and Zheng, J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Bayaa, M., Vulesevic, B., Esbaugh, A., Braun, M., Ekker, M. E., Grosell, M., et al. (2009). The involvement of SLC26 anion transporters in chloride uptake in zebrafish (*Danio rerio*) larvae. *J. Exp. Biol.* 212, 3283–3295. doi: 10.1242/jeb.033910

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bonga, S. E., Flik, G., Balm, P. H. M., and Van Der Meij, J. C. A. (1990). The ultrastructure of chloride cells in the gills of the teleost *Oreochromis mossambicus* during exposure to acidified water. *Cell Tissue Res.* 259, 575–585. doi: 10.1007/BF01740786

Breves, J. P., McCormick, S. D., and Karlstrom, R. O. (2014). Prolactin and teleost ionocytes: new insights into cellular and molecular targets of prolactin in vertebrate epithelia. *Gen. Comp. Endocrinol.* 203, 21–28. doi: 10.1016/j.ygcen.2013.12.014

Campbell, P. G., Twiss, M. R., and Wilkinson, K. J. (1997). Accumulation of natural organic matter on the surfaces of living cells: implications for the interaction of toxic solutes with aquatic biota. *Can. J. Fish. Aquat. Sci.* 54, 2543–2554. doi: 10.1139/f97-161

Chang, W. J., Horng, J. L., Yan, J. J., Hsiao, C. D., and Hwang, P. P. (2009). The transcription factor, glial cell missing 2, is involved in differentiation and functional regulation of H^+-ATPase-rich cells in zebrafish (*Danio rerio*). *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 296, R1192–R1201. doi: 10.1152/ajpregu.90973.2008

De Pinna, M. (2006). "Diversity of Tropical Fishes," in *The Physiology of Tropical Fishes*, eds A. L. Val, V. M. F. Almeida-Val, and D. J. Randall (London: Elsevier), 47–84.

Duarte, R. M., Ferreira, M. S., Wood, C. M., and Val, A. L. (2013). Effect of low pH exposure on Na^+ regulation in two cichlid fish species of the Amazon. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 166, 441–448. doi: 10.1016/j.cbpa.2013.07.022

Duarte, R. M., Smith, D. S., Val, A. L., and Wood, C. M. (2016). Dissolved organic carbon from the upper Rio Negro protects zebrafish (*Danio rerio*) against ionoregulatory disturbances caused by low pH exposure. *Sci. Rep.* 6:20377. doi: 10.1038/srep20377

Duncan, W. P., and Fernandes, M. N. (2010). Physicochemical characterization of the white, black, and clearwater rivers of the Amazon Basin and its implications on the distribution of freshwater stingrays (Chondrichthyes, Potamotrygonidae). *Pan Am. J. Aquat. Sci.* 5, 454–464.

Ertel, J. R., Hedges, J. I., Devol, A. H., Richey, J. E., Nazarg, M., and Ribeiro, J. E. (1986). Dissolved humic substances of the Amazon River system. *Limnol. Oceanogr.* 31, 739–754. doi: 10.4319/lo.1986.31.4.0739

Evans, D. H., Piermarini, P. M., and Choe, K. P. (2005). The multifunctional fish gill: dominant site of gas exchange, osmoregulation, acid-base regulation, and excretion of nitrogenous waste. *Physiol. Rev.* 85, 97–177. doi: 10.1152/physrev.00050.2003

Flik, G., and Perry, S. F. (1989). Cortisol stimulates whole body calcium uptake and the branchial calcium pump in freshwater rainbow trout. *J. Endocrinol.* 120, 75–82. doi: 10.1677/joe.0.1200075

Freda, J., Sanchez, D. A., and Bergman, H. L. (1991). Shortening of branchial tight junctions in acid-exposed rainbow-trout (*Oncorhynchus-Mykiss*). *Can. J. Fish. Aquat. Sci* 48, 2028–2033. doi: 10.1139/f91-241

Furch, K. (1982). Jahreszcltliche chemische Veranderungen in einem Várzea-See des mittleren Amazonas (Lago Calado. Brasilien). *Arch. Hydrobiol.* 95, 47–67.

Furch, K. (1984). "Water chemistry of the Amazon Basin: the distribution of chemical elements among freshwaters," in *The Amazon Limnology and Landscape Ecology of a Mighty Tropical River and its Basin*, ed. H. Sioli (Dordrecht: W. Junk), 167–169.

Gaillardet, J., Dupre, B., Allegre, C. J., and Négrel, P. (1997). Chemical and physical denudation in the Amazon River Basin. *Chem. Geol.* 142, 141–173. doi: 10.1016/S0009-2541(97)00074-0

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 5:R80. doi: 10.1186/gb-2004-5-10-r80

Gibbons, T. C., Metzger, D. C. H., Healy, T. M., and Schultee, P. M. (2017). Gene expression plasticity in response to salinity acclimation in threespine stickleback ecotypes from different salinity habitats. *Mol. Ecol.* 26, 2711–2725. doi: 10.1111/mec.14065

Gilmour, K. M., and Perry, S. F. (2009). Carbonic anhydrase and acid-base regulation in fish. *J. Exp. Biol.* 212, 1647–1661. doi: 10.1242/jeb.029181

Gonzalez, R. J., and Wilson, R. W. (2001). Patterns of ion regulation in acidophilic fish native to the ion-poor, acidic Rio Negro. *J. Fish Biol.* 58, 1680–1690. doi: 10.1006/jfbi.2001.1577

Gonzalez, R. J., Wilson, R. W., Wood, C. M., Patrick, M. L., and Val, A. L. (2002). Diverse strategies for ion regulation in fish collected from the ion-poor, acidic Rio Negro. *Physiol. Biochem. Zool.* 75, 37–47. doi: 10.1086/339216

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Holland, A., Duivenvoorden, L. J., and Kinnear, S. H. W. (2012). Naturally acidic waterways: conceptual food webs for better management and understanding of ecological functioning. *Aquat. Conserv.* 22, 836–847. doi: 10.1002/aqc.2267

Holland, A., Wood, C. M., Smith, D. S., Correia, T. G., and Val, A. L. (2017). Nickel toxicity to cardinal tetra (*Paracheirodon axelrodi*) differs seasonally and among the black, white and clear river waters of the Amazon basin. *Water Res.* 123, 21–29. doi: 10.1016/j.watres.2017.06.044

Ito, Y., Kobayashi, S., Nakamura, N., Miyagi, H., Esaki, M., Hoshijima, K., et al. (2013). Close association of carbonic anhydrase (CA2a and CA15a), Na^+/H^+ exchanger (Nhe3b), and ammonia transporter Rhcg1 in zebrafish ionocytes responsible for Na^+ uptake. *Front. Physiol.* 4:59. doi: 10.3389/fphys.2013.00059

Jesus, T. F., Grosso, A. R., Almeida-Val, V. M. F., and Coelho, M. M. (2016). Transcriptome profiling of two Iberian freshwater fish exposed to thermal stress. *J. Therm. Biol.* 55, 54–61. doi: 10.1016/j.jtherbio.2015.11.009

Johannsson, O. E., Smith, D. S., Sadauskas-Henrique, H., Cimprich, G., Wood, C. M., and Val, A. L. (2016). Photo-oxidation processes, properties of DOC, reactive oxygen species (ROS), and their potential impacts on native biota and carbon cycling in the Rio Negro (Amazonia. Brazil). *Hydrobiologia* 789, 7–29. doi: 10.1007/s10750-016-26879

Junk, W. J., and Furch, K. (1980). Química da água e macrófitas de rios e igarapés na bacia Amazônica e nas áreas adjacentes. Partel: trecho cuiaba - porto velho - manaus. *Acta Amazonica* 10, 611–633. doi: 10.1590/1809-43921980103611

Kelly, S. P., and Wood, C. M. (2001). Effect of cortisol on the physiology of cultured pavement cell epithelia from freshwater trout gills. *Am. J. Physiol. Regul. Integr. Physiol. Comp* 281, R811–R820.

Konhauser, K. O., Fyfe, W. S., and Kronberg, B. I. (1994). Multi-element chemistry of some Amazonian waters and soils. *Chem. Geol.* 111, 155–175. doi: 10.1016/0009-2541(94)900884

Küchler, I. L., Miekeley, N., and Forsberg, B. R. (2000). A contribution to the chemical characterization of rivers in the Rio Negro Basin. Brazil. *J. Braz. Chem. Soc.* 11, 286–292. doi: 10.1590/S0103-50532000000300015

Kumai, Y., Nesan, D., Vijayan, M. M., and Perry, S. F. (2012). Cortisol regulates Na+ uptake in zebrafish, *Danio rerio*, larvae via the glucocorticoid receptor. *Mol. Cell. Endocrinol.* 364, 113–125. doi: 10.1016/j.mce.2012.08.017

Kumai, Y., and Perry, S. F. (2011). Ammonia excretion via Rhcg1 facilitates Na+ uptake in larval zebrafish, *Danio rerio*, in acidic water. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 301, R1517–R1528. doi: 10.1152/ajpregu.00282.2011

Kwong, R. W. M., Kumai, Y., and Perry, S. F. (2014). The physiology of fish at low pH: the zebrafish as a model system. *J. Exp. Biol.* 217, 651–662. doi: 10.1242/jeb.091603

Kwong, R. W. M., and Perry, S. F. (2013). Cortisol regulates epithelial permeability and sodium losses in zebrafish exposed to acidic water. *J. Endocrinol.* 217, 253–264. doi: 10.1530/JOE-12-0574

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Lemgruber, R. D. S. P., Marshall, N. A. D. A., Ghelfi, A., Fagundes, D. B., and Val, A. L. (2013). Functional categorization of transcriptome in the species *Symphysodon aequifasciatus* Pellegrin 1904 (Perciformes: Cichlidae) exposed to benzo[a]pyrene and phenanthrene. *PLoS ONE* 8:e81083. doi: 10.1371/journal.pone.0081083

Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323

Li, S., Liu, C., Huang, J., Liu, Y., Zhang, S., Zheng, G., et al. (2016). Transcriptome and biomineralization responses of the pearl oyster Pinctada fucata to elevated CO_2 and temperature. *Sci. Rep.* 6:18943. doi: 10.1038/srep18943

Lin, H., and Randall, D. J. (1993). H^+ -ATPase activity in crude homogenates of fish gill tissue: inhibitor sensitivity and environmental and hormonal regulation. *J. Exp. Biol.* 174, 163–174.

Lin, L. Y., Horng, J. L., Kunkel, J. G., and Hwang, P. P. (2006). Proton pump-rich cell secretes acid in skin of zebrafish larvae. *Am. J. Physiol. Cell Physiol.* 290, C371–C378. doi: 10.1152/ajpcell.00281.2005

Lin, T. Y., Liao, B. K., Horng, J. L., Yan, J. J., Hsiao, C. D., and Hwang, P. P. (2008). Carbonic anhydrase 2-like a and 15a are involved in acid-base regulation and Na^+ uptake in zebrafish H^+-ATPase-rich cells. *Am. J. Physiol. Cell Physiol.* 294, C1250–C1260. doi: 10.1152/ajpcell.00021.2008

Logan, C. A., and Buckley, B. A. (2015). Transcriptomic responses to environmental temperature in eurythermal and stenothermal fishes. *J. Exp. Biol.* 218, 1915–1924. doi: 10.1242/jeb.114397

Marshall, W. S. (2002). Na^+, Cl^-, Ca^{2+} and Zn^{2+} transport by fish gills: retrospective review and prospective synthesis. *J. Exp. Zool.* 293, 264–283. doi: 10.1002/jez.10127

Matsuo, A. Y. O., and Val, A. L. (2007). Acclimation to humic substances prevents whole body sodium loss and stimulates branchial calcium uptake capacity in cardinal tetras *Paracheirodon axelrodi* (Schultz) subjected to extremely low pH. *J. Fish Biol.* 70, 989–1000. doi: 10.1111/j.1095-8649.2007.01358.x

McDonald, D. G. (1983). The effects of H^+ upon the gills of freshwater fish. *Can. J. Zool.* 61, 691–703. doi: 10.1139/z83-093

Parks, S. K., Tresguerres, M., and Goss, G. G. (2008). Theoretical considerations underlying Na^+ uptake mechanisms in freshwater fishes. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* 148, 411–418. doi: 10.1016/j.cbpc.2008.03.002

Perry, S. F., Vulesevic, B., Grosell, M., and Bayaa, M. (2009). Evidence that SLC26 anion transporters mediate branchial chloride uptake in adult zebrafish (*Danio rerio*). *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 297, R988–R997. doi: 10.1152/ajpregu.00327.2009

Prado-Lima, M., and Val, A. L. (2016). Transcriptomic characterization of tambaqui (*Colossoma macropomum*, Cuvier, 1818) exposed to three climate change scenarios. *PLoS ONE* 11:e0152366. doi: 10.1371/journal.pone.0152366

Queiroz, M. M. A., Horbe, A. M. C., Seyler, P., Augusto, C., and Moura, V. (2009). Hidroquímica do rio solimões na região entre manacapuru e alvarães – Amazonas – Brasil. *Acta Amazonica* 39, 943–952. doi: 10.1590/S0044-59672009000400022

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Schneider, R. F., and Meyer, A. (2017). How plasticity, genetic assimilation and cryptic genetic variation may contribute to adaptive radiations. *Mol. Ecol.* 26, 330–350. doi: 10.1111/mec.13880

Sioli, H. (ed.) (1984). *The AmazonLimnology and Landscape Ecology of a Mighty Tropical River and Its Basin*. Dordrecht: Springer.

Val, A. L., and Almeida-Val, V. M. F. (1995). *Fishes of the Amazon and Their Environment*. Berlin: Springer. doi: 10.1007/978-3-642-79229-8

Vidal-Dupiol, J., Zoccola, D., Tambutté, E., Grunau, C., Cosseau, C., Smith, K. M., et al. (2013). Genes related to ion-transport and energy production are upregulated in response to CO_2-driven pH decrease in corals: new insights from transcriptome analysis. *PLoS ONE* 8:e58652. doi: 10.1371/journal.pone.0058652

Walker, I. (1995). "Amazonian streams and small rivers," in *Limnology in Brazil*, eds J. G. Tundisi, T. MatsumuraTundisi, and C. E. Bicudo (Rio de Janeiro: ABC/SBL), 167–193.

Wood, C. M., Al-Reasi, H. A., and Smith, D. S. (2011). The two faces of DOC. *Aquat. Toxicol.* 105, 3–8. doi: 10.1016/j.aquatox.2011.03.007

Wood, C. M., Kajimura, M., Sloman, K. A., Scott, G. R., Walsh, P. J., Almeida-Val, V. M. F., et al. (2007). Rapid regulation of Na^+ fluxes and ammonia excretion in response to acute environmental hypoxia in the Amazonian oscar, *Astronotus ocellatus*. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 292, R2048–R2058. doi: 10.1152/ajpregu.00640.2006

Wood, C. M., Robertson, L. M., Johannsson, O. E., and Val, A. L. (2014). Mechanisms of Na^+ uptake, ammonia excretion, and their potential linkage in native Rio Negro tetras (*Paracheirodon axelrodi, Hemigrammus rhodostomus*, and *Moenkhausia diktyota*). *J. Comp. Physiol. B.* 184, 877–890. doi: 10.1007/s00360-014-08477

Wood, C. M., Wilson, R. W., Gonzalez, R. J., Patrick, M. L., Bergman, H. L., Narahara, A., et al. (1998). Responses of an Amazonian teleost, the tambaqui (Colossoma macropomum), to low pH in extremely soft water. *Physiol. Zool.* 71, 658–670. doi: 10.1086/515977

Wright, P. A., and Wood, C. M. (2012). Seven things fish know about ammonia and we don't. *Respir. Physiol. Neurobiol.* 184, 231–240. doi: 10.1016/j.resp.2012.07.003

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

Copyright © 2017 Araújo, Ghelfi and Val. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 21 September 2017
doi: 10.3389/fgene.2017.00120

The *Bunocephalus coracoideus* Species Complex (Siluriformes, Aspredinidae). Signs of a Speciation Process through Chromosomal, Genetic and Ecological Diversity

Milena Ferreira[1], Caroline Garcia[2], Daniele A. Matoso[3], Isac S. de Jesus[4], Marcelo de B. Cioffi[5], Luiz A. C. Bertollo[5], Jansen Zuanon[6] and Eliana Feldberg[1]*

[1] Laboratório de Genética Animal, Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil, [2] Laboratório de Citogenética, Departamento de Ciências Biológicas, Universidade Estadual do Sudoeste da Bahia, Jequié, Brazil, [3] Laboratório de Citogenômica Animal, Instituto de Ciências Biológicas, Departamento de Genética, Universidade Federal do Amazonas, Manaus, Brazil, [4] Laboratório de Fisiologia Comportamental e Evolução, Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil, [5] Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, Brazil, [6] Laboratório de Sistemática e Ecologia de Peixes, Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Vladimir Pavan Margarido, Universidade Estadual do Oeste do Paraná, Brazil
Mauro Nirchio, Universidad de Oriente, Venezuela

***Correspondence:**
Milena Ferreira
milena_fro@hotmail.com

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

Received: *17 July 2017*
Accepted: *29 August 2017*
Published: *21 September 2017*

Citation:
Ferreira M, Garcia C, Matoso DA, de Jesus IS, Cioffi MB, Bertollo LAC, Zuanon J and Feldberg E (2017) The Bunocephalus coracoideus Species Complex (Siluriformes, Aspredinidae). Signs of a Speciation Process through Chromosomal, Genetic and Ecological Diversity. Front. Genet. 8:120. doi: 10.3389/fgene.2017.00120

Bunocephalus is the most species-rich Aspredinidae genus, corresponding to a monophyletic clade with 13 valid species. However, many species have their classification put in question. Here, we analyzed individuals from four Amazonian populations of *Bunocephalus coracoideus* by cytogenetic and molecular procedures. The geographic distribution, genetic distances and karyotype data indicate that each population represents an Evolutionary Significant Unit (ESU). Cytogenetic markers showed distinct 2n and karyotype formulas, as well as different numbers and locations of the rDNA sites among ESUs. One of such populations (ESU-D) highlighted an extensive polymorphic condition, with several cytotypes probably due to chromosomal rearrangements and meiotic non-disjunctions. This resulted in several aneuploid karyotypes, which was also supported by the mapping of telomeric sequences. Phylograms based on Maximum Likelihood (ML) and Neighbor Joining (NJ) analyses grouped each ESU on particular highly supported clades, with the estimation of evolutionary divergence indicating values being higher than 3.8–12.3% among them. Our study reveals a huge degree of chromosomal and genetic diversity in *B. coracoideus* and highly points to the existence of four ESUs in allopatric and sympatric speciation processes. In fact, the high divergences found among the ESUs allowed us to delimitate lineages with taxonomic uncertainties in this nominal species.

Keywords: chromosomal differentiation, molecular taxonomy, ecological adaptations, evolutionary units, banjo catfish

INTRODUCTION

Bunocephalus is the most species-rich Aspredinidae genus, corresponding to a monophyletic clade (Cardoso, 2008, 2010) with 13 valid species (Eschmeyer et al., 2017). However, many species have their classification put in question (Carvalho et al., 2015), and, consequently, with questionable taxonomy. The genetic divergence among morphologically indistinguishable specimens within a

single "species" rises doubts on their taxonomic status, once such variations imply in cladogenesis processes (Bickford et al., 2007), which may end up in speciation. Some approaches based only on morphological data may sometimes underestimate such variation due to the phenotypic plasticity evidenced by a significant number of species. Thus, an integrative cytotaxonomic, molecular and morphological analysis is required, attempting to elucidate the real taxonomic status of polymorphic species.

Cytogenetic studies, including the molecular organization and cytogenetic mapping of repetitive DNAs might be a significant data set for the characterization of particular segments of biota, providing important information for phylogenomics (Cioffi et al., 2012). Besides, these sequences seem to escape the selective pressure that acts in the non-repetitive segments, thus representing good evolutionary markers to detect recent events of evolution, once the number and location of these sequences may reveal polymorphisms, with intra- and inter-specific variations due to rearrangements, even in conserved karyotypes (Cioffi et al., 2009; Matoso et al., 2011; Motta-Neto et al., 2012; Oliveira et al., 2015). Up to now, only two *Bunocephalus* species had cytogenetic studies already conducted, *B. doriae* and *B. coracoideus*. While the first one has $2n = 50$ chromosomes, the later presents $2n = 42$, in addition to a multiple $X_1X_1X_2X_2/X_1Y_1X_2Y_2$ sex chromosome system (Fenocchio and Swarça, 2012; Ferreira et al., 2016).

Besides, genetic data can unmask distinct populations covered by a same taxonomic status, which are identified as Significant Evolutionary Units (ESUs) by the conservation biology area. Thus, ESU corresponds to a population, or even to a group of populations, genetically distinct within a given species that contribute to biodiversity (Hey et al., 2003). The ESUs recognition is a task that requires the accordance of a group of procedures as an identification criterion.

TABLE 1 | Estimates of evolutionary divergence between sequences using the COI gene and the K2P model.

	ESU-A	ESU-B	ESU-C	ESU-D
ESU-A	0.2			
ESU-B	12.2	0.2		
ESU-C	10.6	4.0	0.7	
ESU-D	12.3	5.7	3.8	2.5

Values are shown as percentages.

The present study is a contribution to the biodiversity presented by *B. coracoideus* using DNA barcoding and conventional and molecular cytogenetic methodologies. It was analyzed four allopatric populations from the Amazonian hydrographic basin and the results were able to highlight a huge cryptic diversity both intra- and inter-populations, pointing out that *B. coracoideus* corresponds to a species complex.

MATERIALS AND METHODS

Specimens

Individuals of *Bunocephalus coracoideus* from four populations of distinct drainages of the Amazon River were analyzed (**Figure 1** and **Table 1**). The specimens were collected under appropriate authorization of the Brazilian environmental agency ICMBIO/SISBIO (License number 48795-1). All specimens were properly identified by morphological criteria and voucher samples were deposited in the fish collections of the National Institute for Amazon Research (INPA: *Instituto Nacional de Pesquisas da Amazônia*). The experiments followed ethical and anesthesia conducts, in accordance with the Ethics Committee for Animal Use of the National Institute of Amazon Research, under the protocol number 010/2015.

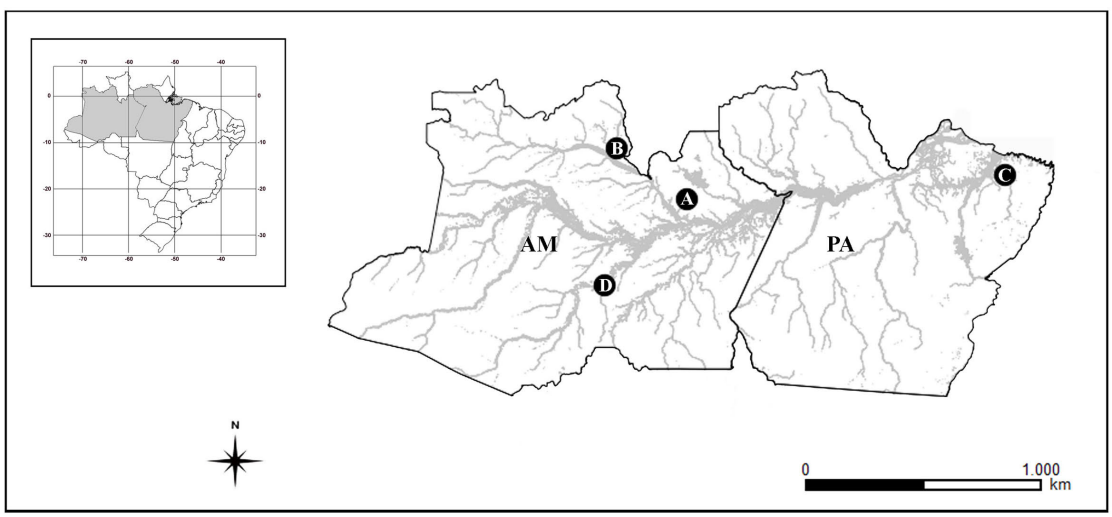


FIGURE 1 | Amazonas (AM) and Pará (PA) Brazilian states map, highlighting the collection sites of the analyzed species. **(A)** ESU-A Igarapé Jundiá – Cuieiras River, **(B)** ESU-B Demini River, **(C)** ESU-C Igarapé Apeú – Guamá River, and **(D)** SEU-D_{1-14} Purus River.

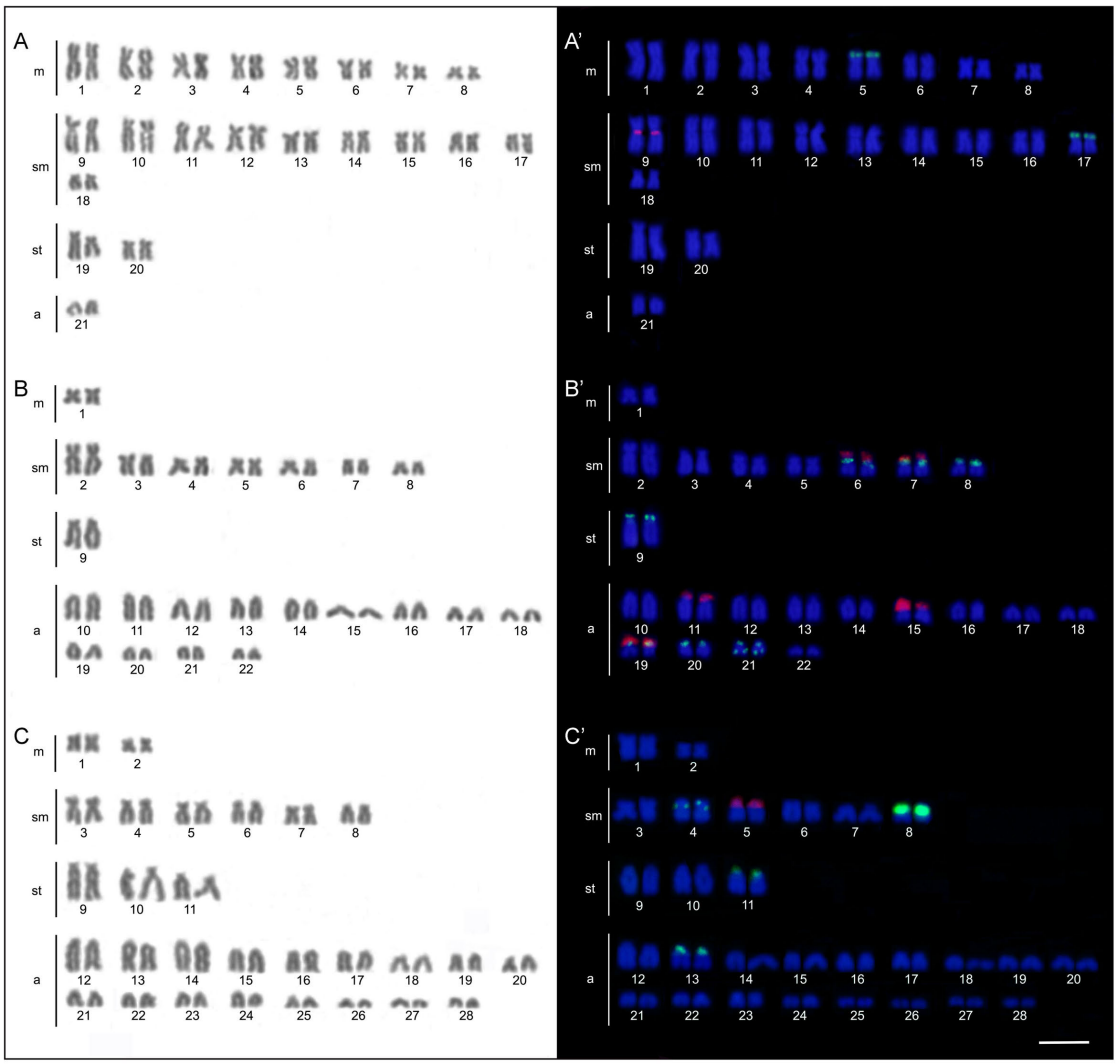


FIGURE 2 | *Bunocephalus coracoideus* karyotypes. On the left, Giemsa staining; on the right. Double-FISH evidencing chromosome pairs bearing the 18S rDNA (red) and 5S rDNA (green) sequences. **(A,A')** ESU-A Igarapé Jundiá – Cuieiras River, **(B,B')** ESU –B Demini River (Cuieiras River) and **(C,C')** ESU-C Igarapé Apeú – Guamá River. Bar = 10 μm.

Mitotic and Meiotic Chromosomal Preparations

Mitotic chromosomes were obtained from the kidney cells according to the protocol described by Gold et al. (1990), using RPMI culture medium (Cultilab). Male testis were used for meiotic preparations, following the protocols described by Bertollo et al. (1978), with changes introduced by Gross et al. (2009). For conventional cytogenetic analysis the chromosomes were stained with 5% Giemsa solution (pH 6.8).

Preparation of FISH Probes

The GoTaq Colorless Master Mix (Promega) was used for the polymerase chain reaction (PCR) amplification of the 18S and 5S rRNA genes and telomeric sequences, using the following primers: 18Sf (5′-CCG CTT TGG TGA CTC TTG AT-3′), 18Sr (5′-CCG AGGACC TCA CTA AAC CA-3′) (Gross et al., 2010), 5S A (5′-TAC GCC CGA TCT CGT CCG ATC-3′), and 5S B (5′-CAGGCT GGT ATG GCC GTA AGC-3′) (Martins and Galetti, 1999). The ribosomal sequence amplification cycles comprised a denaturation for 2 min at 95°C; 35 cycles of 1 min at 94°C, 30 seg. at 56°C, and 1.5 min at 72°C; a final extension of 5 min at 72°C; and a cooling period at 4°C. The primers (TTAGGG)5 and (CCCTAA)5 (Ijdo et al., 1991), were used to obtaining telomeric sequences. PCR was performed with the following profile: 4 min at 94°C; 12 cycles of 1 min at 94°C, 45 s at 52°C, and 1.5 min at 72°C; and 35 cycles of 1 min at 94°C, 1.5 min at 60°C, and 1.5 min at 72°C. The 18S rDNA and the telomeric probes were labeled with digoxigenin-11 dUTP using a DIG-Nick Translation Mix kit (Roche), while the 5S rDNA probe was labeled with biotin-14-dATP using a Biotin-Nick Translation Mix kit (Roche) according to the manufacturer's instructions.

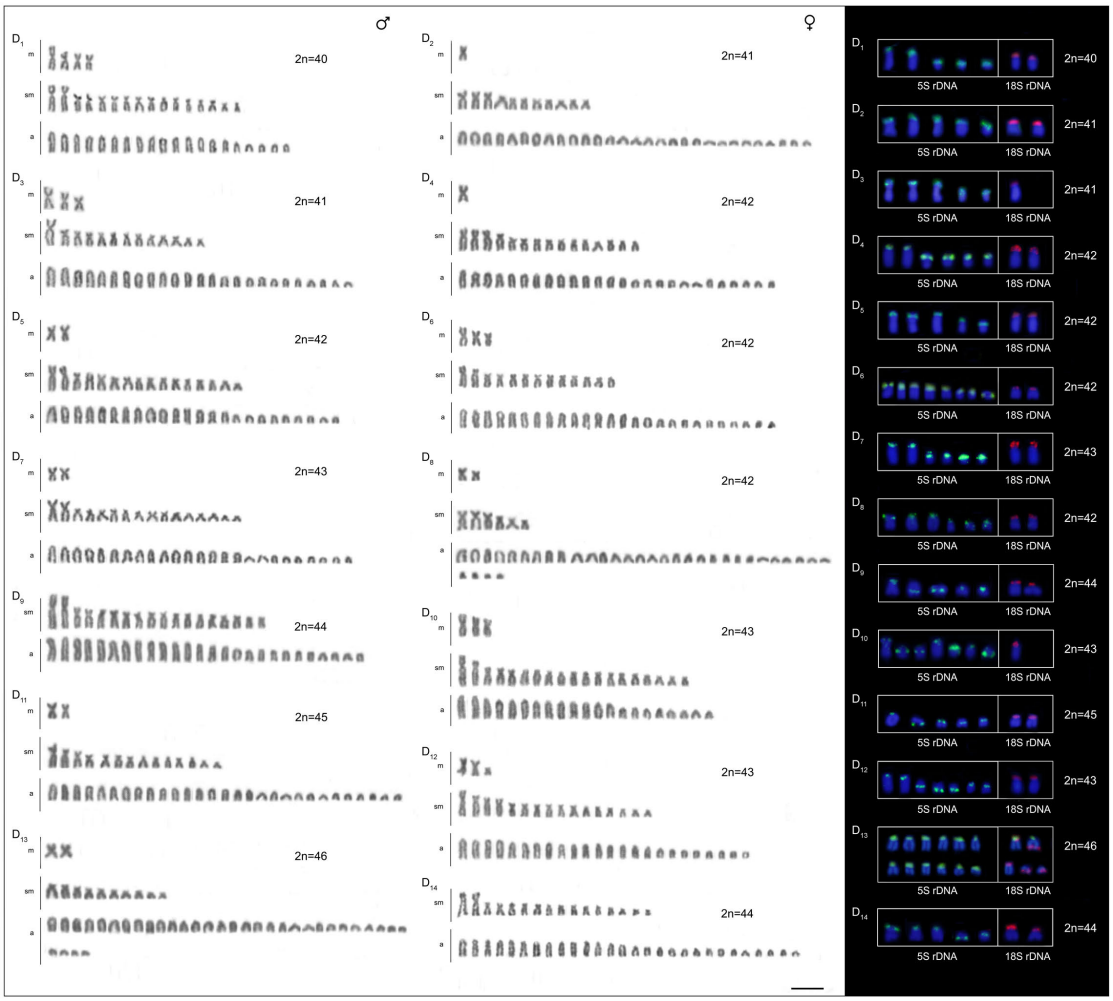


FIGURE 3 | Cytotypes of *B. coracoideus* ESU-D_{1-14} Purus River, with Giemsa staining. For this ESU chromosomes were not paired, once no homeology was usually found between the correspondent pairs among the distinct cytotypes. On the right, double-FISH showing the chromosome pairs bearing the 18S rDNA (red) and 5S (green) sequences. Bar = 10 μm.

Detection of Repetitive DNA Sequences by FISH

Fluorescence *in situ* hybridization (FISH) was performed according to the protocol described by Pinkel et al. (1986), with some modifications. The 18S rDNA and telomeric probes were detected with Anti-digoxigenin-rhodamin (Roche, Basel, Switzerland), while the 5S rDNA probe was detected with avidin-FITC (Sigma). Chromosomes were counterstained with DAPI (1.2 μg/ml) and slides mounted with antifade solution (Vector, Burlingame, CA, United States).

Microscopic Analysis and Image Processing

At least 30 metaphase spreads per individual were analyzed to confirm the 2n, karyotype structure and FISH results. Images were captured using an Olympus BX50 microscope (Olympus Corporation, Ishikawa, Japan) with CoolSNAP camera and the images processed using Image Pro Plus 4.1 software (Media Cybernetics, Silver Spring, MD, United States). The chromosome classification followed the method proposed by Levan et al. (1964), with the following limits for the arms relationship (AR): AR = 1.00–1.70, metacentric (m); AR = 1.71–3.00, submetacentric (sm); AR = 3.01–7.00, subtelocentric (st); and AR > 7.00, acrocentric (a). For the number of chromosome arms [fundamental number (FN)], the metacentric, submetacentric, and subtelocentric chromosomes was considered having two chromosomal arms and the acrocentric chromosomes a single one.

DNA Barcoding Analysis

Representatives of each population were used (**Table 1**). *Bunocephalus* cf. *aloikae, B. amaurus,* and *Amaralia hypsiura* species were used as out groups. Tissues of liver and muscle were stored in absolute ethanol for the acquisition of Cytochrome *C* Oxidase Subunit 1 (COI) sequences. Total DNA was obtained

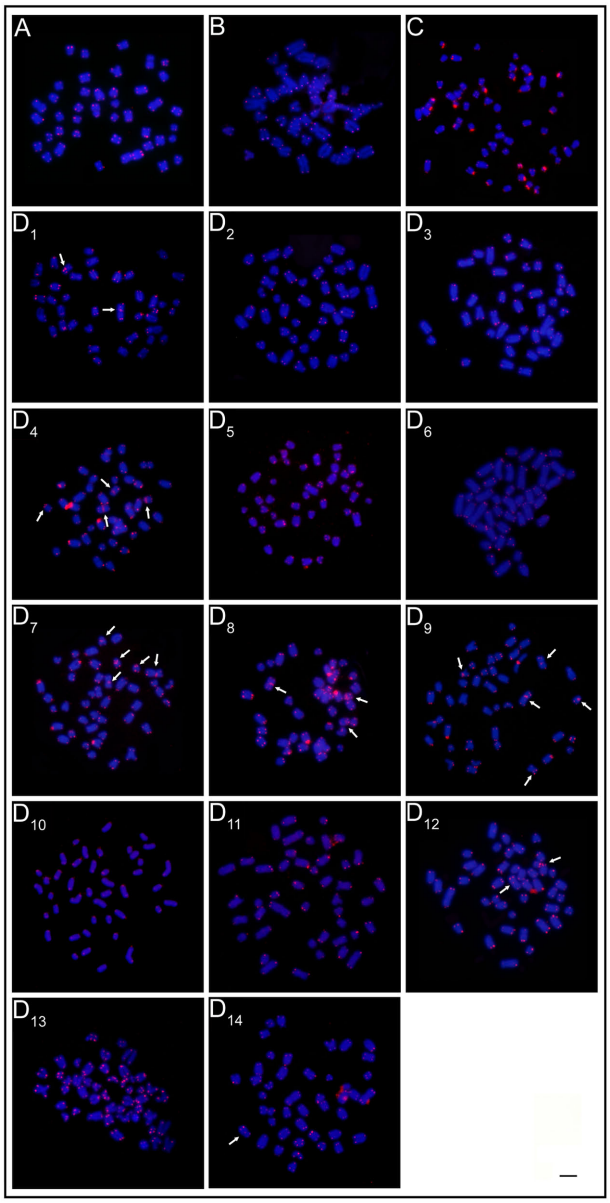


FIGURE 4 | FISH with telomeric probe. **(A)** ESU-A Igarapé Jundiá – Cuieiras River, **(B)** SEU-B Demini River, **(C)** SEU-C Igarapé apeú – Guamá River, and **(D_1 – D_{14})** SEU-D_{1-14} Purus River. Arrows indicate the occurence of interstitial telomeric sites (ITS). Bar = 10 μm.

by Wizard® Genomic DNA Purification Kit. The pair of primers used for the COI mitochondrial region amplification in PCR reactions was VF1_t1 (TGT AAA ACG GCC AGT CAA CCA ACC ACA AAG ACA TTG G) + VR1_t1 (CAG GAA ACA GCT ATG ACT AGA CTT CTG GGT GGC CAA AGA ATC A) (Ivanova et al., 2006). Each PCR reaction presented a final volume of 25 μl containing 1 μl of DNA template [250 ng/μl] + 1 μl of each primer [5 pM]. It was used the GoTaq Colorless Master Mix® (Promega) for the PCR. The amplification cycles comprised denaturation, 2 min at 95°C; 35 cycles of 1 min at 94°C, 30 seg. at 56°C, and 1.5 min at 72°C; a final extension of 5 min at 72°C; and a cooling period at 4°C. PCR products were visualized on a 1.7% agarose gel and purified with 20% PEG (Lis, 1980). For sequencing it was used a "Big Dye Sequence Terminator v.3.1" kit (Applied Biosystems), according to the manufacturer instructions. The amplification conditions were comprised 25 cycles at 96°C for 30 seg.; 15 seg. at 50°C; and 4 min at 60°C. After the reaction, the products were precipitated and sequenced (sequencer model ABI PRISM 3100 Genetic Analyzer from Applied Biosystems/made by HITACHI).

Sequence Alignment and Phylogenetic Analysis

Sequences with 690 pb were used to perform the barcoding analyses, by using COI gene, which were aligned using the Geneious® 10.1.3 software. The distance model of Kimura 2-parameters (Kimura, 1980), was used to build a Neighbor-Joining (NJ) dendrogram and a bootstrap analyses was performed (Felsenstein, 1985) with 1,000 replicates. All the aligned sequences were translated into amino acids to detect possible alignment errors. The Maximum Likelihood (ML) model (Tamura et al., 2004) was performed to recover the phylogenetic topology. All positions containing gaps and missing data were eliminated. There were a total of 461 positions in the final dataset. Pairwise genetic distance calculations and NJ tree analysis were implemented using Molecular Evolutionary Genetics Analysis version 5 (MEGA5) software (Tamura et al., 2011) and applying 1,000 bootstrap replicates.

RESULTS

Cytogenetic Data

The four *B. coracoideus* populations presented distinct karyotypes, which classified as evolutionary significant units (ESUs), with the following characteristics ESU-A: 42 chromosomes (16m+20sm+4st+2a, NF = 82) from Igarapé Jundiá – Cuieiras River (**Figure 2A**), ESU-B: 44 chromosomes (2m+14sm+2st+26a, NF = 62) from Demini River (**Figure 2B**) and ESU-C: 56 chromosomes (4m+12sm+6st+34a, NF = 78) from Igarapé Apeú – Guamá River (**Figure 2C**). There was no karyotype differentiation among males and females in these populations. For the ESU-D, from Purus River, 14 distinct cytotypes bearing variant chromosomes in number and morphology were observed, with 2n varying from 40 to 46. Due to such variation, chromosomes were not grouped in pairs once homeology was not usually found among the distinct cytotypes (**Figure 3**). Significant variations were also found concerning the 18S rDNA carrier chromosomes, localization and number of sites. For ESUs A, B, and C such sequences were found in one or five chromosomes pairs, in the pericentromeric or telomeric regions, in the short or long arms (**Figure 2**). Similar results were also detected among the cytotypes of the ESU-D. However, in this case, although the telomeric position of the sites was consistent for all cytotypes, they were found in the short arms of two chromosomes in the cytotypes D_1, D_2, D_4, D_5, D_6, D_7,

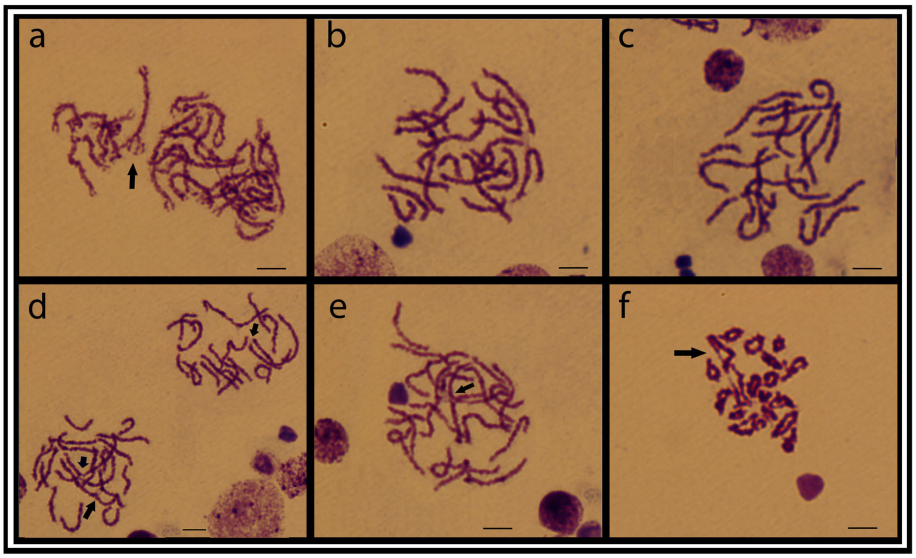


FIGURE 5 | Testicular meiotic cell plates from the ESU-D in Giemsa staining. **(a)** Zigotene cells, with a possible tetravalent arrangement; **(b–e)** Aneuploid pachytene cells and arrows evidencing chiasmata formation; **(f)** Diplotene cells, with associated chromosomes likely in a chromosomal chain organization. Bar = 10 μm.

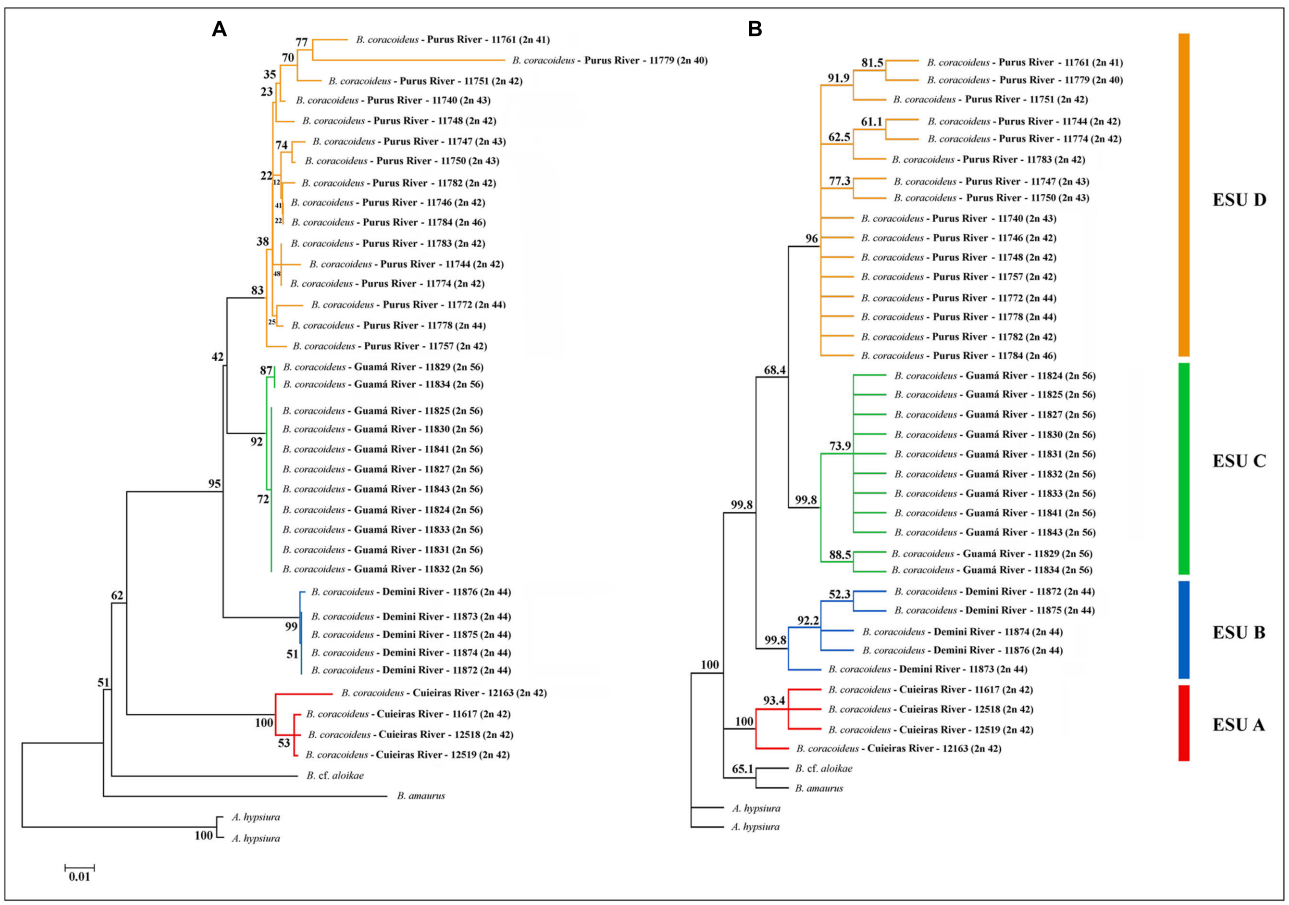


FIGURE 6 | Phylogeny of *B. coracoideus* inferred by the analysis of **(A)** Maximum Likelihood (ML) and **(B)** Neighbor-Joining (NJ) using the mitochondrial gene COI. The bootstrap values for 1000 replications are evidenced above the branches.

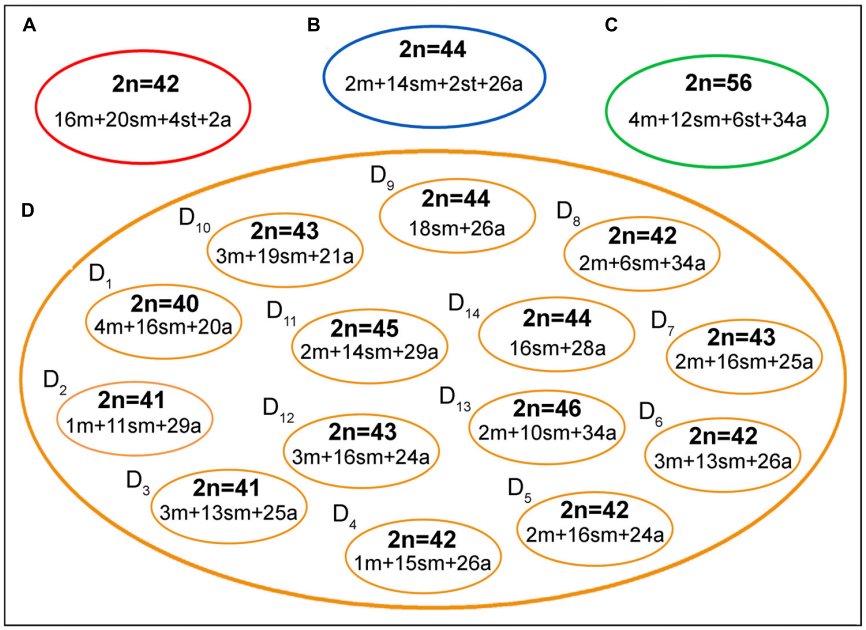


FIGURE 7 | A general overview on the chromosomal heterogeneity on *B. coracoideus*. **(A)** ESU-A Igarapé Jundiá – Cuieiras River, **(B)** SEU-B Demini River, **(C)** SEU-C Igarapé Apeú – Guamá River, and **(D)** SEU-D_{1-14} Purus River.

D_8, D_9, D_{11}, D_{12} and D_{14}, but in only one chromosome in the cytotypes D3 and D_{10}. In turn, the cytotype D_{13} highlighted two chromosomes bearing sites in the short arms and three chromosomes in the long arms (**Figure 3**).

Likewise, the 5S rDNA also showed great variation among ESUs. Like for 18S rDNA, with distinct chromosomes carrying these sequences and with variations in number and localization on the chromosomes. ESU-A showed only two chromosome pairs carrying 5S sequences, while ESUs B, C, and D presented a higher number of these sites. Although maintaining the preferential telomeric localization, interstitial positions were also highlighted mainly among ESU-D cytotypes. In addition, the syntenic localization with the 18 rDNA was evidenced in three chromosome pairs of the ESU-B, as well as in two chromosomes of the ESU-D – cytotype D_{13} (**Figures 2**, **3**).

The mapping of telomeric sequences evidenced only the usual terminal marks on the chromosomes of the ESUs A, B, and C. In turn, the ESU-D exhibited additional interstitial sites (ITS) in seven cytotypes, D_1, D_4, D_7, D_8, D_9, D_{12}, and D_{14} (**Figure 4**).

Meiotic plates from individuals of the ESU-D, showed a variable number of chromosomes corroborating the diversity found in the mitotic chromosomes. From 18 to 22 bivalents were evidenced, in addition to interstitial chiasmata and synaptic points, and probable tetravalent and chromosomal chain formations (**Figure 5**).

DNA Barcoding Analysis

Topologies obtained with the Neighbor-Joining (NJ) and Maximum Likelihood (ML) algorithms were congruent. The major clades were well-supported and it was confirmed that *Bunocephalus* represents a monophyletic group (**Figure 5**). Each population was also grouped as a monophyletic and well-supported clade, justifying them as four ESUs. ESU-A occupies the more basal position in relation to the other ones, and the ESUs C and D are more related to each other and with a more recent divergence (**Figure 6**).

DISCUSSION

The geographic distribution, genetic distances, and karyotype data indicated that each *B. coracoideus* population represents an ESU. In fact, these populations differed by conspicuous karyotypes variability, where each ESU shows specificities on their 2n, karyotype formula and ribosomal sites distribution in the genome (**Figures 7**, **8**). In addition, they have possibly evolved in allopatry due to vicariant events, making their natural contact unfeasible. Oliveira and Gosztonyi (2000) proposed that the ancestral karyotype of Siluriformes contained 2n = 56 chromosomes, mainly two-armed ones. According to our phylogenetic data, ESU-A (2n = 42; 16m+20sm+4st+2a) corresponds to the firstly differentiated karyotype among the four populations analyzed. In this way, ESU- A probably retains an ancestral feature of Siluriformes by the large number of bi-armed chromosomes they have, but with the reduction of the 2n due to chromosomal fusions. In this sense, the other ESUs share a synapomorphic condition by presenting karyotypes mostly composed by acrocentric chromosomes, where pericentric inversions and/or centric fissions may have played a role. Such feature is also found in other *Bunocephalus* species, such as *B. doriae* (Fenocchio and Swarça, 2012) and *B. coracoideus* population from the Negro River (Ferreira et al., 2016), in

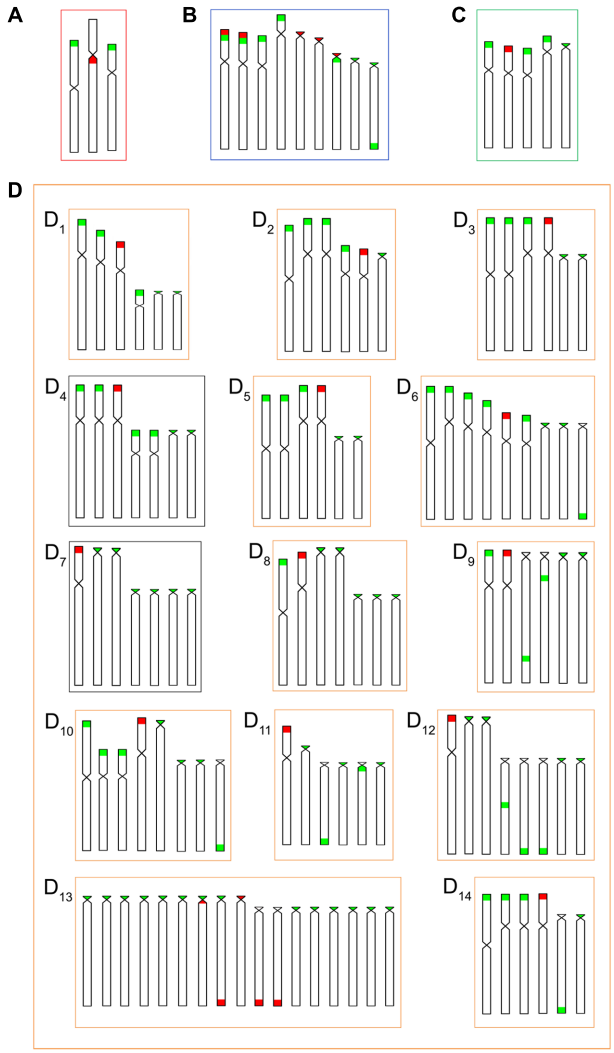


FIGURE 8 | Idiograms for numbers and locations of the 18S (red) and 5S (green) rDNA sequences in the chromosomes of the *B. coracoideus* populations **(A)** ESU-A Igarapé Jundiá – Cuieiras River, **(B)** SEU-B Demini River, **(C)** SEU-C Igarapé Apeú – Guamá River, and **(D)** SEU-D_{1-14} Purus River.

which acrocentric chromosomes are also mainly composing the karyotype. Thus, taking in account the Oliveira and Gosztonyi (2000) proposition, it can be considered that the *Bunocephalus* ESUs A, B, and D presents a trend toward the reduction of the chromosome number in relation to other Siluriformes, while ESU-C maintained the probable ancestral 2n = 56 chromosomes.

Chromosomal rearrangements can play a role on speciation processes as they may act in the reproductive isolation (Mayr, 1995), generating an useful investigation area concerning genetic variability (Faria and Navarro, 2010), Indeed, in some distinct Neotropical fish species, such as *Hoplias malabaricus* (Bertollo et al., 2000), *Astyanax fasciatus* (Pazza et al., 2006), *Astyanax scabripinnis* (Moreira-Filho and Bertollo, 1991; Maistro et al., 2000), and *Hoplerythinus unitaeniatus* (Giuliano-Caetano et al., 2001), the chromosomal diversification raised the hypothesis that they may encompass different species under a same nomenclature.

Additionally, both number and location of rDNA sites were highly variable among ESUs (**Figures 2**, **3**), highlighting their dynamic behavior in the genomes and in generation of the genetic diversity among populations. Besides, it seems that multiple 5S rDNA sites represent a synapomorphy in *B. coracoideus*, since all populations analyzed present such condition. Accordingly, ESU-A presents the lowest number of such sites (in only two pairs of chromosomes), thus representing a basal condition (**Figure 8**). Since the accumulation of repetitive sequences in particular genomic areas can cause chromosomal rearrangements (Lim and Simmons, 1994; Dimitri et al., 1997), the dynamic behavior of rRNA genes might also be linked with the huge karyotype diversity presented by this nominal species.

The occurrence of synteny between the 5S and 18S rRNA genes in the ESUs B and D (**Figure 8B**, D_{13}) represents an uncommon condition among vertebrates (Martins and Galetti, 2001), once these genes are transcribed by distinct RNA polymerases, suggesting the need to be distant from each other or allocated in different chromosomes, avoiding possible harmful rearrangements between them (Amarasinghe and Carlson, 1998; Martins and Galetti, 1999). However, in Siluriformes, the syntenic condition for such both rDNA classes was already found for several species, such as *Imparfinis mirini* and *I. minutus* (Ferreira et al., 2014), *Ancistrus maximus, A. ranunculus, A. dolichopterus, Ancistrus* aff. *dolichopterus* (Favarato et al., 2016), *Hemibagrus wyckii* (Supiwong et al., 2014), *Corydoras carlae* (Rocha et al., 2016), *Panaqolus* sp. (Ayres-Alves et al., 2017) and in *B. coracoideus* (present study). Nevertheless, the evolutionary paths taking to the selection of this apparently non-advantageous condition are not revealed yet but, at first, it appears to be not a deleterious character.

An outstanding finding in our study is the huge karyotype diversity found in the ESU D. In fact, this population highlighted many varying cytotypes living in sympatry. Apparently, such polymorphism does not appear to represent effective reproductive barriers capable to impair crosses, at least among some different cytotypes, increasing the chromosome diversity inside the population. Similar intrapopulacional features were also highlighted in the Characidae, *Astyanax fasciatus,* which presented two well-defined cytotypes, 2n = 46 and 2n = 48, but with numeric and structural chromosome variants when they occur in sympatry (Pazza et al., 2006, 2007, 2008). It is noteworthy that in both cases, ESU-D and *Astyanax fasciatus*, the variant karyotypes apparently do not demonstrate deleterious phenotypic effects on the carriers. However, the hypothesis that such degree of chromosomal diversity may affect, in some way, the homeostasis of the segregation cannot be fully discarded.

To better investigate the extension of the polymorphism inside ESU-D, we extended our analyzes to the chromosomal behavior during meiosis, since it was found monosomies and trisomies in nearly all cytotypes. In order to confirm this condition, meiotic plates of three individuals were analyzed, and attested that,

TABLE 2 | Aspredinidae specimens analyzed in the present study, with their respective collection places, number of individuals, diploid number (2n), and identification.

Species	Locality	Drainage	GPS data	Sampling N°	2n	ESU	GenBank access number	Voucher
B. coracoideus	Fazenda Dimona (Reserva PDBFF*), Amazonas, Brazil	Igarapé Jundiá – Cuieiras River	2°20′59.9″ S 60°05′50.9″ W	3♂2♀	42	A	MF416164–MF416167	INPA-ICT 053204
B. coracoideus	Barcelos, Amazonas, Brazil	Demini River	00°23.624′ S 062°48.187′ W	2♂3♀	44	B	MF416168–MF416172	INPA-ICT 053205
B. coracoideus	Castanhal, Pará, Brazil	Igarapé Apeú – Guamá River	1°23′20.4″ S 47°59′07.5″ W	2♂9♀	56	C	MF416173–MF416183	INPA-ICT 052188
B. coracoideus	Tapauá, Amazonas, Brazil	Purus River	5°37′21.7″ S 63°15′01.5′ W	2♂	40	D	MF416184–MF416199	INPA-ICT 052185
				1♂1♀	41	D		
				4♂6♀	42	D		
				3♂4♀	43	D		
				2♂2♀	44	D		
				1♂	45	D		
				1♂	46	D		
B. amaurus	Castanhal, Pará, Brazil	Igarapé Apeú – Guamá River	1°23′20.4″ S 47°59′07.5″ W	1?	–	Out group	MF416200	INPA-ICT 052187
B. cf. *aloikae*	São Gabriel da Cachoeira Amazonas, Brazil	Curicuriari River, Igarapé Bucu – Negro River	0°14′39.0″ S 67°03′31.1″ W	1?	–	Out group	MF416162	INPA-ICT 053207
A. hypsiura	Cametá, Pará, Brazil	Jutuba Island – Tocantins River	02°14′46.5″ S 49°24′59.7″ W	2?	–	Out group	MF416151–MF416152	INPA-ICT 052186

*Biological Dynamics of Forest Fragments Project.

during meiosis I, a clear numeric variation can be observed. In fact, different bivalent numbers were found in pachytene cells of the same individual, as well as probable trivalents with synapses points. In addition, a typical tetravalent formation and an apparent chromosomal chain were also observed in zygotene and diplotene cells, respectively (**Figure 5**), and such events might have contributed for irregular segregations. It is known that chromosomal rearrangements can alter the homologs pairing during meiosis and, as a consequence, provide unbalanced gametes (Davisson and Akeson, 1993; Navarro and Ruiz, 1997; Spirito, 1998). In this way, non-disjunction events during meiosis may result in aneuploid individuals, a factor that may, at least in part, explain the polymorphic condition found in the ESU-D population. In addition, a second factor probably related to such biodiversity relies on the ecological conditions in the Purus River basin, where ESU-D occurs. This region is located in a lowland area subjected to water flooding, influenced by the seasonality of the river level (Haugaasen and Peres, 2006). These flooded forests, which appear on the rainy season, form complexes labyrinths made by tree logs, rocks and every type of vegetation common to such environments (Luize et al., 2015). This particular habitat favors fish dispersion and the consequent subpopulations segregation until their future reconnection during the dry periods.

Thus, the evolutionary scenario for the ESU-D is that chromosomal rearrangements have occurred and that geographic isolation periods, due to flood pulse cycles may have favored their fixation in the population. During the flood periods, the reestablishment of the physical connection among the previously isolated aquatic environments allowed gene flow among them and, as a consequence, the variety of the cytotypes observed among the population. This hypothesis is reinforced by the ITS found in several cytotypes, indicating the occurrence of chromosomal rearrangements (**Figure 4**).

The DNA barcoding analysis is a very informative tool for biodiversity studies. In *Salminus* fish, for example, it was evidenced eight distinct lineages increasing its current diversity, nowadays limited to four species (Machado et al., 2016). *Rhamdia voulezi* and *Rhamdia branneri*, considered synonyms of *Rhamdia quelen*, are currently argued to constitute valid species supported by karyotype, ecomorphology and morphometric data (Abucarma and Martins-Santos, 2001; Garcia et al., 2010; Mise et al., 2013; Garavello and Shibatta, 2016), as well as by the barcoding DNA analysis (Ribolli et al., 2017).

Facing the karyotype diversity found in *B. coracoideus,* the DNA barcoding methodology was also useful for analyzing the relationships among populations. In fact, this procedure is a helpful tool for analyzing the occurrence of cryptic species (Smith et al., 2008). Theoretically, the nucleotide divergences between populations of a single species (intraspecific variations) are smaller than the ones between distinct species (interspecific variations), the "barcoding gap" (Ward et al., 2005; Hajibabaei et al., 2006). Most congeneric species have showed substantial nucleotide divergences by means of this molecular marker (Hebert et al., 2003). Intraspecific divergences are rarely superior to 2%, and usually do not overcome 1% (Avise, 2000). For *B. coracoideus* the intra-population genetic distance did not overcome the value of 2%, except for the Purus population (ESU-D), which presented divergences among the sequences from 0.2 to 10.3%.

Such molecular data corroborated the karyotype diversity allowed us to infer that there is a probable ongoing sympatric speciation process within this population. From the NJ analysis, all the ESUs were supported with bootstrap values higher than 96%. The same occurred with the phylogeny based on ML, except for the ESU-D, which presented a bootstrap value of 83%, reflecting once more the karyotype variation present among the specimens of this population. However, the high value observed supports its identity as an ESU.

The ESU-A presented a mean distance of 10.6% from the other ESUs (**Table 2**), which is a value equivalent to species differences. The bootstrap value of 62% of ML between the ESU-A and the super clade including ESU-B, ESU-C, and ESU-D is much lower to the 95% ML and 99.8% NJ to grouping them. Besides, in the NJ phylogram, the ESU-A is more related to *B.* cf. *aloikae* and *B. amaurus* than to the other ESUs. In addition to its particular karyotype features, ESU-A presents a high value of allopatric speciation and the potential of being a new species. According to Avise and Walker (1999), the high divergences among the ESUs of *B. coracoideus* allowed us to delimitate lineages with taxonomic uncertainties in this nominal species.

The genetic variability and the natural selection are important conditions for evolutionary changes. Thus, understanding the neutralization of the gene flow or the locking for factors that prevent gene exchanges, such as vicariance, gene mutations and chromosomal rearrangements, are important steps to explain evolutionary processes that frequently lead to speciation (Turelli et al., 2001; Kawakami et al., 2011). Indeed, it is well-known that mutations and chromosome rearrangements can be fixated by genetic drift and, more easily, in small and isolated populations (Jesus et al., 2016), as is the case for the *B. coracoideus* populations here investigated. However, the great challenge for genetic biodiversity analyzes is to preserve the connection with the natural history and the species nomenclature, with consequent implications on their management and conservation (Pellens et al., 2016). In this sense, a key question that emerges is how to classify the evolutionary history of a specific population concerning their genetic diversification. In fact, the description of new species, based on genetic diversity, still finds some resistance and is not yet fully adopted. In this way, many cryptic species remain undescribed, even after their identification by genetic markers (Schlick-Steiner et al., 2007).

CONCLUSION

The diversity of Neotropical freshwater fishes is still largely underestimated (Reis et al., 2016) and requires additional investigations. Nevertheless, a previous challenge remains still to be overcome: "what is a species and what new information is needed to solve this issue?" (Hey, 2001). Our study reveals a huge degree of chromosomal and genetic diversity in *B. coracoideus* and highly suggests the existence of four ESUs in allopatric and sympatric speciation processes. We believe that they were enough to reveal the occurrence of a *B. coracoideus* species complex. It indicates that new available methods, such as the genetic variability, can be definitely used in taxonomic procedures.

AUTHOR CONTRIBUTIONS

MF and CG performed techniques and analyzed the data; MF, JZ, and EF contributed with reagents, materials and analysis tools; MF, CG, DM, IdJ, MC, LB, JZ, and EF wrote the paper.

FUNDING

This study was supported by the Brazilian agencies, Instituto Nacional de Pesquisas da Amazônia (INPA), Fundação de Amparo à Pesquisa do Estado do Amazonas (UNIVERSAL AMAZONAS/FAPEAM 030/2013-062.00663/2015); CAPES pró Amazônia; ADAPTA-II INCT para Adaptação da Biota Aquática da Amazônia; Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 2016/22196-2); and received a productivity grant from CNPq (#313183/2014-7).

ACKNOWLEDGMENTS

The authors thank to Douglas Bastos for the collaboration during the acquirement of the specimens and to Leonardo Goll for the help in meiosis analysis. This is the study number 725 of the technical series of the Biological Dynamics of Forest Fragments Project – PDBFF (INPA/STRI).

REFERENCES

Abucarma, M., and Martins-Santos, I. C. (2001). Karyotype and B chromosome of *Rhamdia* species (Pisces, Pimelodidae) endemic in the River Iguaçu basin. *Cytologia* 66, 299–306. doi: 10.1508/cytologia.66.299

Amarasinghe, V., and Carlson, J. E. (1998). Physical mapping and characterization of 5S rRNA genes in douglas-fir. *Amer. Gen. Assoc.* 89, 495–500. doi: 10.1093/jhered/89.6.495

Avise, J. C. (2000). *Phylogeography: The History and Formation of Species*. Cambridge, MA: Harvard University Press.

Avise, J. C., and Walker, D. (1999). Species realities and numbers in sexual vertebrates: perspectives from an asexually transmitted genome. *Proc. Natl. Acad. Sci. U.S.A.* 96, 992–995. doi: 10.1073/pnas.96.3.992

Ayres-Alves, T., Cardoso, A. L., Nagamachi, C. Y., de Sousa, L. M., Pieczarka, J. C., and Noronha, R. C. R. (2017). Karyotypic evolution and chromosomal organization of repetitive DNA sequences in species of *Panaque*, *Panaqolus*, and *Scobinancistrus* (Siluriformes and Loricariidae) from the Amazon Basin. *Zebrafish* 14, 251–260. doi: 10.1089/zeb.2016.1373

Bertollo, L. A. C., Born, G. G., Dergam, J. A., Fenocchio, A. S., and Moreira-Filho, O. (2000). A biodiversity approach in the Neotropical fish, *Hoplias malabaricus*. Karyotypic survey, geographic distribution of cytotypes and cytotaxonomic considerations. *Chromosome Res.* 8, 603–613. doi: 10.1023/A:1009233907558

Bertollo, L. A. C., Takahashi, C. S., and Moreira-Filho, O. (1978). Cytotaxonomic considerations on *Hoplias lacerdae* (Pisces, Erytrinidae). *Braz. J. Biol.* 1, 103–120.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K., Meier, R., Winker, K., et al. (2007). Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* 22, 148–155. doi: 10.1016/j.tree.2006.11.004

Cardoso, A. R. (2008). *Filogenia da Família Aspredinidae Adams, 1854 e Revisão Taxonômica de Bunocephalinae Eigenmann & Eigenmann, 1888 (Teleostei: Siluriformes: Aspredinidae)*. doctoral's thesis, Pontifícia Universidade Católica do Rio Grande Do Sul, Porto Alegre.

Cardoso, A. R. (2010). Bunocephalus erondinae, a new species of banjo catfish from southern Brazil (Siluriformes: Aspredinidae). *Neotrop. Ichthyol.* 8, 607–613. doi: 10.1590/S1679-62252010000300005

Carvalho, T. P., Cardoso, A. R., Friel, J. P., and Reis, R. E. (2015). Two new species of the banjo catfish *Bunocephalus* Kner (Siluriformes: Aspredinidae) from the upper and middle rio São Francisco basins, Brazil. *Neotrop. Ichthyol.* 13, 499–512. doi: 10.1590/1982-0224-20140152

Cioffi, M. B., Kejnovský, E., Marquioni, V., Poltronieri, J., Molina, W. F., Diniz, D., et al. (2012). The key role of repeated DNAs in sex chromosome evolution in two fish species with ZW sex chromosome system. *Mol. Cytogenet.* 5:28. doi: 10.1186/1755-8166-5-28

Cioffi, M. B., Martins, C., and Bertollo, L. A. C. (2009). Comparative chromosome mapping of repetitive sequences. Implications for genomic evolution in the fish, *Hoplias malabaricus*. *BMC Genetics* 10:34. doi: 10.1186/1471-2156-10-34

Davisson, M. T., and Akeson, E. C. (1993). Recombination suppression by heterozygous Robertsonian chromosomes in the mouse. *Genetics* 133, 649–667.

Dimitri, P., Arcà, B., Berghella, L., and Mei, E. (1997). High genetic instability of heterochromatin after transposition of the LINE-like I factor in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 94, 8052–8057. doi: 10.1073/pnas.94.15.8052

Eschmeyer, W. N., Fricke, R., and van der Laan, R. (eds) (2017). *Catalog of Fishes: Genera, Species, References*. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp [accessed June 22, 2017].

Faria, R., and Navarro, A. (2010). Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends Ecol. Evol.* 25, 660–669. doi: 10.1016/j.tree.2010.07.008

Favarato, R. M., Silva, M., Oliveira, R. R., Artoni, R. F., Feldberg, E., and Matoso, D. A. (2016). Cytogenetic diversity and the evolutionary dynamics of rDNA genes and telomeric sequences in the *Ancistrus* genus (Loricariidae: Ancistrini). *Zebrafish* 13, 103–111. doi: 10.1089/zeb.2015.1140

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791. doi: 10.1111/j.1558-5646.1985.tb00420.x

Fenocchio, A. S., and Swarça, A. C. (2012). Cytogenetic description of *Bunocephalus doriae* Boulenger, 1902 (Siluriformes: Aspredinidae) from the Paraná River (Misiones, Argentina). *Neotrop. Ichthyol.* 10, 461–464. doi: 10.1590/S1679-62252012005000014

Ferreira, M., Garcia, C., Matoso, D. A., Jesus, I. S., and Feldberg, E. (2016). A new multiple sex chromosome system X1X1X2X2/X1Y1X2Y2 in Siluriformes: cytogenetic characterization of *Bunocephalus coracoideus* (Aspredinidae). *Genetica* 44, 591–599. doi: 10.1007/s10709-016-9927-9

Ferreira, M., Kavalco, K. F., De Almeida-Toledo, L. F., and Garcia, C. (2014). Cryptic diversity between two imparfinis species (Siluriformes, Heptapteridae) by cytogenetic analysis and DNA barcoding. *Zebrafish* 11, 306–317. doi: 10.1089/zeb.2014.0981

Garavello, J. C., and Shibatta, O. A. (2016). Reappraisal of Rhamdia branneri Haseman, 1911 and R. voulezi Haseman, 1911 (Siluriformes: Heptapteridae) from the rio Iguaçu with notes on their morphometry and karyotype. *Neotrop. Ichthyol.* 14:e140111. doi: 10.1590/1982-0224-20140111

Garcia, C., Oliveira, C., and Almeida-Toledo, L. F. (2010). Karyotypic evolution trends in *Rhamdia quelen* (Siluriformes, Heptapteridae) with considerations about the origin and differentiation of its supernumerary chromosomes. *Genet. Mol. Res.* 9, 365–384. doi: 10.4238/vol9-1gmr750

Giuliano-Caetano, L., Jorge, L. C., Moreira-Filho, O., and Bertollo, L. A. C. (2001). Comparative cytogenetic studies on *Hoplerythrinus unitaeniatus* populations (Pisces, Erythrinidae). *Cytologia* 66, 39–43. doi: 10.1508/cytologia.66.39

Gold, J. R., Li, C., Shipley, N. S., and Powers, P. K. (1990). Improved methods for working with fish chromosomes with a review of metaphase chromosome banding. *J. Fish Biol.* 37, 563–575. doi: 10.1111/j.1095-8649.1990.tb05889.x

Gross, M. C., Feldberg, E., Cella, D. M., Schneider, M. C., Schneider, C. H., Porto, J. I. R., et al. (2009). Intriguing evidence of translocations in *Discus* fish (Symphysodon, Cichlidae) and a report of the largest meiotic chromosomal chain observed in vertebrates. *Heredity* 102, 435–441. doi: 10.1038/hdy.2009.3

Gross, M. C., Schneider, C. H., Valente, G. T., Martins, C., and Feldberg, E. (2010). Variability of 18S rDNA locus among *Symphysodon* fishes: chromosomal rearrangements. *J. Fish Biol.* 76, 1117–1127. doi: 10.1111/j.1095-8649.2010.02550.x

Hajibabaei, M., Janzen, D. H., Burns, J. M., Hallwachs, W., and Hebert, P. D. N. (2006). DNA barcodes distinguish species of tropical Lepidoptera. *Proc. Natl. Acad. Sci. U.S.A.* 103, 968–971. doi: 10.1073/pnas.0510466103

Haugaasen, T., and Peres, C. A. F. (2006). Edaphic and structural characteristics of flooded and unflooded forests in the lower Rio Purus region of central Amazonia, Brazil. *Acta Amaz.* 36, 25–35. doi: 10.1590/S0044-59672006000100005

Hebert, P. D. N., Ratnasingham, S., and de Waard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. B Biol. Sci.* 270, S96–S99. doi: 10.1098/rsbl.2003.0025

Hey, J. (2001). The mind of the species problem. *Trends Ecol. Evol.* 16, 326–329. doi: 10.1016/S0169-5347(01)02145-0

Hey, J., Waples, R. S., Arnold, M. L., Butlin, R. K., and Harrison, R. G. (2003). Understanding and confronting species uncertainty in biology and conservation. *Trends Ecol. Evol.* 18, 597–603. doi: 10.1016/j.tree.2003.08.014

Ijdo, J. W., Wells, R. A., Baldini, A., and Reeders, S. T. (1991). Improved telomere detection using a telomere repeat probe (TTAGGG)n generated by PCR. *Nucleic Acids Res.* 19:4780. doi: 10.1093/nar/19.17.4780

Ivanova, N. V., Dewaard, J. R., and Hebert, P. D. N. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Mol. Ecol. Notes* 6, 998–1002. doi: 10.1111/j.1471-8286.2006.01428.x

Jesus, I. S., Ferreira, M., Garcia, C., Ribeiro, L. B., Alves-Gomes, J. A., and Feldberg, E. (2016). First cytogenetic description of *Microsternarchus bilineatus* (Gymnotiformes: Hypopomidae) from Negro River (Brazilian Amazon). *Zebrafish* 13, 571–577. doi: 10.1089/zeb.2016.1281

Kawakami, T., Butlin, R. K., and Cooper, S. J. (2011). Chromosomal speciation revisited: modes of diversification in Australian morabine grasshoppers (*Vandiemenella viatica* species Group). *Insects* 2, 49–61. doi: 10.3390/insects2010049

Kimura, M. (1980). A simple method of estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–120. doi: 10.1007/BF01731581

Levan, A., Fredga, K., and Sandberg, A. A. (1964). Nomenclature for centromeric position on chromosomes. *Hereditas* 52, 201–220. doi: 10.1111/j.1601-5223.1964.tb01953.x

Lim, J. K., and Simmons, M. J. (1994). Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster*. *Bioessays* 16, 269–275. doi: 10.1002/bies.950160410

Lis, J. T. (1980). Fractionation of DNA fragments by polyethylene glycol induced precipitation. *Methods Enzymol.* 65, 347–353. doi: 10.1016/S0076-6879(80)65044-7

Luize, B. G., Silva, T. S., Wittmann, F., Assis, R. L., and Venticinque, E. M. (2015). Effects of the flooding gradient on tree community diversity in várzea forests of the Purus River, Central Amazon, Brazil. *Biotropica* 47, 137–142. doi: 10.1111/btp.12203

Machado, C. D. B., Ishizuka, T. K., Freitas, P. D. D., Valiati, V. H., Galetti, P. M. Jr., et al. (2016). DNA barcoding reveals taxonomic uncertainty in *Salminus* (Characiformes). *Syst. Biodivers* 15, 372–382. doi: 10.1080/14772000.2016.1254390

Maistro, E. L., Oliveira, C., and Foresti, F. (2000). Sympatric occurrence of two cytotypes of *Astyanax scabripinnis* (Characiformes, Characidae). *Genet. Mol. Biol.* 23, 365–369. doi: 10.1590/S1415-47572000000200021

Martins, C., and Galetti, P. M. Jr. (1999). Chromosomal localization of 5S rDNA genes in *Leporinus* fish (Anostomidae, Characiformes). *Chromosome Res.* 7, 363–367. doi: 10.1023/A:1009216030316

Martins, C., and Galetti, P. M. Jr. (2001). Two 5S rDNA arrays in Neotropical fish species: is it a general rule for fishes? *Genetica* 111, 439–446.

Matoso, D. A., Almeida-Val, V. M. F., Silva, M., Moraes-Neto, A., Almeida, M. C., Vicari, M. R., et al. (2011). Chromosomal polymorphism in *Steindachneridion melanodermatum* Garavello, 2005 (Siluriformes, Pimelodidae) - a reappraisal the existence of sex chromosomal system in the species. *Rev. Fish Biol. Fish.* 21, 497–508. doi: 10.1007/s11160-011-9201-2

Mayr, E. (1995). "Species, classification, and evolution," in *Biodiversity and Evolution*, eds R. Arai and Y. Doi (Tokyo: Natural Science Museum Foundation), 3–12.

Mise, F. T., Tencatt, L. F. C., and Souza, F. (2013). Ecomorphological differences between Rhamdia (Bleeker, 1858) populations from the Iguaçu River basin. *Biota Neotrop.* 13, 99–104. doi: 10.1590/S1676-06032013000400010

Moreira-Filho, O., and Bertollo, L. A. C. (1991). *Astyanax scabripinnis* (Pisces, Characidae): a species complex. *Brazil J. Genet.* 14, 331–357.

Motta-Neto, C. C., Lima-Filho, P. A., Araujo, W., Bertollo, L. A. C., and Molina, W. (2012). Differentiated evolutionary pathways in Haemulidae (Perciformes): karyotype stasis versus morphological differentiation. *Rev. Fish Biol. Fish.* 22, 457–465. doi: 10.1007/s11160-011-9236-4

Navarro, A., and Ruiz, A. (1997). On the fertility effects of pericentric inversions. *Genetics* 147, 931–933.

Oliveira, C., and Gosztonyi, A. E. (2000). A cytogenetic study of *Diplotnystes mesembrinus* (Teleostei, Siluriformes, Diplomystidae) with a discussion of chromosome evolution in siluriforms. *Caryologia* 53, 31–37. doi: 10.1080/00087114.2000.10589178

Oliveira, E. A., Bertollo, L. A. C., Yano, C. F., Liehr, T., and Cioffi, M. B. (2015). Comparative cytogenetics in the genus *Hoplias* (Characiformes, Erythrinidae) highlights contrasting karyotype evolution among congeneric species. *Mol. Cytogenet.* 8:56. doi: 10.1186/s13039-015-0161-4

Pazza, R., Kavalco, K. F., and Bertollo, L. A. C. (2006). Chromosome polymorphism in *Astyanax fasciatus* (Teleostei, Characidae). 1. Karyotype analysis, Ag-NORs and mapping of the 18S and 5S ribosomal genes in sympatric karyotypes and their possible hybrid forms. *Cytogenet. Genome Res.* 112, 313–319. doi: 10.1159/000089886

Pazza, R., Kavalco, K. F., and Bertollo, L. A. C. (2008). Chromosome polymorphism in *Astyanax fasciatus* (Teleostei, Characidae). 2 – Chromosomal location of a satellite DNA. *Cytogenet. Genome Res.* 122, 61–66. doi: 10.1159/000151317

Pazza, R., Kavalco, K. F., Prioli, S. M. A. P., Prioli, A. J., and Bertollo, L. A. C. (2007). Chromosome polymorphism in *Astyanax fasciatus* (Teleostei, Characidae), Part 3: analysis of the RAPD and ISSR molecular markers. *Biochem. Syst. Ecol.* 35, 843–851. doi: 10.1016/j.bse.2007.03.018

Pellens, R., Faith, D. P., and Grandcolas, P. (2016). "The future of phylogenetic systematics in conservation biology: linking biodiversity and society," in *Biodiversity Conservation and Phylogenetic Systematics*, eds R. Pellens and P. Grandcolas (Basel: Springer International Publishing), 375–383.

Pinkel, D., Straume, T., and Gray, J. W. (1986). Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc. Natl. Acad. Sci.* 83, 2034–2038. doi: 10.1073/pnas.83.9.2934

Reis, R. E., Albert, J. S., Di Dario, F., Mincarone, M. M. M., Petry, P. L., and Rocha, L. R. (2016). Fish biodiversity and conservation in South America. *J. Fish Biol.* 89, 12–47. doi: 10.1111/jfb.13016

Ribolli, J., Scaranto, B. M., Shibatta, O. A., Bombardelli, R. A., and Zaniboni-Filho, E. (2017). DNA barcoding confirms the occurrence of *Rhamdia branneri* and *Rhamdia voulezi* (Siluriformes: Heptapteridae) in the Iguaçu River Basin. *Neotrop. Ichthyol.* 15:e160147. doi: 10.1590/1982-0224-20160147

Rocha, R. H., Baumgätner, L., Paiz, L. M., Margarido, V. P., Fernandes, C. A., and Gubiani, E. A. (2016). An uncommon co-localization of rDNA 5S with major rDNA clusters in Callichthyidae (Siluriformes): a report case in *Corydoras carlae* Nijssen & Isbrücker, 1983. *Compar. Cytogenet.* 10, 603–613. doi: 10.3897/CompCytogen.v10i4.9507

Schlick-Steiner, B. C., Seifert, B., Stauffer, C., Christian, E., Crozier, R. H., and Steiner, F. M. (2007). Without morphology, cryptic species stay in taxonomic crypsis following discovery. *Trends Ecol. Evol.* 22, 391–392. doi: 10.1016/j.tree.2007.05.004

Smith, M. A., Rodriguez, J. J., Whitfield, J. B., Deans, A. R., Janzen, D. H., Hallwachs, W., et al. (2008). Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proc. Natl. Acad. Sci. U.S.A.* 105, 12359–12364. doi: 10.1073/pnas.0805319105

Spirito, F. (1998). "The role of chromosomal change in speciation," in *Endless Forms: Species and Speciation*, eds D. J. Howard and S. H. Berlocher (New York, NY: Oxford University Press), 320–329.

Supiwong, W., Liehr, T., Cioffi, M. B., Chaveerach, A., Kosyakova, N., Fan, X., et al. (2014). Comparative cytogenetic mapping of rRNA genes among naked catfishes: implications for genomic evolution in the Bagridae family. *Genet. Mol. Res.* 13, 9533–9542. doi: 10.4238/2014.November.12.2

Tamura, K., Nei, M., and Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11030–11035. doi: 10.1073/pnas.0404206101

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Turelli, M., Barton, N. H., and Coyne, J. A. (2001). Theory and speciation. *Trends Ecol. Evol.* 16, 330–343. doi: 10.1016/S0169-5347(01)02177-2

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., and Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 360, 1–11. doi: 10.1098/rstb.2005.1716

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

Copyright © 2017 Ferreira, Garcia, Matoso, de Jesus, Cioffi, Bertollo, Zuanon and Feldberg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

METHODS
published: 25 September 2017
doi: 10.3389/fgene.2017.00131

A Flow Cytometry Protocol to Estimate DNA Content in the Yellowtail Tetra *Astyanax altiparanae*

*Pedro L. P. Xavier[1], José A. Senhorini[1], Matheus Pereira-Santos[2], Takafumi Fujimoto[3], Eduardo Shimoda[4], Luciano A. Silva[5], Silvio A. dos Santos[6] and George S. Yasui[1]**

[1] National Center for Research and Conservation of Continental Fish, Chico Mendes Institute of Biodiversity Conservation, Pirassununga, Brazil, [2] Aquaculture Center, Sao Paulo State University, Jaboticabal, Brazil, [3] Faculty of Fisheries Sciences, Hokkaido University, Hakodate, Japan, [4] Department of Pharmacy, Cândido Mendes University, Rio de Janeiro, Brazil, [5] Department of Veterinary Medicine, University of Sao Paulo, Pirassununga, Brazil, [6] AES Tietê, Promissão, Brazil

The production of triploid yellowtail tetra *Astyanax altiparanae* is a key factor to obtain permanently sterile individuals by chromosome set manipulation. Flow cytometric analysis is the main tool for confirmation of the resultant triploids individuals, but very few protocols are specific for *A. altiparanae* species. The current study has developed a protocol to estimate DNA content in this species. Furthermore, a protocol for long-term storage of dorsal fins used for flow cytometry analysis was established. The combination of five solutions with three detergents (Nonidet P-40 Substitute, Tween 20, and Triton X-100) at 0.1, 0.2, and 0.4% concentration was evaluated. Using the best solution from this first experiment, the addition of trypsin (0.125, 0.25, and 0.5%) and sucrose (74 mM) and the effects of increased concentrations of the detergents at 0.6 and 1.2% concentration were also evaluated. After adjustment of the protocol for flow cytometry, preservation of somatic tissue or isolated nuclei was also evaluated by freezing (at −20°C) and fixation in saturated NaCl solution, acetic methanol (1:3), ethanol, and formalin at 10% for 30 or 60 days of storage at 25°C. Flow cytometry analysis in yellowtail tetra species was optimized using the following conditions: lysis solution: 9.53 mM $MgCl_2.7H_20$; 47.67 mM KCl; 15 mM Tris; 74 mM sucrose, 0.6% Triton X-100, pH 8.0; staining solution: Dulbecco's PBS with DAPI 1 $\mu g\ mL^{-1}$; preservation procedure: somatic cells (dorsal fin samples) frozen at −20°C. Using this protocol, samples may be stored up to 60 days with good accuracy for flow cytometry analysis.

Keywords: yellowtail tetra, ploidy status, fish, flow cytometry, sample preservation

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Lenin Arias Rodriguez, Universidad Juárez Autónoma de Tabasco, Mexico
Diogo Teruo Hashimoto, Sao Paulo State University, Brazil

****Correspondence:***
George S. Yasui
yasui@usp.br

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

Received: *27 July 2017*
Accepted: *06 September 2017*
Published: *25 September 2017*

Citation:
Xavier PLP, Senhorini JA, Pereira-Santos M, Fujimoto T, Shimoda E, Silva LA, dos Santos SA and Yasui GS (2017) A Flow Cytometry Protocol to Estimate DNA Content in the Yellowtail Tetra Astyanax altiparanae. Front. Genet. 8:131. doi: 10.3389/fgene.2017.00131

INTRODUCTION

The yellowtail tetra, *Astyanax altiparanae,* has been highlighted as a great model fish for laboratory studies, aquaculture (Prioli et al., 2002; dos Santos et al., 2016), basics and applied research for reproductive technologies, conservation of valuable genetic resources and establishment of gene banks for endangered species using surrogate technologies (Yamaha et al., 2007).

Surrogate propagation permits a fish to produce gametes from another species, and in this case, the yellowtail tetra is strategic because this species can be manipulated to produce gametes from related endangered species. This technology requires sterile host individuals to ensure the production of exogenous gametes. Some protocols established the sterilization of the yellowtail

tetra by chronic exposure to high temperatures. However, sterilization was maintained only for a few weeks (de Siqueira-Silva et al., 2015). On the other hand, artificial induction of triploid by chromosome set manipulation using temperature shock has been promising in several fish species like tilapia and yellowtail tetra may be an efficient method to obtain permanent sterile individuals (Mair, 1993; Piferrer et al., 2009). In previous work (Adamov et al., 2016; do Nascimento et al., 2017), a protocol was established to produce triploid yellowtail tetra, and the ploidy status was confirmed by flow cytometry using a general protocol for plants and other animal species, but several samples could not be analyzed due to inaccurate analysis. It suggests that there is a specific protocol needed for this fish species.

Identification of triploids can be achieved using different methods including nuclear volume of fish erythrocytes (Allen and Stanley, 1978, 1979) and chromosome counts by karyotype (Allen et al., 1982). However, flow cytometry analysis is currently the most effective, rapid, and accurate method to identify the ploidy status. Flow cytometry analysis is a valuable tool for chromosomal studies in fish species, including cell cycle analysis, confirmation of the ploidy status, and determination of genome size in different fish species including the loach (Zhou et al., 2008), Atlantic salmon (Allen, 1983), *Poecilia formosa* (Lampert et al., 2008), and other teleost species (Ciudad et al., 2002; Zhu et al., 2012). However, the most protocols used for those analysis employs especially propidium iodide protocols and commercial kits that despite exhibit successfully results are usually expensive and not tested for fish species.

In such analysis, tissue sampling is critical since samples rapidly loss the viability if they are fresh and quality if previous fixatives was used for flow cytometry analysis. However, some times such samples are usually obtained in field conditions and have to be transported alive into the laboratory facilities for subsequent analysis. Furthermore, it is usually necessary to process large number of samples and such a situation require the development of protocols for long-term storage of these samples during the analysis. Some protocols are available for species such as the grass carp (Brown et al., 2000), Chinese grass carp (Burns et al., 1986), Atlantic salmon and shellfish (Allen, 1983), but there is not a protocol for Neotropical species like the yellowtail tetra.

Therefore, the aim of this study was firstly to establish a two-step protocol to estimate the DNA content in the yellowtail tetra, *A. altiparanae* by flow cytometry. In addition, an efficient protocol for long-term storage of tissue used for subsequent flow cytometry analysis was established. The combination of both procedures may give rise to a reliable determination of ploidy status in the yellowtail tetra and other important related species.

MATERIALS AND METHODS

All the procedures were accepted by the Guide for the Care and Use of Laboratory Animals of CEPTA (CEUA #02031.000033/2015-11).

General Procedure for Flow Cytometry

Adult yellowtail tetra *A. altiparanae* were collected from the Mogi-Guassu river, São Paulo State, Brazil (21.925706 S,

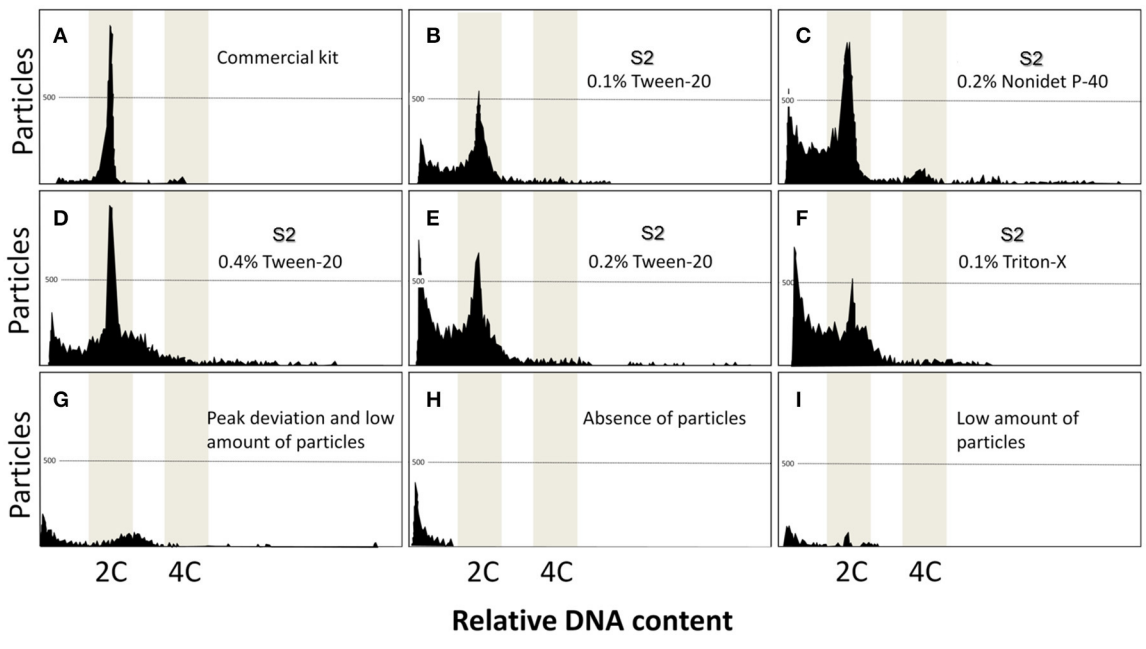


FIGURE 1 | Flow cytometric charts of fin samples using various lysing treatments and staining with DAPI. A combination of five solutions, and three detergents (Triton-X, Tween-20, and Nonidet *P*-40 at 0.1, 0.2, and 0.4%) gave rise to 45 treatments. Additionally, a commercial kit was used as control group. Commercial kit **(A)** produced the clear peaks. S2 associated with Tween-20 **(B,D,E)**, Nonidet **(C)** and Triton-X **(F)** gave some unclear peaks, demonstrating some potencial for analysis. Other treatments produced poor results including peak deviation **(G)**, absence of particles **(H)**, and low amount of particles **(I)**.

TABLE 1 | Evaluation of solutions and detergent concentrations for flow cytometry in fish.

Solution	Detergent	Concentration (%)	Relative DNA content	Number of Particles
Commercial kit (Control)	Unknown	Unknown	1.0 ± 0.0^{a}	9058.3 ± 9238.0^{a}
S1	Nonidet P-40	0.1	0.7 ± 0.6^{a}	217.0 ± 271.2^{abc}
		0.2	1.1 ± 0.0^{a}	451.7 ± 254.5^{abc}
		0.4	0.0 ± 0.0^{b}	0.0 ± 0.0^{c}
	Triton-X	0.1	0.7 ± 0.6^{a}	114.3 ± 125.1^{abc}
		0.2	0.7 ± 0.6^{a}	275.3 ± 396.7^{abc}
		0.4	1.0 ± 0.0^{a}	43.3 ± 27.3^{bc}
	Tween 20	0.1	0.0 ± 0.0^{b}	0.0 ± 0.0^{c}
		0.2	0.0 ± 0.0^{b}	0.0 ± 0.0^{c}
		0.4	0.7 ± 0.6^{a}	116.7 ± 108.3^{abc}
S2	Nonidet P-40	0.1	1.0 ± 0.0^{a}	235.0 ± 193.0^{abc}
		0.2	1.0 ± 0.0^{a}	980.0 ± 1249.7^{abc}
		0.4	1.0 ± 0.0^{a}	565.3 ± 298.1^{abc}
	Triton-X	0.1	1.0 ± 0.0^{a}	1342.3 ± 261.2^{abc}
		0.2	1.0 ± 0.0^{a}	389.0 ± 374.2^{abc}
		0.4	1.0 ± 0.0^{a}	43.7 ± 23.4^{bc}
	Tween 20	0.1	1.0 ± 0.0^{a}	820.3 ± 542.3^{abc}
		0.2	1.0 ± 0.0^{a}	1249.3 ± 978.6^{abc}
		0.4	1.0 ± 0.0^{a}	1555.0 ± 604.6^{abc}
S3	Nonidet P-40	0.1	1.0 ± 0.0^{a}	205.0 ± 120.2^{abc}
		0.2	1.0 ± 0.0^{a}	222.7 ± 107.6^{abc}
		0.4	0.9 ± 0.0^{a}	652.0 ± 275.0^{abc}
	Triton-X	0.1	0.3 ± 0.6^{a}	125.3 ± 217.1^{abc}
		0.2	0.7 ± 0.6^{a}	127.3 ± 145.6^{abc}
		0.4	0.0 ± 0.0^{b}	0.0 ± 0.0^{c}
	Tween 20	0.1	1.0 ± 0.0^{a}	1501.3 ± 827.4^{ab}
		0.2	1.0 ± 0.0^{a}	203.7 ± 203.0^{abc}
		0.4	0.3 ± 0.6^{a}	317.3 ± 148.1^{abc}
S4	Nonidet P-40	0.1	0.4 ± 0.6^{a}	155.3 ± 269.1^{abc}
		0.2	1.1 ± 0.0^{a}	594.7 ± 53.5^{abc}
		0.4	1.0 ± 0.0^{a}	820.3 ± 931.3^{abc}
	Triton-X	0.1	0.7 ± 0.6^{a}	176.7 ± 172.2^{abc}
		0.2	0.7 ± 0.6^{a}	538.0 ± 589.4^{abc}
		0.4	0.7 ± 0.6^{a}	596.3 ± 525.4^{abc}
	Tween 20	0.1	0.3 ± 0.6^{a}	294.3 ± 509.8^{ab}
		0.2	0.7 ± 0.6^{a}	741.3 ± 657.5^{abc}
		0.4	1.0 ± 0.0^{a}	352.7 ± 128.7^{abc}
S5	Nonidet P-40	0.1	1.1 ± 0.0^{a}	220.0 ± 86.3^{abc}
		0.2	1.0 ± 0.0^{a}	295.7 ± 213.3^{abc}
		0.4	1.0 ± 0.0^{a}	327.0 ± 350.9^{abc}
	Triton-X	0.1	1.0 ± 0.0^{a}	201.0 ± 170.6^{abc}
		0.2	1.1 ± 0.0^{a}	161.3 ± 78.1^{abc}
		0.4	1.0 ± 0.0^{a}	130.7 ± 154.7^{abc}
	Tween 20	0.1	1.1 ± 0.0^{a}	671.3 ± 469.8^{abc}
		0.2	0.7 ± 0.6^{a}	959.0 ± 1239.8^{abc}
		0.4	1.0 ± 0.0^{a}	590.7 ± 481.8^{abc}

Pieces of fin (~2 mm^2) were treated with the solutions and detergents combinations below, stained with DAPI, and analyzed by flow cytometry. Different letters within the columns indicates significant differences by Tukey test ($P < 0.05$).

47.369496 W). Males and females were used for the analysis, but no differences in DNA content between sexes were found for this species (Martinez et al., 2012). A small piece of dorsal fin (~2 mm^2) was clipped from the fish and then placed in a 1.5 mL microtube containing 120 μL of lysing solution. This lysing solution was prepared using detergents and other components (see following experiment sections). The samples were incubated at room temperature for 30 min with occasional mixing. Staining was achieved by addition 800 μL of Calcium-Free Dulbecco's PBS (Sigma #D5773) containing DAPI at 1 μg mL^{-1}. The samples were then filtered by a 30-μm mesh (Celltrics, Partec GmBH, Germany). Stained samples were then analyzed by a Partec CyFlow Plody Analyzer (Partec GmBH, Germany) with a specific filter set for DAPI excitation (358 nm). As a control group, a commercial kit specific for aquaculture and plants (Partec CyStain DNA 2-step, Partec GmBH, Germany) was used.

Solutions and Detergents for Cell Lysing

In this study, five (1–5) solutions (S) for cell lysing based on plant flow cytometry were compared: **S1** (Dpoolezel et al., 1989): 15 mM Tris, 2 mM Na_2EDTA, 80 mM KCl, 20 mM NaCl; **S2** (Arumuganathan and Earle, 1991): 9.53 mM $MgCl_4$, 47.67 mM KCl, 15 mM Tris; **S3** (Galbraith et al., 1983): 45 mM $MgCl_2$, 30 mM sodium citrate, 2 mM $NaHCO_3$; **S4** (Marie and Brown, 1993): 50 mM glucose, 15 mM KCl, 15 mM NaCl, 5 mM Na_2EDTA, 50 mM sodium citrate; **S5** (Pfosser et al., 1995): 200 mM tris, 4 mM $MgCl_2.6H_2O$. In each solution mentioned above, three detergents (Tween-20, triton-X, and nonidet P-40 substitute) were added at three concentrations each: 0.1, 0.2, and 0.4%. The combination of five solutions, three detergents at three concentrations, and the control group (commercial kit) gave rise to 46 lysing treatments. Dorsal fin samples were processed in each solution as mentioned above and then were checked for the best results according to the peak quality. This experiment was performed in triplicates.

Evaluation of High Concentrations of Detergents, Trypsin, and Sucrose Additions

Using the best combination of detergents and the solution obtained above (i.e., Arumuganathan and Earle, 1991), detergent concentrations at 0.15, 0.3, 0.6, and 1.2% were evaluated with or without the addition of sucrose (74 mM). Using the best result,

the effects of trypsin addition (at 0.125, 0.25, and 0.5%) were also evaluated on peak quality.

Fixatives and Preservation Strategies

In this experiment, two kinds of samples (fish fin or isolated nuclei) were evaluated. For each of them, five procedures were evaluated: freezing at −20°C, ethanol 70%, acetic methanol (1 part of acetic acid:3 parts of methanol), saline solution (saturated NaCl solution), and formalin at 10%. For preservation of tissue, about ∼2 mm^2 of dorsal fin from the *A. altiparanae* were placed in 1.5 mL microtubes containing 1 mL of each of fixatives and maintained at 25°C in a BOD incubator under dark conditions. For freezing, the sample was placed in a 1.5 mL microtube containing 200 μL of 0.9% NaCl and then directly frozen at −20°C in a biomedical freezer. Isolated nuclei were obtained from dorsal fins (∼2 mm^2) using 100 μL of lysing solution obtained in previous experiment. For preservation of nuclei, 1 mL of each fixative was added to each nuclei suspension and then maintained at 25°C. Nuclei suspensions were also frozen at −20°C. Flow cytometry analysis were performed at 30 and 60 days of storage. Fixed tissue was collected and washed in Dulbecco's PBS and then processed in lysing solution before staining. Nuclei suspensions were centrifuged at 12,000 G for 2 min, and the fixatives were removed by pipetting. The nuclei pellet was re-suspended in staining solution and then analyzed on flow cytometer. As control groups, fresh (non-fixed) samples from day 0, 30, and 60 were used. For each of the 21 treatments, 10 replications were used.

Statistical Analysis

Data are shown as mean ± SD. Data were checked for normality using the Liliefor test and also compared using ANOVA followed by Tukey's multiple range test ($P < 0.05$). Statistica version 11 were used for statistical analysis.

RESULTS

Solutions and Detergents for Cell Lysing

In the control group that used commercial kit, two clear peaks arose: one with 2C content and the other with 4C, originating from dividing cells (**Figure 1A**). On the other hand, none of the combinations of the solutions and detergents evaluated generated clear peaks. Problems included noisy peaks (**Figures 1B–F**), deviation in the DNA content (**Figure 1G**), absence of peaks (**Figure 1H**), and in most cases, low amounts of particles were found (**Figure 1I**). As observed in **Table 1**, the only solution that kept the same relative DNA content in comparison to the control, independent of the detergent or concentration, was the

TABLE 2 | Evaluation of trypsin additions for flow cytometry in fish.

Solution	Detergent	Concentration (%)	Trypsin addition (%)	Relative DNA content	Number of Particles
S2 + 74 mM sucrose	Triton-X	0.6	0	1.00 ± 0.01^{a}	715.0 ± 318.2^{a}
			0.125	1.04 ± 0.01^{a}	400.7 ± 358.5^{ab}
			0.25	1.03 ± 0.02^{a}	139.7 ± 86.6^{ab}
			0.5	0.99 ± 0.03^{a}	74.3 ± 24.0^{b}

S2 with 0.6% Triton-X was combined with trypsin at 0.125, 0.25, and 0.5% and 74 mM Sucrose and measured the relative DNA content and number of isolated particles. Pieces of fin (2 mm^2) were treated with the solutions, stained with DAPI and analyzed by flow cytometry. Different letters within the columns indicates significant differences by Tukey test ($P < 0.05$).

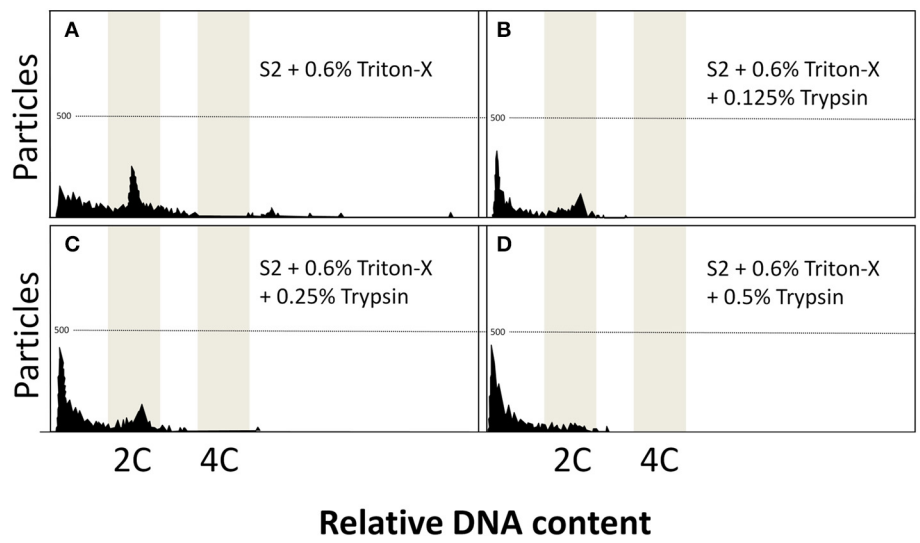


FIGURE 2 | Flow cytometric charts of fin samples using various lysing treatments and staining with DAPI. Solution S2 was combined with three detergents (Triton-X, Tween-20, and Nonidet P-40 at 0.15, 0.3, 0.6, and 1.2%). Trypsin additions at 0.125, 0.25, and 0.5% was also evaluated. Triton-X at 0.6% gave some noisy peaks **(A)**. Trypsin additions produced reduced peaks **(B–D)**.

S2. Furthermore, the S2 solution with Triton at 0.1% presented a good number of particles isolated and the best peak quality in terms of the histograms (**Table 1**). All other solutions presented poor results in regards to quality of peaks, DNA content, and number of particles. Considering the quality of peaks, DNA content, and the number of particles, S2 presented the best results and was used in later experiments.

Evaluation of High Concentrations of Detergents, Trypsin and Sucrose Additions

As observed in **Figure 2**, the addition of trypsin at 0.125 (**Figure 2B**), 0.25 (**Figure 2C**), and 0.5% (**Figure 2D**) presented detrimental effects on the flow cytometry analysis, and lower peaks arose with increasing concentrations of trypsin. Such data coincided with lower amounts of particles (**Table 2**), although it was significantly decreased only at 0.5% (74.3 ± 24.0, $P = 0.0491$). The DNA content did not present any statistical differences ($P = 0.0578$) and ranged from 0.99 ± 0.03 (at 0.5% trypsin) to 1.04 ± 0.01 (at 0.125% trypsin).

Flow cytometry using S2 solution associated with triton-X at 0.15, 0.3, and 0.6% gave rise to noisy peaks (**Figures 3A–C**). The addition of sucrose at 74 mM associated with triton-X at the same concentrations produced clearer peaks, and at 0.3% (**Figure 3E**) and 0.6% (**Figure 3F**), but not at 0.15% (**Figure 3D**), two clear 2 and 4C peaks arose. Relative DNA content (**Table 3**) was decreased in Triton-X at 1.2% (0.85 ± 0.03 C, $P = 0.0155$), and sucrose with Triton-X at 0.15% (0.87 ± 0.02 C, $P = 0.0492$), 0.3% (0.85 ± 0.01 C, $P = 0.0155$), and 0.6% (0.84 ± 0.02 C; $P = 0.1054$). The number of particles from treatment S2 with triton-X at 0.125% was 927.0 ± 132.4 and did not present statistical differences ($P = 0.0538$) when compared with the control (2,716.3 ± 1,822.2). However, other treatments had a significant decrease and ranged from 194.7 ± 37.1 (S2 and triton-X at 0.6%) to 826.3 ± 67.0 (S2 with 0.3% triton X and 74 mM sucrose).

As there were promising results with 0.6% triton-X, higher concentrations of other detergents at 0.6 and 1.2% were also evaluated. The results using S2 and sucrose presented good results and Triton-X, Tween-20, and Nonidet P-40 presented similar results, both at 0.6 and 1.2% (**Figure 4**), with a main peak at 2C and a secondary peak at 4C. The DNA content did not present significant differences ($P = 0.0779$) within all the detergents and concentrations evaluated in this study and ranged from 0.99 ± 0.01 C (1.2% Nonidet P-40) to 1.03 ± 0.01 C (1.2% Tween-20). Similarly, all the treatments did not present significant differences regarding the number of particles ($P = 0.0602$), and this parameter ranged from 725.3 ± 119.0 (0.6% triton) to 1095.3 ± 183.4 (1.2% Tween-20).

Based on the results above, Triton-X at 0.6% gave better results on nuclei isolation (**Table 4**) and subsequent flow cytometric analysis (**Figure 4C**) and, thus, was then chosen for later experiments.

Fixatives and Preservation Strategies

After 30 and 60 days of fixation, only frozen dorsal fin samples maintained the ability for flow cytometric analysis (**Figure 5**).

TABLE 3 | Evaluation of sucrose addition for flow cytometry in fish.

Solution	Detergent	Concentration (%)	Relative DNA content	Number of Particles
Commercial kit (Control)	Unknown	Unknown	1.00 ± 0.08^{a}	2716.3 ± 1822.2^{a}
S2	Triton-X	0.15	0.92 ± 0.04^{ab}	927.0 ± 132.4^{ab}
		0.30	0.88 ± 0.02^{ab}	343.3 ± 209.2^{b}
		0.60	0.88 ± 0.04^{ab}	194.7 ± 37.1^{b}
		1.20	0.85 ± 0.03^{c}	176.0 ± 257.3^{b}
S2 + 74 mM Sucrose	Triton-X	0.15	0.87 ± 0.02^{c}	369.3 ± 337.8^{b}
		0.30	0.85 ± 0.01^{c}	826.3 ± 67.0^{b}
		0.60	0.84 ± 0.02^{c}	776.3 ± 100.2^{b}
		1.20	0.93 ± 0.09^{ab}	341.7 ± 139.5^{b}

S2 was associated Triton-X at 0.15, 0.30, 0.60, 1.20% and sucrose at 74 mM and measured the DNA content and number of particles. Pieces of fin (2 mm^2) were treated with the solutions, stained with DAPI and analyzed by flow cytometry. Different letters within the columns indicates significant differences by Tukey test ($P < 0.05$).

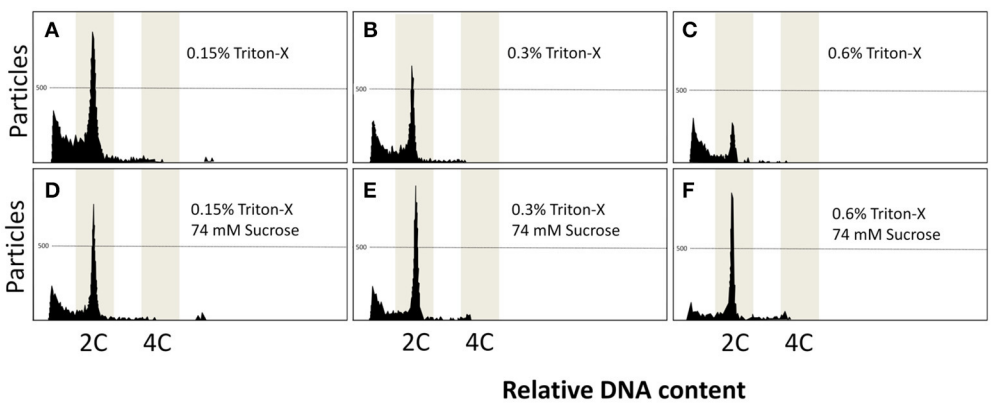


FIGURE 3 | (A–F) Flow cytometric charts of fin samples using various lysing treatments and staining with DAPI. A combination of Triton-X 0.15, 0.3, and 0.6% was associated with sucrose at 74 mM. Triton-X produced noisy peaks at all concentrations. The addition of sucrose improved the results, giving rise to clear peaks.

Very clear peaks arose (**Figures 5B,C**), like samples from the control groups with fresh tissue (**Figure 5A**). Preserved dorsal fin using 70% ethanol produced peaks after 30 and 60 days (**Figures 5D,E**), but deviation of the peak occurred in both cases. Other fixatives gave rise to noisy peaks or, in most cases, did not show any peak (**Figure 5F**). The preservation of isolated nuclei as a new strategy failed in all cases giving rise to noisy or very few particles.

After preservation in 70% ethanol, fin samples presented increased DNA content in flow cytometry analysis after 30 days (1.14 ± 0.04; $P = 0.0001$) and 60 days (1.21 ± 0.06 C; $P = 0.0001$) when compared with control group (1.00 ± 0.02 C) (**Table 5**). Nucleus preserved in formaldehyde after 30 days presented decreased DNA content (0.86 ± 0.04; $P = 0.0001$). A similar result was observed for fin samples preserved with formaldehyde for 60 days (0.80 ± 0.07; $P = 0.0001$) and nucleus preserved for 60 days with salt (0.89 ± 0.10 C; $P = 0.0011$), formaldehyde (0.78 ± 0.04 C; $P = 0.0001$), ethanol (0.72 ± 0.03 C; $P = 0.0001$), and methanol (0.72 ± 0.05 C; $P = 0.0001$).

The number of particles was reduced in all cases when compared with fresh samples (2215.6 ± 1015.9; $P = 0.0001$). However, the best results were achieved with frozen dorsal fins after 30 days (892.9 ± 357.9) and 60 days (862.2 ± 467.8).

DISCUSSION

Successful flow cytometry analysis depends on several aspects including the interaction of the sample and the components of lysing solutions (Loureiro et al., 2007). An adequate combination of osmolality, buffers, stabilizers, and detergents is then required to achieve good results. A two-step procedure with cell lysing and nuclear staining, solubilization of cell membrane and subsequent preservation of nuclear envelope is the key point for flow cytometry, and combination of detergents, solutions, and membrane stabilizer was successful for the studied species. The current results indicates that solutions containing magnesium, presented better results during flow cytometry analysis. This is assumed to be the case because of magnesium ions work as a cromatin stabilizer (Galbraith et al., 1983) and then we conclude that this component is important in order to preserve nuclear membrane of the yellowtail tetra. Similarly, sucrose additions presented good results. Sucrose affects osmolality and presents membrane stabilizer ability and is also widely used in several process including cell lysis and cryopreservation ir order to maintain the membrane characteristics (Medina-Robles et al., 2005). Furthermore, sucrose is important for the maintenance of nuclear integrity (Marie and Brown, 1993). Non-ionic detergents used in this study efficiently solubilized cell membrane, but the nuclei were kept intact, indicating that lysing and stabilizing process is membrane-specific.

In this work, a protocol for sample storage and subsequent flow cytometry analysis was achieved. Such a procedure is applicable for long-term preservation for future analysis. The preservation of marine bacteria under fixation with paraformaldehyde and storage in liquid nitrogen reduced the cell count and generated multiple peaks in flow cytometry analysis (Kamiya et al., 2007). Similar decrease in quality was observed in marine bacteria and algae (Troussellier et al., 1993). In this study, when sub-zero temperatures at -20° C were used the peak

TABLE 4 | Evaluation of S2 solution associated with high concentrations of detergents (0.6 and 1.2%) regarding the relative DNA content and number of isolated particles.

Solution	Detergent	Concentration (%)	Relative DNA content	Number of particles
S2	Nonidet P-40	0.6	0.99 ± 0.01^{a}	829.0 ± 45.9^{a}
		1.2	0.99 ± 0.01^{aa}	763.0 ± 175.3^{a}
	Triton-X	0.6	1.02 ± 0.02^{a}	725.3 ± 119.0^{a}
		1.2	0.99 ± 0.03^{a}	805.3 ± 189.6^{a}
	Tween 20	0.6	1.01 ± 0.01^{a}	1090.7 ± 228.7^{a}
		1.2	1.03 ± 0.01^{ac}	1095.3 ± 183.4^{a}

Pieces of fin (2 mm^2) were treated with the solutions, stained with DAPI, and analyzed by flow cytometry. Different letters within the columns indicates significant differences by Tukey test ($P < 0.05$).

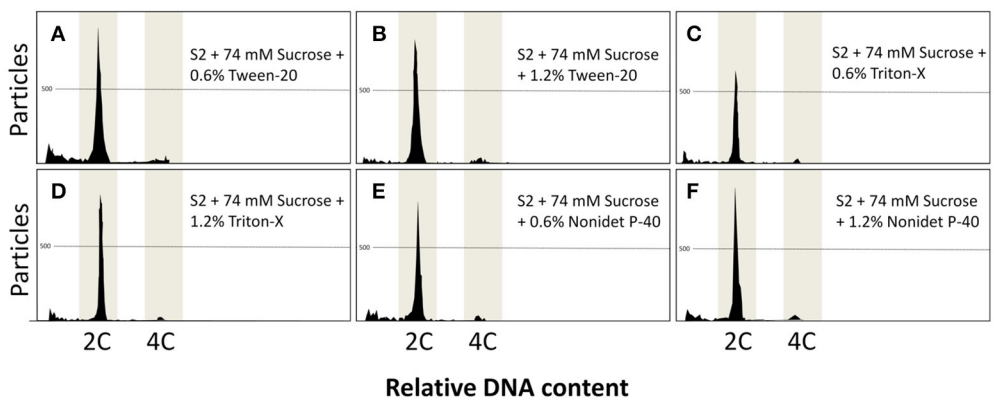


FIGURE 4 | (A–F) Flow cytometric charts of fin samples using various lysing treatments and staining with DAPI. A combination of solution S2 with Triton-X, Tween-20, and Nonidet P-40 at 0.6 and 1.2% and 74 mM sucrose was evaluated. All treatments gave rise to clear peaks.

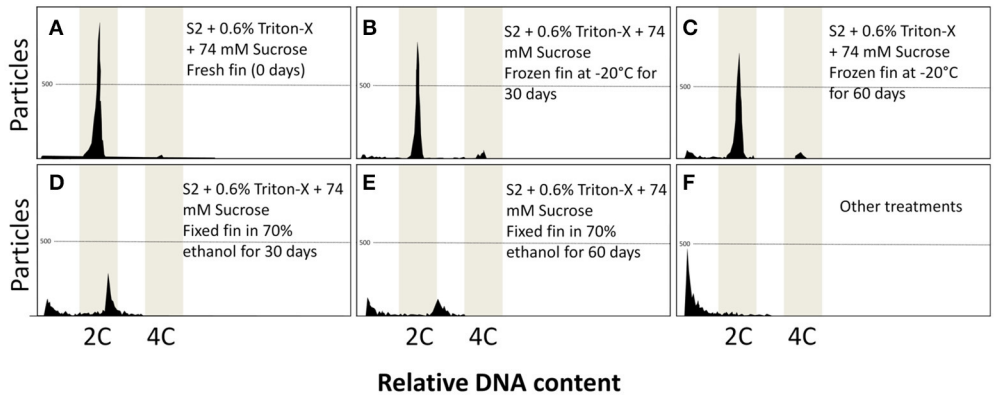


FIGURE 5 | Flow cytometric charts of fin samples and isolated nuclei preserved using various procedures at 0, 30, and 60 days of storage. Frozen fins at −20°C gave the best results and storage at 30 days **(B)** and 60 days **(C)** did not differ from fresh samples (0 days) **(A)**. Fixation of fin in 70% ethanol gave peaks with low amount of particles and peak deviation at 30 days **(D)** and 60 days **(E)**. Other treatments did not generate any data **(F)**.

TABLE 5 | Preservations strategies (Salt, Formaldehyde, Ethanol, Methanol, and Freeze) for flow cytometry in fish.

Storage	Sample	Treatment	RelativeDNA content	Number of particles
0 days	Fin	Control	1.00 ± 0.02^{cde}	2215.60 ± 1015.94^{a}
30 days	Fin	Salt	0.98 ± 0.10^{def}	13.80 ± 17.27^{c}
		Formaldehyde	0.94 ± 0.03^{efgh}	32.00 ± 21.58^{c}
		Ethanol	1.14 ± 0.04^{ab}	13.30 ± 18.96^{c}
		Methanol	1.07 ± 0.02^{bc}	8.10 ± 8.32^{c}
		Freeze	1.01 ± 0.02^{cde}	892.90 ± 357.95^{b}
	Nucleus	Salt	0.99 ± 0.06^{cdef}	8.90 ± 11.11^{c}
		Formaldehyde	0.86 ± 0.04^{hi}	48.20 ± 24.50^{c}
		Ethanol	1.03 ± 0.03^{cd}	102.80 ± 27.47^{c}
		Methanol	0.94 ± 0.05^{efgh}	222.90 ± 207.72^{c}
		Freeze	0.99 ± 0.05^{cdef}	42.90 ± 60.24^{c}
0 days	Fin	Control		
60 days	Fin	Salt	0.96 ± 0.03^{defg}	0.60 ± 0.70^{c}
		Formaldehyde	0.80 ± 0.07^{ij}	15.30 ± 9.45^{c}
		Ethanol	1.21 ± 0.06^{a}	25.70 ± 28.86^{c}
		Methanol	0.98 ± 0.08^{def}	4.90 ± 8.84^{c}
		Freeze	0.98 ± 0.02^{def}	862.20 ± 467.83^{b}
	Nucleus	Salt	0.89 ± 0.10^{gh}	3.50 ± 4.33^{c}
		Formaldehyde	0.78 ± 0.04^{ij}	36.90 ± 15.15^{c}
		Ethanol	0.72 ± 0.03^{j}	146.20 ± 44.50^{c}
		Methanol	0.72 ± 0.05^{j}	21.20 ± 10.67^{c}
		Freeze	0.91 ± 0.06^{fgh}	5.40 ± 4.17^{c}

Pieces of fin (2 mm^2) and isolated nuclei were fixed with different procedures. The samples were analyzed at 0 (control), 30 and 60 days by flow cytometry and obtained the relative DNA content and number of isolated particles. Different letters within the columns indicates significant differences by Tukey test ($P < 0.05$).

quality and cell concentration was maintained for 60 days of storage providing the good results in flow cytometry. Such a new procedure for preservation of fish samples may be used in many fields including aquaculture and biomedicine.

A new procedure for fixation of samples for flow cytometry was evaluated, including isolated nuclei instead of tissue, however, further work is required to improve the method, results as well as the alternative comparisons of different fixatives and ways of storage. However, such an approach of preserving an isolated nucleus may be interesting because fixation and storage of tissue may affect membrane characteristics (Suganuma and Morioka, 1979). It can also reduce the efficiency of lysing and staining procedures, and, thus, the nucleus cannot be isolated in some cases. Nuclear isolation and preservation may be applicable as a preservation procedure for flow cytometry analysis.

In conclusion, a simple and inexpensive protocol for DNA content analysis by flow cytometry in yellowtail tetra *A. altiparanae* and sample preservation was established, as follows: lysing solution for fish samples composed by 9.53 mM $MgCl_2.7H_2O$, 47.67 mM KCl, 15 mM Tris, 74 mM sucrose, and 0.6% Triton-X. Tissue samples of fish may be preserved at −20°C for 60 days for future analysis using such a protocol.

AUTHOR CONTRIBUTIONS

PX: Acquisition, analysis, and interpretation of data, draft of the work, development of intellectual content, writing of the manuscript, final approval of the version. JS: Draft of the work, developmente of intellectual content, final approval of the version. MP: Analysis and interpretation of data, draft of the work, development of intellectual content, final approval of the version. TF: Draft of the work, development of intellectual content, final approval of the version. ES: Analysis of data, development of intellectual content, final approval of the version. LS: Interpretation of data, development of intellectual content, writing of the manuscript, final approval of the version. SdS: Interpretation of data, development of intellectual content, writing of the manuscript and final approval of the version. GY: Analysis and interpretation of data, draft of the work, development of intellectual content, writing of the manuscript, final approval of the version.

ACKNOWLEDGMENTS

Authors are grateful to Sao Paulo Research Foundation (FAPESP) for the financial support of this research (Young Investigators Award Grant #2010/17429-1 and Young Researcher Scholarship #2011/11664-1); AES Tietê (Research & Development Project #0064-1052/2014) and FUNDIBIO; Michael Stablein for the english review of the manuscript. We also acknowledge CEPTA-ICMBio for gently provide the facilities and experimental fish.

REFERENCES

Adamov, N. S. M., Nascimento, N. F., Maciel, C. S. E., Santos, M. P., Senhorini, J. A., Calado, L. L., et al. (2016). Triploid induction in the *Yellowtail tetra*, *Astyanax altiparanae*, using temperature shock: tools for conservation and aquaculture. *J. World Aquac. Soc.* doi: 10.1111/jwas.12390. [Epub ahead of print].

Allen, S. K. (1983). Flow cytometry: assaying experimental polyploid fish and shellfish. *Aquaculture* 33, 317–328. doi: 10.1016/0044-8486(83)90412-X

Allen, S. K., and Stanley, J. G. (1978). Reproductive sterility in polyploid brook trout, salvelinus fontinalis. *Trans. Am. Fish. Soc.* 107, 473–478. doi: 10.1577/1548-8659(1978)107<473:RSIPBT>2.0.CO;2

Allen, S. K., and Stanley, J. G. (1979). Polyploid mosaics induced by cytochalasin b in landlocked atlantic salmon salmo salar. *Trans. Am. Fish. Soc.* 108, 462–466. doi: 10.1577/1548-8659(1979)108<462:PMIBCB>2.0.CO;2

Allen, S. K., Gagnon, P. S., and Hidu, H. (1982). Induced triploidy in the soft-shell clam: cytogenetic and allozymic confirmation. *J. Hered.* 73, 421–428. doi: 10.1093/oxfordjournals.jhered.a109691

Arumuganathan, K., and Earle, E. D. (1991). Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* 9, 208–218. doi: 10.1007/BF02672069

Brown, B. L., Schultz, S. L. W., and White, F. K. H. (2000). A convenient field method of tissue preservation for flow cytometric ploidy assessment of grass carp. *Trans. Am. Fish. Soc.* 129, 1354–1359. doi: 10.1577/1548-8659(2000)129<1354:ACFMOT>2.0.CO;2

Burns, E. R., Anson, J. F., Hinson, W. G., Pipkin, J. L., Kleve, M. G., and Goetz, R. C. (1986). Effect of fixation with formalin on flow cytometric measurement of DNA in nucleated blood cells. *Aquaculture* 55, 149–155. doi: 10.1016/0044-8486(86)90070-0

Ciudad, J., Cid, E., Velasco, A., Lara, J. M., Aijón, J., and Orfao, A. (2002). Flow cytometry measurement of the DNA contents of G0/G1 diploid cells from three different teleost fish species. *Cytometry* 48, 20–25. doi: 10.1002/cyto.10100

de Siqueira-Silva, D. H., Silva, A. P. S., Ninhaus-Silveira, A., and Veríssimo-Silveira, R. (2015). The effects of temperature and busulfan (Myleran) on the *Yellowtail tetra Astyanax altiparanae* (Pisces, Characiformes) spermatogenesis. *Theriogenology* 84, 1033–1042. doi: 10.1016/j.theriogenology.2015.06.004

do Nascimento, N. F., Pereira-Santos, M., Piva, L. H., Manzini, B., Fujimoto, T., Senhorini, J. A., et al. (2017). Growth, fatty acid composition, and reproductive parameters of diploid and triploid yellowtail tetra *Astyanax altiparanae*. *Aquaculture* 471, 163–171. doi: 10.1016/j.aquaculture.2017.01.007

dos Santos, M. P., Yasui, G. S., Xavier, P. L. P., Adamov, N. S. M., Nascimento, N. F., Fujimoto, T., et al. (2016). Morphology of gametes, post-fertilization events and the effect of temperature on the embryonic development of *Astyanax altiparanae* (Teleostei, Characidae). *Zygote* 24, 1–13. doi: 10.1017/S0967199416000101

Dpoolezel, J., Binarova, P., and, S., Lucretti (1989). Analysis of nuclear DNA content in plant cells by flow cytometry. *Biol. Plant.* 31, 113–120. doi: 10.1007/BF02907241

Galbraith, D. W., Harkins, K. R., Maddox, J. M., Ayres, N. M., Sharma, D. P., and Firoozabady, E. (1983). Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* 220, 1049–1051. doi: 10.1126/science.220.4601.1049

Kamiya, E., Izumiyama, S., Nishimura, M., Mitchell, J. G., and Kogure, K. (2007). Effects of fixation and storage on flow cytometric analysis of marine bacteria. *J. Oceanogr.* 63, 101–112. doi: 10.1007/s10872-007-0008-7

Lampert, K. P., Lamatsch, D. K., Fischer, P., and Schartl, M. (2008). A tetraploid amazon molly, *Poecilia formosa*. *J. Hered.* 99, 223–226. doi: 10.1093/jhered/esm102

Loureiro, J., Rodriguez, E., Dolezel, J., and Santos, C. (2007). Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. *Ann. Bot.* 100, 875–888. doi: 10.1093/aob/mcm152

Mair, G. C. (1993). Chromosome-set manipulation in tilapia - techniques, problems and prospects. *Aquaculture* 111, 227–244. doi: 10.1016/0044-8486(93)90040-6

Marie, D., and Brown, S. C. (1993). A cytometric exercise in plant DNA histograms, with 2C values for 70 species. *Biol. Cell* 78, 41–51. doi: 10.1016/0248-4900(93)90113-S

Martinez, E. R. M., Alves, A. L., Silveira, S. M., Foresti, F., and Oliveira, C. (2012). Cytogenetic analysis in the incertae sedis species *Astyanax altiparanae* Garutti and Britzki, 2000 and *Hyphessobrycon eques* Steindachner, 1882 (Characiformes, Characidae) from the upper paraná river basin. *Comp. Cytogenet.* 6, 41–51. doi: 10.3897/compcytogen.v6i1.1873

Medina-Robles, V. M., Velasco-Santamaría, Y. M., and Cruz-Casallas, P. E. (2005). Aspectos generales de la crioconservación espermática en peces teleósteos. *Rev. Col. Cienc. Pec.* 18, 34–48. Available online at: http://www.scielo.org.co/pdf/rccp/v18n1/v18n1a05.pdf

Pfosser, M., Amon, A., Lelley, T., and Heberle-Bors, E. (1995). Evaluation of sensitivity of flow cytometry in detecting aneuploidy in wheat using disomic and ditelosomic wheat-rye addition lines. *Cytometry* 21, 387–393. doi: 10.1002/cyto.990210412

Piferrer, F., Beaumont, A., Falguière, J. C., Flajšhans, M., Haffray, P., and Colombo, L. (2009). Polyploid fish and shellfish: production, biology and applications to aquaculture for performance improvement and genetic containment. *Aquaculture* 293, 125–156. doi: 10.1016/j.aquaculture.2009.04.036

Prioli, S. M. A. P., Prioli, A. J., Júlio, F. J., Pavanelli, C. S., Oliveira, A. V., Carrer, H., et al. (2002). Identification of *Astyanax altiparanae* (Teleostei, Characidae) in the Iguaçu River, Brazil, based on mitochondrial, DNA and RAPD markers. *Genet. Mol. Biol.* 25, 421–430. doi: 10.1590/S1415-47572002000400011

Suganuma, A., and Morioka, H. (1979). Morphological changes in membrane systems of staphylococci after different fixation procedures. *J. Elect. Micros.* 28, 29–35.

Troussellier, M., Courties, C., and Vaquer, A. (1993). Recent applications of flow-cytometry in aquatic microbial ecology. *Biol. Cell* 78, 111–121. doi: 10.1016/0248-4900(93)90121-T

Yamaha, E., Saito, T., Kazeto, R. G., and Arai, K. (2007). Developmental biotechnology for aquaculture, with special reference to surrogate production in teleost fishes. *J. Sea Res.* 58, 8–22. doi: 10.1016/j.seares.2007.02.003

Zhou, X. Y., Li, M. Y., Abbas, K., Gao, Z. X., and Wang, W. M. (2008). Comparison of ploidy level screening methods in chinese dojo loach (*Misgurnus anguillicaudatus*). *J. Appl. Ichthyol.* 24, 664–669. doi: 10.1111/j.1439-0426.2008.01116.x

Zhu, D., Song, W., Yang, K., Cao, X., Gul, Y., and Wang, W. (2012). Flow cytometric determination of genome size for eight commercially important fish species in China. *In Vitro Cell. Dev. Biol.* 48, 507–517. doi: 10.1007/s11626-012-9543-7

Conflict of Interest Statement: SdS was employed by company AES Tietê. The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declare that this study received funding from AES Tietê (Research & Development Project #0064-1052/2014). The funder was not involved in the study design or collection, analysis, or interpretation of the data. The reviewer DH declared a shared affiliation, with no collaboration, with one of the authors MP to the handling Editor.

Copyright © 2017 Xavier, Senhorini, Pereira-Santos, Fujimoto, Shimoda, Silva, dos Santos and Yasui. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 12 October 2017
doi: 10.3389/fgene.2017.00149

Revealing Hidden Diversity of the Underestimated Neotropical Ichthyofauna: DNA Barcoding in the Recently Described Genus *Megaleporinus* (Characiformes: Anostomidae)

Jorge L. Ramirez[1], Jose L. Birindelli[2], Daniel C. Carvalho[3], Paulo R. A. M. Affonso[4], Paulo C. Venere[5], Hernán Ortega[6], Mauricio Carrillo-Avila[7], José A. Rodríguez-Pulido[8] and Pedro M. Galetti Jr.[1]*

[1] *Laboratório de Biodiversidade Molecular e Conservação, Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Paulo, Brazil,* [2] *Departamento de Biologia Animal e Vegetal, Universidade Estadual de Londrina, Londrina, Brazil,* [3] *Laboratório de Genética da Conservação, Programa de Pós-Graduação em Biologia de Vertebrados, PUC Minas, Belo Horizonte, Brazil,* [4] *Departamento de Ciências Biológicas, Universidade Estadual do Sudoeste da Bahia, Jequié, Brazil,* [5] *Departamento de Biologia e Zoologia, Universidade Federal de Mato Grosso, Cuiabá, Brazil,* [6] *Departamento de Ictiología, Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Lima, Peru,* [7] *Facultad de Ciencias Exactas y Naturales, Universidad Surcolombiana, Huila, Colombia,* [8] *Grupo de Investigación en Genética y Reproducción Animal, Universidad de los Llanos, Villavicencio, Colombia*

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Federal University of Pernambuco, Brazil

Reviewed by:
Fábio Fernandes Roxo, Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), Brazil
Henrik R. Nilsson, University of Gothenburg, Sweden
Uedson Pereira Jacobina, Federal University of Alagoas, Brazil

***Correspondence:**
Jorge L. Ramirez
jolobio@ufscar.br

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

Received: *09 August 2017*
Accepted: *27 September 2017*
Published: *12 October 2017*

Citation:
Ramirez JL, Birindelli JL, Carvalho DC, Affonso PRAM, Venere PC, Ortega H, Carrillo-Avila M, Rodríguez-Pulido JA and Galetti PM Jr. (2017) Revealing Hidden Diversity of the Underestimated Neotropical Ichthyofauna: DNA Barcoding in the Recently Described Genus Megaleporinus (Characiformes: Anostomidae). Front. Genet. 8:149. doi: 10.3389/fgene.2017.00149

Molecular studies have improved our knowledge on the neotropical ichthyofauna. DNA barcoding has successfully been used in fish species identification and in detecting cryptic diversity. *Megaleporinus* (Anostomidae) is a recently described freshwater fish genus within which taxonomic uncertainties remain. Here we assessed all nominal species of this genus using a DNA barcode approach (Cytochrome Oxidase subunit I) with a broad sampling to generate a reference library, characterize new molecular lineages, and test the hypothesis that some of the nominal species represent species complexes. The analyses identified 16 (ABGD and BIN) to 18 (ABGD, GMYC, and PTP) different molecular operational taxonomic units (MOTUs) within the 10 studied nominal species, indicating cryptic biodiversity and potential candidate species. Only *Megaleporinus brinco*, *Megaleporinus garmani*, and *Megaleporinus elongatus* showed correspondence between nominal species and MOTUs. Within six nominal species, a subdivision in two MOTUs was found, while *Megaleporinus obtusidens* was divided in three MOTUs, suggesting that DNA barcode is a very useful approach to identify the molecular lineages of *Megaleporinus*, even in the case of recent divergence (< 0.5 Ma). Our results thus provided molecular findings that can be used along with morphological traits to better define each species, including candidate new species. This is the most complete analysis of DNA barcode in this recently described genus, and considering its economic value, a precise species identification is quite desirable and fundamental for conservation of the whole biodiversity of this fish.

Keywords: cryptic species, freshwater fishes, allopatric speciation, South American basins, cytochrome oxidase subunit I

INTRODUCTION

Neotropical freshwater fishes have a remarkable diversity, exceeding 8000 species (Reis et al., 2016), however, much taxonomic uncertainty exists leading to underestimated diversity (Pereira et al., 2013; Reis et al., 2016). Molecular studies have been crucial to improve our knowledge on the ichthyofauna, and DNA barcoding has successfully been used in fish species identification and in detecting species of taxonomic concerns or cryptic diversity (Pereira et al., 2013; Gomes et al., 2015; Ramirez and Galetti, 2015; Machado et al., 2016). Within the neotropical freshwater fishes, the order Characiformes represents more than 30% of the known species, and Anostomidae is one of the most species-rich families, occurring in all major hydrographic basins, with *trans*- and *cis*-Andean distribution in South America (Reis et al., 2003).

Comprising approximately 150 described species, distributed in 15 genera (Garavello and Britski, 2003; Sidlauskas and Vari, 2008; Ramirez et al., 2017), the known diversity of the Anostomidae has increased in recent years. For instance, 14 species and 1 genus were described only in the last 5 years (Birindelli et al., 2013; Burns et al., 2014). DNA barcoding has revealed taxonomic uncertainties within the genus *Laemolyta* (Ramirez and Galetti, 2015), and molecular phylogeny has helped to provide an understanding of the evolutionary history of the Anostomidae (Ramirez and Galetti, 2015; Ramirez et al., 2016, 2017).

Recently, the genus *Megaleporinus* (Ramirez et al., 2017) was described to include 16 lineages, corresponding to 10 nominal species, previously recognized in *Leporinus* or *Hypomasticus* (Ramirez et al., 2017). *Megaleporinus* is supported by cytogenetic, molecular, and morphological data. It is characterized by having a unique ZZ/ZW sex chromosome system (Galetti et al., 1995), while most cytogenetically known *Leporinus* species have no sex chromosomes (Galetti et al., 1981, 1991). Its monophyly is also well supported by mitochondrial and nuclear markers, which identified it as the sister group to *Abramites* (Ramirez et al., 2017). Concerning its morphology, *Megaleporinus* is characterized by being relatively large (adults usually reaching more than 35 cm standard length, including the largest species of the family), three teeth on each premaxillary and dentary bones, and a color pattern of one to three dark mid-lateral blotches (Ramirez et al., 2017). Because of its large size, *Megaleporinus* has an economic importance in subsistence fisheries and aquaculture (Garavello and Britski, 2003).

Recent studies indicate that there is a hidden biodiversity within *Megaleporinus* that needs to be better understood (Avelino et al., 2015; Ramirez et al., 2017). A study based on mitochondrial and nuclear markers, but using few individuals for each species, showed that several nominal species allocated to this genus comprise two or more molecular lineages allopatrically distributed in different basins (Ramirez et al., 2017).

In this study, we used a DNA barcoding approach to generate a reference library for *Megaleporinus*, assessing all nominal species and lineages previously described. We included a broad sampling for most of the species. Our hypothesis is that DNA barcoding support the observation that some of the nominal species represent species complexes with most molecular operational taxonomic units (MOTUs) allopatrically distributed in different basins, as proposed by Ramirez et al. (2017). Identifying such hidden biodiversity within this genus, this paper will contribute to a more complete understanding of its diversity and to the conservation of this important fish group.

MATERIALS AND METHODS

Sampling

Animals were collected on public land, handled and killed under permission (ICMBIO/MMA N° 32215) provided by the Environment Ministry (MMA). This study did not involve endangered or protected species. Fish were collected by fishing rods and gillnets. No ethics committee approval is required for these organisms in Brazil. Fish were killed in the field using cold water and immediately transferred onto ice. Tissue samples were collected after fish death was confirmed through lack of operculum movement.

Specimens from several populations of all *Megaleporinus* species were used in this study, totaling 79 samples of the 10 nominal species, and comprising the 16 molecular lineages described by Ramirez et al., 2017 (**Figures 1**, **2** and **Table 1**). Voucher numbers are provided for the specimens (**Table 1**). Additionally, previous DNA barcode sequences of specimens from the São Francisco (Carvalho et al., 2011), Paraná (Pereira et al., 2013), Paranapanema (Frantine-Silva et al., 2015), and lower Paraná basins (Díaz et al., 2016) were included in our data set giving a total of 116 sequences (**Figures 1**, **2** and **Table 1**).

DNA Extraction, Amplification, and Sequencing

Total DNA was extracted from tissues (fins, muscle, or liver) by the standard phenol–chloroform method (Sambrook et al., 1989). A fragment of Cytochrome Oxidase subunit I (COI; 698 bp) was amplified via polymerase chain reaction (PCR) using primers AnosCOIF and AnosCOIR (Ramirez and Galetti, 2015). PCR products were sequenced for both strands using an ABI 3730xl (Applied Biosystems, Waltham, MA, United States) automatic sequencer. Contigs were assembled and edited using BioEdit (Hall, 1999). All sequences were evaluated manually, deleting regions of low quality. All sequences were verified to represent the COI gene and were checked for indels and stop codons. GenBank (Benson et al., 2017) accession numbers are given in **Table 1**. All information about specimen, sequences, and electropherograms were deposited in a data set of The Barcode of Life Database platform (BOLD) with code DS-MGLEP.

DNA Barcode Analysis

The general mixed Yule coalescent (GMYC) model (Pons et al., 2006) with a single threshold, implemented in the *splits* packages

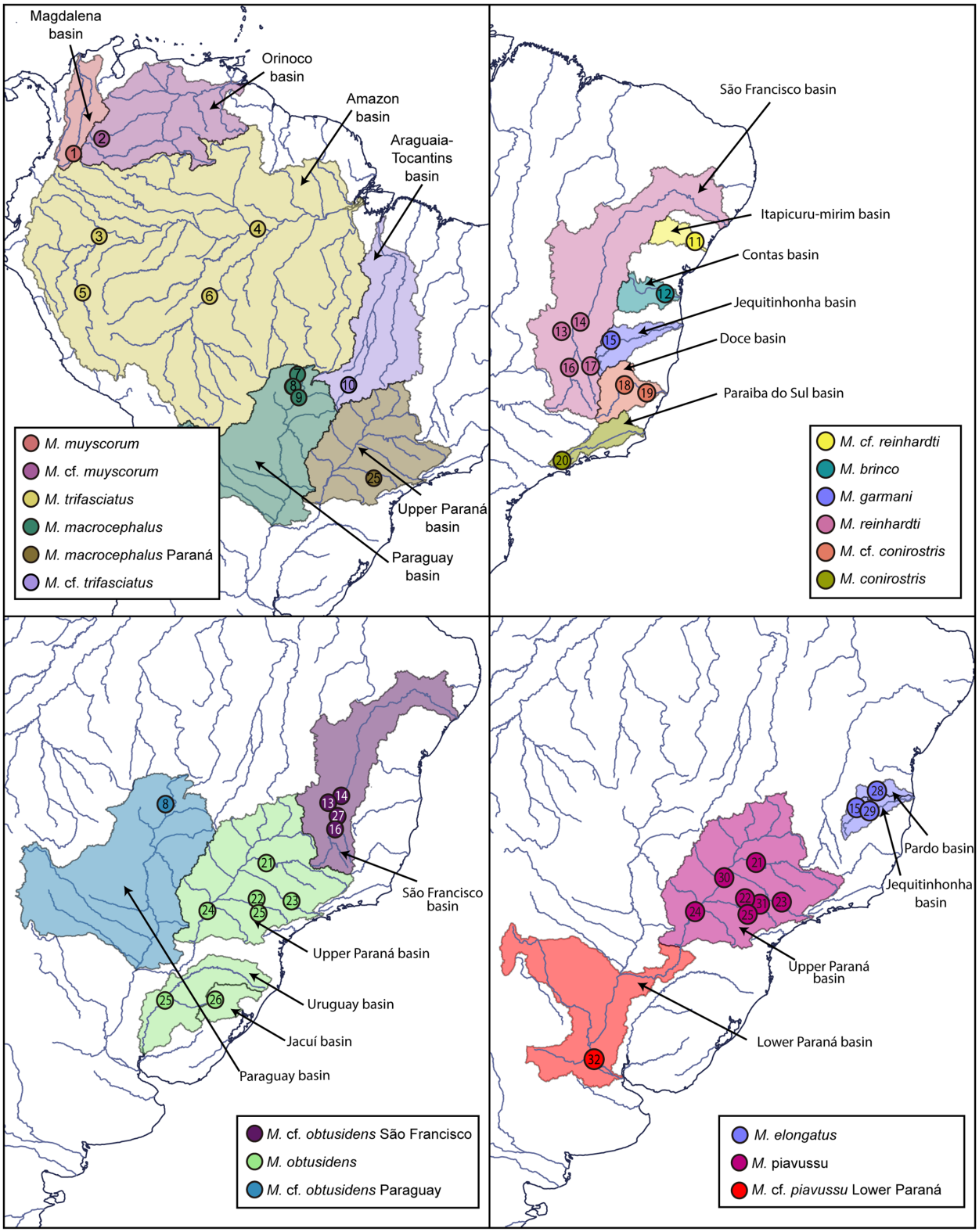


FIGURE 1 | Collection sites (circles) and hydrographic basin of occurrence of *Megaleporinus* MOTUs. Localities' numbers according to **Table 1**.

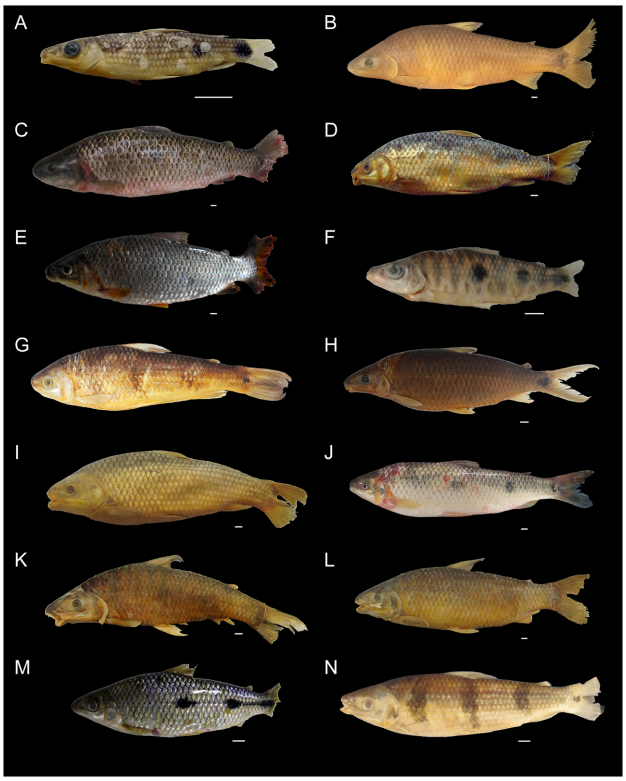


FIGURE 2 | Studied specimens of *Megaleporinus*. **(A)** *M. brinco*, MZUSP 118670; **(B)** *M.* cf. *conirostris*, LISDEBE 6971; **(C)** *M.* cf. *muyscorum*; **(D)** *M.* cf. *obtusidens* Paraguay, MZUSP 118668; **(E)** *M.* cf. *obtusidens* São Francisco, MCP 44805; **(F)** *M.* cf. *reinhardti,* UESB-8206; **(G)** *M.* cf. *trifasciatus*, GEPEMA 5095; **(H)** *M. garmani*, MCNI-PUCMG-0020; **(I)** *M. macrocephalus,* MZUSP 118667; **(J)** *M. muyscorum*, ICN-19074; **(K)** *M. obtusidens*, MZUSP 113982; **(L)** *M. piavussu*, MZUSP 113981; **(M)** *M. reinhardti*; **(N)** *M. trifasciatus*, MUSM – 47351. Scale bars equal 1 cm.

in the R 3.3.3 statistical software (R Core Team, 2017), was used to infer MOTUs. For the GMYC input, an ultrametric tree was generated using Beast 2.4.3 (Bouckaert et al., 2014), with a lognormal relaxed clock, a birth and death model, and a GTR+G substitution model, chosen using jModeltest 2 (Darriba et al., 2012), using 50 million MCMC generations and a burn-in of 10%. Poisson tree processes (PTP) model (Zhang et al., 2013) was used for MOTUs delimitation through the bPTP server[1], using default values. The bPTP server includes a Bayesian implementation of the PTP model and the original maximum likelihood PTP. For the PTP input, a tree was generated using Beast 2.4.6 (Bouckaert et al., 2014), with a strict clock, a birth and death model, and the GTR+G substitution model, using 50 million MCMC generations and a burn-in of 10%.

Additionally, two cluster algorithms were used, the Barcode Index Number System (BIN) (Ratnasingham and Hebert, 2013) and Automatic Barcode Gap Discovery (ABGD) (Puillandre et al., 2012). The BIN was automatically determined in the BOLD Workbench, while the ABGD was performed using Kimura-2-parameter (K2P) distance and default values through the web interface[2].

COI intraspecific and interspecific genetic distances were estimated using the K2P model implemented in Mega 6.0 (Tamura et al., 2013). These values were used to calculate the mean, minimum, and maximum values for intra- and inter-MOTU distances, and intra- and interspecific distances (nominal species). A genetic distance neighbor-joining (NJ) tree analysis was performed based on the K2P substitution model in Mega 6.0 (Tamura et al., 2013).

RESULTS

The alignment of COI sequences resulted in 600 characters with 158 parsimony informative sites (included in the Supplementary Material). The GMYC analysis resulted in 18 MOTUs (Confidence interval: 16–18) (**Table 2**). The GMYC model was preferred over the null model (likelihood ratio = 73.49, $P < 0.0001$), indicating that GMYC results are reliable. The PTP analyses (maximum likelihood and Bayesian implementation) resulted in the same 18 MOTUs obtained in GMYC. The ABGD analysis found six partitions with 27 ($P = 0.001$) to 16 groups ($P = 0.01$), including a partition with the same 18 MOTUs ($P = 0.005$) obtained in the GMYC and PTP analyses. The BOLD system determined 16 BINs (**Table 2**), showing discordance with our MOTUs in only two BINs, AAB8569 [*M. piavussu* (Britski et al., 2012) and *M.* cf. *piavussu* lower Paraná] and AAD1729 [*M. reinhardti* (Lütken, 1875) and *M.* cf. *reinhardti*]. The clustering of the MOTUs obtained by the analyses is shown in **Figure 3**.

Only *Megaleporinus brinco* (Birindelli and Britski, 2013), *Megaleporinus garmani* (Borodin, 1929), and *Megaleporinus elongatus* (Valenciennes, 1850) showed correspondence between nominal species and MOTUs. Within six nominal species, a subdivision in two MOTUs was found, while *Megaleporinus obtusidens* (Valenciennes, 1837) was divided in three MOTUs (**Table 2**).

The mean of intra-MOTU and maximum intra-MOTU distances, the nearest neighbor (NN), and the minimum distance to the NN are shown in **Table 2**, for both GMYC MOTUs and nominal species.

The overall mean of intra-MOTU distances was 0.03%, the maximum intra-MOTU distance was 0.5% (*M. obtusidens*), and the mean of inter-MOTU distances was 9.19%. The lowest and highest values of inter-MOTU distances were 0.67 and 15.31%, respectively. Considering these values, there is a barcode gap that allowed identifying successfully all MOTUs using COI distance. In contrast, when only the nominal species were considered, the maximum intraspecific distance increased to 15.31% [*M. muyscorum* (Steindachner, 1900)], and, in addition, no barcode gap was found.

[1]http://species.h-its.org/ptp/

[2]http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html

TABLE 1 | Sampling information and GenBank accession for specimens included in the analysis.

MOTU	River (Locality)	Basin	BIN	GenBank	Museum ID
Megaleporinus brinco	Contas[12]	Contas	ADB0463	KU134850	MZUSP – 118670
M. brinco	Contas[12]	Contas	ADB0463	KX925449	MZUSP – 118670
M. brinco	Contas[12]	Contas	ADB0463	KX925450	MZUSP – 118670
M. cf. *conirostris*	Doce (Governador Valadares)[18]	Doce	ACL4264	KF568977	MCNI-PUCMG-0186
M. cf. *conirostris*	Doce (Governador Valadares)[18]	Doce	ACL4264	KX925451	–
M. cf. *conirostris*	Doce (Governador Valadares)[18]	Doce	ACL4264	KX925452	–
M. cf. *conirostris*	Doce (Baixo Guandú)[19]	Doce	ACL4264	KX925453	LISDEBE 6971
M. cf. *muyscorum*	Meta (Puerto Lopez)[2]	Orinoco	ADB0512	KU134851	–
M. cf. *obtusidens* Paraguay	Cuiaba (Santo Antonio de Leverger)[8]	Paraguai	ACL3942	KU134861	MZUSP – 118668
M. cf. *obtusidens* São Francisco	São Francisco (Tres Marias)[20]	São Francisco	ABZ0928	HM405029	*
M. cf. *obtusidens* São Francisco	Pandeiros[14]	São Francisco	ABZ0928	HM405142	*
M. cf. *obtusidens* São Francisco	Urucuia (Urucuia)[13]	São Francisco	ABZ0928	HM906022	*
M. cf. *obtusidens* São Francisco	Urucuia (Urucuia)[13]	São Francisco	ABZ0928	HM906023	*
M. cf. *obtusidens* São Francisco	São Francisco (Tres Marias)[16]	São Francisco	ABZ0928	HM405028	*
M. cf. *obtusidens* São Francisco	São Francisco (Pirapora)[27]	São Francisco	ABZ0928	KX925498	LISDEBE 6973
M. cf. *obtusidens* São Francisco	São Francisco (Pirapora)[27]	São Francisco	ABZ0928	KX925499	–
M. cf. *obtusidens* São Francisco	São Francisco (Pirapora)[27]	São Francisco	ABZ0928	KX925500	–
M. cf. *obtusidens* São Francisco	São Francisco (Pirapora)[27]	São Francisco	ABZ0928	KX925501	–
M. cf. *obtusidens* São Francisco	São Francisco (Pirapora)[27]	São Francisco	ABZ0928	KX925502	–
M. cf. *obtusidens* São Francisco	Pandeiros (Pandeiros)[14]	São Francisco	ABZ0928	KU134862	MCP – 44805
M. cf. *obtusidens* São Francisco	Urucuia (Urucuia)[13]	São Francisco	ABZ0928	KX925503	–
M. cf. *obtusidens* São Francisco	Urucuia (Urucuia)[13]	São Francisco	ABZ0928	KX925504	MCP – 44076
M. cf. *reinhardti*	Itapicurú[11]	Itapicurú-mirim	AAD1729	KU134849	–
M. cf. *reinhardti*	Itapicurú[11]	Itapicurú-mirim	AAD1729	KX925454	–
M. cf. *reinhardti*	Itapicurú[11]	Itapicurú-mirim	AAD1729	KX925455	–
M. cf. *reinhardti*	Itapicurú[11]	Itapicurú-mirim	AAD1729	KX925456	UESB-8206
M. cf. *trifasciatus*	Araguaia (Ouro fino)[10]	Tocantins	ACL3074	KX925457	GEPEMA – 4975
M. cf. *trifasciatus*	Araguaia (Barra do Garças)[10]	Tocantins	ACL3074	KF568998	GEPEMA – 5095
M. cf. *trifasciatus*	Araguaia (Barra do Garças)[10]	Tocantins	ACL3074	KX925458	GEPEMA – 5594
M. conirostris	Paraibuna[20]	Paraiba do Sul	ACL3731	KU134852	–
M. conirostris	Paraibuna[20]	Paraiba do Sul	ACL3731	KX925459	–
M. elongatus	Itacambiruçu (Grão Mogol)[15]	Jequitinhonha	ABY2894	KX925463	–
M. elongatus	Itacambiruçu (Grão Mogol)[15]	Jequitinhonha	ABY2894	KU134853	MCNI-PUCMG-0375
M. elongatus		Jequitinhonha	ABY2894	KX925464	MCNI-PUCMG-0221
M. elongatus	Jequitinhonha (UHE Irapé)[29]	Jequitinhonha	ABY2894	KU134854	–
M. elongatus	Jequitinhonha (UHE Irapé)[29]	Jequitinhonha	ABY2894	KX925465	MCNI-PUCMG-0299
M. elongatus	Jequitinhonha (UHE Irapé)[29]	Jequitinhonha	ABY2894	KX925466	MCNI-PUCMG-0300
M. elongatus	Rio Pardo (Águas Vermelhas)[28]	Pardo	ABY2894	KX925460	MCNI-PUCMG-4451
M. elongatus	Rio Pardo (Águas vermelhas)[28]	Pardo	ABY2894	KX925461	MCNI-PUCMG-5175
M. elongatus	Rio Pardo (Águas vermelhas)[28]	Pardo	ABY2894	KX925462	MCNI-PUCMG-5176
M. garmani	Itacambiruçu (Grão Mogol)[15]	Jequitinhonha	ACL3227	KU134855	MCNI-PUCMG-0021
M. garmani	Itacambiruçu (Grão Mogol)[15]	Jequitinhonha	ACL3227	KX925467	MCNI-PUCMG-0020
M. garmani	Itacambiruçu (Grão Mogol)[15]	Jequitinhonha	ACL3227	KX925468	MCNI-PUCMG-0021
M. garmani	Itacambiruçu (Grão Mogol)[15]	Jequitinhonha	ACL3227	KX925469	MCNI-PUCMG-0021
M. garmani	Itacambiruçu (Grão Mogol)[15]	Jequitinhonha	ACL3227	KX925470	MCNI-PUCMG-0374
M. macrocephalus	Manhuaçu (São José do Mantimento)	Doce	AAE5328	KX925475	MCNI-PUCMG-0460
M. macrocephalus	Cuiaba (Santo Antonio de Leverger)[8]	Paraguai	AAE5328	KU134856	MZUSP – 118667
M. macrocephalus	Cuiaba (Santo Antonio de Leverger)[8]	Paraguai	AAE5328	KX925471	LISDEBE 6972
M. macrocephalus	Cuiaba (Cuiaba)[7]	Paraguai	AAE5328	KX925474	–
M. macrocephalus	Cuiaba (Barão de Melgaço)[9]	Paraguai	AAE5328	KX925476	LISDEBE 6974
M. macrocephalus	Cuiaba (Barão de Melgaço)[9]	Paraguai	AAE5328	KX925477	LISDEBE 6974
M. macrocephalus	Tiete (Barra Bonita)	Paraná	AAE5328	KX925473	–
M. macrocephalus	Pandeiros	São Francisco	AAE5328	HM906021	–

(Continued)

TABLE 1 | Continued

MOTU	River (Locality)	Basin	BIN	GenBank	Museum ID
M. macrocephalus	Araguaia (Ouro fino)	Tocantins	AAE5328	KX925472	GEPEMA – 4974
M. macrocephalus Paraná	Cinzas (Bandeirantes)[25]	Paraná	ACO1303	KM897611	*
M. macrocephalus Paraná	Cinzas (Bandeirantes)[25]	Paraná	ACO1303	KM897537	*
M. macrocephalus Paraná	Cinzas (Bandeirantes)[25]	Paraná	ACO1303	KM897575	*
M. macrocephalus Paraná	Cinzas (Bandeirantes)[25]	Paraná	ACO1303	KM897296	*
M. macrocephalus Paraná	Piracicaba (Tamanduá)[23]	Paraná	ACO1303	JN988999	LBPV-19469
M. muyscorum	Magdalena (Neiva)[1]	Magdalena	ADB0701	KX925478	ICN-19072
M. muyscorum	Magdalena (Neiva)[1]	Magdalena	ADB0701	KX925479	ICN-19073
M. muyscorum	Magdalena (Neiva)[1]	Magdalena	ADB0701	KU134857	ICN-19074
M. obtusidens	Jacuí (Jacuizinho Foz)[26]	Jacuí	AAB8578	KU134859	MCP-25476
M. obtusidens	Piracicaba (Tamanduá)[23]	Paraná	AAB8578	JN988985	LBPV-19849
M. obtusidens	Piracicaba (Tamanduá)[23]	Paraná	AAB8578	JN988984	LBPV-19850
M. obtusidens	Piracicaba (Tamanduá)[23]	Paraná	AAB8578	JN988983	LBPV-19852
M. obtusidens	Paranapanema (Canoas)[22]	Paraná	AAB8578	KM897227	*
M. obtusidens	Paranapanema (Canoas)[22]	Paraná	AAB8578	KM897138	*
M. obtusidens	Paranapanema (Canoas)[22]	Paraná	AAB8578	KM897434	*
M. obtusidens	Cinzas (Bandeirantes)[25]	Paraná	AAB8578	KM897597	*
M. obtusidens	Turvo (Icem)[21]	Paraná	AAB8578	KX925480	LISDEBE 6969
M. obtusidens	Turvo (Icem)[21]	Paraná	AAB8578	KX925481	–
M. obtusidens	Turvo (Icem)[21]	Paraná	AAB8578	KX925482	–
M. obtusidens	Turvo (Icem)[21]	Paraná	AAB8578	KX925483	–
M. obtusidens	Turvo (Icem)[21]	Paraná	AAB8578	KU134858	MZUSP – 113982
M. obtusidens	Turvo (Icem)[21]	Paraná	AAB8578	KF568987	–
M. obtusidens	Paraná (Porto Camargo)[24]	Paraná	AAB8578	KX925484	–
M. obtusidens	Paraná (Porto Camargo)[24]	Paraná	AAB8578	KX925485	–
M. obtusidens	Ibicui (BR 472)[25]	Uruguay	AAB8578	KU134860	MCP-28917
M. piavussu	Piracicaba (Tamanduá)[23]	Paraná	AAB8569	JN989005	LBPV-15587*
M. piavussu	Piracicaba (Tamanduá)[23]	Paraná	AAB8569	JN989004	LBPV-19851*
M. piavussu	Piracicaba (Tamanduá)[23]	Paraná	AAB8569	JN989003	LBPV-19854*
M. piavussu	Paranapanema (Canoas)[22]	Paraná	AAB8569	KM897529	*
M. piavussu	Paranapanema (Canoas)[22]	Paraná	AAB8569	KM897489	*
M. piavussu	Paranapanema (Canoas)[22]	Paraná	AAB8569	KM897621	*
M. piavussu	Cinzas (Bandeirantes)[25]	Paraná	AAB8569	KM897419	*
M. piavussu	Paranapanema (Canoas II)[31]	Paraná	AAB8569	KM897506	*
M. piavussu	Paranapanema (Canoas II)[31]	Paraná	AAB8569	KM897347	*
M. piavussu	Paranapanema (Canoas II)[31]	Paraná	AAB8569	KM897192	*
M. piavussu	Turvo (Icem)[21]	Paraná	AAB8569	KF568991	MZUSP – 113981
M. piavussu	Turvo (Icem)[21]	Paraná	AAB8569	KX925486	LISDEBE 6968
M. piavussu	Turvo (Icem)[21]	Paraná	AAB8569	KX925487	LISDEBE 6970
M. piavussu	Turvo (Icem)[21]	Paraná	AAB8569	KX925488	LISDEBE 6970
M. piavussu	Turvo (Icem)[21]	Paraná	AAB8569	KX925489	–
M. piavussu	Turvo (Icem)[21]	Paraná	AAB8569	KX925490	–
M. piavussu	Turvo (Icem)[21]	Paraná	AAB8569	KX925491	–
M. piavussu	Paraná (Porto Camargo)[24]	Paraná	AAB8569	KX925492	–
M. piavussu	Paraná (Pauliceia)[30]	Paraná	AAB8569	KX925493	–
M. cf. *piavussu* lower Paraná	Paraná (Rosario)[32]	Paraná	AAB8569	KU288864	*
M. cf. *piavussu* lower Paraná	Paraná (Rosario)[32]	Paraná	AAB8569	KU288865	*
M. cf. *piavussu* lower Paraná	Paraná (Rosario)[32]	Paraná	AAB8569	KU288866	*
M. cf. *piavussu* lower Paraná	Paraná (Rosario)[32]	Paraná	AAB8569	KU289030	*
M. reinhardti	Pandeiros[14]	São Francisco	AAD1729	HM906025	*
M. reinhardti	Curimataí[17]	São Francisco	AAD1729	HM405147	MCP – 44776*
M. reinhardti	Curimataí[17]	São Francisco	AAD1729	HM906026	*
M. reinhardti	Urucuia (Urucuia)[13]	São Francisco	AAD1729	HM906027	*

(Continued)

TABLE 1 | Continued

MOTU	River (Locality)	Basin	BIN	GenBank	Museum ID
M. reinhardti	Urucuia (Urucuia)[13]	São Francisco	AAD1729	HM906028	*
M. reinhardti	São Francisco (Três Marias)[16]	São Francisco	AAD1729	KX925494	–
M. reinhardti	Urucuia (Urucuia)[13]	São Francisco	AAD1729	KX925495	–
M. reinhardti	Urucuia (Urucuia)[13]	São Francisco	AAD1729	KX925496	–
M. reinhardti	Curimataí[17]	São Francisco	AAD1729	KX925497	MCP - 44770
M. trifasciatus	Madeira[6]	Amazonas	ACL3073	KU134864	UFRO-I 4902
M. trifasciatus	Ucayali (Pucallpa)[5]	Amazonas	ACL3073	KU134865	MUSM - 47351
M. trifasciatus	Ucayali (Pucallpa)[5]	Amazonas	ACL3073	KX925505	MUSM - 47351
M. trifasciatus	Amazonas (Belen)[3]	Amazonas	ACL3073	KX925506	MUSM - 47364
M. trifasciatus	Lago Catalão (Manaus)[4]	Amazonas	ACL3073	KX925507	INPA 11641

**Obtained from BOLD.*

DISCUSSION

Our hypothesis that some of the nominal species represent species complexes separated in different basins could not be rejected by DNA barcoding analysis, revealing taxonomic uncertainties and a hidden diversity within this recently described genus. The DNA barcode analyses identified 16 (ABGD and BIN) to 18 (ABGD, GMYC, and PTP) different MOTUs (**Figure 3**), with two new MOTUs (*M. macrocephalus* Paraná and *M.* cf. *piavussu* lower Paraná) not analyzed by

TABLE 2 | Genetic K2P distances of *Megaleporinus* species.

	Mean intra-	Maximum intra-	NN	Distance to NN
MOTUs				
Megaleporinus brinco	0	0	*M. obtusidens*	6.78
M. conirostris	0	0	*M.* cf. *conirostris*	3.99
M. cf. *conirostris*	0	0	*M. conirostris*	3.99
M. elongatus	0.04	0.17	*M.* cf. *obtusidens* São Francisco	2.74
M. garmani	0	0	*M. obtusidens*	7.68
M. macrocephalus	0	0	*M. macrocephalus* Paraná	1.86
M. macrocephalus Paraná	0	0	*M. macrocephalus*	1.86
M. muyscorum	0	0	*M. reinhardti*	11.6
M. cf. *muyscorum*	–	–	*M. trifasciatus*	7.48
M. obtusidens	0.14	0.5	*M.* cf. *obtusidens* São Francisco	2.84
M. cf. *obtusidens* Paraguay	–	–	*M. piavussu* Lower Paraná	2.9
M. cf. *obtusidens* São Francisco	0	0	*M. elongatus*	2.74
M. piavussu	0.06	0.17	*M. piavussu* Lower Paraná	0.67
M. piavussu Lower Paraná	0.08	0.17	*M. piavussu*	0.67
M. reinhardti	0	0	*M.* cf. *reinhardti*	0.67
M. cf. *reinhardti*	0	0	*M. reinhardti*	0.67
M. trifasciatus	0	0	*M. macrocephalus*	4.52
M. cf. *trifasciatus*	0	0	*M. trifasciatus*	6.33
Nominal				
Megaleporinus brinco	0	0	*M. obtusidens*	6.78
M. conirostris	2.13	3.99	*M. obtusidens*	5.6
M. elongatus	0.04	0.17	*M. obtusidens*	2.74
M. garmani	0	0	*M. obtusidens*	7.68
M. macrocephalus	0.86	1.86	*M. trifasciatus*	4.52
M. muyscorum	7.66	15.31	*M. trifasciatus*	7.48
M. obtusidens	1.94	6.72	*M. elongatus*	2.74
M. piavussu	0.26	1.01	*M. obtusidens*	2.9
M. reinhardti	0.31	0.7	*M. conirostris*	6.14
M. trifasciatus	3.39	6.33	*M. macrocephalus*	4.52

The mean and the maximum of intra-group distances, the nearest neighbor (NN), and the minimum distance to the NN for MOTUs (ABGD, GMYC, and PTP) and Nominal species.

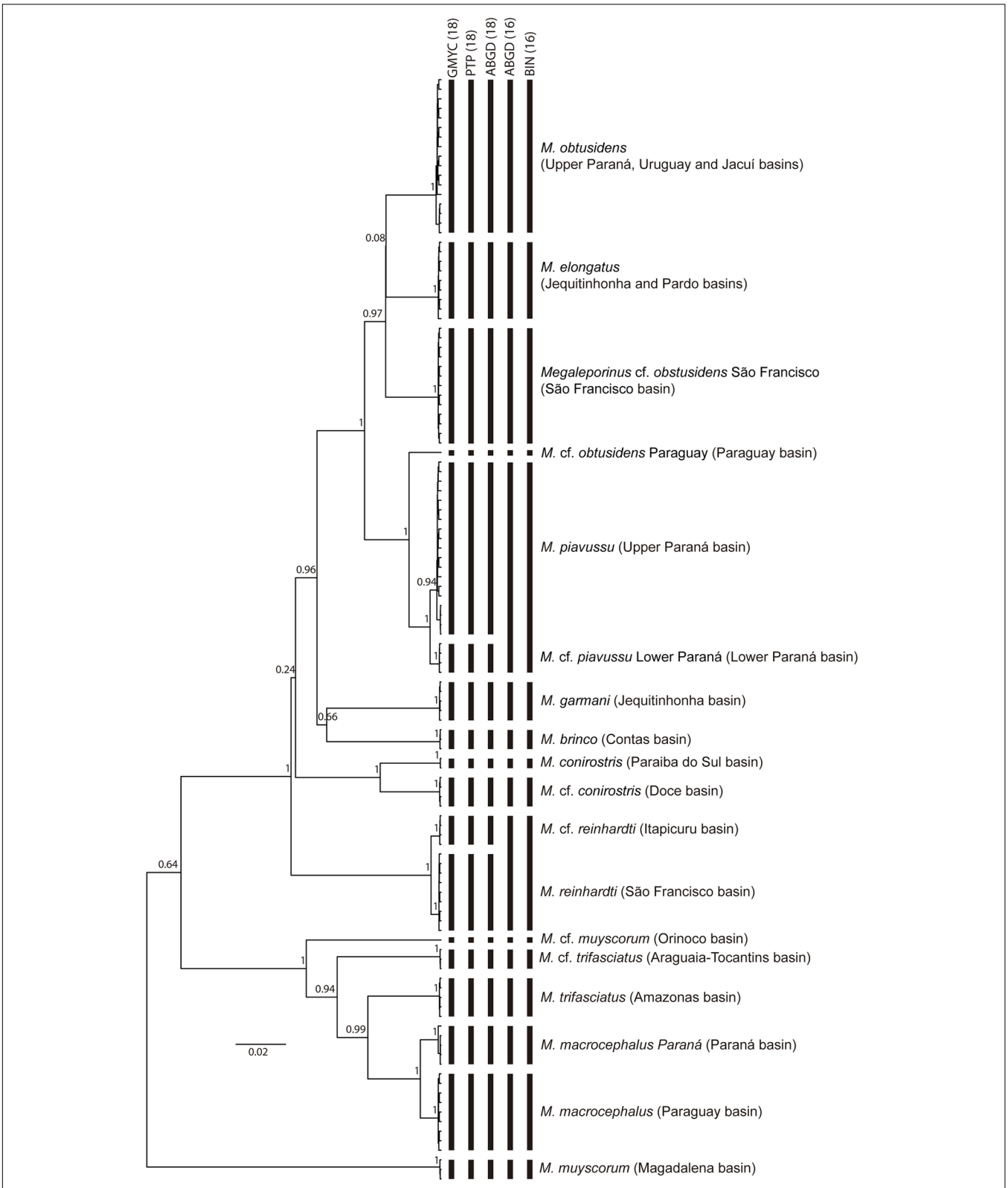


FIGURE 3 | Bayesian tree showing the clustering of the MOTUs obtained by the species delimitation analyses.

Ramirez et al. (2017). This high number of MOTUs contrasts with the 10 nominal species recognized in the genus thus far, showing several potential target for cryptic species to be described, reinforcing the general idea that there is still a lot of undocumented diversity within the neotropical ichthyofauna (Reis et al., 2016). The difference between the number of MOTUs detected is due to the lower genetic distance value (0.67%) between two pairs of MOTUs: *M. reinhardti* and *M.* cf. *reinhardti*, separating the genetic lineages from São Francisco and Itapicuru, respectively, and between *M. piavussu* and *M.* cf. *piavussu* lower Paraná. These lower genetic distance values are likely due to a recent divergence between these MOTUs [<0.5 Ma for *M. reinhardti* and *M.* cf. *reinhardti* according to Ramirez et al. (2017)]. Of note, besides presenting an allopatric distribution, these MOTUs were also recovered by the monophyly criterion (**Figure 3**). MOTUs with recent origin have less time to accumulate genetic differences than species with ancient origin, hindering their identification. Despite this low genetic distance, the species delimitation methods could delimit these MOTUS, especially those based on phylogenetic trees (GMYC and PTP).

A key aspect implicit in the DNA barcoding analysis is the genetic distance threshold used to define MOTUs. COI distances of 1% (Hubert et al., 2008) to 2% (Pereira et al., 2013) have been claimed as threshold to fish DNA barcode analysis. However, such values were derived from comparative analyses among phylogenetically diverse groups. For instance, 2% was used to characterize DNA barcoding of a fish community of a given river (Pereira et al., 2013). However, when the DNA barcoding analyses have focused within a group of species closely related (e.g., a genus), lower threshold values have been reported (Carvalho et al., 2011; Pereira et al., 2011, 2013; Ramirez and Galetti, 2015). Particularly in Anostomidae, a lower threshold of 0.92% was reported to distinguish MOTUs within the genus *Laemolyta* (Ramirez and Galetti, 2015). Although most of the values obtained herein were above 2% (13 out of 18 MOTUs, **Table 2**), a maximum threshold of 0.67% for *Megaleporinus* was detected between the MOTUs obtained. It reinforces that lower genetic distance values might be obtained when intra-genus MOTUs are analyzed, mainly between recent divergent lineages.

Five nominal species, *M. conirostris* (Steindachner, 1875), *M. macrocephalus* (Garavello and Britski, 1988), *M. muyscorum, M. obtusidens,* and *M. trifasciatus* (Steindachner, 1876), showed high COI distance values (> 1.8%, **Table 2**) between individuals from different basins, indicating a scenario of potential allopatric speciation within these species.

In contrast to previous results (Avelino et al., 2015), evidence of local differentiation was not found here and all cryptic diversity correspond to inter-basin differentiation. Analyzing only two samples of *M. reinhardti* from the Três Marias (MG, Brazil) region (São Francisco basin), Avelino et al. (2015) reported an intraspecific distance of 3.8% between them, suggesting a local differentiation. Here we analyzed nine individuals, representing four different localities, including Três Marias region, and we found no genetic distance (0%) among them. Mitochondrial pseudogenes, sequencing errors, or misidentification could explain such discrepancies, and it would be more cautious to consider *M. reinhardti* from São Francisco as a single MOTU, as recovered here.

Similar discordance is observed for *M. piavussu* (upper Paraná). Avelino et al. (2015) included four samples from a single locality and reported a mean intraspecific distance of 2.8%. Our present data set for this species included 18 individuals obtained from six localities and showed a lower maximum intraspecific distance of 0.17%. It is strongly suggested that *M. piavussu* is also a single MOTU.

Incongruences were also observed within the nominal *M. obtusidens*. While four groups (A–D), showing 0.7–4.1% mean intraspecific distances, were previously reported (Avelino et al., 2015), we found three MOTUs showing 0–0.5% COI distances. The group D mentioned as part of *M. obtusidens* by Avelino et al. (2015), which included individuals caught downstream the Itaipú dam (Paraná basin), was recovered here as a sister group of *M. piavussu*, and was named *M.* cf. *piavussu* lower Paraná (**Figure 3**).

One particular aspect was highlighted in our results. Several individuals clustered in the *M. macrocephalus* clade were caught in different hydrographic basins, as Doce, São Francisco, Tocantins, and Paraná, outside of its original distribution in the Paraguay basin likely due to aquaculture releasing. Similar findings had already been described in the São Francisco basin (Carvalho et al., 2011). This species is a commercial important fish being extensively farmed throughout the Brazilian territory, and accidental or intentional releasing can occur (e.g., Langeani et al., 2007; Vieira, 2010). In such case, the use of DNA barcoding provides a rapid and accurate identification of this species and can be used in management and monitoring potential ecosystem disturbance caused by an invasive species.

In summary, the use of DNA barcoding points at the need for a taxonomic revision of this genus. A search for morphological traits able to support a taxonomic delimitation could be facilitated whether the MOTUs identified here are considered. A morphological trait showing a range of variation when searched within a given nominal species perhaps could be more informative if studied in each MOTU separately. In such case, our results would give an important contribution for the taxonomy of *Megaleporinus* facilitating the search for decisive taxonomic characters. This is the most complete analysis of DNA barcode in this recently described genus, and considering the economic value of this group, a precise species identification is quite desirable and fundamental for conservation of the whole biodiversity of this genus.

AUTHOR CONTRIBUTIONS

JR and PG designed the research. JR, DC, PA, PV, HO, MC-A, and JR-P collected data. JR performed the analyses. All authors contributed to the writing of the manuscript.

FUNDING

The authors thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support (Universal

473474/2011-5 and 405309/2016-3 to PG, Universal 420255/2016-8 to JB and Rede BrBOL 564953/2010-5). JR received a fellowship grant from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 2011/21836-4). Authors received productivity research grants from CNPq (304440/2009-4 to PG) and Fundação Araucária (641/2014 to JB).

ACKNOWLEDGMENTS

We are grateful to C. Cramer, C. Doria, C. Nolorbe, D. Motta, H. Sanchez, J.C. Riofrio, and W. Troy for help to obtain part of the tissue samples and MMA/ICMBIO for collection authorization (32215-1). The authors thank the three reviewers for suggestion and comments which improved the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2017.00149/full#supplementary-material

REFERENCES

Avelino, G. S., Britski, H. A., Foresti, F., and Oliveira, C. (2015). Molecular identification of *Leporinus* from the south portion of South America. *DNA Barcodes* 3, 98–109. doi: 10.1515/dna-2015-0013

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., et al. (2017). GenBank. *Nucleic Acids Res.* 45, D37–D42. doi: 10.1093/nar/gkw1070

Birindelli, J. L. O., and Britski, H. A. (2013). Two new species of *Leporinus* (Characiformes: Anostomidae) from the Brazilian Amazon, and redescription of *Leporinus striatus* Kner 1858. *J. Fish Biol.* 83, 1128–1160. doi: 10.1111/jfb.12206

Birindelli, J. L. O., Britski, H. A., and Garavello, J. C. (2013). Two new species of *Leporinus* Agassiz (Characiformes?: Anostomidae) from eastern basins of Brazil, and redescription of *L. melanopleura* Günther. *Neotrop. Ichthyol.* 11, 9–23. doi: 10.1590/S1679-62252013000100002

Borodin, N. A. (1929). Notes on some species and subespecies of the genus *Leporinus* Spix. *Mem. Museum Comp. Zool.* 50, 269–290.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., et al. (2014). BEAST 2: a software platform for bayesian evolutionary analysis. *PLOS Comput. Biol.* 10:e1003537. doi: 10.1371/journal.pcbi.1003537

Britski, H. A., Birindelli, J. L. O., and Garavello, J. C. (2012). A new species of *Leporinus* Agassiz, 1829 from the Upper Rio Paraná basin (Characiformes, Anostomidae) with redescription of *L. elongatus* Valenciennes, 1850 and L. obtusidens (Valenciennes, 1837). *Pap. Avulsos Zool.* 52, 441–475.

Burns, M. D., Frable, B. W., and Sidlauskas, B. L. (2014). A new species of *Leporinus* (Characiformes: Anostomidae), from the Orinoco Basin, Venezuela. *Copeia* 2014, 206–214. doi: 10.1643/CI-13-071

Carvalho, D. C., Oliveira, D. A. A., Pompeu, P. S., Leal, C. G., Oliveira, C., and Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River basin. *Mitochondrial DNA* 22, 80–86. doi: 10.3109/19401736.2011.588214

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772. doi: 10.1038/nmeth.2109

Díaz, J., Villanova, G. V., Brancolini, F., del Pazo, F., Posner, V. M., Grimberg, A., et al. (2016). First DNA barcode reference library for the identification of South American freshwater fish from the Lower Paraná river. *PLOS ONE* 11:e0157419. doi: 10.1371/journal.pone.0157419

Frantine-Silva, W., Sofia, S. H., Orsi, M. L., and Almeida, F. S. (2015). DNA barcoding of freshwater ichthyoplankton in the Neotropics as a tool for ecological monitoring. *Mol. Ecol. Resour.* 15, 1226–1237. doi: 10.1111/1755-0998.12385

Galetti, P. M. Jr., Cesar, A. C. G., and Venere, P. C. (1991). Heterochromatin and NORs variability in *Leporinus* fish (Anostomidae, Characiformes). *Caryologia* 44, 287–292. doi: 10.1080/00087114.1991.10797193

Galetti, P. M. Jr., Lima, N. R. W., and Venere, P. C. (1995). A monophyletic ZW sex chromosome system in *Leporinus* (Anostomidae, Characiformes). *Cytologia (Tokyo).* 60, 375–382. doi: 10.1508/cytologia.60.375

Galetti, PM Jr, Foresti, F., Bertollo, L. A., and Moreira, Filho O (1981). Heteromorphic sex chromosomes in three species of the genus *Leporinus* (Pisces, Anostomidae). *Cytogenet. Genome Res.* 29, 138–142. doi: 10.1159/000131562

Garavello, J. C., and Britski, H. A. (1988). *Leporinus* macrocephalus sp. n. da bacia do rio Paraguai (Ostariophysi, Anostomidae). *Naturalia* 13, 67–74.

Garavello, J. C., and Britski, H. A. (2003). "Family Anostomidae," in *Check List of the Freshwater Fishes of South and Central America*, eds R. E. Reis, S. O. Kullander, and C. J. Ferraris Jr. (Porto Alegre: EDIPUCRS), 71–84.

Gomes, L. C., Pessali, T. C., Sales, N. G., Pompeu, P. S., and Carvalho, D. C. (2015). Integrative taxonomy detects cryptic and overlooked fish species in a neotropical river basin. *Genetica* 143, 581–588. doi: 10.1007/s10709-015-9856-z

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., Taylor, E., Burridge, M., et al. (2008). Identifying Canadian freshwater fishes through DNA barcodes. *PLOS ONE* 3:e2490. doi: 10.1371/journal.pone.0002490

Langeani, F., Corrêa e Castro, R. M., Oyakawa, O. T., Shibatta, O. A., Pavanelli, C. S., and Casatti, L. (2007). Diversidade da ictiofauna do Alto Rio Paraná: composição atual e perspectivas futuras. *Biota Neotrop.* 7, 181–197. doi: 10.1590/S1676-06032007000300020

Lütken, C. R. (1875). Velhas Flodens Fiske. Et bidrag til Brasiliens Ichthyologi. Elfter Prof. J. Reinhardt Indsamlinger og Optegnelser. *Danske Vidensk. Selsk. Skr. Kjøbenhavn* 12, 122–254.

Machado, C. D. B., Ishizuka, T. K., Freitas, P. D., De Valiati, V. H., and Galetti, P. M. (2016). DNA barcoding reveals taxonomic uncertainty in *Salminus* (Characiformes). *Syst. Biodivers.* 15, 372–382. doi: 10.1080/14772000.2016.1254390

Pereira, L. H. G., Hanner, R., Foresti, F., and Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genet.* 14:20. doi: 10.1186/1471-2156-14-20

Pereira, L. H. G., Maia, G. M. G., Hanner, R., Foresti, F., and Oliveira, C. (2011). DNA barcodes discriminate freshwater fishes from the Paraíba do Sul River Basin, São Paulo, Brazil. *Mitochondrial DNA* 22(Suppl. 1), 71–79. doi: 10.3109/19401736.2010.532213

Pons, J., Barraclough, T., Gomez-Zurita, J., Cardoso, A., Duran, D., Hazell, S., et al. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55, 595–609. doi: 10.1080/10635150600852011

Puillandre, N., Lambert, A., Brouillet, S., and Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Mol. Ecol.* 21, 1864–1877. doi: 10.1111/j.1365-294X.2011.05239.x

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ramirez, J. L., Birindelli, J. L. O., and Galetti, P. M. (2017). A new genus of Anostomidae (Ostariophysi: Characiformes): diversity, phylogeny and biogeography based on cytogenetic, molecular and morphological data. *Mol. Phylogenet. Evol.* 107, 308–323. doi: 10.1016/j.ympev.2016.11.012

Ramirez, J. L., Carvalho-Costa, L. F., Venere, P. C., Carvalho, D. C., Troy, W. P., and Galetti, P. M. (2016). Testing monophyly of the freshwater fish *Leporinus*

(Characiformes, Anostomidae) through molecular analysis. *J. Fish Biol.* 88, 1204–1214. doi: 10.1111/jfb.12906

Ramirez, J. L., and Galetti, P. M. Jr. (2015). DNA barcode and evolutionary relationship within *Laemolyta* Cope 1872 (Characiformes: Anostomidae) through molecular analyses. *Mol. Phylogenet. Evol.* 93, 77–82. doi: 10.1016/j.ympev.2015.07.021

Ratnasingham, S., and Hebert, P. D. N. (2013). A DNA-based registry for all animal species: the Barcode Index Number (BIN) System. *PLOS ONE* 8:e66213. doi: 10.1371/journal.pone.0066213

Reis, R. E., Albert, J. S., Di Dario, F., Mincarone, M. M. M., Petry, P. L., and Rocha, L. R. (2016). Fish biodiversity and conservation in South America. *J. Fish Biol.* 89, 12–47. doi: 10.1111/jfb.13016

Reis, R. E., Kullander, S. O., and Ferraris, C. J. Jr. (2003). *Check List of the Freshwater Fishes of South and Central America*. Porto Alegre: EDIPUCRS.

Sambrook, J., Fritish, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sidlauskas, B. L., and Vari, R. P. (2008). Phylogenetic relationships within the South American fish family Anostomidae (Teleostei, Ostariophysi,Characiformes). *Zool. J. Linn. Soc.* 154, 70–210. doi: 10.1111/j.1096-3642.2008.00407.x

Steindachner, F. (1875). Die Süsswasserfische des südöstlichen Brasilien (II). *Sitzungsber. Akad. Wiss. Wien* 71, 211–245.

Steindachner, F. (1876). Ichthyologische Beiträge (V). *Sitzungsber. Akad. Wiss. Wien* 74, 49–240.

Steindachner, F. (1900). Erstattungen eines vorlaüfigen Berichtes über einige von Ihrer königlichen Hoheit Frau Prinzessin Therese von Bayeren während einer Reise nach Südamerika 1898 gesammelte neue Fischarten. *Anz. Akad. Wiss. Wien* 37, 3.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Valenciennes, A. (1837). "Poissons [plates]," in *Voyage Dans l'Amérique Méridionale*, ed. A. d'Orbigny (Paris: Pitois-Levrault), 1834-42.

Valenciennes, M. A. (1850). "Suite du livre vingt-deuxième. Suite de la famille des Salmonoïdes," in *Histoire Naturelle des Poissons. Tome Vingt-Deuxième*, eds M. Cuvier and M. A. Valenciennes (Malden, MA: Blackwell Science), 1–91.

Vieira, F. (2010). Distribuição, impactos ambientais e conservação da fauna de peixes da bacia do rio Doce. *MG Biota* 2, 5–22.

Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29, 2869–2876. doi: 10.1093/bioinformatics/btt499

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

Copyright © 2017 Ramirez, Birindelli, Carvalho, Affonso, Venere, Ortega, Carrillo-Avila, Rodríguez-Pulido and Galetti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 24 October 2017
doi: 10.3389/fgene.2017.00159

Low Genetic Diversity and Structuring of the *Arapaima* (Osteoglossiformes, Arapaimidae) Population of the Araguaia-Tocantins Basin

*Carla A. Vitorino[1], Fabrícia Nogueira[2], Issakar L. Souza[3], Juliana Araripe[2] and Paulo C. Venere[1]**

[1] Instituto de Biociências, Universidade Federal de Mato Grosso, Cuiabá, Brazil, [2] Instituto de Estudos Costeiros, Universidade Federal do Pará, Bragança, Brazil, [3] Departamento de Biologia Celular, Embriologia e Genética, Universidade Federal de Santa Catarina, Florianópolis, Brazil

The arapaima, *Arapaima gigas*, is a fish whose populations are threatened by both overfishing and the ongoing destruction of its natural habitats. In the Amazon basin, varying levels of population structure have been found in *A. gigas*, although no data are available on the genetic diversity or structure of the populations found in the Araguaia-Tocantins basin, which has a topographic profile, hydrological regime, and history of fishing quite distinct from those of the Amazon. In this context, microsatellite markers were used to assess the genetic diversity and connectivity of five wild *A. gigas* populations in the Araguaia-Tocantins basin. The results of the analysis indicated low levels of genetic diversity in comparison with other *A. gigas* populations, studied in the Amazon basin. The AMOVA revealed that the *Arapaima* populations of the Araguaia-Tocantins basin are structured significantly. No correlation was found between pairwise F_{ST} values and the geographical distance among populations. The low level of genetic variability and the evidence of restricted gene flow may both be accounted for by overfishing, as well as the other human impacts that these populations have been exposed to over the years. The genetic fragility of these populations demands attention, given that future environmental changes (natural or otherwise) may further reduce these indices and eventually endanger these populations. The results of this study emphasize the need to take the genetic differences among the study populations into account when planning management measures and conservation strategies for the arapaima stocks of the Araguaia-Tocantins basin.

Keywords: genetic structure, microsatellite, conservation, habitat fragmentation, genetics population

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

Reviewed by:
Łukasz Kajtoch, Institute of Systematics and Evolution of Animals (PAN), Poland
Daniel Cardoso Carvalho, Pontifícia Universidade Católica de Minas Gerais, Brazil

****Correspondence:***
Paulo C. Venere
pvenere@uol.com.br

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

Received: *22 May 2017*
Accepted: *10 October 2017*
Published: *24 October 2017*

Citation:
Vitorino CA, Nogueira F, Souza IL, Araripe J and Venere PC (2017) Low Genetic Diversity and Structuring of the Arapaima (Osteoglossiformes, Arapaimidae) Population of the Araguaia-Tocantins Basin. Front. Genet. 8:159. doi: 10.3389/fgene.2017.00159

INTRODUCTION

Worldwide, the populations of many fish species are declining rapidly (Allan et al., 2005), with the communities occupying lakes, rivers and floodplains being the most affected (Abell et al., 2008). In the Neotropical region, the biodiversity of aquatic ecosystems is under intense pressure, primarily from human activities, such as overfishing, pollution, habitat fragmentation, deforestation and the introduction of exotic species (Agostinho et al., 2005).

The arapaima, *Arapaima gigas* (Schinz, 1822), is a fish of considerable economic importance in the neotropics, and is one of the species listed in the Convention of International Trade in Endangered Species of Wild Fauna and Flora II. The species is listed as "Data Deficient" by the IUCN (World Conservation Monitoring Centre, 2014). As the arapaima prefers the lentic environments, such as flooded forests, rivers and lakes, of the Amazon, Araguaia-Tocantins and Essequibo basins (Castello and Stewart, 2010; Castello et al., 2013), its populations become increasingly concentrated during the dry season, and the high densities of fish that accumulate during this period greatly increase their vulnerability to capture. Given this vulnerability of the species to fishing and the reduction of its stocks in recent years, researchers have focused increasingly on its genetic diversity and population structure (Farias et al., 2003; Hrbek et al., 2005, 2007; Hrbek and Farias, 2008; Hamoy et al., 2008; Araripe et al., 2013; Vitorino et al., 2015), chromosomal evolution (Marques et al., 2006), and other aspects of its biology (Gomes, 2007; Castello, 2008a,b; Arantes et al., 2010; Fernandes et al., 2012; Farias et al., 2015).

Despite this recent interest, many features of the biology of arapaima are still relatively poorly known, in particular the structure of its natural populations. Up until now, in addition, most studies have focused on the populations of the Amazon basin, and little is known about the genetic diversity of the populations that inhabit the Araguaia-Tocantins basin, even though this system is considered to be a priority area for the conservation of the aquatic biodiversity of the Cerrado biome.

As the characteristics of the Araguaia-Tocantins basin are quite distinct from those of the Amazon basin, the data available for Amazonian *Arapaima gigas* cannot be extrapolated reliably to the Araguaia-Tocantins. In this way, Latrubesse and Stevaux (2002) identified three principal topographical sectors in the Araguaia basin (the principal focus of the present study): (i) the upper Araguaia River, which extends from the headwaters to Registro do Araguaia, located mainly on Pre-Cambrian rocks, with a V-shaped valley and many rapids, (ii) the middle Araguaia River, between Registro do Araguaia and Conceição do Araguaia, which is characterized by a well-developed alluvial plain located on the lowland Bananal Plain, a prominent geomorphological and sedimentary unit, with some isolated rock formations along the main channel, which form small rapids, and (iii) the lower Araguaia River, which runs downstream from the Bananal Plain, over an area of crystalline rocks until its confluence with Tocantins River, which has no well-defined alluvial plain (Latrubesse and Stevaux, 2002).

Because of these characteristics, the hydrological regime of the Araguaia River is also quite unique, and is determined by well-defined dry and rainy seasons (Aquino et al., 2008). The flood pulse is extremely rapid, with floodplain lakes being connected for only short periods of time. The Araguaia-Tocantins basin also includes portions of the two principal Brazilian biomes, the Amazon, to the north and the Cerrado, to the south (Aquino et al., 2005). However, these important centers of biodiversity have been impacted intensively by industrial-scale farming operations and extensive cattle ranching, the construction of reservoirs for the production of hydroelectric energy (Aquino et al., 2008, 2009; Latrubesse et al., 2009), and the establishment of the Tocantins-Araguaia waterway (Almeida and Peres, 2007). All these impacts have resulted in extensive alterations of the region's natural environments, in particular the fragmentation of habitats, which reduces gene flow among populations through processes such as siltation and the restriction of aquatic environments (Agostinho et al., 2005).

While arapaima stocks have been reduced by both fishing and habitat loss, few populations have been studied in the Araguaia-Tocantins basin, and there is a pressing need for the expansion of the database to provide a reliable assessment of the conservation status of these populations. The first study to use genetic markers in arapaima from the Araguaia-Tocantins basin was conducted by Vitorino et al. (2015), who found low levels of expected heterozygosity and a distinct structure among the four study populations. The present study provides further advances in the understanding of the genetic characteristics of this important species of Neotropical fish, through the investigation of wild arapaima populations in the Araguaia-Tocantins basin using microsatellite markers. The present study aimed to confirm the findings of Vitorino et al. (2015), and in particular, determine the existence of significant structuring in the arapaima populations of this basin. The prediction that significant population structure exists within the study area is based on the observation that, under natural conditions, arapaima typically form relatively stable family groups, which migrate laterally over short distances (Castello and Stewart, 2010; Castello et al., 2013). In this context, habitat structure is also expected to contribute to genetic diversity, given that marginal lakes increase in number and size moving downstream from the sampling point furthest upstream (point 1, **Figure 1**), with a larger number of environments been expected to support larger numbers of fish and family groups.

While the study of Vitorino et al. (2015) analyzed the genetic diversity of the arapaima populations of the Araguaia-Tocantins basin using ISSR markers, microsatellites are more appropriate for population analyses, given their co-dominant inheritance (vs. dominant alleles in the ISSR markers) and high degree of polymorphism. Microsatellite markers thus provide a much greater potential for statistical analyses, and were expected to provide a detailed database for the more reliable definition of the genetic diversity of populations. These analyses have important implications for the definition of priority areas for conservation and the implementation of management units for this species.

MATERIALS AND METHODS

Two hundred and ninety-four samples of the muscle tissue and fins of *Arapaima gigas* were obtained from five different locations on the Araguaia and Tocantins rivers, three in the state of Mato Grosso (Araguaiana, Novo Santo Antônio, and São Félix do Araguaia), one in Goiás (Luiz Alves), and one in Pará, Itupiranga (**Figure 1** and **Table 1**). The municipality of Araguaiana represents the region furthest upstream of the middle Araguaia River, where the river first forms marginal lakes along its course, appropriate for *A. gigas*. Araguaiana is also the sector with the smallest number of lakes (Alves and Carvalho, 2007).

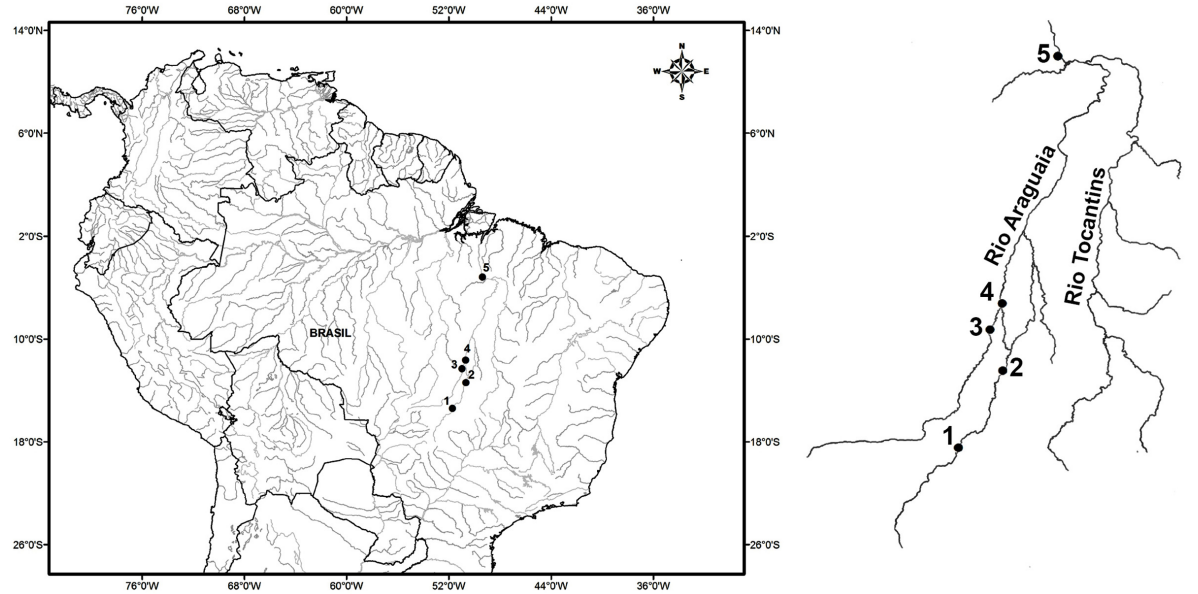


FIGURE 1 | Map of Brazil, with emphasis on the Araguaia-Tocantins basin, indicating the collection sites of *Arapaima gigas*: (1) Araguaiana-MT, (2) Luís Alves, (3) Novo Santo Antônio-MT, (4) São Felix do Araguaia-MT and (5) Itupiranga-PA.

Luis Alves is just upstream from the bifurcation in the river that creates the Javaés River, and forms Bananal Island. It is important to note that this sector of the river is very popular with sport fisherman, due to the large number of lakes found in this area (Alves and Carvalho, 2007). Novo Santo Antônio, a municipality located between the Mortes and Araguaia rivers, is the region's principal arapaima fishery center, and the lakes found along the course of the Mortes River provide the bulk of the catch landed on the Araguaia River (Kirsten et al., 2012). The municipality of São Félix do Araguaia is located at the confluence of the Mortes and Araguaia river, on the Bananal plain (de Melo et al., 2007). Most of the arapaima fishermen of Mato Grosso are resident in this municipality (Kirsten et al., 2012). The municipality of Itupiranga is located downstream of the confluence of Araguaia with Tocantins River, 170 km upstream from the Tucurui hydroelectric dam.

The samples collected in Araguaiana were obtained during a rescue operation, which translocated fish from seasonal lakes to a perennial body of water. Small pieces of fin were collected from the live fish during this operation. All the other tissue samples were obtained from specimens caught and marketed by local fishermen at each site, so no specimens were euthanized specifically for the purposes of the study. All the samples (fin and muscle) were preserved in 100% alcohol, and deposited at the Cytogenetics and Animal Genetics Laboratory, of the Federal University of Mato Grosso (LabGen/GEPEMA/UFMT) in Cuiabá, Brazil. When it was necessary to handle animals, all procedures adhered to the recommendations of the Guide for the Care and Use of Laboratory Animals. The collection and transportation of biological specimens by the authors of the present study is authorized by the Brazilian Institute for the Environment and Renewable Resources (IBAMA) through permanent license number 15226-1, issued by the Chico Mendes Institute for the conservation of Biodiversity (ICMBio).

TABLE 1 | Localization and number of individual (*n*) of Arapaima analyzed in this study.

Locality	Coordinates	*n*
Araguaiana – MT	15°23′58″S/51°43′00″W	60
Luís Alves – GO	13°22′361″S/50°40′08″W	76
Novo Santo Antônio – MT	12°18′56″S/50°58′24″W	86
São Félix do Araguaia – MT	11°38′26″S/50°40′43″W	50
Itupiranga – PA	5°10′46″S/49°21′44″W	14

The total DNA was extracted according to the saline extraction protocol of Aljanabi and Martinez (1997), with minor modifications. The amount and quality of the DNA obtained through this procedure were analyzed in a Biophotometer Plus (Eppendorf Hamburg, Hamburg, Germany), and the samples were diluted to a final concentration of 5 ng/μL.

Seven primer pairs (AgCTm4, AgCTm7, AgCAm2, AgCAm15, AgCAm16, AgCAm20, and AgCAm26) were selected based on previous studies and also because they are highly polymorphic for the *Arapaima* populations of the Amazon Basin (Farias et al., 2003, 2015; Araripe et al., 2013). These primers fluorescently labeled were used for amplification of the microsatellite regions by Polymerase Chain Reaction (PCR), following the conditions described by Farias et al. (2003). For genotyping, the PCR product was mixed with a standard molecular weight marker (MegaBACE ET-550R Size Standard), which was injected into a MegaBace 1000 automatic sequencer (Amersham Biosciences). The alleles were identified using Fragment Profiler 1.2 (Amersham Biosciences).

Once the database was assembled, the existence of possible genotyping errors, null alleles or scoring was verified in MicroChecker (Van Oosterhout et al., 2004). To test whether the populations were in Hardy-Weinberg equilibrium, Genepop (Raymond and Rousset, 1995) was used to estimate the intrapopulation fixation index or coefficient of inbreeding (F_{IS}) for each locus, which was compared with the null hypothesis ($F_{IS} = 0$). Genepop was also used to determine allele frequencies, and the presence of polymorphic loci, and exclusive and rare alleles. Allele richness was estimated in Fstat 2.9.3.2 (Goudet, 2002). Expected and observed heterozygosity were determined by Arlequin 3.5.1.2 (Excoffier and Lischer, 2010). The significance of the differences between expected and observed heterozygosity was evaluated using a one-way ANOVA, run in PAST 2.17c (Hammer et al., 2001).

Genetic differentiation among populations was evaluated using an Analysis of Molecular Variance (AMOVA), the molecular fixation index (F_{ST}), and the divergence parameter (R_{ST}), which were all determined by Arlequin 3.5.1.2 (Excoffier and Lischer, 2010). A Mantel test was used to verify possible correlation between genetic differentiation (F_{ST}, R_{ST} and F_{ST}/1-F_{ST}) and geographical distance, which was measured following the main channel of the river.

The probability of a given number of stocks was based on a Bayesian approach run in STRUCTURE 2.3.3 (Pritchard et al., 2000). The number of presumed populations (K) was set from 1 to 7. The analyses had a burn-in of 100,000 runs, and a Monte Carlo Markov Chain (MCMC) of 1,000,000, using a model without admixture and allele frequencies. The number of populations was defined by the delta *k*-value, obtained using the STRUCTURE HARVESTER program (Earl and VonHoldt, 2012).

The estimated effective size (N_e) of each population was derived from the theta values (θ) generated in MIGRATE-n 3.2.6 (Beerli and Felsenstein, 2001), using the formula $N_e = \theta/4\mu$, with a microsatellite mutation rate (μ) of 5.56×10^{-4} per locus per generation (Whittaker et al., 2003; Yue et al., 2007).

The MIGRATE-n program was also used to calculate migration rates between pairs of populations by the coalescence method. The number of migrants per generation (Nm) was obtained by multiplying the migration rates obtained by the program ($M = m/\mu$, where m is the fraction of new immigrants from the population per generation) by the θ values of the receptor population in each pairwise comparison. This analysis was performed using the Brownian model, with a uniform distribution and a constant mutation rate between the loci. The standard search strategy values of the MIGRATE program were used, except for the use of 20 short chains and 5 long chains.

The BOTTLENECK program, version 1.2.02 (Cornuet and Luikart, 1996) was used to verify the existence of recent demographic events, such as population bottlenecks. The Wilcoxon test was run using the Two-Phased Mutation (T.P.M) model, established with 30% for the Infinite Allele Mutation (IAM) model and 70% for the Stepwise Mutation Model (SMM).

RESULTS

Genetic Diversity

A total of 25 alleles were identified, with an average of 3.57 alleles per locus. While all the loci were polymorphic, four were monomorphic in at least one of the arapaima populations analyzed from the Araguaia-Tocantins basin. The number of alleles, observed and expected heterozygosity, allelic richness, and the fixation index (F_{IS}) were calculated for all seven microsatellite loci analyzed, together with the mean parameters for all loci (**Table 2**).

The population from Novo Santo Antônio had the highest genetic diversity, with 72% of the identified alleles and an allelic richness of 2.429. At the opposite extreme, the arapaima from Araguaiana were the least diverse (44.0%) with an allelic richness of only 1.521. The specimens from Araguaiana also returned the lowest heterozygosity (Ho = 0.133, He = 0.133), while those from Itupiranga had the highest values, with observed heterozygosity of 0.428 and an expected heterozygosity of 0.449 (**Table 2**).

Five loci deviated from Hardy-Weinberg equilibrium, with heterozygote frequencies being either lower or higher than expected (**Table 2**). Only four of the 28 values estimated for the inbreeding coefficient (F_{IS}) were significantly positive, which indicates a deficit of heterozygotes. Three of these loci were recorded in the population from Novo Santo Antônio, while the other was from São Felix do Araguaia. A single locus from Itupiranga was significantly negative, indicating an excess of heterozygotes. When the whole set of loci is considered, a significant F_{IS} value ($p < 0.001$) was only for the Novo Santo Antônio presented, indicating the occurrence of inbreeding in this population.

One to three alleles, with high frequencies, were recorded for most loci. All loci except AgCAm15 presented exclusive alleles for the different study populations, at frequencies ranging from 0.012 to 1.00. No exclusive alleles was detected in the population from São Felix do Araguaia (**Table 3**). No new alleles were detected in this study, given that all the alleles recorded here had been reported previously by Farias et al. (2003) and Araripe (2008).

The MicroChecker analysis (**Table 3**) indicated the presence of null alleles in the populations from Luís Alves (for loci AgCAm20 and AgCTm4), Novo Santo Antônio (loci AgCAm2, AgCTm4, AgCAm16 and AgCAm20), São Félix do Araguaia (AgCAm15), and Itupiranga (AgCTm4). No systematic pattern was observed in the occurrence of null alleles in the different study populations, however. No evidence was found of either the misidentification of stutters as alleles or dropouts (dominance of small alleles).

Population Structure

The values recorded for Wright index (F_{ST}) ranged from 0.061 to 0.669, while those for the Slatkin divergence parameter (R_{ST}), varied from 0.025 to 0.688. Both indices were significant for all pairs of locations (**Table 4**). The results of the Mantel test rejected the hypothesis that genetic differentiation was related to geographical distance, considering the values of either F_{ST}

($r^2 = 0.466115$, $p = 0.236$), R_{ST} ($r^2 = 0.452354$, $p = 0.240$) or F_{ST}/1-F_{ST} ($r^2 = 0.219112$, $p = 0.221000$). The results of the Bayesian inference based on the mean likelihood [Ln (K)] of the Δk (Evanno et al., 2005) indicate the presence of two clusters within the Araguaia-Tocantins basin, one formed by the populations of Luís Alves, Novo Santo Antônio, São Félix do Araguaia, and Itupiranga, and the other by that of Araguaiana, which forms a distinct group (**Figure 2**). These findings were used to establish the clusters for the analysis of molecular variance (AMOVA).

The results of the AMOVA indicated a lack of significant variation between the two groups ($\Phi_{ST} = 44.15\%$, $R_{ST} = 40.07\%$). The variation among populations within each group contributed only 12.35% (Φ_{ST}) and 11.17% (R_{ST}) of the variance, while the variance within populations is 43.51% for Φ_{ST} and 48.76 for R_{ST}.

TABLE 2 | Genetic diversity indexes for *Arapaima gigas*, using microsatellite markers.

Loci		Population				
		Ara	LAI	NSA	SFA	Itu
AgCTm7	*N*	2	2	2	2	2
	A_R	1.999	1.993	1.997	2.000	3.000
	H_O	0.271	0.254	0.294	0.520	0.714*
	H_E	0.305	0.346	0.284	0.504	0.679
	Fis	0.115	−0.168	−0.033	−0.032	−0.053*
AgCAm20	*N*	1	1	1	1	3
	A_R	1.000	1.000	1.000	1.000	3.000
	H_O	Mono	Mono	Mono	Mono	0.428
	H_E	Mono	Mono	Mono	Mono	0.574
	Fis	–	–	–	–	0.261
AgCAm2	*N*	1	3	5	4	3
	A_R	1.000	2.247	2.676	2.913	3.000
	H_O	Mono	0.260	0.131*	0.286	0.215
	H_E	Mono	0.236	0.188	0.253	0.315
	Fis	–	−0.103	0.304*	0.004	0.328
AgCAm26	*N*	2	2	2	2	2
	A_R	1.998	2.000	2.000	2.000	2.000
	H_O	0.356	0.587	0.442	0.480	0.571
	H_E	0.295	0.501	0.449	0.504	0.518
	Fis	−0.208	−0.172	0.017	0.049	−0.106
AgCAm16	*N*	2	2	2	2	1
	A_R	1.995	1.335	1.894	1.738	1.000
	H_O	0.271	0.026	0.094*	0.0802	Mono
	H_E	0.261	0.026	0.132	0.078	Mono
	Fis	−0.04	−0.006	0.288*	−0.032	–
AgCTm4	*N*	1	3	3	3	2
	A_R	1.000	2.203	2.174	2.094	2.000
	H_O	Mono	0.123	0.059*	0.120	0.286
	H_E	Mono	0.118	0.144	0.115	0.508
	Fis	–	−0.04	0.592*	−0.044	0.477
AgCAm15	*N*	2	2	3	3	3
	A_R	1.689	2.000	2.171	2.280	3.000
	H_O	0.036	0.343	0.4393	0.3203*	0.786
	H_E	0.069	0.339	0.4643	0.4793	0.553
	Fis	0.488	−0.011	0.055	0.359*	−0.444
Mean	*N*	1.571	2.143	2.571	2.429	2.429
	A_R	1.526	1.825	1.987	2.004	2.429
	H_O	0.133	0.228	0.209	0.258	0.429
	H_E	0.133	0.190	0.176	0.276	0.450
Total	P_L	4	6	6	6	6
	Allele	11	15	18	17	17
	% allele	44%	60%	72%	68%	68%

*N, number of alleles; A_R, allelic richness; H_O, observed heterozygosity; H_E, expected heterozygosity; Fis, coefficient of inbreeding computed as in Weir and Cockerham (1984); P_L, polymorphic locus; –, monomorphic locus. *P < 0.05. Ara, Araguaiana; LAI, Luís Alves; NSA, Novo Santo Antônio; SFA, São Felix do Araguaia and Itu, Itupiranga.*

TABLE 3 | Allele frequency and presence of exclusive alleles.

Loci	Allele	Population				
		Ara	LAI	NSA	SFA	Itu
AgCTm7	279	0.186	0.851	0.171	0.520	0.250
	285	0.814	0.148	0.829	0.480	0.357
	295	0.000	0.000	0.000	0.000	0.393*
AgCAm20	265	0.000	0.000	0.000	0.000	0.393*
	267	0.000	0.000	0.000	0.000	0.071*
	269	1.000	1.000	1.000	1.000	0.536
AgCAm2	313	0.000	0.000	0.006	0.000	0.821
	315	0.000	0.000	0.000	0.010	0.036
	317	1.000*	0.000	0.000	0.000	0.000
	321	0.000	0.055	0.012	0.0,31	0.000
	323	0.000	0.868	0.899	0.837	0.143
	325	0.000	0.076	0.071	0.122	0.000
	333	0.000	0.000	0.012*	0.000	0.000
AgCAm26	213	0.822	0.532	0.663	0.520	0.500
	217	0.178	0.468	0.337	0.480	0.500
AgCAm16	251	0.153*	0.000	0.000	0.000	0.000
	255	0.000	0.986	0.929	0.960	1.000
	257	0.847	0.013	0.071	0.040	0.000
AgCTm4	275	0.000	0.041*	0.000	0.000	0.000
	277	1.000	0.938	0.924	0.940	0.571
	279	0.000	0.020	0.065	0.050	0.000
	283	0.000	0.000	0.012	0.010	0.429
AgCAm15	228	0.036	0.785	0.646	0.650	0.464
	230	0.000	0.000	0.006	0.010	0.036
	244	0.964	0.214	0.348	0.340	0.500

*Ara, Araguaiana; LAI, Luís Alves; NSA, Novo Santo Antônio; SFA, São Félix do Araguaia and Itu, Itupiranga. *Exclusive allele.*

TABLE 4 | F_{ST}/R_{ST} values and geographic distance in kilometers (above the diagonal) between *A. gigas* populations from Araguaia-Tocantins basin.

Population	Ara	LAI	NSA	SFA	Itu
Ara	–	329.8	653.2	594.9	1589.3
LAI	0.669/0.688	–	323.4	265.1	1259.5
NSA	0.563/0.512	0.245/0.146	–	111	1105.4
SFA	0.577/0.548	0.061/0.044	0.072/00251	–	994.4
Itu	0.600/0.4384	0.425/0.510	0.396/0.3224	0.322/0.3431	–

Ara, Araguaiana; LAI, Luís Alves; NSA, Novo Santo Antônio; SFA, São Félix do Araguaia and Itu, Itupiranga.

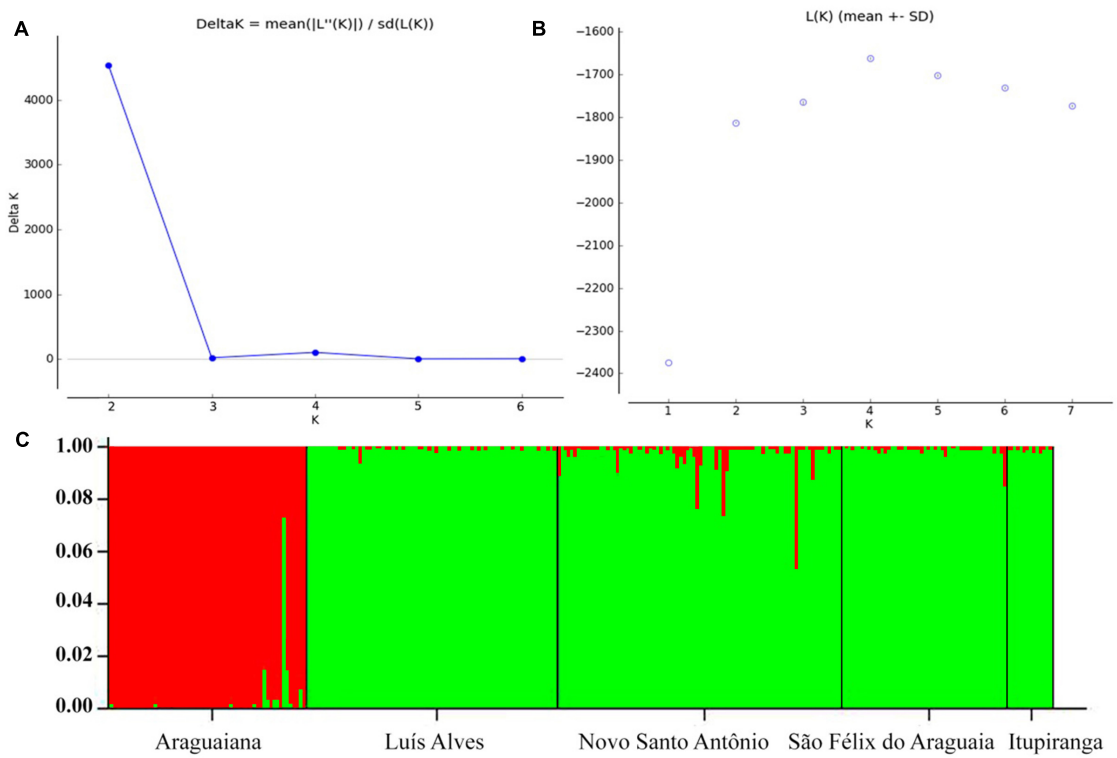


FIGURE 2 | Genetic structuring of *Arapaima gigas* from the Araguaia-Tocantins basin, according to Bayesian analysis. Estimating the number of K groups based on **(A)** ΔK and **(B)** the mean maximum likelihood. **(C)** Each column represents a different individual and the colors represent genetic stocks. The collection sites are separated by black lines.

The Φ_{ST} ($\Phi_{ST} = 0.564$; $p = 0.0000$) and ($R_{ST} = 0.512$; $p = 0.000$) were both relatively high among the study populations.

Effective population size ranged from 128 to 351 individuals, with the lowest value being recorded at Araguaiana and the highest at São Félix do Araguaia (**Table 5**). The MIGRATE-n coalescence analysis (Beerli and Felsenstein, 2001) indicated low levels of gene flow between populations, and only six of the 20 estimates of migrant numbers were higher than one migrant per generation.

The Bottleneck program did not detect any significant deficit of heterozygotes. However, a significant ($p < 0.05$) excess of heterozygotes was found in the Itupiranga population using the TPM (Wilcoxon test) model.

DISCUSSION

This study is the first to use microsatellite markers to examine the genetic diversity of the arapaima populations of the Araguaia-Tocantins basin. Overall, the results indicated significantly lower levels of genetic diversity and heterozygosity than those found in previous studies of the same genetic markers in populations from other areas (Farias et al., 2003, 2015; Hamoy et al., 2008; de Alencar Leão, 2009; Araripe et al., 2013).

The arapaima populations that inhabit the study region are affected by natural processes, in particular the hydrological regime, that are quite distinct in comparison with the populations found in the Amazon basin, the region in which the species has been investigated in most detail (Vitorino et al., 2015). The unique features of the Araguaia-Tocantins basin may influence the reduced genetic diversity (number of alleles, heterozygosity, and allelic richness) of its population in comparison with those from the Amazonian basin. This is also true of the allelic richness of the arapaima populations of Tucuruí (Araripe, 2008) and Bananal Island (de Alencar Leão, 2009), two other areas located within the Araguaia-Tocantins basin, which present low values in comparison with population from the Amazon basin. The levels of genetic diversity recorded in the *Arapaima* populations of the Araguaia-Tocantins basin are lower than expected for the species, but are consistent with the findings of Vitorino et al. (2015), based on ISSR markers. Using mitochondrial markers, da Penha (2014) also found relatively low genetic variability in arapaima populations of the Araguaia-Tocantins basin in comparison with those of the Amazon basin.

The points sampled in the present study are exploited intensively by local fisheries. In fact, most of the arapaima marketed locally are harvested from natural populations (Çiftci and Okumus, 2002), including not only adults, but also the capture of fry as stock for rearing on fish farms. This type of overexploitation often results in a decline in effective population size (N_e), provoking the loss of genetic diversity, which leads to a reduced resilience of populations to environmental stresses and

climate change, and a loss of resistance to pathogens (Allendorf and Luikart, 2007; Hughes et al., 2015).

The variation in the N_e values estimated for the different populations analyzed in the present study indicates potential differences in their population dynamics, with distinct patterns of fluctuation over time in the number of individuals reproducing, a process that will have knock-on effects for the sustainability of the population (Carson et al., 2011). In other words, each arapaima population may have a distinct demographic history, based on the differences in variables such as the number of mature individuals, the sex ratio, offspring survival rates, and the availability and quality of habitats. However, the results of the Bottleneck analysis indicate a significant impact on genetic structure only in the case of the population from Itupiranga. This may be related to the construction of the hydroelectric dam at Tucuruí, and the creation of one of the world's largest reservoirs, which caused considerable impacts on the local fish fauna of the Tocantins River.

The reduced genetic variability found in the arapaima populations of the Araguaia-Tocantins basin may also be related to the anthropogenic impacts that have altered the natural features of the basin extensively throughout most of its length. The hydrological cycle of this basin is more intense than that of the Amazon, and this cycle has been modified by local farming and ranching activities, and the establishment of the Araguaia-Tocantins waterway (Latrubesse and Stevaux, 2006). These modifications have had a major impact, principally on the region's lacustrine environments, causing a reduction in the size of its lakes, and altering their flood cycle, which affects the lateral migrations typical of this fish (Castello and Stewart, 2010; Castello et al., 2013). These changes may force the arapaima to remain in the same lakes, unable to migrate in search of reproductive partners, eventually creating small, isolated populations susceptible to inbreeding. Genetic drift tends to have a greater impact in smaller populations, which also favor inbreeding, exacerbating the loss of genetic diversity. It seems likely that the combined effects of these processes have reduced the reproductive potential of the local arapaima populations, impacting their effective size. The geomorphology and flood cycle of the Araguaia-Tocantins and Amazon basins are quite distinct, and this appears to be the principal factor determining the variation in the genetic diversity of the stocks analyzed from the Amazon (Farias et al., 2003, 2015; Hamoy et al., 2008; de Alencar Leão, 2009; Araripe et al., 2013) and the Tocantins-Araguaia basin (Vitorino et al., 2015).

However, the possibility that low genetic variability is a natural characteristic of the arapaima populations of the Araguaia-Tocantins basin cannot be ruled out altogether. But whatever the determining factors, this genetic fragility is a cause for concern, given that future environmental impacts (natural or otherwise) may further reduce the diversity of these populations, and threaten their long-term viability.

This fragility is further reinforced by the differences in the genetic variability of each population in the Araguaia-Tocantins basin. Specimens collected at Araguaiana, for example, had the lowest genetic diversity of any population, indicating the smallest effective size of any population, which implies that this population is the most vulnerable to anthropogenic interference from the expansion of the agricultural frontier occurring within the basin (Latrubesse and Stevaux, 2006), as well as being the sector of the basin that has the smallest number of lakes, the preferred habitat of the arapaima (Alves and Carvalho, 2007).

The most likely explanation for the deviations from Hardy-Weinberg equilibrium detected in the arapaima populations of the Araguaia-Tocantins basin is inbreeding, given that significant F_{IS} values (indicating a deficiency of heterozygotes) were found in six of the eight deviations recorded. A number of studies have recorded a deficiency of heterozygotes in fish populations (Castric et al., 2002; Langen et al., 2011; O'Leary et al., 2013; Ferreira et al., 2015), indicating that inbreeding may be relatively common in these vertebrates. In the arapaima this may be related to the relatively sedentary behavior of the species and its parental care, as observed in a number of other species of Neotropical fish (Sofia et al., 2006, 2008; Ferreira et al., 2015).

TABLE 5 | Analysis of the MIGRATE program showing the estimates of gene flow peer-to-peer among the Arapaima populations of the Araguaia-Tocantins region.

Emigrant population	θ	N_e	Nm [Receiving population]				
	($4N_e\mu$)	($\theta/4\mu$)	Ara	LAl	NSA	SFA	Itu
Ara	0.2847 (0.258–0.315)	128.0126 (116.0–141.63)	–	0.7053 (0.444–1.074)	0.8615 (0.543–1.308)	0.8624 (0.549–1.306)	0.3151 (0.184–0.546)
LAl	0.5284 (0.485–0.576)	237.5899 (218.07–258.99)	0.0632 (0.023–0.139)	–	0.5272 (0.295–0.875)	1.5895 (1.117–2.220)	0.2170 (0.116–0.400)
NSA	0.5573 (0.516–0.602)	250.5845 (232.01–270.68)	0.3663 (0.234–0.553)	1.1346 (0.776–1.618)	–	1.2733 (0.860–1.831)	0.0904 (0.036–0.203)
SFA	0.7820 (0.713–0.860)	351.6187 (320.59–386.69)	0.3529 (0.225–0.534)	1.0905 (0.742–1,559)	2.7187 (2.039–3.575)	–	0.1386 (0.065–0.280)
Itu	0.4237 (0.362–0.524)	190.5126 (162.76–235.61)	0.3330 (0.210–0.507)	1.2991 (0.905–1.821)	0.6836 (0.379–1.102)	0.1635 (0.061–0.357)	–

The same pair of populations may have different numbers of immigrant and migrant individuals. Ln(L): −2473.954. θ = Population genetic parameter defined as $4N_e\mu$ (Effective population size multiplied by the mutation rate), Ne, estimate of effective population size. Nm, estimate of the number of migrants defined as NeM (effective population size multiplied by the migration rate). Ln (L) = Ln(L), Natural log of the probability of the estimated parameters. Ara, Araguaiana; LAl, Luís Alves; NSA, Novo Santo Antônio; SFA, São Félix do Araguaia and Itu, Itupiranga.

In addition to these behavioral traits, arapaima is intolerant of lotic environments, so areas of strong rapids represent effective barriers to the dispersal of this species (Castello and Stewart, 2010; Castello et al., 2013). This may lead to the formation of family groups in the different micro-regions of the Araguaia-Tocantins basin. Genetic differences between subpopulations related to geographic distance or the presence of physical barriers, such as waterfalls, have been identified in Amazonian arapaima (Araripe et al., 2013) and in other fish species (Hughes et al., 2015).

The pairwise F_{ST} values indicate moderate (0.05–0.15) to extreme (>0.25) genetic differentiation between populations. However, the Bayesian analysis points to the presence of only two genetic stocks in the Araguaia-Tocantins basin, one at Araguaiana, and the other formed by the remaining populations, at Luís Alves, Novo Santo Antônio, Sao Félix do Araguaia and Itupiranga (**Figure 2**). This arrangement contrasts with that recorded by Vitorino et al. (2015), based on ISSR markers, which indicated that the Araguaiana and Novo Santo Antônio populations shared the same genetic stock, while São Félix do Araguaia and Itupiranga form separate populations (Luís Alves was not included in this analysis).

The structure proposed by the Bayesian analysis is also supported by the estimates of gene flow and the number of migrants between populations, given that the populations that share the same genetic stock had the highest migration rates. It is important to note that the gene flow detected here may reflect ancestral processes, rather than the recent exchange of individuals, because the number of migrants is a somewhat abstract quantity, which cannot be distinguished from the effective population size (Balloux and Lugon-Moulin, 2002).

While the genetic variation found in the present study is not related systematically to geographic distance, the geographically closer populations (Luís Alves, Novo Santo Antônio and São Félix do Araguaia) returned the lowest F_{ST} and R_{ST} values. Reduced genetic differentiation was expected between the samples from Novo Santo Antônio and São Félix do Araguaia, which are the geographically closest sites (separated by a distance of only 111 km), although the F_{ST} values indicate that the populations of São Félix do Araguaia and Luís Alves (265.1 km apart) are the most similar. In addition to the basic differences in comparison with the Amazonian arapaima populations, then, the relationship between population structure and geographic distance is also distinct from that recorded by Araripe et al. (2013).

The present study confirmed the low genetic diversity of the *Arapaima gigas* populations of the Araguaia-Tocantins basin, which can be linked directly to the environmental fragility of this river system, which reinforces the need for a better understanding of the processes that may further reduce the viability of these populations. The findings of this study also indicated that the genetic diversity of the populations is distributed heterogeneously within the study area, and that the establishment of a single protected area would be insufficient for the preservation of the genetic diversity of the arapaima populations of the Araguaia-Tocantins basin as a whole. In particular, future management measures should consider the population from Araguaiana as an independent unit, distinct from the other Araguaia-Tocantins populations. The Itupiranga region should also be defined as a priority area for conservation, given the high allelic richness found in this population.

ETHICS STATEMENT

This study was carried out in strict accordance with the recommendations provided in the Guide for the Care and Use of laboratory Animals. During the development of this work, no animals were sacrificed. Collection was authorized by SEMA (license number 104358/2011/SEMA-MT), Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis - IBAMA and Instituto Chico Mendes de Conservação da Biodiversidade – ICMBio (License number 15226-1).

AUTHOR CONTRIBUTIONS

CV performed the molecular genetic studies, performed the statistical analyzes and drafted the manuscript. FN performed some molecular genetic studies and contributed to the correction of the text. JA and PV conceived and coordinated the study, participated in its elaboration and helped to draft the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

The authors are grateful to FAPEMAT (Fundação de Amparo à Pesquisa do Estado de Mato Grosso – Process: 841147/2009/FAPEMAT/PRONEX), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico/INAU), and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for financial support; Drs. Fabio Porto Foresti and Claudio Oliveira for providing biological samples. We are also very grateful to Mr. Irineu Pirani and Wagner Alves de Santana, for permitting the collection of arapaima tissue samples during rescue operations.

REFERENCES

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., et al. (2008). Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. *Bioscience* 58, 403–414. doi: 10.1641/B580507

Agostinho, A. A., Thomaz, S. M., and Gomes, L. C. (2005). Conservação da biodiversidade em águas continentais do Brasil. *Megadiversidade* 1, 70–78.

Aljanabi, S. M., and Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* 25, 4692–4693. doi: 10.1093/nar/25.22.4692

Allan, J. D., Abell, R., Hogan, Z., Revenga, C., Taylor, B. W., Welcomme, R. L., et al. (2005). Overfishing of inland waters. *Bioscience* 55, 1041–1051. doi: 10.1641/0006-3568(2005)055[1041:OOIW]2.0.CO;2

Allendorf, F. W., and Luikart, G. H. (2007). *Conservation an the Genetics of Populations*. Oxford: Blackwell Publishing Ltd.

Almeida, A., and Peres, F. C. (2007). Hidrovia Tocantins - Araguaia: importância e impactos econômicos, sociais e ambientais, segundo a percepção dos agentes econômicos locais. *Rev. Bras. Recur. Hídricos* 12, 169–177. doi: 10.21168/rbrh.v12n2.p169-177

Alves, T. M., and Carvalho, T. M. (2007). Técnicas de sensoriamento remoto para classificação e quantificação do sistema lacustre do Rio Araguaia entre Barra do Garças e foz do Rio Cristalino. *Rev. Geogr. Acad.* 1, 79–94.

Aquino, S., Latrubesse, E., and Bayer, M. (2009). Assessment of wash load transport in the Araguaia River (Aruanã gauge station), central Brazil. *Lat. Am. J. Sedimentol. Basin Anal.* 16, 119–128.

Aquino, S., Latrubesse, E. M., and de Souza Filho, E. E. (2008). Relações entre o regime hidrológico e os ecossistemas aquáticos da planície aluvial do Rio Araguaia. *Acta Sci. Biol. Sci.* 30, 361–369. doi: 10.4025/actascibiolsci.v30i4.5866

Aquino, S., Stevaux, J. C., and Latrubesse, E. M. (2005). Regime hidrológico e aspectos do comportamento morfohidráulico do Rio Araguaia. *Rev. Bras. Geomorfol.* 6, 29–41. doi: 10.20502/rbg.v6i2.49

Arantes, C. C., Castello, L., Stewart, D. J., Cetra, M., and Queiroz, H. L. (2010). Population density, growth and reproduction of arapaima in an Amazonian river-floodplain. *Ecol. Freshw. Fish* 19, 455–465. doi: 10.1111/j.1600-0633.2010.00431.x

Araripe, J. (2008). *Genética de Populações de Pirarucus (Arapaima gigas) da Reserva de Mamirauá e Considerações sobre a Estrutura Genética para a Espécie*. Belém: Universidade Federal do Pará.

Araripe, J., Rêgo, P. S., Queiroz, H., Sampaio, I., and Schneider, H. (2013). Dispersal capacity and genetic structure of *Arapaima gigas* on different geographic scales using microsatellite markers. *PLOS ONE* 8:e54470. doi: 10.1371/journal.pone.0054470

Balloux, F., and Lugon-Moulin, N. (2002). The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* 11, 155–165. doi: 10.1046/j.0962-1083.2001.01436.x

Beerli, P., and Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4563–4568. doi: 10.1073/pnas.081068098

Carson, E. W., Saillant, E., Renshaw, M. A., Cummings, N. J., and Gold, J. R. (2011). Population structure, long-term connectivity, and effective size of mutton snapper (*Lutjanus analis*) in the Caribbean Sea and Florida keys. *Fish. Bull.* 109, 416–428.

Castello, L. (2008a). Lateral migration of *Arapaima gigas* in floodplains of the Amazon. *Ecol. Freshw. Fish* 17, 38–46. doi: 10.1111/j.1600-0633.2007.00255.x

Castello, L. (2008b). Nesting habitat of *Arapaima gigas* (Schinz) in Amazonian floodplains. *J. Fish Biol.* 72, 1520–1528. doi: 10.1111/j.1095-8649.2007.01778.x

Castello, L., and Stewart, D. J. (2010). Assessing CITES non-detriment findings procedures for *Arapaima* in Brazil. *J. Appl. Ichthyol.* 26, 49–56. doi: 10.1111/j.1439-0426.2009.01355.x

Castello, L., Stewart, D. J., and Arantes, C. C. (2013). "'O quê sabemos e precisamos fazer a respeito da conservação do pirarucu (*Arapaima* spp.) na Amazônia,'" in *Biologia, Conservação e Manejo Participativo de Pirarucu na Pan-Amazônia*, ed. E. Amaral (Tefé: Instituto de Desenvolvimento Sustentável Mamirauá), 17–31.

Castric, V., Bernatchez, L., Belkhir, K., and Bonhomme, F. (2002). Heterozygote deficiencies in small lacustrine populations of brook Charr *Salvelinus fontinalis* Mitchill (Pisces, Salmonidae): a test of alternative hypotheses. *Heredity* 89, 27–35. doi: 10.1038/sj.hdy.6800089

Çiftci, Y., and Okumus, I. (2002). Fish population genetics and applications of molecular markers to fisheries and aquaculture: I- basic principles of fish population genetics. *Turk. J. Fish. Aquat. Sci.* 2, 145–155.

Cornuet, J. M., and Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144, 2001–2014.

da Penha, F. N. (2014). *Distintas Linhagens do Gigante Arapaima (Osteoglossiformes: Osteoglossidae) Revelam Diferenças Entre Populações das Bacias Amazônica e Araguaia-Tocantins*. Master's thesis, Universidade Federal do Pará, Pará.

de Alencar Leão, A. S. (2009). *Análise da Variabilidade Genética das Populações de Pirarucu (Arapaima gigas, Schinz 1822) dos Principais Tributários do Rio Amazonas Através do uso de Marcadores Microssatélites*. Master's thesis, Universidade Federal do Pará, Pará.

de Melo, T. L., Tejerina-Garro, F. L., and Melo, C. E. (2007). Diversidade biológica da comunidade de peixes no baixo Rio das Mortes, Mato Grosso, Brasil. *Rev. Bras. Zool.* 24, 657–665. doi: 10.1590/S0101-81752007000300017

Earl, D. A., and VonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Farias, I. P., Hrbek, T., Brinkmann, H., Sampaio, I., and Meyer, A. (2003). Characterization and isolation of DNA microsatellite primers for *Arapaima gigas*, an economically important but severely over-exploited fish species of the Amazon basin. *Mol. Ecol. Resour.* 1, 128–130. doi: 10.1046/j.1471-8286

Farias, I. P., Leão, A., Almeida, Y. S., Verba, J. T., Crossa, M. M., Honczaryk, A., et al. (2015). Evidence of polygamy in the socially monogamous Amazonian fish *Arapaima gigas* (Schinz, 1822) (Osteoglossiformes, Arapaimidae). *Neotrop. Ichthyol.* 13, 195–204. doi: 10.1590/1982-0224-20140010

Fernandes, M. N., da Cruz, A. L., da Costa, O. T. F., and Perry, S. F. (2012). Morphometric partitioning of the respiratory surface area and diffusion capacity of the gills and swim bladder in juvenile Amazonian air-breathing fish. *Arapaima gigas*. *Micron* 43, 961–970. doi: 10.1016/j.micron.2012.03.018

Ferreira, D. G., Galindo, B. A., Frantine-Silva, W., Almeida, F. S., and Sofia, S. H. (2015). Genetic structure of a Neotropical sedentary fish revealed by AFLP, microsatellite and mtDNA markers: a case study. *Conserv. Genet.* 16, 151–166. doi: 10.1007/s10592-014-0648-2

Gomes, L. D. C. (2007). Physiological responses of pirarucu (*Arapaima gigas*) to acute handling stress. *Acta Amazon.* 37, 629–634. doi: 10.1590/S0044-59672007000400019

Goudet, G. J. (2002). *FSTAT, A Program for Windows to Estimate and Test Gene Diversities and Fixation Indices (Version 2.9.3.2)*. Lausanne: Institute of Ecology.

Hammer, Ø., Harper, D. A. T., and Ryan, P. D. (2001). Paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 4, 9–18. doi: 10.1016/j.bcp.2008.05.025

Hamoy, I. G., Santos, E. J. M., and Santos, S. E. B. (2008). Rapid and inexpensive analysis of genetic variability in *Arapaima gigas* by PCR multiplex panel of eight microsatellites. *Genet. Mol. Res.* 7, 29–32. doi: 10.4238/vol7-1gmr394

Hrbek, T., Crossa, M., and Farias, I. P. (2007). Conservation strategies for *Arapaima gigas* (Schinz, 1822) and the Amazonian várzea ecosystem. *Braz. J. Biol.* 67, 909–917. doi: 10.1590/S1519-69842007000500015

Hrbek, T., and Farias, I. P. (2008). The complete mitochondrial genome of the pirarucu. *Genet. Mol. Biol.* 31, 293–302. doi: 10.1590/S1415-47572008000200024

Hrbek, T., Farias, I. P., Crossa, M., Sampaio, I., Porto, J. I. R., and Meyer, A. (2005). Population genetic analysis of *Arapaima gigas*, one of the largest freshwater fishes of the Amazon basin: implications for its conservation. *Anim. Conserv.* 8, 297–308. doi: 10.1017/S1367943005002210

Hughes, J. M., Schmidt, D. J., Huey, J. A., Real, K. M., Espinoza, T., Mcdougall, A., et al. (2015). Extremely low microsatellite diversity but distinct population structure in a long-lived threatened species, the Australian lungfish *Neoceratodus forsteri* (Dipnoi). *PLOS ONE* 10:e0121858. doi: 10.1371/journal.pone.0121858

Kirsten, I. F., Puerta, L. R., de Mateus, L. A. F., Catella, A. C., and Lima, I. S. (2012). A pesca do pirarucu (*Arapaima* sp.) na bacia do Rio Araguaia em Mato Grosso - Brasil. *Bol. Inst. Pesca* 38, 131–144.

Langen, K., Schwarzer, J., Kullmann, H., Bakker, T. C. M., and Thünken, T. (2011). Microsatellite support for active inbreeding in a cichlid fish. *PLOS ONE* 6:e24689. doi: 10.1371/journal.pone.0024689

Latrubesse, E. M., Amsler, M. L., de Morais, R. P., and Aquino, S. (2009). The geomorphologic response of a large pristine alluvial river to tremendous deforestation in the South American tropics: the case of the Araguaia River. *Geomorphology* 113, 239–252. doi: 10.1016/j.geomorph.2009.03.014

Latrubesse, E. M., and Stevaux, J. C. (2002). Geomorphology and environmental aspects of the Araguaia fluvial basin, Brazil. *Z. Geomorphol.* 129, 109–127.

Latrubesse, E. M., and Stevaux, J. C. (2006). Características físico-bióticas e problemas ambientais associados à planície aluvial do Rio Araguaia. Brasil Central. *Rev. Geociências* 10, 67–75.

Marques, D. K., Venere, P. C., and Galetti, P. M. Jr. (2006). Chromosomal characterization of the bonytongue *Arapaima gigas* (Osteoglossiformes: Arapaimidae). *Neotrop. Ichthyol.* 4, 215–218. doi: 10.1590/S1679-62252006000200007

O'Leary, S. J., Hice, L. A., Feldheim, K. A., Frisk, M. G., McElroy, A. E., Fast, M. D., et al. (2013). Severe inbreeding and small effective number of breeders in a formerly abundant marine fish. *PLOS ONE* 8:e66126. doi: 10.1371/journal.pone.0066126

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Raymond, M., and Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* 86, 248–249. doi: 10.1093/oxfordjournals.jhered.a111573

Sofia, S. H., Galindo, B. A., Paula, F. M., Sodré, L. M. K., and Martinez, C. B. R. (2008). Genetic diversity of *Hypostomus ancistroides* (Teleostei, Loricariidae) from an urban stream. *Genet. Mol. Biol.* 1, 317–323. doi: 10.1111/jfb.12675

Sofia, S. H., Silva, C. R. M., Galindo, B. A., Almeida, F. S., Sodré, L. M. K., and Martinez, C. B. R. (2006). Population genetic structure of *Astyanax scabripinnis* (Teleostei, Characidae) from an Urban Stream. *Hydrobiologia* 553, 245–254. doi: 10.1007/s10750-005-1106-4

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Resour.* 4, 535–538. doi: 10.1111/j.1471-8286.2004.00684.x

Vitorino, C. A., Oliveira, R. C. C., Margarido, V. P., and Venere, P. C. (2015). Genetic diversity of *Arapaima gigas* (Schinz, 1822) (Osteoglossiformes: Arapaimidae) in the Araguaia-Tocantins basin estimated by ISSR marker. *Neotrop. Ichthyol.* 13, 557–568. doi: 10.1590/1982-0224-20150037

Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi: 10.1111/j.1558-5646.1984.tb05657.x

Whittaker, J. C., Harbord, R. M., Boxall, N., Mackay, I., Dawson, G., and Sibly, R. M. (2003). Likelihood-based estimation of microsatellite mutation rates. *Genetics* 164, 781–787.

World Conservation Monitoring Centre (2014). *The IUCN Red List of Threatened Species. Version: 2014.3*. Available at: www.iucnredlist.org [accessed May 4, 2015].

Yue, G. H., David, L., and Orban, L. (2007). Mutation rate and pattern of microsatellites in common carp (*Cyprinus carpio* L.). *Genetica* 129, 329–331. doi: 10.1007/s10709-006-0003-8

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

Copyright © 2017 Vitorino, Nogueira, Souza, Araripe and Venere. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 24 November 2017
doi: 10.3389/fgene.2017.00185

Hidden Diversity in the Populations of the Armored Catfish *Ancistrus* Kner, 1854 (Loricariidae, Hypostominae) from the Paraná River Basin Revealed by Molecular and Cytogenetic Data

Ana C. Prizon[1], Daniel P. Bruschi[2], Luciana A. Borin-Carvalho[1], Andréa Cius[1], Ligia M. Barbosa[1], Henrique B. Ruiz[3], Claudio H. Zawadzki[3], Alberto S. Fenocchio[4] and Ana L. de Brito Portela-Castro[1]*

[1] *Laboratório de Citogenética de Vertebrados, Departamento de Biotecnologia, Genética e Biologia Celular, Universidade Estadual de Maringá, Maringá, Brazil,* [2] *Laboratório de Citogenética Animal e Mutagênese Ambioental, Departamento de Genética, Universidade Federal do Paraná, Curitiba, Brazil,* [3] *Departamento de Biologia/Nupélia, Universidade Estadual de Maringá, Maringá, Brazil,* [4] *Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones, Posadas, Argentina*

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Sandra Mariotto, Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso, Brazil
Shane Lavery, University of Auckland, New Zealand

***Correspondence:**
Daniel P. Bruschi
danielpachecobruschi@gmail.com

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

Received: *12 August 2017*
Accepted: *07 November 2017*
Published: *24 November 2017*

Citation:
Prizon AC, Bruschi DP, Borin-Carvalho LA, Cius A, Barbosa LM, Ruiz HB, Zawadzki CH, Fenocchio AS and Portela-Castro ALdB (2017) Hidden Diversity in the Populations of the Armored Catfish Ancistrus Kner, 1854 (Loricariidae, Hypostominae) from the Paraná River Basin Revealed by Molecular and Cytogenetic Data. Front. Genet. 8:185. doi: 10.3389/fgene.2017.00185

Only one species of armored catfish, *Ancistrus cirrhosus* Valenciennes 1836, has been historically described in the basin of the Paraná River, from Misiones (Argentina). However, the ample variation found in the morphology and coloration of the populations sampled in the tributaries of the Brazilian state of Paraná makes it difficult to establish the real taxonomic status and evolutionary history of the *Ancistrus* specimens, suggesting that *A. cirrhosus* is not the only species found in this basin. By combining data on mitochondrial DNA (COI gene) and chromosomal markers from different *Ancistrus* populations, totaling 144 specimens, in the tributaries of the Paraná, and specimens from Misiones (type-locality of *A. cirrhosus*), we detected five distinct evolutionary lineages. All the specimens were 2n = 50, but had four distinct karyotype formulae. The results of the Generalized Mixed Yule Coalescent (GYMC) and the genetic distances (uncorrected *P*-values) between lineages ranged from 3 to 5%. Clusters of 18S rDNA were observed in a single chromosome pair in seven populations of *Ancistrus,* but at different positions, in some cases, in synteny with the 5S rDNA sites. Multiple 5S sites were observed in all populations. Overall, the cytogenetic data reinforce the genetic evidence of the diversification of lineages, and indicate the existence of candidate species in the study region. The evidence indicates that at least four candidate species of the *Ancistrus* may coexist in the Paraná basin besides *A. cirrhosus*. Overall, our results provide a comprehensive scenario for the genetic variation among *Ancistrus* populations and reinforce the conclusion that the true diversity of the freshwater fish of the Neotropical regions has been underestimated.

Keywords: *Ancistrus* genus, cytotaxonomy, species delimitation, candidate species, chromosomal evolution

INTRODUCTION

The inclusion of genetic data in studies of taxonomy and evolution has had a profound influence on our understanding of the unique diversity of species in the Neotropical region (Pereira et al., 2011, 2013). The combination of approaches has altered our perceptions of the region's biological diversity and contributed to an increase in the rate of discovery of new species, especially in cryptic lineages. In this context, the delimitation of species using DNA sequences can be highly efficient, enabling the more systematic documentation of species diversity (e.g., Yang and Rannala, 2010; Ence and Carstens, 2011; Fujisawa and Barraclough, 2013). For example, the Generalized Mixed Yule Coalescent (GMYC) method has been designed to delimit potential lineages using information on a single locus, that is, this method considers that the mutations arising in one species cannot spread readily into another species (Ahearn and Templeton, 1989; Barraclough et al., 2003; De Queiroz, 2007). The premise of the GMYC method is that independent evolution leads to the emergence of distinct genetic clusters, separated by longer internal branches, optimizing the set of nodes that defines the transition between inter- and intra-specific processes (Barraclough et al., 2003).

Fish are an excellent candidate group for the application of integrative taxonomical methods, which combine different lines of evidence (DNA sequences, chromosomal data, and morphological features) to define taxonomic status. Most surveys of Neotropical freshwater fish have focused on major river basins, even though a large proportion of the total diversity comprises small species, found in minor rivers and streams. These species are often highly endemic and occupy a wide variety of microhabitats, providing enormous potential for diversification (Viana et al., 2013). Castro (1999) referred to the identification of the cryptic diversity of small freshwater fish as a major challenge for Neotropical Ichthyology. The reliable definition of species and their ranges is also essential to conservation strategies (Angulo and Icochea, 2010). In this context, the rivers and streams of the Paraná River basin provide an interesting study area for the evaluation of the degree to which the taxonomic diversity of these environments has been underestimated.

The Ancistrini is composed of 29 genera, with a total of 217 recognized species (Fisch-Muller, 2003). With 69 species, *Ancistrus* Kner, 1854 is one of the most diverse Ancistrini groups, the second richest in species of the Loricariidae (Ferraris, 2007; Bifi et al., 2009; Froese and Pauly, 2017). Cytogenetic data on *Ancistrus* are still scarce, and restricted to species found in the basins of the Paraguay River in Mato Grosso, and the Amazon, in Manaus (de Oliveira et al., 2009; Mariotto et al., 2011, 2013; Favarato et al., 2016; Prizon et al., 2016). While the cytogenetics of *Ancistrus* species from other river basins are still unknown, considerable variability has been found in this genus, with diploid (2n) numbers of 34, 38, 40, 42, 44, 48, 50, and 54 chromosomes. Surveys of the upper Paraná River have revealed the presence of a single species of *Ancistrus* in this basin, identified as *Ancistrus cirrhosus* (Langeani et al., 2007). However, the considerable variation in the morphology and coloration observed in the specimens collected in the tributaries of the Paraná River hampers the reliable identification of the *Ancistrus* species found in this region. Thus, given the karyotypic diversity of *Ancistrus* and the assumption that cryptic diversity exists in the rivers and streams of the Upper Paraná River basin, we combine chromosomal data and DNA sequences to evaluate the taxonomic status of these populations.

MATERIALS AND METHODS

Biological Samples

A total of 144 *Ancistrus* specimens were collected in ten rivers of the Paraná River basin (**Table 1**, **Figure 1**). Specimen collection was authorized by the Brazilian Environment Ministry through its Biodiversity Information and Authorization System (SISBIO), under license number 36575-1. The protocols used in this study were submitted to the Ethics Committee on the use of animals in research (CEUA) of the Universidade Estadual de Maringá (UEM) and approved under case number 013/2009. Voucher specimens were deposited in the ichthyological collection of the Limnology, Ichthyology and Aquaculture Research Center (Nupélia) at Universidade Estadual de Maringá, Paraná, Brazil. The catalog numbers are provided in **Table 1**.

TABLE 1 | Details of the *Ancistrus* populations and specimens sampled in the study area of the upper Paraná basin.

Code	Molecular sample	Cytogenetic sample	River	Locality/State or Province/Country	Geographical coordinates	NUP
L1	1♂+1♀	13♂+36♀	Mourão	Campo Mourão/Paraná/Brazil	25°04′46″S, 53°54′45″W	11,993
L2	2♂	7♂+4♀	19 Stream	Paraiso do Norte/Paraná/Brazil	52°38′17″S, 23°16′08″W	13,646
L3	2♂+2♀	2♂+2♀	Keller	Marialva/Paraná/Brazil	23°38′48″S, 51°52′51″W	18,794
L4	1	–	Patos	Prudentópolis/Paraná/Brazil	25°09′59″S, 50°56′29″W	15,537
L5	1	–	São João	Prudentópolis/ Paraná/Brazil	25°05′10″S, 51°00′11″W	15,812
L6	1♂+1 Immature	2♂+2 Immature	São Francisco Verdadeiro	Toledo/Paraná/Brazil	24°46′50″S, 53°43′00″W	15,146
L7	1	–	Arroyo Iguaçu	Marechal Cândido Rondon/ Paraná/Brazil	24°25′18″S, 54°01′09″W	15,250
L8	4	2♂+10♀	Arroyo San Juan	Misiones/Posadas/Argentina	27°22.623′S, 55°53.571′W	18,795
L9	1♂+1 Immature	10♀+8♂+3 Immature	Ocoí	Medianeira/Paraná/Brazil	25°15′12″S, 54°01′55″W	4,729
L10	1♂+1♀	10♂+30♀	São Francisco Falso	Vera Cruz do Oeste/Paraná/Brazil	25°04′46″S, 53°54′45″W	15,145

NUP, catalog number of the voucher specimen in the Nupélia collection; ♂, male; ♀, female.

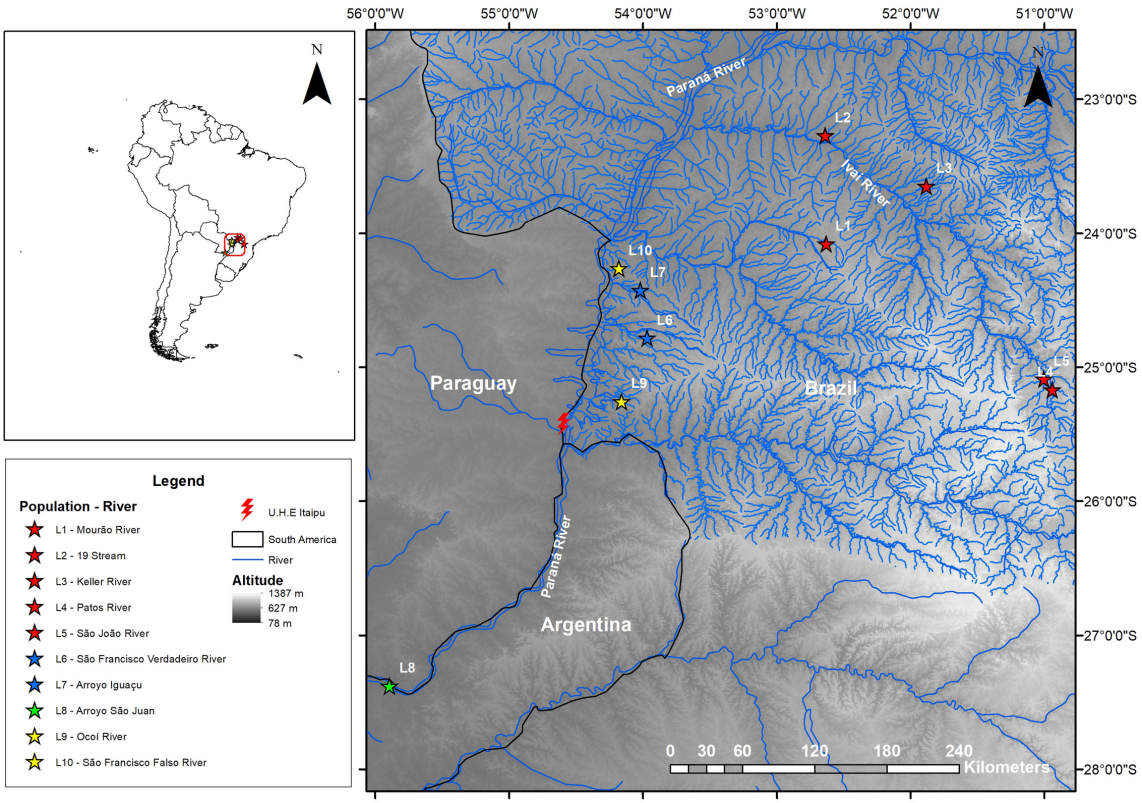


FIGURE 1 | Topographic map showing the sites for the collection of specimens from the *Ancistrus* populations sampled in the Paraná River basin, in the present study. Each color represents a different clade recovered in our phylogenetic inferences (see **Figure 2**): red star, populations in clade I; blue star, populations in clade II; green star, populations in clade IV; and yellow star, populations in clade V. L1, Mourão River; L2, 19 Stream; L3, Keller River; L4, Patos River; L5, São João River; L6, São Francisco Verdadeiro River; L7, Arroyo Iguaçu; L8, *Ancistrus cirrhosus*; L9, Ocoí River; L10, São Francisco Falso River.

Isolation, Amplification, and Sequencing of the DNA

Genomic DNA was extracted from the liver, muscle tissue, or a cell suspension of a subset of the sample (**Table 1**) using the TNES method, as applied by Bruschi et al. (2012). A fragment of the mitochondrial cytochrome C oxidase subunit I (COI) gene was amplified by polymerase chain reaction (PCR) using the primers: FishF1 (5′-TCAACCAACCACAAAGACATTGGCAC-3′), FishR1 (5′-TAGACTTCTGGGTGGCCAAAGAATCA-3′) (Ward et al., 2005). The solution for the amplification reaction included 20 ng/μl of the DNA template, 7 pmol of the forward and reverse primers, 10 mM of dNTPs, 1 U *Taq* DNA Polymerase, 1.5 mM $MgCl_2$, and 1x PCR buffer (200 mM Tris, pH 8.4, 500 mM KCL). The amplification protocol was 5 min−94°C/(94°C/30 s−60°C/1 min−72°C/30 s) 35 cycles/10 min−72°C. The amplified products were purified using Exonuclease I (10 units) and SAP (1 unit), incubated for 45 min 37°C, followed by denaturation at 85°C for 10 min (Applied Biosystems, Santa Clara, CA, USA), as recommended by the manufacturer. The samples were then used directly as templates for sequencing in an automatic ABI/Prism DNA sequencer (Applied Biosystems, Foster City, CA, USA) with the BigDye Terminator kit (Applied Biosystems, Foster City, CA, USA), as recommended by the manufacturer. The DNA samples were sequenced bidirectionally and were edited in Bioedit version 7.2.5 (http://www.mbio.ncsu.edu/bioedit/page2.html) (Hall, 1999).

Phylogenetic Inferences and the Delimitation of Species

The phylogenetic relationships among the populations were inferred from the matrix of the 459-bp sequence of the COI gene. The dataset was complemented with 35 sequences of *Ancistrus* and one sequence of the sister group *Lasiancistrus* available in GenBank (Supplementary File 1). The outgroup was *Pseudolithoxus* sp., which was chosen based on the arrangement reported by Lujan et al. (2015). The sequence was aligned using Clustal W in BioEdit, version 7.2.5.0 (Thompson et al., 1994). The initial alignments were checked visually and adjusted wherever necessary. The dataset was used for phylogenetic reconstruction by Bayesian inference (BI) and the Maximum Parsimony (MP) approach.

Bayesian Inference (BI) method was applied to the dataset, which were divided into three partitions according to codon position for *mit*-COI. The best model of nucleotide evolution for each nucleotide partition was determined using Akaike

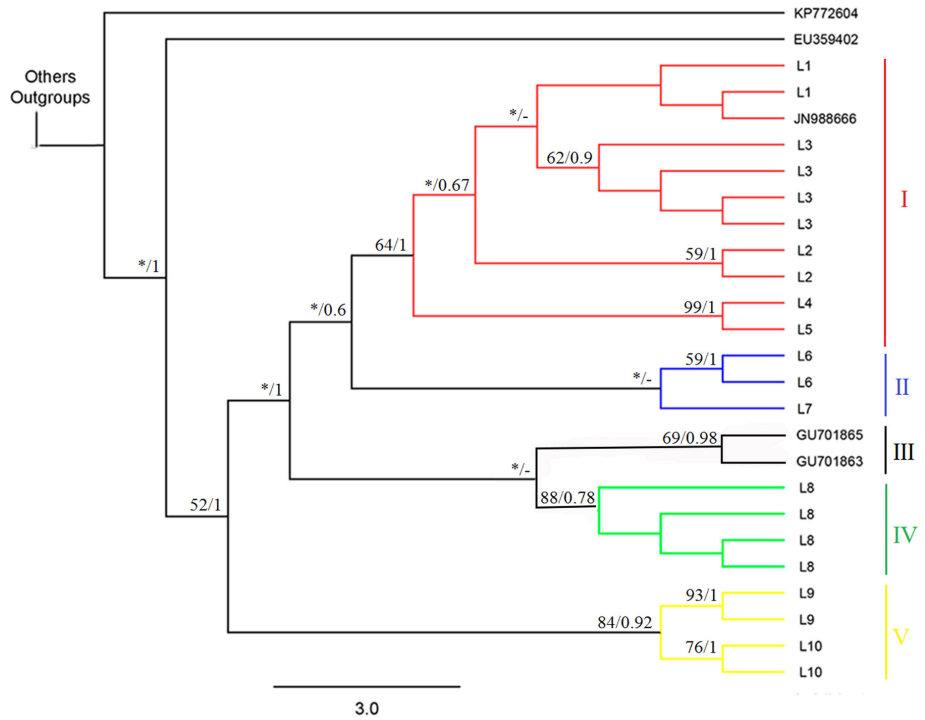


FIGURE 2 | Strict consensus cladogram produced by the Maximum Parsimony analysis showing the intrageneric relationships of the *Ancistrus* species from the Paraná basin, based on a 459-bp sequence of the COI gene. The numbers above the branches show the bootstrap support and Bayesian posterior probabilities, respectively. Each color represents one of the five evolutionary lineages recovered in the analysis. The terminal species with alphanumerical identifiers were obtained from GenBank (Supplementary File 1). The asterisks indicate nodes with bootstrap values of <50. The trace indicates that the node has not been recovered by Bayesian inference. Codes: L1, Mourão River; L2, 19 Stream; L3, Keller River; L4, Patos River; L5, São João River; L6, São Francisco Verdadeiro River; L7, Arroyo Iguaçu; L8, *Ancistrus cirrhosus*; L9, Ocoí River; L10, São Francisco Falso River.

Information Criterion (AIC) with the software jModelTest v2.1.6 (Guindon and Gascuel, 2003; Darriba et al., 2012). The BI was performed with the software Mr. Bayes 3.2.6 (Ronquist and Huelsenbeck, 2003), as available in the CIPRES Science Gateway 3.1 (Miller et al., 2010). BI was implemented using two independent runs, each starting from random trees, with four simultaneous independent chains, and performed 10,000,000 generations, keeping one tree every 1,000th generation. Of all trees sampled, 20% were discarded as burn-in and checked by the convergence criterion (frequencies of average standard deviation of split <0.01) with Tracer v.1.6 (Rambaut et al., 2014), while the remaining were used to reconstruct a 50% majority-rule consensus tree and to estimate Bayesian posterior probabilities (BPP) of the branches. A node was considered to be strongly supported if it had a BPP $\geq$ 0.95, while moderate support was considered when BPP $\geq$ 0.9.

The MP analysis was implemented in TNT v1.1 (Goloboff et al., 2003) using a heuristic search method with tree bisection-reconnection (TBR) swapping and 100 random additional replicates. The bootstrap values of the branches inferred in this analysis were calculated with 1,000 non-parametric pseudoreplicates.

The genetic distances among and within species were calculated using the Kimura-2-Parameter (K2P) and *p*-distance model (Kimura, 1980) implemented in MEGA v 6.0. A neighbor-joining (NJ) tree of K2P distances was created to provide a graphic representation of the patterning of divergence between species with the software MEGA v 7.0 (Kumar et al., 2016). We also applied the General Mixed Yule-coalescent (GMYC) method to delineate species using single-locus sequence data. The GMYC requires a fully resolved and ultrametric tree as input for the analysis and combines a coalescence model of intraspecific branching with a Yule model for interspecific branching to estimate species boundaries and provide statistical confidence intervals to evaluate the sequences of the clusters recovered. Ultrametric trees were constructed by a BI tree in BEAST2 2.4.0 (Drummond et al., 2006; Drummond and Rambaut, 2007). We conducted three independent runs using different priors, that is, the Yule, relaxed clock, and constant coalescent models. An ultrametric gene tree was obtained for each prior. An XML file was produced using the BEAUti2 v2.4.5 interface with the following settings: GTR+G+I substitution model, previously inferred by MrMODELTEST (Nylander et al., 2004), empirical base frequencies, four gamma categories, all codon positions partitioned with unlinked base frequencies and substitution rates. The MCMC chain was 10 million generations long, and was logged every 1,000 generations. The Estimated Sample Sizes (ESS) and trace files of the runs were

TABLE 2 | Uncorrected pairwise distances between the mitochondrial COI sequences of the *Ancistrus* populations from the Paraná basin.

Populations*	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	6 (%)	7 (%)	8 (%)	9 (%)	10 (%)	11 (%)	12 (%)
1. *Ancistrus* sp. L1												
2. *A.cirrhosus* JN988666	0.5											
3. *Ancistrus* sp. L2	0.5	0.4										
4. *Ancistrus* sp. L3	0.5	0.4	0.4									
5. *Ancistrus* sp. L4	1.6	1.8	1.8	1.8								
6. *Ancistrus* sp. L5	1.6	1.8	1.8	1.8	0.0							
7. *Ancistrus* sp. L6	1.8	2.0	1.8	2.0	2.7	2.7						
8. *Ancistrus* sp. L7	2.2	2.5	2.5	2.5	3.3	3.3	2.0					
9. *Ancistrus cirrhosus* L8	2.5	2.7	2.7	2.7	2.7	2.7	2.2	3.1				
10. *A.cirrhosus* GU701865	2.2	2.4	2.4	2.4	2.7	2.7	1.8	2.7	2.2			
11. *A.cirrhosus* GU701863	2.0	2.2	2.2	2.2	2.9	2.9	2.0	2.9	2.4	1.3		
12. *Ancistrus* sp. L9	3.2	3.4	3.4	3.0	3.4	3.4	2.8	3.7	3.2	2.6	2.8	
13. *Ancistrus* sp. L10	2.7	2.9	2.9	2.9	3.1	3.1	2.4	3.3	2.5	2.4	2.5	1.2

**Codes: L1, Mourão River; L2, 19 Stream; L3, Keller River; L4, Patos River; L5, São João River; L6, São Francisco Verdadeiro River; L7, Arroyo Iguaçu; L8, Ancistrus cirrhosus; L9, Ocoí River; L10, São Francisco Falso River. JN988666, GU701865, and GU701865: GenBank Access Number.*

evaluated in Tracer v1.6. The resulting logs were analyzed in TREEANNOTATOR 2.4.4, with 25% burn-in, maximum clade credibility trees with a 0.5 posterior probability limit, and node heights of the target tree. The splits package (https://r-forge.r-project.org/R/?group_id=333) in R was used for the GMYC calculations, using the single-threshold strategy and default scaling parameters.

Cytogenetic Analysis

All specimens were anesthetized and euthanized by an overdose of clove oil (Griffiths, 2000). Mitotic chromosomes were obtained from kidney cells according to Bertollo et al. (1978). The AgNORs were revealed by the silver nitrate impregnation technique (Howell and Black, 1980). The regions of heterochromatin were determined by the C-banding technique (Sumner, 1972) and stained with propidium iodide according to the method of Lui et al. (2012). Physical mapping of the 5S rDNA and 18S rDNA sequences was carried out by fluorescence *in situ* hybridization (FISH) according to Pinkel et al. (1986), with probes obtained from *Leporinus elongatus* Valenciennes, 1850 (Martins and Galetti, 1999) and *Prochilodus argenteus* Spix et Agassiz, 1829 (Hatanaka and Galetti, 2004). We also isolated and cloned the rDNA 5S gene of the *Ancistrus* sample from Keller River using a more specific DNA probe for this gene. The genomic DNA extracted for the molecular analyses was used as the template for the amplification reaction, using the primers 5S-A (5′-TACGCCCGATCTCGTCCGATC-3′) and 5S-B (5′-CAGGCTGGTATGGCCGTAAGC-3′) (Pendas et al., 1994). The products of this amplification were isolated in the 1.5% agarose gel, then purified with an EasyPure® Quick Gel Extraction kit. These sequences were inserted into a linearized cloning vector (pJET1.2/blunt) by the CloneJET PCR Cloning kit (Thermo Scientific) and cloned in *Escherichia coli*. Ten clones with the insert containing the 5S rDNA sequence were selected for sequencing. The 5S rDNA nucleotide sequences were edited using BioEdit software and compared with sequences from GenBank database (www.ncbi.nlm.nih.gov).

Hybridization was conducted under high stringency conditions (77%), and the probes were labeled by nick translation with digoxigenin-11-dUTP (5S rDNA) and biotin-16-dUTP (18S rDNA). The 5S rDNA probes obtained from the recombinant plasmids were labeled by PCR using the fluorochrome digoxigenin-11-dUTP. The solution for labeling reaction included 20 ng/μl of the DNA template, 7 pmol of the forward and reverse primers, 4 mM of dNTPs, 1 mM of dig-11-dUTP, 1 U *Taq* DNA Polymerase, 1.5 mM $MgCl_2$, and 1x PCR buffer (200 mM Tris, pH 8.4, 500 mM KCL). The hybridization signals were detected using anti-digoxigenin-rhodamine for the 5S rDNA probe and avidin-FITC (fluorescein isothiocyanate) for the 18S rDNA probe. The chromosomes were counterstained with DAPI. Double staining was carried out with chromomycin A3 (CMA3) and DAPI, according to Schweizer (1976). The metaphases were photographed using an epifluorescence microscope and adjusted for best contrast and brightness using the Adobe Photoshop CS6 software.

RESULTS

Phylogenetic Inferences and Delimitation of Species

The phylogenetic reconstructions based on the MP and BI approaches produced the basic topology of the dataset (**Figures 2**, **3**). Topology inferred from the Bayesian analysis performed with the software Mr. Bayes 3.2.6. and from Neighbor-Joining were presented in the supplementary material (**Figures S1**, **S2**, respectively). These analyses recovered five clades from the *Ancistrus* populations of the Paraná basin. The first clade (clade I) comprises five populations: *Ancistrus* sp. "Mourão River" (L1) + *Ancistrus* sp. "19 Stream" (L2) + *Ancistrus* sp. "Keller River" (L3) + *Ancistrus* sp. "Patos River" (L4) + *Ancistrus* sp. "João River" (L5). The genetic distance

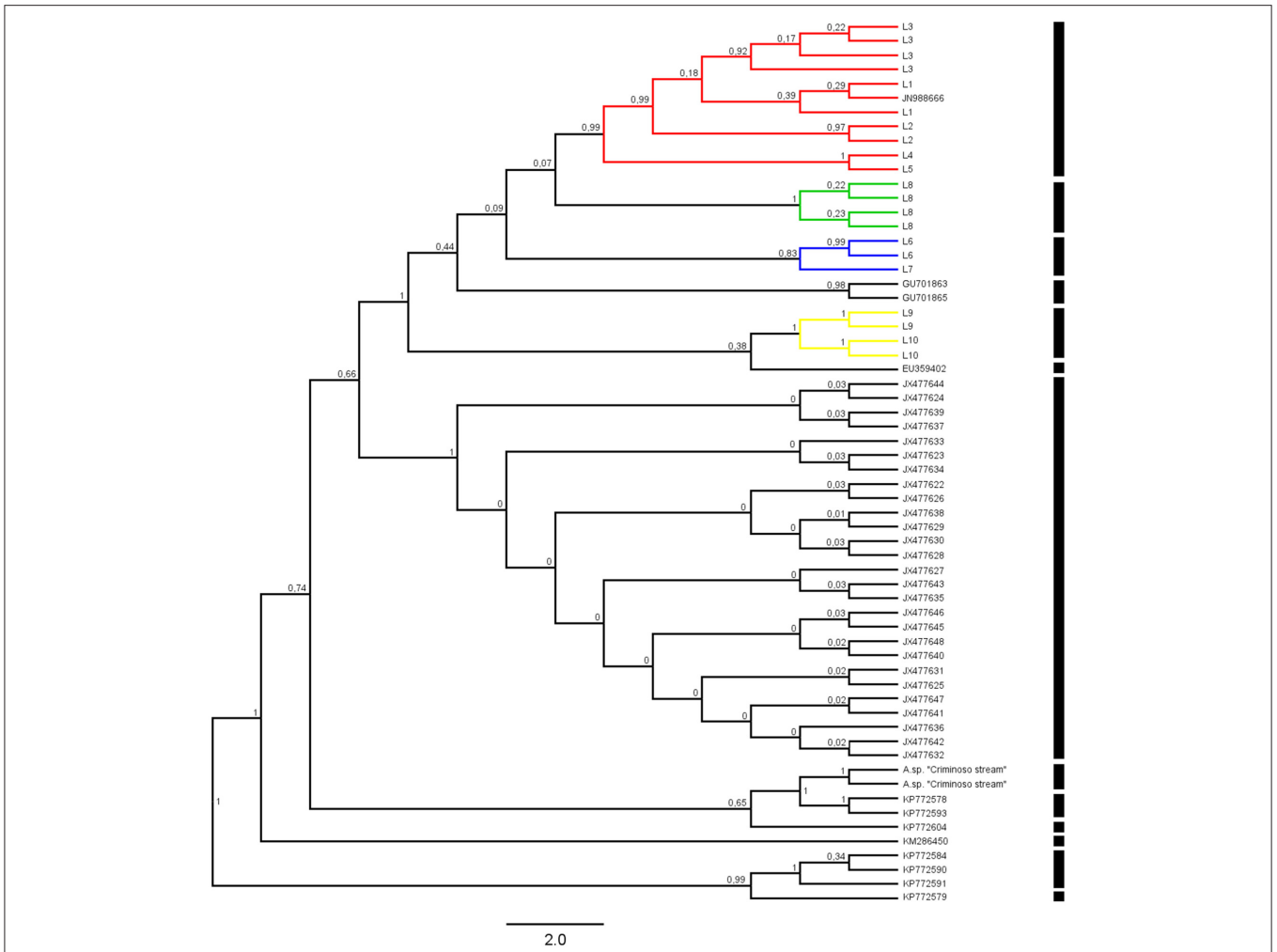


FIGURE 3 | Bayesian consensus tree for the *Ancistrus* species of the Paraná River basin produced in BEAST2 2.4.0 and obtained from 10 million generations. Posterior probabilities are shown above the branches. Each color represents one of the five evolutionary lineages recovered. The results of the GMYC method using the Yule model of species delimitation are shown as black bars to the right of the phylogeny. Species access numbers are specified in the supplementary material (Supplementary File 1). Codes: L1, Mourão River; L2, 19 Stream; L3, Keller River; L4, Patos River; L5, São João River; L6, São Francisco Verdadeiro River; L7, Arroyo Iguaçu; L8, *Ancistrus cirrhosus*; L9, Ocoí River; L10, São Francisco Falso River.

analysis returned low uncorrected *P*-distances among these populations, ranging from 0.0 to 1.8% (**Table 2**). The second clade (clade II) included the *Ancistrus* populations from São Francisco Verdadeiro River (L6) and *Ancistrus* sp. from Arroyo Iguaçu (L7), separated by a genetic distance of 2.0% (**Table 2**). Two sequences deposited in Genbank as *A.* "*cirrhosus*" included in our dataset were recovered in the clade III. Clade IV included only the population from Arroyo San Juan (L8), Argentina, and was considered to represent the nominal *A. cirrhosus* due to its proximity to the type-locality of this species. Clade V consisted of the *Ancistrus* sp. "Ocoí River" (L9) + *Ancistrus* sp. "São Francisco Falso River" (L10) populations. The uncorrected *P*-distance between these populations (1.2%) was also relatively low (**Table 2**). In all cases, by contrast, the genetic distances (uncorrected *P*-values) between clades were at least 3%, being 3% between clades I–II, I–III, II–III, II–IV, III–IV; 4% between clades I–IV, II–V, III–V, IV–V and finally a genetic distance of 5% between clades I–V.

The ultrametric trees obtained using the Yule and Constant Coalescent priors were congruent in the GMYC analysis, recognizing 13 separate entities and identifying nine clusters in our dataset (including outgroup sequences). Considering only the *Ancistrus* populations, the focus of the present study, the sequences were grouped into five well-supported clusters (**Figure 3**), which corresponded exactly to the major clades recovered in our phylogenetic inferences (BI and MP). The Yule and Constant Coalescent priors of the branching rates indicated that the likelihood of the null model (i.e., that all the sequences belong to the same species) was 412.5381 and 421.5504, respectively, with a likelihood of 421.9409 and 430.6968, respectively, for the GMYC model (i.e., the existence of distinct species). The difference is highly significant, indicating

TABLE 3 | Cytogenetic parameters of the *Ancistrus* populations sampled in the different rivers of the Paraná basin.

Ancistrus population	Clade	Karyotype formula	Ag-NOR	rDNA sites	
				18S	5S
Mourão River (L1)	I	12m+18sm+12s+8a	12 (sm)	12 (sm)	1 (m), 14 (sm), 19 (st)**, 20 (st)
Stream 19 (L2)	I	12m+18sm+12s+8a-♀	12 (sm)	12 (sm)*	1 (m), 12 (sm)*, 15 (sm), 20 (st), 25(a)
		11m+18sm+13s+8a-♂			
Keller River (L3)	I	12m+18sm+12s+8a-♀	12 (sm)	12 (sm)*	1 (m), 12 (sm)*, 15 (sm), 20 (st), 22(a), 25(a)
		11m+18sm+13s+8a-♂			
São Francisco Verdadeiro River (L6)	II	14m+16sm+14st+6a	18 (st)	18 (st)*	1 (m), 15 (sm), 18 (st)*, 21 (st)
Arroyo San Juan (L8)	III	10m+14sm+12st+14a	17 (st)	17 (st)	1 (m), 18 (st), 23(a)
Ocoí River (L9)	IV	10m+18sm+16st+6a	18 (st)	18 (st)*	18 (st)*, 21 (st), 22 (st)
São Francisco Falso River (L10)	IV	10m+18sm+16st+6a	18 (st)	18 (st)*	11 (sm), 14 (sm)18 (st)*, 19 (st)

*Synteny between the 18S and 5S rDNA sites; **in only one homolog; m, metacentric; sm, submetacentric; st, subtelocentric; a, acrocentric; ♂, male; ♀, female.

the presence of more than one species in our sample. The analysis based on the Relaxed clock priors presented low scores in Tracer and did not produce a reliable interpretation of the phylogenetic relationships among the populations, and was not considered in the species delimitation tests.

Cytogenetic Analysis

Cytogenetic data were obtained for the populations from the Mourão, Keller, São Francisco Verdadeiro, São Francisco Falso, and Ocoi Rivers, and 19 Stream (Brazil), and Arroyo San Juan (Argentina). All the specimens analyzed presented a diploid number of 2n = 50 chromosomes, although four distinct karyotype formulae were detected (**Table 3**, **Figure 4**), together with variation in the location of the 18S and 5S rDNA sites (**Figure 5**). Interestingly, these formulae corresponded strongly with the major clades recovered in our genetic analysis, with exception of the clade III that have no karyotype available. A chromosome heteromorphism found in all the males of Stream 19 (L2) and Keller River (L3) populations were consistent with an XX/XY system. The X chromosome is a large metacentric (second pair) while the Y chromosome is a small subtelocentric (**Figure 4**).

The analysis of the heterochromatin revealed C-positive blocks at the pericentromeric and subterminal positions in a number of different chromosomal pairs of the populations allocated to clade I, with a conspicuous block coinciding with the NOR site of pair 12 in all karyotypes (**Figures 6A–C**). The X chromosome (pair 2, metacentric) from 19 Stream had a C-positive block in the pericentromeric region, while we detected heterochromatin blocks in the pericentromeric, interstitial, and subterminal regions of the X chromosome from Keller River. The Y chromosome of the Keller River population had a weak heterochromatin block in the subterminal region of the long arm, which was not found in the Y chromosome from 19 Stream (**Figures 6B,C**). In clades II, IV, and V, considerable variation was found in the amount and distribution of constitutive heterochromatin, which was concentrated primarily in the pericentromeric and subterminal regions of the chromosomes (**Figures 6D–G**). In the population from the São Francisco Falso River (clade V), in addition, we detected an interstitial heterochromatin block in pair 17, as well as a much larger amount of heterochromatin distributed throughout the chromosomes in comparison with the other populations (**Figure 6G**). Conspicuous blocks of heterochromatin were also found in the nucleolar pairs of all the populations analyzed in clades II and IV (pair 18) (**Figures 6D,E**).

The base-specific fluorochrome staining enhanced the distinctive composition of some heterochromatin blocks. In particular, staining with Chromomycin A3 revealed a richness of G and C in the subterminal and pericentromeric heterochromatin blocks of all the populations analyzed, and provides additional features for the comparative analysis. As expected, conspicuous C-positive areas were detected in the NOR-bearing chromosomes (pair 12, 17, and 18) (**Figure S3**).

DISCUSSION

Molecular Phylogenetic Inferences Reveal a Number of Distinct Lineages in the Paraná Basin

The phylogenetic reconstructions and cytogenetic analyses presented in this study both detected the presence of five major clades among the *Ancistrus* populations surveyed, pointing to the existence of at least five lineages within the Paraná basin in Brazil previously undetected. Historically, all *Ancistrus* from the Paraná River basin were assigned to a single species, *A. cirrhosus*, which was described by Valenciennes (1836) from specimens collected in the Province of Misiones and Buenos Aires (Argentina), although the diagnostic traits are weakly defined (Langeani et al., 2007). Despite the morphological variations and wide geographic distribution, the *Ancistrus* populations of the Paraná basin have been identified invariably as *A. cirrhosus*. This fact has intrigued the ichthyologists with regard to the taxonomic status of many specimens, using only morphological and meristic characters. Therefore, the data obtained in the present study recovered at least five independent lineages (clades I–V) in Paraná basin. Given the proximity of the Arroyo San Juan to the type-locality of *A. cirrhosus,* this population (clade IV) was identified as nominal species *A. cirrhosus*. In this context, the other four clades (I, II, III, and V) can be categorized as "candidate species" in the terminology of Vieites et al. (2009). This interpretation is also

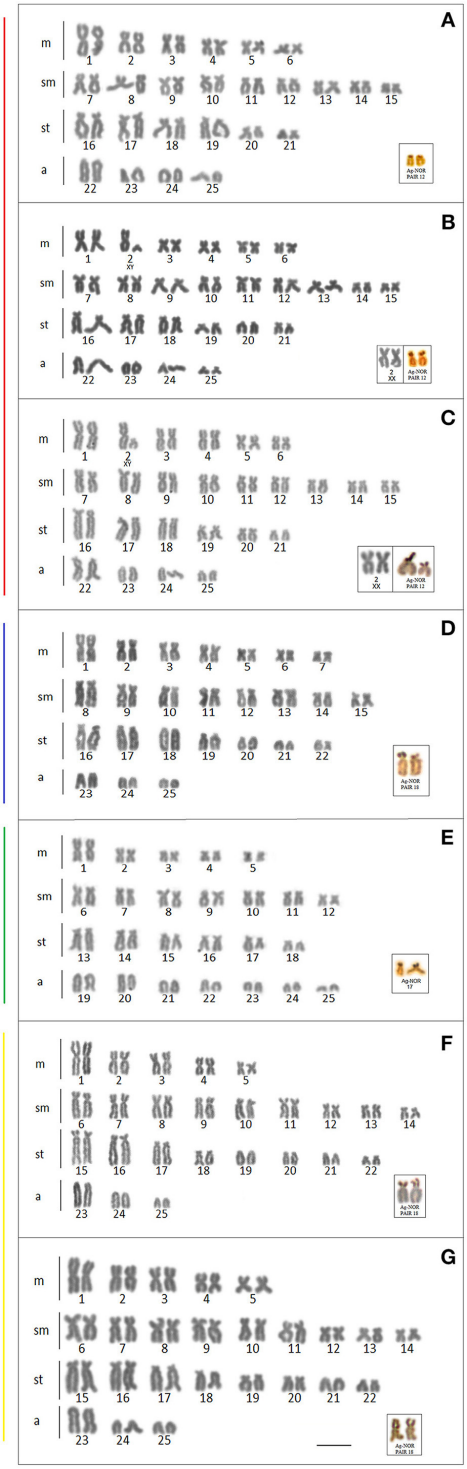


FIGURE 4 | Karyotype of the *Ancistrus* populations sampled in different rivers of the Paraná basin, stained with Giemsa. The configuration of the silver nitrate-stained nucleolar organizing regions (Ag-NORs) are shown in the box. Each color of the side bars represents one of the evolutionary lineages recovered in the analysis, according to **Figure 2**. **(A)** L1: Mourão River; **(B)** L2: 19 Stream; **(C)** L3: Keller River; **(D)** L6: São Francisco Verdadeiro River; **(E)** L8: *Ancistrus cirrhosus*; **(F)** L9: Ocoí River; **(G)** L10: São Francisco Falso River. Bar = 10 μm.

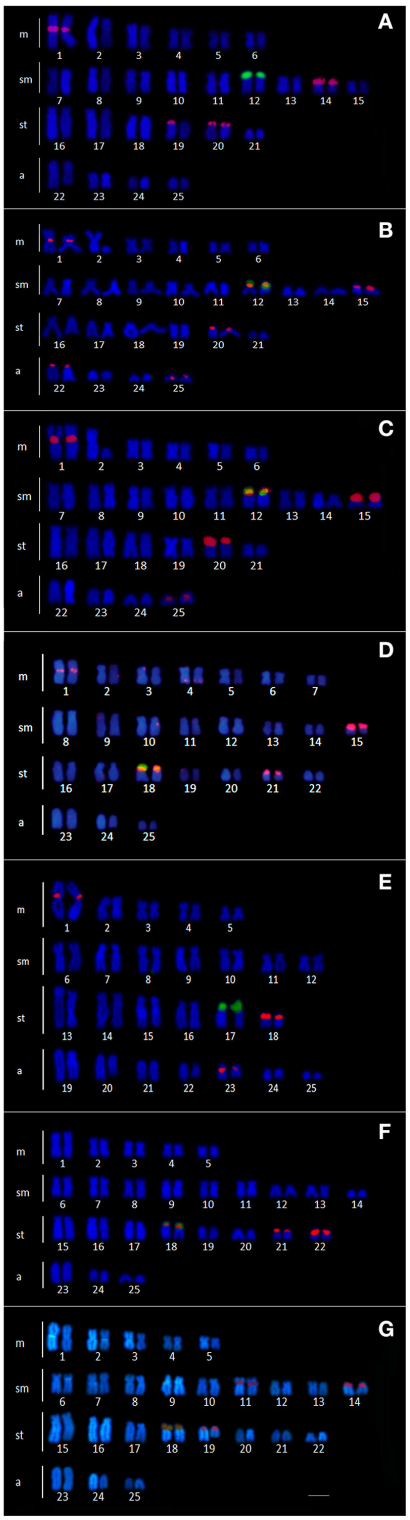


FIGURE 5 | Chromosomes of the *Ancistrus* populations of species after dual color-FISH showing the 5S rDNA (red) and 18S rDNA (green) sites. Each color of the side bars represents one of the evolutionary lineages recovered in the analysis, according to **Figure 2**. **(A)** L1: Mourão River; **(B)** L2: 19 Stream; **(C)** L3: Keller River; **(D)** L6: São Francisco Verdadeiro River; **(E)** L8: *Ancistrus cirrhosus*; **(F)** L9: Ocoí River; **(G)** L10: São Francisco Falso River. Bar = 10 μm.

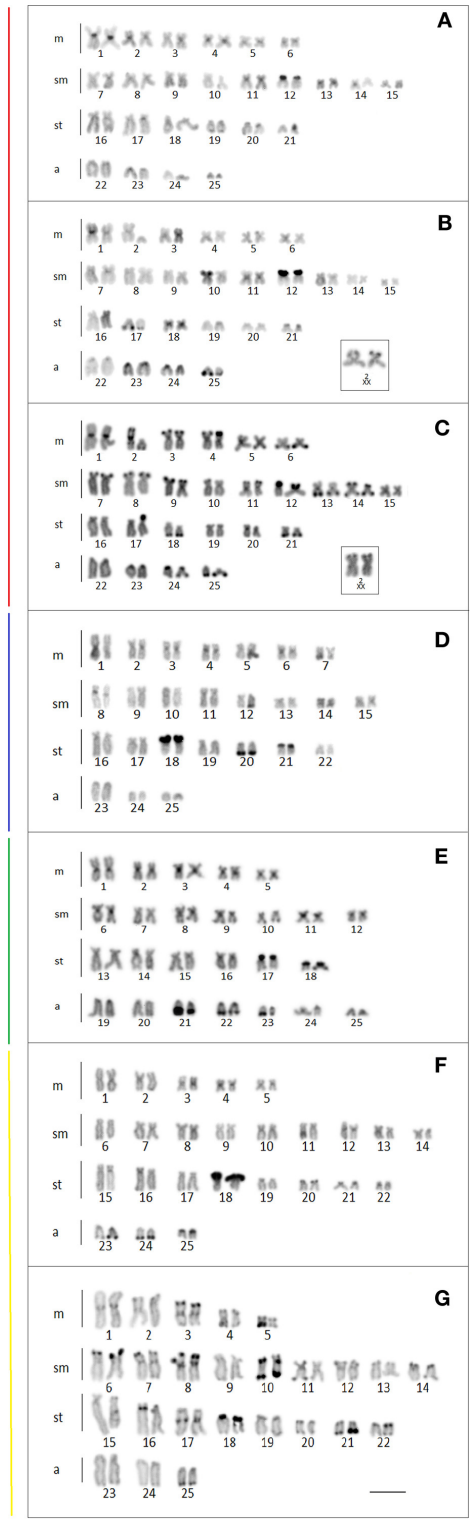


FIGURE 6 | Karyotypes of the *Ancistrus* populations showing the distribution of the heterochromatin after C-banding. Each color of the side bars represents one of the evolutionary lineages recovered in the analysis, according to **Figure 2**. **(A)** L1, Mourão River; **(B)** L2, 19 Stream; **(C)** L3, Keller River; **(D)** L6, São Francisco Verdadeiro River; **(E)** L8, *Ancistrus cirrhosus*; **(F)** L9, Ocoí River; **(G)** L10, São Francisco Falso River. Bar = 10 μm.

supported by the results of the GMYC method, and the genetic distance detected among populations, considering a 2% threshold for interspecific differentiation.

The GMYC method has become one of the most popular tools for the delimitation of species based on single-locus data, and has been applied to the analysis of a number of poorly-known groups of organism (Barraclough et al., 2009; Monaghan et al., 2009; Marshall et al., 2011; Vuataz et al., 2011; Roxo et al., 2015). The GMYC method uses an ultrametric tree derived from the sequences to identify shifts in the branching rate from the Yule model (species) to the coalescent (population) process, with the algorithm computing the probability of splits between lineages in relation to speciation rates, thus identifying a cutoff value, at which species and populations split from one another (Powell, 2012). This approach was highly effective in the present study, allowing us to delimit five groups, which four corresponded exactly with the chromosomal data (unfortunately, no chromosomal data available to clade III), reinforcing the hypothesis that four "candidate species" exist in the Paraná River basin, besides *A. cirrhosus* as previously reported.

Armbruster (2004) found a sister group relationship between the Ancistrini and Pterygoplichthini tribes, which both belong to the subfamily Hypostominae. *Ancistrus* is the most species-rich genus of the tribe Ancistrini, although phylogenetic analyses of this genus are still scarce. Lujan et al. (2015) adopted a broad approach to the phylogenetic relationships of the Loricariidae, a Neotropical catfish family, but included few *Ancistrus* species, which reinforces the need for further research. Studies in systematics based on osteological (Schaefer, 1987) and molecular (Montoya-Burgos et al., 1998) data show that *Ancistrus* constitutes a monophyletic group of species. The lack of any robust phylogenetic tree for *Ancistrus* still limits our understanding of its intrageneric relationships, and more detailed analyses, with a more representative dataset, are needed to provide a more comprehensive understanding of the phylogenetic relationships of this genus.

The Chromosomal Data Reinforce the Hypothesis of Complete Lineage Divergence

Cytogenetic studies, together with molecular and biochemical analyses, may be useful for the identification of cryptic species (Nakayama et al., 2001; Milhomen et al., 2007). In the present study, in fact, the inclusion of a cytogenetic approach was crucial to the recognition of the "candidate species," given that the distinct chromosomal formulae found in the four lineages (clades I, II, IV, and V) emphasizes their reciprocal monophyly and the lack of gene flow between them. The results obtained also complement the available cytogenetic data to the *Ancistrus* species, since these studies are currently restricted to the taxa found in the Paraguay (Mato Grosso) and Amazon (Manaus) basins (de Oliveira et al., 2007, 2008, 2009; Mariotto et al., 2011, 2013; Favarato et al., 2016; Prizon et al., 2016).

In the present study, a diploid number of 2n = 50 chromosomes was recorded in all the samples from the Paraná River basin. This number is within the range recorded for the

genus, which vary from 2n = 34 in *Ancistrus cuiabae* to 2n = 54 in *Ancistrus claro* (Mariotto et al., 2009, 2013; Favarato et al., 2016; Prizon et al., 2016). Despite the homogeneity of the diploid number, the karyotype formula varied considerably among populations. The number of acrocentric chromosomes, for example, varied from three to four pairs in the populations of clades I, II, and V, to seven pairs in clade IV (Arroyo San Juan). This points evidences to the occurrence of chromosomal rearrangements, such as translocation and pericentric inversions, which did not affect the diploid number. Unfortunately, while we were able to identify these features, we were unable to trace the pathways of the transformations due to the lack of resolution in the internal topology of the *Ancistrus* lineages recognized here. In this case, further phylogenetic studies, based on a larger set of characters and a multi-locus dataset, may provide more definitive insights into the evolution of this group. Among all the ancistrinis species studied so far, the karyotypic evolution of *Ancistrus* is invariably associated with chromosomal rearrangements, which typically involve variation in the diploid number, given that this number ranges from 34 to 54 in this genus. As Artoni and Bertollo (2001) considered 2n = 54 to be the ancestral diploid number of the Loricariidae, karyotypic evolution in *Ancistrus* appears to have been associated with a reduction in the chromosome number. In fact, de Oliveira et al. (2009) suggested that the karyotypic evolution of this genus was predominantly involves by centric fusions.

In *Ancistrus* the occurrence of heteromorphic sex chromosomes has been well-documented in some species, including simple systems (Mariotto et al., 2004; Alves et al., 2006; Mariotto and Miyazawa, 2006) and multiples (de Oliveira et al., 2007, 2008) which also contributed to the karyotype evolution in the genus (Favarato et al., 2016). One peculiar feature observed in clade I was the heteromorphic sex chromosomes found in the populations of 19 Stream (L2) and the Keller River (L3), which are consistent with an XX/XY system. Surprisingly, however, this feature was not observed in the specimens from the Mourão River (L1), which were included in clade I and separated by low genetic distances from the populations of Stream 19 and the Keller River. The karyotypic formula of the Mourão River population differs from these two other populations only by the absence of the pair of sex chromosomes, and the GYMC analysis also identified these populations as a unique taxonomical unit. A similar result was obtained by Henning et al. (2011), combining cytogenetic and molecular data for different species of *Eigenmannia*, included two populations of *E. virescens* (Mogi-Guaçu and Tietê rivers), whose karyotypes with 2n = 38 differ by the presence of a pair sexual XX/XY (Tietê river population). The acrocentric X chromosome possesses a heterochromatinized distal region (Almeida-Toledo et al., 2001) and according to Henning et al. (2011), both populations (Mogi-Guaçu e Tietê rivers) were considered sister species. Furthermore, these authors concluded that seems likely that suppression of recombination in the homologous pair of acrocentric chromosomes and accumulation of heterochromatin on the X chromosome occurred after a recent geographical separation. In this context, we hypothesized that the heteromorphic sex chromosome found in the *Ancistrus* populations of 19 Stream and the Keller River represent a recent event which may have occurred after the geographical isolation of these populations from that of the Mourão River, together with the behavioral characteristics of these fish, which occupy specific microhabitats, form territories, and do not normally migrate (Power, 1984, 1990; Buck and Sazima, 1995). All these characteristics favor the fixation of chromosomal rearrangements and could be contribute to allopatric speciation on a micro scale (de Oliveira et al., 2009).

The hypothesis of the recent differentiation of the sex chromosome pair in the *Ancistrus* populations of clade I was also supported by the C-banding data. Heterochromatin is widely used in the identification of the sex chromosomes, and the addition or deletion of heterochromatin or the occurrence of a pericentric inversion involving one of the chromosomes have been postulated as important mechanisms in the origin of simple sexual chromosome systems in Neotropical fish (Almeida-Toledo et al., 2000). If we compare the Y chromosome of males from 19 Stream (L2) and Keller River (L3), it is possible to detect a discreet heterochromatin block in the subterminal region of the long arm of the Keller River males, which was not detected in the Y chromosomes from 19 Stream. The presence of the heteromorphic pair in the *Ancistrus* population of 19 Stream and Keller River suggests that chromosomal rearrangements (inversions), the loss of chromosomal material and, in the specific case of the Keller River, the presence of constitutive heterochromatin in the heteromorphic pair, may all be evidence of the recent origin of the Y chromosome, derived from a large metacentric, similar to the X chromosome. Pericentric inversions, followed by a loss of chromosomal material, have been suggest as a mechanism to explain the origin of the ZZ/ZW chromosome system in *Ancistrus* sp. Piagaçu (de Oliveira et al., 2007) and the XX/XY chromosome systems in *Ancistrus* sp. Purus and *Ancistrus* sp. Macoari (de Oliveira et al., 2009), given absence of the heterochromatin blocks in either the X or Y chromosomes.

The NOR mapping provided an excellent marker in the present study, being found in a single chromosome pair in each clade, a condition shared with most other *Ancistrus* species (Medeiros et al., 2016). The variation among clades in the NOR-bearing chromosome may be the result of chromosomal rearrangements occurring during chromosomal evolution. We also recorded synteny between the 18S and 5S rDNA sites in most populations of clades I, II, and V, except for the population from Mourão River (clade I). This synteny of the rDNA may represent the basal condition for the genus (Mariotto et al., 2011), as it is found in *A. claro* (2n = 54), although synteny of the ribosomal sites was not found in the population from Arroyo San Juan (clade IV). The position and distribution of the 5S rDNA sites varied considerably among the four clades, as they do in other *Ancistrus* species, occupying multiple sites in pericentromeric, interstitial, or terminal positions. This variation is considered to be an important reflection of the enormous karyotypic diversity found in the *Ancistrus*, which is seen as evidence of the apomorphic condition of this group (Medeiros et al., 2016). This variability, together with the disjunction of the ribosomal sites caused by rearrangements or mobile genetic elements appears to be a common condition among

Neotropical fish species (de Oliveira et al., 2009). We observed heterochromatin in association with the ribosomal DNA sites (18 and 5S), recurrent characteristic in the karyotypes of Neotropical fishes (Vicari et al., 2003). The presence of heterochromatin, which contains large quantities of transposable elements (Dimitri et al., 2009) may facilitate transposition events, moving ribosomal genes around the genome (Moreira-Filho et al., 1984; Vicari et al., 2008; Gross et al., 2009, 2010). This may be one of the factors responsible for the presence of the multiple 5S ribosomal DNA sites found in the present study.

CONCLUSION

The cytogenetic data available for the genus *Ancistrus* indicate a highly heterogeneous pattern of chromosome evolution, marked by Robertsonian and non-Robertsonian rearrangements. While we do not have an exact understanding of the mechanisms that determine these rearrangements in natural populations, their fixation may either initiate or contribute to the divergence process, with specific implications for the utility of chromosomal characters for phylogenetic inference (Sites and Kent, 1994). As in other fish groups, the sex chromosomes, present in some *Ancistrus* species, may have contributed to high rates of evolution. The inferences obtained in the present study from a combined approach of molecular and cytogenetic analyses further corroborate the taxonomic complexity of this genus. This approach was especially important due to the lack of diagnostic features in the morphology of these fishes. The hidden diversity of the study populations was nevertheless decoded successfully by the combined approach, which allowed us to differentiate five distinct lineages of *Ancistrus*, reinforcing the hypothesis of the presence of at least four candidate species in the upper Paraná River basin, besides of the *A. cirrhosus,* previously described. Finally, our findings reinforce the observation that the true diversity of the freshwater fish of the Neotropical has been underestimated and improve our understanding of a regional diversity within *Ancistrus* genus.

AUTHOR CONTRIBUTIONS

AP: provided chromosomal and molecular data, and drafted the manuscript; DB: designed and coordinated the study of molecular data and helped draft the manuscript; HR and CZ: collected specimens from Paraná state and helped to identify the specimens; AF: collected and processed material of specimens from Arroyo San Juan, Argentina; LB-C, AC, and LB: assisted in the execution and analysis of chromosomal banding; AdBP-C: designed and coordinated the study of cytogenetic data and helped draft the manuscript. All authors have read and approved the final manuscript.

FUNDING

This work was funded by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil) and Fundação Araucária.

ACKNOWLEDGMENTS

We thank the Brazilian agency CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil) for financial support, Maringá State University (UEM) and Limnology, Ichthyology and Aquaculture Research Center (Nupélia) for the logistic support, the collection, and identification of species. We also thank Prof. Dr. Mateus Arduvino Reck and Prof. Dr. Alessandra Valéria de Oliveira for their assistance and contribution in the programs of phylogeny.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2017.00185/full#supplementary-material

Figure S1 | Topology inferred from the Bayesian analysis performed with the software Mr. Bayes 3.2.6. posterior probabilities are shown at each node. Scale bar represents the number of substitutions per site. The terminal species with alphanumerical identifiers were obtained from GenBank (Supplementary File 1). Codes: L1, Mourão River; L2, 19 Stream; L3, Keller River; L4, Patos River; L5, São João River; L6, São Francisco Verdadeiro River; L7, Arroyo Iguaçu; L8, *Ancistrus cirrhosus*; L9, Ocoí River; L10, São Francisco Falso River.

Figure S2 | Chromosomes of the *Ancistrus* populations of species after CMA3 staining. Each color of the side bars represents one of the evolutionary lineages recovered in the analysis, according to **Figure 2**. **(a)** L1, Mourão River; **(b)** L2, 19 Stream; **(c)** L3, Keller River; **(d)** L6, São Francisco Verdadeiro River; **(e)** L8, *Ancistrus cirrhosus*; **(f)** L9, Ocoí River; **(g)** L10, São Francisco Falso River. Bar = 10 μm.

Figure S3 | NJ dendrograma of the *Ancistrus* specimens. Node values = bootstrap test (1,000 pseudo replicas) are shown next to the branches. The terminal species with alphanumerical identifiers were obtained from GenBank (Supplementary File 1). Codes: L1, Mourão River; L2, 19 Stream; L3, Keller River; L4, Patos River; L5, São João River; L6, São Francisco Verdadeiro River; L7, Arroyo Iguaçu; L8, *Ancistrus cirrhosus*; L9, Ocoí River; L10, São Francisco Falso River.

REFERENCES

Ahearn, J. N., Templeton, A. R. (1989). Interspecific hybrids of *Drosophila heteroneura* and D. silvestris I. Courtship success. *Evolution* 43, 347–361. doi: 10.1111/j.1558-5646.1989.tb04232.x

Almeida-Toledo, L. F., Foresti, F., Daniel, M. F., and Toledo-Filho, S. A. (2000). Sex chromosome evolution in fish: the formation of the neo-Y chromosome in *Eigenmannia* (*Gymnotiformes*). *Chromosoma* 109, 197–200. doi: 10.1007/s004120050428

Almeida-Toledo, L. F., Foresti, F., Pequignot, E. V., and Daniel-Silva, M. F. Z. (2001). XX:XY sex chromosome system with X heterocromatinization: an early stage of sex chromosome differentiation in the Neotropic eletric eel *Eigenmannia virescens*. *Cytogenet. Cell Genet. Res.* 95, 73–78. doi: 10.1159/000057020

Alves, Á. L., Oliveira, C., Nirchio, M., Granado, A., and Foresti, F. (2006). Karyotypic relationships among the tribes of Hypostominae (Siluriformes: Loricariidae) with description of XO sex chromosome system in a Neotropical fish species. *Genetica* 128, 1–9. doi: 10.1007/s10709-005-0715-1

Angulo, A., and Icochea, J. (2010). Cryptic species complexes, widespread species and conservation: lessons from Amazonian frogs of the *Leptodactylus marmoratus* group (*Anura: Leptodactylidae*). *Syst. Biodivers.* 8, 357–370. doi: 10.1080/14772000.2010.507264

Armbruster, J. W. (2004). Phylogenetic relationships of the suckermouth armoured catfishes (*Loricariidae*) with emphasis on the *Hypostominae* and the Ancistrinae. *Zool. J. Linn. Soc.* 141, 1–80. doi: 10.1111/j.1096-3642.2004.00109.x

Artoni, R. F., and Bertollo, L. A. (2001). Trends in the karyotype evolution of Loricariidae fish (Siluriformes). *Hereditas* 134, 201–210. doi: 10.1111/j.1601-5223.2001.00201.x

Barraclough, T. G., Birky, C. W., and Burt, A. (2003). Diversification in sexual and asexual organisms. *Evolution* 57, 2166–2172. doi: 10.1111/j.0014-3820.2003.tb00394.x

Barraclough, T. G., Hughes, M., Ashford-Hodges, N., and Fujisawa, T. (2009). Inferring evolutionarily significant units of bacterial diversity from broad environmental surveys of single-locus data. *Biol. Lett.* 5, 425–428. doi: 10.1098/rsbl.2009.0091

Bertollo, L. A. C., Takahashi, C. S., and Moreira-Filho, O. (1978). Cytotaxonomic considerations on *Hoplias lacerdae* (Pisces, *Erythrinidae*). *Rev. Bras. Genet*. 1, 103–120.

Bifi, A., Pavanelli, G. C. S., and Zawadzki, C. H. (2009). Three new species of *Ancistrus* Kner, 1854 (Siluriformes: *Loricariidae*) from the Rio Iguaçu basin, Paraná State, Brazil. *Zootaxa* 2275, 41–59. doi: 10.5281/zenodo.191064

Bruschi, D. P., Busin, C. S., Siqueira, S., and Recco-Pimentel, S. M. (2012). Cytogenetic analysis of two species in the *Phyllomedusa hypochondrialis* group (Anura, Hylidae). *Hereditas* 149, 34–40. doi: 10.1111/j.1601-5223.2010.02236.x

Buck, S., and Sazima, I. (1995). An assemblage of mailed catfishes (Loricariidae) in southeastern Brazil: distribuition, activity, and feeding. *Ichthyol. Explor. Freshw.* 6, 325–332.

Castro, R. M. C. (1999). "Evolução da ictiofauna de riachos sul-americanos: padrões gerais e possíveis processos gerais e possíveis processos casuais," in *Ecologia de Peixes de Riachos*, eds E. P. Caramashi and R. Mazzoni e Peres-Neto (Rio de Janeiro: PPGE Instituto de Biologia), 139–155.

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9:772. doi: 10.1038/nmeth.2109

de Oliveira, R. R., Feldberg, E., Anjos, M. B., and Zuanon, J. (2007). Karyotype characterization and ZZ/ZW sex chromosome heteromorphism in two species of the catfish genus *Ancistrus* Kner, 1854 (Siluriformes: *Loricariidae*) from the Amazon basin. *Neotrop. Ichthyol.* 5, 301–306. doi: 10.1590/S1679-62252007000300010

de Oliveira, R. R., Feldberg, E., Anjos, M. B., and Zuanon, J. (2008). Ocurrence of multiple sexual chromosomes (XX/XY1Y2 and Z1Z1Z2Z2/Z1Z2W1W2) in catfishes of the genus *Ancistrus* (Siluriformes, Loricariidae) from the Amazon Basin. *Genetica* 134, 243–249. doi: 10.1007/s10709-007-9231-9

de Oliveira, R. R., Feldberg, E., Anjos, M. B., and Zuanon, J. (2009). Mechanisms of chromosomal evolution and its possible relation to natural history characteristics in *Ancistrus* catfishes (Siluriformes: *Loricariidae*). *J. Fish Biol.* 75, 2209–2225. doi: 10.1111/j.1095-8649.2009.02450.x

De Queiroz, K. (2007). Species concepts and species delimitation. *Syst. Biol.* 56, 879–886. doi: 10.1080/10635150701701083

Dimitri, P., Caizzi, R., Giordano, E.,Carmela Accardo, M., Lattanzi, G., and Biamonti, G. (2009). Constitutive heterochromatin: a surprising variety of expressed sequences. *Chromosoma* 118, 419–435. doi: 10.1007/s00412-009-0211-y

Drummond, A. J., Ho, S. Y., Phillips, M. J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88. doi: 10.1371/journal.pbio.0040088

Drummond, A. J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214. doi: 10.1186/1471-2148-7-214

Ence, D. D., and Carstens, B. C. (2011). SpedeSTEM: a rapid and accurate method for species delimitation. *Mol. Ecol. Resour.* 11, 473–480. doi: 10.1111/j.1755-0998.2010.02947.x

Favarato, R. M., Silva, M., Oliveira, R. R., Artoni, R. F., Feldberg, E., and Matoso, D. A. (2016). Cytogenetic diversity and the evolutionary dynamics of rDNA genes and telomeric sequences in the *Ancistrus* genus (*Loricariidae*: Ancistrini). *Zebrafish* 13, 103–111. doi: 10.1089/zeb.2015.1140

Ferraris, C. J., Jr. (2007). Checklist of catfishes, recent and fossil (Osteichthyes: Siluriformes), and catalogue of Siluriform primary types. *New Zealand. Zootaxa* 1418:628. doi: 10.11646/zootaxa.1418.1.1

Fisch-Muller, S. (2003). "Subfamily Ancistrinae (Armored catfishes)," in *Check List of the Freshwater Fishes of South and Central America*, eds R. E. Reis, S. O. Kullander Jr., and C. J. Ferraris (Porto Alegre: Museu de Ciências e Tecnologia), 373–400.

Froese, R., and Pauly, D. (2017). *FishBase: World Wide Web Electronic Publication* [Dataset]. Available online at: http://www.fishbase.org

Fujisawa, T., and Barraclough, T. G. (2013). Delimiting species using single-locus data and the generalized mixed yule coalescent (GMYC) approach: a revised method and evaluation on simulated datasets. *Syst. Biol.* 62, 707–724. doi: 10.1093/sysbio/syt033

Goloboff, P. A., Farris, J. S., and Nixon, K. C. (2003). *T.N.T.: Tree Analysis Using New Technology*. Program and Documentation. Available online at: http://www.lillo.org.ar/phylogeny/

Griffiths, S. P. (2000). The use clove oil as an anaesthetic and method for sampling intertidal rockpool fishes. *J. Fish Biol.* 57, 1453–1464. doi: 10.1111/j.1095-8649.2000.tb02224.x

Gross, M. C., Schneider, C. H., Valente, G. T., Martins, C., and Feldberg, E. (2010). Variability of 18S rDNA locus among *Symphysodon* fishes: chromosomal rearrangements. *J. Fish Biol.* 76, 1117–1127. doi: 10.1111/j.1095-8649.2010.02550.x

Gross, M. C., Schneider, C. H., Valente, G. T., Porto, J. I., Martins, C., and Feldberg, E. (2009). Comparative cytogenetic analysis of the genus *Symphysodon* (Discus Fishes, *Cichlidae*): chromosomal characteristics of retrotransposons and minor ribosomal DNA. *Cytogenet. Genome Res.* 127, 43–53. doi: 10.1159/000279443

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704. doi: 10.1080/10635150390235520

Hall, T. (1999). *Bioedit Program*. Available online at: http://www.mbio.ncsu.edu/BioEdit/bioedit.html

Hatanaka, T., and Galetti, P. M., Jr. (2004). Mapping of the 18S and 5S ribosomal RNA genes in the fish *Prochilodus argenteus* Agassiz, 1829 (*Characiformes*, Prochilodontidae). *Genetica* 122, 239–244. doi: 10.1007/s10709-004-2039-y

Henning, F., Moyses, C. B., Calcagnotto, D., Meyer, A., and Almeida-Toledo, L. F. (2011). Independent fusions and recent origins of sex chromosomes in the evolution and diversification of glass knife fishes (*Eigenmannia*). *Heredity* 106, 391–400. doi: 10.1038/hdy.2010.82

Howell, W. M., and Black, D. A. (1980). Controlled silver staining of nucleolus organizer regions with a protective colloidal developer: a 1-step method. *Experientia* 36, 1014–1015. doi: 10.1007/BF01953855

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–120. doi: 10.1007/BF01731581

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Langeani, F., Castro, R. M. C., Oyakawa, O. T., Shibatta, O. A., Pavanelli, C. S., and Casatti, L. (2007). Diversidade da ictiofauna do Alto Rio Paraná: composição atual e perspectivas futuras. *Biota Neotrop.* 7, 181–197. doi: 10.1590/S1676-06032007000300020

Lui, R. L., Blanco, D. R., Moreira-Filho, O., and Margarido, V. P. (2012). Propidium iodide for making heterochromatin more evident in the C-banding technique. *Biotech. Histochem.* 87, 433–438. doi: 10.3109/10520295.2012.696700

Lujan, N. K., Armbruster, J., Lovejoy, N. R., and López-Férnandez, H. (2015). Multilocus molecular phylogeny of the suckermouth armored catfishes (Siluriformes: *Loricariidae*) with a focus on subfamily Hypostominae. *Mol. Phylogenet. Evol.* 82, 269–288. doi: 10.1016/j.ympev.2014.08.020

Mariotto, S., Artoni, R. F., and Miyazawa, C. S. (2004). Occurence of sexual chromosome, of the type ZZ/ZW, in *Ancistrus* cf. *dubius* (Loricariidae, Ancistrinae) of the Paraguay River Basin, Mato Grosso, Brazil. *Caryologia* 57, 327–331. doi: 10.1080/00087114.2004.10589413

Mariotto, S., Centofante, L., Miyazawa, C. S., Bertollo, L. A. C., and Moreira-Filho, O. (2009). Chromosome polymorphism in *Ancistrus cuiabae* Knaack, 1999 (Siluriformes: Loricariidae: Ancistrini). *Neotrop. Ichthyol.* 7, 595–600. doi: 10.1590/S1679-62252009000400006

Mariotto, S., Centofante, L., and Moreira-Filho, O. (2013). Diversity and chromosomal evolution in the genus *Ancistrus* Kner, 1854 (*Loricariidae*: Ancistrini) from three hydrographic basins of Mato Grosso State, Brazil. *Neotrop. Ichthyol.* 11, 125–131. doi: 10.1590/S1679-62252013000100015

Mariotto, S., Centofante, L., Vicari, M. R., Artoni, R. F., and Moreira-Filho, O. (2011). Chromosomal diversification in ribosomal DNA sites in *Ancistrus*

Kner, 1854 (*Loricariidae*, Ancistrini) from three hydrographic basins of Mato Grosso, Brazil. *Comp. Cytogenet.* 5, 289–300. doi: 10.3897/compcytogen.v5i4.1757

Mariotto, S., and Miyazawa, C. S. (2006). *Ancistrus* cf. *dubius* (Siluriformes, Ancistrinae), a complex of species. 1. Chromosomal characterization of four populations and occurrence of sex chromosomes of the type XX/XY, in the Pantanal Basin of Mato Grosso, Brazil. *Caryologia* 59, 299–304. doi: 10.1080/00087114.2006.10797929

Marshall, D. C., Hill, K. B., Cooley, J. R., and Simon, C. (2011). Hybridization, mitochondrial DNA phylogeography, and prediction of the early stages of reproductive isolation: lessons from New Zealand cicadas (Genus *Kikihia*). *Syst. Biol.* 60, 482–502. doi: 10.1093/sysbio/syr017

Martins, C., and Galetti, P. M., Jr. (1999). Chromosomal localization of 5S rDNA genes in *Leporinus* fish (*Anostomidae*, *Characiformes*). *Chromosome Res.* 7, 363–367. doi: 10.1023/A:1009216030316

Medeiros, L. A., Ginani, E. G., Sousa, L. M., Py-Daniel, L. H. R., and Feldberg, E. (2016). Cytogenetic analysis of *Baryancistrus xanthellus* (Siluriformes: *Loricariidae*: Ancistrini), an ornamental fish endemic to the Xingu River, Brazil. *Neotrop. Ichthyol.* 14:e150108. doi: 10.1590/1982-0224-20150108

Milhomen, S. S. R., Pieczarka, J. C., Crampton, W. G. R., Souza, A. C. P., Carvalho, J. R., Jr., et al. (2007). Differences in karyotype between two sympatric species of *Gymnotus* (*Gymnotiformes*: *Gymnotidae*) from the easern amazon of Brazil. *Zootaxa* 1397, 55–62. doi: 10.11646/zootaxa.1397.7

Miller, G. H., Pfeiffer, W., and Schwartz, T. (2010). "Creating the CIPRES Science Gateway for inference of large phylogenetic trees," in *Proceeding of the Gatway Computing Environments Workshop* (New Orleands, LA), 1–8.

Monaghan, M. T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D. J., et al. (2009). Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Syst. Biol.* 58, 298–311. doi: 10.1093/sysbio/syp027

Montoya-Burgos, J. I., Fisch-Muller, S., Weber, C., and Pawlowski, J. (1998). "Phylogenetic relationships of the Loricariidae (Siluriformes) based on mitochondrial rRNA gene sequences," in *Phylogeny and Classification of Neotropical Fishes*, eds L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M. S. Lucena, and C. A. S. Lucena (Porto Alegre: Edipucrs), 363–374.

Moreira-Filho, O. L., Bertollo, A. C., and Galetti, P. M., Jr. (1984). Structure and variability of nucleolar organizer regions in Parodontidae fish. *Can. J. Genet. Cytol.* 26, 564–568. doi: 10.1139/g84-089

Nakayama, C. M., Jégu, M., Porto, J. I. R., and Feldberg, E. (2001). Karyological evidence for a cryptic species of piranha within *Serrasalmus rhombeus* (*Characidae*, Serrasalminae) in the Amazon. *Copeia* 3, 866–869. doi: 10.1643/0045-8511(2001)001[0866:KEFACS]2.0.CO;2

Nylander, J. A., Ronquist, F., Huelsenbeck, J. P., and Nieves-Aldrey, J. L. (2004). Bayesian phylogenetic analysis of combined data. *Syst. Biol.* 53, 47–67. doi: 10.1080/10635150490264699

Pendas, A. M., Moran, P., Freije, J. P., and Garcia-Vazquez, E. (1994). Chromosomal mapping and nucleotide sequence of two tandem repeats of Atlantic salmon 5S rDNA. *Cytogenet. Genome Res.* 67, 31–36. doi: 10.1159/000133792

Pereira, L. H., Hanner, R., Foresti, F., and Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genet.* 14:e20. doi: 10.1186/1471-2156-14-20

Pereira, L. H. G., Pazian, M. F., Hanner, R., Foresti, F., and Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (*Characiformes*: *Characidae*) from the Upper Paraná Basin of Brazil. *Mitochondrial DNA* 22, 87–96. doi: 10.3109/19401736.2011.588213

Pinkel, D., Straume, T., and Gray, J. W. (1986). Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc. Natl. Acad. Sci. U.S.A.* 83, 2934–2938. doi: 10.1073/pnas.83.9.2934

Powell, J. R. (2012). Accounting for uncertainty in species delineation during the analysis of environmental DNA sequence data. *Methods Ecol. Evol.* 3, 1–11. doi: 10.1111/j.2041-210X.2011.00122.x

Power, M. E. (1984). The importance of sediment in the grazing ecology and size class interactions of an armored catfish, *Ancistrus spinosus*. *Environ. Biol. Fishes* 10, 173–181. doi: 10.1007/BF00001124

Power, M. E. (1990). Resource enhancement by indirect effects of grazers: armored catfish, algae, a sediment. *Ecology* 71, 897–904. doi: 10.2307/1937361

Prizon, A. C., Borin-Carvalho, L. A., Bruschi, D. P., Ribeiro, M. O., Barbosa, L. M., Ferreira, G. E., et al. (2016). Cytogenetic data on *Ancistrus* sp. *(Siluriformes, Loricariidae)* of the Paraguay River basin (MS) sheds light on intrageneric karyotype diversification. *Comp. Cytogenet.* 10, 625–636. doi: 10.3897/compcytogen.v10i4.8532

Rambaut, A., Suchard, M. A., Xie, D., and Drummond, A. J. (2014). *Tracer v1.6*. Available online at: http://tree.bio.ed.ac.uk/software/tracer/

Ronquist, F., and Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574. doi: 10.1093/bioinformatics/btg180

Roxo, F. F., Ochoa, L. E., Costa-Silva, G. J., and Oliveira, C. (2015). Species delimitation in *Neoplecostomus* (Siluriformes: *Loricariidae*) using morphologic and genetic approaches. *DNA Barcodes* 3, 110–117. doi: 10.1515/dna-2015-0014

Schaefer, S. A. (1987). Osteology of *Hypostomus plecostomus* (Linnaeus), with a phylogenetic analysis of the loricariid subfamilies (Pisces: Siluroidei). *Contribut. Sci.* 394, 1–31.

Schweizer, D. (1976). Reverse fluorescent chromosome banding with chromomycin and DAPI. *Chromossoma* 58, 307–324. doi: 10.1007/BF00292840

Sites, J. W., and Kent, M. (1994). Chromosomal evolution, speciation, and systematics: some relevant issues. *Herpetologica* 50, 237–249.

Sumner, A. T. (1972). A simple technique for demonstrating centromeric heterochromatin. *Exp. Cell Res.* 75, 304–306. doi: 10.1016/0014-4827(72)90558-7

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Viana, D., Zawadzki, C. H., Oliveira, E. F., Vogel, H. F., and Graça, W. J. (2013). Structure of the ichthyofauna of the Bonito river, Ivaí river basin, upper Paraná river system, Brazil. *Biota Neotrop.* 13, 219–226. doi: 10.1590/S1676-06032013000200021

Vicari, M. R., Artoni, R. F., and Bertollo, L. A. C. (2003). Heterochromatin polymorphism associated with 18S rDNA. A diferrential pathway among the fish *Hoplias malabaricus* from Southern Brazil. *Cytogenet. Genome Res.* 101, 24–28. doi: 10.1159/000073413

Vicari, M. R., Noleto, R. B., Artoni, R. F., Moreira-Filho, O., and Bertollo, L. A. C. (2008). Comparative cytogenetics among species of the *Astyanax scabripinnis* complex. Evolutionary and biogeographical inferences. *Genet. Mol. Biol.* 31, 173–179. doi: 10.1590/S1415-47572008000200002

Vieites, D. R., Wollenberg, K. C., Andreone, F., Köhler, J., Glaw, F., and Vences, M. (2009). Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8267–8272. doi: 10.1073/pnas.0810821106

Vuataz, L., Sartori, M., Wagner, A., and Monaghan, M. T. (2011). Toward a DNA taxonomy of alpine Rhithrogena (Ephemeroptera: Heptageniidae) using a mixed Yule-coalescent analysis of mitochondrial and nuclear DNA. *PLoS ONE* 6:e19728. doi: 10.1371/journal.pone.0019728

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., and Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1847–1857. doi: 10.1098/rstb.2005.1716

Yang, Z., and Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9264–9269. doi: 10.1073/pnas.0913022107

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

Copyright © 2017 Prizon, Bruschi, Borin-Carvalho, Cius, Barbosa, Ruiz, Zawadzki, Fenocchio and Portela-Castro. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 05 December 2017
doi: 10.3389/fgene.2017.00199

Headwater Capture Evidenced by Paleo-Rivers Reconstruction and Population Genetic Structure of the Armored Catfish (*Pareiorhaphis garbei*) in the Serra do Mar Mountains of Southeastern Brazil

Sergio M. Q. Lima[1,2], Waldir M. Berbel-Filho[1,3], Thais F. P. Araújo[1], Henrique Lazzarotto[4,5], Andrey Tatarenkov[2] and John C. Avise[2]*

[1] Laboratório de Ictiologia Sistemática e Evolutiva, Departamento de Botânica e Zoologia, Centro de Biociências, Universidade Federal do Rio Grande do Norte, Natal, Brazil, [2] Department of Ecology and Evolutionary Biology, University of California, Irvine, Irvine, CA, United States, [3] Department of Biosciences, College of Science, Swansea University, Swansea, United Kingdom, [4] Laboratório de Ecologia de Peixes, Departamento de Ecologia, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, [5] California Academy of Sciences, San Francisco, CA, United States

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Tomas Hrbek, Federal University of Amazonas, Brazil
Peter A. Ritchie, Victoria University of Wellington, New Zealand

***Correspondence:**
Sergio M. Q. Lima
smaialima@gmail.com

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

Received: *01 September 2017*
Accepted: *21 November 2017*
Published: *05 December 2017*

Citation:
Lima SMQ, Berbel-Filho WM, Araújo TFP, Lazzarotto H, Tatarenkov A and Avise JC (2017) Headwater Capture Evidenced by Paleo-Rivers Reconstruction and Population Genetic Structure of the Armored Catfish (Pareiorhaphis garbei) in the Serra do Mar Mountains of Southeastern Brazil. Front. Genet. 8:199. doi: 10.3389/fgene.2017.00199

Paleo-drainage connections and headwater stream-captures are two main historical processes shaping the distribution of strictly freshwater fishes. Recently, bathymetric-based methods of paleo-drainage reconstruction have opened new possibilities to investigate how these processes have shaped the genetic structure of freshwater organisms. In this context, the present study used paleo-drainage reconstructions and single-locus cluster delimitation analyses to examine genetic structure on the whole distribution of *Pareiorhaphis garbei*, a 'near threatened' armored catfish from the Fluminense freshwater ecoregion in Southeastern Brazil. Sequences of two mitochondrial genes (cytochrome b and cytochrome c oxidase subunit 1) were obtained from five sampling sites in four coastal drainages: Macaé (KAE), São João (SJO), Guapi-Macacu [sub-basins Guapiaçu (GAC) and Guapimirim (GMI)], and Santo Aleixo (SAL). Pronounced genetic structure was found, involving 10 haplotypes for *cytB* and 6 for *coi*, with no haplotypes shared between localities. Coalescent-based delineation methods as well as distance-based methods revealed genetic clusters corresponding to each sample site. Paleo-drainage reconstructions showed two putative paleo-rivers: an eastern one connecting KAE and SJO; and a western one merging in the Guanabara Bay (GAC, GMI, and SAL). A disagreement was uncovered between the inferred past riverine connections and current population genetic structure. Although KAE and SJO belong to the same paleo-river, the latter is more closely related to specimens from the Guanabara paleo-river. This discordance between paleo-drainage connections and phylogenetic structure may indicate an ancient stream-capture event in headwaters of this region. Furthermore, all analyses showed high divergence between KAE and the other lineages, suggesting at least one cryptic species in the latter, and that the nominal species should be restricted to the Macaé river basin, its type locality. In this drainage, impacts such as the invasive species and habitat loss can be especially threatening

for such species with a narrow range. Our results also suggest that freshwater fishes from headwaters in the Serra do Mar mountains might have different biogeographical patterns than those from the lowlands, indicating a complex and dynamic climatic and geomorphological history.

Keywords: biogeography, Atlantic Forest, phylogeography, conservation genetics, Loricariidae

INTRODUCTION

Two alternative hypotheses are typically invoked to explain disjoint distributions encompassing more than one drainage basin of strictly freshwater species: past anastomoses of coastal basins due to marine regression episodes that resulted in paleo-drainage connections or headwater stream-captures caused by tectonic adjustments (Albert and Reis, 2011). Stream capture occurs when a river changes its course and connects with another drainage system as a result of geomorphological changes (Bishop, 1995; Wilkinson et al., 2006), whereas paleo-drainage connections happen when sea levels are lowered by climatic changes, causing basins that previously were isolated to coalesce (Voris, 2000; Dias et al., 2014). To assess how these processes have influenced the genetic structure of freshwater fishes, recently developed GIS-based methods have been used to reconstruct paleo-river systems (Dias et al., 2014; Thomaz et al., 2015, 2017). These methods use topographical and bathymetrical data to reconstruct land exposure, fine-scale depth and steepness of exposed areas (which represent putative riverbeds), as well as putative flow and basin limits (Thomaz et al., 2015). Bathymetric-based reconstructions proved to be highly concordant with the available geological record for paleo-drainages (Dias et al., 2014), supporting GIS-based reconstruction as a powerful and reliable tool for paleo-drainage reconstructions, especially in regions with scarce geological data (Thomaz et al., 2015).

Headwater regions usually exhibit high levels of endemism and microhabitat specialization (Albert et al., 2011; Buckup, 2011). One such example of headwater taxa involves loricariid catfishes of the genus *Pareiorhaphis*. These rheophilic (fast-water) fishes usually show a narrow distribution, with 20 of 25 described species (80%) restricted to a single basin (Pereira et al., 2017). One of the exceptions within the genus is *Pareiorhaphis garbei* (Ihering, 1911), a species found in four coastal basins of the Rio de Janeiro State in southeastern Brazil: Macaé, São João, Guapi-Macacu, and Santo Aleixo drainages (Maia et al., 2013). This armored catfish is restricted to clear-water streams with fast-flowing high-oxygen waters and predominance of rocky (bouldered) substrate (Lazzarotto et al., 2007), and these habitats only occur in the headwaters or upper reaches of coastal basins (Maia et al., 2013). The geographic distribution of *P. garbei* lies entirely within the Atlantic Forest biodiversity hotspot (Myers et al., 2000), which in turn is situated within the Fluminense freshwater ecoregion, a small area comprising about 110 freshwater fish species and a high proportion (42%) of endemic species (Albert et al., 2011).

Due to the intense modification of natural areas by rural and urban expansion, many freshwater fish species in the Fluminense ecoregion are threatened. Previously, *P. garbei* was classified as an 'endangered' species [Ministério do Meio Ambiente [MMA], 2004], but it is currently considered 'near threatened' and is, thus, a priority species for research [Instituto Chico Mendes de Conservação da Biodiversidade [ICMBio], 2014]. Deforestation, chemical pesticides, and the introduction of rainbow trout *Oncorhynchus mykiss* (Walbaum, 1792) were shown to be threats for *P. garbei* (Pereira and Brito, 2008), even though most records for the species are in protected areas (Maia et al., 2013). However, molecular studies indicate that some taxa putatively distributed along coastal basins of the Fluminense ecoregion may be species complexes with individually narrow distributions (Villa-Verde et al., 2012; Cherobim et al., 2016; Roxo et al., 2017).

Application of molecular phylogenetic analyses and paleo-drainage reconstruction of headwater taxa might reveal their population histories, including information relevant for taxonomy and conservation. Single-locus cluster delineation, often based on mitochondrial DNA, is drastically changing investigations of global biodiversity (Hajibabaei et al., 2007; Pereira L.H. et al., 2013). Clusters revealed by this analysis are considered operational taxonomic units (OTUs). Units that are supported by several clustering methods with different assumptions offer a starting point for taxonomic reevaluation (Kekkonen and Hebert, 2014). This study addresses the genetic structure of *P. garbei* to infer historical processes that may have influenced the species' disjointed distribution along coastal basins of the Serra do Mar. In particular, we will address possible paleo-drainage connections and headwater captures. If the lineage patterns are consistent with reconstructed paleo-rivers, then anastomosis of coastal basins by regressive glacial episodes would likely explain the current distribution. However, if the observed phylogeographic pattern is incongruent with the paleo-river reconstruction, then stream capture would likely explain lineage patterns along the headwaters of distinct basins. Furthermore, the species delimitation analyses will provide complementary data for systematic studies, as well as useful information for evaluating the conservation status of *P. garbei*.

MATERIALS AND METHODS

Sampling

Specimens of *P. garbei* were captured in five localities of four hydrographic basins covering the entire known species distribution (Maia et al., 2013): Macaé (KAE), São João (SJO), Guapi-Macacu [sub-basins Guapiaçu (GAC) and Guapimirim (GMI)], and Santo Aleixo (SAL) (**Table 1**). Whereas the rivers Santo Aleixo and Guapi-Macacu flow into the Guanabara Bay in the same mangrove area (and their river mouths are less than 2 km apart), the São João and Macaé rivers flow to the east

TABLE 1 | Sampling sites of *Pareiorhaphis garbei* in the Serra do Mar coastal basins, Rio de Janeiro State, southeastern Brazil.

Locality	Basin	Geographic Coordinates	Altitude (m)	CU	Vouchers (UFRN)	*cytB*	*coi*
Macaé (KAE)	Macaé	22°25′50.1″ S 42°32′20.2″ W	1090	1	098, 108	6	6
São João (SJO)	São João	22°29′54.5″ S 42°30′34.6″ W	90	2	010, 1025	4	10
Guapiaçu (GAC)	Guapi-Macacu	22°24′53″ S 42°43′37″ W	180	3	DEPR J8483	3	5
Guapimirim (GMI)	Guapi-Macacu	22°29′32.1″S 43°00′03.3″ W	400	3	092	4	4
Santo Aleixo (SAL)	Santo Aleixo	22°31′29.7″ S 43°01′54.1″ W	270	3	103, 111	1	–

Conservation Units (CU) are as follows: (1) Três Picos State Park, (2) Bacia do Rio São João Environmental Protection Area, and (3) Petrópolis Environmental Protection Area/Serra dos Órgãos National Park. Vouchers: (UFRN) Universidade Federal do Rio Grande do Norte fish collection, (DEPRJ) Departamento de Ecologia Universidade Federal do Rio de Janeiro fish collection. Columns cytB and coi show number of specimens studied at each gene.

coast of Rio de Janeiro State directly into the Atlantic Ocean, about 100 km from Guanabara Bay (Costa, 2014) (**Figure 1**). Muscle tissues or fin clips were obtained from specimens preserved in absolute ethanol. These individuals are stored at the ichthyological collections of the Universidade Federal do Rio Grande do Norte and of the Departamento de Ecologia of the Universidade Federal do Rio de Janeiro (**Table 1**).

Paleo-Drainages Reconstruction

To access paleo-drainages connections during the Last Glacial Maximum (LGM), topographical and bathymetrical data were retrieved from the digital elevation model at 30 arc-second resolution[1]. This layer was uploaded into ArcGIS10 software, and the *Hydrological* tools add-in was used to reconstruct the paleo-drainages, following the steps described in Thomaz et al. (2015). First, the area exposed during the LGM (−125 m) was identified using the *Contour* followed by *Mask*. Afterwards, the tools *Fill*, *Flow Direction*, and *Basin* were used, respectively, to fill depressions on the surface, identify the steepness within each cell, and define the basin borders. Finally, the *Stream order* function was used to estimate putative paleo-rivers.

DNA Extraction and Amplification

Genomic DNA samples were extracted using phenol/chloroform/isopropanol with ethanol precipitation protocol (Milligan, 1998), and mitochondrial DNA fragments of genes encoding cytochrome oxidase subunit I (*coi*) and cytochrome b (*cytB*) were amplified and sequenced using primers and conditions proposed by Villa-Verde et al. (2012). Overall, 18 and 25 specimens were amplified and sequenced for *cytB* (1056 bp) and *coi* (878 bp), respectively (GenBank MG251217-251259) (**Table 1**).

Phylogenetic Analysis and Species Delimitation

Sequences were edited in MEGA6 (Tamura et al., 2013), then aligned using the MUSCLE algorithm (Edgar, 2004). *P*-distance (*cytB*) and Kimura-2-parameter (K2P) (*coi*) distances for markers were also calculated in MEGA6. A Bayesian phylogenetic reconstruction for both *cytB* and *coi* data was done in BEAST v. 1.7 (Drummond et al., 2012) using HKY+G (*cytB*) and K2P+G (*coi*) as nucleotide substitution models, as suggested by jModeltest 2 (Darriba et al., 2012). An uncorrelated relaxed lognormal model with estimated rate was used, with ucld.mean parameter set and uniform distribution (0 and 10 as lower and upper boundaries). Remaining parameters were set as default. The length of the MCMC chain was 10,000,000 runs with sampling every 1000 runs. ESS (> 200) values were checked using Tracer v. 1.5 (Rambaut, 2009). The initial 2000 trees were discarded as burn-in period, and a final tree was reconstructed using TreeAnnotator v.1.50. Each haplotype network was inferred using the TCS method on PopART (Leigh and Bryant, 2015). *Pareiorhaphis* cf. *bahianus* from Contas river basin in Bahia State was used as out-group in an additional Bayesian phylogeny, using the same parameters described above, for each gene to corroborate the relationships among *P. garbei* lineages (**Supplementary Figure S1**).

For increased robustness in delimiting OTUs, five single-locus species-delimitation analyses were performed using both markers separately: single-threshold of Generalized Mixed Yule-Coalescent (sGMYC) (Fujisawa and Barraclough, 2013); multiple-threshold GMYC (mGMYC) (Fujisawa and Barraclough, 2013); Bayesian implementation of Poisson Tree Process (bPTP) (Zhang et al., 2013); multiple rate PTP (mPTP) (Kapli et al., 2017); and Automatic Barcode Gap Discovery (ABGD) (Puillandre et al., 2012). Ultrametric trees generated from *cytB* and *coi* data of the Bayesian phylogeny were used as an input file for sGMYC and mGMYC. Previous studies have shown that the clock model and tree prior have low impact on the results of both sGMYC and mGMYC (Tavalera et al., 2013). These two analyses were conducted in R (R Core Team, 2017), using the package SPLITS (Ezard et al., 2009). We re-ran the phylogenetic reconstructions under the same parameters used in BEAST using MrBayes v. 3.2.6 (Huelsenbeck and Ronquist, 2001) to generate a bifurcating phylogram (with branch length representing the number of substitutions) to be the input file for PTPs analyses. bPTP analyses were performed in the online server[2]. The analysis length was 500,000 generations with sampling of 500 and burn-in of 0.1. Convergence was visualized on the MCMC interactions plots vs. log-likelihood. mPTP was run using the online server[3], under the same parameters as for bPTP. ABGD distance-based analyses were run through the software command line, with a gap width value of 1.0 for all the distances available (*p*-distance, K2P, and Jukes–Cantor). The ABGD delineation taken into account

[1]http://www.gebco.net/

[2]http://species.h-its.org

[3]http://mptp.h-its.org

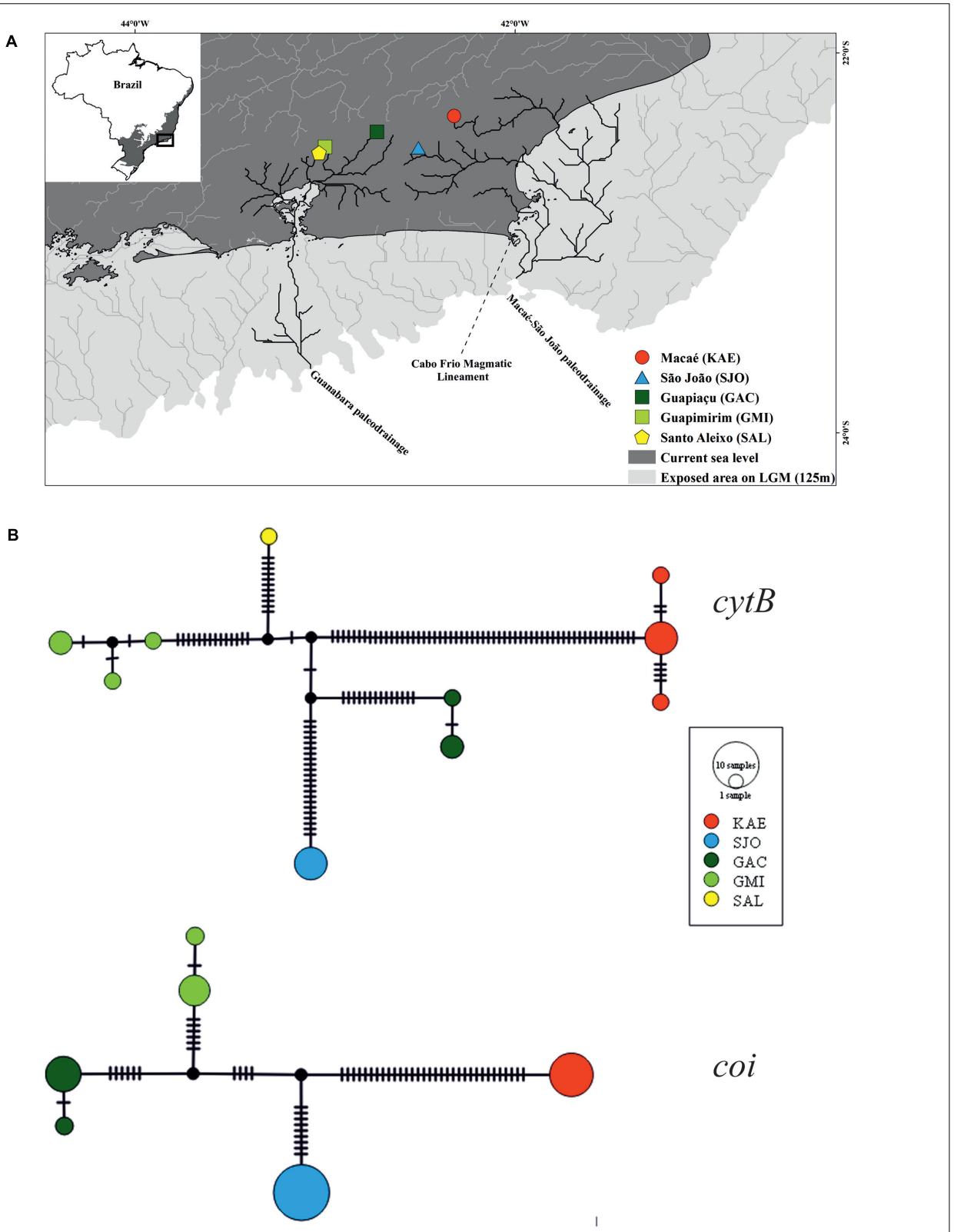


FIGURE 1 | (A) Paleo-drainages reconstruction of coastal basins from Serra do Mar mountains, Rio de Janeiro, southeastern Brazil. Dark gray represents the continent at current sea level and light gray a putative exposure of the continental shelf during a sea regression of 125 m. The Atlantic Forest distribution is represented in grey on the inset map. **(B)** Haplotypes networks of cytochrome b (*cytB*) and cytochrome oxidase I (*coi*) in *Pareiorhaphis garbei*.

was the one with the *P*-value of ∼0.01, as advocated by previous studies (Puillandre et al., 2012; Blair and Bryson, 2017). To have a species-delimitation partition based on a genetic distance threshold value, we used cut-off values of 5% divergence for *cytB* and 2% divergence for *coi* as indicators of distinct species. These values were based on previous reviews on genetic distances on fish species (Ward, 2009; Kartavtsev, 2011).

RESULTS

A total of 10 *cytB* and six *coi* haplotypes were found, none of which were shared among localities. For both mtDNA markers, the haplotype networks showed multiple mutational steps among all haplotypes (**Figure 1**). Mutations were especially abundant along the branch connecting KAE and the other populations (**Figure 1**). Genetic *p*-distances and K2P distances are presented in **Table 2**. Between KAE and the other localities, divergence values ranged from 7.5 to 8.0% for *cytB* and 4.9 to 5.0% for *coi*; between SJO and rivers that flow to the Guanabara Bay (GAC, GMI, and SAL) from 3.6 to 4.4% (*cytB*) and 2.2% (*coi*); and, among the watersheds that run into Guanabara Bay, from 2.6 to 3.2% (*cytB*) and 1.4% (*coi*) (**Table 2**).

Our phylogenetic reconstructions showed high genetic structuring within *P. garbei*, with all sampling sites representing monophyletic clusters in both the *cytB* and the *coi* data. According to this result, SJO is more closely related to the Guanabara Bay populations (GAC, GMI, and SAL) than to the KAE. Within the Guanabara Bay watersheds, GMI is more closely related to SAL (from an adjacent basin) than it is to GAC (that belongs to the same drainage, Guapi-Macacu), although supported by lower posterior probabilities (**Figure 2**).

The sGMYC and ABGD delimitation analyses indicated five distinct OTUs, corresponding to each sampling site for *cytB*. mGMYC split the SJO clade into two different OTUs, resulting in six genetic clusters. Both PTP methods (bPTP and mPTP) indicated potential distinct lineages within the KAE clade. The strict species threshold (*p*-distance > 5%) separates KAE from the other sampling sites (**Figure 2A**). Regarding to the *coi* reconstruction, four OTUs corresponding to sampling sites were shown as the best genetic partition in three of the five species-delineation methods (sGMYC, bPTP, and ABGD). However, some incongruences were found. Similarly to the *cytB* reconstruction, mGMYC tended to split lineages, dividing the GAC clade into two different clusters. Furthermore, mPTP found putative distinct genetic lineages within the SJO and KAE clades, resulting in a total of seven OTUs. Using the K2P distance species threshold value (>2%), the Guanabara Bay drainages represent one species while SJO and KAE each represent different putative species (**Figure 2B**).

Our paleo-drainage reconstruction places Macaé (KAE) and São João (SJO) in the same paleo-drainage (the Macaé-São João paleo-river), whereas Guapiaçu (GAC), Guapimirim (GMI), and Santo Aleixo (SAL) all reside in the Guanabara paleo-river (**Figure 1A**). This paleo-reconstruction does not corroborate the phylogenetic result, according to which SJO is somewhat more closely related to Guanabara Bay rivers populations than to KAE.

TABLE 2 | *P*-distances of cytochrome b gene (below diagonal) and K2P distances of cytochrome oxidase subunit I gene (above diagonal) for *Pareiorhaphis garbei* in the Serra do Mar coastal basins, Rio de Janeiro State, southeastern Brazil.

	KAE	SJO	GAC	GMI	SAL
Macaé (KAE)		0.049	0.050	0.050	–
São João (SJO)	0.080		0.022	0.022	–
Guapiaçu (GAC)	0.080	0.037		0.014	–
Guapimirim (GMI)	0.080	0.044	0.032		–
Santo Aleixo (SAL)	0.075	0.036	0.028	0.026	

DISCUSSION

Our results revealed a discordance between past paleo-river connections and current genetic structure, which might indicate an ancient stream capture event in the headwaters of the Serra do Mar mountains. According to the paleo-drainage scenario, it was expected that SJO would be closely related to KAE; however, it appears to be a sister group to lineages from the

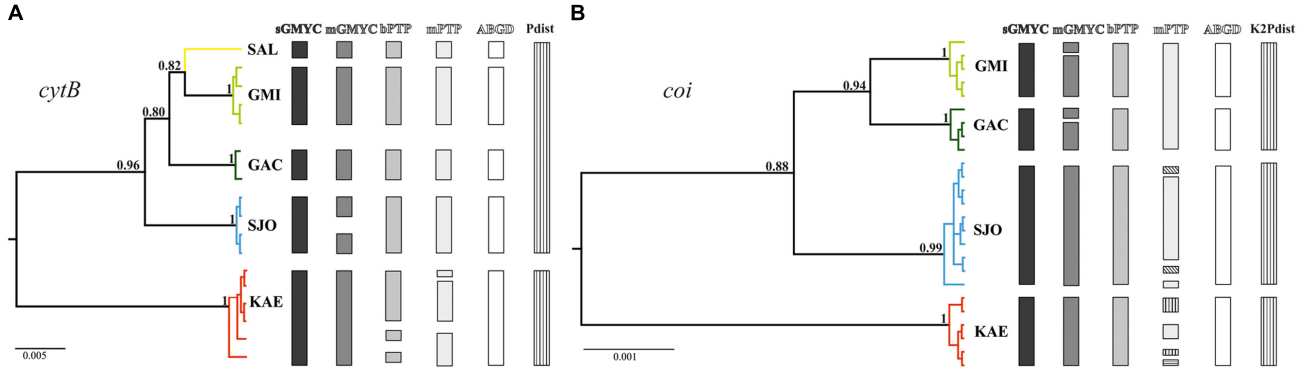


FIGURE 2 | Ultrametric trees generated by the Bayesian phylogenetic reconstructions for **(A)** cytochrome b (*cytB*) and **(B)** cytochrome oxidase I (*coi*) in *Pareiorhaphis garbei*. The bars in different colors and patterns represent the operational taxonomic units (OTUs) partitions obtained by the different species delimitation methods.

Guanabara paleo-river. This result corroborates the river-capture event suggested by Maia et al. (2013) based on the occurrence of *P. garbei* in a single tributary adjacent to the headwaters of Macaé river basin. These drainage rearrangements are the result of tectonic reactivation that started in the Paleogene and continues to the present (Ribeiro et al., 2006; Lima and Ribeiro, 2011; Lima et al., 2016).

Although located within the Guapi-Macacu river basin, the water path connecting GAC and GMI is a lowland area where micro- and meso-habitats suitable for *P. garbei* are absent. The fact that GAC and GMI share no haplotypes, and belong to distinct lineages, reveals that any dispersal of *P. garbei* through the lower portions of rivers is very limited (and might explain in part why the current genetic structure of *P. garbei* does not perfectly reflect the paleo-drainage reconstruction).

Along the geographic range of *P. garbei*, a geographic barrier known as the Cabo Frio Magmatic Lineament (CFML) (Riccomini et al., 2005) may have influenced the evolution and diversification of many freshwater fishes. Based on molecular data, Pereira T.L. et al. (2013) suggested that this barrier had a vicariant effect isolating an eastern lineage (including a clade in the São João and Macaé basins) from a southeastern lineage (from Guanabara Bay to Paranaguá Bay) in the trahira, *Hoplias malabaricus* (Bloch, 1794). This pattern is also corroborated by the geographic distribution of *Atlantirivulus* rivulids in lowland areas (Costa, 2014). Although the CFLM seems to be an effective barrier to *P. garbei* (because different lineages are on its alternate sides), specimens from the São João basin are more closely related to Guanabara Bay drainages than to Macaé drainage, suggesting that different biogeographic forces may apply to headwater versus lowland fishes. Although not exclusively a headwater species, a biogeographic pattern similar to that found in *P. garbei* is also shown by *Hisonotus* loricariids, with *H. notatus* Eigenmann and Eigenmann, 1889 occurring from the São João river basin to the south (including Guanabara Bay's drainages), whereas *H. thayeri* Martins and Langeani, 2016 occurs from the Macaé river northward (Martins and Langeani, 2016). Altogether, these various patterns indicate a complex and dynamic history of the coastal basins of the Fluminense ecoregion. Curiously, the CFLM is also the boundary between two marine provinces: the Tropical Southwestern Atlantic and the Warm Temperate Southwestern Atlantic (Spalding et al., 2007).

Pareiorhaphis garbei was the subject of a taxonomic review (Pereira and Reis, 2002; Pereira, 2005) that did not include molecular data, preventing the discovery of deep genetic structure. Several other demersal Neotropical freshwater fishes that display only subtle morphological differences have proved to show substantial genetic divergences that support the description of new and often endemic species (Benine et al., 2009; Melo et al., 2011; Cherobim et al., 2016), and some of these support an ancient split between rivers flowing to the Guanabara Bay versus the São João and Macaé drainages (Villa-Verde et al., 2012; Costa, 2014; Roxo et al., 2017). All molecular analyses herein performed indicate high divergence of the KAE population of *P. garbei* from all others, strongly suggesting that individuals from these two groups represent different species. If this outcome is formally recognized, *P. garbei* (*sensu stricto*) would be confined to the Macaé river basin (its type locality), with another species distributed along the headwaters of São João, Guapi-Macacu, and Santo Aleixo river basins. Moreover, according to high divergence at *coi*, the São João lineage could also be a distinct species from the one inhabiting drainages entering Guanabara Bay. In either case, *P. garbei* itself would be limited to Macaé river basin, in which the invasive rainbow trout poses a risk factor (Lazzarotto et al., 2007; Pereira and Brito, 2008). Currently, *P. garbei* is classified as 'near threatened' [Instituto Chico Mendes de Conservação da Biodiversidade [ICMBio], 2014], a category that probably stems from the recent record of this species in the São João river and the confirmation of its occurrence in protected areas (Maia et al., 2013). However, if the deep phylogenetic lineages found here corroborate a species complex in *P. garbei*, the conservation status of the entire assemblage may need reexamination, mainly due to a reduction in the range of each species due to the taxonomic splitting.

In summary, our molecular analysis together with paleo-drainage reconstructions revealed that the current phylogeographic patterns of the rheophilic catfish *P. garbei* were importantly impacted by headwater stream-captures. This taxon proved to encompass at least two highly divergent lineages, and furthermore, each headwater seems to represent a genetically diagnosable OTU. The integrative approach employed here helps to introduce a useful way to test hypotheses regarding to the distribution and conservation of Neotropical freshwater fishes.

AUTHOR CONTRIBUTIONS

SL had substantial contribution to the acquisition of samples, analyses, and interpretation of the results, writing the manuscript. WB-F and TA also contributed in the analyses, writing, and figures. HL helped in the fish collection, design, and writing the manuscript. AT and JA provided logistic support for molecular data during SL's postdoctoral activities at UCI. All authors contributed in conception and elaboration of the manuscript, and read and approved its final version.

FUNDING

SL and HL received postdoctoral scholarships from Science without Borders Program/CNPq (203476/2014-0 and 202253/2015-5, respectively). WB-F receives a Ph.D. scholarship from Science without Borders Program/CNPq (233161/2014-7).

ACKNOWLEDGMENTS

We are grateful to Edson H. L. Pereira for receiving and analyzing specimens for morphological differences.

We are further indebted to CNPq (Conselho Nacional de Desenvolvimento Cientifico e Tecnologico) for supporting the research and to ICMBio (Instituto Chico Mendes de Conservação da Biodiversidade) and INEA/RJ (Instituto Estadual do Ambiente do Rio de Janeiro) for providing collecting permits to SL (28332-1, 30532-1, and 049/2011). Thiago Barros and Érica Caramaschi were responsible for providing specimens from Guapiaçu sub-basin, and therefore, we are also thankful to them. We are thankful to Andrea Thomaz for helpful advices on paleo-drainages reconstruction.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2017.00199/full#supplementary-material

FIGURE S1 | Bayesian phylogenetic reconstructions for cytochrome b (*cytB*) and cytochrome oxidase I (*coi*) in *Pareiorhaphis garbei*, using HKY+I (*cytB*) and TrN+G (*coi*) as nucleotide substitution models. *Pareiorhaphis* cf. *bahianus* (MG496258 and MG496259) from Contas river basin (northeastern Brazil) was used as out-group. Posterior probabilities are shown above the nodes.

REFERENCES

Albert, J. S., Petry, P., and Reis, R. E. (2011). "Major biogeographic and phylogenetic patterns," in *Historical Biogeography of Neotropical Freshwater Fishes*, eds J. S. Albert and R. Reis (Berkeley, CA: University of California Press), 21–58. doi: 10.1525/california/9780520268685.003.0002

Albert, J. S., and Reis, R. E. (2011). "Introduction to neotropical freshwaters," in *Historical Biogeography of Neotropical Freshwater Fishes*, eds J. S. Albert and R. Reis (Berkeley, CA: University of California Press), 3–19. doi: 10.1525/california/9780520268685.003.0001

Benine, R. C., Mariguela, T. C., and Oliveira, C. (2009). New species of *Moenkhausia* Eigenmann, 1903 (Characiformes: Characidae) with comments on the *Moenkhausia oligolepis* species complex. *Neotrop. Ichthyol.* 7, 161–168. doi: 10.1590/S1679-62252009000200005

Bishop, P. (1995). Drainage rearrangement by river capture, beheading and diversion. *Prog. Phys. Geogr.* 19, 449–473. doi: 10.1177/030913339501900402

Blair, C., and Bryson, R. W. Jr. (2017). Cryptic diversity and discordance in single-locus species delimitation methods within horned lizards (Phrynosomatidae: *Phrynosoma*). *Mol. Ecol. Resour.* 17, 1168–1182. doi: 10.1111/1755-0998.12658

Buckup, P. A. (2011). "The Eastern Brazilian shield," in *Historical Biogeography of Neotropical Freshwater Fishes*, eds J. S. Albert and R. Reis (Berkeley, CA: University of California Press), 203–210. doi: 10.1525/california/9780520268685.003.0012

Cherobim, A. M., Lazzarotto, H., and Langeani, F. (2016). A new species of the catfish *Neoplecostomus* (Loricariidae: Neoplecostominae) from a coastal drainage in southeastern Brazil. *Neotrop. Ichthyol.* 14:e160015. doi: 10.1590/1982-0224-20160015

Costa, W. J. E. M. (2014). Four new species of the genus *Atlantirivulus* (Cyprinodontiformes: Rivulidae) from the Brazilian Atlantic Forest. *Vertebr. Zool.* 64, 9–21.

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772–772. doi: 10.1038/nmeth.2109

Dias, M. S., Oberdorff, T., Hugueny, B., Leprieur, F., Jézéquel, C., Cornu, J. F., et al. (2014). Global imprint of historical connectivity on freshwater fish biodiversity. *Ecol. Lett.* 17, 1130–1140. doi: 10.1111/ele.12319

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. doi: 10.1093/molbev/mss075

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Ezard, T., Fujisawa, T., and Barraclough, T. G. (2009). *Splits: Species' Limits by Threshold Statistics. R Package Version 1*. Available at: http://r-forge.r-project.org/projects/splits

Fujisawa, T., and Barraclough, T. G. (2013). Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Syst. Biol.* 62, 707–724. doi: 10.1093/sysbio/syt033

Hajibabaei, M., Singer, G. A. C., Hebert, P. D. N., and Hickey, D. A. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends Genet.* 23, 167–172. doi: 10.1016/j.tig.2007.02.001

Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755. doi: 10.1093/bioinformatics/17.8.754

Instituto Chico Mendes de Conservação da Biodiversidade [ICMBio] (2014). *Lista de Espécies Quase Ameaçadas*. Available at: http://www.icmbio.gov.br/portal/images/stories/biodiversidade/fauna-brasileira/quase-ameacadas/Lista_NT_site.xls

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., et al. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics* 33, 1630–1638. doi: 10.1101/063875

Kartavtsev, Y. P. (2011). Divergence at Cyt-b and Co-1 mtDNA genes on different taxonomic levels and genetics of speciation in animals. *Mitochondrial DNA* 22, 55–65. doi: 10.3109/19401736.2011.588215

Kekkonen, M., and Hebert, P. D. N. (2014). DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. *Mol. Ecol. Resour.* 14, 706–715. doi: 10.1111/1755-0998.12233

Lazzarotto, H., Brito, M. F., and Caramaschi, E. P. (2007). Threatened fishes of the world: *Pareiorhaphis garbei* (Ihering, 1911) (Ostariophysii: Loricariidae). *Environ. Biol. Fishes* 78, 91–92. doi: 10.1007/s10641-006-9036-1

Leigh, J. W., and Bryant, D. (2015). Popart: full-feature software for haplotype network construction. *Methods Ecol. Evol.* 6, 1110–1116. doi: 10.1111/2041-210x.12410

Lima, F. C. T., and Ribeiro, A. C. (2011). "Continental-scale tectonic controls of biogeography and ecology," in *Historical Biogeography of Neotropical Freshwater Fishes*, eds J. S. Albert and R. Reis (Berkeley, CA: University of California Press), 145–164. doi: 10.1525/california/9780520268685.003.0009

Lima, S. M. Q., Vasconcellos, A. V., Berbel-Filho, W. M., Lazoski, C., Russo, C. A., Sazima, I., et al. (2016). Effects of Pleistocene climatic and geomorphological changes on the population structure of the restricted-range catfish *Trichogenes longipinnis* (Siluriformes: Trichomycteridae). *Syst. Biodivers.* 14, 155–170. doi: 10.1080/14772000.2015.1104398

Maia, G. F., de Paiva, R. E. C., dos Santos, A. B., Lazzarotto, H., and Lima, S. M. Q. (2013). New distribution record for the threatened armoured suckermouth catfish *Pareiorhaphis garbei* (Ihering, 1911) (Siluriformes: Loricariidae) in São João river basin, Rio de Janeiro, southeastern Brazil. *Check List* 9, 1113–1116. doi: 10.15560/9.5.1113

Martins, F. O., and Langeani, F. (2016). Redescription of *Hisonotus notatus* Eigenmann & Eigenmann, 1889 (Loricariidae: Hypoptopomatinae), the type species of the genus, and description of a new species from coastal drainages of southeastern Brazil. *Neotrop. Ichthyol.* 14:e150100. doi: 10.1590/1982-0224-20150100

Melo, B. F., Benine, R. C., Mariguela, T. C., and Oliveira, C. (2011). A new species of *Tetragonopterus* Cuvier, 1816 (Characiformes: Characidae: Tetragonopterinae) from the rio Jari, Amapá, northern Brazil. *Neotrop. Ichthyol.* 9, 49–56. doi: 10.1590/s1679-62252011000100002

Milligan, B. G. (1998). "Total DNA isolation," in *Molecular Genetic Analysis of Populations*, ed. A. R. Hoelzel (Oxford: Oxford University Press), 28–64.

Ministério do Meio Ambiente [MMA] (2004). *Lista Nacional das Espécies de Invertebrados Aquáticos e Peixes Ameaçados de Extinção com Categorias da IUCN. Instrução Normativa no. 5 de 21 de Maio de 2004. Diário Oficial da União 102: 102-142*. Available at: http://www.mma.gov.br/estruturas/179/_arquivos/in_mma_005_04_179.pdf

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858. doi: 10.1038/35002501

Pereira, E. H., Pessali, T. C., Andrade, F. D., and Reis, R. E. (2017). Description of a new species of *Pareiorhaphis* (Loricariidae: Neoplecostominae) from the rio Jequitinhonha basin, Minas Gerais, eastern Brazil. *Neotrop. Ichthyol.* 15:e170007. doi: 10.1590/1982-0224-20170007

Pereira, E. H., and Reis, R. E. (2002). Revision of the loricariid genera *Hemipsilichthys* and *Isbrueckerichthys* (Teleostei: Siluriformes), with descriptions of five new species of *Hemipsilichthys*. *Ichthyol. Explor. Freshw.* 13, 97–146.

Pereira, E. H. L. (2005). Resurrection of *Pareiorhaphis* Miranda Ribeiro, 1918 (Teleostei: Siluriformes: Loricariidae), and description of a new species from the rio Iguaçu basin, Brazil. *Neotrop. Ichthyol.* 3, 271–276. doi: 10.1590/s1679-62252005000200004

Pereira, E. H. L., and Brito, M. F. J. (2008). "*Hemipsilichthys garbei* Ihering, 1911," in *Livro Vermelho da Fauna Brasileira Ameaçada de Extinção [Brazil Red Book of Threatened Species of Fauna]*, Vol. 2, eds A. B. M. Machado, G. M. M. Drummond, and A. P. Paglia (Brasília: Ministério do Meio Ambiente), 217–218.

Pereira, L. H., Hanner, R., Foresti, F., and Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genet.* 14:20. doi: 10.1186/1471-2156-14-20

Pereira, T. L., Santos, U., Schaefer, C. E., Souza, G. O., Paiva, S. R., Malabarba, L. R., et al. (2013). Dispersal and vicariance of *Hoplias malabaricus* (Bloch, 1794) (Teleostei, Erythrinidae) populations of the Brazilian continental margin. *J. Biogeogr.* 40, 905–914. doi: 10.1111/jbi.12044

Puillandre, N., Lambert, A., Brouillet, S., and Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Mol. Ecol.* 21, 1864–1877. doi: 10.1111/j.1365-294x.2011.05239.x

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rambaut, A. (2009). *Tracer Version v1. 5.0*. Available at: http://tree.bio.ed.ac.uk/software/tracer/

Ribeiro, A. C., Lima, F. C. T., Riccomini, C., and Menezes, N. A. (2006). Fishes of the Atlantic Rainforest of Boracéia: testimonies of the Quaternary fault reactivation within a Neoproterozoic tectonic province in Southeastern Brazil. *Ichthyol. Explor. Freshw.* 17, 157–164.

Riccomini, C., Velázquez, V. F., and Gomes, C. B. (2005). "Tectonic controls of the Mesozoic and Cenozoic alkaline magmatism in central-southeastern Brazilian platform," in *Mesozoic to Cenozoic Alkaline Magmatism in the Brazilian Platform*, eds P. Comin-Chiaramonti and C. B. Gomes (São Paulo: Edusp), 31–56. doi: 10.2138/am.2007.485

Roxo, F. F., Melo, B. F., Silva, G. S., and Oliveira, C. (2017). New species of *Parotocinclus* (Siluriformes: Loricariidae) from coastal drainages of Rio de Janeiro, southeastern Brazil. *Zootaxa* 4232, 260–270. doi: 10.11646/zootaxa.4232.2.9

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M. A. X., et al. (2007). Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience* 57, 573–583. doi: 10.1641/b570707

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tavalera, G., Dimca, V., and Vila, R. (2013). Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods Ecol. Evol.* 4, 1101–1110. doi: 10.1111/2041-210x.12107

Thomaz, A. T., Malabarba, L. R., Bonatto, S. L., and Knowles, L. L. (2015). 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. *J. Biogeogr.* 42, 2389–2401. doi: 10.1111/jbi.12597

Thomaz, A. T., Malabarba, L. R., and Knowles, L. L. (2017). Genomic signatures of paleodrainages in a freshwater fish along the southeastern coast of Brazil: genetic structure reflects past riverine properties. *Heredity* 119, 287–294. doi: 10.1038/hdy.2017.46

Villa-Verde, L., Lazzarotto, H., and Lima, S. M. Q. (2012). A new glanapterygine catfish of the genus *Listrura* (Siluriformes: Trichomycteridae) from southeastern Brazil, corroborated by morphological and molecular data. *Neotrop. Ichthyol.* 10, 527–538. doi: 10.1590/s1679-62252012000300005

Voris, H. K. (2000). Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *J. Biogeogr.* 27, 1153–1167. doi: 10.1046/j.1365-2699.2000.00489.x

Ward, R. D. (2009). DNA barcode divergence among species and genera of birds and fishes. *Mol. Ecol. Resour.* 9, 1077–1085. doi: 10.1111/j.1755-0998.2009.02541.x

Wilkinson, M. J., Marshall, L. G., and Lundberg, J. G. (2006). River behavior on megafans and potential influences on diversification and distribution of aquatic organisms. *J. South Am. Earth Sci.* 21, 151–172. doi: 10.1016/j.jsames.2005.08.002

Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29, 2869–2876. doi: 10.1093/bioinformatics/btt499

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

Copyright © 2017 Lima, Berbel-Filho, Araújo, Lazzarotto, Tatarenkov and Avise. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 12 December 2017
doi: 10.3389/fgene.2017.00196

Genetic Diversity of an Imperiled Neotropical Catfish and Recommendations for Its Restoration

*Fernando S. Fonseca[1,2], Rodrigo R. Domingues[3], Eric M. Hallerman[4] and Alexandre W. S. Hilsdorf[1]**

[1] Laboratório de Genética de Organismos Aquáticos e Aquicultura, Unidade de Biotecnologia, Núcleo Integrado de Biotecnologia, Universidade de Mogi das Cruzes, Mogi das Cruzes, Brazil, [2] Instituto de Pesca, Agência Paulista de Tecnologia dos Agronegócios, Secretaria de Agricultura e Abastecimento do Estado de São Paulo, São José do Rio Preto, Brazil, [3] Instituto do Mar, Universidade Federal de São Paulo, Santos, Brazil, [4] Department of Fish and Wildlife Conservation, Virginia Polytechnic Institute and State University, Blacksburg, VA, United States

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

Reviewed by:
Carolina Machado, Federal University of São Carlos, Brazil
Natalia Martinkova, Academy of Sciences of the Czech Republic, Czechia

***Correspondence:**
Alexandre W. S. Hilsdorf
wagner@umc.br

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

Received: *01 September 2017*
Accepted: *20 November 2017*
Published: *12 December 2017*

Citation:
Fonseca FS, Domingues RR, Hallerman EM and Hilsdorf AWS (2017) Genetic Diversity of an Imperiled Neotropical Catfish and Recommendations for Its Restoration. Front. Genet. 8:196. doi: 10.3389/fgene.2017.00196

The long-whiskered catfish *Steindachneridion parahybae* (Family Pimelodidae) is endemic to the Paraíba do Sul River basin in southeastern Brazil. This species was heavily exploited by artisanal fisheries and faces challenges posed by dams, introduced species, and deterioration of critical habitat. The remaining populations are small and extirpated from some locales, and the species is listed as critically endangered in Brazil. Screening variation at a partial mitochondrial control region sequence (mtCR) and 20 microsatellite loci, we: (i) describe the patterns of genetic diversity along its current distributional range; (ii) test the null hypothesis of panmixia; (iii) investigate the main factors driving its current population structure, and (iv) propose management of broodstock for fostering recovery of wild populations through genetically cognizant restocking. Our microsatellite data for 70 individuals from five collections indicate moderate levels of heterozygosity ($H_O = 0.45$) and low levels of inbreeding ($F_{IS} = 0.016$). Individual-based cluster analyses showed clear genetic structure, with three clusters of individuals over the collection area with no mis-assigned individuals, suggesting no recent migration among the three clusters. Pairwise D_{EST} values showed moderate and significant genetic differentiation among all populations so identified. The MUR population may have suffered a recent demographic reduction. mtCRs for 70 individuals exhibited 36 haplotypes resulting from 38 polymorphic sites. Overall, mitochondrial haplotype diversity was 0.930 (±0.023) and nucleotide diversity was 0.011 (±0.002). Significant population structure was observed, with $\phi_{ST} = 0.226$. Genetic markers could be used in a hatchery-based restoration program emphasizing breeding of pairs with low kinship values in order to promote retention of genetic diversity and avoid inbreeding. Individual average kinship relationships showed 87.3% advised matings, 11.0% marginal matings, and 1.7% advised against. While these results comprise

a contribution toward planning better breeding management and monitoring, parallel actions to be undertaken include surveying healthy riverine habits for reintroduction and continued searching for wild individuals to introduce new variation into the captive broodstock to avoid adaptation to captivity and to minimize inbreeding.

Keywords: *Steindachneridion parahybae*, surubim-do-paraíba, Paraíba do Sul River, Siluriformes, population structure, conservation aquaculture

INTRODUCTION

The Atlantic Forest is a globally important biome; with high freshwater fish diversity, a high degree of endemism, and threats of many impending extinctions, it is considered a conservation hotspot (Myers et al., 2000; Hubert and Renno, 2006). An important component of the fish fauna within this biome is the genus *Steindachneridion* (Order Siluriformes: Family Pimelodidae), which encompasses six recognized species of migratory long-whiskered catfishes that inhabit deep water in mid-sized streams with rocky bottoms (Agostinho et al., 2003; Garavello, 2005).

Commonly referred to "surubim-do-paraíba," *Steindachneridion parahybae* (Steindachner, 1877) is a medium-sized, dorsoventrally flattened, whiskered catfish endemic to the Paraíba do Sul River basin in southeast Brazil (Garavello, 2005). This species was heavily exploited by artisanal fisheries along the Paraíba do Sul River in the early 1950s (Machado and Abreu, 1952). Records of historical harvest, however, are scarce, and survival of this species faces ongoing anthropogenic impacts (Hilsdorf and Petrere, 2002; Honji et al., 2009). Because of overfishing, construction of dams, introduced species, and deterioration of critical habitat (Loris, 2008; Moraes et al., 2017), the species is listed as "Critically Endangered" in the official lists of imperiled species in Brazil (Caneppele et al., 2008; Oyakawa et al., 2009).

Population decline is a prominent concern for conservation genetics because of potentially harmful consequences for species. Indeed, small and fragmented populations can undergo changes such as loss of genetic variation, fixation of deleterious alleles, inbreeding, and reduced fitness (Frankham et al., 2009). Critically, such outcomes can limit the ability of a given species to adapt to future environmental changes, increasing the risk of extinction (Frankham, 2005). Molecular markers are powerful tools for quantifying genetic variation at the individual, population, and species levels, and screening of markers can contribute to identification of units of relevance for the management and conservation of target species (Allendorf et al., 2010; Chauhan and Rajiv, 2010). Screening of molecular markers in studies of Neotropical freshwater fishes has contributed to assessment and conservation of fisheries genetic resources (Piorski et al., 2008; Hilsdorf and Hallerman, 2017). In particular, assessing genetic diversity and characterizing population structure is of primary importance for framing management strategies to minimize the likelihood of population extinction.

Once abundant in artisanal fisheries, *S. parahybae* is currently on the verge of extinction, with only a few remaining small and isolated populations (Honji et al., 2017). Management to foster recovery of the species is imperative due to its ecological importance as an apex predator and its economic importance for artisanal fisheries. Against this background, screening variation at a partial mitochondrial control region sequence (mtCR) and 20 microsatellite loci, we address for the first time issues important to the genetic conservation of the endemic Neotropical catfish *S. parahybae*. Specifically, we: (i) describe the patterns of genetic diversity of the *S. parahybae* along its current distributional range; (ii) test the null hypothesis of panmixia in this region; (iii) investigate the main factors driving its current population structure, and (iv) propose *ex situ* genetic breeding management for recovery of wild populations through genetically cognizant restocking programs.

MATERIALS AND METHODS

Ethics Statement

All field work for fish sampling complied with the legal regulations of Brazil (collection permit SISBIO: 22140, Project: P&D CESP/ANEEL: 0061-017/2006).

Sampling Sites and DNA Extraction

A total of 70 individual *S. parahybae* were captured through an extensive occurrence survey along the tributaries of the Paraíba do Sul River basin in Brazil. Sites of occurrence were located with the help of local artisanal fishers and the Companhia Energética de São Paulo (São Paulo State Electric Power Company, CESP). Wild *S. parahybae* were collected at five different locations (**Figure 1**) within the Paraíba do Sul River basin in 2004 and between 2010 and 2015, consisting of: 1 individual from the Paraíba do Sul River (PSP) in São Paulo State (22°34′S; 44°53′W), 24 individuals from the Paraíba do Sul River (PRJ) in Rio de Janeiro State (22°13′S; 43°25′W), 5 individuals from the Preto River (PRE) in Rio de Janeiro State (22°00′S; 43°20′W), 3 individuals from the Pomba River (POM) in Rio de Janeiro State (21°32′S; 42°09′W), and 37 individuals from the Muriaé River (MUR) in Rio de Janeiro State (21°12′S; 41°55′W). All individuals were transferred live to aquaculture facilities belonging to CESP and kept as an *ex situ* germplasm bank (Supplementary Images S1, S2), where they were fin-clipped and individually marked with passive integrated transponder tags.

Total genomic DNA was extracted from approximately 20 mg of fin-clip tissue using a phenol-chloroform protocol adapted from Taggart et al. (1992) using STE (0.1 M NaCl, 0.05 M Tris-HCl, 0.01 M EDTA) with a lower concentration of EDTA buffer.

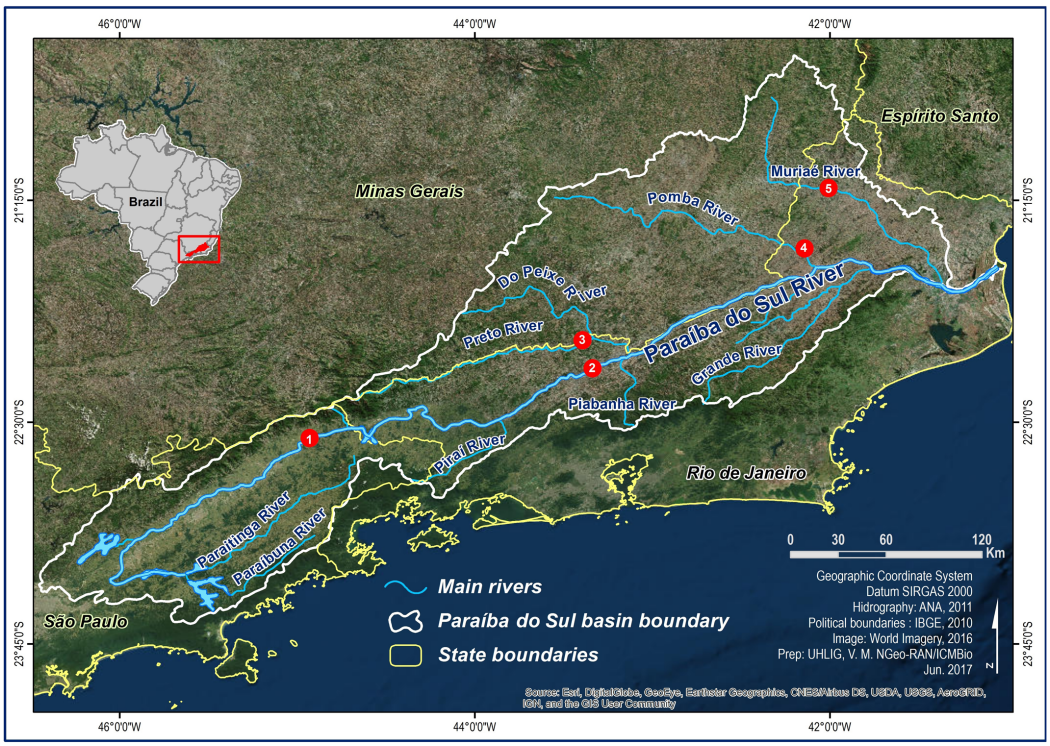


FIGURE 1 | Sampling sites for *Steindachneridion parahybae* in the Paraíba do Sul River basin. 1 – Paraíba do Sul River – SP (PSP), 2 – Paraíba do Sul River – RJ (PRJ), 3 – Preto River - RJ (PRE), 4 – Pomba River – RJ (POM), 5 – Muriaé River – RJ (MUR).

Microsatellite Genotyping and Data Analysis

Genetic variability was screened at 20 di-, tri-, and tetra-nucleotide microsatellite loci (*Spa2, Spa3, Spa4, Spa5, Spa6, Spa7, Spa8, Spa11, Spa12, Spa14, Spa15, Spa16, Spa17, Spa18, Spa19, Spa20, Spa22, Spa23, Spa28,* and *Spa42*) described by Ojeda et al. (2016) using primer pairs developed by those authors. Sequences of all microsatellites were deposited in GenBank under accession numbers KU821557 to KU821572 and KX578067 to KX578070. Microsatellite alleles were genotyped using a 6.5% denatured polyacrylamide gel matrix (KB plus 6.5% Matrix Gel, Li-Color Biosciences, Lincoln, NE, United States) and a Li-Cor 4300 DNA Analyzer (IR2, Lincoln, NE, United States) using the IRDye® 700 marker and universal M13 tail primer described by Schuelke (2000). The sizes of alleles were estimated by interpolating their position relative to molecular weight markers (50–350 bp DNA Sizing Standard IRDye® 700) using the SagaGT Client program (Li-Cor Biosciences, Lincoln, NE, United States). All recommendations for minimizing genotyping errors originating from DNA quality and handling procedures (Pompanon et al., 2005) were implemented, including semi-automated scoring followed by visual inspection by two independent people.

The presence of null alleles and allelic dropout were tested for using MICRO-CHECKER 2.2.3 (Van Oosterhout et al., 2004). Deviations of genotype frequencies from Hardy–Weinberg expectations (HWE) were calculated by an exact test using the Markov chain randomization approach (Guo and Thompson, 1992) with HW-QUICKCHECK (Kalinowski, 2006). The significance for both types of tests was assessed employing a Markov chain Monte Carlo (MCMC) procedure based on 1,000 dememorizations, 100 batches, and 10,000 iterations per batch, and the critical value for significance was adjusted for multiple tests using the Holm–Bonferroni method (Holm, 1979). The nuclear diversity for each collection area and locus was quantified as the number of alleles per locus (A), allelic richness (Ar), observed heterozygosity (H_O), expected heterozygosity (H_E), and fixation index (F_{IS}) using the Adegenet package (Jombart, 2008) in R v 3.1.2 (R Development Core Team, 2014). An individual assignment approach was performed using the Neighbor-Joining (NJ) algorithm using Euclidian distance between matrices of allele frequencies using the Adegenet package in R. In addition, discriminant analysis of principal components (DAPC, Jombart et al., 2010) was used to assign individuals to genetic cluster(s), using the Adegenet package in R. This multivariate approach was designed to identify and describe clusters of genetically related individuals in a manner similar to that performed by the Structure package (Pritchard et al., 2000). DAPC has the advantage that it does not require populations to be in Hardy–Weinberg equilibrium or linkage equilibrium between loci (Jombart et al., 2010). The optimal number of principal components retained followed the indication of α scores. The cluster assignments were pre-defined to correspond to *a priori* defined sample locations. Population genetic structure was estimated among clusters identified by DAPC using the

population pairwise differentiation index D_{EST} (Jost, 2008) using the DEMETICS package in R (R Development Core Team, 2014). Holm–Bonferroni (Holm, 1979) adjustments of α were used to correct for multiple tests.

The program BOTTLENECK 1.2.02 (Cornuet and Luikart, 1997) was used to evaluate the possibility of changes in the recent effective size of a population. In this analysis, heterozygote excesses were checked using three different models: the infinite alleles model (IAM), stepwise mutation model (SMM), and the two-stage intermediate model (two-phase microsatellite evolution mode – TPM) with 70% SMM and 30% IAM. The two-phase model (TPM) is most powerful when fewer than 20 loci are used (Piry et al., 1999). The probability of significant heterozygosity excess was determined using 10,000 replications and a one-tailed Wilcoxon signed-rank tests ($\alpha = 0.05$).

To guide design of matings for the restocking program, we performed average kinship assessment using the program COANCESTRY version 1.0.1.5 (Wang, 2011). This analysis is based on allele frequencies in the generation of matrices (F_0) using means and variances of kinship coefficients. Three relationship categories (unrelated, half-sibling, and full-sibling) were used, with 5,000 pairs of simulated individuals for each category. In this evaluation, seven r_{xy} (KE) kinship estimators were used as follows: trioml (triadic likelihood estimator – Wang, 2007), wang (Wang, 2002), lynchLi (Li et al., 1993), lynchrd (Lynch and Ritland, 1999), ritland (Ritland, 1996), quellergt (Queller and Goodnight, 1989), and dyadml (dyadic likelihood estimator – Milligan, 2003). In addition, we applied the ML-Relate package (Kalinowski et al., 2006) to make a conservative estimation of the level of kinship. We then compared the results of both approaches.

Mitochondrial DNA Amplification, Sequencing, and Data Analysis

A partial mitochondrial DNA control region (mtCR) was amplified by polymerase chain reaction (PCR) using two external primers, SPf (5′TCTAACTCCCAAAGCTAGAATC3′) and SPr (5′GGAACTTTCTAGGGTTCATCTTAAC3′), developed in this study. Amplification was performed in a 20-μl reaction containing: 20 ng/μl of genomic DNA, 0.5 IU *Taq* DNA polymerase (Thermo Fisher Scientific, Inc., São Paulo, Brazil), 1x buffer (100 mM Tris-HCl pH 8.8, 500 mM KCl), 1.5 mM $MgCl_2$, 2.5 mM dNTPs, 10 mM of each primer, and ultrapure water. The PCR thermal-cycling profile consisted of initial denaturation at 94°C for 3 min; followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 53°C for 1 min, and extension at 72°C for 1:30 min; and a final extension step at 72°C for 10 min.

The amplicons were purified using ExoSAP enzymes (Affymetrix, Cleveland, OH, United States) following the manufacturer's instructions, except that the shortest time for deactivation of the enzymes was 15 min. The amplicons were sequenced on an Applied Biosystems 3130 Genetic Analyzer, and bases were called using Applied Biosystems Sequencing Analysis Software version 5.2. The sequences were edited manually when necessary using CodonCode (v.5.1.5, Codon Code Corporation, Centerville, MA, United States).

Summary statistics for genetic diversity – including haplotype (*Hd*) and nucleotide (π) diversities, number of haplotypes (*h*) and polymorphic sites (*S*) – were calculated in DnaSP v. 5.10 (Librado and Rozas, 2009). The relationships among haplotypes and their geographic distribution were assessed using the median-joining algorithm in NETWORK 4.6.1.1. To test the null hypothesis of matrilineal panmixia of *S. parahybae* along the Paraíba do Sul River basin, ϕ_{ST} (Weir and Cockerham, 1984) was calculated by analysis of molecular variance (AMOVA, Excoffier et al., 1992). To assess genetic structure across the collections, 10,000 permutations of the data were used to test the significance of hierarchical differentiation using Arlequin version 3.5.1.3 (Excoffier and Lischer, 2010).

The estimated probabilities were corrected using Holm–Bonferroni sequential adjustments for multiple tests (Holm, 1979). Due to small sample sizes at the PSP ($n = 1$), PRE ($n = 5$), and POM ($n = 3$) sites, data from those collection areas were included only to provide preliminary insights into how these collections may be related to the others.

Two neutrality metrics were calculated to infer the population demographic history of *S. parahybae*, R_2 (Ramos-Onsins and Rozas, 2002) and F_S (Fu, 1997). Both metrics were calculated and their statistical significance tested using 10,000 permutations under the coalescent process implemented in DnaSP v.5.10 (Librado and Rozas, 2009).

RESULTS

Microsatellite Intra- and Inter-population Diversity

All samples of *S. parahybae* from the five collection areas along the Paraíba do Sul River basin were successfully genotyped at 20 microsatellite loci. All loci within all collections were in Hardy–Weinberg equilibrium after correction for multiple tests. The linkage disequilibrium tests presented significant values after sequential Bonferroni correction ($p < 0.05$) for seven pairs of loci in a single population (MUR), which may be the consequence of many factors, including limited numbers of parents in the preceding generation, as well as natural selection, gene flow, assortative mating, and linkage. There was no evidence of genotyping errors, such as stuttering, null alleles, or large allele drop-out. The highest number of alleles (13) was amplified at the *Spa18* locus, while *Spa3*, *Spa7*, *Spa16*, *Spa23*, and *Spa28* all exhibited only two alleles. The number of alleles per locus within particular populations ranged between 2 and 13, with a mean of 4.7 ± 3.2. The allelic richness values ranged between 6.1 and 10.4, with average of 2.65 ± 2.07. The observed heterozygosity values ranged between 0.0 and 0.89, with average of 0.45 ± 0.24. Private alleles were observed in the MUR (14), PRE (1), and PRJ (2) collections. The F_{IS} inbreeding coefficient values ranged between −0.002 and 0.356, with average of 0.016. The summary statistics for all loci are shown in Supplementary Table S1.

Results of the individual-based DAPC and NJ cluster analyses showed clear genetic structure, with three clusters of individuals over the entire collection (**Figure 2**). Overall, individuals from the MUR and PRJ sites showed a 100% likelihood of assignment

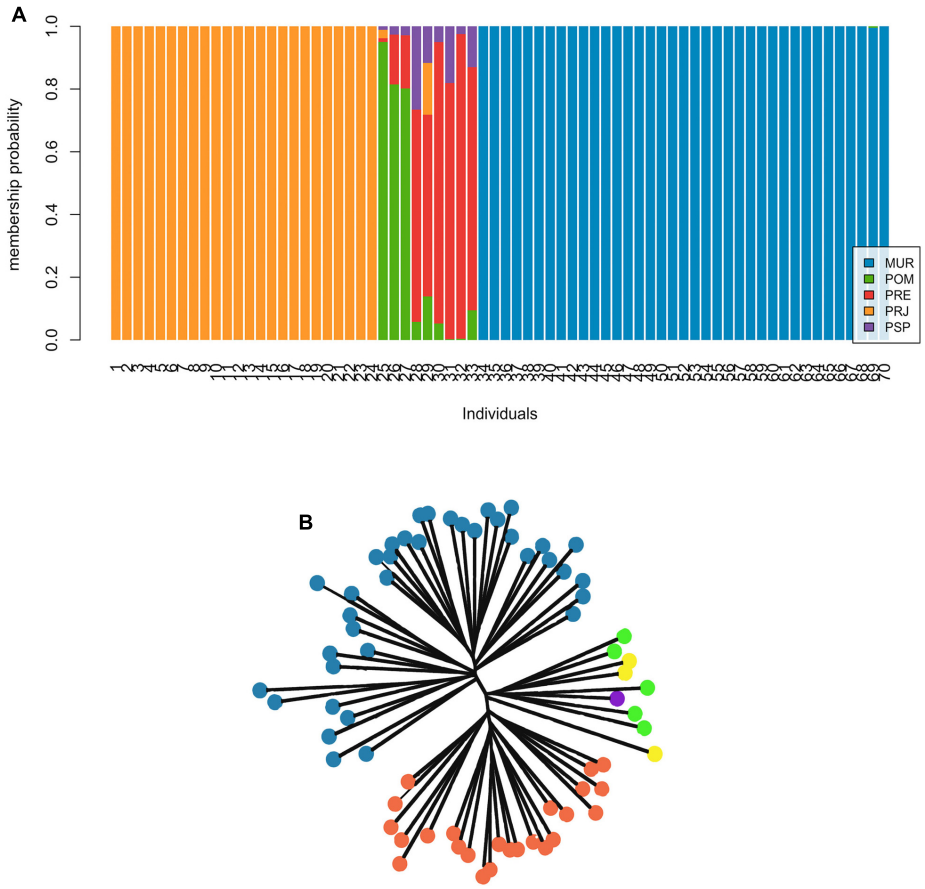


FIGURE 2 | Different individual assignment approaches in *Steindachneridion parahybae*. **(A)** Results of the Bayesian approach with the use of discriminant analysis of principal components (DAPC) to investigate genetic structure. The ideal number of major components found for the analysis was seven based on the α scoring trend. Each vertical bar represents an individual in the DAPC diagram and each color represents the probability of belonging to one of the genetic groups. **(B)** Unrooted tree from a Neighbor-joining analysis of distance data.

to their original collection area (hereafter termed populations). Individuals from POM, PSP, and PRE collections present a group of individuals of mixed ancestry (**Figure 2A**). Even though they do not form a same genetic cluster, we decided to assume a single cluster due to the low sample number. Neighbor-joining analysis (**Figure 2B**) also corroborated the three clusters of individuals: MUR, PRJ, and the POM/PSP/PRE sites. No mis-assigned individuals were observed, suggesting that there may have been no recent migration among these well-defined clusters.

The genetic differentiation indices were highly significant ($D_{EST} = 0.082$, $p = 0.001$). The pairwise D_{EST} values for populations identified previously by NJ and DAPC analysis were congruent with the results of individual-based analysis, showing moderate and significant genetic differentiation among all populations (**Table 1**).

In silico analyzes performed in BOTTLENECK 1.2.02 showed significant values ($p < 0.05$) for the MUR population using the SMM mutation model (**Table 2**). However, there were no significant values for TPM model, which is more powerful for fewer than 20 loci (Piry et al., 1999). These results suggest that among all populations, only the MUR population may have suffered recent demographic reduction. Because the POM, PSP, and PRE collections populations had few specimens, assessment of demographic bottlenecks in these populations had insufficient power.

TABLE 1 | Pairwise mtDNA ϕ_{ST} values (above diagonal) and pairwise Jost's D_{EST} values using microsatellites (below the diagonal) among three population clusters of *Steindachneridion parahybae*.

Localities	PRJ	POM, PSP, and PRE	MUR
PRJ	–	0.1860*	0.3044*
POM, PSP, and PRE	0.0401*	–	−0.0256^{ns}
MUR	0.0688*	0.0529*	–

^{ns}*Not significant, *highly significant ($p < 0.01$).*

Mitochondrial Partial Control Region Sequence Variation

A total of 791 bp of the mtCR were resolved from 70 *S. parahybae*. Sequences of all haplotypes were deposited in GenBank under accession numbers MG012754–MG012789. The

TABLE 2 | Results of the BOTTLENECK 1.2.02 program (Cornuet and Luikart, 1997) used to evaluate the possibility of changes in the effective size of the population.[1]

Localities	p IAM	*SD*	p SMM	*SD*	p TPM	*SD*
PRJ	0.387	0.075	0.074	0.034	0.344	0.448
POM	0.563	0.491	0.408	0.294	0.564	0.442
PRE	0.434	0.316	0.119	0.175	0.391	0.373
MUR	0.365	0.047	**0.034**	0.000	0.356	0.404

[1] *p is the probability value, with significant values ($p < 0.05$) shown in bold; IAM, the infinite alleles model; SMM, the stepwise mutation model; TPM, two-phase model of microsatellite evolution mode; and SD, standard deviation.*

sequences exhibited 36 haplotypes (Supplementary Tables S2, S3) resulting from 38 polymorphic sites and 41 mutations (34 transitions and 7 transversions). The nucleotide composition was 34.14% A, 19.66% C, 20.98% G, and 25.23% T. The overall *h* and π diversities were 0.930 (±0.023) and 0.011 (±0.002), respectively. The haplotype diversity values ranged from 0.868 ± 0.049 (MUR) to 0.99 ± 0.272 (POM), and nucleotide diversity values ranged from 0.005 ± 0.004 (POM) to 0.013 ± 0.007 (PRJ) (**Table 3**). Haplotype H2 ($n = 13$) was the most common, occurring mainly in MUR, and shared by all populations except for PSP (Supplementary Table S3). A total of 33 haplotypes were singletons, 16 occurring in MUR and 12 in PRJ. An unrooted median-joining tree of mtCR haplotypes (**Figure 3**) showed mainly singleton haplotypes in populations MUR and PRJ. Haplotype H2 was most frequent haplotype, shared by all *S. parahybae* populations except for PSP. Significant population structure was observed for *S. parahybae* along the Paraíba do Sul River basin based on mtCR data ($\phi_{ST} = 0.226$, $p < 0.05$).

Fu's F_S metric presented high values, but there were no significant values (**Table 3**). In contrast, the R_2 values were high and significant ($p < 0.05$), indicating a significant departure from neutrality. According to Ramos-Onsins and Rozas (2002), the R_2 test is more powerful in detecting demographic events in small sample sizes, whereas Fu's F_S is better for large sample sizes.

TABLE 3 | Mitochondrial DNA sampling and descriptive statistics for *Steindachneridion parahybae*.

Sampling localities	*N*	*S*	*H*	*h* ± *SD*	π ± *SD*	F_S	R_2
MUR	37	25	19	0.868 ± 0.049	0.008 ± 0.004	0.353^{ns}	0.162*
PRE	5	17	4	0.900 ± 0.161	0.011 ± 0.007	0.388^{ns}	0.162*
PRJ	244	34	14	0.935 ± 0.031	0.013 ± 0.007	0.377^{ns}	0.163*
PSP	1	0	1	–	–	–	–
POM	37	6	3	0.99 ± 0.272	0.005 ± 0.004	0.353^{ns}	0.161*

(N) = number of sequenced individuals; (S) = number of polymorphic sites; H = haplotypes; (h) haplotype diversity; (π) nucleotide diversity; (SD) standard deviation; and F_S (Fu, 1997) and R_2 (Ramos-Onsins and Rozas, 2002) indices of neutrality. ^{ns}*Not significant, *highly significant ($p < 0.01$).*

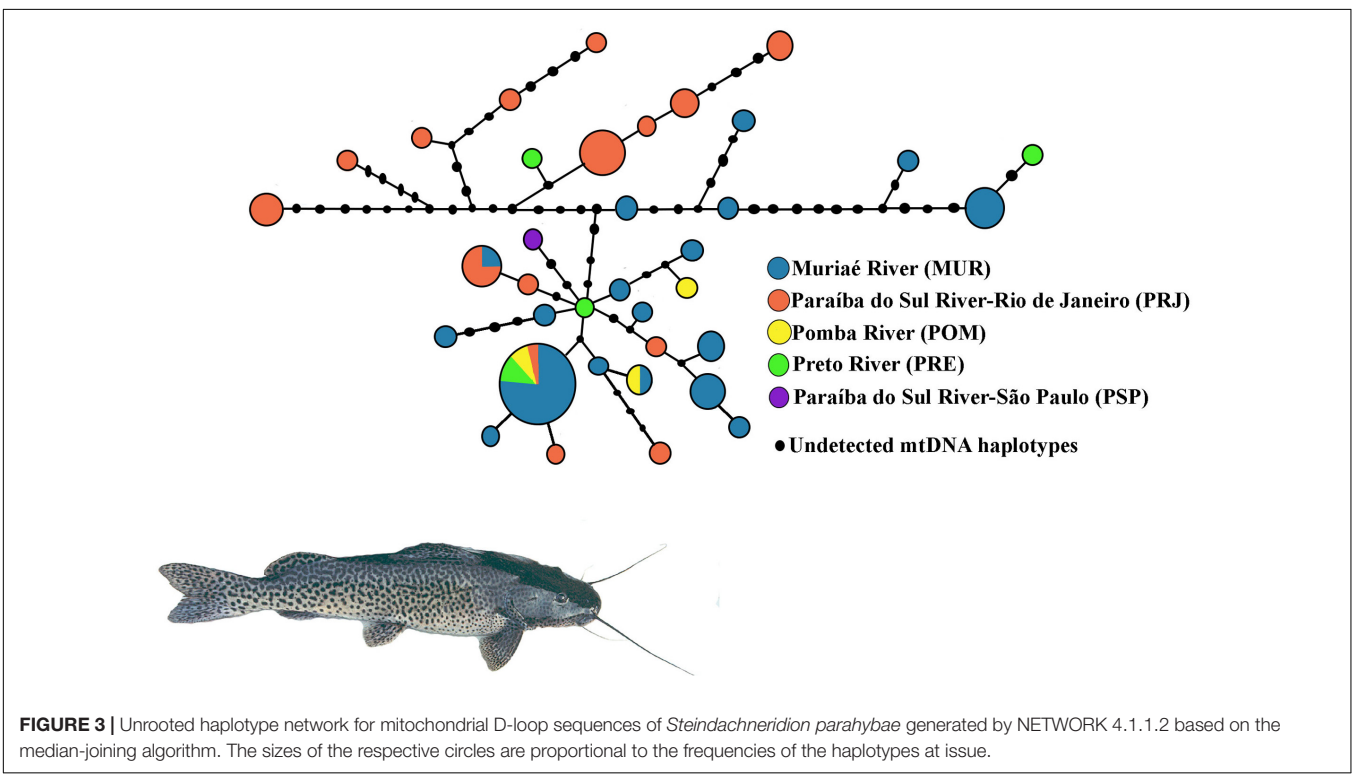


FIGURE 3 | Unrooted haplotype network for mitochondrial D-loop sequences of *Steindachneridion parahybae* generated by NETWORK 4.1.1.2 based on the median-joining algorithm. The sizes of the respective circles are proportional to the frequencies of the haplotypes at issue.

Genetic Relatedness Among *Steindachneridion parahybae* in the *Ex Situ* Germplasm Bank

Among prospective broodstock in the *ex situ* germplasm bank, sampling variances for the kinship estimators (KE) ranged from 0.0198 to 0.1046. The lowest variance was found in the triadic likelihood estimator (TrioML). Considering that the KEs had high correlations and little influence on the inference of relatedness, paired kinship coefficients are reported only for the TrioML. Values estimated by the TrioML of kinship were low, indicating a low inbreeding index for 67 individuals, and thus, a range of permissible pairings among the broodstock candidates. Only three individuals with the highest values of inbreeding were classified as moderately endogamous. Means and variances for kinship likelihood estimators are showcased in Supplementary Table S4.

When we matched the 44 females with the 21 males among the broodstock candidates, we obtained 924 interactions that were added to the possible interactions with the five individuals of indeterminate sex (MUR12, MUR15, MUR20, MUR28, and MUR37), obtaining 1,269 possible matings. The index obtained for kinship values was split into three categories according to expected average theoretical values (Wang, 2011). That is, the relatedness values for simulated pairs were considered to be high above 0.5 (full-sib and parent-offspring), intermediate between 0.25 and 0.5 (half-sib or other kinship), and low below 0.25. Considering all possible matings, the calculated kinship relationships indicated 87.3% (1,108/1,269) advisable matings, 11.0% (140/1,269) marginal matings, and 1.7% (21/1,269) inadvisable matings. A table of consensus outcomes from the COANCESTRY and ML-Relate analyses is showcased in **Table 4**. Supplementary Tables S5, S6 show the results of the respective analyses.

DISCUSSION

Genetic Diversity and Population Structure

Our assessment of the extant *S. parahybae* populations reflected the endangered status of this Neotropical catfish in its area of occurrence. Attempts to locate wild populations proved the first indicator of the imperiled status of *S. parahybae* in the Paraíba do Sul watershed. Many efforts have been made during recent years to find and collect *S. parahybae* individuals to be kept in an *ex situ* gene bank so that a genetically cognizant restocking program (Miller and Kapuscinski, 2003) could be implemented. However, relatively few specimens have been found and collected until now. Although intensive efforts were made to collect specimens in localities where their occurrence was recorded by local anglers, at sites POM, PRE, and PSP, just a few individuals were collected; in the case of PSP, just one individual was collected in 2004. According to Honji et al. (2009), this species is regionally extinct in São Paulo State. At one collection site (PRJ), 24 individuals were collected in 2004 before an organochlorine insecticide (Endosulfan) spill into the Paraíba do Sul River killed fish and other riverine animals (Brito et al., 2012). After that, several unsuccessful attempts to capture *S. parahybae* have been made; hence, the species was likely extirpated at this site. The MUR is a locality where the species seems to be present, although in a small population. Despite the difficulty of catching this species in artisanal fisheries, local fishermen report the rare presence of *S. parahybae* in the Muriaé River.

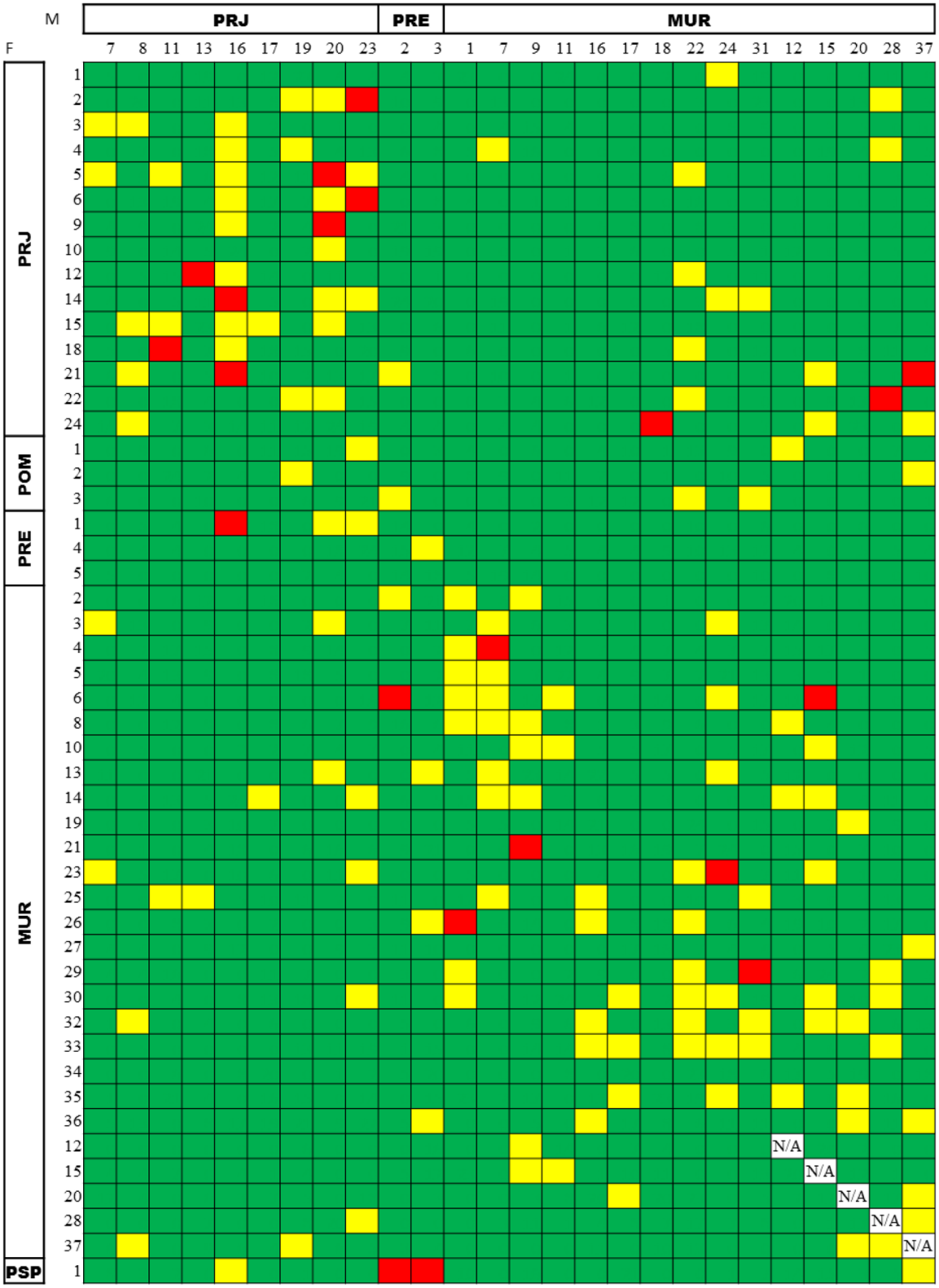
Assessing the genetic differentiation of wild populations and individuals kept in *ex situ* gene banks is pivotal for developing conservation strategies to save a given species from extinction (FAO, 2008). Summary statistics for genetic diversity for the 20 microsatellite loci screened are shown in Supplementary Table S1. Heterozygosity ($H_E = 0.47$) and number of alleles per locus ($A = 4.7$) for wild *S. parahybae* populations were low across the 20 loci screened in this study.

Despite this low genetic variability, the inbreeding index ($F_{IS} = 0.016$) was below the average value ($F_{IS} = 0.05$) for 18 endangered animal species kept in *ex situ* conservation programs (Witzenberger and Hochkirch, 2011). The relatively low values of genetic diversity at microsatellite loci, a molecular marker type whose dynamics are more closely associated with contemporary events (Selkoe and Toonen, 2006; Weese et al., 2012), could be associated with the current state of fragility of populations likely impacted by anthropogenic habitat alterations and overfishing. These findings were corroborated with *in silico* analyses performed in BOTTLENECK 1.2.02.

The discreteness of populations of *S. parahybae* currently along the Paraíba do Sul River basin evident by analyses of microsatellite loci and mitochondrial DNA sequencing may be the result of isolation due to environmental alterations or to the characteristic short-distance reproductive migration behavior of this catfish species. No data are specifically available about the migration intensity of *S. parahybae*, but studies on catfishes in the Uruguay River, including the congeneric *S. inscripta*, show apparent short-distance migration (Zaniboni-Filho and Schulz, 2003). As the mobility of species is a factor determining the degree of homogenization of allele frequencies by gene flow (Avise, 2004), the contemporary genetic structure found for this species may then be a combination of short migration and river disruption by dams, which may act as barriers for gene flow (Dehais et al., 2010; Roberts et al., 2013; Gouskov et al., 2016). Indeed, this recent effect may be supported by the observed differentiation values for the nuclear markers, indicating low to moderate differentiation within the contemporary period.

The differentiation shown for mitochondrial D-loop sequences ($\phi_{ST} = 0.226$) and microsatellites ($D_{EST} = 0.082$) can be interpreted as reflecting historical and recent structuring, especially for the female component of the population, with some degree of homogenizing gene flow for males, with or without the effects of recent anthropogenic isolation. Cluster analyses (**Figure 2**) showed considerable population structure, indicating the division of the *S. parahybae* gene pool into three distinct groups, MUR, PRJ, and a mixed group formed by POM, PSP, and PRE. The strongly supported assigned clusters for MUR and PRJ ($q > 0.9$) denoted virtual absence of migration between the two groups. The identification and characterization of an Evolutionary Significant Unit (ESU), based on population genetic

TABLE 4 | Consensus pairwise relatedness of broodstock candidates among the *Steindachneridion parahybae* individuals maintained in the germplasm bank, as assessed by both the COANCESTRY (TrioML estimator) and the ML-Relate packages.



Crosses in red and yellow cells should be avoided due to high or moderate level of kinship. Shown in green cells are the consensus best crosses for producing offspring for restocking due to low kinship. Females are denoted (F) and males (M). N/A – data not available, for individuals of indeterminate sex.

structure and adaptive differentiation, would provide a useful basis for conservation, as well as a more complete understanding of the role of adaptive evolution in the natural history of the species (Casacci et al., 2014). Carvalho et al. (2012) reported two ESUs within the widely distributed *Pseudoplatystoma corruscans*, a commercially important Neotropical siluriform catfish. However, despite the presence of private alleles in the MUR sample, populations of *S. parahybae* may more appropriately be considered Management Units (MUs, Moritz, 1994), which can be defined as populations with significant divergence of allele frequencies at nuclear or mitochondrial loci, regardless of the phylogenetic distinctiveness of the alleles. Taking another approach, Palsbøll et al. (2007) proposed that MUs should be determined by the level of interpopulation genetic divergence in general; however, the threshold values to establish populations' genetic connectivity must be flexible to reflect each specific biological and conservation context. This is particularly the case for the Muriaé River population, where *S. parahybae* individuals yet seem to occur.

Phylogeography and Demographic History

Despite the difficulty of comparing among studies the average divergence between mtDNA sequences due to lack of standardization of sequence length, small nucleotide diversity combined with considerable haplotype diversity (Grant and Bowen, 1998) might suggest a demographic expansion after a period of low effective population size for *S. parahybae* populations. Mesquita et al. (2001) assessed *Chondrostoma lusitanicumem,* an endangered fish species in Portugal, and found the same combination of molecular genetic variation among marker types, suggesting that high mitochondrial haplotype diversity with low nucleotide divergence may indicate non-equilibrium evolutionary action or non-neutral forces such as "founder-flush" population differentiation or even the persistence of slightly deleterious mutations in small populations due to ineffectiveness of "purifying" selection. In the present study, the hypothesis of demographic expansion is supported by results of the R_2 test, a test to detect population growth. The positive, significant values refuted the null hypothesis of constant population size under the neutral model. In addition, the haplotype network (**Figure 3**) does not have the star-like shape typical of populations that show a signature of a species that has expanded its numbers from a modest number of founders (Avise, 2000). It is likely that throughout the evolutionary history of *S. parahybae*, different mitochondrial lineages colonized and developed in each locality of the basin, and that carriers of distinct mitochondrial lineages migrated along the basin when no artificial barriers, such as dams, were present. Such dams are barriers that prevent the reproductive migration of many fish species (Pelicice et al., 2015). As human occupation expanded along the watershed, overfishing, pollution, introduction of species, and dam construction for electric power generation impacted *S. parahybae* population size and geographic range, accompanied by the local and global loss of many mitochondrial haplotypes. For instance, the only specimen captured from PSP has a unique haplotype at the terminus of a branch of the network, not present in any other sampling site in this survey. In addition, the existence of many different mitochondrial haplotypes in the MUR and PRJ populations, many separated by inferred mutational differences not represented by observations in living individuals, supports this interpretation.

On the other hand, the result of the BOTTLENECK test suggests that the MUR population was subject to a recent population bottleneck. This result is supported by observation of excess heterozygosity in the population as a whole. The bottleneck process, which the extant *S. parahybae* population may have undergone, suggests the loss of rare alleles during the population reduction. The loss of alleles may ultimately jeopardize the population due to loss genetic variability crucial for meeting future ecological challenges; however, the presence of private alleles may also suggest local adaptation (Santos et al., 2016). The absence of *S. parahybae* individuals in the artisanal fisheries in localities where they were common as recently as the early 1950s (Machado and Abreu, 1952) and in scientific samples at sites of previous occurrence in the Paraíba do Sul River basin clearly suggests that *S. parahybae* populations underwent a genetic bottleneck process and that some populations were extirpated.

A Restoration Genetics Approach to Recovery of *S. parahybae* Populations

Restocking of freshwater fishes is often conducted to restore populations impacted by the construction of hydroelectric power dams (Marmulla, 2001; Palmer et al., 2007; Piorski et al., 2008). While the effectiveness of the approach has been questioned (Snyder et al., 1996; Fraser, 2008; Neff et al., 2011), restoration of imperiled or extirpated populations must be considered when other efforts have proven unsuccessful (Attard et al., 2016). Restoration genetics attempts to use the tools of genomics to support proper management of captive populations and propagation of genetically healthy fingerlings (Frankham, 2010). Human-driven environmental changes in the Paraíba do Sul watershed date to the onset of agricultural production (sugarcane in the 17th century, coffee in the late 18th–19th centuries), and industrialization from the 1950s (Pádua, 2017). More than 120 hydropower stations in the Paraíba do Sul River and its tributaries have disrupted fluvial connectivity, leading to losses of local populations and species (Agostinho et al., 2008; Loris, 2008). *Steindachneridion parahybae* is among the endemic fishes of Paraíba do Sul negatively impacted by human actions, and is listed as threatened by the environmental agencies of the Brazilian government (Honji et al., 2017). The recovery of these imperiled species has been promoted by establishing *ex situ* germplasm banks and genetics-based captive breeding programs. The reintroduction programs for *S. parahybae* follow to some degree the framework proposed by Attard et al. (2016), which accounts for past, present, and future components of captive-breeding and reintroduction. The geographical localization and genetic characterization of extant *S. parahybae* populations described herein, and their tagging and maintenance in the *ex situ* gene bank facilities encompass the past and present components of the holistic framework for successful reintroductions.

Successful captive breeding depends upon keeping levels of inbreeding low and optimizing genetic variability so that fingerlings intended for reintroduction reflect as much as possible the genetic makeup of the species. To achieve this goal, we assessed the genetic relatedness of all captively bred individuals to generate guidance for the hatchery manager to select genetically unrelated fish to be crossed (Saura et al., 2013). A set of 20 species-specific microsatellite markers (Ojeda et al., 2016) was used successfully to assess the mean kinship within and among the remaining populations of *S. parahybae*, generating relatedness indices between each captive individual. Keeping the minimum number of founders to minimize a loss of genetic diversity, avoiding both inbreeding and outbreeding depression, and searching for new wild populations to constantly renew the captive stock are all pivotal to the success of *ex situ* conservation projects (Miller and Kapuscinski, 2003; Witzenberger and Hochkirch, 2011; IUCN/SSC, 2013). Neff et al. (2011) argued that just taking into account multilocus heterozygosity to measure the likely effectiveness of reintroduction is too simplistic. According to these authors, it is more important to come to better comprehend the complexity of the genetic architecture of fitness, so that the importance of genetic diversity to the evolutionary potential of populations of fishes and their subsequent adaptation after reintroduction can be gauged. Therefore, they recommended that effective management and conservation of populations in imminent peril of extirpation should include: (i) rehabilitation and maintenance of habitat and ecological function; (ii) incorporation of natural ecological processes into artificial breeding programs, and (iii) use of a full or partial factorial breeding design in artificial breeding programs. This last recommendation regarding breeding design is intended to maximize the amount of genetic differentiation throughout the species' occurrence.

Use of estimators for pairwise relatedness and individual inbreeding coefficients based on multilocus microsatellite markers has proven an efficient strategy to achieve better responses in reintroduction programs. Sriphairoj et al. (2007), working with the critically endangered Mekong giant catfish *Pangasianodon gigas* tested different scenarios regarding different broodstock recruitment and mating strategies. The authors suggested that using minimal kinship in a random mating scheme could keep the effective population size (N_e) at 100 so that allelic diversity can be retained at above 90% of current allelic diversity for at least four generations. In the present study of endangered *S. parahybae* populations, we used this same approach, also based on multilocus microsatellite markers, to genetically characterize the few remaining wild populations and broodstock currently maintained in an *ex situ* germplasm bank. These genetic data comprise a contribution toward planning better breeding management and monitoring any changes in the genetic makeup of the broodstock, as well as in the reintroduced fingerlings in their new environments. Parallel actions to be taken would include surveying healthy riverine habits for reintroduction and continued searching for wild individuals to introduce new variation into the captive broodstock to avoid adaptation to captivity and to prevent inbreeding. Successes in other genetic marker-assisted restoration programs suggest that application of this approach might also prove effective for *S. parahybae*. Olsen et al. (2000) used microsatellite DNA markers to identify a targeted stock of pink salmon *Oncorhynchus gorbuscha* in a supportive breeding program on the Dungeness River in Washington State, United States, providing the basis for hatchery-based supplementation. Following cessation of stocking from outside sources, microsatellite marker-assisted hatchery-based supplementation of walleye *Sander vitreus* (Palmer et al., 2007) contributed to a rebuilding of the native population in the New River, Virginia, United States and its recognition as a premier recreational fishery for the species.

AUTHOR CONTRIBUTIONS

AH conceived and supervised the project. FF performed the experiments, analyzed data, and drafted the manuscript, RD performed statistical analyses and contributed to manuscript writing, AH and EH wrote and critically edited the final manuscript. All authors read and approved the final version of the manuscript.

FUNDING

The authors thank the São Paulo Research Foundation (FAPESP # 11/23752-2) and AGEVAP (#10/2012) for granting funds to support this project, as well as FUNDEPAG and Instituto de Pesca for funding FF to develop this work.

ACKNOWLEDGMENTS

The authors thank Danilo Caneppele from the Companhia Energética de São Paulo (CESP) for access to the *S. parahybae* samples kept in the CESP *ex situ* germplasm bank and Dr. Guilherme Gui from Fundação Piabanha for the great effort to find and sample *S. parahybae* in rivers of the Paraíba do Sul watershed in the State of Rio de Janeiro. They also thank Vívian Uhlig, M.Sc. from ICMBIO, for the map drawing, and Leo Menino for helping to format figures. Funding for EH's participation in this work was provided in part by the Virginia Agricultural Experiment Station and the United States Department of Agriculture National Institute of Food and Agriculture. This work was developed as part of the full requirements for the doctoral dissertation of FF in biotechnology at the University of Mogi das Cruzes/FAEP. The manuscript was strengthened by attention to the comments to two peer reviewers.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2017.00196/full#supplementary-material

REFERENCES

Agostinho, A. A., Gomes, L. C., Suzuki, H. I., and Júlio, H. F. Jr. (2003). "Migratory fishes of the upper Paraná river basin," in *Migratory Fishes of South America: Biology, Fisheries and Conservation Status*, eds J. Carolsfeld, B. Harvey, C. Ross, and A. Baer (Victoria, BC: International Development Research Centre), 23–75.

Agostinho, A. A., Pelicice, F. M., and Gomes, L. C. (2008). Dams and the fish fauna of the Neotropical region: impacts and management related to diversity and fisheries. *Braz. J. Biol.* 68(Suppl. 4), 1119–1132. doi: 10.1590/S1519-69842008000500019

Allendorf, F. W., Hohenlohe, P. A., and Luikart, G. (2010). Genomics and the future of conservation genetics. *Nat. Rev. Genet.* 11, 698–709. doi: 10.1038/nrg2844

Attard, C. R. M., Möller, L. M., Sasaki, M., Hammer, M. P., Bice, C. M., Brauer, C. J., et al. (2016). A novel holistic framework for genetic-based captive-breeding and reintroduction programs. *Conserv. Biol.* 30, 1060–1069. doi: 10.1111/cobi.12699

Avise, J. C. (2000). *Phylogeography: The History and Formation of Species*. Cambridge, MA: Harvard University Press.

Avise, J. C. (2004). *Molecular Markers, Natural History, and Evolution*, second Edn. Sunderland, MA: Sinauer.

Brito, I. A., Freire, C. A., Yamamoto, F. Y., Assis, H. C. S., Souza-Bastos, L. R., Cestari, M. M., et al. (2012). Monitoring water quality in reservoirs for human supply through multi-biomarker evaluation in tropical fish. *J. Environ. Monit.* 14, 615–625. doi: 10.1039/C2EM10461J

Caneppele, D., Pompeu, P., and Garavello, J. C. (2008). "Surubim do Paraíba (*Steindachneridion parahybae*," in *Livro Vermelho da Fauna Brasileira Ameaçada de Extinção*, first Edn, eds A. B. M. Machado, G. M. Drummond, and A. P. Paglia (Brasília: Ministério do Meio Ambiente), 234–238.

Carvalho, D. C., Oliveira, D. A. A., Beheregaray, L. B., and Torres, R. A. (2012). Hidden genetic diversity and distinct evolutionarily significant units in a commercially important Neotropical apex predator, the catfish *Pseudoplatystoma corruscans*. *Conserv. Genet.* 13, 1671–1675. doi: 10.1007/s10592-012-0402-6

Casacci, L. P., Barbero, F., and Balletto, E. (2014). The "Evolutionarily Significant Unit" concept and its applicability in biological conservation. *Ital. J. Zool.* 81, 182–193. doi: 10.1080/11250003.2013.870240

Chauhan, T., and Rajiv, K. (2010). Molecular markers and their applications in fisheries and aquaculture. *Adv. Biosci. Biotechnol.* 1, 281–291. doi: 10.4236/abb.2010.14037

Cornuet, J. M., and Luikart, G. (1997). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144, 2001–2014.

Dehais, C., Eudeline, R., Berrebi, P., and Argillier, C. (2010). Microgeographic isolation in chub (Cyprinidae: *Squalius cephalus*) population of the durance river: estimating fragmentation by dams. *Ecol. Freshw. Fish* 19, 267–278. doi: 10.1111/j.1600-0633.2010.00411.x

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver. 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Excoffier, L. G., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.

FAO (2008). *Aquaculture Development. 3. Genetic Resource Management*. FAO Technical Guidelines for Responsible Fisheries 5. Rome: Food and Agriculture Organization of the United Nations.

Frankham, R. (2005). Stress and adaptation in conservation genetics. *Biol. Evol. J.* 18, 750–755. doi: 10.1111/j.1420-9101.2005.00885.x

Frankham, R. (2010). Challenges and opportunities of genetic approaches to biological conservation. *Biol. Conserv.* 143, 1919–1927. doi: 10.1016/j.biocon.2010.05.011

Frankham, R., Ballou, J. D., and Briscoe, D. A. (2009). *Introduction to Conservation Genetics*, 2nd Edn. Cambridge: Cambridge University Press.

Fraser, D. J. (2008). How well can captive breeding programs conserve biodiversity? A review of salmonids. *Evol. Appl.* 1, 535–586. doi: 10.1111/j.1752-4571.2008.00036.x

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.

Garavello, J. C. (2005). Revision of Genus Steindachneridion (Siluriformes: Pimelodidae). *Neotrop. Ichthyol.* 3, 607–623. doi: 10.1590/S1679-62252005000400018

Gouskov, A., Reyes, M., Wirthner-Bitterlin, L., and Vorburger, C. (2016). Fish population genetic structure shaped by hydroelectric power plants in the upper Rhine catchment. *Evol. Appl.* 9, 394–408. doi: 10.1111/eva.12339

Grant, W. S., and Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: insights for sardines and anchovies and lessons for conservation. *J. Hered.* 89, 415–426. doi: 10.1093/jhered/89.5.415

Guo, S. W., and Thompson, E. A. (1992). Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics* 48, 361–372. doi: 10.2307/2532296

Hilsdorf, A. W. S., and Hallerman, E. M. (2017). *Genetic Resources of Neotropical Fishes*. Cham: Springer International Publishing, doi: 10.1007/978-3-319-55838-7

Hilsdorf, A. W. S., and Petrere, M. (2002). Peixes da bacia do Rio Paraíba do Sul: aspecto de sua diversidade e conservação. *Ciênc. Hoje* 180, 62–65.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.

Honji, R. M., Caneppele, D., Hilsdorf, A. W. S., and Moreira, R. G. (2009). Threatened fishes of the world: *Steindachneridion parahybae* (Steindachner, 1877) (Siluriformes: Pimelodidae). *Environ. Biol. Fish* 85, 207–208. doi: 10.1007/s10641-009-9480-9

Honji, R. M., Tolussi, C. E., Caneppele, D., Polaz, C. N. M., Hilsdorf, A. W. S., and Moreira, R. G. (2017). Biodiversidade e conservação da ictiofauna ameaçada de extinção da bacia do Rio Paraíba do Sul. *Rev. Biol.* 17, 18–30. doi: 10.7594/revbio.17.02.05

Hubert, N., and Renno, J. F. (2006). Historical biogeography of South American freshwater fishes. *J. Biogeogr.* 33, 1414–1436. doi: 10.3410/f.1047378.497321

IUCN/SSC (2013). *Guidelines for Reintroductions and Other Conservation Translocations, Version 1.0*. Gland: International Union for Conservation of Nature IUCN.

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11:94. doi: 10.1186/1471-2156-11-94

Jost, L. (2008). G_{ST} and its relatives do not measure differentiation. *Molec. Ecol.* 17, 4015–4026. doi: 10.1111/j.1365-294X.2008.03887.x

Kalinowski, S. T. (2006). HW-QUICKCHECK: an easy-to-use computer program for checking genotypes for agreement with Hardy-Weinberg expectations. *Mol. Ecol. Notes* 6, 974–979. doi: 10.1111/j.1471-8286.2006.01456.x

Kalinowski, S. T., Wagner, A. P., and Taper, M. L. (2006). ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol. Ecol. Notes* 6, 576–579. doi: 10.1111/j.1471-8286.2006.01256.x

Li, C. C., Weeks, D. E., and Chakravarti, A. (1993). Similarity of DNA fingerprints due to chance and relatedness. *Hum. Hered.* 43, 45–52. doi: 10.1159/000154113

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187

Loris, A. (2008). The limits of integrated water resources management: a case study of Brazil's Paraíba do Sul River Basin. *Sustainability* 4, 4–11.

Lynch, M., and Ritland, K. (1999). Estimation of pairwise relatedness with molecular markers. *Genetics* 152, 1753–1766.

Machado, C. E. M., and Abreu, H. C. F. (1952). Notas preliminares sobre a caça e a pesca no Estado de São Paulo - A pesca no Vale do Paraíba. *Bol. Ind. Anim.* 13, 145–160.

Marmulla, G. (2001). *Dams, Fish and Fisheries Opportunities, Challenges and Conflict Resolution.* FAO Technical Paper No. 419. Rome: FAO.

Mesquita, N., Carvalho, G., Shaw, P., Crespo, E., and Coelho, M. M. (2001). River basin-related genetic structuring in an endangered fish species, *Chondrostoma lusitanicum*, based on mtDNA sequencing and RFLP analysis. *Heredity* 86, 253–264. doi: 10.1046/j.1365-2540.2001.00776.x

Miller, L., and Kapuscinski, A. R. (2003). "Genetic guidelines for hatchery supplementation programs," in *Population Genetics: Principles and Applications for Fishery Scientists*, ed. E. M. Hallerman (Bethesda, MD: American Fisheries Society), 329–355.

Milligan, B. G. (2003). Maximum-likelihood estimation of relatedness. *Genetics* 163, 1153–1167.

Moraes, M. B., Polaz, C. N. M., Caramaschi, E. P., Santos, S. Jr., Souza, G., and Carvalho, F. L. (2017). Espécies exóticas e alóctones da bacia do Rio Paraíba do Sul: implicações para a conservação. Biodiversidade Brasileira (conservação de peixes continentais e manejo de unidades de conservação). *BioBrasil* 1, 34–54.

Moritz, C. (1994). Defining 'Evolutionarily Significant Units' for conservation. *Trends Ecol. Evol.* 9, 373–375. doi: 10.1016/0169-5347(94)90057-4

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Fonseca, G. A. B., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858. doi: 10.1038/35002501

Neff, B. D., Garner, S. R., and Pitcher, T. E. (2011). Conservation and enhancement of wild fish populations: preserving genetic quality versus genetic diversity. *Can. J. Fish. Aquat. Sci.* 68, 1139–1154. doi: 10.1139/f2011-029

Ojeda, A. P., Hilsdorf, A. W. S., Leite, A. C., Yang, A.-A., Izuno, A., He, C., et al. (2016). Development and characterization of microsatellite markers for the critically endangered Neotropical catfish *Steindachneridion parahybae*. *Conserv. Genet. Resour.* 8, 587–594. doi: 10.1007/s12686-016-0635-7

Olsen, J. B., Bentzen, P., Banks, M. A., Shaklee, J. B., and Young, S. (2000). Microsatellites reveal population identity of individual pink salmon to allow supportive breeding of a population at risk of extinction. *Trans. Am. Fish. Soc.* 129, 232–242. doi: 10.1577/1548-8659

Oyakawa, T. O., Menezes, N. A., Shibatta, O. A., Lima, F. C. T., Langeani, F., Pavanelli, C. S., et al. (2009). "Peixes de água doce," in *Fauna Ameaçada de Extinção no Estado de São Paulo: Vertebrados*, eds P. M. Bressan, M. C. M. Kierulff, and A. M. Sugieda (Sao Paulo: Fundação Parque Zoológico de São Paulo), 401.

Pádua, J. A. (2017). "Brazil in the history of the Anthropocene," in *Brazil in the Anthropocene: Conflicts Between Predatory Development and Environmental Policies*, eds L.-R. Issberner and P. Léna (New York, NY: Routledge), 19–40.

Palmer, G. C., Williams, J., Scott, M., Finne, K., Johnson, N., Dutton, D., et al. (2007). "Genetic marker-assisted restoration of the presumptive native walleye fishery in the New River, Virginia and West Virginia," in *Proceedings of the Annual Conference Southeastern Association of Fish and Wildlife Agencies*, Vol. 61, Montgomery, AK, 17–22.

Palsbøll, P. J., Bérubé, M., and Allendorf, F. W. (2007). Identification of management units using population genetic data. *Trends Ecol. Evol.* 22, 11–16. doi: 10.1016/j.tree.2006.09.003

Pelicice, F. M., Pompeu, P. S., and Agostinho, A. A. (2015). Large reservoirs as ecological barriers to downstream movements of Neotropical migratory fish. *Fish Fish.* 16, 697–715. doi: 10.1111/faf.12089

Piorski, N. M., Sanches, A., Carvalho-Costa, L. F., Hatanaka, T., Carrillo-Avila, M., Freitas, P. D., et al. (2008). Contribution of conservation genetics in assessing Neotropical freshwater fish biodiversity. *Braz. J. Biol.* 68, 1039–1050. doi: 10.1590/S1519-69842008000500011

Piry, S., Luikart, G., and Cornuet, J. M. (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *J. Hered.* 90, 502–503. doi: 10.1093/jhered/90.4.502

Pompanon, F., Bonin, A., Bellemain, E., and Taberlet, P. (2005). Genotyping errors: causes, consequences and solutions. *Nat. Rev. Genet.* 6, 847–859. doi: 10.1038/nrg1707

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Queller, D. C., and Goodnight, K. F. (1989). Estimating relatedness using molecular markers. *Evolution* 43, 258–275. doi: 10.1111/j.1558-5646.1989.tb04226.x

R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ramos-Onsins, S. E., and Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* 19, 2092–2100. doi: 10.1093/oxfordjournals.molbev.a004034

Ritland, K. (1996). Estimators for pairwise relatedness and inbreeding coefficients. *Genet. Res.* 67, 175–186. doi: 10.1017/S0016672300033620

Roberts, J. H., Angermeier, P. L., and Hallerman, E. M. (2013). Distance, dams, and drift: what structures populations of an endangered, benthic stream fish? *Freshwat. Biol.* 58, 2050–2064. doi: 10.1111/fwb.12190

Santos, C. H. A., Santana, G. X., Sá-Leitão, C. S., Paula-Silva, M. N., and Almeida-Val, V. M. F. (2016). Loss of genetic diversity in farmed populations of *Colossoma macropomum* estimated by microsatellites. *Anim. Genet.* 47, 373–376. doi: 10.1111/age.12422

Saura, M., Fernández, A., Rodríguez, M. C., Toro, M. A., Barragán, C., Fernández, A. I., et al. (2013). Genome-wide estimates of coancestry and inbreeding in a closed herd of ancient Iberian pigs. *PLOS ONE* 8:e78314. doi: 10.1371/journal.pone.0078314

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* 18, 233–234. doi: 10.1038/72708

Selkoe, K. A., and Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* 9, 615–629. doi: 10.1111/j.1461-0248.2006.00889.x

Snyder, N. F. R., Derrickson, S. R., Beissinger, S. R., Wiley, J. W., Smith, T. B., Toone, W. D., et al. (1996). Limitations of captive breeding in endangered species recovery. *Conserv. Biol.* 10, 338–348. doi: 10.1046/j.1523-1739.1996.10020338.x

Sriphairoj, K., Kamonrat, W., and Na-Nakorn, U. (2007). Genetic aspect in broodstock management of the critically endangered Mekong giant catfish, *Pangasianodon gigas* in Thailand. *Aquaculture* 264, 36–46. doi: 10.1016/j.aquaculture.2006.12.046

Taggart, J. B., Hynes, R. A., Prodohl, P. A., and Fergusson, A. (1992). A simplified protocol for routine total DNA isolation from salmonid fishes. *J. Fish. Biol.* 40, 963–965. doi: 10.1111/j.1095-8649.1992.tb02641.x

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Res.* 4, 535–538. doi: 10.1111/j.1471-8286.2004.00684.x

Wang, J. (2002). An estimator for pairwise relatedness using molecular markers. *Genetics* 160, 1203–1215.

Wang, J. (2007). Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genet. Res.* 89, 135–153. doi: 10.1017/S0016672307008798

Wang, J. (2011). COANCESTRY: a program for simulating, estimating and analyzing relatedness and inbreeding coefficients. *Mol. Ecol. Res.* 11, 141–145. doi: 10.1111/j.1755-0998.2010.02885.x

Weese, D. J., Ferguson, M. M., and Robinson, B. W. (2012). Contemporary and historical evolutionary processes interact to shape patterns of within-lake phenotypic divergences in polyphenetic pumpkinseed sunfish, *Lepomis gibbosus*. *Ecol. Evol.* 2, 574–592. doi: 10.1002/ece3.72

Weir, B. S., and Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.

Witzenberger, K. A., and Hochkirch, A. (2011). *Ex situ* conservation genetics: a review of molecular studies on the genetic consequences of captive breeding programmes for endangered animal species. *Biodivers. Conserv.* 20, 1843–1861. doi: 10.1007/s10531-011-0074-4

Zaniboni-Filho, E., and Schulz, U. H. (2003). "Migratory fishes of the Uruguay River," in *Migratory Fishes of South America. Biology, Fisheries and Conservation Status*, eds J. Carolsfeld, B. Harvey, C. Ross, and A. Baer (Victoria, BC: International Development Research Centre), 157–194.

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

Copyright © 2017 Fonseca, Domingues, Hallerman and Hilsdorf. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 12 December 2017
doi: 10.3389/fgene.2017.00203

First Chromosomal Analysis in Hepsetidae (Actinopterygii, Characiformes): Insights into Relationship between African and Neotropical Fish Groups

*Pedro C. Carvalho[1], Ezequiel A. de Oliveira[1,2], Luiz A. C. Bertollo[1], Cassia F. Yano[1], Claudio Oliveira[3], Eva Decru[4], Oladele I. Jegede[5], Terumi Hatanaka[1], Thomas Liehr[6], Ahmed B. H. Al-Rikabi[6] and Marcelo de B. Cioffi[1]**

[1] Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, Brazil, [2] Secretaria de Estado de Educação de Mato Grosso (Seduc-MT), Cuiabá, Brazil, [3] Departamento de Morfologia, Instituto de Biociências, Universidade Estadual Paulista, Botucatu, Brazil, [4] Section Vertebrates, Ichthyology, Royal Museum for Central Africa, Tervuren, Belgium, [5] Department of Fisheries and Aquaculture, Adamawa State University, Mubi, Nigeria, [6] Institute of Human Genetics, University Hospital Jena, Jena, Germany

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni,
Ponta Grossa State University, Brazil

Reviewed by:
Maelin Silva,
Universidade Federal da Fronteira Sul, Brazil
Lukas Kratochvil,
Charles University, Czechia

***Correspondence:**
Marcelo de B. Cioffi
mbcioffi@ufscar.br

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

Received: 18 October 2017
Accepted: 22 November 2017
Published: 12 December 2017

Citation:
Carvalho PC, de Oliveira EA, Bertollo LAC, Yano CF, Oliveira C, Decru E, Jegede OI, Hatanaka T, Liehr T, Al-Rikabi ABH and Cioffi MB (2017) First Chromosomal Analysis in Hepsetidae (Actinopterygii, Characiformes): Insights into Relationship between African and Neotropical Fish Groups. Front. Genet. 8:203. doi: 10.3389/fgene.2017.00203

Hepsetidae is a small fish family with only the genus *Hepsetus*, with six described species distributed throughout the South, Central and Western regions of Africa, showing a close relationship with the Alestidae and some Neotropical fish families. However, no cytogenetic information is available for both Hepsetidae and Alestidae species, thus preventing any evolutionary comparative studies at the chromosomal level. In the present study, we are providing new cytogenetic data for *Hepsetus odoe*, including the standard karyotype, C-banding, repetitive DNAs mapping, comparative genomic hybridization (CGH) and whole chromosome painting (WCP), providing chromosomal patterns and subsidies for comparative cytogenetics with other characiform families. Both males and females *H. odoe* have 2n = 58 chromosomes (10m + 28sm + 20st/a), with most of the C-band positive heterochromatin localized in the centromeric and subtelomeric regions. Only one pair of chromosomes bears proximal 5S rDNA sites in the short arms, contrasting with the 18S rDNA sequences which are located in the terminal regions of four chromosome pairs. Clear interstitial hybridization signals are evidenced for the U1 and U2 snDNA probes, but in only one and two chromosome pairs, respectively. Microsatellite motifs are widely distributed in the karyotype, with exception for the $(CGG)_{10}$, $(GAA)_{10}$ and $(GAG)_{10}$ probes, which highlight conspicuous interstitial signals on an unique pair of chromosomes. Comparative data from conventional and molecular cytogenetics, including CGH and WCP experiments, indicate that *H. odoe* and some Erythrinidae species, particularly *Erythrinus erythrinus*, share similar chromosomal sequences suggesting some relatedness among them, although bearing genomic specificities in view of their divergent evolutionary histories.

Keywords: fishes, molecular cytogenetics, chromosomal painting, comparative genomic hybridization (CGH), karyotype evolution

INTRODUCTION

Characiformes comprises 24 families and more than 2100 species (Eschmeyer and Fong, 2017), distributed in many Neotropical and Ethiopian rivers (Nelson et al., 2016). As they are exclusively freshwater fishes, their evolutionary history is related with continents fragmentations and settlement and, with the development of natural barriers during their dispersion throughout secondary habitats (Vari and Malabarba, 1998; Oliveira et al., 2007).

The most primitive characiforms are the African citharinoids (Arroyave et al., 2013), and the relationship between the Neotropical and Ethiopian species may be closely linked with the Gondwana break-up, with a fast diversification established in a new habitat free of competition (Calcagnotto et al., 2005). Despite significant efforts on morphological and molecular analyzes, phylogenetic relationships are currently still uncertain for several groups and even the monophyly of Characiformes is still debated (Arcila et al., 2017).

The wide diversification of the characiforms is highlighted by the high karyotype variability found within distinct Neotropical groups, showing the fast evolution of these fishes as expected by the high fragmentation observed in the South American rivers, in contrast with the African ones, which presents lower fragmentation and variability (Ortí and Meyer, 1997; Oliveira et al., 2007). One example of such scenario concerns the Erythrinidae, a small family widely distributed throughout South America, consisting of the genus *Erythrinus*, *Hoplerythrinus,* and *Hoplias* (Oyakawa, 2003). Cytogenetics of the Erythrinidae fishes have been quite investigated over years, especially for *H. malabaricus* and *E. erythrinus*, where a variety of chromosomal features occurs even within a same nominal species, thus supporting the presence of species complexes (Bertollo, 2007; Cioffi et al., 2012). In fact, erythrinids hold a variety of different karyomorphs, with diploid numbers (2n) varying from 39 in *Hoplias malabaricus* (karyomorph D) to 2n = 54 in *Erythrinus erythrinus* (karyomorph A), in addition to distinct sex chromosomes systems with independent origins and particular evolutionary trajectories (Cioffi et al., 2013). The diploid number found for most *Erythrinus* species (2n = 54) is also the common one observed for Characiformes, which possibly represents the ancestral condition for this order (Oliveira et al., 2007; Cioffi et al., 2012). However, the full comprehension of the evolutionary relationships of its families is not clear until now. A recent phylogeny based on 1,051 genetic markers showed that both African Hepsetidae and Alestidae families have closer relationship and, in a lower scale, to other Neotropical families, such as Erythrinidae, Cynodontidae and Hemiodotidae, but not with Lebiasinidae and Ctenoluciidae (Arcila et al., 2017). This result is not fully consensual with some previous phylogenetic proposals (Ortí and Meyer, 1997; Buckup, 1998; Calcagnotto et al., 2005), where some of above families were found to be related.

Notwithstanding, except for the Erythrinidae (see above), most of these families remain with kayotypes poorly analyzed, thus limiting any evolutionary comparative studies among them at the chromosomal level. In this sense, karyological data for Hemiodontidae are mainly restricted to chromosome numbers although all species presenting the same diploid number (2n = 54) and bi-armed chromosomes (Arefjev, 1990; Porto et al., 1992, 1993; Arai, 2011). Concerning Cynodontidae, the only species analyzed up to now (*Rhaphiodon vulpinus*) also presented the same 2n and karyotype structure (Pastori et al., 2009). In turn, the available chromosome data for Lebiasinidae are also mainly restricted to chromosome numbers (Scheel, 1973; Arai, 2011), with exception for a few species (Arefjev, 1990; Oliveira et al., 1991). Despite such largely limitation, a high diversity characterizes their diploid numbers, which ranges from 2*n* = 22 in *Nannostomus unifasciatus*, to 2n = 46 in *N. trifasciatus* (Oliveira et al., 2007; Arai, 2011). Occasional occurrence of large metacentric pairs, such as in *N. unifasciatus* (Arefjev, 1990) points to Robertsonian fusions in the karyotype differentiation. *Pyrrhulina australis* and *Pyrrhulina* aff. *australis* share 2n = 40 (4st + 36a), however, a significant genomic divergence was found between them, evidencing that they correspond to distinct evolutionary units (Moraes et al., 2017). As regards to Ctenoluciidae, four species of the *Boulengerella* genus from the Amazon River basin (Brazil), showed 2n = 36 and a very similar karyotype organization. A conspicuous chromosomal heteromorphism in male specimens point to a possible XX/XY sex chromosome system in such species (de Souza E Sousa et al., 2017). Besides, *Ctenolucius hujeta* (2n = 36) is the only additional species of Ctenoluciidae that has its chromosomal number already analyzed (Arefjev, 1990), coinciding with those found for the *Boulengerella* species.

The Hepsetidae family contains only a single genus (*Hepsetus*) and, for a long time, *H. odoe* was considered the only valid species. However, five additional species have been described by recent studies: *H. kingsleyae*, *H. lineatus*, *H. occidentalis*, *H. cuvieri,* and *H. microlepis*, distributed throughout the South, Central and Western regions of Africa (Decru et al., 2012, 2013a,b, 2015), where they have great significance for local economy (Kareem et al., 2016). Despite the economic and evolutionary importance of this group, no chromosome data are available for any Hepsetidae species.

In the present study, we provide, for the first time, cytogenetic data for *Hepsetus odoe*, including the standard karyotype, C-banding, repetitive DNAs mapping, comparative genomic hybridization (CGH) and whole chromosome painting (WCP), in order to investigate its chromosomal patterns and provide subsidies for comparative analyzes with some Neotropical fish families. In this sense, this study represents the first one of a series focusing on the cytogenetics and cytogenomics of the African species toward their karyoevolutionary processes.

MATERIALS AND METHODS

Specimens, Chromosome Preparations, C-banding and DNA Samples

Eleven specimens of *Hepsetus odoe* (06 males and 05 females) from the Opa Reservoir, Obafemi Awolowo University, Nigeria

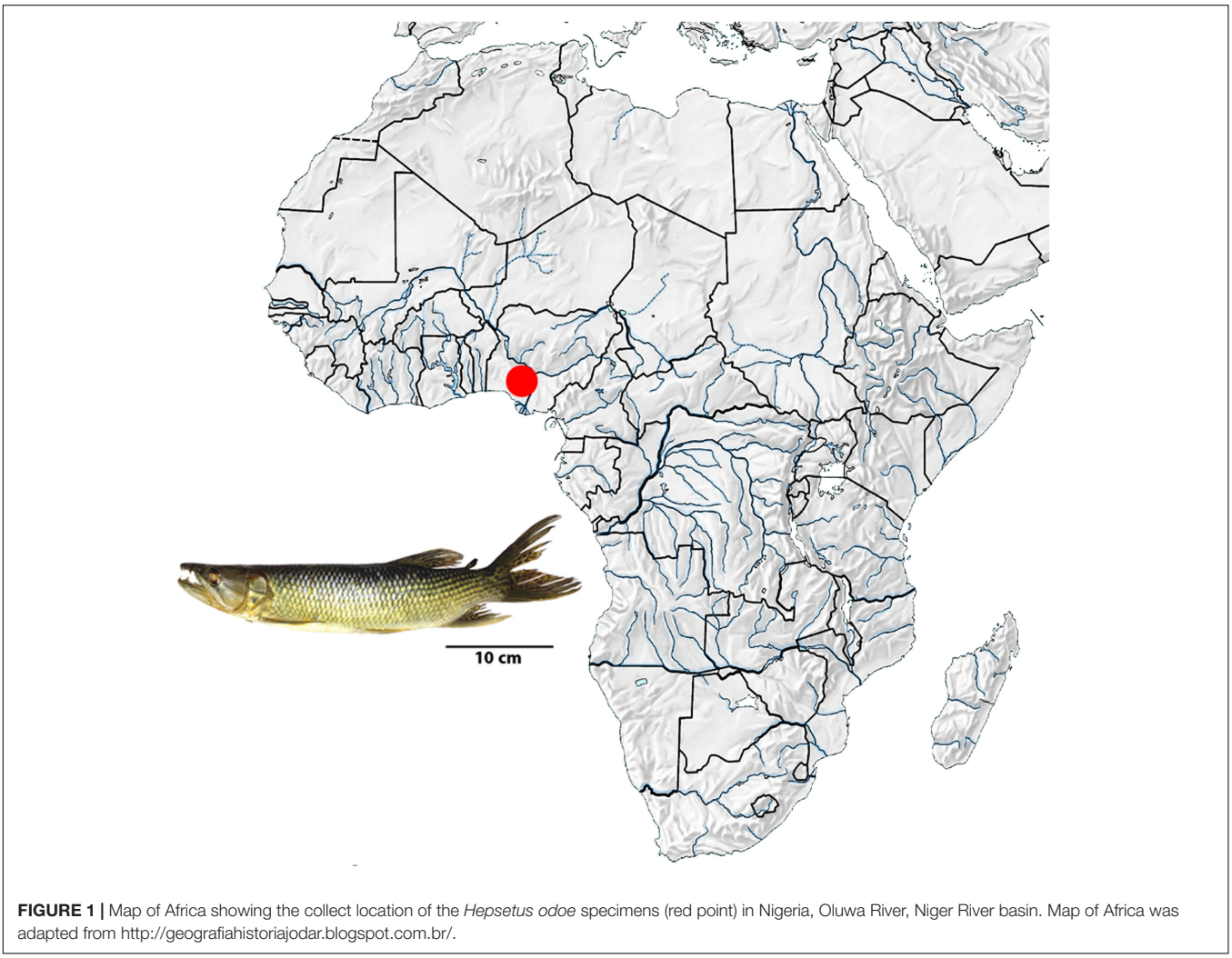


FIGURE 1 | Map of Africa showing the collect location of the *Hepsetus odoe* specimens (red point) in Nigeria, Oluwa River, Niger River basin. Map of Africa was adapted from http://geografiahistoriajodar.blogspot.com.br/.

(6°51′45″ N, 4°79′00″ E) were analyzed (**Figure 1**). The specimens were transferred to laboratory aquaria and kept under standard conditions for 1 day prior to the experiments. As *H. odoe* represent a non-CITES threatened species, no proper authorization was required for their sampling and/or transportation. All specimens were deposited in the Museu de Zoologia of the Universidade de São Paulo (MZUSP), under the accession No. 119844. Mitotic chromosomes were obtained by the protocols described in Bertollo et al. (2015) and experiments followed ethical conducts in accordance with the Ethics Committee on Animal Experimentation of the Universidade Federal de São Carlos (Process number CEUA 1853260315). The C-positive heterochromatin was detected using the Barium hydroxide protocol (Sumner, 1972). The genomic DNA was extracted according to standard phenol–chloroform procedures (Sambrook and Russell, 2001).

Probes for Chromosome Hybridization

A total of 11 repetitive DNA sequences, including four multigene families (U1 and U2 snDNA, 5S and 18S rDNAs) and seven microsatellite repeat motifs $(A)_{30}$, $(CA)_{15}$, $(GA)_{15}$, $(CAC)_{10}$, $(CGG)_{10}$, $(GAA)_{10}$ and $(GAG)_{10}$, were used as probes for FISH experiments. The oligonucleotide probes were directly labeled with Cy3 during synthesis according to Kubat et al. (2008). The other four tandemly arrayed DNA sequences were obtained via PCR from the nuclear DNA of *H. odoe*. The 5S rDNA repeat copy included 120 base pairs (bp) of the 5S rRNA transcribing gene and 200 bp of the non-transcribed spacer (NTS), produced according to Pendás et al. (1994). The second probe contained 1,400-bp repeats of the 18S rRNA gene, obtained according to Cioffi et al. (2009). Both rDNA probes were cloned into plasmid vectors and propagated in DH5α *Escherichia coli* competent cells (Invitrogen, San Diego, CA, United States). The U1 and U2 snDNA sequences were produced by PCR, according to Cross et al. (2005) and Silva et al. (2015), respectively. All these probes were directly labeled with Spectrum Orange-dUTP by nick translation, according to manufacturer's recommendations (Roche, Mannheim, Germany), with the exception of 5S rDNA, which was directly labeled with Spectrum Green-dUTP, also by nick translation (Roche, Mannheim, Germany).

Fluorescence *in Situ* Hybridization (FISH) for Repetitive DNA Mapping

Fluorescence *in situ* hybridization (FISH) was performed under high stringency conditions on metaphase chromosome spreads, as described in Yano et al. (2017a). The chromosome slides were incubated with RNAse (10 μg/mL) for 1 h at 37°C in a wet chamber and then washed for 5 min in 1x PBS and incubated with pepsin 0,005% for 10 min at room temperature. It was followed a wash in 1x PBS, a fixation with 1% formaldehyde for 10 min at room temperature, and another 1x PBS wash. The slides were then set for an alcoholic series of 70, 85, and 100% 2 min each, followed by the DNA denaturation in 70% formamide/2x SSC for 3 min at 75°C. After denaturation, the chromosome spreads were dehydrated in an ethanol series of 70, 85, and 100% at room temperature, 2 min each. 20 μL of the hybridization mixture (100 ng probes, 50% deionized formamide, 10% dextran sulfate) were then dropped on the slides, and the hybridization was performed for 16–18 h at 37°C in a wet chamber containing 2x SSC. A post-hybridization wash was carried out with 2x SSC for 5 min followed by another wash in 1x SSC at 42°C, 5 min. A final washing series was then performed at room temperature, consisting of 1x PBS for 5 min, and ethanol 70, 85, and 100% for 2 min each. Finally, the chromosomes were counterstained with DAPI (1.2 μg/mL) and the slides mounted with an antifading solution (Vector, Burlingame, CA, United States).

Chromosomal Microdissection, Probe Preparation and Labeling

Fifteen copies of the following chromosomes were isolated by microdissection and amplified using the procedure described in Yang et al. (2009): (i) X chromosome of *Hoplias malabaricus* karyomorph B (HMB-X); (ii) Y_1 chromosome of *H. malabaricus* karyomorph G (HMG-Y1) and (iii) Y chromosome of *Erythrinus erythrinus* karyomorph D (ERY-Y). These probes were labeled with Spectrum Orange-dUTP (ERY-Y) or Spectrum Green-dUTP (HMB-X and HMG-Y1) (Vysis, Downers Grove, IL, United States) in a secondary DOP PCR using 1 μL of the primarily amplified product as a template DNA, following Yang et al. (2009).

FISH of Whole Chromosome Specific Probes (W)

Chromosomal preparations of males and females of *H. odoe* were used for Zoo-FISH experiments with all the above mentioned probes. The hybridization procedures followed Yano et al. (2017a). To block the hybridization of high-copy repeat sequences 60 μg of C_0t-1 DNA directly isolated from *H. malabaricus* (karyomorphs B and G) and *E. erythrinus* (karyomorph D) male genomes were prepared according to Zwick et al. (1997). Hybridization was performed for 144 h at 37°C in a moist chamber. The post-hybridization wash was carried out with 1x SSC for 5 min at 65°C, and in 4x SSC/Tween using a shaker at RT and then rinsed quickly in 1x PBS. Subsequently, the slides were dehydrated in an ethanol series (70, 85, and 100%), 2 min each. Finally, the chromosomes were counterstained with DAPI (1.2 μg/mL) and mounted in antifade solution (Vector).

Probes for Comparative Genomic Hybridization (CGH)

The gDNA of *H. odoe* was used for comparative analyzes with the gDNAs of several Erythrinidae species, namely *E. erythrinus* (karyomorph D), *Hoplias lacerdae, H. malabaricus* (karyomorph A) and *Hoplerythrinus unitaeniatus* (karyomorph D). The gDNA of *H. odoe* was labeled with biotin-16-dUTP using BIO-nick-translation Mix (Roche), while the male-derived gDNAs of *E. erythrinus*, *H. malabaricus, H. unitaeniatus,* and *H. lacerdae* were labeled with digoxigenin-11-dUTP using DIG-nick-translation Mix (Roche, Manheim, Germany). In all experiments it was utilized C_0t-1 DNA (i.e., fraction of genomic DNA enriched for highly and moderately repetitive sequences), prepared according to Zwick et al. (1997), for blocking common genomic repetitive sequences. The final probe was composed of 500 ng of *H. odoe* gDNA plus 500 ng of the corresponding gDNA for each Erythrinidae species. The probe was ethanol-precipitated and the dry pellet dissolved in a hybridization buffer (20 μL per slide) containing 50% formamide + 2x SSC + 10% SDS+ 10% dextran sulfate and Denhardt's buffer, pH 7.0).

Fluorescence *in Situ* Hybridization for CGH

CGH experiments were performed according to Symonová et al. (2013). Slides with the metaphase plates were stored overnight in a freezer, being submitted to an alcoholic series of 70, 85, and 100%, 3 min each, before and after the storage. After that, the slides were aged for 1–2 h at 60°C, washed in 2x SSC for 5 min, treated with RNAse (200 μg/mL) for 90 min at 37°C in a wet chamber and them washed in 2x SSC for 30 s. It was followed another alcoholic series treatment, a wash in 1x PBS for 5 min, a Pepsin (50 μg/mL) treatment, a wash in 1x PBS for 5 min and an additional alcoholic series treatment. Finally, the material was denaturated in 75% formamide/2x SSC at 74°C for 3 min, followed by an alcoholic series being the first 70% cold ethanol. 20 μL of the probes were spotted to the slides, which were them incubated at room temperature (37°C) in a dark humid chamber for 3 days, with rubber-sealed coverslips. The rubber cement and coverslips were removed in a solution of 4x SSC/0.1% Tween. The slides were then washed twice in 50% formamide/2x SSC for 10 min each, three times in 1x SSC, rinsed in 2x SSC at room temperature, and incubated 20 min. in a humid chamber with 500 μL of 3%BSA/4x SSC/Tween, with coverslips. The hybridization signal was detected with anti-digoxigenin-Rhodamin (Roche) diluted in 0.5% bovine serum albumin (BSA) in PBS, and avidin-FITC (Sigma) diluted in PBS containing 10% normal goat serum (NGS). Four final washes were performed at 44°C in 4x SSC/0.1% Tween, 7 min each. Finally, the chromosomes were counterstained with DAPI (1.2 μg/mL) and mounted in an antifade solution (Vector, Burlingame, CA, United States).

Microscopic Analyses

At least 30 metaphase spreads per individual were analyzed to confirm the diploid number, karyotype structure and FISH results. Images were captured using an Olympus BX50 microscope (Olympus Corporation, Ishikawa, Japan) with CoolSNAP and the images processed using Image Pro Plus 4.1 software (Media Cybernetics, Silver Spring, MD, United States). Chromosomes were classified as metacentric (m), submetacentric (sm), subtelocentric (st) or acrocentric (a), according to their arm ratios (Levan et al., 1964).

RESULTS

Karyotype Composition and C-banding

All specimens, both males and females, have 2n = 58 (10m + 28sm + 20st/a). The C-positive heterochromatic is most localized in the centromeric and subtelomeric regions, with a more conspicuous block present in the 28th chromosome pair of the karyotype (**Figure 2**).

Chromosomal Mapping of Repetitive DNAs

The 5S rDNA occurs in the proximal region of the short arms of only one sm chromosome pair, while the 18S rDNA is located in the terminal region of the long arms of four chromosome pairs (1m + 2sm + 1st) (**Figure 2**). Clear interstitial hybridization signals were observed in one pair of chromosomes for the U1 snDNA, and in two chromosome pairs for the U2 snDNA, being interstitial and telomeric located in each one of them, respectively (**Figure 3**). Widely distributed marks were evidenced by the microsatellite motifs. Signals were mainly telomeric, but some also interstitial, as for $(A)_{30}$, $(CA)_{15}$, $(GA)_{15}$, $(CAC)_{10}$ and $(GAG)_{10}$ probes. Exceptions for these general patterns were presented by the $(CGG)_{10}$, $(GAA)_{10}$ and $(GAG)_{10}$ probes, which highlighted a conspicuous interstitial signal on a unique pair of chromosomes (**Figure 3**).

Comparative Genomic Hybridization (CGH)

The comparative genomic hybridization showed that the gDNA of *H. odoe* shares some homologies with those of the Erythrinidae species analyzed. Despite some scattered hybridization, labeled telomeric and pericentromeric regions were evidenced according to each species. However, it stands out the hybridization pattern with *E. erythrinus*, where some whole chromosome pairs were labeled, in addition to telomeric overlaps in other ones. An exclusive acrocentric chromosome of *H. odoe*, that presented hybridization signals only with the gDNA of *H. unitaeniatus*, was also highlighted (**Figure 4**).

Detection of Chromosomal Homeologies by Zoo-FISH Experiments

Hybridization performed with HMB-X (X chromosome from *Hoplias malabaricus* karyomorph B) probe highly painted one small st/a chromosome of *H. odoe* (**Figure 5a**). Both HMG-Y1 (Y_1 chromosome from *H. malabaricus* karyomorph G) and ERY-Y (Y chromosome from *Erythrinus erythrinus* karyomorph D) probes painted the p arms of medium-sized st/a chromosomes (**Figures 5b,c**) of *H. odoe*. Besides, ERY-Y probe also produced faint scattered hybridization pattern on several other chromosomes of *H. odoe* (**Figure 5c**).

DISCUSSION

General Chromosome Features of *Hepsetus odoe*

The lack of karyotype data for several fish groups impairs comparative analyzes on their evolutionary trends and chromosomal relationships. This is the case for the African Hepsetidae family for which chromosomal characteristics are completely unknown. In this sense, this study is the first one providing classical and molecular cytogenetic data for one of its representative species, *H. odoe*.

Both male and female specimens of *H. odoe* have the same karyotype structure, with 2n = 58 (5m + 14sm + 10st/a), with no evidence of differentiated sex chromosomes. The heterochromatin distribution follows the general pattern usually found in many other fish species, with preferential centromeric localization. Only one chromosome pair bears proximal 5S rDNA sites in their short arms, in contrast to the 18S rDNA sequences that are located in the telomeric regions of four different pairs in the karyotype. The distribution of these multigene families is shared among many fish groups (Pendás et al., 1994; Gornung, 2013) where the clustering of the 5S and 18S rDNAs in different chromosomes may avoid unwanted chromosomal changes between them (Martins and Galetti, 1999). In addition, this differential clustering is also true for the U2 snDNA sequences, since the cytogenetic mapping for different genes that composed this multigene family, although scarce among fishes, shows a preferential distribution among distinct chromosomes (reviewed in Yano et al., 2017b), as also observed in *H. odoe*.

With respect to microsatellites, although the scattered distribution of some of them is not so useful for comparative approaches, the conspicuous interstitial bands that $(GAG)_{10}$, $(CGG)_{10}$ and $(GAA)_{10}$ probes highlighted in the genome of *H. odoe,* constitute important markers for comparative evolutionary analyzes with other Hepsetidae and also close related species. In fact, the clustering of microsatellites represents important evolutionary stages by composing non-coding genome regions, as well as relevant steps in the sex chromosome's differentiation process (Bergero and Charlesworth, 2009).

DNA sequence analysis strongly suggests that Hepsetidae and Alestidae are phylogenetic close related families (Oliveira et al., 2011; Arcila et al., 2017). In this sense, the present data set for *H. odoe* are useful tools for complementary investigations covering other Hepsetidae and Alestidae species. In fact, this study represents the first one of a series focusing on the cytogenetics and cytogenomics of such African families, toward the investigation of their karyoevolutionary processes and relatedness.

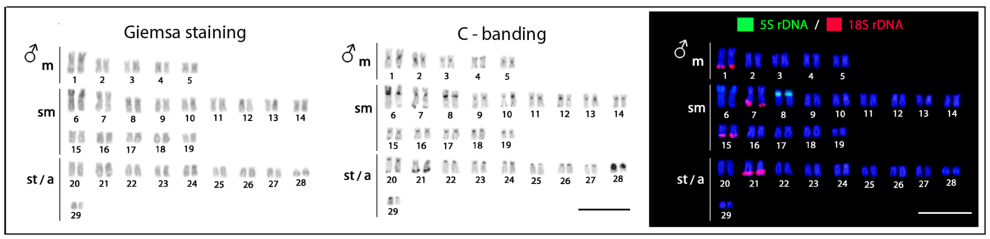


FIGURE 2 | *Hepsetus odoe* male karyotypes under standard Giemsa staining, C-banding and double-FISH with 18S rDNA (red) and 5S rDNA (green) probes. Both males and females have the same karyotypes. Bar = 5 μm.

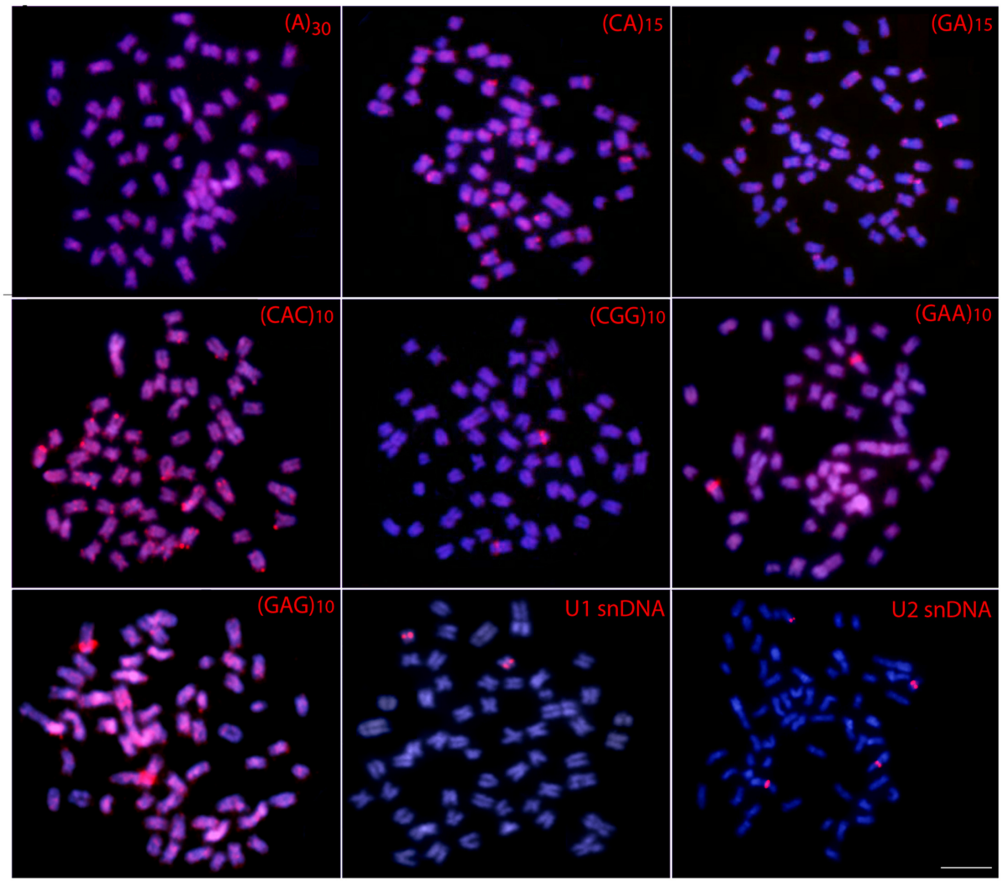


FIGURE 3 | Metaphase plates of *Hepsetus odoe* hybridized with repetitive DNA sequences, including mono-, di- and trinucleotide microsatellites and the multigene families U_1 and U_2 snDNAs. Bar = 5 μm.

Comparative Cytogenetics of *Hepsetus odoe* with Other Characiformes Species

Some previous phylogenetic studies (Ortí and Meyer, 1997; Buckup, 1998; Calcagnotto et al., 2005) have suggested a relationship between Hepsetidae and some other Neotropical groups, such as the Erythrinidae, Ctenoluciidae and Lebiasinidae, although without a full consensus among them. Using new sequencing technology together with phylogenetic reconstructions, a new scenario was evidenced, discarding relationships of Hepsetidae with Lebiasinidae and Cnetoluciidae and, instead off, placing Hepsetidae and Alestidae in a closer clade which has a near position in the phylogenetic tree to some other Neotropical families, such as Erythrinidae, Cynodontidae and Hemiodontidae (Oliveira et al., 2011; Arcila et al., 2017). In this way, as the cytogenetic studies among Cynodontidae and Hemiodontidae families are until now restricted to 2n descriptions in few species, and Alestidae species are still unavailable in spite of recent collecting efforts, we performed a comparative analysis among *H. odoe* and Erythrinidae, Ctenoluciidae and Lebiasinidae

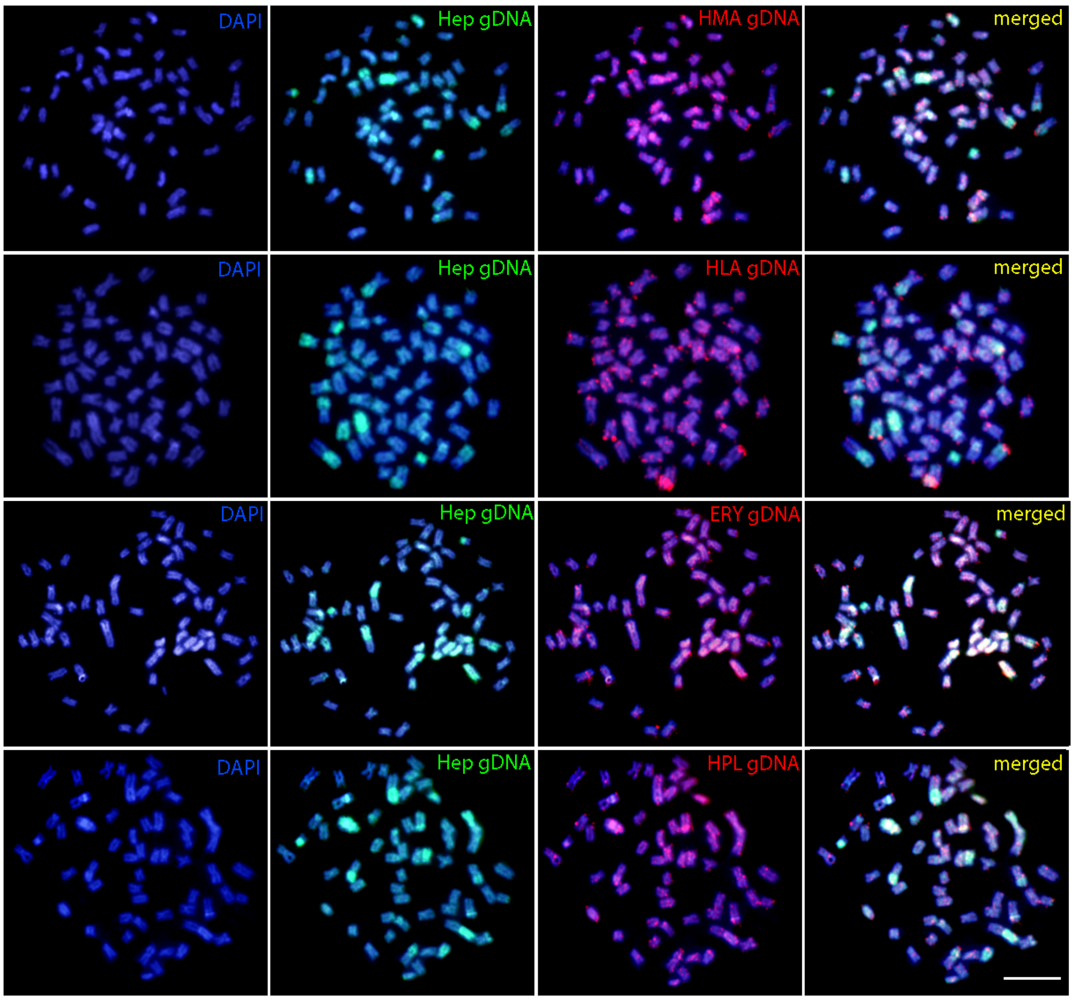


FIGURE 4 | Comparative genomic hybridization (CGH) in metaphase plates of *Hepsetus odoe*. First column: DAPI images (blue); Second column: hybridization pattern with *Hepsetus odoe* (Hep) gDNA probe; Third column: Hybridization patterns with *Hoplias malabaricus* (HMA) gDNA, *Hoplias lacerdae* (HLA) gDNA, *Erythrinus erythrinus* (ERY) gDNA and *Hoplerythrinus unitaeniatus* (HPL) gDNA probes; Fourth column: merged images of each genomic probes and DAPI staining. The common genomic regions are depicted in yellow. Bar = 5 μm.

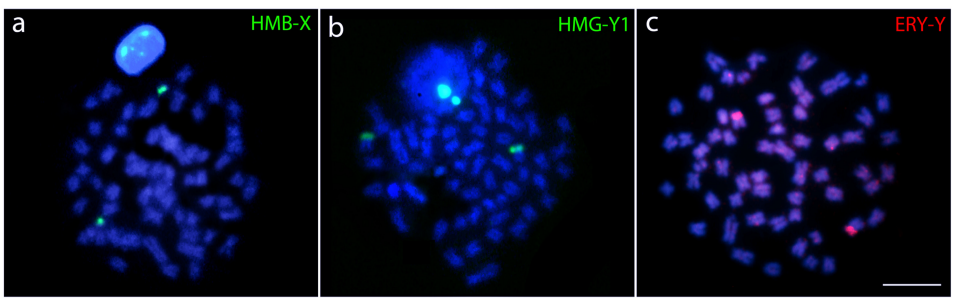


FIGURE 5 | Whole chromosome painting (WCP) in metaphase plates of *Hepsetus odoe* showing the chromosomes hybridized with **(a)** the X chromosome of *Hoplias malabaricus* karyomorph B (HMB-X), **(b)** the Y_1 chromosome of *Hoplias malabaricus* karyomorph G (HMG-Y_1) and **(c)** the Y chromosome of *Erythrinus erythrinus* karyomorph D (ERY-Y).

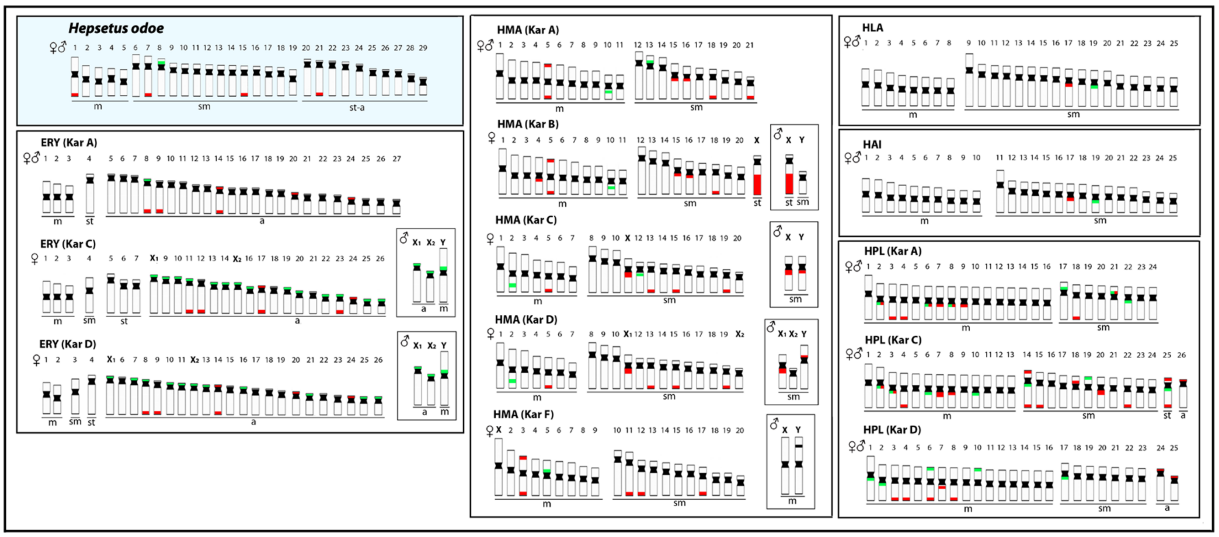


FIGURE 6 | Representative idiograms of *Hepsetus odoe* and Erythrinidae species: *Erythinus erythrinus* (ERY) karyomorphs (Kar) A, C, D; *Hoplias malabaricus* (HMA) karyomorphs A. B, C, D, F; *Hoplias lacerdae* (HLA); *Hoplias aimara* (HAI) and *Hoplerythrinus unitaeniatus* (HPL) karyomorphs A.C, D. The distribution of the 18S and the 5S rDNAs for each species are highlighted in red and green, respectively. The sex chromosomes are boxed. Data from Cioffi et al. (2009), Martins et al. (2013), Martinez et al. (2015), and Oliveira et al. (2015).

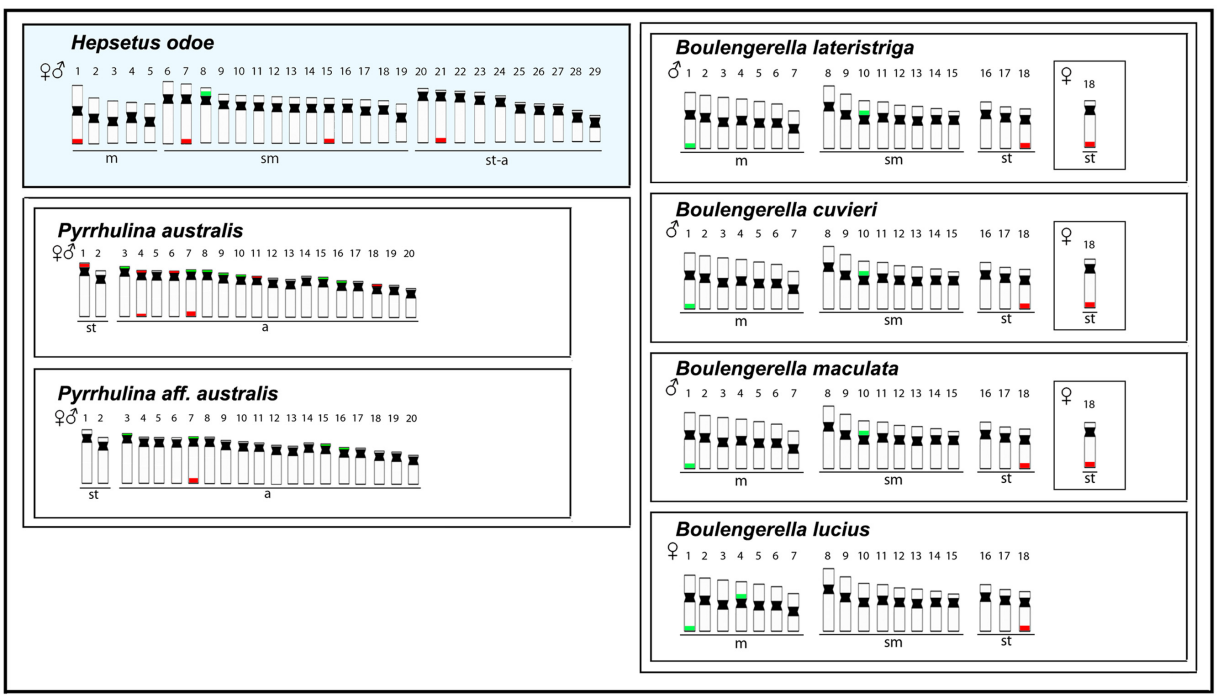


FIGURE 7 | Representative idiograms of *Hepsetus odoe* and *Pyrrhulina* (Lebiasinidae) (data from Moraes et al., 2017) and *Boulengerella* (Ctenoluciidae) (data from de Souza E Sousa et al., 2017) species. The distribution of the 18S and the 5S rDNAs for each species are highlighted in red and green, respectively.

representatives. In this sense, **Figures 6**, **7** depict some data, including chromosome number, karyotype organization, sex chromosome systems and distribution of the major and minor rDNA sequences in some Erytrinidae, Lebiasinidae, and Ctneoluciidae species. A general overview clearly indicates that Erythrinidae retains the highest amount of characters resembling those of *H. odoe* than Lebiasinidae and Ctenoluciidae species. Indeed, *Erythrinus erythrinus* (2n = 54/52), *Hoplias lacerdae* and *H. aimara* (2n = 50) and *Hoplerythrinus unitaeniatus* (2n = 48/52) show diploid numbers

closer to that of *H. odoe* (2n = 58) then *Pyrrhulina* (2n = 40; Lebiasinadae) and *Boulengerella* (2n = 36; Ctenoluciidae) species.

Particularly, inside Erythrinidae, *E. erythrinus* stand out as having more chromosomal similarities with *H. odoe* than the other ones, taking into account the broad organization of the karyotype and the amount of mono-armed chromosomes that they have. In fact, *E. erythrinus* karyomorph A shows the most basal karyotype inside this genus, considering that the other *Erythrinus* karyomorphs highlight clearly derived features, such as the differentiation of a multiple $X_1X_1X_2X_2/X_1X_2Y$ sex chromosome system (Bertollo et al., 2004) and the huge dispersion of the 5S rDNA sequences in the genome (Cioffi et al., 2010; Martins et al., 2013). In addition, like *H. odoe*, *E. erythrinus* karyomorph A presents only one chromosome pair bearing 5S rDNA sequences at a similar position on the chromosomes, as well as a number of exclusive telomeric 18S rDNA sites. However, whereas in *H. odoe* the major rDNA sequences are only distributed in the long arms of the chromosomes, in *E. erythrinus* they are found both in the short as well as in the long arms (Cioffi et al., 2010). This is not an unexpected condition in view of differential distributions that can be set up along the evolutionary history of the species. In fact, repetitive DNAs have played a particular role on fish karyotyping (Cioffi and Bertollo, 2012), and variations in amount and types of several classes of repetitive DNAs are expected considering the inherent dynamism of these sequences during the evolutionary history of different taxa (Kubat et al., 2008; Cioffi et al., 2010, 2012; Pokorná et al., 2011; Yano et al., 2016). In spite of this, the distribution pattern of the $(GAG)_{10}$ microsatellites in *H. odoe* also shows a significant accumulation on the *E. erythrinus* chromosomes (Yano et al., 2014).

Considering the above correlations between *Hepsetus* and Erythrinidae, comparative genomic hybridization (CGH) and whole chromosome painting (WCP) were also performed to obtain additional informative markers for comparative cytogenetics. Among fishes, CGH has been already applied for several purposes, such as to compare genomes of closely related species (Zhu and Gui, 2007; Knytl et al., 2013; Majtánová et al., 2016; Moraes et al., 2017), to detect parental genomes in hybrids (Symonová et al., 2013; Pereira et al., 2014), and to elucidate the origin and evolution of B and sex chromosomes (Fantinatti et al., 2011; Freitas et al., 2017; Yano et al., 2017c), among others. In our present case, CGH with four Erythrinidae species evidenced the co-localization of scattered signals in almost all chromosomes of *H. odoe*, together with the preferential signals in the terminal parts of some chromosomes, thus indicating the shared repetitive content of such regions. However, it stands out the hybridization pattern with *E. erythrinus*, where some whole chromosome pairs were painted, in addition to telomeric overlaps in other ones. Furthermore, the hybridization with *H. odoe* gDNA revealed the occurrence of conspicuous species-specific regions, very likely as a result of its particular evolutionary history, given that the resolution of the CGH method predominantly relies on the presence of species-specific (or sex-specific) repetitive DNA sequences and the evolutionary distance of the compared genomes.

Besides CGH, WCP experiments were also performed using microdissected sex chromosomes from *H. malabaricus* karyomorphs B (HMB-X) and G ($HMG\text{-}Y_1$) and *E. erythrinus* karyomorph D (ERY-Y) as probes, in order to verify the occurrence of putative sex chromosomes in *H. odoe*. As a control experiment, all probes were previously hybridized in male chromosomal preparations of *H. malabaricus* (karyomorphs B and G) and *E. erythrinus* (karyomorph D), clearly demonstrating the hybridization signals on the sex chromosomes of these karyomorphs, thus corroborating previous data (Cioffi et al., 2013; Oliveira et al., 2017). When these probes were hybridized to chromosomal preparations of *H. odoe*, HMB-X highly painted one small st/a chromosome, while HMG-Y1 and ERY-Y probes painted the p arms of medium-sized st/a chromosomes. This way, these results highlight that such linkage groups are shared by *H. odoe* and Erythrinidae species, corroborating the CGH experiments which also demonstrated the sharing of a considerable genomic fraction among such groups. The maintenance of such linkage groups is somehow surprising considering the phylogenetic distance between these clades. However, chromosome homology across widely phylogenetically distributed clades have been also detected in several mammals (Balmus et al., 2007; Dementyeva et al., 2010; Kulemzina et al., 2011), birds (Oliveira et al., 2008, 2010; Tagliarini et al., 2011) and lizard (Pokorná et al., 2011) species. In the later, Zoo-FISH experiments using a Z-derived probe from *Gallus gallus* showed that the fraction of the reptile genome that is homologous to the avian Z chromosome exhibits a conserved synteny, despite the very ancient times (∼275 Mya) of their divergence (Pokorná et al., 2011).

CONCLUSION

This study, focusing on standard and molecular cytogenetic approaches of *H. odoe*, represents the first data set for an Hepsetidae species. Our data supports the likely proximity between African and Neotropical families, such as Hepsetidae and Erythrinidae. In fact, our experiments, including CGH and WCP, indicate that *H. odoe* and some Erythrinidae species, in special from the genus *Erythrinus*, share similar chromosomal sequences, thus reflecting some degree of relationship among them. In fact, *Erythrinus* seems to carry the most basal karyotype organization within Erythrinidae, and likely the most proximal to that highlighted by *H. odoe*. This study represents the first one of a series of further investigations focusing on the African Characiformes chromosomal and genomic characteristics, allowing a broader and more detailed view on the evolutionary history of this group through a cytogenetic approach. Such additional data will securely improve our knowledge about the relatedness of the African and the Neotropical characiform families.

AUTHOR CONTRIBUTIONS

PC carried out the cytogenetic analysis and drafted the manuscript. EdO, CY, AA-R, and TH helped in the cytogenetic analysis, drafted and revised the manuscript. CO, ED, OJ, and TL drafted and revised the manuscript. MC and LB coordinated the study, drafted and revised the manuscript. All authors read and approved the final version of the manuscript

ACKNOWLEDGMENTS

This study was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (Proc. nos. 304992/2015-1 and 401575/2016-0) and Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP (Proc. nos. 2016/21411-7, 2016/17556-0 and 2017/08471-3) (FAPESP 2014/26508-3 and CNPq 306054/2006-0 to CO).

REFERENCES

Arai, R. (2011). *Fish Karyotypes: A Check List*. Tokyo: Springer. doi: 10.1007/978-4-431-53877-6

Arcila, D., Ortí, G., Vari, R., Armbruster, J. W., Stiassny, M. L., Ko, K. D., et al. (2017). Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. *Nat. Ecol. Evolut.* 1:0020. doi: 10.1038/s41559-016-0020

Arefjev, V. A. (1990). Problems of karyotypic variability in the family Characidae (Pisces, Characiformes) with the description of somatic karyotypes for six species of tetras. *Caryologia* 43, 305–319. doi: 10.1080/00087114.1990.10797009

Arroyave, J., Denton, J. S., and Stiassny, M. L. (2013). Are characiform fishes Gondwanan in origin? Insights from a time-scaled molecular phylogeny of the Citharinoidei (Ostariophysi: Characiformes). *PLOS ONE* 8:e77269. doi: 10.1371/journal.pone.0077269

Balmus, G., Trifonov, V. A., Biltueva, L. S., O'Brien, P. C. M., Alkalaeva, E. S., Fu, B., et al. (2007). Cross species painting among camel, cattle, pig and human: further insights into the putative Cetartiodactyla ancestral karyotype. *Chromosome Res.* 15, 499–514. doi: 10.1007/s10577-007-1154-x

Bergero, R., and Charlesworth, D. (2009). The evolution of restricted recombination in sex chromosomes. *Trends Ecol. Evolut.* 24, 94–102. doi: 10.1016/j.tree.2008.09.010

Bertollo, L. A. C. (2007). "Chromosome evolution in the Neotropical Erythrinidae fish family: an overview," in *Fish Cytogenetics*, eds E. Pisano, C. Ozouf-Costaz, F. Foresti, and B. G. Kapoor (Enfield, NH: Science Publishers), 195–211.

Bertollo, L. A. C., Cioffi, M. B., and Moreira-Filho, O. (2015). "Direct chromosome preparation from freshwater teleost fishes," in *Fish Cytogenetic Techniques*, eds C. Ozouf-Costaz, E. Pisano, F. Foresti, and L. F. Almeida Toledo (Enfield, CT: CRC Press), 21–26.

Bertollo, L. A. C., Oliveira, C., Molina, W. F., Margarido, V. P., Fontes, M. S., Pastori, M. C., et al. (2004). Chromosome evolution in the erythrinid fish, *Erythrinus erythrinus* (Teleostei: Characiformes). *Heredity* 93, 228–233. doi: 10.1038/sj.hdy.6800511

Buckup, P. A. (1998). "Relationships of the Characidiinae and phylogeny of characiform fishes (Teleostei: Characiformes)," in *Phylogeny and Classification of Neotropical Fishes*, eds L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M. S. Lucena, and C. A. S. Lucena (Porto Alegre: EDIPUCRS), 193–234.

Calcagnotto, D., Schaefer, S. A., and DeSalle, R. (2005). Relationships among characiform fishes inferred from analysis of nuclear and mitochondrial gene sequences. *Mol. Phylogenet. Evol.* 36, 135–153. doi: 10.1016/j.ympev.2005.01.004

Cioffi, M. B., and Bertollo, L. A. C. (2012). "Chromosomal distribution and evolution of repetitive DNAs in fish," in *Repetitive DNA*, ed. M. A. Garrido-Ramos (Basel: Karger), 197–221. doi: 10.1159/000337950

Cioffi, M. B., Liehr, T., Trifonov, V., Molina, W. F., and Bertollo, L. A. C. (2013). Independent sex chromosome evolution in lower vertebrates: a molecular cytogenetic overview in the Erythrinidae fish family. *Cytogenet. Genome Res.* 141, 186–194. doi: 10.1159/000354039

Cioffi, M. B., Martins, C., and Bertollo, L. A. (2010). Chromosome spreading of associated transposable elements and ribosomal DNA in the fish *Erythrinus erythrinus*. Implications for genome change and karyoevolution in fish. *BMC Evolut. Biol.* 10:271. doi: 10.1186/1471-2148-10-271

Cioffi, M. B., Martins, C., and Bertollo, L. A. C. (2009). Comparative chromosome mapping of repetitive sequences. Implications for genomic evolution in the fish, *Hoplias malabaricus*. *BMC Genetics* 10:34. doi: 10.1186/1471-2156-10-34

Cioffi, M. B., Molina, W. F., Artoni, R. F., and Bertollo, L. A. C. (2012). "Chromosomes as tools for discovering biodiversity – the case of Erythrinidae fish family," in *Recent Trends in Cytogenetic Studies – Methodologies and Applications*, ed. P. Tirunilai (Rijeka: InTech Publisher), 125–146.

Cross, I., Díaz, E., Sánchez, I., and Rebordinos, L. (2005). Molecular and cytogenetic characterization of *Crassostrea angulata* chromosomes. *Aquaculture* 247, 135–144. doi: 10.1016/j.aquaculture.2005.02.039

de Souza E Sousa, J. S., Viana, P. F., Bertollo, L. A. C., Cioffi, M. B., and Feldberg, E. (2017). Evolutionary relationships among *Boulengerella* species (Ctenoluciidae, Characiformes): genomic organization of repetitive DNAs and highly conserved karyotypes. *Cytogenet. Genome Res.* 152, 194–203. doi: 10.1159/000480141

Decru, E., Snoeks, J., and Vreven, E. (2013a). The true identity of the holotype of *Hepsetus odoe* and the names of the two west african species of *Hepsetus* (Teleostei: Hepsetidae). *Ichthyol. Explor. Freshw.* 24, 187–192.

Decru, E., Snoeks, J., and Vreven, E. (2015). Taxonomic evaluation of the *Hepsetus* from the Congo basin with the revalidation of *H. microlepis* (Teleostei: Hepsetidae). *Ichthyol. Explor. Freshw.* 26, 273–287.

Decru, E., Vreven, E., and Snoeks, J. (2012). A revision of the West African *Hepsetus* (Characiformes: Hepsetidae) with a description of *Hepsetus akawo* sp. nov. and a redescription of *Hepsetus odoe* (Bloch, 1794). *J. Nat. Hist.* 46, 1–23. doi: 10.1080/00222933.2011.622055

Decru, E., Vreven, E., and Snoeks, J. (2013b). A revision of the Lower Guinean *Hepsetus* species (Characiformes; Hepsetidae) with the description of *Hepsetus kingsleyae* sp. nov. *J. Fish Biol.* 82, 1351–1375. doi: 10.1111/jfb.12079

Dementyeva, P. V., Trifonov, V. A., Kulemzina, A. I., and Graphodatsky, A. S. (2010). Reconstruction of the putative Cervidae ancestral karyotype by chromosome painting of Siberian roe deer (*Capreolus pygargus*) with dromedary probes. *Cytogenet. Genome Res.* 128, 228–235. doi: 10.1159/000298878

Eschmeyer, W. N., and Fong, J. D. (2017). *Species by Family/Subfamily*. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp [accessed March 30, 2017].

Fantinatti, B. E., Mazzuchelli, J., Valente, G. T., Cabral-de-Mello, D. C., and Martins, C. (2011). Genomic content and new insights on the origin of the B chromosome of the cichlid fish *Astatotilapia latifasciata*. *Genetica* 139, 1273–1282. doi: 10.1007/s10709-012-9629-x

Freitas, N. L., Al-Rikabi, A. B. H., Bertollo, L. A. C., Ezaz, T., Yano, C. F., Oliveira, E. A., et al. (2017). Early stages of XY sex chromosomes differentiation in the fish *Hoplias malabaricus* (Characiformes, Erythrinidae) revealed by DNA repeats accumulation. *Curr. Genomics* 18. doi: 10.2174/1389202918666170711160528

Gornung, E. (2013). Twenty years of physical mapping of major ribosomal RNA genes across the teleosts: a review of research. *Cytogenet. Genome Res.* 141, 90–102. doi: 10.1159/000354832

Kareem, O. K., Olanrewaju, A. N., Osho, E. F., Orisasona, O., and Akintunde, M. A. (2016). Growth patterns and condition factor of *Hepsetus odoe* (Bloch, 1794) captured in eleyele lake, Southwest Nigeria. *Fish. Aquacult. J.* 7:178. doi: 10.4172/2150-3508.1000178

Knytl, M., Kalous, L., Symonová, R., Rylková, K., and Ráb, P. (2013). Chromosome studies of european cyprinid fishes: cross-species painting reveals natural allotetraploid origin of a *Carassius* female with 206 chromosomes. *Cytogenet. Genome Res.* 139, 276–283. doi: 10.1159/000350689

Kubat, Z., Hobza, R., Vyskot, B., and Kejnovsky, E. (2008). Microsatellite accumulation on the Y chromosome in *Silene latifolia*. *Genome* 51, 350–356. doi: 10.1139/G08-024

Kulemzina, A. I., Yang, F., Trifonov, V. A., Ryder, O. A., Ferguson-Smith, M. A., and Graphodatsky, A. S. (2011). Chromosome painting in Tragulidae facilitates the reconstruction of Ruminantia ancestral karyotype. *Chromosome Res.* 19, 531–539. doi: 10.1007/s10577-011-9201-z

Levan, A., Fredga, K., and Sandberg, A. A. (1964). Nomenclature for centromeric position on chromosomes. *Hereditas* 52, 201–220. doi: 10.1111/j.1601-5223.1964.tb01953.x

Majtánová, Z., Choleva, L., Symonová, R., Ráb, P., Kotusz, J., Pekárik, L., et al. (2016). Asexual reproduction does not apparently increase the rate of chromosomal evolution: karyotype stability in diploid and triploid clonal hybrid fish (*Cobitis*, Cypriniformes, Teleostei). *PLOS ONE* 11:e0146872. doi: 10.1371/journal.pone.0146872

Martins, C., and Galetti, P. M. (1999). Chromosomal localization of 5S rDNA genes in *Leporinus* fish (Anostomidae, Characiformes). *Chromosome Res.* 7, 363–367. doi: 10.1023/A:1009216030316

Martins, N. F., Bertollo, L. A. C., Troy, W. P., Feldberg, E., de Souza Valentin, F. C., and Cioffi, M. B. (2013). Differentiation and evolutionary relationships in *Erythrinus erythrinus* (Characiformes, Erythrinidae): comparative chromosome mapping of repetitive sequences. *Rev. Fish Biol. Fish.* 23, 261–269. doi: 10.1007/s11160-012-9292-4

Martinez, J. D. F., Lui, R. L., Traldi, J. B., Blanco, D. R., and Moreira-Filho, O. (2015). Occurrence of natural hybrids among sympatric karyomorphs in *Hoplerythrinus unitaeniatus* (Characiformes, Erythrinidae). *Zebrafish* 12, 281–287. doi: 10.1089/zeb.2015.1083

Moraes, R. L. R., Bertollo, L. A. C., Marinho, M. M. F., Yano, C. F., Hatanaka, T., Barby, F. F., et al. (2017). Evolutionary relationships and cytotaxonomy considerations in the genus *Pyrrhulina* (Characiformes, Lebiasinidae). *Zebrafish* doi: 10.1089/zeb.2017.1465 [Epub ahead of print].

Nelson, J. S., Grande, T. C., and Wilson, M. V. (2016). *Fishes of the World*. Somerset, NJ: John Wiley & Sons Inc. doi: 10.1002/9781119174844

Oliveira, C., Almeida-Toledo, L. F., and Foresti, F. (2007). "Karyotypic evolution in Neotropical fishes," in *Fish Cytogenetics*, eds E. Pisano, C. Ozouf-Costaz, F. Foresti, and B. G. Kapoor (Enfield, CT: Science Publishers), 111–164.

Oliveira, C., Andreata, A. A., Toledo, L. F. A., and Toledo, S. A. (1991). Karyotype and nucleolus organizer regions of *Pyrrhulina* cf australis (Pisces, Characiformes, Lebiasinidae). *Brazilian Journal of Genetics* 14, 685–690.

Oliveira, C., Avelino, G. S., Abe, K. T., Mariguela, T. C., Benine, R. C., Ortí, G., et al. (2011). Phylogenetic relationships within the speciose family Characidae (Teleostei: Ostariophysi: Characiformes) based on multilocus analysis and extensive ingroup sampling. *BMC Evol. Biol.* 11:275. doi: 10.1186/1471-2148-11-275

Oliveira, E. A., Bertollo, L. A. C., Yano, C. F., Cioffi, M. B. (2015). Comparative cytogenetics in the genus *Hoplias* (Characiformes, Erythrinidae) highlights contrasting karyotype evolution among congeneric species. *Mol. Cytogenet.* 8:56. doi: 10.1186/s13039-015-0161-4

Oliveira, E. A., Sember, A., Bertollo, L. A. C., Yano, C. F., Ezaz, T., Moreira-Filho, O., et al. (2017). Tracking the evolutionary pathway of sex chromosomes among fishes: characterizing the unique XX/XY1Y2 system in *Hoplias malabaricus* (Teleostei, Characiformes). *Chromosoma* 1–14. doi: 10.1007/s00412-017-0648-3

Oliveira, E. H. C., De Moura, S. P., Dos Anjos, L. J. S., Nagamachi, C. Y., Pieczarka, J. C., and Ferguson-Smith, M. A. (2008). Comparative chromosome painting between chicken and spectacled owl (*Pulsatrix perspicillata*): implications for chromosomal evolution in the Strigidae (Aves, Strigiformes). *Cytogenet. Genome Res.* 122, 157–162. doi: 10.1159/000163093

Oliveira, E. H. C., Tagliarini, M. M., Rissino, J. D., Pieczarka, J. C., Nagamachi, C. Y., O'Brien, P. C., et al. (2010). Reciprocal chromosome painting between white hawk (*Leucopternis albicollis*) and chicken reveals extensive fusions and fissions during karyotype evolution of Accipitridae (Aves, Falconiformes). *Chromosome Res.* 18, 349–355. doi: 10.1007/s10577-010-9117-z

Ortí, G., and Meyer, A. (1997). The radiation of characiform fishes and the limits of resolution of mitochondrial ribosomal DNA sequences. *Syst. Biol.* 46, 75–100. doi: 10.1093/sysbio/46.1.75

Oyakawa, O. T. (2003). "Family erythrinidae," in *Check List of the Freshwater Fishes of South and Central America*, eds R. E. Reis, S. O. Kullander, and C. J. Ferraris Jr. (Porto Alegre: EDIPUCRS), 238–240.

Pastori, M. C., Roncati, H. C., Aichino, D. R., Ledesma, M. A., and Fenocchio, A. S. (2009) Chromosome characterization and cytotaxonomic considerations on Characidae, Acestrorhynchidae and Cynodontidae (Pisces, Characiformes) from the Paraná River (Argentina). *Caryologia* 62, 69–74. doi: 10.1080/00087114.2004.10589668

Pendás, A. M., Moran, P., Freije, J. P., and Garcia-Vazquez, E. (1994). Chromosomal mapping and nucleotide sequence of two tandem repeats of Atlantic salmon 5S rDNA. *Cytogenet. Genome Res.* 67, 31–36. doi: 10.1159/000133792

Pereira, C. S. A., Aboim, M. A., Ráb, P., and Collares-Pereira, M. J. (2014). Introgressive hybridization as a promoter of genome reshuffling in natural homoploid fish hybrids (Cyprinidae, Leuciscinae). *Heredity* 112, 343–350. doi: 10.1038/hdy.2013.110

Pokorná, M., Kratochvíl, L., and Kejnovský, E. (2011). Microsatellite distribution on sex chromosomes at different stages of heteromorphism and heterochromatinization in two lizard species (Squamata: Eublepharidae: *Coleonyx elegans* and Lacertidae: *Eremias velox*). *BMC Genet.* 12:90. doi: 10.1186/1471-2156-12-90

Porto, J. I. R., Feldberg, E., das Neves Falcao, J., and Nakayama, C. M. (1993). Cytogenetic studies in hemiodidae (Ostariophysi, Characiformes) fishes from the central Amazon. *Cytologia* 58, 397–402. doi: 10.1508/cytologia.58.397

Porto, J. I. R., Feldberg, E., Nakayama, C. M., and Falcão, J. (1992). A checklist of chromosome numbers and karyotypes of Amazonian freshwater fishes. *Rev. D'hydrobiol. Trop.* 25, 287–299.

Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Scheel, J. J. (1973). *Fish Chromosomes and Their Evolution*. Internal Report of Danmarks Akvarium 22, NAID: 10006435147, Charlottenlund.

Silva, D. M., Utsunomia, R., Pansonato-Alves, J. C., Oliveira, C., and Foresti, F. (2015). Chromosomal mapping of repetitive DNA sequences in five species of Astyanax (Characiformes, Characidae) reveals independent location of U1 and U2 snRNA sites and association of U1 snRNA and 5S rDNA. *Cytogenet. Genome Res.* 146, 144–152. doi: 10.1159/000438813

Sumner, A. T. (1972). A simple technique for demonstrating centromeric heterochromatin. *Exp. Cell Res.* 75, 304–306. doi: 10.1016/0014-4827(72)90558-7

Symonová, R., Majtánová, Z., Sember, A., Staaks, G. B. O., Bohlen, J., Freyhof, J., et al. (2013). Genome differentiation in a species pair of coregonine fishes: an extremely rapid speciation driven by stress-activated retrotransposons mediating extensive ribosomal DNA multiplications. *BMC Evol. Biol.* 13:42. doi: 10.1186/1471-2148-13-42

Tagliarini, M. M., O'Brien, P., Ferguson-Smith, M. A., and de Oliveira, E. H. (2011). Maintenance of syntenic groups between Cathartidae and *Gallus gallus* indicates symplesiomorphic karyotypes in new world vultures. *Genet. Mol. Biol.* 34, 80–83. doi: 10.1590/S1415-47572010005000117

Vari, R. P., and Malabarba, L. R. (1998). "Neotropical ichthyology: an overview," in *Phylogeny and Classification of Neotropical Fishes*, eds L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M. S. Lucena, and C. A. S. Lucena (Porto Alegre: EDIPUCRS), 1–12.

Yang, F., Trifonov, V., Ng, B. L., Kosyakova, N., and Carter, N. P. (2009). "Generation of paint probes by flow-sorted and microdissected chromosomes," in *Fluorescence In Situ Hybridization (FISH)–Application Guide*, ed. T. Liehr (Berlin: Springer), 35–52. doi: 10.1007/978-3-540-70581-9_3

Yano, C. F., Bertollo, L. A. C., and Cioffi, M. B. (2017a). "Fish-FISH: molecular cytogenetics in fish species," in *Fluorescence in situ Hybridization (FISH) - Application Guide*, ed. T. Liehr (Berlin: Springer), 429–444.

Yano, C. F., Bertollo, L. A. C., Ezaz, T., Trifonov, V., Sember, A., Liehr, T., et al. (2017c). Highly conserved Z and molecularly diverged W chromosomes in the fish genus *Triportheus* (Characiformes, Triportheidae). *Heredity* 118, 276–283. doi: 10.1038/hdy.2016.83

Yano, C. F., Bertollo, L. A. C., Liehr, T., Troy, W. P., and Cioffi, M. B. (2016). W chromosome dynamics in *Triportheus* species (Characiformes, Triportheidae).

An ongoing process narrated by repetitive sequences. *J. Heredity* 107, 342–348. doi: 10.1093/jhered/esw021

Yano, C. F., Bertollo, L. A. C., Molina, W. F., Liehr, T., and Cioffi, M. B. (2014). Genomic organization of repetitive DNAs and its implications for male karyotype and the neo-Y chromosome differentiation in *Erythrinus erythrinus* (Characiformes, Erythrinidae). *Comp. Cytogenet*. 8, 139–151. doi: 10.3897/CompCytogen.v8i2.7597

Yano, C. F., Bertollo, L. A. C., Rebordinos, L., Merlo, M. A., Liehr, T., Portela-Bens, S., et al. (2017b). Evolutionary dynamics of rDNAs and U2 small nuclear DNAs in *Triportheus* (Characiformes, Triportheidae): high variability and particular syntenic organization. *Zebrafish* 14, 146–154. doi: 10.1089/zeb.2016.1351

Zhu, H. P., and Gui, J. F. (2007). Identification of genome organization in the unusual allotetraploid form of *Carassius auratus* gibelio. *Aquaculture* 265, 109–117. doi: 10.1016/j.aquaculture.2006.10.026

Zwick, M. S., Hanson, R. E., McKnight, T. D., Islam-Faridi, M. N., Stelly, D. M., Wing, R. A., et al. (1997). A rapid procedure for the isolation of C0t-1 DNA from plants. *Genome* 40, 138–142. doi: 10.1139/g97-020

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

Copyright © 2017 Carvalho, de Oliveira, Bertollo, Yano, Oliveira, Decru, Jegede, Hatanaka, Liehr, Al-Rikabi and Cioffi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 19 December 2017
doi: 10.3389/fgene.2017.00214

Population Genetic Structure of *Cnesterodon decemmaculatus* (Poeciliidae): A Freshwater Look at the Pampa Biome in Southern South America

*Aline M. C. Ramos-Fregonezi[1,2], Luiz R. Malabarba[3] and Nelson J. R. Fagundes[1]**

[1] Laboratory of Medical and Evolutionary Genetics, Department of Genetics, Federal University of Rio Grande do Sul, Porto Alegre, Brazil, [2] Laboratory of Bioinformatics and Evolution, Department of General Biology, Federal University of Viçosa, Viçosa, Brazil, [3] Laboratory of Ichthyology, Department of Zoology, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

OPEN ACCESS

Edited by:
Rodrigo A. Torres,
Universidade Federal
de Pernambuco, Brazil

Reviewed by:
Paulo Affonso,
Universidade Estadual do Sudoeste
da Bahia, Brazil
Tomas Hrbek,
Federal University of Amazonas, Brazil

***Correspondence:**
Nelson J. R. Fagundes
nelson.fagundes@ufrgs.br

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

Received: 31 August 2017
Accepted: 04 December 2017
Published: 19 December 2017

Citation:
Ramos-Fregonezi AMC,
Malabarba LR and Fagundes NJR
(2017) Population Genetic Structure
of Cnesterodon decemmaculatus
(Poeciliidae): A Freshwater Look
at the Pampa Biome in Southern
South America. Front. Genet. 8:214.
doi: 10.3389/fgene.2017.00214

The Pampas is a Neotropical biome formed primarily by low altitude grasslands and encompasses the southernmost portion of Brazil, Uruguay, and part of Argentina. Despite the high level of endemism, and its significant environmental heterogeneity, Pampean species are underrepresented in phylogeographic studies, especially aquatic organisms. The Pampean hydrological system resulted from a long history of tectonism, climate, and sea level changes since the Neogene. In this study, we examined the population genetic structure of *Cnesterodon decemmaculatus*, a freshwater fish species that occurs throughout most of the Pampa biome. We characterized mitochondrial and autosomal genetic lineages in populations sampled from Southern Brazil and Uruguay to investigate (1) the correspondence between current drainage systems and evolutionary lineages, (2) the demographic history for each genetic lineage, and (3) the temporal depth of these lineages. Overall, we found that the major evolutionary lineages in this species are strongly related to the main Pampean drainage systems, even though stream capture events may have affected the distribution of genetic lineages among drainages. There was evidence for recent population growth in the lineages occupying drainages closest to the shore, which may indicate the effect of quaternary sea-level changes. In general, divergence time estimates among evolutionary lineages were shallow, ranging from 20,000 to 800,000 years before present, indicating a geologically recent history for this group, as previously reported in other Pampean species. A Bayesian phylogeographical reconstruction suggested that an ancestral lineage probably colonized the Uruguay River Basin, and then expanded throughout the Pampas. This evolutionary scenario may represent useful starting models for other freshwater species having a similar distribution.

Keywords: Bayesian phylogeography, mitochondrial DNA, neotropical ichthyology, Pampa biome, stream capture

INTRODUCTION

The South American Pampa, or "Pampas," is a Neotropical biome dominated by natural grasslands spreading over plains of Uruguay, Northern Argentina, Southern Brazil, and part of Paraguay (Pallarés et al., 2005). However, from the biological standpoint, the Pampas are far from homogeneous. Indeed, many proposals for small biological "provinces" have been made, especially regarding on the different plant communities (Cabrera and Willink, 1980; Overbeck et al., 2007), which reflects the mosaic of soil types resulted from its complex geological history. At a smaller scale, such heterogeneity leads to high species endemism and significant genetic structure in Pampean species studied so far (e.g., Freitas et al., 2012; Turchetto et al., 2014; Felappi et al., 2015). While some phylogeographic studies in the Pampas have found little geographic structure and shallow gene trees [<0.1 million years ago (mya)] (Speranza et al., 2007; Turchetto et al., 2014), other studies found strong geographic structure and a deep mitochondrial gene tree (>2.5 mya) (Felappi et al., 2015). However, these studies did not include freshwater organisms, and, therefore, it is difficult to predict the level and depth of the genetic structure exhibited by these species. Species occurring in the Pampa biome have been underrepresented in studies of phylogeography and conservation genetics (Lawler et al., 2006; Beheregaray, 2008; Turchetto-Zolet et al., 2013), which further complicates the understanding of general drivers of biological diversification in this biome.

The Pampas hydrological system resulted from a long history of tectonism, climate, and sea level changes since the Neogene (Casciotta et al., 1999). During the Quaternary, soil erosion and other marine regression/transgression cycles have continued to shape hydrological systems (Tonni and Cione, 1997; Quattrocchio et al., 2008) by altering the relationship among tributaries, isolated drainages, lagoons, and estuaries (Martin and Dominguez, 1994; Thomaz et al., 2015). These geomorphological processes promoted successive stream captures events, allowing obligate freshwater species to disperse using temporary connections (e.g., Loureiro et al., 2011; Candela et al., 2012; Grassi et al., 2017). Stream capture (also known as river capture, headwater capture, or drainage rearrangement) is a geomorphological process that consists in the contact between neighboring drainages. It occurs when the tributary of a river basin starts flowing toward a neighbor basin, and results in icthyofauna dispersal between them (Bishop, 1995). Temporary connections between adjacent basins caused by fluctuations in the sea level during glacial–interglacial cycles may have also influenced the current distribution of several freshwater fish genera (e.g., Ketmaier et al., 2004; Swartz et al., 2007, 2009; Thomaz et al., 2015), including the Pampean taxa *Australoheros*, *Cnesterodon*, *Jenynsia*, and *Corydoras* (e.g., Bruno et al., 2011, 2016; Ponce et al., 2011).

It is well accepted that the distribution of freshwater fish lineages mainly reflects the paleogeography of a specific region (Bermingham and Avise, 1986; Bernatchez and Wilson, 1998; Avise, 2000; Lévêque et al., 2008). *Cnesterodon decemmaculatus* (Jenyns 1842) is endemic to the Pampa biome and is one of the most widespread freshwater fish species in this region. It is found in the freshwater Ecoregions 332 (lower Uruguay) and 334 (laguna dos Patos basin) described by Abell et al. (2008), but exclusively in grassland environments in the southern part of these ecoregions, associated to the Pampa biome, being absent in the northern portion of these ecoregions located in the Atlantic Forest Biome. It is further found in the freshwater ecoregions 345 (Subtropical Potamic axis) and 347 (Bonaerensean Atlantic), also in the Pampas. Thus, this species constitutes an excellent model to examine historical processes that played a major role in shaping the genetic structure of freshwater fishes in the Pampas. In addition, this species can be found in a wide range of habitats, including rivers, ponds, and shallow wetlands, even though it has a poor swimming capacity in high-speed currents (Trenti et al., 1999). A recent study using *C. decemmaculatus* populations from the southern Pampean region showed low genetic divergence and evidences of recent colonization of this area (Bruno et al., 2016). However, this study focused on the coastal drainages from Argentina (ecoregion 347 – Bonaerensean Atlantic), which represent a small portion of the natural distribution of *C. decemmaculatus* in the Pampas.

In this study, we used mitochondrial and nuclear genetic markers to evaluate the genetic structure of *C. decemmaculatus* over its distribution. More specifically, we were interested in answer the following questions: (1) How strong is the correspondence between current drainage systems and evolutionary lineages? (2) Is there evidence of ancient population growth or reduction for each genetic lineage in this species? (3) What is the temporal depth for the evolutionary divergence among lineages and populations? (4) What is the colonization history for this taxon in the Pampas?

MATERIALS AND METHODS

We sampled 99 individuals from 37 localities (**Figure 1**), covering a significant part of the species' range. Our sampling ranged from one to eight individuals per locality (Supplementary Table S1). All individuals included in our analyses were fixed in 96% ethanol and deposited in the fish collection at the Federal University of Rio Grande do Sul (Universidade Federal do Rio Grande do Sul – UFRGS). All collections were performed under the approval of the government authorities of Brazil and Uruguay (Ministério do Meio Ambiente, Brazil – SISBIO, license number 12038-2, and Dirección de Recursos Naturales del Ministerio de Ganadería, Agricultura y Pesca, Uruguay). The collection and euthanasia of the specimens were approved by the ethics committee of the UFRGS, license number 24434.

Total genomic DNA was isolated from muscle tissue with cetyltrimethyl ammonium bromide (CTAB) as described by Doyle and Doyle (1987). We tested two mitochondrial (mtDNA) genes: cytochrome oxidase subunit I (*COI*) (Herbert et al., 2003) and NADH dehydrogenase 2 (*ND2*) (Sorenson, 2003) and two nuclear (nDNA) genes: SH3 and PX domain containing 3-like protein (*SH3PX3*) and myosin heavy chain 6 (*Myh6*) (Li et al., 2007). However, *SH3PX3* and *COI* had insufficient variation in a preliminary sample of 10 individuals per basin and, therefore, were excluded from further analysis. PCR amplification

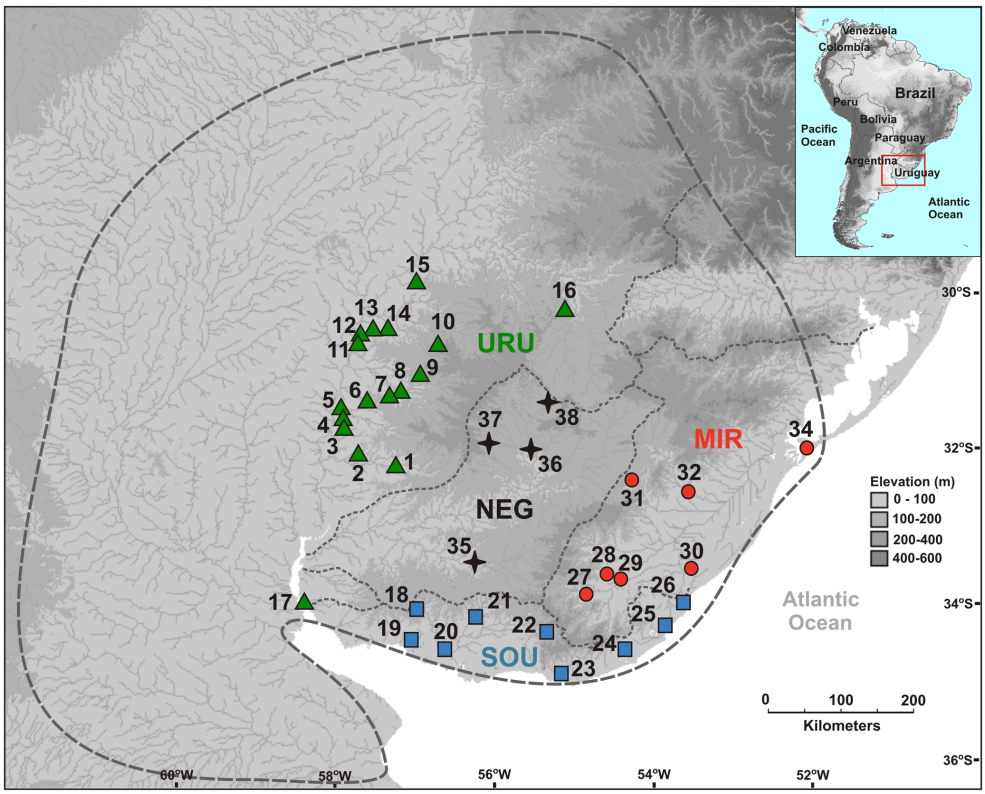


FIGURE 1 | Geographical location of sampled sites showing the limits of the major drainage systems (assumed as different populations). The dashed gray line indicates the distribution of *Cnesterodon decemmaculatus* according to Lucinda, 2005. Vouchers of the specimens used in the study, and the geographical coordinates for sampling sites are available in Supplementary Table S1.

protocols for *ND2* and *Myh6* followed Sorenson (2003) and Li et al. (2007), respectively. PCR products were checked on a 1% agarose gel, purified with Exonuclease I and Shrimp Alkaline Phosphatase (GE Healthcare®) and sequenced using the Sanger method at Macrogen Inc. (Seoul, South Korea). Both DNA strands were sequenced, checked, and aligned with GeneiousPro v.4.8 (Drummond et al., 2007). Haplotypes of the *Myh6* gene were estimated in PHASE 2.1 (Stephens et al., 2001) using 10,000 steps sampling every 10 steps and discarding the first 1000 steps as burnin. We ran PHASE several times using different starting seeds to ensure the reliability of the final estimate. Mitochondrial haplotypes were defined using DnaSP 5.10 (Rozas et al., 2003). All sequences obtained in the present study were deposited in GenBank (KU214332–KU214430 and KU214252–KU214331, for *ND2* and *Myh6* genes, respectively).

Dispersal of obligatory freshwater fish requires connections of aquatic habitats between adjacent basins and, most of taxa have their distribution range narrowly coincident, if not equal, with the hydrographic basins boundaries in which they live (Ward et al., 1994; Albert and Reis, 2011). Based on this, and also because we found a general genetic structure associated to major Pampean drainages, we defined four geographical groups for population analysis: Uruguay River Basin (URU), Negro River Basin (NEG), Mirim Lagoon Basin (MIR), and Southern coastal Basins of Uruguay (SOU) (**Figure 1** and Supplementary Table S1). Because independent drainage systems represent plausible boundaries for isolated biological populations for freshwater organisms, individuals collected in different localities within each basin were merged into a single population, resulting in a sample of 62 individuals for URU, 11 for NEG, 15 for MIR, and 11 for SOU. Following Abell et al. (2008), URU and NEG represent the freshwater Ecoregion 332 (lower Uruguay), MIR represents Ecoregion 334 (laguna dos Patos), while SOU have populations located in both Ecoregions 334 and 345 (lower Paraná).

To determine the correspondence between current drainage systems and evolutionary lineages, we inferred the evolutionary relationship among haplotypes using a median-joining network (Bandelt et al., 1999) estimated in NETWORK 4.1.0.9[1]. The resulting genetic structure was quantified using an analysis of molecular variance (AMOVA) (Excoffier et al., 1992) and pairwise Φ_{ST}s calculated in Arlequin 3.5 (Excoffier and Lischer, 2010). In addition, we tested for population structure without assuming any hypothesis based on hydrography using the Bayesian clustering algorithm implemented in BAPS (Corander et al., 2008). We used the mixture population model and set the maximum number of populations (K) to 37, the number of sampled local populations. Because stream capture events may affect gene genealogies and molecular diversity patterns

[1]www.fluxus-engineering.com

within drainages, we decided to run several analyses excluding localities for which we may have evidence for stream capture. We inferred a possible stream capture event whenever we found shared haplotypes between neighbor drainages, especially when populations closely related to watersheds were involved (see the section "Discussion" for further details for each case).

To test for evidence of ancient population growth or reduction for each population (URU, NEG, MIR, and SOU) we first estimated descriptive statistics of genetic diversity, such as nucleotide (π) and haplotype diversity (H), followed by Tajima's D (Tajima, 1983) and Fu's F_S (Fu, 1997) neutrality tests, which were calculated in Arlequin 3.5 (Excoffier and Lischer, 2010). We performed neutrality tests under two schemes: considering all individuals for a given basin, and excluding localities for which there was evidence of headwater capture. We also estimated the effective population size (N_E) and the population growth parameter (G) for each population without the localities with evidence for headwater capture (namely URU 7, URU 8, NEG 35, NEG 36, MIR 31), using both markers in the program LAMARC v. 2.1.6 (Kuhner, 2006). We performed a maximum-likelihood search based on three replicates of 10 initial chains and two final chains. The initial chains consisted of 250 samples drawn every 20 steps, and a burn-in of 1000 samples for each chain. The two final chains used the same burn-in and sampling interval, but consisted of 10,000 samples. We assumed a generation time of 1 year based on estimates of sexual maturity for the poeciliid *Poecilia reticulata* (Reznick and Bryga, 1987).

Finally, we inferred the divergence times among populations and the colonization history for *C. decemmaculatus* based on two approaches. First, we estimated a time-calibrated gene-tree genealogy for mtDNA lineages using Bayesian inference in BEAST 1.7.5 (Drummond and Rambaut, 2007). We used a coalescent constant-size tree prior with a random starting tree and the TN93+G model of sequences evolution, as determined by the corrected Akaike Information Criterion (AICc) in jModelTest2 (Darriba et al., 2012). We used each drainage as a discrete trait to allow the estimation of the most likely location of all ancestors in the mtDNA tree. While this may be indicative of dispersal events, it should not be viewed as a formal test of stream capture hypotheses. The MCMC was run twice for 10 million generations each, and 10% of the samples was discarded as burn-in. A normal strict molecular clock was assumed, calibrated with a normal distribution for molecular substitution rate of $8.6 \pm 0.1 \times 10^{-9}$ substitutions per site per year (s/s/y), which has been suggested in the literature for Cyprinodontiformes based on a dataset which included the *ND2* gene (Hrbek and Meyer, 2003).

Second, we used a coalescent-based species-tree analysis (Heled and Drummond, 2010) to infer the colonization history of each river basin using the StarBeast model in the program BEAST 1.7.5 (Drummond and Rambaut, 2007). For this analysis, we used populations as terminals, and used both markers with unlinked gene-trees (mitochondrial – *ND2* and nuclear – *Myh6*). We excluded localities showing evidence for headwater capture, as the species-tree method does not allow gene flow among terminals. We used a rate-reference prior to calibrate the molecular clock for *Myh6* (Ferreira and Suchard, 2008). Under this approach, a time-calibrated genealogy (i. e., *ND2*) is taken as a reference to estimate the relative molecular clock associated with alternative gene trees (i.e., *Myh6*). This is a useful approach when different partitions evolve at different rates, but there is no prior information on the molecular clock for all independent partitions in the matrix (Ferreira and Suchard, 2008). All remaining parameters and priors followed the Bayesian phylogenetic analyses mentioned above. In all Bayesian analysis, sampling sufficiency was evaluated by monitoring effective sample size (ESS) and ensuring that all values were higher than 200.

RESULTS

Ninety-nine sequences were obtained for the mtDNA *ND2* gene (980 bp) yielding 31 different haplotypes defined by 42 variable sites (28 parsimoniously informative). For the autosomal *Myh6* gene we had 188 sequences (681 bp) resulting in 25 different haplotypes defined by 21 variable sites (16 parsimoniously informative). In general, the mtDNA haplotype network showed little haplotype sharing among populations, and a close relationship among haplotypes from the same population (**Figure 2A**). Populations of MIR and SOU exhibited related haplotypes, except for two SOU localities that had haplotypes H21 and H30, which are more closely related to haplotypes found in URU. Concerning *Myh6* haplotypes (**Figure 2B**), URU, MIR, and SOU shared the most frequent haplotypes (H5 and H1), and H9 was the only haplotype present in populations of all drainage systems. The close relationship among haplotypes for both markers is suggestive of a shallow genealogical history for this species (see below).

Considering all samples, the Bayesian clustering analysis found K = 6 genetic populations (**Figure 2C**). There was a marked difference in the number of different genetic population occurring in each major drainage. Five genetic populations occurred in URU, three in NEG, and two in both SOU and MIR. Not considering samples possibly affected by stream capture events, we found four genetic populations in URU, two in SOU, and one in both NEG and MIR, suggesting significant genetic structure within URU mainly due to three genetic groups with a restricted geographical distribution in this drainage (**Figure 2C**).

The mtDNA genealogy for *C. decemmaculatus* (**Figure 3**) suggests the most recent common ancestor (MRCA) of all mitochondrial lineages dates from 0.8 mya [95% credible interval (CI) 0.5–1.2 mya], and were most probably located in URU (PP = 0.58). A well-supported clade (posterior probability PP = 1.00), sister to all other lineages, occurred in isolated populations inhabiting elevated plains from three basins: URU, MIR, and NEG (haplotypes H11, H12, H14, and H19). These haplotypes represented one of the Bayesian genetic populations. However, the location of the MRCA for this clade was uncertain (PP < 0.5 for all locations). Its sister clade, which had moderate support (PP = 0.82) had URU as the most likely location (PP = 0.75), and showed H25 (another Bayesian genetic population) sister to a well-supported clade (PP = 0.99) that also has URU as its most likely location (PP = 0.80). In turn, this clade harbored four Bayesian genetic populations that corresponded to

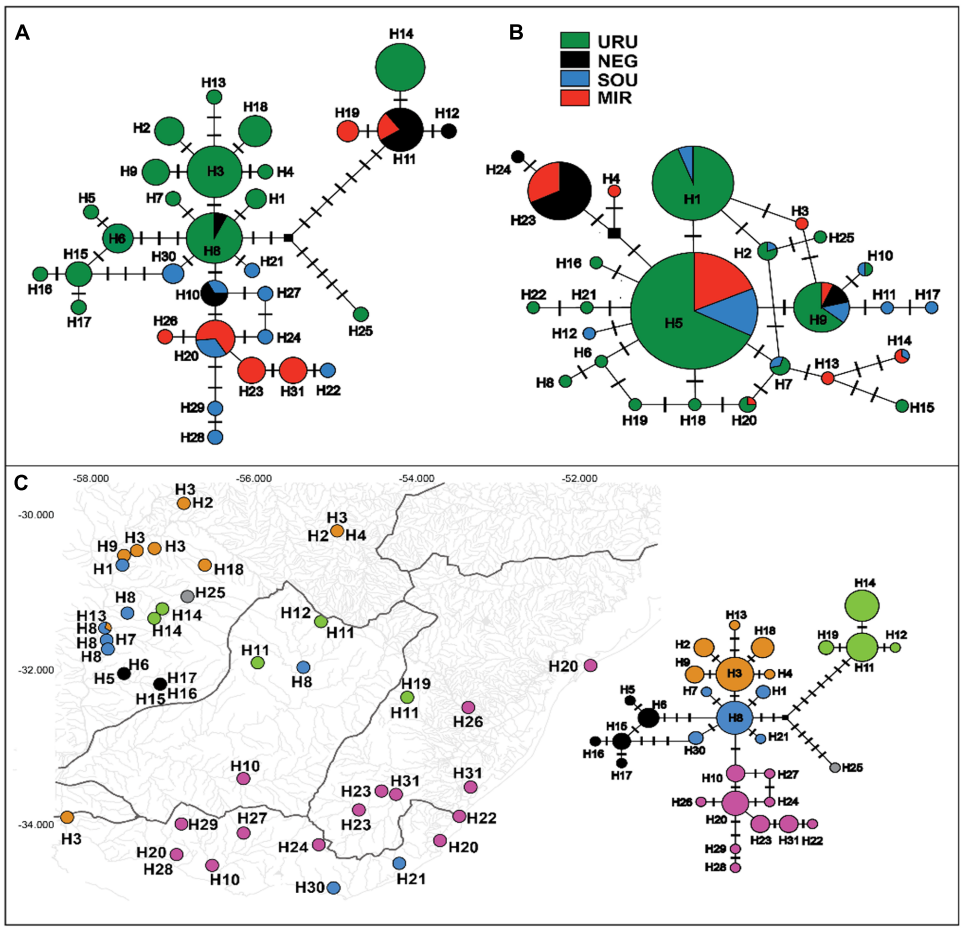


FIGURE 2 | Median-joining networks for **(A)** mtDNA *ND2* haplotypes, and **(B)** nDNA *Myh6* haplotypes. Circles size are proportional to the observed frequency of each haplotype. Cross marks represent mutational differences between haplotypes. **(C)** Geographic location of each mtDNA haplotype. Different colors represent the six genetic groups identified in BAPS analysis.

three well-supported clades ($PP > 0.90$) plus a "mixed" genetic population that included H1, H7, H8, H30, and H21. Regarding the most likely ancestral location, all remaining genetic groups had URU as the most likely location ($PP > 0.85$), with one exception. The clade, which was almost exclusive to MIR and SOU (though H10 also occurred in NEG), had SOU or MIR as its most likely ancestral location ($PP = 0.46$, $PP = 0.30$, respectively).

Considering the full dataset, both markers showed high haplotype diversities ($H = 0.94$ for *ND2* and $H = 0.79$ for *Myh6*). Both haplotype and nucleotide diversity were higher for URU and SOU compared with NEG and MIR (**Table 1**). There was significant genetic population structure (*ND2* $\Phi_{ST} = 0.26$, and *Myh6* $\Phi_{ST} = 0.30$, $P < 0.001$). Excluding localities that may represent stream capture events increased Φ_{ST} estimates (*ND2* $\Phi_{ST} = 0.59$, and *Myh6* $\Phi_{ST} = 0.35$, $P < 0.001$). For *ND2*, this was still lower than the Φ_{ST} estimate among the genetic populations delimited with BAPS ($\Phi_{ST} = 0.75$, $P < 0.001$), which is expected, given that Bayesian populations were defined based on ND2 data. However, this trend was not followed by *Myh6* ($\Phi_{ST} = 0.19$, $P < 0.001$). Pairwise Φ_{ST} values indicated that geographic groups were significantly different from each other except MIR vs. SOU (**Table 2**). Taking into account possible events of stream capture resulted in increased Φ_{ST} values between NEG and the remaining populations for both genetic markers, and between the other pairs only for *ND2* (**Table 2**). Because localities from URU might represent different genetic populations, we also quantified the genetic structure between them. For both *ND2* and *Myh6*, we found significant genetic structure in URU (0.81 and 0.39, respectively, $P < 0.001$), which was robust to the exclusion of some localities due to stream capture events (0.64 and 0.31, respectively, $P < 0.001$).

Neutrality tests provided mixed evidence for population growth in some populations. However, they were dependent on which statistic or genetic marker were used (**Table 1**). On the other hand, maximum likelihood-based demographic estimates revealed population growth for SOU and MIR (**Table 3**). Effective population size was also larger for SOU and MIR ($\sim 3 \times 10^6$ effective individuals), while NEG had the lowest value ($\sim$180,000 effective individuals; **Table 3** and **Figure 4**). The combined analysis of mtDNA and nDNA markers suggested that the split

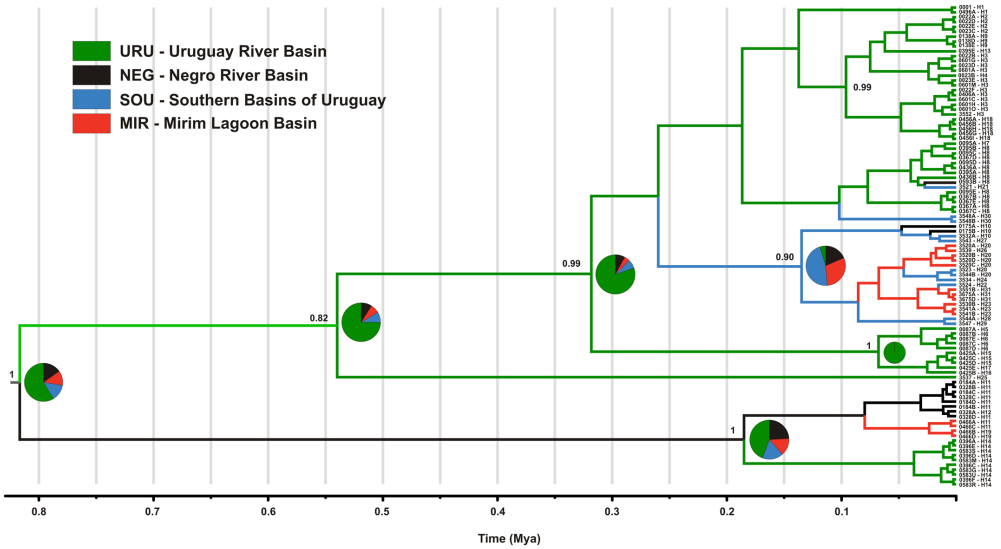


FIGURE 3 | Maximum clade credibility tree among *ND2* lineages found in *Cnesterodon decemmaculatus* individuals. Posterior probabilities (*PP*) are shown above the branches for selected nodes with $PP > 0.8$. Pie charts beside selected ancestral node represent the *PP* of ancestral location among each of the four populations according to the color scheme shown in the inlet.

TABLE 1 | Diversity indexes obtained for the mitochondrial and nuclear markers, respectively (*ND2* values/*Myh6* values), for the four populations considered in the study.

Population	*N*	*H*	π	Tajima's *D*	Fu's F_S	Tajima's *D*‡	Fu's F_S‡
URU	62/60	0.89/0.74	0.006/0.001	−0.48/−1.40	0.17/−9.36**	−1.43/−1.49*	−1.90/−10.17**
NEG	11/9	0.60/0.46	0.004/0.001	0.78/−0.62	2.93/0.93	−1.05/0.15	−0.18/0.55
MIR	15/14	0.86/0.71	0.005/0.002	1.06/−0.29	2.27/−1.67	0.46/−0.71	−0.41/−1.94
SOU	11/11	0.96/0.79	0.003/0.002	−0.63/−1.13	−4.39**/−5.24**	–	–
All samples	99/94	0.94/0.79	0.007/0.002	−0.75/−1.46*	−6.40/−16.76**	–	–

*N, number of individuals sampled for ND2/Myh6 markers; H, haplotype diversity; π, nucleotide diversity; *P < 0.05; **P < 0.02; ‡excluding localities showing evidence of stream capture (see text).*

between NEG and the clade URU+MIR+SOU, which was also the root of the tree, dates back from ∼0.6 mya (95% CI 0.27–0.95 mya). This was followed by the divergence between URU and MIR+SOU around 0.12 mya (95% CI 0.04–0.19 mya), and the subsequent divergence between MIR and SOU <30,000 years ago (95% CI 23,000–57,000) (**Figure 4**).

DISCUSSION

In this study, we characterized the major genetic structure patterns of populations of *C. decemmaculatus* in most of its distribution in the Pampa biome. Differently from the star-like pattern found for Argentinean populations of *C. decemmaculatus* in the southern Pampas (Bruno et al., 2016), which indicates a very recent evolutionary history, our results reveal a significant genetic structure that, in general, paralleled broad drainage systems, especially considering mtDNA lineages (**Figures 2A**, **3**). However, there is no simple relationship between clades and basins, such that a single basin may contain more than one mtDNA clade, and different basins may share the same mtDNA haplotype (**Figures 2A**, **3**). Thus, understanding population structure imposed by the drainage system is necessary, but not sufficient, to account for the genetic patterns observed in our study. This is reinforced by the Bayesian clustering of populations, which suggested $K = 6$ as the bet number of genetic populations. Three alternative (but not exclusive)

TABLE 2 | Pairwise Φ_{ST} values for mtDNA *ND2* (lower diagonal) and nDNA *Myh6* (upper diagonal) markers.

Populations	URU	NEG	MIR	SOU
All locations				
URU	–	0.59	0.10	0.04
NEG	0.32	–	0.37	0.55
MIR	0.19	0.33	–	0.04*
SOU	0.18	0.58	0.11	–
Locations excl. SC[1]				
URU	–	0.65	0.01*	0.02*
NEG	0.71	–	0.67	0.66
MIR	0.41	0.93	–	−0.02*
SOU	0.33	0.88	0.14	–

$P < 0.05$ for all values except when noted by an asterisk (*). [1]Excluding locations showing evidence of stream capture (see text).

TABLE 3 | Estimates for the effective population size (N_E) and growth parameter (G) in all populations.

Population	Theta (θ) (95% CI)	N_E (95% CI)	G (95% CI)
URU	1.8×10^{-2}	1.38×10^{6}	–
	$(1.5 \times 10^{-2}$–$2.1 \times 10^{-2})$	$(1.18 \times 10^{6}$–$1.63 \times 10^{6})$	
NEG	2.4×10^{-3}	1.82×10^{5}	–
	$(1.6 \times 10^{-3}$–$3.7 \times 10^{-3})$	$(1.22 \times 10^{5}$–$2.86 \times 10^{5})$	
MIR	4.4×10^{-2}	3.39×10^{6}	2994.07
	$(1.5 \times 10^{-2}$–$9.7 \times 10^{-2})$	$(1.12 \times 10^{6}$–$7.44 \times 10^{6})$	(1314.72–5313.02)
SOU	4.7×10^{-2}	3.62×10^{6}	1015.58
	$(2.3 \times 10^{-2}$–$10.1 \times 10^{-2})$	$(1.78 \times 10^{6}$–$7.73 \times 10^{6})$	(393.40–1579.73)

Theta (θ) = $4N_E\mu$, where μ is the mutation rate per site per generation. 95% CI–95% credible interval for the estimates. There was no evidence for growth in URU and NEG (see the section "Materials and Methods").

scenarios that may explain the discrepancy between drainage and genetic structure are: (1) migration among neighbor populations (active dispersal), (2) headwater capture (geodispersal), or (3) shared ancestral polymorphism. The first scenario seems unlikely because *C. decemmaculatus* has a poor swimming capacity in high-speed currents (Trenti et al., 1999), being unable to disperse actively across long distances. On the other hand, *Cnesterodon* species are commonly found in shallow wetland habitats (Maltchik et al., 2014), suggesting that geodispersal may represent a more likely alternative for these species. In this case, wetland habitats close to the watersheds between different drainages may facilitate headwater capture events, as minor environmental changes may reshape the hydrological network at these sites, impeding a clear-cut relationship between genetic lineages and drainages.

However, distinguishing between the geodispersal and ancestral polymorphism hypotheses may be more nuanced. This is further complicated by the fact that it is very difficult to set up an explicit test of stream capture hypotheses. One example of a likely case of stream capture involving closely related haplotypes occurring in different drainages is represented by mtDNA haplotypes H11, H14, and H19 (**Figure 3**). Even though URU has received higher support as the most probable location of the MRCA, mtDNA network analysis indicates that this clade evolved in NEG with subsequent stream capture events toward URU and MIR. First, there are seven exclusive substitutions in this clade (**Figure 2A**), suggesting it may have evolved in relative isolation. Second, the sampling points in URU and MIR containing these lineages are adjacent to NEG. Third, H14 and H19 are probably descendent from H11, which occurs in NEG (**Figure 2A**). Under this interpretation, this clade would reflect three independent stream capture events: NEG would have been colonized from URU between 0.2 and 0.8 mya (see below), while MIR and URU would have received migrants from NEG more recently (<0.2 mya). Even in the alternative scenario in which URU is the true location for the MRCA of this clade, two headwater capture events would have occurred, since this lineage would have reached NEG from URU by ∼0.2 mya, dispersing to MIR later on. Two other examples of haplotype sharing that may reflect stream capture toward NEG include H8 (from URU) and H10 (from SOU). In both cases, these haplotypes are more closely related to other lineages from specific drainages, and their sequence identity suggests a recent arrival in NEG. However,

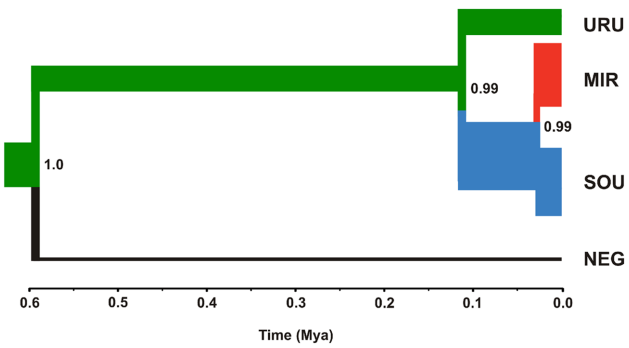


FIGURE 4 | Maximum clade credibility tree considering *ND2* and *Myh6* under a coalescent approach. The thickness of the branches is proportional to effective population size (N_E) estimates for each geographical group (see **Table 3** for more details). The color scheme for each branch followed the estimates of the most likely location for ancestral nodes, shown in **Figure 3**.

while H10 occurs in a site close to SOU, H8 occurs relatively distant from the watershed with URU. An alternative hypothesis for the occurrence of H8 in NEG would involve upstream active dispersal from lower URU populations despite the low dispersal ability of *Cnesterodon* (Trenti et al., 1999).

The tectonic activity associated with stream capture events in eastern South America may be as old as 1.6 mya (Saadi et al., 2002). The presence of H11 and H19 in MIR and H8 in NEG is in agreement with a pattern in which more eastern drainages capture upland shield drainages. However, the presence of H14 in URU and H10 in NEG indicates that these events could also occur in the opposite direction, suggesting a highly dynamic scenario in the Pampas, possibly because this biome is largely flat. Indeed, Loureiro et al. (2011) suggested drainage rearrangements in both directions across the MIR/NEG watershed based on the distribution pattern of the killifish *Austrolebias* in Uruguay. Similarly to *Cnesterodon*, *Austrolebias* inhabit shallow wetland habitats, and have little capabilities for active dispersal (Maltchik et al., 2014), reinforcing the relevance of passive geodispersal in the Pampas. In contrast, the most parsimonious explanation for the presence of haplotypes H21 and H30 in SOU is incomplete lineage sorting, given the low phylogenetic signal associated with location and clade structure for these haplotypes, and because the sampling points

in which these haplotypes were found are distant from the URU watershed.

Our results suggest that the MRCA of *C. decemmaculatus* reached the lower Uruguay before 0.6 myr (**Figure 4**). The mtDNA genealogy suggests that from URU, *C. decemmaculatus* would have colonized NEG and SOU basins independently (**Figure 3**). The high genetic diversity in URU is consistent with its role as the ancestral location. Even though we have found a case of headwater capture from NEG to MIR, the most likely colonization route leading to MIR is through SOU, given that most haplotypes found in SOU and MIR belong to the same mtDNA haplogroup. The Bayesian delimitation of genetic populations also suggested that most localities in MIR and SOU belong to the same genetic population, which is also in agreement with pairwise indices of genetic structure (**Table 2**). These finding also highlight that, for this species, the genetic diversity in the lower Uruguay ecoregion would be much higher than in Laguna dos Patos (*sensu* Abell et al., 2008). Furthermore, given that URU is neighbor to SOU, but not MIR, a colonization route from URU to MIR and then SOU is less likely. Indeed, Bruno et al. (2016) suggested that the Argentinean coastal populations of *C. decemmaculatus* descend from upstream URU populations associated with the Rio de la Plata mouth, which highlight the mouth of the Uruguay river as a putative source for populations currently inhabiting isolated drainages flowing to Río de la Plata. Alternatively, SOU/MIR could have been colonized from NEG. This hypothesis would be strengthened if we assume that H10 evolved *in situ* in NEG before dispersing eastward. Additional sampling efforts, especially in NEG, but also in regions not sampled by this study, such as the Argentinian coast, will be required to refine these scenarios, as well as explicitly testing for hypotheses of stream capture or incomplete lineage sorting.

While the Pampean region has been always dominated by grasslands during the whole Pleistocene (Behling et al., 2005), changes in precipitation regimes and in the sea-level may have affected populations of *C. decemmaculatus*. For example, it could be expected that periods of increased humidity would have favored the formation of new wetland areas and promoted population expansion (Zemlak et al., 2008; Jones and Johnson, 2009). On the other hand, marine transgressions would have caused local extinctions (Villwock and Tomazelli, 1995). In this regard, the high genetic diversity in SOU was surprising, since this area suffered at least three marine transgression events since the Pleistocene (Villwock and Tomazelli, 1995) that could have led to declines in population size and local extinction. The high N_E and the significant negative values of Fu's neutrality test for mtDNA indicate that either local extinction did not affect these populations, or that the high genetic diversity resulted from recent and strong population growth (**Table 3**). We also found evidence for population growth in MIR. However, differences in genetic diversity and N_E estimates between MIR and SOU may reflect different population histories after marine transgressions. Alternatively, some founder effect may have reduced genetic diversity in MIR following its colonization from SOU. Coastal populations from SOU may have also benefited from an extended coastal plain during marine retractions, facilitating population growth during these periods, as have been suggested for other species occurring along the coastal rivers of southern and southeastern Brazil (Thomaz et al., 2015).

AUTHOR CONTRIBUTIONS

AR-F, LM, and NF designed the study, interpreted the results, and contributed to the final version of the manuscript. AR-F did the laboratory work. AR-F and NF analyzed the data.

FUNDING

This research was supported by Conselho Nacional de Desenvolvimento Científico (CNPq), Programa de Pós-graduação em Biologia Animal (PPGBAN-UFRGS), and Programa de Pós Graduação em Genética e Biologia Molecular (PPGBM-UFRGS).

ACKNOWLEDGMENTS

The authors would like to thank Marcelo Loureiro Barrela (Universidad de la República, Uruguay), Juliana Wingert, Alice Hirschmann, and Juliano Ferrer, who provided specimens for study or assistance during field work. They also thank Jéferson Fregonezi, Andréa Thomaz, and three reviewers for contributions in a preliminary version of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2017.00214/full#supplementary-material

REFERENCES

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., et al. (2008). Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. *BioScience* 58, 403–414. doi: 10.1641/B580507

Albert, J. S., and Reis, R. (2011). *Historical Biogeography of Neotropical Freshwater Fishes*. Berkeley, CA: University of California Press. doi: 10.1525/california/9780520268685.001.0001

Avise, J. C. (2000). *Phylogeography: The History and Formation of Species*. Cambridge: Harvard University Press.

Bandelt, H. J., Forster, P., and Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48. doi: 10.1093/oxfordjournals.molbev.a026036

Beheregaray, L. B. (2008). Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Mol. Ecol.* 17, 3754–3774. doi: 10.1111/j.1365-294X.2008.03857.x

Behling, H., Pillar, V., and Bauermann, S. G. (2005). Late quaternary grassland (Campos), gallery forest, fire and climate dynamics, studied by pollen, charcoal and multivariate analysis of the São Francisco de Assis core in western Rio Grande do Sul (southern Brazil). *Rev. Palaeobot. Palynol.* 133, 235–248. doi: 10.1016/j.revpalbo.2004.10.004

Bermingham, E., and Avise, J. C. (1986). Molecular zoogeography of freshwater fishes in the southeastern United States. *Genetics* 113, 939–965.

Bernatchez, L., and Wilson, C. C. (1998). Comparative phylogeography of nearctic and palearctic freshwater fishes. *Mol. Ecol.* 7, 431–452. doi: 10.1046/j.1365-294x.1998.00319.x

Bishop, P. (1995). Drainage rearrangement by river capture, beheading and diversion. *Progr. Phys. Geogr.* 19, 449–473. doi: 10.1177/030913339501900402

Bruno, M. C., Lizarralde, M., Almirón, A., and Casciotta, J. (2011). Presence of *Australoheros facetus* (Teleostei: Cichlidae) in the southern Pampean area. Considerations about the ichthyofaunal settlement and distribution in the Southern boundary of the Brazilian Subregion. *Ichthyol. Contrib. Peces Criollos.* 23, 1–3.

Bruno, M. C., Mapelli, F. J., Casciotta, J. R., Almirón, A. E., and Lizarralde, M. S. (2016). Phylogeography of *Cnesterodon decemmaculatus* (Cyprinodontiformes: poeciilidae) in Southern Pampas, Argentina: ancient versus recent patterns in freshwater fishes. *Environ. Biol. Fish.* 99, 293–307. doi: 10.1007/s10641-016-0474-0

Cabrera, A. L., and Willink, A. (1980). *Biogeografía de América Latina*. Washington, DC: OEA.

Candela, A. M., Bonini, R. A., and Noriega, J. I. (2012). First continental vertebrates from the marine Parana Formation (Late Miocene, Mesopotamia, Argentina): chronology, biogeography, and paleoenvironments. *Geobios* 45, 515–526. doi: 10.1016/j.geobios.2012.05.003

Casciotta, J., Almirón, A., Cione, A., and Azpelicueta, M. (1999). Brazilian freshwater fish assemblages from southern Pampean area, Argentina. *Biogeographica.* 75, 67–78.

Corander, J., Marttinen, P., Sirén, J., and Tang, J. (2008). Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9:539. doi: 10.1186/1471-2105-9-539

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9:772. doi: 10.1038/nmeth.2109

Doyle, J. J., and Doyle, J. L. (1987). A rapid isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 9, 11–15.

Drummond, A. J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., et al. (2007). *Geneious v4.8*.

Drummond, A. J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214. doi: 10.1186/1471-2148-7-214

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Excoffier, L., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes – Application to human mitochondrial-DNA restriction data. *Genetics* 131, 479–491.

Felappi, J. F., Vieira, R. C., Fagundes, N. J., and Verrastro, L. V. (2015). So far away, yet so close: strong genetic structure in *Homonota uruguayensis* (Squamata, Phyllodactylidae), a species with restricted geographic distribution in the Brazilian and uruguayan pampas. *PLOS ONE* 10:e0118162. doi: 10.1371/journal.pone.0118162

Ferreira, M. A., and Suchard, M. A. (2008). Bayesian analysis of elapsed times in continuous-time Markov chains. *Can. J. Stat.* 36, 355–368. doi: 10.1002/cjs.5550360302

Freitas, T. R. O., Fernandes, F. A., Fornel, R., and Roratto, P. A. (2012). An endemic new species of tuco-tuco, genus *Ctenomys* (Rodentia: Ctenomyidae), with a restricted geographic distribution in Southern Brazil. *J. Mammal.* 93, 1355–1367. doi: 10.1644/12-MAMM-A-007.1

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.

Grassi, D. J., Swarça, A. C., Dergam, J. A., Pastori, M. C., and Fenocchio, A. S. (2017). Cytogenetic characterization of *Hoplias malabaricus* (Bloch, 1794) from the Ctalamochita River (Córdoba, Argentina): first evidence for southernmost populations of this species complex and comments on its biogeography. *Comp. Cytogenet.* 11, 15–28. doi: 10.3897/CompCytogen.v11i1.10262

Heled, J., and Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27, 570–580. doi: 10.1093/molbev/msp274

Herbert, P. D. N., Cywinska, A., Ball, S., and Dewaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc. R. Soc. B* 270, 313–321. doi: 10.1098/rspb.2002.2218

Hrbek, T., and Meyer, A. (2003). Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). *J. Evol. Biol.* 16, 17–36. doi: 10.1046/j.1420-9101.2003.00475.x

Jones, C. P., and Johnson, J. B. (2009). Phylogeography of the livebearer *Xenophallus umbratilis* (Teleostei: Poeciliidae): glacial cycles and sea level change predict diversification of a freshwater tropical fish. *Mol. Ecol.* 18, 1640–1653. doi: 10.1111/j.1365-294X.2009.04129.x

Ketmaier, V., Bianco, P. G., Cobolli, M., Krivokapic, M., Caniglia, R., and Matthaeis, E. (2004). Molecular phylogeny of two lineages of Leuciscinae cyprinids (Telestes and Scardinius) from the peri-Mediterranean area based on cytochrome b data. *Mol. Phylogenet. Evol.* 32, 1061–1071. doi: 10.1016/j.ympev.2004.04.008

Kuhner, M. K. (2006). LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics* 22, 768–770. doi: 10.1093/bioinformatics/btk051

Lawler, J. J., Aukema, J. E., Grant, J. B., Halpern, B. S., Kareiva, P., Nelson, C. R., et al. (2006). Conservation science: a 20-year report card. *Front. Ecol. Environ.* 4:473–480. doi: 10.1890/1540-9295(2006)4[473:CSAYRC]2.0.CO;2

Lévêque, C., Oberdorff, T., Paugy, D., Stiassny, M. L. J., and Tedesco, P. A. (2008). Global diversity of fish (Pisces) in freshwater. *Hydrobiologia* 595, 545–567. doi: 10.1007/s10750-007-9034-0

Li, C., Ortí, G., Zhang, G., and Lu, G. (2007). A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evol. Biol.* 7:44. doi: 10.1186/1471-2148-7-44

Loureiro, M., Duarte, A., and Zarucki, M. (2011). A new species of *Austrolebias* costa (cyprinodontiformes: rivulidae) from northeastern Uruguay, with comments on distribution patterns. *Neotrop. Ichthyol.* 9, 335–342. doi: 10.1590/S1679-62252011000200010

Maltchik, L., Lanés, L. E. K., Keppeler, F. W., Rolon, A. S., and Stenert, C. (2014). Landscape and habitat characteristics associated with fish occurrence and richness in southern Brazil palustrine wetland systems. *Environ. Biol. Fish.* 97, 297–308. doi: 10.1007/s10641-013-0152-4

Martin, L., and Dominguez, J. M. L. (1994). "Geological history of coastal lagoons," in *Coastal Lagoon Processes*, ed. B. Kjerfve (Amsterdam: Elsevier), 41–68.

Overbeck, G. E., Muller, S. C., Fidelis, A., Pfadenhauer, J., Pillar, V. D., Blanco, C. C., et al. (2007). Brazil's neglected biome: the South Brazilian campos. *Perspect. Plant Ecol. Evol. Syst.* 9, 101–116. doi: 10.1016/j.ppees.2007.07.005

Pallarés, O. R., Beretta, E. J., and Maraschin, G. E. (2005). "The south american campos ecosystem," in *Grasslands of the World*, eds J. Suttie, S. G. Reynolds, and C. Batello (Rome: Food and agriculture organization of the United Nations), 171–219.

Ponce, J. F., Rabassa, J., Coronato, A., and Borromei, A. M. (2011). Palaeogeographical evolution of the Atlantic coast of Pampa and Patagonia from the last glacial maximum to the Middle Holocene. *Biol. J. Linn. Soc.* 103, 363–379. doi: 10.1111/j.1095-8312.2011.01653.x

Quattrocchio, M. E., Borromei, A. M., Deschamps, C. M., Grill, S. C., and Zavala, C. A. (2008). Landscape evolution and climate changes in the Late Pleistocene–Holocene, southern Pampa (Argentina): evidence from palynology, mammals and sedimentology. *Quat. Int.* 181, 123–138. doi: 10.1016/j.quaint.2007.02.018

Reznick, D. N., and Bryga, H. (1987). Life-history evolution in guppies (*Poecilia reticulata*): 1. Phenotypic and genetic changes in an introduction experiment. *Evolution* 41, 1370–1385. doi: 10.1111/j.1558-5646.1987.tb02474.x

Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X., and Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496–2497. doi: 10.1093/bioinformatics/btg359

Saadi, A., Machette, M. N., Haller, K. M., Dart, R. L., Bradley, L., and Souza, A. M. P. D. (2002). *Map and Database of Quaternary Faults and Lineaments in Brazil. U.S. Geological Survey. Open-File Report* 02–230. Available at http://pubs.usgs.gov/of/2002/ofr-02-230

Sorenson, M. D. (2003). *Avian mtDNA Primers*. Available at: http://people.bu.edu/msoren/Bird.mt.Primers.pdf

Speranza, P. R., Seijo, G., Grela, I. A., and Neffa, V. G. S. (2007). Chloroplast DNA variation in the *Turnera sidoides* L. complex (Turneraceae): biogeographical implications. *J. Biogeogr.* 34, 427–436. doi: 10.1111/j.1365-2699.2006.01622.x

Stephens, M., Smith, N. J., and Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68, 978–989. doi: 10.1086/319501

Swartz, E. R., Skelton, P. H., and Bloomer, P. (2007). Sea-level changes, river capture and the evolution of populations of the Eastern Cape and fiery redfins (*Pseudobarbus afer* and *Pseudobarbus phlegethon*, Cyprinidae) across multiple river systems in South Africa. *J. Biogeogr.* 34, 2086–2099. doi: 10.1111/j.1365-2699.2007.01768.x

Swartz, E. R., Skelton, P. H., and Bloomer, P. (2009). Phylogeny and biogeography of the genus *Pseudobarbus* (Cyprinidae): shedding light on the drainage history of rivers associated with the Cape Floristic Region. *Mol. Phylogenet. Evol.* 51, 75–84. doi: 10.1016/j.ympev.2008.10.017

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105, 437–460.

Thomaz, A. T., Malabarba, L. R., Bonatto, S. L., and Knowles, L. L. (2015). 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. *J. Biogeogr.* 42, 2389–2401. doi: 10.1111/jbi.12597

Tonni, E. P., and Cione, A. L. (1997). Did the Argentine Pampean ecosystem exist in the Pleistocene. *Curr. Res. Pleistocene* 14, 131–133.

Trenti, P. S., Gómez, S. E., and Ferriz, R. A. (1999). Capacidad de natación en tres peces pampásicos. *Aprona* 38, 2–9.

Turchetto, C., Fagundes, N. J. R., Segatto, A. L. A., Kuhlemeier, C., Neffa, V. G. S., Speranza, P. R., et al. (2014). Diversification in the South American Pampas: the genetic and morphological variation of the widespread *Petunia axillaris* complex (Solanaceae). *Mol. Ecol.* 23, 374–389. doi: 10.1111/mec.12632

Turchetto-Zolet, A. C., Pinheiro, F., Salgueiro, F., and Palma-Silva, C. (2013). Phylogeographical patterns shed light on evolutionary process in South America. *Mol. Ecol.* 22, 1193–1213. doi: 10.1111/mec.12164

Villwock, J. A., and Tomazelli, L. J. (1995). Geologia costeira do Rio Grande do Sul. *Notas Técnicas* 8, 1–45.

Ward, R. D., Woodwark, M., and Skibinski, D. O. F. (1994). A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *J. Fish. Biol.* 44, 213–232. doi: 10.1111/j.1095-8649.1994.tb01200.x

Zemlak, T. S., Habit, E. M., Walde, S. J., Battini, M. A., Adams, E. D., and Ruzzante, D. E. (2008). Across the southern Andes on fin: glacial refugia, drainage reversals and a secondary contact zone revealed by the phylogeographical signal of *Galaxias platei* in Patagonia. *Mol. Ecol.* 17, 5049–5061. doi: 10.1111/j.1365-294X.2008.03987.x

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

Copyright © 2017 Ramos-Fregonezi, Malabarba and Fagundes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 17 January 2018
doi: 10.3389/fgene.2018.00001

Genetic Pattern and Demographic History of *Salminus brasiliensis*: Population Expansion in the Pantanal Region during the Pleistocene

Lívia A. de Carvalho Mondin[1], Carolina B. Machado[2], Emiko K. de Resende[3], Debora K. S. Marques[3] and Pedro M. Galetti Jr.[2]*

[1] Departamento de Ciências Biológicas, Universidade do Estado de Mato Grosso, Tangará da Serra, Brazil, [2] Laboratório de Biodiversidade Molecular e Conservação, Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, Brazil, [3] Embrapa Pantanal, Empresa Brasileira de Pesquisa Agropecuária, Corumbá, Brazil

Pleistocene climate changes were major historical events that impacted South American biodiversity. Although the effects of such changes are well-documented for several biomes, it is poorly known how these climate shifts affected the biodiversity of the Pantanal floodplain. Fish are one of the most diverse groups in the Pantanal floodplains and can be taken as a suitable biological model for reconstructing paleoenvironmental scenarios. To identify the effects of Pleistocene climate changes on Pantanal's ichthyofauna, we used genetic data from multiple populations of a top-predator long-distance migratory fish, *Salminus brasiliensis*. We specifically investigated whether Pleistocene climate changes affected the demography of this species. If this was the case, we expected to find changes in population size over time. Thus, we assessed the genetic diversity of *S. brasiliensis* to trace the demographic history of nine populations from the Upper Paraguay basin, which includes the Pantanal floodplain, that form a single genetic group, employing approximate Bayesian computation (ABC) to test five scenarios: constant population, old expansion, old decline, old bottleneck following by recent expansion, and old expansion following by recent decline. Based on two mitochondrial DNA markers, our inferences from ABC analysis, the results of Bayesian skyline plot, the implications of star-like networks, and the patterns of genetic diversity (high haplotype diversity and low-to-moderate nucleotide diversity) indicated a sudden population expansion. ABC allowed us to make strong quantitative inferences about the demographic history of *S. brasiliensis*. We estimated a small ancestral population size that underwent a drastic fivefold expansion, probably associated with the colonization of newly formed habitats. The estimated time of this expansion was consistent with a humid and warm phase as inferred by speleothem growth phases and travertine records during Pleistocene interglacial periods. The strong concordance between our genetic inferences and this historical data could represent the first genetic record of a humid and warm phase in the Pantanal in the period since the Last Interglacial to 40 ka.

Keywords: approximate Bayesian computation, mitochondrial markers, interglacial period, neotropical fish, Upper Paraguay basin

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

Reviewed by:
Rubens Pazza, Federal University of Viçosa, Brazil
Silvia Helena Sofia, Universidade Estadual de Londrina, Brazil

****Correspondence:***
Carolina B. Machado
carolbioms@gmail.com

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

Received: *01 September 2017*
Accepted: *03 January 2018*
Published: *17 January 2018*

Citation:
Mondin LAC, Machado CB, Resende EK, Marques DKS and Galetti PM Jr. (2018) Genetic Pattern and Demographic History of Salminus brasiliensis: Population Expansion in the Pantanal Region during the Pleistocene. Front. Genet. 9:1. doi: 10.3389/fgene.2018.00001

INTRODUCTION

The genetic diversity and population structure of species are largely determined by intrinsic traits, contemporary factors that include anthropogenic activities and historical events that have affected species over geological time (Avise, 2000; Osborne et al., 2014). Phylogeographic studies have shown that the genetic diversity patterns of several populations were shaped by past geological and climatic events (Hewitt, 1996; Bonaccorso et al., 2006; Carnaval et al., 2009; Thomaz et al., 2015; Cabanne et al., 2016).

The Pleistocene (2.58–0.01 million years ago) was marked by worldwide climatic changes involving multiple successive cycles of glacial and interglacial events, associated with abrupt and dramatic changes in temperature and precipitation (Webb and Bartlein, 1992). Overall, the environmental conditions varied from warm and wet during the interglacial to cold and dry during the glacial periods (EPICA Community Members, 2004).

Although there are no records of large ice sheets covering the landscape, as happened in the Northern Hemisphere, climatic changes during the Pleistocene also drastically affected Southern Hemispheric biomes (Vuilleumier, 1971; Hewitt, 1996). Several studies have reported contractions and range expansions in the Amazon and the Atlantic Forest (Harris and Mix, 1999; Carnaval et al., 2009; Ribas et al., 2012; Cabanne et al., 2016; Ledo and Colli, 2017), replacement by savannas and arid lands (Wang et al., 2004; Carnaval and Bates, 2007; Thomé et al., 2016), climatic disturbances in humid and arid areas in the Andes (Bräuning, 2009), and fluctuations in sea levels affecting the Atlantic coast (Thomaz et al., 2015; Ponce and Rabassa, 2016). The magnitude of climatic oscillations varied in different parts of South America. However, all of these events impacted the biodiversity, driving extinction episodes, speciation, intraspecific divergence, and demographic oscillations (Vuilleumier, 1971; Hewitt, 1996; Beheregaray, 2008). However, it remains poorly understood how these historical events affected the Pantanal biome and its biodiversity.

The Pantanal is a large seasonal floodplain located in the center of South America. This region is entirely contained in the Upper Paraguay basin, which is a complex drainage network due to its geological history (Assine, 2015). Pantanal was formed recently and has mainly been associated with the Andean uplift (Assine, 2015). However, geomorphological (Assine and Soares, 2004) and palynological (Ferraz-Vicentini and Salgado-Labouriau, 1996) data have shown that the paleoclimatic fluctuations have promoted considerable landscape changes since the Late Pleistocene. Although there are historical records of environmental changes, little is known of the effects of these changes on biodiversity in the Pantanal (Márquez et al., 2006; Lopes I.F. et al., 2007; Santos et al., 2008).

Fish are one of the most diverse groups in the Pantanal floodplains (Alho, 2008). Because their distribution is restricted to freshwater drainages, which in turn reflect geomorphological or climatic changes, these animals are a suitable biological model for reconstructing paleoenvironmental scenarios (Montoya-Burgos, 2003). To try to identify any effects of Pleistocene climate changes in fish, we used genetic data from multiple sites of *Salminus brasiliensis*. As a top predator distributed throughout the La Plata basin (Lima and Britski, 2007), this species represents an important resource in the Pantanal floodplains. Although it plays an important ecological role in controlling the structure of the ecosystem, the genetic diversity of *S. brasiliensis* in the Upper Paraguay basin is unknown. Assessing the population genetic pattern of the existing species can identify how its genetic variation was affected by past climate changes and can also provide insights into its demographic response (Pil et al., 2017).

Here, we focused on assessing the genetic diversity and pattern of *S. brasiliensis* sampled in nine rivers to trace its demographic history in the Upper Paraguay basin using a statistical framework to test the following five scenarios: constant population (null hypothesis), old expansion, old decline, old bottleneck followed by recent expansion, and old expansion followed by recent decline. We specifically investigated whether Pleistocene climate changes affected past *S. brasiliensis* demography. If so, we expected to find a population reduction signature during glacial periods due to climatic conditions that reduced the habitat, when the Pantanal region was almost dry. Further, a population expansion signature was expected in the interglacial periods, when the wetlands were re-established, increasing fluvial discharge and suitable habitats for the biological model. After reconstructing the most probable demographic history for *S. brasiliensis* based on two mitochondrial markers (CytB and Dloop), we associated changes in population size over time with known regional climatic fluctuations. Our findings might contribute to the conservation and management of natural populations of *S. brasiliensis* from the Upper Paraguay basin by being used to predict whether these populations will be able to persist in future scenarios of environmental change.

MATERIALS AND METHODS

Ethics Statement, Sampling Collection, DNA Extraction, and Amplification

This study was developed in accordance with Brazilian law, approved by the ethic committee on animal use at Federal University of São Carlos (CEUA number 5765010416), and carried out under a temporary scientific collection license number 32217-4 provided by ICMBio-SISBIO. A total of 52 *S. brasiliensis* individuals were sampled in nine rivers from the Upper Paraguay basin (**Figure 1** and Supplementary Table S1). We sampled a small fragment of the caudal fin and then returned the fish to the river. All tissue samples were preserved in absolute ethanol and stored at −20°C. Total genomic DNA was extracted using the saline precipitation protocol (Aljanabi and Martinez, 1997).

We amplified partial sequences of two mitochondrial DNA markers: Cytochrome B (CytB) and Dloop (see Supplementary Table S2 for primers and references). Polymerase Chain Reactions for both markers (PCR) occurred in 25 μL of final volume containing 16.9 μL of mili-Q water, 2.5 μL of 10x Invitrogen PCR buffer (1x), 2.5 μL of dNTPs mix (0.25 mM), 1 μL of $MgCl_2$ (2 mM), 0.5 μL of each primer (0.2 mM), 0.1 μL of Platinum®Taq Polymerase (0.5 unit) and 1 μL of DNA

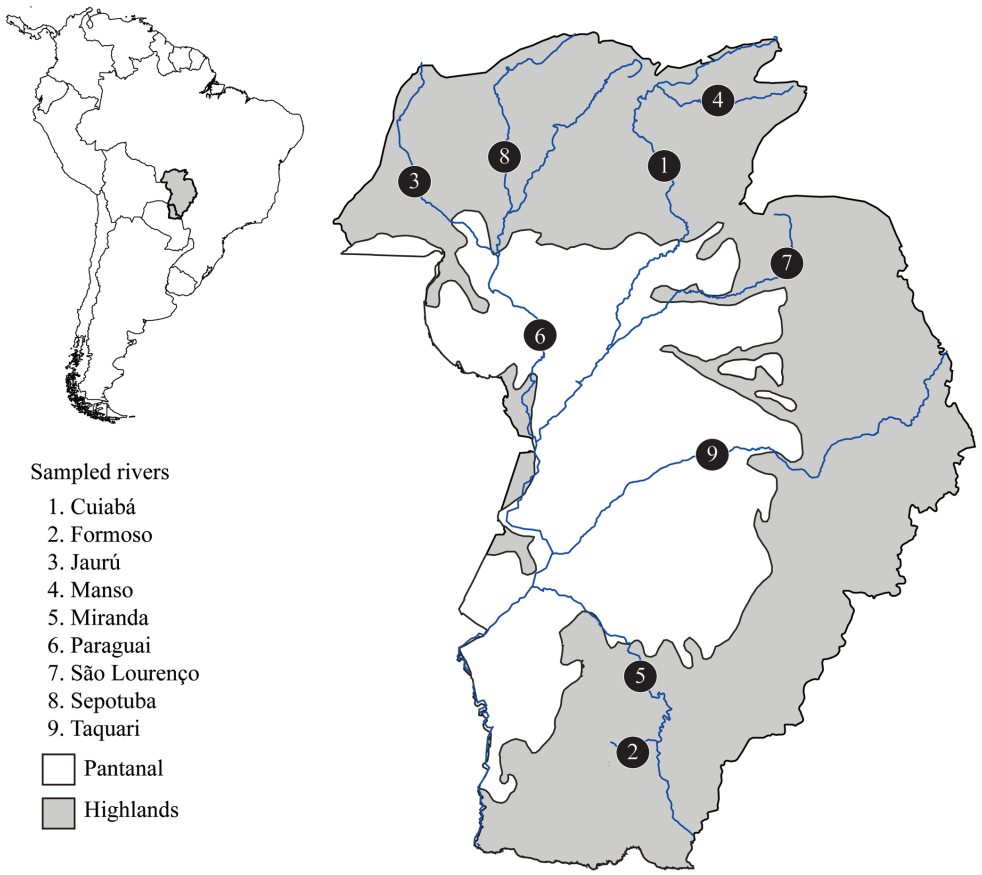


FIGURE 1 | Map of the Upper Paraguay basin in Brazilian territory indicating rivers where *Salminus brasiliensis* were sampled. Populations are labeled with numbers. The Pantanal area is in white and represents the lowlands (below 60 m above sea level).

template (50 ng/μL). The annealing temperature of PCR thermal profile was described in Supplementary Table S2. All PCR products were purified using the 20% Polyethylene glycol protocol (Lis and Schleif, 1975). Fragments were sequenced using ABI3730XL automatic sequencer. All the sequences were aligned automatically using the multiple alignment ClustalW and minor manual adjustments were made to improve it in BioEdit (Hall, 1999). The GenBank accession numbers are given in Supplementary Table S1.

Genetic Diversity and Population Structure

The genetic diversity indices for each population, including number of haplotypes, haplotype diversity, and nucleotide diversity, were calculated in DnaSP 5.0 (Librado and Rozas, 2009). We also reconstructed the haplotype networks for each molecular marker using the median-joining method (Bandelt et al., 1999) included in PopART (Population Analysis with Reticulate Trees – Leigh and Bryant, 2015).

We investigated whether there were distinct mitochondrial genetic clusters by assigning individuals to populations using a Bayesian model-based approach implemented in Bayesian Analysis of Population Structure (BAPS) v. 6.0 (Corander et al., 2008). BAPS identifies clusters based on nucleotide frequencies in DNA sequences. We concatenated mitochondrial markers and conducted a mixing analysis without spatialization (option chosen "clustering of group of individuals"). This method determines which combination of predetermined samples is best supported by the data. The program was run for 10 replicates for each K (1–10). The best clustering partition was determined by the highest value of likelihood and highest probability determined by the program.

Inference and Hypotheses Tests of Demographic Histories

Because no genetic population structuring was found in *S. brasiliensis* using BAPS (Supplementary Figure S1 and Supplementary Table S3), we drawn all scenarios based on historical demographic changes in a single population. Demographic history was inferred through a Bayesian skyline plot (BSP) analysis (Drummond et al., 2005) implemented in BEAST 2.4.4 (Bouckaert et al., 2014). BSP assumes a single panmictic population and uses inferred patterns of coalescence to fit a demographic model to a set of sequence data (Drummond et al., 2005).

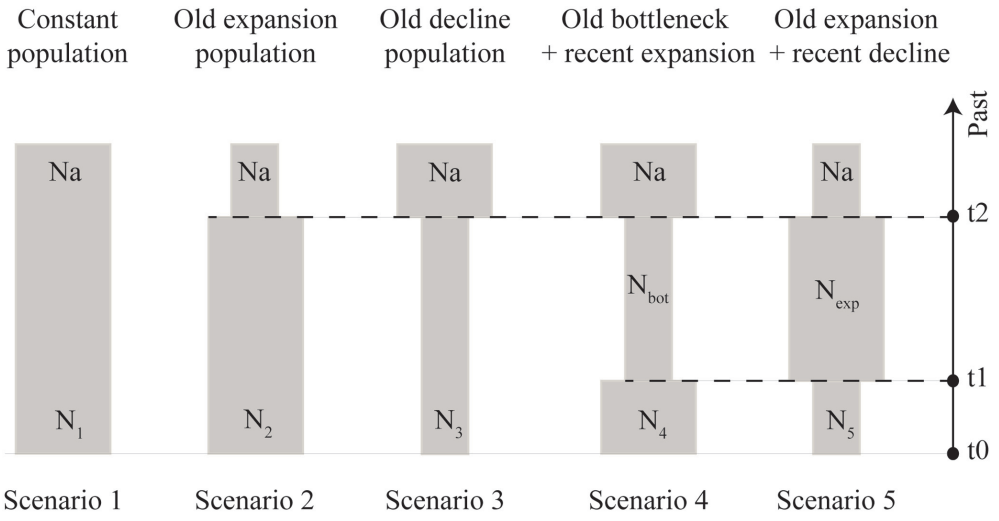


FIGURE 2 | Schematic representation of five demographic scenarios, including model parameters of *S. brasiliensis*, tested by approximate Bayesian computation (ABC). Time and effective population size are not strictly to scale.

First, we determined the best evolutionary model for each molecular marker in JModelTest 2.1.4 (Posada, 2008) using the Bayesian information criterion. The CytB and Dloop markers were set to evolve according to a HKY and HKY+I models, respectively. In BEAST 2.4.4, we chose as the tree prior the coalescent Bayesian skyline with both linked mitochondrial DNA markers to perform the BSP analysis. Priors used in this analysis were kept in default. To date the past demographic events, a molecular clock was calibrated using the mutation rate defined for the *Salminus* genus (Machado et al., unpublished), using the relaxed lognormal model: 6.31×10^{-3} (CytB) and 1.935×10^{-2} (Dloop) mutations per site per million years. This rate was based on a previous topology calibrated from the genus rooted with *Brycon* specimens (Machado et al., unpublished). This topology was reconstructed using two mitochondrial and two nuclear markers and calibrated using a biogeographic event, the uplift of the Eastern Andes Cordillera (Hoorn et al., 1995), which separated the Magdalena and Amazon paleobasins and is consistent with the divergence of *S. affinis* from all other *Salminus*.

Three independent Markov chains were initiated from random trees, run separately for 10 million generations, and sampled at intervals of 10000 steps, 25% of which were discarded as burn-in. The log and tree files for each independent Markov chain were combined using LogCombiner in BEAST 2.4.4. Tracer v1.5 (Rambaut et al., 2014) was used to check the convergence of the runs according to the effective sampling size (>200) and reconstruct the population dynamic over time through the Bayesian Skyline reconstruction option.

To corroborate the findings on the population expansion of *S. brasiliensis* in the Upper Paraguay basin suggested by the BSP analysis, we tested five demographic scenarios that could have happened during the Late Pleistocene (expansions and bottlenecks) using an approximate Bayesian computation (ABC) framework in DIYABC v2.1.0 (Cornuet et al., 2014): constant population (null hypothesis), old expansion, old decline, old bottleneck followed by expansion, and old expansion followed by decline (**Figure 2**). ABC approach allows a quantitative evaluation of the demographic and evolutionary history by strictly contrasting realistic models defined *a priori* and estimating relevant parameters (Beaumont et al., 2002).

All scenarios assumed an initial ancestral population (Na), demographic changes occurring during the Late Pleistocene (2–120 ka for t1, and 4–120 ka for t2), and recent populations represented in t0 (**Figure 2**). The prior distribution of demographic parameters and the mutation rate are listed in Supplementary Table S4. For each tested scenario, we calculated the following summary statistics: the number of haplotypes, number of segregating sites, variance of pairwise differences, and variance of numbers of the rarest nucleotides at segregating sites. The summary statistics were used for comparison between simulated and observed datasets.

We ran one million simulations for each scenario and used summary statistics and principal component analysis to confirm the good fit of all scenarios with the observed data (Supplementary Figure S2). Then we compared the competing scenarios by estimating their posterior probabilities using polychotomous logistic regression on 1% of simulated datasets closest to the observed data (Cornuet et al., 2008). The best scenario was the one that showed the highest posterior probability with a 95% confidence interval (CI) non-overlapping with other scenarios' posterior probabilities. In order to evaluate confidence in each scenario, we also calculated the posterior predictive error. Subsequently, we estimated the posterior probabilities of each parameter under the best scenario using a local linear regression on 1% closest simulated data sets and applying a logit transformation to parameter values as suggested by Cornuet et al. (2008). We assessed their precision of parameter estimation through the relative median of the absolute error (RMAE) based on 500 pseudo-observed datasets (pods) under

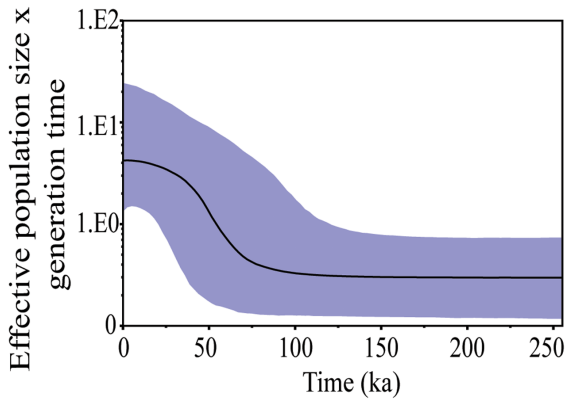


FIGURE 3 | Bayesian skyline plot (BSP) showing change in effective population size of *S. brasiliensis* from the Upper Paraguay basin over time based on CytB and Dloop markers. The y-axis, population size × generation time*; x-axis, time (indicated in thousand of years ago). *Generation time measured in million years.

the best scenario. Low RMAE values indicated that the parameter estimated is trustworthy (Cornuet et al., 2008). After these steps, we performed a model verification by evaluating the goodness-of-fit of the best scenario with respect to empirical dataset. This was carried out by simulating 1000 pods under the best scenario. In this step, we used different summary statistics to avoid over-estimating the fit by using the same statistics twice (Cornuet et al., 2010). We selected mean of pairwise differences, Tajima's D, private segregating sites and mean of number of the rarest nucleotide at segregating sites.

RESULTS

Fragments of 710 base pairs (bp) of CytB and 571 bp of Dloop were obtained from 52 individuals of *S. brasiliensis* from the Upper Paraguay basin. CytB had 20 polymorphic sites, of which 9 were parsimoniously informative, while Dloop had 37 polymorphic sites, of which 26 were parsimoniously informative. No insertions, deletions, or stop codons (for CytB sequences) were observed for either marker.

We identified 21 and 33 haplotypes for CytB and Dloop, respectively (Supplementary Table S5 and Supplementary Figure S1). No common haplotype was shared among all populations in any molecular markers, but a high number of private haplotypes (61.9% for CytB and 66.7% for Dloop) was observed. The distribution of the haplotypes in the network indicated no population structuring in the rivers sampled (Supplementary Figure S1). The haplotype and nucleotide diversity values for each population are shown in Supplementary Table S5. Overall, despite the nucleotide diversity exhibited was low (0.00365 for CytB) to moderate (0.01434 for Dloop), the level of haplotype diversity was high (0.928 for CytB, and 0.975 for Dloop).

The BAPS approach clustered all individuals into only one mitochondrial genetic group (Supplementary Figure S1) and the BSP analysis gave the effective population size, which underwent a process of expansion that started approximately 75 ka (**Figure 3**). The greatest expansion was underway through around 40 ka and continued gradually until 20 ka, when the effective population size became stable.

Approximate Bayesian computation analysis also corroborated the old expansion scenario as the most probable one to explain the current genetic diversity of *S. brasiliensis* from the Upper Paraguay basin (Supplementary Figure S3). The best scenario showed a posterior probability (PP) of 0.5466 (95% CI: 0.5341–0.5590), while the second PP was observed for scenario 4 (old bottleneck followed by expansion), with 0.2589 (95% CI: 0.2466–0.2712) (Supplementary Table S6). The posterior error rate was 0.303. Features such as the non-overlapping 95% CI among the scenarios' PPs, and the signal expansion found by BSP analysis and genetic pattern (see section "Discussion") suggest a greater confidence that scenario 2 (old population expansion) is the best of the proposed scenarios.

The best scenario recovered the ancestral population of *S. brasiliensis* (Na = 161000) expanding at 31000 generations ago (i.e., about 62 ka). At present, the median effective population size is estimated at N_2 = 811000 (**Table 1**). Large confidence intervals are common in such demographic inferences. All our RMAE values are <0.2, except by current effective population size (RMAE = 0.299), indicating the parameters estimated by ABC analysis were reliable values. This was expected because mitochondrial is reliable to detect past demographic parameter estimates (Na and t2) (Cornuet et al., 2010). When we applied the model checking option to the best scenario, we observed on each PCA plane a wide cloud of data sets simulated from the prior, with the observed dataset in the middle of a small cluster of datasets from the posterior predictive distribution (Supplementary Figure S4). This indicate that the scenario 2 fits well the empirical data.

DISCUSSION

High haplotype of genetic diversity and no spatial genetic structuring among populations of *S. brasiliensis* were here observed. The results also revealed that its current genetic diversity was strongly influenced by an expansion event that happened during the Pleistocene, likely associated with warm and humid weather conditions during interglacial periods. All of these findings are discussed in detail below.

TABLE 1 | Parameter estimates generated from a local linear regression on simulated datasets generated from the best scenario.

Parameter	Median	95% Confidence interval	RMAE
N_2 (Present population)	811000	495000–982000	0.299
Na (Ancestral population)	161000	41500–539000	0.194
t2 (Expansion time)	62 (ka)	9.860–116.6 (ka)	0.093

The precision of parameter estimations was assessed by computing the relative median of absolute error (RMAE). ka, 1000 years before present.

Genetic Diversity and Structure of *S. brasiliensis*

Previous studies in different hydrographic systems and using other molecular markers, such as RAPD and microsatellites, have determined a great genetic variability for *S. brasiliensis* (Ramella et al., 2006; Lopes C.M. et al., 2007; Gomes et al., 2013; Ribeiro et al., 2016; Ribolli et al., 2017). High levels of haplotype diversity seem to be common in migratory freshwater fishes with large populations that occupy heterogeneous environments (see Oliveira et al., 2009). A large effective population size and high migration rates may minimize the effects of genetic drift as a force that decreases intraspecific genetic diversity (Allendorf and Luikart, 2007). Furthermore, the pattern of genetic diversity of *S. brasiliensis* also can be attributed to its demographic history. According to Grant and Bowen (1998), high haplotype diversity and low nucleotide diversity is associated with an expansion event after a period of small effective population size. Thus, this process could have introduced and maintained new mutations in the population of *S. brasiliensis*, increasing its genetic variability.

No mtDNA genetic structuring was detected in *S. brasiliensis*. Its higher migratory capacity (it can move 1000 km during the reproductive season; Petrere, 1985) probably played an important role in causing the lack of spatial genetic differentiation between populations. It is expected that long-distance migratory fishes within a hydrographic basin such as the Upper Paraguay, without physical, abiotic, and/or biotic barriers, will not show a spatial genetic structuring (Pil et al., 2017). However, it seems that *S. brasiliensis* has undergone a different process of population differentiation. Ribolli et al. (2017) identified distinct temporal genetic populations of *S. brasiliensis* from the Upper Uruguay basin during the reproductive season. According to those authors, individuals or populations of *S. brasiliensis* that occupy the same geographic location might be adapted to spawning at different times during the same reproductive season (Ribolli et al., 2017). Because our sample lacks more precise temporal information for some sampled individuals, no time-related genetic population structuring can be inferred here.

Revisiting the Pleistocene in the Pantanal Region

The demographic history of *S. brasiliensis* populations in the Upper Paraguay basin was addressed for the first time in this study, and it was found that a demographic expansion event probably occurred in its early evolutionary history. This took place in the late Pleistocene, before the Last Glacial Maximum (LGM, 23–19 ka). Our inference from ABC analysis, the results of BSP, the implications of star-like networks, and the patterns of genetic diversity (high haplotype diversity and low to moderate nucleotide diversity) indicated a sudden population expansion.

The ABC approach was an efficient tool for reconstructing and testing alternative demographic scenarios; this allowed us to determine the most plausible evolutionary history of *S. brasiliensis*, beyond making strong quantitative inferences about its demographic history. We estimated a small ancestral population size that underwent a drastic fivefold expansion, probably associated with the colonization of newly formed habitats. The time estimated for this expansion was consistent with the humid and warm phase inferred from speleothem growth phases and travertine records during Pleistocene interglacial periods (Wang et al., 2004).

Pleistocene climatic oscillations in South America drastically altered freshwater systems by changing hydrological flow and lake distributions as well as rerouting rivers (Stevaux, 2000; Cross et al., 2001; Fritz et al., 2004). As an alluvial plain dominated by rivers, the Pantanal region reveals geomorphological relicts overprinted on its modern landscape that indicate the occurrence of intense historical hydrological rearrangement (Assine and Soares, 2004). Paleochannels, changes in the levels of marginal lakes, and discontinuous sedimentation caused by intermittent flows are examples of relict landforms, which reflect environmental and climatic conditions different from those currently observed (Assine, 2015). Radiocarbon dating and palynological data support the conclusion that during colder and arid events in the Late Pleistocene/Holocene, the Pantanal experienced desert-like conditions with sparse dry vegetation, intermittent flow along the alluvial fan, and the development of paleochannels and erosional surfaces. Further, humid and warm conditions would have influenced the development of lakes and river systems and consequently expanded freshwater surface area (Ferraz-Vicentini and Salgado-Labouriau, 1996; Assine and Soares, 2004; Assine, 2015).

The pulses of the reduction and expansion of the landscape left a phylogenetic signature in regional fauna, leading to a population expansion in bird (Lopes I.F. et al., 2007; Santos et al., 2008) and mammal species (Márquez et al., 2006) stemming from the humid cycles of the Pleistocene. The signature population expansion in *S. brasiliensis* was dated to 62 ka. Although there is a lack of regional paleoenvironmental data extending further than 40 ka, speleothem data from the Brazilian Northeast region reveal a pluvial maxima phase at about 61.2–59.1 ka (Wang et al., 2004), reinforcing our findings from *S. brasiliensis*. The strong concordance between our genetic inferences and this historical data could represent the first genetic record of a humid and warm phase in the Pantanal in the period since the Last Interglacial to 40 ka.

CONCLUSION

In summary, we focused to identify the effect of Pleistocene climatic changes on demographic history of *S. brasiliensis* from the Upper Paraguay basin. Our findings suggest that Pleistocene climate fluctuations fundamentally shaped the genetic diversity and pattern of this species in this region. Coalescent-based analyses, particularly the statistical framework provided by the ABC method, supported an ancient expansion event before the LGM as the best scenario to explain the current genetic diversity of *S. brasiliensis* based on two mitochondrial markers.

Understanding how historical events impact genetic diversity is important for predicting whether populations will persist under future environmental changes (Pauls et al., 2013). High genetic variation in populations is desirable, because it provides the

basis for evolutionary change in a species, thereby improving the chances of survival in a dynamic environment (Frankham, 1995). Although populations of *S. brasiliensis* from the study area show high levels of genetic diversity, our concern over its conservation status cannot wane, because its genetic pool is unique and completely different from the *S. brasiliensis* found in the Upper Paraná basin (Machado et al., 2017). We suggest that conservation management of this species should concentrate on maximizing the retention of diversity, and preventing severe population decline due to anthropogenic influence.

AUTHOR CONTRIBUTIONS

LCM, CM, ER, DM, and PG conceived the ideas and designed the experiments. LCM, ER, and DM collected the samples. CM and PG performed the bioinformatics analyses. LCM, CM, and PG led the writing, with assistance from ER and DM. All authors read and approved the final manuscript.

FUNDING

This study was funded by the Brazilian agencies Fundação de Amparo à Pesquisa do Estado de Mato Grosso (FAPEMAT), Fundação CAPES, and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

ACKNOWLEDGMENTS

The authors thank EMBRAPA Pantanal for the use of some tissue samples.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00001/full#supplementary-material

REFERENCES

Alho, C. J. R. (2008). Biodiversity of the Pantanal: response to seasonal flooding regime and to environmental degradation. *Braz. J. Biol.* 68, 957–966. doi: 10.1590/S1519-69842008000500005

Aljanabi, S. M., and Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* 25, 4692–4693. doi: 10.1093/nar/25.22.4692

Allendorf, F. W., and Luikart, G. H. (2007). "Population subdivision," in *Conservation and the Genetics of Populations*, (Hoboken, NJ: John Wiley & Sons), 197–232.

Assine, M. L. (2015). "Brazilian Pantanal: a large pristine tropical wetland," in *Landscapes and Landforms of Brazil*, eds B. Vieira, A. Salgado, and L. Santos (Dordrecht: Springer), 135–146. doi: 10.1007/978-94-017-8023-0_12

Assine, M. L., and Soares, P. C. (2004). Quaternary of the Pantanal, west-central Brazil. *Quat. Int.* 114, 23–34. doi: 10.1016/S1040-6182(03)00039-9

Avise, J. C. (2000). *Phylogeography: The History and Formation of Species*. Cambridge, MA: Harvard University Press.

Bandelt, H.-J., Forster, P., and Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48. doi: 10.1093/oxfordjournals.molbev.a026036

Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate Bayesian Computation in population genetics. *Genetics* 162, 2025–2035.

Beheregaray, L. B. (2008). Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Mol. Ecol.* 17, 3754–3774. doi: 10.1111/j.1365-294X.2008.03857.x

Bonaccorso, E., Koch, I., and Peterson, A. T. (2006). Pleistocene fragmentation of Amazon species' ranges. *Divers. Distrib.* 12, 157–164. doi: 10.1111/j.1366-9516.2005.00212.x

Bouckaert, R. R., Heled, J., Kuehnert, D., Vaughan, T. G., Wu, C.-H., Xie, D., et al. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLOS Comput. Biol.* 10:e1003537. doi: 10.1371/journal.pcbi.1003537

Bräuning, A. (2009). Climate variability of the tropical Andes since the late Pleistocene. *Adv. Geosci.* 22, 13–25. doi: 10.5194/adgeo-22-13-2009

Cabanne, G. S., Calderón, L., Trujillo Arias, N., Flores, P., Pessoa, R., d'Horta, F. M., et al. (2016). Effects of Pleistocene climate changes on species ranges and evolutionary processes in the Neotropical Atlantic Forest. *Biol. J. Linn. Soc.* 119, 856–872. doi: 10.1111/bij.12844

Carnaval, A. C., and Bates, J. M. (2007). Amphibian DNA shows marked genetic structure and tracks Pleistocene climate change in Northeastern Brazil. *Evolution* 61, 2942–2957. doi: 10.1111/j.1558-5646.2007.00241.x

Carnaval, A. C. O. Q., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T., and Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science* 323, 785–789. doi: 10.1126/science.1166955

Corander, J., Marttinen, P., Sirén, J., and Tang, J. (2008). Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9:539. doi: 10.1186/1471-2105-9-539

Cornuet, J. M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., et al. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30, 1187–1189. doi: 10.1093/bioinformatics/btt763

Cornuet, J. M., Ravigné, V., and Estoup, A. (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* 11:401. doi: 10.1186/1471-2105-11-401

Cornuet, J. M., Santos, F., Beaumont, M. A., Robert, C. P., Marin, J. M., Balding, D. J., et al. (2008). Inferring population history with DIYABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24, 2713–2719. doi: 10.1093/bioinformatics/btn514

Cross, S. L., Baker, P. A., Seltzer, G. O., Fritz, S. C., and Dunbar, R. B. (2001). Late quaternary climate and hydrology of tropical South America inferred from an isotopic and chemical model of Lake Titicaca, Bolivia, and Peru. *Quat. Res.* 56, 1–9. doi: 10.1006/qres.2001.2244

Drummond, A. J., Rambaut, A., Shapiro, B., and Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22, 1185–1192. doi: 10.1093/molbev/msi103

EPICA Community Members (2004). Eight glacial cycles from an Antarctic ice core. *Nature* 429, 623–662. doi: 10.1038/nature02599

Ferraz-Vicentini, K. R., and Salgado-Labouriau, M. L. (1996). Palynological analysis of a palm swamp in central Brazil. *J. South Am. Earth Sci.* 9, 207–219. doi: 10.1016/0895-9811(96)00007-7

Frankham, R. (1995). Conservation genetics. *Annu. Rev. Genet.* 29, 305–327. doi: 10.1146/annurev.ge.29.120195.001513

Fritz, S. C., Baker, P. A., Lowenstein, T. K., Seltzer, G. O., Rigsby, C. A., Dwyer, G. S., et al. (2004). Hydrologic variation during the last 170,000 years in the southern hemisphere tropics of South America. *Quat. Res.* 61, 95–104. doi: 10.1016/j.yqres.2003.08.007

Gomes, P. C., Lopera-Barrero, N. M., Vargas, L., Streit Junior, D. P., Povh, J. A., Nardez Sirol, R., et al. (2013). Genetic diversity of *Salminus brasiliensis* (Characiformes: Characidae) collected in the passage ladder of the Canoas I hydropower plant in the Paranapanema River, Brazil. *Semin. Cienc. Agrar.* 34, 1421–1432. doi: 10.5433/1679-0359.2013v34n3p1421-1432

Grant, W. A. S., and Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *J. Hered.* 89, 415–426. doi: 10.1093/jhered/89.5.415

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Sympo. Ser.* 41, 95–98.

Harris, S. E., and Mix, A. C. (1999). Pleistocene precipitation balance in the Amazon Basin recorded in deep sea sediments. *Quat. Res.* 51, 14–26. doi: 10.1006/qres.1998.2008

Hewitt, G. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* 58, 247–276. doi: 10.1111/j.1095-8312.1996.tb01434.x

Hoorn, C., Guerrero, J., Sarmiento, G. A., and Lorente, M. A. (1995). Andean tectonic as a cause for changing drainage patterns in Miocene northern South America. *Geology* 23, 237–240. doi: 10.1130/0091-76131995023<0237:ATAACF<2.3.CO;2

Ledo, R. M. D., and Colli, G. R. (2017). The historical connections between the Amazon and the Atlantic Forest revisited. *J. Biogeogr.* 44, 2551–2563. doi: 10.1111/jbi.13049

Leigh, J. W., and Bryant, D. (2015). PopArt: full-feature software for haplotype network construction. *Methods Ecol. Evol.* 6, 1110–1116. doi: 10.1111/2041-210X.12410

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187

Lima, F. C. T., and Britski, H. A. (2007). *Salminus franciscanus*, a new species from the rio São Francisco basin, Brazil (Ostariophysi: Characiformes: Characidae). *Neotrop. Ichthyol.* 5, 237–244. doi: 10.1590/S1679-62252007000300001

Lis, J. T., and Schleif, R. (1975). Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucleic Acids Res.* 2, 383–389. doi: 10.1093/nar/2.3.383

Lopes, C. M., Almeida, F. S. D., Orsi, M. L., Britto, S. G. D. C., Sirol, R. N., and Sodré, L. M. K. (2007). Fish passage ladders from Canoas Complex–Paranapanema River: evaluation of genetic structure maintenance of *Salminus brasiliensis* (Teleostei: Characiformes). *Neotrop. Ichthyol.* 5, 131–138. doi: 10.1590/S1679-62252007000200006

Lopes, I. F., Miño, C. I., and Del Lama, S. N. (2007). Genetic diversity and evidence of recent demographic expansion in waterbird populations from the Brazilian Pantanal. *Braz. J. Biol.* 67, 849–857. doi: 10.1590/S1519-69842007000500007

Machado, C. B., Ishizuka, T. K., Freitas, P. D., Valiati, V. H., and Galetti, P. M. (2017). DNA barcoding reveals taxonomic uncertainty in *Salminus* (Characiformes). *Syst. Biodivers.* 15, 372–382. doi: 10.1080/14772000.2016.1254390

Márquez, A., Maldonado, J. E., González, S., Beccaceci, M. D., Garcia, J. E., and Duarte, J. M. B. (2006). Phylogeography and Pleistocene demographic history of the endangered marsh deer (*Blastocerus dichotomus*) from the Río de la Plata Basin. *Conserv. Genet.* 7:563. doi: 10.1007/s10592-005-9067-8

Montoya-Burgos, J. I. (2003). Historical biogeography of the catfish genus *Hypostomus* (Siluriformes: Loricariidae), with implications on the diversification of Neotropical ichthyofauna. *Mol. Ecol.* 12, 1855–1867. doi: 10.1046/j.1365-294X.2003.01857.x

Oliveira, C., Foresti, F., and Hilsdorf, A. W. S. (2009). Genetics of neotropical fish: from chromosomes to populations. *Fish Physiol. Biochem.* 35, 81–100. doi: 10.1007/s10695-008-9250-1

Osborne, M. J., Perkin, J. S., Gido, K. B., and Turner, T. F. (2014). Comparative riverscape genetics reveals reservoirs of genetic diversity for conservation and restoration of Great Plains fishes. *Mol. Ecol.* 23, 5663–5679. doi: 10.1111/mec.12970

Pauls, S. U., Nowak, C., Bálint, M., and Pfenninger, M. (2013). The impact of global climate change on genetic diversity within populations and species. *Mol. Ecol.* 22, 925–946. doi: 10.1111/mec.12152

Petrere, M. Jr. (1985). Migraciones de peces de agua dulce en America Latina: Algunos comentarios. *COPESCAL Doc. Ocasional* 1, 1–17.

Pil, M. W., Baggio, R. A., Tschá, M. K., Marteleto, F. M., Orélis-Ribeiro, R., Patella, L., et al. (2017). The influence of paleoclimate on the distribution of genetic variability and demography of fishes in a large and highly fragmented neotropical river. *Hydrobiologia* 805, 97–112. doi: 10.1007/s10750-017-3285-1

Ponce, J. F., and Rabassa, J. (2016). "Paleogeographic evolution of the Atlantic coast of South America during marine isotope stage 3 (MIS 3)," in *Marine Isotope Stage 3 in Southern South America, 60 ka BP–30 ka BP*, eds G. M. Gasparini, J. Rabassa, C. Deschamps, and E. P. Tonni (Cham: Springer), 155–165. doi: 10.1007/978-3-319-40000-6_9

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25, 1253–1256. doi: 10.1093/molbev/msn083

Rambaut, A., Suchard, M. A., Xie, D., and Drummond, A. J. (2014). *Tracer v1.6*. Available at: http://tree.bio.ed.ac.uk/software/tracer/

Ramella, M. S., Kroth, M. A., Meurer, S., Nuñer, A. P. D. O., Zaniboni-Filho, E., and Ariai, A. C. M. (2006). Genetic variability in four fish species (*Pimelodus maculatus, Prochilodus lineatus, Salminus brasiliensis* and *Steindachneridion scripta*) from Uruguay River Basin. *Braz. Arch. Biol. Technol.* 49, 589–598. doi: 10.1590/S1516-89132006000500008

Ribas, C. C., Aleixo, A., Nogueira, A. C. R., Miyaki, C. Y., and Cracraft, J. (2012). A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. *Proc. R. Soc. B Biol. Sci.* 279, 681–689. doi: 10.1098/rspb.2011.1120

Ribeiro, R., Lopera-Barrero, N. M., Povh, J. A., Rodriguez-Rodriguez, M. D. P., Fornari, D. C., Baumgartner, G., et al. (2016). Genetic diversity of Salminus brasiliensis wild populations in downstream and upstream Cachoeira Branca, Verde River MS Brazil: a preliminary view. *Semin. Cienc. Agrar.* 37, 507–516. doi: 10.5433/1679-0359.2016v37n1p507

Ribolli, J., Hoeinghaus, D. J., Johnson, J. A., Zaniboni-Filho, E., de Freitas, P. D., and Galetti, P. M. (2017). Isolation-by-time population structure in potamodromous *Dourado Salminus brasiliensis* in southern Brazil. *Conserv. Genet.* 18, 67–76. doi: 10.1007/s10592-016-0882-x

Santos, M. H., Lopes, I. F., and Del Lama, S. N. (2008). Population genetic study of mitochondrial DNA in Roseate spoonbill (Aves; *Platalea ajaja*) breeding colonies from the Pantanal wetlands, Brazil. *Biochem. Genet.* 46, 492–505. doi: 10.1007/s10528-008-9165-7

Stevaux, J. C. (2000). Climatic events during the late Pleistocene and Holocene in the upper Parana River: Correlation with NE Argentina and South-Central Brazil. *Quat. Int.* 72, 73–85. doi: 10.1016/S1040-6182(00)00023-9

Thomaz, A. T., Malabarba, L. R., Bonatto, S. L., and Knowles, L. L. (2015). Testing the effects of palaeodrainages versus habitat stability on genetic divergence in riverine systems: study of a neotropical fish of the Brazilian coastal Atlantic Forest. *J. Biogeogr.* 42, 2389–2401. doi: 10.1111/jbi.12597

Thomé, M. T. C., Sequeira, F., Brusquetti, F., Carstens, B., Haddad, C. F. B., Rodrigues, M. T., et al. (2016). Recurrent connections between Amazon and Atlantic forests shaped diversity in Caatinga four-eyed frogs. *J. Biogeogr.* 43, 1045–1056. doi: 10.1111/jbi.12685

Vuilleumier, B. S. (1971). Pleistocene changes in the fauna and flora of South America. *Science* 173, 771–779. doi: 10.1126/science.173.3999.771

Wang, X., Auler, A. S., Edwards, R. L., Cheng, H., Cristalli, P. S., Smart, P. L., et al. (2004). Wet periods in northeastern Brazil over the past 210 kyr linked to distant climate anomalies. *Nature* 432, 740–743. doi: 10.1038/nature03067

Webb, T., and Bartlein, P. J. (1992). Global changes during the last 3 million years: climatic controls and biotic response. *Annu. Rev. Eco. Evo. Syst.* 23, 141–173. doi: 10.1146/annurev.es.23.110192.001041

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 18 January 2018
doi: 10.3389/fgene.2018.00003

Microsatellites Associated with Growth Performance and Analysis of Resistance to *Aeromonas hydrophila* in Tambaqui *Colossoma macropomum*

*Raquel B. Ariede[1], Milena V. Freitas[1], Milene E. Hata[1], Vito A. Mastrochirico-Filho[1], Fabiana Pilarski[1], Sergio R. Batlouni[1], Fábio Porto-Foresti[2] and Diogo T. Hashimoto[1]**

[1] Aquaculture Center of Unesp, São Paulo State University (Unesp), Jaboticabal, Brazil, [2] School of Sciences, São Paulo State University (Unesp), Bauru, Brazil

Tambaqui, *Colossoma macropomum*, is the main native fish species produced in Brazil, and is an important species for genetic improvement in aquaculture. In addition, breeding studies on this species can be optimized with the use of molecular markers associated with productive phenotypes. The objective of the present study was to test the performance of growth traits and resistance to the bacteria, *Aeromonas hydrophila*, in association with microsatellite markers in *C. macropomum*. In this study, three full-sib families were subjected to bacterial challenge and morphometric growth assessments. Tambaqui families subjected to the bacterial challenge differed significantly in death time and mortality rate. There was, however, no association between resistance to bacteria and microsatellite markers. In relation to growth traits, we observed a marker/phenotype association in two microsatellites. The marker in the 6b isoform x5 gene (*TNCRC6b*) was associated with length, whereas an anonymous marker was associated with height. The present study highlighted the evaluation of molecular markers associated with growth traits, and can serve as the basis for future marker-assisted selection (MAS) of tambaqui.

Keywords: animal improvement, aquaculture, bacterial challenge, MAS, QTL

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Luis Antonio Inoue, Embrapa Western Agriculture, Brazil
Carolina Sousa Sá Leitão, National Institute of Amazonian Research, Brazil

***Correspondence:**
Diogo T. Hashimoto
diogo@caunesp.unesp.br

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

Received: *01 September 2017*
Accepted: *04 January 2018*
Published: *18 January 2018*

Citation:
Ariede RB, Freitas MV, Hata ME, Mastrochirico-Filho VA, Pilarski F, Batlouni SR, Porto-Foresti F and Hashimoto DT (2018) Microsatellites Associated with Growth Performance and Analysis of Resistance to Aeromonas hydrophila in Tambaqui Colossoma macropomum. Front. Genet. 9:3. doi: 10.3389/fgene.2018.00003

INTRODUCTION

Tambaqui (*Colossoma macropomum*) is a freshwater fish belonging to the family Serrasalmidae, with natural occurrence in the Amazon and Orinoco basins (Jégu, 2003). Its body shape is similar to that of piranhas (round-shaped fish), reaching up to 1 m in length and a weight of up to 30 kg (Goulding and Carvalho, 1982). The natural food of tambaqui, as opposed to the carnivorous piranhas, is composed mainly of fruits and seeds (Lucas, 2008). Aquaculture production of tambaqui has increased over the past 10 years in several countries of South America, and particularly in Bolivia, Brazil, Colombia, Ecuador, Peru, and Venezuela (reviewed in Valladão et al., 2016). Tambaqui is currently considered the most important native fish produced in Brazilian aquaculture, with approximately 135 thousand tons produced in 2015 (IBGE, 2016). Production of tambaqui is also widespread throughout Brazil, particularly in the north, northeast, midwest, and southeast regions of the country (MPA, 2013).

In South America, most of the fish production by aquaculture is resulting from exotic species such as tilapia, *Oreochromis niloticus*, and salmon, *Salmo salar* (Valladão et al., 2016). However, different from tilapia and salmon, which has 20 and 13 breeding programs for genetic selection, tambaqui has few studies in order to select individuals of better performance for aquaculture (Gjedrem et al., 2012; Marcos et al., 2016). Tilapia is the main fish produced in Brazil (approximately 220 thousand tons per year) (IBGE, 2016), and most of this production is a result of the genetically improved farmed tilapia (GIFT) program, a genetic selection project developed over several generations (Ponzoni et al., 2011). Knowledge about growth performance and disease resistance in aquaculture is often derived from research using model fish, such as rainbow trout (*Oncorhynchus mykiss*), salmon, and tilapia (Gjedrem, 2000, 2010; LaFrentz et al., 2016; Shoemaker et al., 2017). Similar studies, however, are scarce for native species such as tambaqui. Any additional information on the genetic improvement of this species will contribute significantly toward increasing its productivity in aquaculture operations.

Growth performance (e.g., weight gain and morphometric parameters) is the only trait that has been studied for improving the breeding of tambaqui (Marcos et al., 2016; Mello et al., 2016). However, disease resistance is another important characteristic that can be evaluated for the genetic selection of this species. In South Brazil, where temperature variations during station changes are frequent, tambaquis suffer from outbreaks of disease, resulting in economic losses to the producers. The bacterium *Aeromonas hydrophila* is one the main opportunistic pathogens that causes disease in tambaqui (Valladão et al., 2016). This microorganism is often associated with hemorrhagic septicemia, fin erosion, and distention of the abdominal cavity, among other pathological conditions (Paniagua et al., 1990), and causes mass mortality in several fish species (Abdel-Tawwab et al., 2008; Sirimanapong et al., 2014).

Molecular markers associated with productive traits are being increasingly used to improve breeding programs (Houston et al., 2012; Yáñez et al., 2014). Marker-assisted selection (MAS) can increase the genetic gain of target species in approximately 25–50% of cases, compared to traditional selection techniques (Ma et al., 2014). The MAS approach associates genotypes with phenotypes to increase accuracy when predicting the genetic values of selection candidates (Yáñez et al., 2014). Moreover, MAS can be performed during the early life stages, and is indicated for traits such as disease resistance that are difficult to measure directly in selection candidates (Ma et al., 2014; Taylor, 2014). Therefore, numerous studies have been performed on fish species to identify, characterize, and validate molecular markers associated with different traits, including growth performance and disease resistance (Song et al., 2012; Rodríguez-Ramilo et al., 2013).

The objective of this study was to evaluate growth parameters and resistance to *A. hydrophila* in three full-sib families of tambaqui, *C. macropomum*. Moreover, we tested a set of microsatellites, including anonymous and gene-associated markers, to verify the association of tambaqui genotypes with its growth performance and resistance to *A. hydrophila*.

MATERIALS AND METHODS

Ethics Statement

This study was approved by the Ethics Committee on Animal Use (CEUA number 18.764/16) of Faculdade de Ciências Agrárias e Veterinárias, UNESP, Campus Jaboticabal, SP, Brazil. Fish were anesthetized during the experiments with benzocaine (40 mg/L) and all efforts were made to minimize suffering.

Family Obtainment

Three full-sib families of tambaqui were obtained by mating three breeding pairs from the Aquaculture Center of UNESP, Campus Jaboticabal, SP, Brazil. Spawning was induced using carp pituitary extract that was dissolved in saline solution (0.9% NaCl) and applied in two dosages at an interval of 12 h (first and second dosage of 0.6 and 5.4 mg/kg, respectively), according to the protocol described by Pinheiro et al. (1988). After hatching in 20 L conical incubators over 15 days, approximately 150 larvae per family were transferred to 60 L tanks and maintained there until they reached approximately 3 g of body weight. Juveniles were identified by pit-tags placed posteriorly, and transferred to three different 750 L tanks and maintained until 6 months after hatching.

Bacterial Challenge

Aeromonas hydrophila was previously isolated from diseased *Colossoma macropomum* and stored in 30% glycerol at −80°C. This pathogenic strain was subsequently cultured in a nutrient tryptic soy agar (TSA), Vegitone (Sigma-Aldrich, India), for 24 h at 28°C. The colony was then transferred to nutrient tryptic soy broth (TSB) (Fluka, Sigma-Aldrich, India) and cultured for 24 h at 28°C. After growth, the culture broth was centrifuged at 3000 × g for 10 min at 4°C (Eppendorf Centrifuge 5810). The supernatant was discarded and the bacterial pellet was resuspended in sterile phosphate-buffered saline (PBS). This washing procedure was repeated twice. The optical density (OD) of the solution (previously determined for LD_{50} tests—lethal dosage in 50% of fish) was adjusted to 0.8 at 600 nm (CDCP, WHO, 2003) in a spectrophotometer (2100 Unico, Japan). A sample of 100 μl of this solution was removed from the inoculum (used for the fish challenge test) for conducting serial dilutions and plate counts in duplicate on TSA. A LD_{50} test was previously conducted on 30 individuals (10 from each family), using the following dosages: 10^6, 10^7, and 10^8 CFU (colony forming units)/ml of *A. hydrophila*.

In the final experiment, 48 full-sib animals were used from each family (3 families × 48 fish = 144 fish). Each family was further separated into triplicate treatment groups of 12 fish (12 fish × 3 treatment replicates = 36 fish), plus an additional 12 fish were used for the control group. Replicate and control groups were composed of 36 individuals (i.e., 12 × 3 families). Each replicate was maintained in 750 L tanks, using a recirculation system fitted with mechanical and biological filters. Water temperature was maintained at 30 ± 1°C.

After adjustment of the LD_{50} (1×10^8 CFU), infection was induced via intraperitoneal inoculation with a dose of 0.1 ml of bacteria per 10 g of fish body weight. Fish in the control

group were inoculated with saline solution. The experiment was conducted over a period of 15 days, in order to evaluate resistance traits of mortality rate and the time of death. To evaluate the resistance among families, death time variance was performed using Tukey's test ($p < 0.05$) and the mortality rate was assessed by a chi-square test ($p < 0.01$).

Growth Traits

Six growth traits were measured in 96 individuals (32 fish from each family). Traits assessed included head length (HL), head height (HH), standard length (SL), height from the beginning of the dorsal fin to the beginning of the ventral fin (H1), height from the end of the dorsal fin to the beginning of the anal fin (H2), and body weight (BW). The five morphometric traits were measured in millimeters using the Image J/Fiji 1.46 software and BW was measured by a precision digital balance (accuracy of 0.01 g). After measurements, correlation of traits was estimated using the Pearson correlation (r). We did not compare growth performance among the families because each family grew in different environments (individual tanks).

Evaluation of Microsatellite Markers

We used 13 gene-associated microsatellite markers in total (c5009, c5837, c4604, c3592, c3842, c4296, c2311, c3818, c841, c4706, C2647, c3843, c3905) and 4 anonymous markers (r1366, r912, r415, r3808) for evaluating associations with growth and resistance traits (Ariede et al., 2018). All microsatellites were previously tested for polymorphism in the progenitors. However, only heterozygous markers in at least one of the progenitors were selected for analysis in the progeny. Only nine microsatellites (c5837, c4604, c3592, c3842, c4296, c2311, c3818, r1366, r912) were selected with the previous characteristics, seven of which were gene-associated and two of which were anonymous markers. Individuals from family 1 were genotyped with the microsatellites c2311, c3592, c5837, c3818, and c4296; family 2 with c3842; and family 3 with c4604, c3842, r1366, and r912.

Genomic DNA was extracted from all full-sibs and parents, following the Wizard Genomic DNA Purification Kit (Promega) protocol. The sequencing strategy adopted in this study followed the protocols described by Schuelke (2010), and used the CAGtag primer (5′-CAGTCGGGCGTCATCA-3′) (Shirk et al., 2013) labeled with the fluorochromes HEX, FAM, or NED. The genotyping by polymerase chain reaction (PCR) was performed with the following reagents: 100 μM of dNTP, 1.5 mM $MgCl_2$, 1 × Taq DNA buffer, 0.1 μM of each primer (F and R), 0.01 μM of CAGtag primer, 0.5 units of Taq Polymerase (Invitrogen), and 10–50 ng of genomic DNA. The cycling program for amplification consisted of: 9 cycles of 95°C for 30 s, 55–60°C for 30 s (adjusted for each primer set), 72°C for 20 s; followed by 30 cycles of 95°C for 30 s, 50°C for 30 s and 72°C for 20 s. During the first nine cycles, the annealing temperature of 55–60°C allowed for the incorporation of the primers (F and R) from the microsatellite loci. In the following 30 cycles, the temperature of 50°C facilitated the annealing of the fluorescent dye-labeled CAGtag primer. PCR products were analyzed by capillary electrophoresis in the equipment ABI3730 XL, using the DS-30 matrix with the GeneScan 500 ROX dye Size Standard (Thermo). We used the program GeneMapper 4.0 (Applied Biosystems) to analyze allele sizes.

Chi-square tests ($p < 0.01$) were performed on progeny genotypes to verify if they were distributed according to the Mendelian pattern. The association between microsatellite markers and growth traits and time of death was evaluated with the generalized linear model procedure of the analysis of variance (PROC-GLM ANOVA), using the statistical program, SAS (Statistical Analysis System, version 9.2). A linear animal model was used as follows: $Y = m + G + e$, where **Y** is the observed value of a given continuous characteristic (growth traits or time of death); **m** is the overall mean of the characteristic; **G** is the effect of the genotype; and **e** the random error effect ($p < 0.05$). The statistical model was based on that described by Ma et al. (2014), with some modifications. The mortality rate was considered as a binary variable (live = 1 and dead = 0); thus, this characteristic was evaluated by means of a contingency table (chi-square test, $p < 0.01$).

RESULTS

Analysis of Resistance to *A. hydrophila*

All infected animals presented clinical signs of *A. hydrophila* infection. The average of death time observed in families 1, 2, and 3 was 17.20, 9.13, and 18.25 h, respectively, indicating significant differences between families 1 and 2, and families 2 and 3 (**Figure 1A**). Moreover, mortality rates were significantly different between the families (**Figure 1B**); the mortality rate of family 2 (57.6%) was higher than the mortality rates of family 1 (30.5%) and family 3 (28.5%). There was no fish mortality observed in the control treatment.

Analysis of Growth Performance

The average BW in families 1, 2, and 3 was 44.22, 38.22, and 30.68 g, respectively. Body weights in families 1, 2, and 3 ranged between 12–70 g, 16–67 g, and 21–57 g, respectively. The average SL in families 1, 2, and 3 was 106.47, 103.69, and 96.01 mm, respectively. Other growth traits are described in **Table 1**. Phenotypic correlation indicated that all characteristics were significantly correlated ($p < 0.01$). Measurements of SL and H1, however, were poorly correlated with a correlation value of 0.37 (**Table 2**).

Analysis of Association with Microsatellite Loci

Genotypes of the parents with their respective microsatellites are described in **Table 3**. The markers c4604, c3842, and r1366 that were applied in family 3 did not follow the Mendelian segregation pattern ($p < 0.01$), and were therefore excluded from the association analysis. In relation to the analysis of resistance to *A. hydrophila*, significant association was not detected between microsatellites and death time or mortality rate. In contrast, growth traits had significant association with markers c3842 and r912 in families 2 and 3, respectively. Progeny genotypes and growth traits averages are described in **Table 4**. The gene-associated marker c3842 showed an association with SL, and the anonymous marker r912 was associated with H2. For the marker c3842, individuals with the genotype 183/183 had an

average SL of 111.56 mm, whereas those with the genotype 183/186 showed an average SL of 98.33 mm (**Table 4**). Moreover, for the marker r912, fish with the genotypes 175/175 and 175/185 demonstrated an average H2 of 27.44 and 29.65 mm, respectively.

DISCUSSION

Challenge experiments of *A. hydrophila* in this study revealed significant variations in death times and mortality rates between the three families of tambaqui. Our results indicated that family 2 was more susceptible to *A. hydrophila* than families 1 and 3; death time was lower and the mortality rate higher in family 2, compared with those observed in the other two families. These results indicate genetic variation for this trait in the families of tambaqui analyzed in this study; selective breeding could thus be applied for resistance to *A. hydrophila* in tambaqui, as has been done in other model aquaculture species such as salmon,

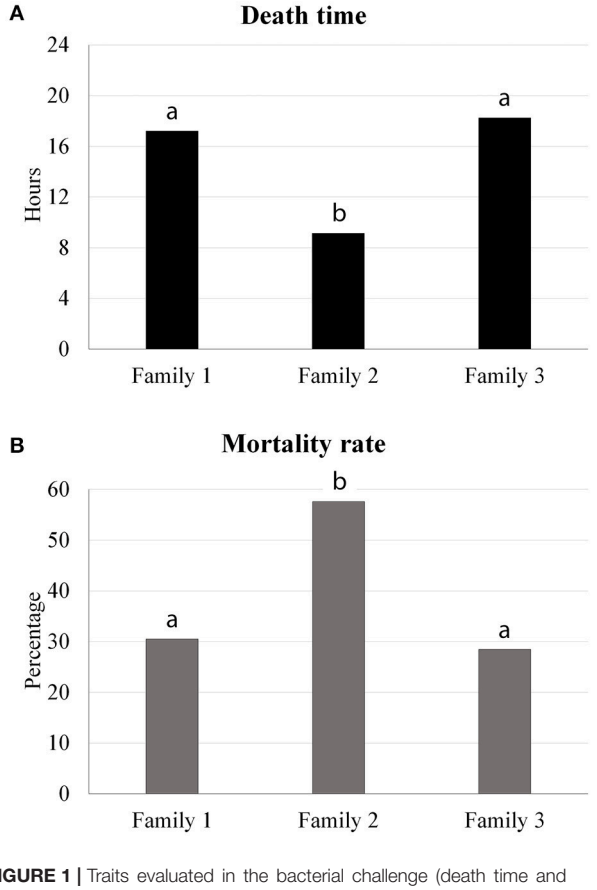


FIGURE 1 | Traits evaluated in the bacterial challenge (death time and mortality rate) indicating significant differences between the tambaqui families.

TABLE 2 | Analysis of Pearson correlation (r) between growth traits (weight and five morphometric measures) in three families of tambaqui *Colossoma macropomum*.

	BW	HL	HH	SL	H1
HL	0.71				
HH	0.70	0.94			
SL	0.73	0.84	0.83		
H1	0.62	0.58	0.62	0.37	
H2	0.73	0.66	0.70	0.73	0.68

All values represented significant correlations ($p < 0.05$). BW, body weight (g); HL, head length (mm); HH, head height (mm); SL, standard length (mm); H1, height from the beginning of the dorsal fin to the beginning of the ventral fin (mm); H2, height from the end of the dorsal fin to the beginning of the anal fin (mm).

TABLE 3 | Characterization of heterozygous microsatellites in at least one of the progenitors in three families of tambaqui *Colossoma macropomum*.

Family	*Loci*	Genotype♂	Genotype♀
1	c5837	174/179	174/174
	c2311	163/169	166/175
	c3818	204/214	204/210
	c3592	166/170	170/170
	c4296	118/122	118/118
2	c3842	183/186	183/183
3	c4604	142/151	148/151
	c3842	183/186	183/186
	r1366	261/277	261/277
	r912	175/175	175/185

Genotypes are represented by allele size in base pairs (bp).

TABLE 1 | Analysis of growth performance in three families of tambaqui *Colossoma macropomum*, including average values and standard deviation for weight and morphometric measures.

Family	Traits					
	BW	HL	HH	SL	H1	H2
1	44.5 ± 12.7	38.0 ± 4.9	34.5 ± 4.0	105.6 ± 11.3	55.4 ± 5.8	32.6 ± 4.7
2	39.8 ± 8.0	38.5 ± 4.8	35.2 ± 3.9	104.3 ± 18.0	55.6 ± 10.1	32.2 ± 4.1
3	31.7 ± 10.7	35.4 ± 4.6	31.1 ± 3.4	96.4 ± 8.2	47.4 ± 4.4	28.5 ± 3.0

BW, body weight (g); HL, head length (mm); HH, head height (mm); SL, standard length (mm); H1, height from the beginning of the dorsal fin to the beginning of the ventral fin (mm); H2, height from the end of the dorsal fin to the beginning of the anal fin (mm).

TABLE 4 | Analysis of the microsatellite loci in the progeny and average values of growth traits (weight and five morphometric measures) by clustering individuals according to the respective genotype in three families of tambaqui *Colossoma macropomum*.

Family	*Loci*	Progeny genotype	Genotype frequency	BW	HL	HH	SL	H1	H2
1	c5837	174/174	0.34	42.55	37.61	34.61	104.80	53.60	34.15
		174/179	0.66	42.52	38.17	34.44	106.07	53.27	31.92
	c2311	163/166	0.25	36.50	36.07	32.96	101.72	51.14	30.39
		163/175	0.40	45.08	38.94	35.45	107.49	54.44	33.92
		166/169	0.19	41.67	37.26	33.49	105.26	53.37	32.32
		169/175	0.16	46.60	39.37	35.67	107.51	54.25	33.61
	c3818	204/204	0.24	39.29	37.61	34.01	102.54	51.52	30.41
		204/210	0.34	37.40	36.74	33.67	101.44	50.86	32.17
		204/214	0.18	43.75	36.88	34.10	106.19	54.52	32.77
		210/214	0.24	51.71	41.36	36.62	114.07	57.56	35.60
	c3592	166/170	0.59	42.47	37.28	34.18	104.84	53.06	32.24
		170/170	0.41	42.62	38.99	34.96	106.78	53.85	33.34
	c4296	118/118	0.62	43.60	38.30	34.83	106.20	53.49	32.60
		118/122	0.38	40.75	37.45	33.94	104.68	53.21	32.82
2	c3842	183/183	0.47	45.29	39.94	36.18	111.56	55.15	32.71
		183/186	0.53	36.50	38.13	34.81	98.33	54.42	32.38
3	c4604*	142/148	0.47	32.53	35.05	30.75	97.50	48.38	29.69
		142/151	0.22	30.43	34.67	31.30	95.06	46.04	27.09
		148/151	0.06	24.50	34.07	28.40	91.31	44.29	26.92
		151/151	0.25	32.88	36.92	32.22	96.39	47.16	27.81
	c3842*	183/183	0.03	34.00	35.74	30.19	86.76	42.95	24.28
		183/186	0.56	30.72	35.60	31.41	96.13	47.79	28.94
		186/186	0.41	32.77	35.03	30.72	97.29	46.98	28.17
	r1366*	261/261	0.06	29.00	32.21	29.85	94.11	46.71	27.93
		261/277	0.66	32.71	36.30	31.41	97.30	47.77	28.84
		277/277	0.28	29.78	33.91	30.62	94.47	46.35	27.76
	r912	175/175	0.53	30.35	34.18	31.02	94.41	46.29	27.44
		175/185	0.47	33.13	36.72	31.16	98.46	48.46	29.65

*Genotypes are represented by allele size in base pairs (bp). BW, body weight (g); HL, head length (mm); HH, head height (mm); SL, standard length (mm); H1, height from the beginning of the dorsal fin to the beginning of the ventral fin (mm); H2, height from the end of the dorsal fin to the beginning of the anal fin (mm). *Markers that did not follow the Mendelian segregation pattern ($p < 0.01$).*

tilapia and rainbow trout (Silverstein et al., 2009; Ødegård et al., 2011; Wiens et al., 2013; Evenhuis et al., 2015; LaFrentz et al., 2016; Shoemaker et al., 2017). For example, Nile tilapia families challenged with the pathogen *Streptococcus iniae* showed an additive genetic component to resistance, indicated by the variation in survival rate; genetic selection methods can be thus applied in the production of animals resistant to this pathogen (LaFrentz et al., 2016).

Standard methods for disease prevention and control in the aquaculture industry include the use of vaccines, antibiotics, and management strategies (fish density, water flow, oxygen control, etc.), all of which are only partially effective. Vaccination is generally expensive and impractical for large-scale fish production, as it requires that all fish are individually treated (Yáñez et al., 2014). The indiscriminate use of antibiotics can lead to the development of populations of antibiotic-resistant bacteria, and pose a risk to human health and the environment (Vivekanandhan et al., 2002; Taylor et al., 2011). Conversely, selective breeding for disease resistance is a more effective, sustainable, and low-cost approach to improving fish health and performance, without the risks and inefficiencies posed by the other standard methods (Stear et al., 2001; Bishop, 2010).

This study did not identify the association of molecular markers with resistance to *A. hydrophila* in tambaqui. In general, as this trait is usually controlled by polygenic architecture, studies to search for association with resistance to pathogens have previously been conducted using hundreds of molecular markers on species such as turbot (*Scophthalmus maximus*), rainbow trout, salmon, Nile tilapia, and cod (*Gadus morhua*) (Pardo et al., 2008; Ødegård et al., 2010; Yáñez et al., 2014; Evenhuis et al., 2015; LaFrentz et al., 2016). Therefore, studies specifically on tambaqui are required to search for quantitative trait loci (QTL) of resistance to *A. hydrophila,* using a higher marker density.

The results of phenotypic correlation between morphometric measures in this study were similar to those previously reported for tambaqui by Mello et al. (2016); these authors demonstrated

that selection for filet weight can be made on the basis of morphometric measures due to positive correlation between these characteristics. In both studies, weight had a higher phenotypic correlation with standard length and head length ($r = 0.70$ in both studies). Moreover, the highest phenotypic correlation was observed between head length and standard length ($r = 0.84$ in the present study, and 0.90 in Mello et al., 2016). This pattern of high correlation between body measurements and weight ($r = 0.90$) was also observed in Nile tilapia (Rutten et al., 2005), indicating that the selection for a specific trait would also result in the gain of other traits, as is the case for tambaqui.

The morphometric data observed in *C. macropomum* indicates that this species shows a large head length, which is not attractive for the commercial sector; a larger head length lowers the filet yield (Mello et al., 2016). As observed in both studies, there is a positive correlation between head length and weight; therefore, it is not possible to select individuals of greater body weight without them also having a larger head length.

Several studies using growth markers were performed on genes that are directly linked to this phenotype, such as the growth hormone (*GH*) gene and the insulin gene (*IGF1*) (Yue and Orban, 2002; Tsai et al., 2014). In a study with two tilapia species, six microsatellites were detected in four genes that were involved in growth and reproduction processes (Yue and Orban, 2002). The c3842 marker, identified in this study in association with growth, is inserted in the region of the repeat-containing 6b isoform x5 gene (*TNCRC6b*), located in the 5'-UTR region. This gene has already been described in mammals (humans and rats) and in fish (*Maylandia zebra*), and its function may be related to various biological processes such as the silencing of genes guided by micro RNAs, and the regulation of positive messenger RNA, among others. It is thus necessary to increase our understanding of the role of this gene and the possible relation between its molecular functions and differential growth in tambaqui. The results obtained from the association analysis with the gene-associated marker c3842 suggest that the allele 183, when homozygous, exerts influence on the increase of the mean SL of the animal.

The second microsatellite that showed association with growth was the anonymous marker r912. In general, anonymous markers are often used in QTL identification studies, because such markers are physically close to these loci and therefore closely linked (Reid et al., 2005; Bouza et al., 2012). Our results suggest that r912 is related to a second morphometric characteristic (Height 2). Moreover, and contrary to that observed in the gene-associated marker, increased height occurs when the r912 marker is heterozygous, suggesting that the presence of allele 185 exerts influence on animal height. To date, the reference genome of tambaqui is not available; therefore, it is not possible to characterize the loci of the r912 marker and the putative linked gene. A similar strategy of gene prospection was carried out in salmon, whose reference genome is sequenced, allowing the prediction of two markers in association with the genes *MEP1A* (meprin A subunit beta-like) and *PCNT* (pericentrin) (Houston et al., 2014; Tsai et al., 2016).

Currently, 100% of the tambaqui production is still based on stocks without genetic improvement. The identification of two markers (c3842 and r912) associated with growth in tambaqui will support and accelerate the processes for selection of superior genotypes using MAS. A classic example of QTL identification is the detection of a haplotype that is associated with approximately 80% of the genetic variation in salmon resistant to pancreatic necrosis infection (Houston et al., 2008, 2012). Owing to their efficacy, these markers are currently being successfully incorporated into breeding programs of Atlantic salmon in Norway and Scotland (Moen et al., 2009; Houston et al., 2010). The results of this study will thus likely increase our knowledge base of tambaqui aquaculture, and facilitate the development of future tambaqui breeding programs.

CONCLUSION

This study provides new information on tambaqui (*C. macropomum*) growth traits to guide future breeding programs for this species. The prospective microsatellites highlighted in this study could be used in the validation of MAS in families of *C. macropomum*. This study also contributes to the development of a bacterial challenge protocol with *Aeromonas hydrophila* in order to characterize pathogen-resistant tambaqui families for selective breeding.

AUTHOR CONTRIBUTIONS

RA–Acquisition, analysis and interpretation of data, draft of the work, development of intellectual content, writing of the manuscript, final approval of the version. MF, VM-F, FP, and SB–Draft of the work, development of intellectual content, final approval of the version. MH–Analysis and interpretation of data, draft of the work, final approval of the version. FP-F–Acquisition, Interpretation of data, development of intellectual content, final approval of the version. DH–Draft of the work, development of intellectual content, writing of the manuscript, final approval of the version.

ACKNOWLEDGMENTS

This work was supported by grants from FAPESP (2014/03772-7) and CNPq (446779/2014-8, 130629/2015-4 and 305916/2015-7).

REFERENCES

Abdel-Tawwab, M., Abdel-Rahman, A. M., and Ismael, N. E. M. (2008). Evaluation of commercial live baker's yeast, *Saccharomyces cerevisiae*, as a growth and immunity promoter for Fry Nile tilapia, *Oreochromis niloticus* challenged *in situ* with *Aeromonas hydrophila*. *Aquaculture* 280, 185–189. doi: 10.1016/j.aquaculture.2008.03.055

Ariede, R. B., Freitas, M. V., Hata, M. E., Matrochirico-Filho, V. A., Utsunomia, R., Mendonça, F. F., et al. (2018). Development of microsatellite markers using next-generation sequencing for the fish *Colossoma macropomum*. *Mol. Biol. Rep*. 45, 9–18. doi: 10.1007/s11033-017-4134-z

Bishop, S. C. (2010). "Disease resistance: genetics," in *Encyclopedia of Animal Science*, ed W. G. Pond and A. W. Bell (New York, NY: Marcel Dekker, Inc.), 288–290.

Bouza, C., Hermida, M., Pardo, B. G., Vera, M., Fernández, C., de la Herrán, R., et al. (2012). An Expressed Sequence Tag (EST) enriched genetic map of turbot (*Scophthalmus maximus*): a useful framework for comparative genomics across model and farmed teleosts. *BMC Genet.* 13:54. doi: 10.1186/1471-2156-13-54

Evenhuis, J. P., Leeds, T. D., Marancik, D. P., LaPatra, S. E., and Wiens, G. D. (2015). Rainbow trout (*Oncorhynchus mykiss*) resistance to columnaris disease is heritable and favorably correlated with bacterial cold water disease resistance. *J. Anim. Sci*. 93, 1546–1554. doi: 10.2527/jas.2014-8566

Gjedrem, T. (2000). Genetic improvement of cold-water fish species. *Aquacult. Res*. 31, 25–33. doi: 10.1046/j.1365-2109.2000.00389.x

Gjedrem, T. (2010). The first family-based breeding program in aquaculture. *Rev. Aquacult*. 2, 2–15. doi: 10.1111/j.1753-5131.2010.01011.x

Gjedrem, T., Robinson, N., and Rye, M. (2012). The importance of selective breeding in aquaculture to meet future demands for animal protein: a review. *Aquaculture* 350–353, 117–129. doi: 10.1016/j.aquaculture.2012.04.008

Goulding, M., and Carvalho, M. L. (1982). Life history and management of the tambaqui (*Colossoma macropomum*, *Characidae*): an important Amazonian food fish. *Rev. Bras. Zool*. 1, 107–133. doi: 10.1590/S0101-81751982000200001

Houston, R. D., Davey, J. W., Bishop, S. C., Lowe, N. R., Mota-Velasco, J. C., Hamilton, A., et al. (2012). Characterisation of QTL-linked and genome-wide restriction site associated DNA (RAD) markers in farmed Atlantic salmon. *BMC Genomics* 13:244. doi: 10.1186/1471-2164-13-244

Houston, R. D., Gheyas, A., Hamilton, A., Guy, D. R., Tinch, A. E., Taggart, J. B., et al. (2008). Detection and confirmation of a major QTL affecting resistance to infectious pancreatic necrosis (IPN) in Atlantic salmon (*Salmo salar*). *Dev. Biol.* 132, 199–204. doi: 10.1159/000317160

Houston, R. D., Haley, C. S., Hamilton, A., Guy, D. R., Mota-Velasco, J. C., Gheyas, A. A., et al. (2010). The susceptibility of Atlantic salmon fry to freshwater infectious pancreatic necrosis is largely explained by a major QTL. *Heredity* 105, 318–327 doi: 10.1038/hdy.2009.171

Houston, R. D., Taggart, J. B., Cézard, T., Bekaert, M., Lowe, N. R., Downing, A., et al. (2014). Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). *BMC Genomics* 15:90. doi: 10.1186/1471-2164-15-90

IBGE (2016). Produção da Pecuária Municipal 2015. *Rio de Janeiro* 41, 1–108.

Jégu, M. (2003). "Subfamily Serrasalminae (Pacus and piranhas)," in *Check List of the Freshwater Fishes of South and Central America*, eds R. E. Reis, S. O. Kullander, and C. J. Ferraris (Porto Alegre: Edipucrs), 182–196.

LaFrentz, B. R., Lozano, C. A., Shoemaker, C. A., García, J. C., Xu, D. H., Løvoll, M., et al. (2016). Controlled challenge experiment demonstrates substantial additive genetic variation in resistance of Nile tilapia (*Oreochromis niloticus*) to *Streptococcus iniae*. *Aquaculture* 458, 134–139. doi: 10.1016/j.aquaculture.2016.02.034

Lucas, C. M. (2008). Within flood season variation in fruit consumption and seed dispersal by two characin fishes of the Amazon. *Biotropica* 40, 581–589. doi: 10.1111/j.1744-7429.2008.00415.x

Ma, H., Jiang, W., Liu, P., Feng, N., Ma, Q., Ma, C., et al. (2014). Identification of transcriptome-derived microsatellite markers and their association with the growth performance of the mud crab (*Scylla paramamosain*). *PLoS ONE* 9:e89134. doi: 10.1371/journal.pone.0089134

Marcos, R., Povh, J. A., Fornari, D. C., Oliveira, C. A. L., Ribeiro, R. P., Lopera-Barrero, N. M., et al. (2016). Weight gain and morphometric growth of genetically improved tambaqui (*Colossoma macropomum*). *Semina: Ciências Agrárias* 37, 2521–2528. doi: 10.5433/1679-0359.2016v37n4Supl1p2521

Mello, F. D., Oliveira, C. A. L., Streit, Jr., D., Resende, E. K., Oliveira, S. N., Fornari, D. C., et al. (2016). Estimation of genetic parameters for body weight and morphometric traits to tambaqui *Colossoma macropomum*. *J. Fish. Sci*. 10, 96–100.

Moen, T., Baranski, M., Sonesson, A. K., and Kjøglum, S. (2009). Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*): population-level associations between markers and trait. *BMC Genomics* 10:368. doi: 10.1186/1471-2164-10-368

MPA (2013). *Ministério da Pesca e Aquicultura. Boletim Estatístico da pesca e Aquicultura—2011*. Brasília: Brasil-Ministério da Pesca e Aquicultura.

Ødegård, J., Baranski, M., Gjerde, B., and Gjedrem, T. (2011). Methodology for genetic evaluation of disease resistance in aquaculture species: challenges and future prospects. *Aquacult. Res*. 42, 103–114. doi: 10.1111/j.1365-2109.2010.02669.x

Ødegård, J., Sommer, A. I., and Praebel, A. K. (2010). Heritability of resistance to viral nervous necrosis in Atlantic cod (*Gadus morhua*). *Aquaculture* 300, 59–64. doi: 10.1016/j.aquaculture.2010.01.006

Paniagua, C., Rivero, O., Anguita, J., and Naharro, G. (1990). Pathogenicity factors and virulence for rainbow trout (*Salmo gairdneri*) of motile *Aeromonas* spp. isolated from a river. *J. Clin. Microbiol*. 28, 350–355

Pardo, B. G., Fernández, C., Millán, A., Bouza, C., Vázquez-López, A., Vera, M., et al. (2008). Expressed sequence tags (ESTs) from immune tissues of turbot (*Scophthalmus maximus*) challenged with pathogens. *BMC Vet. Res*. 4:37. doi: 10.1186/1746-6148-4-37

Pinheiro, J. L. P., Silva, M. C. N., Silva, M. S., Soares, M. A. A. Q., and Souza, N. H. (1988). *Produção de Alevinos: Tecnologia Aplicada nas estações de Piscicultura da CODEVASF no Baixo São Francisco*. Brasília: CODEVASF.

Ponzoni, R. W., Nguyen, N. H., Khaw, H. L., Hamzah, A., Bakar, K. R. A., and Yee, H. Y. (2011). Genetic improvement of *Nile tilapia* (*Oreochromis niloticus*) with special reference to the work conducted by the WorldFish center with the Gift strain. *Rev. Aquacult*. 3, 27–41. doi: 10.1111/j.1753-5131.2010.01041.x

Reid, D. P., Szanto, A., Glebe, B., Danzmann, R. G., and Ferguson, M. M. (2005). QTL for body weight and condition factor in Atlantic salmon (*Salmo salar*): comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Arctic charr (*Salvelinus alpinus*). *Heredity* 94, 166–172. doi: 10.1038/sj.hdy.6800590

Rodríguez-Ramilo, S. T., Fernández, J., Toro, M. A., Bouza, C., Hermida, M., Fernández, C., et al. (2013). Uncovering QTL for resistance and survival time to *Philasterides dicentrarchi* in turbot (*Scophthalmus maximus*). *Anim. Genet*. 44, 149–157. doi: 10.1111/j.1365-2052.2012.02385.x

Rutten, M. J. M., Bovenhuis, H., and Komen, H. (2005). Genetic parameters for fillet traits and body measurements in *Nile tilapia* (*Oreochromis niloticus*). *Aquaculture* 246, 125–132. doi: 10.1016/j.aquaculture.2005.01.006

Schuelke, M. (2010). An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol*. 18, 233–234. doi: 10.1038/72708

Shirk, R. Y., Glenn, T. C., Chang, S., and Hamrick, J. L. (2013). Development and characterization of microsatellite primers in *Geranium carolinianum* (Geraniaceae) with 454 sequencing. *Appl. Plant. Sci*. 1:1300006 doi: 10.3732/apps.1300006

Shoemaker, C. A., Lozano, C. A., LaFrentz, B. R., García, J. C., Soto, E., Xu, D. H., et al. (2017). Additive genetic variation in resistance of Nile tilapia (*Oreochromis niloticus*) to *Streptococcus iniae* and S. *agalactiae capsular* type Ib: is genetic resistance correlated? *Aquaculture* 468, 193–198. doi: 10.1016/j.aquaculture.2016.10.022

Silverstein, J. T., Vallejo, R. L., Palti, Y., Leeds, T. D., Rexroad, C. E., Welch, T. J., et al. (2009). Rainbow trout resistance to bacterial cold-water disease is moderately heritable and is not adversely correlated with growth. *J. Anim. Sci*. 87, 860–867. doi: 10.2527/jas.2008-1157

Sirimanapong, W., Thompson, K. D., Kledmanee, K., Thaijongrak, P., Collet, B., Ooi, E. L., et al. (2014). Optimisation and standardisation of functional immune assays for striped catfish (*Pangasianodon hypophthalmus*) to compare their immune response to live and heat killed *Aeromonas hydrophila* as models of infection and vaccination. *Fish Shellfish Immunol*. 40, 374–383. doi: 10.1016/j.fsi.2014.07.021

Song, W., Li, Y., Zhao, Y., Liu, Y., Niu, Y., Pang, R., et al. (2012). Construction of a high-density microsatellite genetic linkage map and mapping of sexual and growth-related traits in half-smooth tongue sole (*Cynoglossus semilaevis*). *PLoS ONE* 7:e52097. doi: 10.1371/journal.pone.0052097

Stear, M. J., Bishop, S. C., Mallard, B. A., and Raadsma, H. (2001). The sustainability, feasibility and desirability of breeding livestock for disease resistance. *Res. Vet. Sci*. 71, 1–7. doi: 10.1053/rvsc.2001.0496

Taylor, J. F. (2014). Implementation and accuracy of genomic selection. *Aquaculture* 420–421(Suppl. 1), S8–S14. doi: 10.1016/j.aquaculture.2013.02.017

Taylor, N. G., Verner-Jeffreys, D. W., and Baker-Austin, C. (2011). Aquatic systems: maintaining, mixing and mobilising antimicrobial resistance? *Trends Ecol. Evol.* 26, 278–284. doi: 10.1016/j.tree.2011.03.004

Tsai, H. Y., Hamilton, A., Guy, D. R., and Houston, R. D. (2014). Single nucleotide polymorphisms in the insulin-like growth factor 1 (IGF1) gene are associated with growth-related traits in farmed Atlantic salmon. *Anim. Genet.* 45, 709–715. doi: 10.1111/age.12202

Tsai, H. Y., Hamilton, A., Guy, D. R., Tinch, A. E., Bishop, S. C., and Houston, R. D. (2016). Verification of snps associated with growth traits in two populations of farmed Atlantic salmon. *Int. J. Mol. Sci.* 17:5. doi: 10.3390/ijms17010005

Valladão, G. M. R., Gallani, S. U., and Pilarski, F. (2016). South american fish for continental aquaculture. *Rev. Aquacult.* doi: 10.1111/raq.12164

Vivekanandhan, G., Savithamani, K., Hatha, A. A. M., and Lakshmanaperumalsamy, P. (2002). Antibiotic resistance of *Aeromonas hydrophila* isolated from marketed fish and prawn of South India. *Int. J. Food Microbiol.* 76, 165–168. doi: 10.1016/S0168-1605(02)00009-0

Wiens, G. D., LaPatra, S. E., Welch, T. J., Evenhuis, J. P., Rexroad, C. E. III., and Leeds, T. D. (2013). On-farm performance of rainbow trout (*Oncorhynchus mykiss*) selectively bred for resistance to bacterial cold water disease: effect of rearing environment on survival phenotype. *Aquaculture* 388–391, 128–136. doi: 10.1016/j.aquaculture.2013.01.018

Yáñez, J. M., Houston, R. D., and Newman, S. (2014). Genetics and genomics of disease resistance in salmonid species. *Front. Genet.* 5:415. doi: 10.3389/fgene.2014.00415

Yue, G. H., and Orban, L. (2002). Microsatellites from genes show polymorphism in two related *Oreochromis* species. *Mol. Ecol. Notes* 2, 99–100. doi: 10.1046/j.1471-8286.2002.00159

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

Copyright © 2018 Ariede, Freitas, Hata, Mastrochirico-Filho, Pilarski, Batlouni, Porto-Foresti and Hashimoto. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Genetics

ORIGINAL RESEARCH
published: 26 January 2018
doi: 10.3389/fgene.2018.00008

Extensive Karyotype Reorganization in the Fish *Gymnotus arapaima* (Gymnotiformes, Gymnotidae) Highlighted by Zoo-FISH Analysis

*Milla de Andrade Machado[1], Julio C. Pieczarka[1], Fernando H. R. Silva[1], Patricia C. M. O'Brien[2], Malcolm A. Ferguson-Smith[2] and Cleusa Y. Nagamachi[1]**

[1] *Laboratório de Citogenética, Centro de Estudos Avançados da Biodiversidade, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém-Pará, Brazil,* [2] *Cambridge Resource Centre for Comparative Genomics, Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom*

The genus *Gymnotus* (Gymnotiformes) contains over 40 species of freshwater electric fishes exhibiting a wide distribution throughout Central and South America, and being particularly prevalent in the Amazon basin. Cytogenetics has been an important tool in the cytotaxonomy and elucidation of evolutionary processes in this genus, including the unraveling the variety of diploid chromosome number ($2n$ = from 34 to 54), the high karyotype diversity among species with a shared diploid number, different sex chromosome systems, and variation in the distribution of several Repetitive DNAs and colocation and association between those sequences. Recently whole chromosome painting (WCP) has been used for tracking the chromosomal evolution of the genus, showing highly reorganized karyotypes and the conserved synteny of the NOR bearing par within the clade *G. carapo*. In this study, painting probes derived from the chromosomes of *G. carapo* (GCA, $2n$ = 42, 30 m/sm + 12 st/a) were hybridized to the mitotic metaphases of *G. arapaima* (GAR, $2n$ = 44, 24 m/sm + 20 st/a). Our results uncovered chromosomal rearrangements and a high number of repetitive DNA regions. From the 12 chromosome pairs of *G. carapo* that can be individually differentiated (GCA1–3, 6, 7, 9, 14, 16, and 18–21), six pairs (GCA 1, 9, 14, 18, 20, 21) show conserved homology with GAR, five pairs (GCA 1, 9, 14, 20, 21) are also shared with cryptic species *G. carapo* $2n$ = 40 (34 m/sm + 6 st/a) and only the NOR bearing pair (GCA 20) is shared with *G. capanema* (GCP $2n$ = 34, 20 m/sm + 14 st/a). The remaining chromosomes are reorganized in the karyotype of GAR. Despite the close phylogenetic relationships of these species, our chromosome painting studies demonstrate an extensive reorganization of their karyotypes.

Keywords: chromosome painting, WCP, *Gymnotus*, FISH, cytotaxonomy, karyotype evolution

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Alexandr Sember, Institute of Animal Physiology and Genetics (ASCR), Czechia
Marcelo De Bello Cioffi, Federal University of São Carlos, Brazil

***Correspondence:**
Cleusa Y. Nagamachi
cleusanagamachi@gmail.com;
cleusa@ufpa.br

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

Received: *23 September 2017*
Accepted: *08 January 2018*
Published: *26 January 2018*

Citation:
Machado MA, Pieczarka JC, Silva FHR, O'Brien PCM, Ferguson-Smith MA and Nagamachi CY (2018) Extensive Karyotype Reorganization in the Fish Gymnotus arapaima (Gymnotiformes, Gymnotidae) Highlighted by Zoo-FISH Analysis. Front. Genet. 9:8. doi: 10.3389/fgene.2018.00008

INTRODUCTION

Gymnotus (Gymnotiformes) is a monophyletic genus of freshwater electric fishes (Albert, 2001; Lovejoy et al., 2010; Tagliacollo et al., 2016) distributed throughout South America (Albert et al., 2005). It represents the most specious genus (40 species; Ferraris et al., 2017) and the widest distribution in the order, with prevalence in the Amazon basin, where several species of *Gymnotus* co-occur in sympatry (Albert and Crampton, 2003; Crampton et al., 2005).

Based on the integrated data from DNA sequencing of six genes, coupled with 223 morphological characters and with Model-Based Total Evidence phylogenetic analyses, Tagliacollo et al. (2016) divided the genus into six clades: *G. pantherinus, G. coatesi, G. anguillaris, G. tigre, G. cylindricus,* and *G. carapo*. The *Gymnotus carapo* group is regarded as monophyletic and is located in a derived position within the genus (Albert, 2001; Lovejoy et al., 2010; Tagliacollo et al., 2016). Craig et al. (2017) described seven subspecies for *G. carapo*.

Cytogenetics has been an important tool in cytotaxonomy and has proved to be very useful in understanding the evolutionary processes behind the diversification of *Gymnotus*. The Gymnotiformes order has considerable variation, not only in diploid number (from $2n = 24$ in *Apteronotus albifrons*, Howell, 1972; Almeida-Toledo et al., 1981; Mendes et al., 2012; to $2n = 74$ in *Rhabdolichops* cf *eastward,* Suárez et al., 2017) but also in the karyotype formula and location of repetitive sequences (Fernandes et al., 2005; Almeida-Toledo et al., 2007; Silva et al., 2009; da Silva et al., 2013; Jesus et al., 2016; Araya-Jaime et al., 2017; Batista et al., 2017; Sousa et al., 2017; Takagui et al., 2017). Recently, fluorescence *in situ* hybridization (FISH), has played an important role in understanding the genome structure of fish species (Yi et al., 2003; Cabral-de-Mello and Martins, 2010; Martins et al., 2011; Vicari et al., 2011; Gornung, 2013; Knytl et al., 2013; Yano et al., 2017) and molecular cytogenetic studies in Gymnotiformes have shown dynamic reorganization, including pericentric inversions observed through repetitive DNA position (Fernandes et al., 2017), sequence dispersion via transposable elements and the association between different repetitive sequences (Utsunomia et al., 2014; da Silva et al., 2016; Machado et al., 2017) and the presence of different sex chromosome systems (Margarido et al., 2007; Henning et al., 2008, 2011; da Silva et al., 2011, 2014; Almeida et al., 2015). This evolutionary plasticity of the karyotype is seen in *Gymnotus* (**Table 1**), a genus that has high interspecific variability in chromosome numbers (**Figure 1**, **Table 1**), ranging from $2n = 34$ in *Gymnotus capanema* (Milhomem et al., 2012a) to $2n = 54$ in *G. carapo* (Foresti et al., 1984), *G. mamiraua* (Milhomem et al., 2007), *G. paraguensis* (Margarido et al., 2007) and *G. inaequilabiatus* (Scacchetti et al., 2011). *Gymnotus arapaima* is located within the *G. carapo* clade, with $2n = 44$ (24 m/sm + 20 st/a; Milhomem et al., 2012b).

Whole chromosome painting (WCP) techniques use specific painting probes of whole chromosomes, chromosomes arms or chromosome regions to find homologous segments in other species (Yang and Graphodatsky, 2017) and Nagamachi et al. (2010) produced whole chromosome probes from *G. carapo* (GCA, $2n = 42$) by chromosome sorting using flow cytometry and made a comparative genomic map against the chromosomal background of the cytotype with $2n = 40$ chromosomes. The results uncovered a high degree of chromosomal repatterning between these cytotypes, with only eight pairs showing conserved synteny (GCA 1, 2, 6, 9, 14, 19, 20, 21). Nagamachi et al. (2013) used the same set of probes for *G. capanema* (GCP, $2n = 34$) and the results showed that the degree of genomic reorganization was much higher, with only four pairs (GCA 6, 7, 19, 20) showing conserved synteny with GCA $2n = 42$ and three pairs (GCA 6, 19, 20) with GCA $2n = 40$. Of these, GCA 7 and 19 are associated with other chromosomes in the karyotype of GCP. The study of Milhomem et al. (2013), with the probe derived from the NOR bearing par of GCA, $2n = 42$, shows that there is a possible synapomorphy of the NOR bearing par within the *G. carapo* clade.

We use the same set of probes produced by Nagamachi et al. (2010) to analyze the karyotype of *G. arapaima* and to compare the results with our previous studies of species in the genus *Gymnotus*. Our findings confirm and extend our understanding of the extensive karyotype reorganization within this genus.

MATERIALS AND METHODS

Sampling

Samples of *G. arapaima* (GAR, $2n = 44$, 24 m/sm + 20 st/a) were collected in the Mamiraua Reserve (Reserva de Desenvolvimento Sustentável Mamiraua) in the Amazon basin, Brazil (03°02′11.8″S 064°51′16.6″W). These samples were previously analyzed by conventional cytogenetic methods (Milhomem et al., 2012b). The animals collected were handled following procedures recommended by the American Fisheries Society. JCP has a permanent field permit, number 13248 from "Instituto Chico Mendes de Conservação da Biodiversidade." The Cytogenetics Laboratory of UFPa has permit number 19/2003 from the Ministry of Environment for sample transport and permit 52/2003 for using the samples for research. The Ethics Committee of the Federal University of Para (Comitê de Ética Animal da Universidade Federal do Pará) approved this research (Permit 68/2015).

WCP

WCP probes from *G. carapo* ($2n = 42$; 30 m/sm + 12 st/a) described in Nagamachi et al. (2010) were hybridized onto metaphases of *G. arapaima* (GAR, $2n = 44$, 24 m/sm + 20 st/a). The chromosomes of GCA, $2n = 42$ were flow-sorted into four regions (R1–R4), from which probes were produced. R1 represented the NOR-bearing chromosome (GCA20), R2 contains the four largest pairs (1–3 and 16); R3 contains the eight medium-sized pairs (4–8 and 17–19) and R4 the eight smallest pairs (9–15 and 21). Additional sorting produced subregion probes (S) from each of the three regions with multiple chromosome pairs included (R2, R3 and R4). R2: S2A (GCA 1, 2 and 16); S2B (GCA 2 and 16) and S2C (GCA 1 and 16). R3: S3A GCA (5–7 and 17); S3B (GCA6 not 7; re-analyzed in Nagamachi et al., 2013, GCA 19); S3C (GCA 7); and S3D (GCA5–7, 17 and 18). R4: S4A (GCA 12, 13 and 15); S4B (GCA 12–15); S4C (GCA 10–13, 15 and 21); and S4D (GCA 12–15 and 21). For details, see **Figure 2** and **Table 1** in Nagamachi et al. (2010).

To find out the corresponding segments between GAR and GCA ($2n = 42$), we used dual-color FISH with probes from R3 and R4. The other non-hybridized chromosomes or segments correspond to R1 (GAR 19, Milhomem et al., 2013) or R2. For a more refined identification of the chromosomes from R2, R3 and R4, we employed dual-color FISH using probes from the subregions as specified in **Table 2** of Nagamachi et al. (2010),

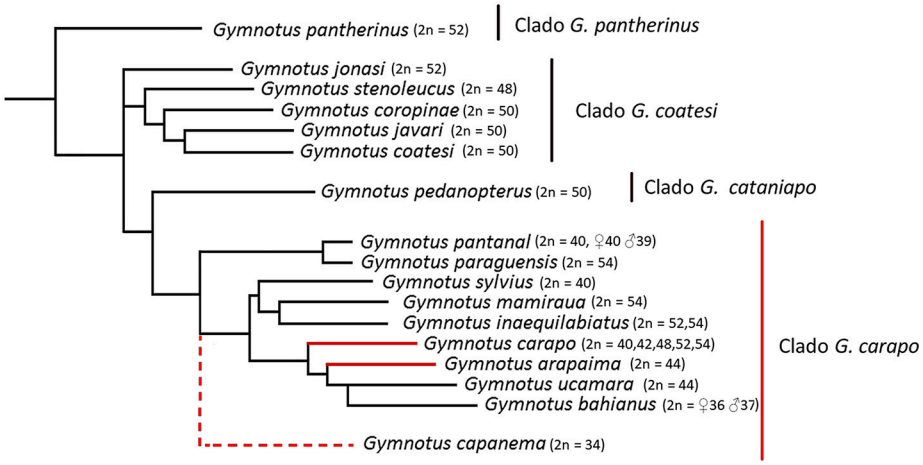


FIGURE 1 | Representative tree of species of *Gymnotus* with diploid number known (Data present in **Table 1**). It was included only species with know phylogenetic relationships, based on data from Albert et al. (2005) and Tagliacollo et al. (2016). *G. capanema* was included in the *G. carapo* clade based on Milhomem et al. (2012a), but has unclear place within the clade.

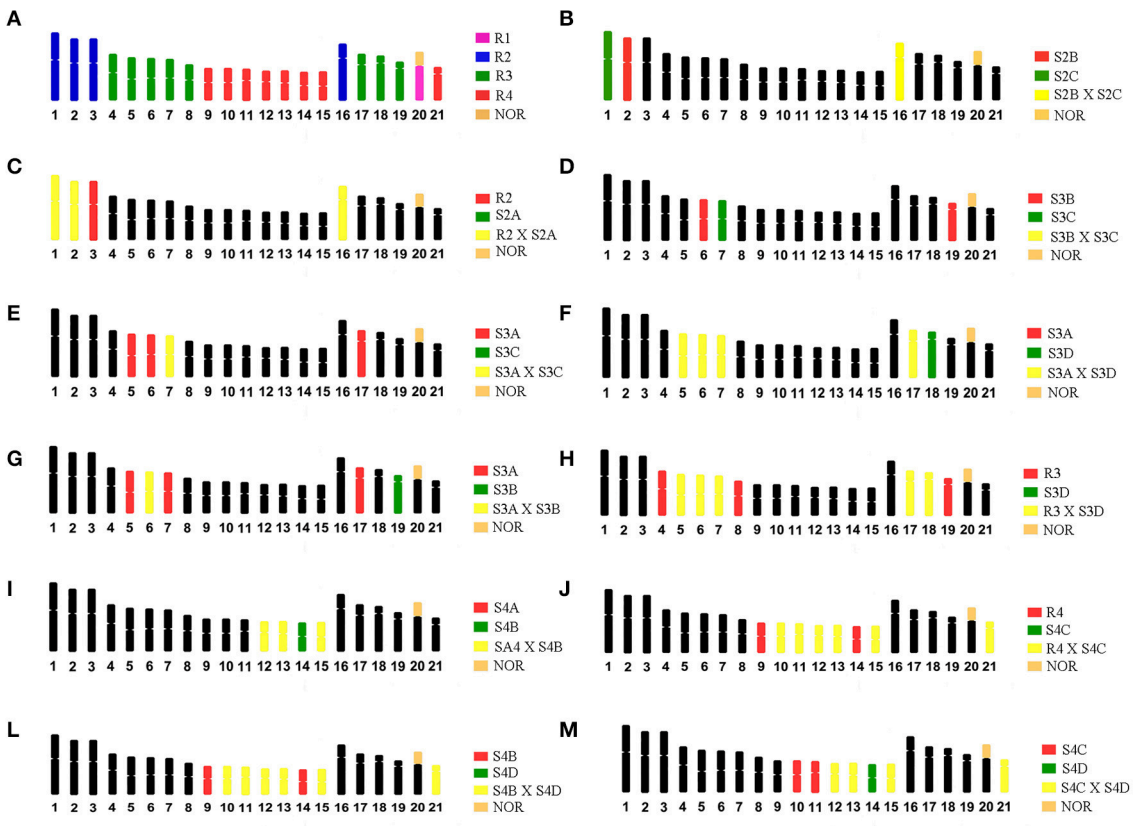


FIGURE 2 | Ideograms of the karyotype of *G. carapo* (*2n* = 42) representing: **(A)** The four chromosome regions (R1, R2, R3, R4) obtained by Nagamachi et al. (2010). **(B–M)** Dual – color FISH experiments based on Nagamachi et al. (2010, 2013), to identify chromosomal homology with GCA (*2n* = 42). S2A, S2B, and S2C represents sub-regions **(A–C)** within region 2; S3A, S3B, S3C, and S3D represents sub-regions **(A–D)** within region 3; S4A, S4B, S4C, and S4D represents sub-regions **(A–D)** within region 4.

TABLE 1 | Cytogenetic data from the genus *Gymnotus*, including 2*n*, karyotype formula (KF), NOR and ribosomal DNA sequences 18 and 5S.

Species	2*n* (KF)	NOR*	18S*	5S*	Authors
Gymnotus arapaima	44 (24 m/sm + 20 st/a)	2	2	–	Milhomem et al., 2012b
Gymnotus bahianus	♀36 (30 m/sm + 6 st) ♂37 (32 m/sm + 5 st)	2	2	2	Almeida et al., 2015
Gymnotus carapo	54 (54 m/sm)	2	–	–	Foresti et al., 1984
	52 (50 m/sm + 2 st/a)	2	–	–	
	48 (34 m/sm + 14 st/a)	–	–	–	
	42 (32 m/sm + 10 st/a)	2	–	–	Fernandes-Matioli et al., 1998
	54 (52 m/sm + 2 st/a)	2	–	–	Claro, 2008
	54 (52 m/sm + 2 st/a)	2	14	2	Milhomem et al., 2007
	42 (30 m/sm + 12 st/a)	2	–	–	Milhomem et al., 2008
	40 (28 m/sm + 12 st/a)	2	–	–	
Gymnotus cf. *carapo*	54 (50 m/sm + 4 st/a)	2	2	≤ 30	Scacchetti et al., 2011
Gymnotus carapo'Catalão'	40 (30 m/sm + 10 st)	–	2	4	da Silva et al., 2014
Gymnotus carapo'Maranhão'	42 (30 m/sm + 12 st/4a)	–	2	14	da Silva, 2015
Gymnotus capanema	34 (20 m/sm + 14 st/a)	2	2	–	Milhomem et al., 2012a
Gymnotus coatesi	50 (24 m/sm + 26 st/a)	8	19	2	Machado et al., 2017
Gymnotus coropinae	♀50 (28 m/sm + 22 st/a) ♂49 (26 m/sm + 23 st/a)	–	2	2	da Silva et al., 2014
Gymnotus inaequilabiatus	52 (50 m/sm + 2 st/a)	2	–	–	Fernandes-Matioli et al., 1998
	54 (52 m/sm + 2 st/a)	–	2	≤34	Scacchetti et al., 2011
Gymnotus javari	50 (20 m/sm + 30 st/a)	–	–	2	Utsunomia et al., 2014
Gymnotus jonasi	52 (12 m/sm + 40 st/a)	6	6	–	Milhomem et al., 2012b
Gymnotus mamiraua	54 (50 m/sm + 4 st/a)	–	–	–	
	54 (38 m/sm + 16 st/a)	2	2	26	Milhomem et al., 2012b da Silva et al., 2016
Gymnotus pantanal	40 (14 m/sm + 26 st/a)	4	–	–	Fernandes et al., 2005.
	♀40 (14 m/sm + 26 st/a) ♂39 (15 m/sm + 24 st/a)	2	–	4	Margarido et al., 2007; da Silva et al., 2011
Gymnotus pantherinus	52 (46 m/sm + 6 st/a)	2	–	–	Fernandes-Matioli et al., 1998
	52 (50 m/sm + 2 st/a)	2	2	4	Scacchetti et al., 2011
Gymnotus paraguensis	54 (52 m/sm + 2 st)	2	–	38	Margarido et al., 2007; da Silva et al., 2011
	54 (50 m/sm + 4 st)	2	–	–	Lacerda and Maistro, 2007
Gymnotus cf. *pedanopterus*	50 (42 m/sm + 8 st/a)	–	2	2	da Silva, 2015
Gymnotus cf. *stenoleucus*	48 (20 m/sm + 28 st/a)	–	2	2	da Silva, 2015
Gymnotus sylvius	40 (38 m/sm + 2 st/a)	2	–	–	Fernandes-Matioli et al., 1998
	40 (30 m/sm + 10 st/a)	2	–	–	Albert et al., 1999
	40 (38 m/sm + 2 st/a)	2	–	–	Claro, 2008
	40 (36 m/sm + 4 st/a)	2	–	–	Lacerda and Maistro, 2007
	40 (36 m/sm + 4 st/a)	2	–	–	Margarido et al., 2007
	40 (34 m/sm + 6 st)	2	2	2	Scacchetti et al., 2011
Gymnotus ucamara	44 (28 m/sm + 16 st/a)	–	2	4	da Silva, 2015
Gymnotus sp.	50 (26 m/sm + 24 st/a)	2	–	–	Lacerda and Maistro, 2007
Gymnotus sp. 'Negro'	♀50 (22 m/sm + 28 st) ♂50 (21 m/sm + 29 st)	–	2	4	da Silva et al., 2014

**Number of chromosome with signals; m, metacentric; sm, submetacentric; st, subtelocentric; a, acrocentric.*

with some modifications related to the identification of the chromosomes of S3B made in Nagamachi et al. (2013). With those experiments (as illustrated in **Figure 2**) it was possible to identify individually GCA pairs 1–3, 6, 7, 9, 14, 16, and 18–21, while it was not possible to distinguish the pairs [4, 8], [10, 11], [5, 17], and [12, 13, 15].

FISH

Chromosome painting techniques followed Yang et al. (1995) with adaptations. Slides were digested with 1% pepsin to remove the excess of cytoplasm, treated with formaldehyde 1%, and dehydrated in ethanol series (2x 2 min 70%, 2x 2 min 90%, and 1x 4 min 100%). Subsequently the slides were aged overnight

TABLE 2 | Chromosome homologies between *G. carapo* (2*n* = 42), *G. carapo* (2*n* = 40), *G. capanema* (2*n* = 34), and *G. arapaima* (2*n* = 44).

Region	*G. carapo* chromosome		*G. capanema* chromosome	*G. arapaima* chromosome
	GCA, $2n = 42$[a]	GCA, $2n = 40$[a]	GCP, $2n = 34$[b]	GAR, $2n = 44$
R 1	20	20	15	19
R 2	1	1	$5q + 9q$	1
	2	2	$3q_{dist} + 16$	$14q_{dist} + 21$
	3	$5q_{dist} + 6\ (p + q_{prox})$	$2p_{dist} + 12q_{dist} + 13q_{prox}$	$13q_{dist} + 18$
	16	$7q + 18\ (p + q_{prox})$	$7q + 14$	$2 + 14q_{prox}$
R 3	[4, 8]	$6q_{dist} + 9q + 10$	$4q_{dist} + 6p$	5+ 20
	6	11	8	$4q + 16q_{int}$
	7	$8p + 9p$	$1q_{dist}$	$3 + 16p$
	[5, 17]	$4, 8q, 18q_{dist}$	$9p + 11 + 12$ (except q_{dist})	$6 + 16q$ (except q_{int})
	18	$3p + 7p$	$1p_{prox} + 2p_{prox} + 4q_{prox}$	15
	19	19	$2q_{dist}$	$7p + 22$
R 4	9	14	$7p + 3q_{prox}$	8
	[10, 11]	$5p + q_{prox}, 12q$	$10 + 17$	$4p + 7q + 12$
	[12, 13, 15]	3q, 12p, 13, 16	$1p_{dist} + 2q_{prox} + 6q_{dist}$	$10 + 11 + 13\ (p + q_{prox})$
	14	17	$13p + 13q_{dist}$	9
	21	15	$6q_{prox} + 1q_{prox}$	17

dist, *distal*; *prox*, *proximal*; *int* , *interstitial*.
[a]*According to Nagamachi et al. (2010).*
[b]*According to Nagamachi et al. (2013).*

at 37°C. The probes were prepared following Nagamachi et al. (2010), denatured for 15 min at 70°C and applied onto a slide with chromosomes that were previously denatured at 70°C for 4 min in 70% formamide/2× SSC [pH 7.0]. The hybridization lasted 72 h at 37°C. The slides were washed once in a solution of 50% formamide/2× SSC, once in 2× SSC and once in 4× Tween, 5 min each.

The dual-color FISH experiments were made with probes that were either directly labeled or biotinylated detected with avidin, (Vector Laboratories, Burlingame, CA, USA) linked to Cy3 or FITC (Amersham, Piscataway, NJ, United States). DAPI (4′,6-diamidino-2-phenylindole) was used as a counterstain.

Microscopy and Image Processing

Image acquisition was made using the software Nis-elements in the microscope Nikon H550S. Chromosomes were morphologically classified according to Levan et al. (1964). The karyotype was organized according to Milhomem et al. (2012b).

RESULTS

The whole chromosome probes from *G. carapo* were hybridized to chromosomes of *G. arapaima*. The regions of homology (hereafter designated as R1-4) obtained with GCA ($2n = 42$) probes against the chromosomes of GAR are indicated on the karyotype of GAR arranged from DAPI-stained chromosomes (**Figure 3**). Dual color FISH with the probes of R3 (red) and R4 (green) defined the chromosome groups in GAR that corresponded to the four groups of regions in GCA (**Figure 3**), as R3 and R4 do not share chromosome pairs. Any chromosome segments hybridizing simultaneously with two colors indicate repetitive DNA sequences that are common to both regions. The chromosomes or segments in blue (DAPI) represent the NOR-bearing chromosomes (R1, GCA20) and the chromosomes corresponding to R2 (pairs 1–3 and 16). **Table 2** shows the correspondence of the GCA ($2n = 42$) chromosomes with the previously published karyotypes of GCA ($2n = 40$) and GCP ($2n = 34$), and GAR ($2n = 44$, present study).

From the 12 chromosome pairs of *G. carapo* that can be individually differentiated (GCA 1–3, 6, 7, 9, 14, 16, and 18–21), six pairs (GCA 1, 9, 14, 18, 20, 21) have conserved homology within GAR. GCA 20 hybridizes to one whole chromosome, pair 19, as described by Milhomem et al. (2013). Six chromosome pairs (GCA 2, 3, 6, 7, 16, and 19) show two signals on GAR chromosomes.

The GCA probes that represent two chromosome pairs [4, 8] revealed two signals, and pairs [10, 11] and [5, 17] revealed three signals and the probe representing three pairs [12, 13, 15] also revealed three signals on GAR chromosomes.

The following associations were found: GAR 4: [10, 11]/C/6, GAR 7: 19/C/[10, 11], GAR 13: [12, 13, 15]/C/ [12, 13, 15]/*/3, GAR 14: */C/16/*/2, GAR 16: 7/C/ [5, 17]/6/ [5, 17] (where C = centromere and * = repetitive sequences).

DISCUSSION

Our results demonstrate that the genomic reorganization in the analyzed species of *Gymnotus* is greater than that assumed by classical cytogenetics (Milhomem et al., 2008, 2012a,b).

Whole chromosome probes from GCA $2n = 42$ have been used for comparative genomic mapping (CGM) of the karyotype

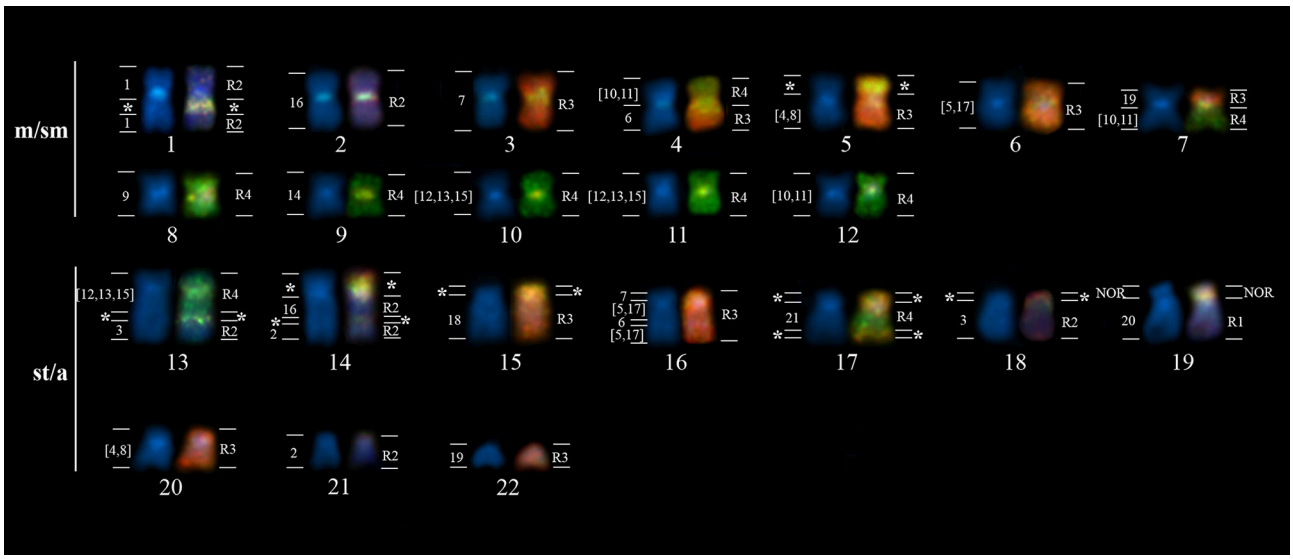


FIGURE 3 | Haploid karyotype of *G. arapaima* (GAR) arranged from mitotic chromosomes after dual-color hybridization with probes derived from Region 3 (R3, red) and Region 4 (R4, green) from the *Gymnotus carapo* (GCA) chromosome complement. Regions R1 and R2 were not subjected to FISH analysis and, therefore, the equivalent homeologous parts on GAR chromosomes are DAPI-stained (blue) only. For each of the 22 GAR chromosome pairs, the DAPI-only stained homolog is depicted on the left, while the dual-color FISH hybridization pattern is present on the right. The correspondence to *G. carapo* (GCA) homeologous chromosomes is indicated by chromosome pair numbers on the left side of the DAPI-stained GAR chromosomes, while the correspondence to the particular GCA regions (R1–4) is indicated on the right side of FISH-painted chromosomes. *Repetitive sequences.

of (i) cryptic species GCA $2n = 40$ (Nagamachi et al., 2010), (ii) GCP $2n = 34$ (Nagamachi et al., 2013) and, in the present work, iii) onto the karyotype of GAR $2n = 44$ (**Figure 4**). Similar to the observations in the two previously mapped species (Nagamachi et al., 2010, 2013), GAR also presents a highly reorganized karyotype (**Figures 3**, **4**, **Table 2**) in relation to GCA $2n = 42$ and also in relation to GCA $2n = 40$ and GCP $2n = 34$. From the 12 chromosome pairs of GCA $2n = 42$ that can be individually differentiated (GCA 1–3, 6, 7, 9, 14, 16, 18–21), GAR shows conserved synteny of six pairs (GCA 1, 9, 14, 18, 20, 21); five pairs (GCA 1, 9, 14, 20, 21) with the cryptic species GCA $2n = 40$ and only one pair with GCP (GCA 20). On the other hand, GCA $2n = 40$ shares with GCA $2n = 42$, eight pairs (GCA 1, 2, 6, 9, 14, 19, 20, 21) and with GCP, three pairs (GCA 6, 19, 20) (**Figure 4**). It is also worth noting that the probes representing GCA [4, 8] and [12, 13, 15] show two and three signals, respectively, in three species (GCA 42, GCP and GAR) indicating that these chromosomes may have retained their homology.

A comparative analysis of the WCP data described above shows that the karyotypes of both GCP and GAR are related to the karyotypes of GCA. GCP, although part of the *carapo* group (Milhomem et al., 2012a), has an uncertain position inside the phylogeny of the clade, while GCA and GAR are closely related. GCP and GAR do not share the same chromosome rearrangements (**Table 2**), meaning that these rearrangements must have occurred after their speciation. The results of the CGM suggest either a divergence prior to that of GAR or a recent divergence characterized by fast karyotype evolution and fixation of a high number of chromosomal rearrangements.

It is also clear that the karyotype of GAR is evolutionary closer to the GCA karyotype than to the GCP karyotype. However, GAR is located 2000 km away from the other species, while GCP and GCA ($2n = 42$) are 200 km apart (**Figure 5**). This might suggest that the karyotypes of GCA and GAR are more conserved while GCP changed over a shorter period of time. Another explanation for this huge differentiation of the GCP karyotype might lie in the fact that this species inhabits Rio Açaiteuazinho drainage from Northeast Para, which is not connected with the Amazon basin, while GCA and GAR are part of the same hydrographic basin, despite the long distance between them (**Figure 5**).

Freshwater fishes in general have a higher rate of chromosomal rearrangements than marine fishes due to the reduced flow with the natural barriers present in the freshwater environment compared to the open marine biome, with bigger populations and high potential for dispersion and higher gene flow, reducing the chance for karyotype changes to fixate in the population (Molina, 2007; Nirchio et al., 2014; Artoni et al., 2015). Lande (1977) theorizes that the rates of chromosomal rearrangement are proportional to selection and inversely proportional to the effective size of the population and Araya-Jaime et al. (2017) suggests that this could be considered a general model of chromosomal evolution within Gymnotiformes, since populations with little or no geneflow may facilitate the fixation of chromosomal rearrangements within a particular species in a shorter evolutionary time. This may be a contributory factor to speciation within the group and may also contribute to the higher number of rearrangements found. It is a valid reminder that the high

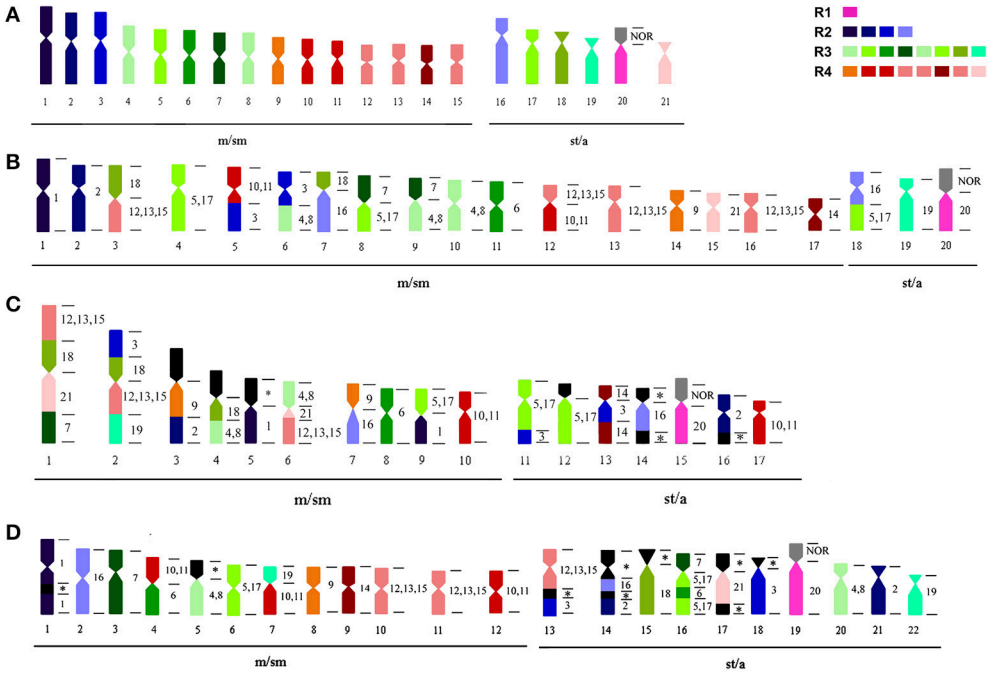


FIGURE 4 | Ideogram with the karyotypes of **(A)** *G. carapo* (2*n* = 42); **(B)** *G. carapo* (2*n* = 40); **(C)** *G. capanema*, and **(D)** *G. arapaima*. The numbers at the right side of chromosomes in **(B–D)** show the homology with the karyotype **(A)** of *G. carapo*. Each color in the karyotypes **(B–D)** represents the correspondent chromosome colored in **(A)**. Chromosomes groups [4, 8]; [5, 17]; [10, 11], and [12, 13, 15] share the same color within each group.

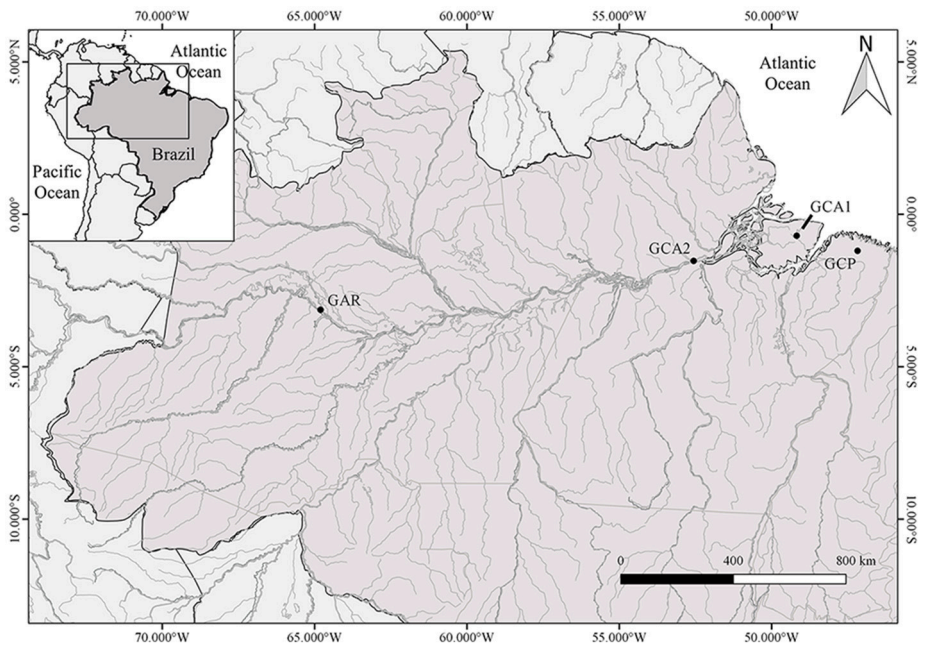


FIGURE 5 | A map of Northern Brazil showing the geographical distribution of the samples from the four species of *Gymnotus* analyzed by whole chromosome painting. *G. carapo* 2*n* = 42 (GCA1, Nagamachi et al., 2010, Santa Cruz do Arari, Marajo Island), *G. carapo* 2*n* = 40 (GCA2, Nagamachi et al., 2010, Almerim, Amazon river drainage), *G. capanema* (GCP, Nagamachi et al., 2013, Capanema, Rio Açaiteuazinho drainage) and *G. arapaima* (GAR, present study, Mamirauá Sustainable Development Reserve, Amazonas, Brazil).

number of rearrangements observed in the present study was possible through WCP, and groups with a more stable diploid number and karyotypic formula potentially could have fixed a higher number of rearrangements that did not cause major structural changes.

As Region 3 was labeled with a red fluorochrome and Region 4 with a green one, all yellow regions in **Figure 3** are the result of hybridization of both probes to the same region. Although R3 and R4 do not share the same chromosome pair, they share the same or highly similar repetitive DNA. The hybridization of both probes to the same regions of GAR chromosomes confirms that this sequence is also present in this species. Since repetitive sequences evolve quickly by concerted evolution with significant differences between species (Pons and Gillespie, 2004), the presence of the highly similar repetitive DNA sequence in different species clearly shows that these species diverged recently, without sufficient time to accumulate sequence differences. Despite the huge amount of rearrangement, the repetitive DNA sequence strongly suggests that these species diverged recently and also that the rearrangements responsible for the karyotypic differences are also recent.

Taken together, the sum of the results might explain the difficulty in finding synapomorphies among the species compared so far, since most of the rearrangements might have become fixed after the species became isolated. On the other hand, because the *G. carapo* clade is a derived one (Tagliacollo et al., 2016, **Figure 1**) and because up until today there are few species of *Gymnotus* studied by chromosome painting, we currently cannot conclusively resolve whether the homologous chromosomes present a symplesiomorphic or synapomorphic character. An example is the NOR bearing pair that maps to GCA 20 using rDNA probes in species of the *carapo* group, but this location is different in species outside this group (Milhomem et al., 2013), which suggests that it is a synapomorphy. This matter will be better understood once species outside the *carapo* group are mapped with all the GCA whole chromosome probes.

AUTHOR CONTRIBUTIONS

MM, JP, FS, PO, MF-S, and CN: gave substantial contributions to the conception of the work; the acquisition, analysis, and interpretation of data for the work; participated in the draft of the work or revised it critically for important intellectual content; gave final approval of the version to be published; and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

FUNDING

This research was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) through the Edital Universal (Proc. 475013/2012-3) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) through the Edital 047/2012 PRÓ-AMAZÔNIA: Biodiversidade e Sustentabilidade on a project coordinated by CYN; by Fundação Amazônia Paraense de Amparo à Pesquisa (FAPESPA) through the National Excellence on Research Program (PRONEX, TO 011/2008) and Banco Nacional de Desenvolvimento Econômico e Social – BNDES (Operação 2.318.698.0001) on a project coordinated by JP.

ACKNOWLEDGMENTS

This study is part of the Master Dissertation of MM who was a recipient of a CAPES Scholarship in Genetics and Molecular Biology, UFPA. CYN (308428/20013-7) and JP (308401/2013-1) are grateful to CNPq for Productivity Grants. The authors are grateful to members of the team of the cytogenetics laboratory UFPA for the fieldwork and chromosomal preparations. To MSc. Jorge Rissino, to MSc. Shirley Nascimento and Maria da Conceição for assistance in laboratory work. We also thank the Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) for the collection permit 020/2005 (Registration: 207419).

REFERENCES

Albert, J. S. (2001). *Species Diversity and Phylogenetic Systematics of American Knifefishes (Gymnotiformes, Teleostei).* Ann Arbor, MI: Miscellaneous Publications. University of Michigan, 190, 1–129.

Albert, J. S., and Crampton, W. G. R. (2003). Seven new species of the neotropical electric fish *Gymnotus* (Teleostei, Gymnotiformes) with a redescription of *G. carapo* (Linnaeus). *Zootaxa* 287, 1–54. doi: 10.11646/zootaxa.287.1.1

Albert, J. S., Crampton, W. G. R., Thorsen, D. H., and Lovejoy, N. R. (2005). Phylogenetic systematics and historical biogeography of the neotropical electric fish *Gymnotus* (Teleostei: Gymnotidae). *Syst. Biodivers*. 2, 375–417. doi: 10.1017/S1477200004001574

Albert, J. S., de Campos Fernandes-Matioli, F. M., and de Almeida-Toledo, L. F. (1999). New species of *Gymnotus* (Gymnotiformes, Teleostei) from southeastern Brazil: toward the deconstruction of *Gymnotus carapo*. *Copeia* 1999, 410–421. doi: 10.2307/1447486

Almeida, J. S., Migues, V. H., Diniz, D., and Affonso, P. R. A. (2015). A unique sex chromosome system in the knifefish *Gymnotus bahianus* with inferences about chromosomal evolution of Gymnotidae. *J. Hered.* 106, 177–183. doi: 10.1093/jhered/esu087

Almeida-Toledo, L. F., Daniel-Silva, M. F. Z., Moyses, C. B., and Foresti, F. (2007). "Chromosome variability in Gymnotiformes (Teleostei: Ostariophysi)," in *Fish Cytogenetics,* eds E. Pisano, C. Ozouf-Costaz, F. Foresti, and B. G. Kapoor (Boca Raton, FL: CRC Press), 16–39.

Almeida-Toledo, L. F., Foresti, F., and de Almeida-Toledo, S. (1981). Constitutive heterochromatin and nucleolus organizer region in the knifefish, *Apteronotus albifrons* (Pisces, Apteronotidae). *Cell Mol. Life Sci.* 37, 953–954. doi: 10.1007/BF01971773

Araya-Jaime, C., Mateussi, N. T. B., Utsunomia, R., Costa-Silva, G. J., Oliveira, C., and Foresti, F. (2017). ZZ/Z0: the new system of sex chromosomes in *Eigenmannia* aff. trilineata (Teleostei: Gymnotiformes: Sternopygidae) characterized by molecular cytogenetics and DNA barcoding. Zebrafish 14, 464–470. doi: 10.1089/zeb.2017.1422

Artoni, R. F., Castro, J. P., Jacobina, U. P., Lima-Filho, P. A., da Costa, G. W. W. F., and Molina, W. F. (2015). Inferring diversity and evolution in fish by means of integrative molecular cytogenetics, *Sci. World J.* 2015:365787. doi: 10.1155/2015/365787

Batista, J. A., Cardoso, A. L., Milhomem-Paixão, S. S. R., Ready, J. S., Pieczarka, J. C., and Nagamachi, C. Y. (2017). The karyotype of *Microsternarchus* aff. bilineatus: a first case of Y chromosome degeneration in Gymnotiformes. Zebrafish 14, 244–250. doi: 10.1089/zeb.2016.1383

Cabral-de-Mello, D. C., and Martins, C. (2010). "Breaking down the genome organization and karyotype differentiation through the epifluorescence microscope lens: insects and fish as models," in *Formatex Microscopy*, eds A. Méndez-Vilas and J. Díaz (Badajoz: Formatex Research Center), 658–669.

Claro, F. L. (2008). *Gymnotus Carapo and Gymnotus sylvius (Teleostei: Gymnotidae): Uma Abordagem Citogenético-Molecular*. Master's dissertation. São Paulo, SP: Universidade de São Paulo.

Craig, J., Crampton, W. G. R., and Albert, J. S. (2017). Revision of the polytypic electric fish *Gymnotus carapo* (Gymnotiformes, Teleostei), with descriptions of seven subspecies. *Zootaxa* 4318, 401–438. doi: 10.11646/zootaxa.4318.3.1

Crampton, W. G. R., Thorsen, D. H., and Albert, J. S. (2005). Three new species from a diverse, sympatric assemblage of the electric fish *Gymnotus* (Gymnotiformes, Gymnotidae) in the lowland Amazon basin, with notes on ecology. *Copeia* 2005, 82–99. doi: 10.1643/CI-03-242R2

da Silva, M. (2015). *Análise Biogeográfica do Gênero Gymnotus (Gymnotidae, Gymnotiformes), Por Meio de Marcadores Cariotípicos e Moleculares*. Dissertation. Manaus, AM: Instituto Nacional de Pesquisas da Amazônia.

da Silva, M., Barbosa, P., Artoni, R. F., and Feldberg, E. (2016). Evolutionary dynamics of 5S rDNA and recurrent association of transposable elements in electric fish of the family Gymnotidae (Gymnotiformes): the case of *Gymnotus mamiraua*. *Cytogenet. Genome Res.* 149, 297–303. doi: 10.1159/000449431

da Silva, M., Matoso, D. A., Artoni, R. F., and Feldberg, E. (2014). New approach data in electric fish (Teleostei: *Gymnotus*): sex chromosome evolution and repetitive DNA. *Zebrafish* 11, 528–535. doi: 10.1089/zeb.2013.0966

da Silva, M., Matoso, D. A., Vicari, M. R., de Almeida, M. C., Margarido, V. P., and Artoni, R. F. (2011). Physical mapping of 5S rDNA in two species of knifefishes: *Gymnotus* pantanal and *Gymnotus* paraguensis (Gymnotiformes). *Cytogenet. Genome Res.* 134, 303–307. doi: 10.1159/000328998

da Silva, P. C., Nagamachi, C. Y., Silva, D. S., Milhomem, S. S. R., Cardoso, A. L., de Oliveira, J. A., et al. (2013). Karyotypic similarities between two species of Rhamphichthys (*Rhamphichthyidae*, Gymnotiformes) from the Amazon basin. *Comp. Cytogenet.* 7:279. doi: 10.3897/compcytogen.v7i4.4366

Fernandes, C. A., Paiz, L. M., Baumgärtner, L., Margarido, V. P., and Vieira, M. M. D. R. (2017). Comparative cytogenetics of the black ghost knifefish (Gymnotiformes: Apteronotidae): evidence of chromosomal fusion and pericentric inversions in karyotypes of two Apteronotus species. *Zebrafish* 14, 471–476. doi: 10.1089/zeb.2017.1432

Fernandes, F. M., Albert, J. S., Daniel-Silva, M. D., Lopes, C. E., Crampton, W. G., and Almeida-Toledo, L. F. (2005). A new *Gymnotus* (Teleostei: Gymnotiformes: Gymnotidae) from the Pantanal Matogrossense of Brazil and adjacent drainages: continued documentation of a cryptic fauna. *Zootaxa* 933, 1–14. doi: 10.11646/zootaxa.933.1.1

Fernandes-Matioli, F. M. C., Marchetto, M. C. N., Almeida-Toledo, L. F., and Toledo-Filho, S. A. (1998). High intraspecific karyological conservation in four species of *Gymnotus* (Pisces: Gymnotiformes) from southeastern brazilian basins. *Caryologia* 51, 221–234. doi: 10.1080/00087114.1998.10797414

Ferraris, C. J. Jr., de Santana, C. D., and Vari, R. P. (2017). Checklist of Gymnotiformes (Osteichthyes: Ostariophysi) and catalogue of primary types. *Neotrop. Ichthyol.* 15:e160067. doi: 10.1590/1982-0224-20160067

Foresti, F., de Almeida Toledo, L. F., and de Almeida Toledo, F. S. (1984). Chromosome studies in *Gymnotus carapo* and *Gymnotus* sp. (Pisces, Gymnotidae). *Caryologia* 37, 141–146. doi: 10.1080/00087114.1984.10797693

Gornung, E. (2013). Twenty years of physical mapping of major ribosomal RNA genes across the teleosts, a review of research. *Cytogenet. Genome Res.* 141, 90–102. doi: 10.1159/000354832

Henning, F., Moysés, C. B., Calcagnotto, D., Meyer, A., and de Almeida-Toledo, L. F. (2011). Independent fusions and recent origins of sex chromosomes in the evolution and diversification of glass knife fishes (Eigenmannia). *Heredity* 106, 391–400. doi: 10.1038/hdy.2010.82

Henning, F., Trifonov, V., Ferguson-Smith, M. A., and de Almeida-Toledo, L. F. (2008). Non-homologous sex chromosomes in two species of the genus Eigenmannia (Teleostei: Gymnotiformes). *Cytogenet. Genome Res.* 121:55. doi: 10.1159/000124382

Howell, W. M. (1972). Somatic chromosomes of the black ghost knifefish, *Apteronotus albifrons* (Pisces: Apteronotidae). *Copeia* 1972, 191–193. doi: 10.2307/1442803

Jesus, I. S., Ferreira, M., Garcia, C., Ribeiro, L. B., Alves-Gomes, J. A., and Feldberg, E. (2016). First cytogenetic description of Microsternarchus bilineatus (Gymnotiformes: Hypopomidae) from Negro river (Brazilian Amazon). *Zebrafish* 13, 571–577. doi: 10.1089/zeb.2016.1281

Knytl, M., Kalous, L., Symonová, R., Rylková, K., and Ráb, P. (2013). Chromosome studies of European cyprinid fishes: cross-species painting reveals natural allotetraploid origin of a Carassius female with 206 chromosomes. *Cytogenet. Genome Res.* 139, 276–283. doi: 10.1159/000350689

Lacerda, M. C. V., and Maistro, E. L. (2007). Cytogenetic analysis of three sympatric *Gymnotus* species (Teleostei: Gymnotidae) from the Fundo stream, MG, Brazil. *Cytologia* 72, 89–93. doi: 10.1508/cytologia.72.89

Lande, R. (1977). On comparing coefficients of variation. *Syst. Zool.* 26, 214–217. doi: 10.2307/2412845

Levan, A., Fredga, K., and Sandberg, A. A. (1964). Nomenclature for centromeric position on chromosomes. *Hereditas* 52, 201–220. doi: 10.1111/j.1601-5223.1964.tb01953.x

Lovejoy, N. R., Lester, K., Crampton, W. G. R., Marques, F. P., and Albert, J. S. (2010). Phylogeny, biogeography, and electric signal evolution of neotropical knifefishes of the genus *Gymnotus* (Osteichthyes: Gymnotidae). *Mol. Phylogenet. Evol.* 54, 278–290. doi: 10.1016/j.ympev.2009.09.017

Machado, M. A., Cardoso, A. L., Milhomem-Paixão, S. S. R., Pieczarka, J. C., and Nagamachi, C. Y. (2017). *Gymnotus coatesi* (Gymnotiformes): a case of colocation of multiple sites of 18S rDNA with telomeric sequences. *Zebrafish* 14, 459–463. doi: 10.1089/zeb.2017.1435

Margarido, V. P., Bellafronte, E., and Moreira-Filho, O. (2007). Cytogenetic analysis of three sympatric *Gymnotus* (Gymnotiformes,Gymnotidae) species verifies invasive species in the Upper Parana river basin, Brazil. *J. Fish Biol.* 70, 155–164. doi: 10.1111/j.1095-8649.2007.01365.x

Martins, C., Cabral-de-Melo, D. C., Targino, G. V., Mazzuchelli, J., and Oliveira, S. G. (2011). "Cytogenetic mapping and contribution to the knowledge of animal genomes," in *Advances in Genetics Research,* Vol. 4, ed K. V. Urbano (New York, NY: Nova Science Publisher), 1–82.

Mendes, V. P., Portela-Castro, A. L. B., and Júlio-Júnior, H. F. (2012). First record of supernumerary (B) chromosomes in electric fish (Gymnotiformes) and the karyotype structure of three species of the same order from the upper Paraná River basin. *Comp. Cytogenet.* 6, 1–16. doi: 10.3897/compcytogen.v6i1.1752

Milhomem, S. S. R., Crampton, W. G. R., Pieczarka, J. C., Shetka, G. H., Silva, D. S., and Nagamachi, C. Y. (2012a). *Gymnotus capanema*, a new species of electric knife fish (Gymnotiformes, Gymnotidae) from eastern Amazonia, with comments on an unusual karyotype. *J. Fish Biol.* 80, 802–815. doi: 10.1111/j.1095-8649.2012.03219.x

Milhomem, S. S. R., Crampton, W. G. R., Pieczarka, J. C., Silva, D. S., Cardoso, A. L., da Silva, P. C., et al. (2012b). Chromosomal and electric signal diversity in three sympatric electric knifefish species (Gymnotus, Gymnotidae) from the central Amazon floodplain. *Rev. Fish Biol. Fisher.* 22, 485–497. doi: 10.1007/s11160-011-9239-1

Milhomem, S. S. R., Pieczarka, J. C., Crampton, W. G. R., Silva, D. S., Souza, A. C. P., Carvalho, J. R., et al. (2008). Chromosomal evidence for a putative cryptic species in the *Gymnotus carapo* species-complex (Gymnotiformes, Gymnotidae). *BMC Genet.* 9:75. doi: 10.1186/1471-2156-9-75

Milhomem, S. S. R., Pieczarka, J. C., Crampton, W. G. R., Souza, A. C. P., Carvalho Jr, J. R., and Nagamachi, C. Y. (2007). Differences in karyotype between two sympatric species of *Gymnotus* (Gymnotiformes: Gymnotidae) from the eastern amazon of Brazil. *Zootaxa* 1397, 55–62. doi: 10.11646/zootaxa.1397.7

Milhomem, S. S. R., Scacchetti, P. C., Pieczarka, J. C., Ferguson-Smith, M. A., Pansonato-Alves, J. C., O'Brien, P. C. M., et al. (2013). Are NORs always located on homeologous chromosomes? A FISH investigation with rDNA and whole chromosome probes in Gymnotus fishes (Gymnotiformes). *PLoS ONE* 8:e55608. doi: 10.1371/journal.pone.0055608

Molina, W. F. (2007). "Chromosome changes and stasis in marine fish groups," in *Fish Cytogenetic*, eds E. Pisano, C. Ozouf-Costaz, F. Foresti, and B. G. Kapoor (Boca Raton, FL: CRC Press), 69–110.

Nagamachi, C. Y., Pieczarka, J. C., Milhomem, S. S. R., Batista, J. A., O'Brien, P. C. M., and Ferguson-Smith, M. A. (2013). Chromosome painting reveals multiple rearrangements between *Gymnotus capanema* and *Gymnotus*

carapo (Gymnotidae, Gymnotiformes). *Cytogenet. Genome Res.* 141, 163–168. doi: 10.1159/000354988

Nagamachi, C. Y., Pieczarka, J. C., Milhomem, S. S. R., O'Brien, P. C. M., Souza, A. C. P., and Ferguson-Smith, M. A. (2010). Multiple rearrangements in cryptic species of electric knifefish, *Gymnotus carapo* (Gymnotidae, Gymnotiformes) revealed by chromosome painting. *BMC Genet.* 11:1. doi: 10.1186/1471-2156-11-28

Nirchio, M., Rossi, A. R., Foresti, F., and Oliveira, C. (2014). Chromosome evolution in fishes: a new challenging proposal from Neotropical species. *Neotrop. Ichthyol.* 12, 761–770. doi: 10.1590/1982-0224-20130008

Pons, J., and Gillespie, R. G. (2004). Evolution of satellite DNAs in a radiation of endemic hawaiian spiders: does concerted evolution of highly repetitive sequences reflect evolutionary history? *J. Mol. Evol.* 59, 632–641. doi: 10.1007/s00239-004-2655-2

Scacchetti, P., Pansonato-Alves, J., Utsunomia, R., Oliveira, C., and Foresti, F. (2011). Karyotypic diversity in four species of the genus *Gymnotus* Linnaeus, 1758 (Teleostei, Gymnotiformes, Gymnotidae): physical mapping of ribosomal genes and telomeric sequences. *Comp. Cytogenet.* 5:223. doi: 10.3897/compcytogen.v5i3.1375

Silva, D. S., Milhomem, S. S. R., Pieczarka, J. C., and Nagamachi, C. Y. (2009). Cytogenetic studies in Eigenmannia virescens (Sternopygidae, Gymnotiformes) and new inferences on the origin of sex chromosomes in the Eigenmannia genus. *BMC Genet.* 10:74. doi: 10.1186/1471-2156-10-74

Sousa, T. P., Marques, D. K. S., Vitorino, C. D. A., Faria, K. D. C., Braga, G. D. S. F., Ferreira, D. C., et al. (2017). Cytogenetic and molecular data Support the occurrence of three *Gymnotus* species (Gymnotiformes: Gymnotidae) used as live bait in Corumbá, Brazil: implications for conservation and management of professional fishing. *Zebrafish* 14, 177–186. doi: 10.1089/zeb.2016.1356

Suárez, P., Barroso, I. C. G. P., Silva, D. S., Milhomem, S. S. R., Cabral-de-Mello, D. C., Martins, C., et al. (2017). Highest diploid number among Gymnotiformes: first cytogenetic insights into Rhabdolichops (Sternopygidae). *Zebrafish* 14, 272–279. doi: 10.1089/zeb.2016.1405

Tagliacollo, V. A., Bernt, M. J., Craig, J. M., Oliveira, C., and Albert, J. S. (2016). Model-based total evidence phylogeny of neotropical electric knifefishes (Teleostei, Gymnotiformes). *Mol. Phylogenet. Evol.* 95, 20–33. doi: 10.1016/j.ympev.2015.11.007

Takagui, F. H., da Rosa, R., Shibatta, O. A., and Giuliano-Caetano, L. (2017). Chromosomal similarity between two species of *Apteronotus albifrons* complex (Apteronotidae–Gymnotiformes) implications in cytotaxonomy and karyotypic evolution. *Caryologia* 70, 1–4. doi: 10.1080/00087114.2017.1306385

Utsunomia, R., Scacchetti, P. C., Pansonato-Alves, J. C., Oliveira, C., and Foresti, F. (2014). Comparative chromosome mapping of U2 snRNA and 5S rRNA genes in *Gymnotus* species (Gymnotiformes, Gymnotidae): evolutionary dynamics and sex chromosome linkage in *G. pantanal. Cytogenet. Genome Res.* 142, 286–292. doi: 10.1159/000362258

Vicari, M. R., de Mello Pistune, H. F., Castro, J. P., de Almeida, M. C., Bertollo, L. A. C., Moreira-Filho, O., et al. (2011). New insights on the origin of B chromosomes in Astyanax scabripinnis obtained by chromosome painting and FISH. *Genetica* 139, 1073–1081. doi: 10.1007/s10709-011-9611-z

Yang, F., Carter, N. P., Shi, L., and Ferguson-Smith, M. A. (1995). A comparative study of karyotypes of muntjacs by chromosome painting. *Chromosoma* 103, 642–652. doi: 10.1007/BF00357691

Yang, F., and Graphodatsky, A. S. (2017). "Animal probes and ZOO-FISH," in *Fluorescence in situ Hybridization (FISH) Application Guide*, ed T. Liehr (Berlin; Heidelberg: Springer-Verlag), 395–415.

Yano, C. F., Bertollo, L. A. C., and Cioffi, M. D. B. (2017). "Fish-FISH: molecular cytogenetics in fish species," in *Fluorescence In Situ Hybridization (FISH) Application Guide*, ed T. Liehr (Berlin; Heidelberg: Springer-Verlag), 429–443.

Yi, M. S., Li, Y. Q., Liu, J. D., Zhou, L., Yu, Q. X., and Gui, J. F. (2003). Molecular cytogenetic detection of paternal chromosome fragments in allogynogenetic gibel carp, *Carassius auratus* gibelio Bloch. *Chromosome Res.* 11, 665–671. doi: 10.1023/A:1025985625706

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 02 February 2018
doi: 10.3389/fgene.2018.00013

Genetic Diversity of the Endangered Neotropical Cichlid Fish (*Gymnogeophagus setequedas*) in Brazil

Lenice Souza-Shibatta[1], Thais Kotelok-Diniz[1], Dhiego G. Ferreira[2], Oscar A. Shibatta[3], Silvia H. Sofia[1], Lucileine de Assumpção[4], Suelen F. R. Pini[4], Sergio Makrakis[4] and Maristela C. Makrakis[4]*

[1] Laboratório de Genética e Ecologia Animal, Departamento de Biologia Geral, Universidade Estadual de Londrina, Londrina, Brazil, [2] Laboratório de Genética e Conservação, Universidade Estadual do Norte do Paraná, Cornélio Procópio, Brazil, [3] Museu de Zoologia, Departamento de Biologia Animal e Vegetal, Universidade Estadual de Londrina, Londrina, Brazil, [4] Grupo de Pesquisa em Tecnologia em Ecohidráulica e Conservação de Recursos Pesqueiros e Hídricos – GETECH, Universidade Estadual do Oeste do Paraná, Toledo, Brazil

Gymnogeophagus setequedas is a rare and rheophilic species of tribe Geophagini, considered endangered in Brazilian red lists. Its previously known geographical distribution range was the Paraná River basin, in Paraguay, and a tributary of the Itaipu Reservoir in Brazil. Since its description no specimens have been collected in the original known distribution area. However, recent records of *G. setequedas* in the lower Iguaçu River, in a region considered highly endemic for the ichthyofauna, extended the known geographical distribution and may represent one of the last remnants of the species. The aim of this study was to estimate the genetic diversity and population structure of *G. setequedas*, using microsatellite markers and mitochondrial haplotypes, in order to test the hypothesis of low genetic diversity in this restricted population. Muscular tissue samples of 86 specimens were obtained from nine locations in the Lower Iguaçu River basin, between upstream of the Iguaçu Falls and downstream of the Salto Caxias Reservoir. Seven microsatellites loci were examined and a total of 120 different alleles were obtained. The number of alleles per locus (N_A) was 17.429, effective alleles (N_E) 6.644, expected heterozygosity (H_E) 0.675, observed (H_O) heterozygosity 0.592, and inbreeding coefficient (F_{IS}) 0.128. Twelve haplotypes in the D-Loop region were revealed, with values of h (0.7642) and π (0.00729), suggesting a large and stable population with a long evolutionary history. Thus, both molecular markers revealed high levels of genetic diversity and indicated the occurrence of a single *G. setequedas* population distributed along a stretch of approximately 200 km. The pattern of mismatch distribution was multimodal, which is usually ascribed to populations in demographic equilibrium. Nevertheless, the construction of a new hydroelectric power plant, already underway between the Salto Caxias Reservoir and Iguaçu Falls, could fragment this population, causing loss of genetic diversity and population decline, and for this reason it is necessary to maintain the Iguaçu River tributaries and downstream area from the Lower Iguaçu Reservoir free of additional dams, to guarantee the survival of this species.

Keywords: conservation, freshwater, Lower Iguaçu River, *Gymnogeophagus*, endangered species

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Paulo Affonso, Universidade Estadual do Sudoeste da Bahia, Brazil
Charles Masembe, Makerere University, Uganda

****Correspondence:***
Lenice Souza-Shibatta
lenicesouza@hotmail.com

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

Received: *31 October 2017*
Accepted: *10 January 2018*
Published: *02 February 2018*

Citation:
Souza-Shibatta L, Kotelok-Diniz T, Ferreira DG, Shibatta OA, Sofia SH, de Assumpção L, Pini SFR, Makrakis S and Makrakis MC (2018) Genetic Diversity of the Endangered Neotropical Cichlid Fish (Gymnogeophagus setequedas) in Brazil. Front. Genet. 9:13. doi: 10.3389/fgene.2018.00013

INTRODUCTION

The largest biodiversity asset in the world is located in Brazil (ICMBio, 2017). However, more than 1,170 species in Brazil have been classified as threatened with extinction. Unfortunately, of these, more than 26% are Actinopterygii fish found in freshwaters (ICMBio, 2017), including *Gymnogeophagus setequedas* Reis et al. (1992), the only one of the 17 species of the genus described so far considered as threatened (Abilhoa and Duboc, 2004; Pavanelli and Reis, 2008). This has led the Brazilian Environmental Ministry to decree the species with the status of Endangered species (EN) (decree #445, International Union for Conservation of Nature [IUCN], 2014).

Gymnogeophagus setequedas was described based on specimens collected in tributaries of the Paraná River in Paraguay and Brazil, near the Sete Quedas region, an area currently submerged due to construction of the Itaipu Hydroelectric Power Plant (Reis et al., 1992; Pavanelli and Reis, 2008). However, since its description no specimens have been collected in the original known distribution area (Abilhoa and Duboc, 2004; Agostinho et al., 2004; Pavanelli and Reis, 2008). Nevertheless, 15 specimens of *G. setequedas* were recently collected in the Lower Iguaçu River, both up and downstream from Iguaçu Falls in the Iguaçu National Park (Paiz et al., 2017). According to the authors, finding this species in that region was quite unexpected as the Iguaçu waterfalls have lead to effective geographic isolation of the Ichthyofauna of the Iguaçu River (Zawadzki et al., 1999), providing an accentuated degree of endemicity, estimated between 51 and 71% (Abell et al., 2008). In addition, due to its high ecological importance, the Iguaçu River basin is considered an ecoregion, separated from the rest of the Paraná River Basin (Abell et al., 2008).

The Iguaçu River basin covers an area of approximately 72,000 km^2, representing part of the landscape of the three Paraná plateaus, subdivided into three regions: Upper Iguaçu (1st plateau, Curitiba region), Middle Iguaçu (2nd plateau, Ponta Grossa region), and Lower Iguaçu (3rd plateau, Guarapuava region) (Maack, 2001). The portion of the 3rd plateau that includes the Lower Iguaçu is characterized by the presence of numerous waterfalls, such as Salto Grande (13 m), Salto Santiago (40 m), Salto Osório (30 m), and Iguaçu Falls (Maack, 1981). The region is very attractive for hydroelectric use due to its high gradient, and, thus, the original rapids and waterfalls, have been transformed into a sequence of reservoirs that flooded approximately 656 km^2, remarkably altering the landscape (Júlio Júnior et al., 1997).

It was believed that *G. setequedas* preferred lentic environments, as in the other species of the genus (Pavanelli and Reis, 2008). However, it seems that this species behaves differently from its congeners, preferring fast waters. This fact was corroborated by its recent capture in the Lower Iguaçu River, in stretches without containment and with fast waters (Paiz et al., 2017). In addition, this species disappeared after construction of the Itaipu reservoir, being collected only twice, suggesting its dependence on lotic environments (Agostinho et al., 2004). Pavanelli and Reis (2008) consider that this species no longer occurs in the Itaipu reservoir, possibly because it did not succeed in colonizing the environment formed after construction of the reservoir. In Paraguay this species is also considered an EN (Liotta, 2010), for the same reasons as in Brazil.

According to Wu et al. (2015), understanding the diversity and genetic structure of endangered species is fundamental to engage effective environmental conservation and management actions. Genetic diversity is essential if populations are evolving in response to environmental changes. For instance, due to the effects of anthropogenic disturbances, a small and isolated population is more likely to lose genetic diversity, and consequently present population decline, than a huge population with high genetic diversity (Frankham et al., 2010; Allendorf et al., 2012).

The genetic diversity status of species is the starting point for systematic planning of actions that should be taken to ensure the survival of species and reduce their risk of extinction. No works are known which focus on the biological (Pavanelli and Reis, 2008) or genetic diversity of *G. setequedas*. In addition, the diploid number has only recently been presented (Paiz et al., 2017). Thus, the aim of this study was to estimate the genetic diversity and population structure of *G. setequedas* along its recently known area of occurrence, using microsatellite markers and mitochondrial haplotypes (D-loop), thus presenting the first data of a population study of this species threatened with extinction.

MATERIALS AND METHODS

Study Area and Sampling

Our study area comprises a stretch of the Lower Iguaçu River basin, between upstream Iguaçu Falls and downstream Salto Caxias Reservoir (**Figure 1**).

Samples of 86 *G. setequedas* were collected at nine different points, some of them located in the Iguaçu National Park (PNI): two points in the main channel of the Iguaçu River (IGU 1 and IGU 2 – PNI), near the Iguaçu falls, and seven tributaries of the Iguaçu River (STO-Santo Antônio, SIL-Silva – PNI, Jardim, FLO-Floriano – PNI, GON-Gonçalves Dias – PNI, CAP-Capanema, AND-Andrada, and COT-Cotegipe) (**Figure 1** and **Table 1**). The samples were collected in 2012 (November), 2013 (November and December), and 2014 (January, February, March, April, July, August, September, November, and December). The specimens were captured using nets of different mesh sizes and electric fishing. Samples of muscle and rayed fins were taken from the fish, stored in microtubes containing 100% ethanol and kept at −20°C. Specimens were fixed in 10% formalin and preserved in 70% ethanol and deposited in the fish collection of the Zoology Museum at the Universidade Estadual de Londrina under catalog number: MZUEL 16332, 16353, 16354, 17094–17096.

DNA Extraction and Quantification

Total DNA was extracted from muscle or rayed fins preserved in 95% EtOH following the phenol/chloroform protocol of Almeida et al. (2001). NanoDrop™ 1000 was used for determination of DNA concentrations and samples were diluted in ultrapure water,

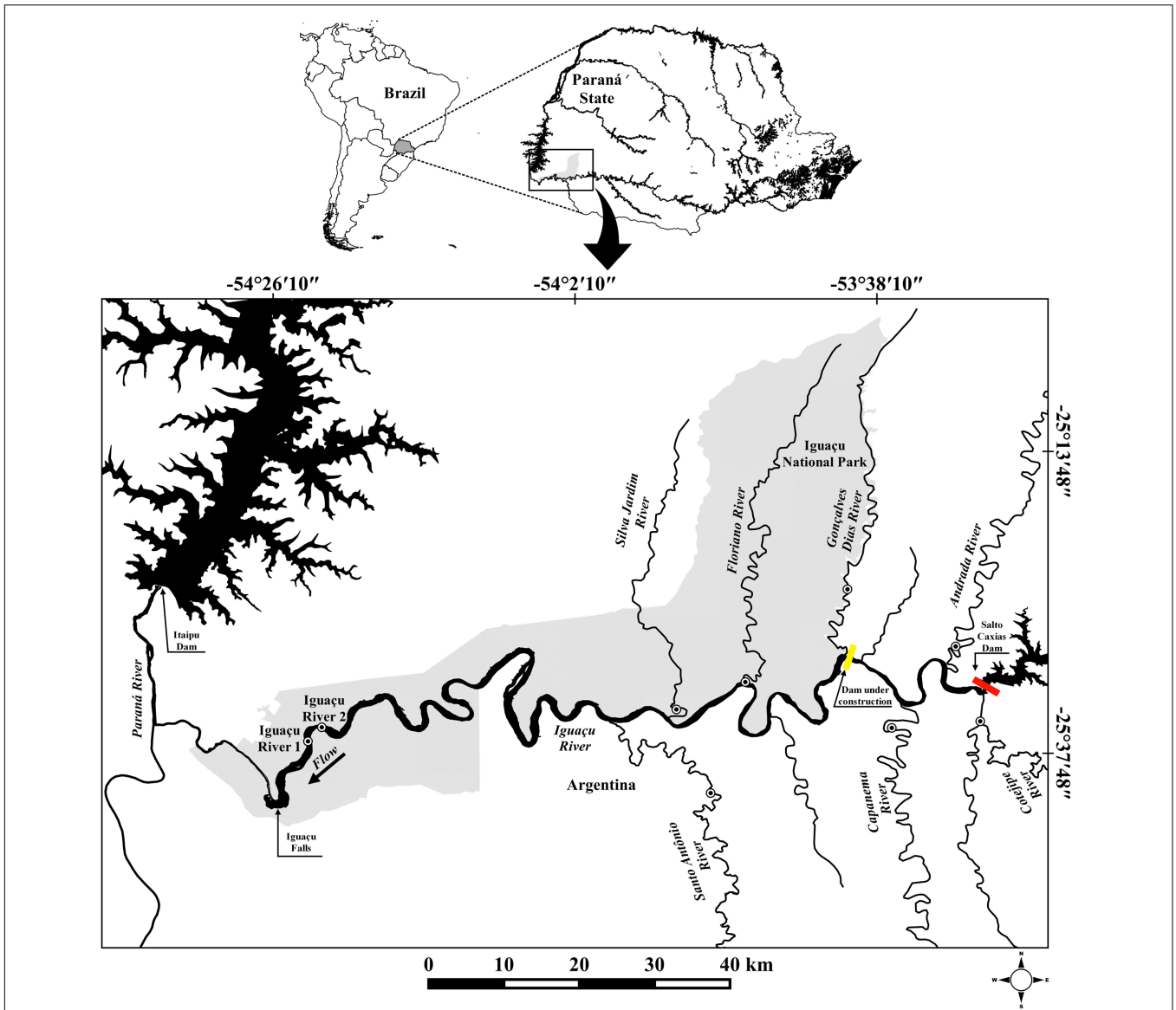


FIGURE 1 | Map of southern South America showing collection points (dotted circle) of *Gymnogeophagus setequedas*. The yellow bar represents the Baixo Iguaçu Hydroelectric Power Plant under construction, and red bar indicates Salto Caxias Hydroelectric Power Plant.

10 ng/μL for microsatellite markers and 5 ng/μL for mtDNA D-loop markers.

Microsatellite Amplification and Genotyping

Cross-amplification tests were conducted from seven loci described for *Geophagus brasiliensis* (Gbra6, Gbra16, Gbra17, Gbra62, Gbra63, Gbra80, and Gbra96) (Ferreira et al., 2013). Reagent concentrations and PCR conditions were performed according to Ferreira et al. (2015), using the modifications proposed by Schuelke (2000). PCR thermal conditions were conducted as follows: initial denaturation step at 94°C for 4 min, followed by 35 cycles of denaturation at 94°C for 40 s. The annealing temperatures of successful cros-amplifications loci were 48°C (Gbra16, Gbra62, Gbra63, Gbra80), 54°C (Gbra06, Gbra17, Gbra70), or 60°C (Gbra96) for 1 min, extension at 72°C for 1 min, followed by a final extension at 72°C for 30 min. PCR products were analyzed on an ABI PRISM 3500-XL automated sequencer (Applied Biosystems) using GeneScan 600 Liz (Applied Biosystems) as a molecular weight marker.

mtDNA (D-Loop) Marker

Part of the control region (D-loop) of the *G. setequedas* mitochondrial DNA was amplified using PCR. The primers used were L 5′-AGAGCGTCGGTCTTGTAAACC-3′ (Cronin et al., 1993) and H 5′-CTGAAGTAGGAACCAGATG-3′ (Meyer et al., 1990). PCR reactions were performed in a 25 μL final volume containing 1X GoTaq Master Mix (Promega), 1 μM

TABLE 1 | Sampling sites for *Gymnogeophagus setequedas*, including sample sizes per site (N) and location.

Collection point	*N*	Specific collecting location	Geographic coordinates
(1) – (IGU 1)	02	Iguaçu River 1	25°35′43.79″S 54°22′17.51″W
(2) – (IGU 2)	50	Iguaçu River 2	25°36′50.77″S 54°23′23.42″W
(3) – (STO)	25	Santo Antônio River	25°40′57.60″S 53°51′22.91″W
(4) – (SIL)	01	Silva Jardim River	25°34′18.64″S 53°54′06.13″W
(5) – (FLO)	01	Floriano River	25°32′08.12″S 53°48′37.46″W
(6) – (GON)	02	Gonçalves Dias River	25°24′45.53″S 53°40′30.30″W
(7) – (CAP)	01	Capanema River	25°35′46.76″S 53°36′58.98″W
(8) – (AND)	03	Andrada River	25°29′18.17″S 53°31′54.16″W
(9) – (COT)	01	Cotegipe River	25°35′15.39″S 53°29′57.11″W

of each primer, 15 ng DNA, and ultrapure water to volume. The thermal profile included an initial denaturation at 94°C for 4 min, followed by 41 cycles at 94°C for 15 s, annealing at 56°C for 30 s, and extension at 72°C for 2 min, with a final extension at 72°C for 10 min. The PCR products were purified using ExoSAP IT (Prodimol Biotecnologia S.A, Belo Horizonte, Minas Gerais, Brazil). The Big Dye Terminator v 3.1 kits (Applied Biosystems) and ABI-PRISM 3500 XL automated sequencer (Applied Biosystems) were used for sequence analysis. Multiple alignment analysis was carried out using the ClustalW application (Thompson et al., 1994) in BioEdit 7.1.3.0 (Hall, 1999). NCBI's BLAST search (Basic Local Alignment Search Tool, Altschul et al., 1990) was used to confirm the origin of the fragment. To search for possible tRNA, was used an online version of the tRNAscan-SE (Lowe and Eddy, 1997), available at http://lowelab.ucsc.edu/tRNAscan-SE. Sequences of the 12 different haplotypes were deposited in GenBank (MG581478 to MG581489).

Genetic Analyses (Microsatellites)

Population Structure

The first step for genetic analyses was to define the number of existing populations. For this we used population analyzes based on Bayesian approaches that mainly include "attribution methods." These methods calculate the probability of the different genotypes being observed in each population and assign the individuals to the populations according to the possibilities of the genotypes belonging to them, without any a priori inference. Thus, such analyzes allow to infer which population an individual belongs to, regardless of their collection site (Beaumont and Rannala, 2004).

In order to evaluate the relationship between samples, we conducted a Bayesian cluster analysis of the population by using STRUCTURE v.2.3.3 (Pritchard et al., 2000) program. The number of populations (K) was estimated by using the admixture model and correlated allele frequencies among populations, with K ranging from 1 to 10 (K = 1–10) (Evanno et al., 2005). A total of 20 independent runs of 100,000 Markov Chain Monte Carlo (MCMC) iterations discarded as burn-in, followed by 1,000,000 MCMC iterations were used for each value of K. The best-fit number of groupings was evaluated using K, ln Pr (X/K) (Pritchard et al., 2000) and ΔK *ad hoc* statistics (Evanno et al., 2005) by Structure Harvester v.0.6.7 (Earl and VonHoldt, 2012). Graphs representing the membership coefficient of each sampled individual were plotted using Distruct 1.1 (Rosenberg, 2004). Genetic differentiation estimates were assessed from pairwise Φ_{ST} values obtained in ARLEQUIN v.3.5.1.3 (Excoffier and Lischer, 2010). Significant estimates were based on 10,000 permutations. Subsequently, P-values corresponding to alpha = 0.05 were adjusted after Holm-Bonferroni correction for multiple tests (Holm, 1979).

Genetic Diversity

Number of alleles per locus (N_A), effective number of alleles (N_E), expected and observed heterozygosity (H_O, H_E) were obtained with POPGEN v. 1.31 (Yeh et al., 2000) software. Inbreeding coefficient (F_{IS}) was obtained with Fstat v2.9.3 program (Goudet, 2001). Deviation from Hardy-Weinberg equilibrium (HWE) and the linkage disequilibrium between pairs of loci with significance (P-value), later adjusted by the Bonferroni sequential correction (Rice, 1989) were tested with the GENEPOP v.1.2 (Raymond and Rousset, 1995). MICRO-CHECKER 2.2.1 (Van Oosterhout et al., 2004) software was used to test for the possible presence of null alleles or other genotyping errors such as allelic dropout and reading errors due to stutter peaks.

Gene Flow

The contemporary migration rates over a few previous generations and the direction of migration among the samples studied, was estimated by using the BayesAss v 3.0.3 program (Wilson and Rannala, 2003), at 95% confidence intervals. Ten runs were analyzed using different random starting seed numbers, with 3,000,000 MCMC iterations, including 999,999 discarded burn-in iterations. After the burn-in, every 2000th iteration was sampled. The delta values (maximum amount by which parameter values are allowed to change between iterations) were 0.15 for allele frequencies, 0.025 and 0.05 for migration rate, and 0.15 for inbreeding value.

Demographic Analyses

Recent population bottleneck signs were evaluated on microsatellite data using Bottleneck v.1.2.02 program (Piry et al., 1999), considering deviations from the mutation-drift equilibrium. Three tests were used, including two tests to indicate bottlenecks in the presence of significant excess heterozygosity: "Sign test" (Cornuet and Luikart, 1996) and the "Wilcoxon sign-rank test" (Luikart and Cornuet, 1998), both based on the Infinite Alleles Model (IAM), Stepwise Mutation Model (SMM), and Two-Phase Model (TPM – with 90% SMM and 10% IAM), with a P-value < 0.05. The third test was the "Mode shift test" that indicates bottlenecks resulting from alterations in allele frequency distributions (Luikart et al., 1998).

Genetic Analyses (mtDNA)

Genetic Diversity

Haplotype number, haplotype diversity (h) and nucleotide diversity (π) were obtained from DnaSP v.5 program (Librado and Rozas, 2009). Arlequin v.3.5.1.3 program (Excoffier and Lischer, 2010) was used to conduct the selective neutrality tests based on the infinite sites model of Tajima (D) (Tajima, 1989) and Fu (Fs) (Fu, 1997). The Network v.4.6.1.1 program (Fluxus Technology Ltd.[1]) was used to construct haplotype networks from mtDNA data based on the median-joining algorithm (Bandelt et al., 1999).

RESULTS

Population Structure

Bayesian clustering analysis (Structure) applied to microsatellite data indicated that the most probable K ($K+$ cluster number) was $K = 1$, from ln Pr(X/K). The graphic representation of ΔK showed that there were no well-defined groups, the ancestral values were distributed homogeneously among individuals and samples, indicating the occurrence of a single *G. setequedas* population (**Figure 2A**). Molecular Variance Analysis (AMOVA), conducted for all samples showed very little variation among them ($\Phi_{ST} = 0.025$, $P > 0.05$). Following this result, all other analyzes were based on a single population.

[1]http://www.fluxus-engineering.com

Genetic Diversity

In the entire sample, a total of 120 different alleles were obtained from seven microsatellite loci. The number of alleles per locus (N_A) was 17.429, effective alleles (N_E) 6.644, expected heterozygosity (H_E) 0.675, observed (H_O) heterozygosity 0.592, and inbreeding coefficient (F_{IS}) 0.128.

After applying the Bonferroni sequential correction, there were no significant deviations ($P < 0.05$) in the Hardy–Weinberg equilibrium (HWE) at the majority of microsatellite loci, only the locus *Gbra*96 showed significant deviation. This correction was also applied to the linkage disequilibrium (LD) tests, and a significant value was found only in locus *Gbra*96. The Micro-Checker program found no null alleles among the samples.

From the amplification and sequencing of mtDNA of 82 *G. setequedas* individuals, a 449 bp fragment from the D-loop region was obtained. Twenty polymorphic sites (17 transition and three transversion mutations) and four indels sites were found. Twelve different haplotypes were revealed, of which four haplotypes (H5, H7, H10, and H12) were singletons (**Figure 2B**). The H2 haplotype was the most frequent, observed in 37 samples from four different locations (IGU1, IGU2, STO, and FLO). Although SIL, FLO, GON, COT, and CAP present only one haplotype each, these haplotypes are shared with other locations, except for H5 found only in CAP. The IGU2 has the highest number of haplotypes ($N = 8$), followed by STO ($N = 6$), AND ($N = 3$), and IGU1 ($N = 2$). Haplotype (h) and nucleotide (π) diversity values were 0.7642 and 0.00729, respectively (**Table 2**).

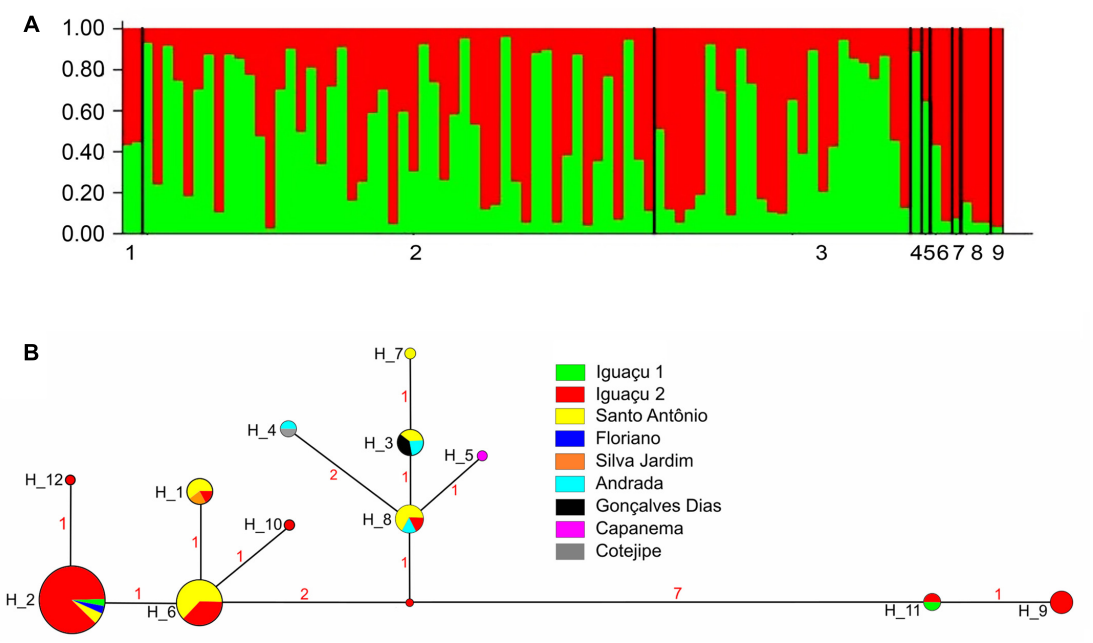


FIGURE 2 | (A) Graphical representation of Bayesian cluster analysis from $K = 1$. Each column represents a different individual of *G. setequedas*, and the colors represent the probability of the ancestral coefficient of individual in each genetic cluster. Numbers represent the collection points: (1) Iguaçu 1; (2) Iguaçu 2; (3) Santo Antônio; (4) Silva Jardim; (5) Floriano; (6) Gonçalves Dias; (7) Capanema; (8) Andrada; (9) Cotegipe. **(B)** Haplotype network based on partial sequencing of the D-loop region (mtDNA) of 82 individuals of *G. setequedas* from the Lower Iguaçu River. Circle sizes are proportional to haplotype frequency. Numbers between haplotypes denote mutational steps between sequences.

Demographic Analyses

The signed-rank test did not produce significant values in any of the mutational models (IAM, SSM, or TPM). In the mode-shift test, the samples showed typical L-shaped distribution (non-bottleneck) in the frequency of the alleles in the mode-shift test (**Table 3**). The mismatch distribution graphic demonstrated a multimodal distribution for haplotypes (**Figure 3**), which is usually ascribed to populations in demographic equilibrium. In the neutrality tests, Tajima test values (D) and Fu test values (F_s) were negative and not significant (**Table 2**).

Gene Flow

Bayesian gene flow analysis of the microsatellite data revealed contemporary migration values among the samples within the confidence interval (95%). Values for non-immigrants in each sample, ranging from 79 to 82.6%. However, migration estimates were also obtained and showed similar values among the majority of samples. The lowest migration estimates were from IGU 1 to COT and IGU 2 to AND, both with 1.9%. On the other hand, the highest migration values were from FLO to IGU2, and to GON, and COT to CAP, with 2.7%. Among all the samples, FLO was the one that obtained the highest percentage of migrants, for the largest number of sites (**Table 4**).

DISCUSSION

Genetic Diversity and Population Structure

The study of molecular markers, such as Microsatellite and mtDNA, generates important information on the genetic variation and structure of fish species and is a significant step toward realizing the goal of conservation of species in their natural populations (Carvalho-Costa et al., 2008; Piorski et al., 2008; Garcez et al., 2011; Abdul-Muneer, 2014). According to Wu et al. (2015), understanding the diversity and genetic structure of endangered species are essential to engage effective environmental conservation and management action. The long-term persistence of species depends on sufficient genetic diversity to adapt and survive in variable or changing environments (Hughes et al., 2008).

According to DeWoody and Avise (2000), based on a meta-analysis of microsatellite polymorphisms, freshwater fish, on average, have 9.1 ± 6.1 alleles and expected heterozygosity of 0.54 ± 0.25 per population. Therefore, based on microsatellites, the genetic diversity of the *G. setequedas* population in terms of allele numbers (17.14) and the expected heterozygosity (0.67) are as expected for freshwater fish. In addition to the high diversity in the nuclear markers, the *G. setequedas* analyzed showed significant variations in mitochondrial DNA, exhibiting high levels of genetic diversity. According to Freeland (2005), h is considered the haploid equivalent of HE in data on diploids. The similarity among these estimates suggests that the current variations in nuclear and mitochondrial DNA are evenly distributed throughout *G. setequedas* population.

A high level of genetic diversity is an important attribute for species and may confer the basis for adaptation to environmental change (Piorski et al., 2008), especially when it comes to endangered species such *as G. setequedas* (Abilhoa and Duboc, 2004; Pavanelli and Reis, 2008; International Union for Conservation of Nature [IUCN], 2014).

Demographic History

A 449 bp test in the D-Loop region, one of the most variable regions of mtDNA (Frankham et al., 2010), revealed 12 haplotypes and high values of π (0.00729) and h (0.750). In addition, analysis of the microsatellite data using Bottleneck program showed no significant recent bottlenecks. The absence of recent bottlenecks is corroborated by the high haplotype ($h > 0.5$) and nucleotide diversity ($\pi > 0.5\%$) values in the mtDNA. According to Grant and Bowen (1998), high haplotypic diversity combined with high nucleotide diversity represents a

TABLE 2 | Genetic diversity of *G. setequedas* in the Lower Iguaçu River basin, based on microsatellite markers and mitochondrial haplotypes (D-Loop).

Microsatellites							mtDNA					
N	***A***	N_A	$\bar{N}_E$	H_O	H_E	F_{IS}	***N***	**Nh**	***h***	**π**	***D***	F_S
86	120	17.143	6.574	0.593	0.673	0.126	82	12	0.7642	0.00729	−0.570	−0.285

N, number of individuals examined; A, total number of alleles; N_A, mean number of alleles; N_E, mean number of effective alleles; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, inbreeding coefficient; Nh, number of haplotypes; h, haplotype diversity; π, nucleotide diversity; D, Tajima's neutrality test; Fs, Fu's neutrality test.

TABLE 3 | Bottleneck tests in 86 samples of *G. setequedas* from the Lower Iguaçu River Basin.

		Sign test/Wilcoxon sign-rank test						Allele frequency distribution
		IAM[a]		TPM[b]		SMM[c]		
	N	H_d/H_e	***P***	H_d/H_e	***P***	H_d/H_e	***P***	
Total	86	3/4	0.710	6/1	0.996	7/0	1.000	*L-shaped*

N, number of individuals; H_e, number of loci exhibiting excess heterozygosity; H_d , number of loci exhibiting deficient heterozygosity. Normal L-shaped distribution = non-bottlenecked population. [a]*Infinite allele model,* [b]*two phase model,* [c]*stepwise mutation model.*

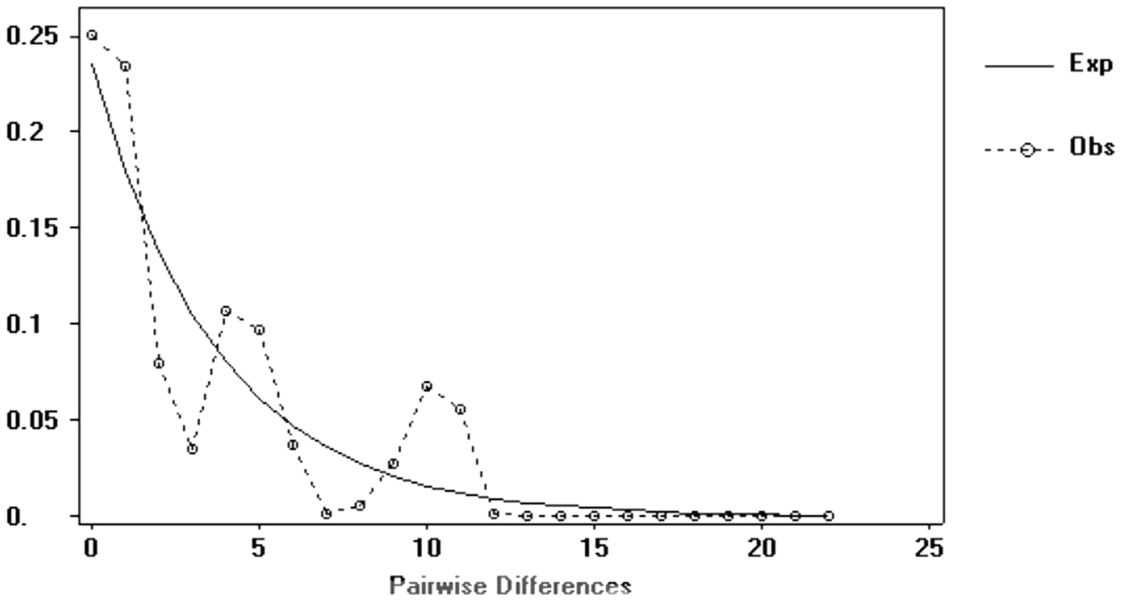


FIGURE 3 | Mismatch distribution of mitochondrial haplotypes obtained from *G. setequedas* samples studied in the Lower Iguaçu River basin.

large and stable population with a long evolutionary history, or secondary contact between different lineages. Stable population with a long evolutionary history seems to be a very plausible possibility for *G. setequedas*, as the Iguaçu waterfalls have exerted effective geographic isolation on ictiofauna of the Iguaçu river (Zawadzki et al., 1999), providing an accentuated degree of endemicity, of more than 70% (Abell et al., 2008). In addition, analysis of the distribution of substitution differences between pairs of haplotypes (mismatch distribution) (Cunha and Solé-Cava, 2012) shows multimodal distributions, which is generally attributed to populations in demographic equilibrium (Rogers and Harpending, 1992).

However, negative values in the Tajima's *D* test and Fu's F_s test, even if not significant, could suggest population expansion after an ancient bottleneck (Slatkin and Hudson, 1991; Grant and Bowen, 1998), indicating that all the current haplotypes are closely related and derived from a single main haplotype (H2). Signs of old bottlenecks may be less evident at microsatellite loci, since they tend to recover from the variation more rapidly than mitochondrial sequences. At the same time, π recovery after a genetic bottleneck is slower than *h* at mtDNA (McCusker and Bentzen, 2010).

Gene Flow

The individuals of the Floriano River (FLO) presented the highest rates of migration, and the highest levels of admixture in samples were found in the Iguaçu 2 and Gonçalves Dias rivers. The specimens from rivers further upstream in the drainage (FLO, GON, CAP, and COT) appear to have more levels of admixture between them. This factor might suggest that entry into the upper tributaries is more likely than the lower. However, the highest rates of migrants to Iguaçu 1 are from most upstream tributaries, suggesting that populations of *G. setequedas* maintain satisfactory gene flow in all stretches of the river studied. According to Palstra and Ruzzante (2008), if local populations are small, as is the case in the present study, gene flow is the key factor to prevent the

TABLE 4 | Gene flow estimates based on Bayesian inferences of migration rates using BayesAss.

Migration	IGU 1	IGU 2	STO	SIL	FLO	GON	CAP	AND	COT
IGU 1	**82.6%**	2.2%	2%	2.1%	2.6%	2.4%	2.4%	2.4%	2.4%
IGU 2	2.1%	**82.2%**	2.2%	2.1%	2.7%	2.2%	2.5%	2.5%	2.5%
STO	2%	2.3%	**81.5%**	2.3%	2.3%	2.4%	2.4%	2.4%	2.4%
SIL	2.2%	2.4%	2.5%	**81.7%**	2.6%	2.4%	2.4%	2.4%	2.4%
FLO	2%	2.2%	2.2%	2.3%	**79%**	2.4%	2.6%	2.4%	2.5%
GON	2.2%	2.2%	2.4%	2.3%	2.7%	**80.7%**	2.6%	2.5%	2.6%
CAP	2.2%	2.2%	2.4%	2.2%	2.5%	2.4%	**79.7%**	2.6%	2.7%
AND	2.4%	1.9%	2.1%	2.4%	2.5%	2.4%	2.3%	**80%**	2.2%
COT	1.9%	2%	2.2%	2.1%	2.6%	2.4%	2.5%	2.4%	**79.8%**

The bold values along the diagonal represent non-migrants within a putative source subpopulation.

stochastic loss of genetic diversity, besides providing the required alleles to subpopulations under selection that lack favorable genotypes (Kinnison and Hairston, 2007). Although these results allow inferring gene flow between localities, according to Wilson and Rannala (2003) a strong estimate can be reach with a higher sample size per locality. It can be the next goal for further studies, but it is a difficult task to solve immediately because the species is not abundant and is mainly distributed in a preservation area.

According to Fagan et al. (2002, 2005), riverine populations are forecasted to be particularly vulnerable to fragmentation due to their dendritic structure, which may be exacerbated by unidirectional migration. Natural barriers (rapids and waterfalls) and man-made structures, such as dams, also fragment riverine populations, influencing in the dispersal rate and migration pattern (Wofford et al., 2005), even of a rheophilic species of fishes with strong swimming abilities such as *G. setequedas* (Paiz et al., 2017). However, the construction of a new hydroelectric power plant (Baixo Iguaçu HPP), already underway between the Salto Caxias Reservoir and Iguaçu Falls, could fragment this population preventing the gene flow. As a consequence, there may be loss of genetic diversity and population decline, especially in the area of future reservoir. Moreover, this separate population can be extinguished, as has already happened with another population of *G. setequedas* after the construction of the Itaipu Hydroelectric Power Plant. The disappearance was attributed to the lentic waters of the Itaipu Reservoir, which isolated populations of this rheophilic species, which previously occurred in tributaries of both river banks, in Paraguay and Brazil, and probably in the Paraná River (Paiz et al., 2017).

Conservation Implications

The abundance, dispersal, and population size are reduced in populations structured by habitat fragmentation due to barriers such as dams, thereby increasing the risk of extinction (Gross et al., 2004; Letcher et al., 2007). This fragmentation can lead to the total or partial isolation of a population, conditioning the response of the individuals. Thus, in the recently found population of *G. setequedas* in the Iguaçu River, a drastic reduction and loss of genetic diversity, due to inbreeding, must be avoided preserving the lotic characteristics of the environment.

For instance, due to the effects of anthropogenic disturbance, small and isolated populations are more likely to suffer loss of genetic diversity and population decline, than a huge population with high genetic diversity (Frankham et al., 2010; Allendorf et al., 2012). According to Frankham (2003), inbreeding reduces reproduction and survival rates, and loss of genetic diversity reduces the ability of populations to evolve to cope with environmental changes, leading to extinction risk.

The type locality and most of the records of *G. setequedas* are in Paraguay, in tributaries of the right bank of the Paraná river, in the region of influence of the Itaipu reservoir and downstream (Reis et al., 1992). Since the species description, despite several attempts, it was not possible to collect new specimens from the known geographic range of occurrence (Agostinho et al., 2004; Pavanelli and Reis, 2008). According to Pavanelli and Reis (2008), this species no longer occurs in the Itaipu reservoir, as well as in the floodplain upstream of the reservoir. Despite several collection efforts on the Iguaçu River (Pavanelli and Reis, 2008), mainly in the Lower Iguaçu upstream the National Park (Baumgartner et al., 2012), this species was not collected. For this reason, the conservation status of *G. setequedas* was invariably attributed to a threatened category (Abilhoa and Duboc, 2004; Pavanelli and Reis, 2008; International Union for Conservation of Nature [IUCN], 2014; Paiz et al., 2017). However, recently Paiz et al. (2017) and the present study report the presence of *G. setequedas* in the Lower Iguaçu in the National Park region. In this way, the population of *G. setequedas* of the Lower Iguaçu River may be one of the last remnants of this species and, according to Pavanelli and Reis (2008), as *G. setequedas* is a naturally rare species, it is advisable that any anthropogenic changes in its original ecosystem be discouraged.

The results presented here demonstrate that the population of *G. setequedas* of the Iguaçu River still maintains satisfactory levels of genetic diversity. However, in terms of conservation management plans, to guarantee the survival of this species, it is necessary to maintain the tributaries of the Iguaçu River and the downstream area from the future reservoir (Baixo Iguaçu Reservoir) without additional dams. Long-term monitoring of genetic diversity and inbreeding could also help conserve this population and provide a basis for future decisions.

ETHICS STATEMENT

This study was carried out in strict accordance with the recommendations provided in the Guide for the Care and Use of Laboratory Animals. Collection was authorized by the System of Authorization and Information on Biodiversity – SISBIO (SISBIO n°. 25648-3 and 25648-4), by the Chico Mendes Institute for Biodiversity Conservation ICMBio 003/2014 and Official SEI n°. 63/2016-DIBIO/ICMBio), and by the Environmental Institute of Paraná – IAP (n°. 37788 and 43394). The sampling protocol was approved by the Ethics Committee on the Use of Animals – CEUA of the Universidade Estadual do Oeste do Paraná (n°. 62/09).

AUTHOR CONTRIBUTIONS

LS-S, DF, MM, and OS designed the research. LA, SP, SM, and MM collected data. LS-S, TK-D, and DF performed the molecular genetic studies. All authors contributed to the writing of the manuscript.

ACKNOWLEDGMENTS

We are grateful to ICMBio-Parque Nacional do Iguaçu; Consórcio Empreendedor Baixo Iguaçu (CEBI); Universidade Estadual do Oeste do Paraná (UNIOESTE); and Universidade Estadual de Londrina (UEL).

REFERENCES

Abdul-Muneer, P. M. (2014). Application of microsatellite markers in conservation genetics and fisheries management: recent advances in population structure analysis and conservation strategies. *Genet. Res. Int.* 1, 1–11. doi: 10.1155/2014/691759

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., et al. (2008). Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. *BioScience* 58, 403–414. doi: 10.1641/B580507

Abilhoa, V., and Duboc, L. F. (2004). "Peixes," in *Livro Vermelho da Fauna Ameaçada no Estado do Paraná*, eds S. B. Mikich and R. S. Bérnils (Curitiba: Instituto Ambiental do Paraná), 581–677.

Agostinho, A. A., Thomaz, S. M., and Gomes, L. C. (2004). Threats for biodiversity in the floodplain of the Upper Paraná River: effects of hydrological regulation by dams. *Ecohydrol. Hydrobiol.* 4, 267–280.

Allendorf, F. W., Luikart, G. H., and Aitken, S. N. (2012). *Conservation and the Genetics of Populations*. Oxford: Wiley Blackwell Press.

Almeida, F. S., Fungaro, M. H. P., and Sodré, L. M. K. (2001). RAPD and isoenzyme analysis of genetic variability in three allied species of catfish (Siluriformes: Pimelodidae) from the Tibagi River, Brazil. *J. Zool.* 253, 113–120. doi: 10.1017/S0952836901000103

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Bandelt, H. J., Forster, P., and Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 36–48. doi: 10.1093/oxfordjournals.molbev.a026036

Baumgartner, G., Pavanelli, C. S., Baumgartner, D., Bifi, A. G., Debona, T., and Frana, V. A. (2012). *Peixes do Baixo Rio Iguaçu*. Maringá: Eduem. doi: 10.7476/9788576285861

Beaumont, M. A., and Rannala, B. (2004). The Bayesian revolution in genetics. *Nat. Rev. Genet.* 5, 251–261. doi: 10.1038/nrg1318

Carvalho-Costa, L. F., Hatanaka, T., and Galetti, P. M. Jr. (2008). Evidence of lack of population substructuring in the Brazilian freshwater fish *Prochilodus costatus*. *Genet. Mol. Biol.* 31, 377–380. doi: 10.1590/S1415-47572008000200036

Cornuet, J. M., and Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144, 2001–2014.

Cronin, M. A., Spearman, W. J., Wilmot, R. L., and Patton, J. C. (1993). Mitochondrial DNA variation in chinook (*Oncorhynchus tshawytscha*) and chum salmon (*O. keta*) detected by restriction enzyme analysis of polymerase chain reaction (PCR) products. *Can. J. Fish. Aquat. Sci.* 50, 708–715. doi: 10.1139/f93-081

Cunha, H. A., and Solé-Cava, A. M. (2012). "Análise filogeográfica," in *Biologia Molecular e Evolução*, eds S. R. Matioli and F. M. C. Fernandes (Ribeirão Preto: Holos), 197–215.

DeWoody, J. A., and Avise, J. C. (2000). Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *J. Fish Biol.* 56, 461–473. doi: 10.1111/j.1095-8649.2000.tb00748.x

Earl, D. A., and VonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing structure output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Excoffier, L., and Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Fagan, W. F., Aumann, C., Kennedy, C. M., and Unmack, P. J. (2005). Rarity, fragmentation, and the scale-dependence of extinction-risk in desert fishes. *Ecology* 86, 34–41. doi: 10.1890/04-0491

Fagan, W. F., Unmack, P. J., Burgess, C., and Minckley, W. L. (2002). Rarity, fragmentation, and extinction risk in desert fishes. *Ecology* 83, 3250–3256. doi: 10.1890/0012-9658(2002)083[3250:RFAERI]2.0.CO;2

Ferreira, D. G., Galindo, B. A., Alves, A. N., Almeida, F. S., Ruas, C. F., and Sofia, S. H. (2013). Development and characterization of 14 microsatellite loci in the Neotropical fish *Geophagus brasiliensis* (Perciformes, Cichlidae). *J. Fish Biol.* 83, 1430–1438. doi: 10.1111/jfb.12227

Ferreira, D. G., Galindo, B. A., Frantine-Silva, W., Almeida, F. S., and Sofia, S. H. (2015). Genetic structure of a Neotropical sedentary fish revealed by AFLP, microsatellite and mtDNA markers: a case study. *Conserv. Genet.* 16, 151–166. doi: 10.1007/s10592-014-0648-2

Frankham, R. (2003). Genetics and conservation biology. *C. R. Biol.* 326, 22–29. doi: 10.1016/S1631-0691(03)00023-4

Frankham, R., Ballou, J. J. D., and Briscoe, D. D. A. (2010). *Introduction to Conservation Genetics*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511809002

Freeland, J. R. (2005). *Molecular Ecology*. Chichester: John Wiley & Sons Ltd.

Fu, Y. X. (1997). Statistical test of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.

Garcez, R., Calcagnotto, D., and Almeida-Toledo, L. F. (2011). Population structure of the migratory fish *Prochilodus lineatus* (Characiformes) from Rio Grande basin (Brazil), an area fragmented by dams. *Aquat. Conserv.* 21, 268–275. doi: 10.1002/aqc.1176

Goudet, J. (2001). *FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices (Version 2.9.3)*. Available at: http://www.unil.ch/dee/page36921_fr.html [accessed September 3, 2017].

Grant, W. S., and Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines e anchovies and lessons for conservation. *J. Hered.* 89, 415–426. doi: 10.1093/jhered/89.5.415

Gross, M. C., Schneider, C. H., Matiello, M. C. A., Leite, M. L., Bertollo, L. A. C., and Artoni, R. F. (2004). Population structure, fluctuating asymmetry and genetic variability in an endemic and highly isolated *Astyanax fish* population (Characidae). *Genet. Mol. Biol.* 27, 529–535. doi: 10.1590/S1415-47572004000400011

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.

Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., and Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecol. Lett.* 11, 609–623. doi: 10.1111/j.1461-0248.2008.01179.x

ICMBio (2017). *Lista de Espécies Ameaçadas*. Available at: http://www.icmbio.gov.br/portal/especies-ameacadas-destaque

International Union for Conservation of Nature [IUCN] (2014). *Standards and Petitions Subcommittee. Guidelines for using the IUCN Red List Categories and Criteria. Version 11*. Available at: http://www.iucnredlist.org/technical-documents/red-list-training/red-list-guidance-docs

Júlio Júnior, H. F., Bonecker, C., and Agostinho, A. A. (1997). "Reservatório de segredo e sua inserção na bacia do rio Iguaçu," in *Reservatório de Segredo: Bases Ecológicas Para o Manejo*, eds A. A. Agostinho and L. C. Gomes (Maringá: Eduem), 1–17.

Kinnison, M. T., and Hairston, N. G. (2007). Eco-evolutionary conservation biology: contemporary evolution and the dynamics of persistence. *Funct. Ecol.* 21, 444–454. doi: 10.1111/j.1365-2435.2007.01278.x

Letcher, B. H., Nislow, K. H., Coombs, J. A., O'donnell, M. J., and And Dubreuil, T. L. (2007). Population response to habitat fragmentation in a stream-dwelling brook trout population. *PLOS ONE* 2:e1139. doi: 10.1371/journal.pone.0001139

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187

Liotta, J. (2010). "Antecedentes y propuestas metodológicas para evaluar el estado de conservación de la ictiofauna de la cuenca del plata," in *Conservación de los Peces de la Cuenca del Plata en Argentina: Enfoques Metodológicos Para su Evaluación y Manejo*, eds J. Cappato, V. De la Balze, J. Peteán, and J. Liotta (Buenos Aires: Fundación para la Conservación y el Uso Sustentable de los Humedales).

Lowe, T. M., and Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25, 955–964. doi: 10.1093/nar/25.5.0955

Luikart, G., Allendorf, F. W., Cornuet, J. M., and Sherwin, W. B. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J. Hered.* 89, 238–247. doi: 10.1093/jhered/89.3.238

Luikart, G., and Cornuet, J. M. (1998). Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conserv. Biol.* 12, 228–237. doi: 10.1046/j.1523-1739.1998.96388.x

Maack, R. (1981). *Geografia Física do Estado do Paraná*. Rio de Janeiro: José Olympio Press.

Maack, R. (2001). Breves notícias sobre a geologia dos estados do Paraná e Santa Catarina. *Braz. Arch. Biol. Technol.* 169–288. doi: 10.1590/S1516-89132001000500010

McCusker, M. R., and Bentzen, P. (2010). Positive relationships between genetic diversity and abundance in fishes. *Mol. Ecol.* 19, 4852–4862. doi: 10.1111/j.1365-294X.2010.04822.x

Meyer, A., Kocher, T. D., Basasibwaki, P., and Wilson, A. C. (1990). Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* 347, 550–553. doi: 10.1038/347550a0

Paiz, L. M., Baumgärtner, L., Graça, W. J., Margarido, V. P., and Pavanelli, C. S. (2017). Cytogenetics of *Gymnogeophagus setequedas* (Cichlidae: Geophaginae), with comments on its geographical distribution. *Neotrop. Ichthyol.* 15:e160035. doi: 10.1590/1982-0224-20160035

Palstra, F. P., and Ruzzante, D. E. (2008). Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Mol. Ecol.* 17, 3428–3447. doi: 10.1111/j.1365-294X.2008.03842.x

Pavanelli, C. S., and Reis, R. E. (2008). "Gymnogeophagus setequedas," in *Livro Vermelho da Fauna Brasileira Ameaçada de Extinção*, eds A. B. M. Machado, G. M. Drummond, and A. P. Paglia (Brasília: Fundação Biodiversitas), 184–186.

Piorski, N. M., Sanches, A., Carvalho-Costa, L. F., Hatanaka, T., Carrillo-Avila, M., Freitas, P. D., et al. (2008). Contribution of conservation genetics in assessing neotropical freshwater fish biodiversity. *Braz. J. Biol.* 68, 1039–1050. doi: 10.1590/S1519-69842008000500011

Piry, S., Luikart, G., and Cornuet, J. M. (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J. Hered.* 90, 502–503. doi: 10.1093/jhered/90.4.502

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Raymond, M., and Rousset, M. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* 86, 248–249. doi: 10.1093/oxfordjournals.jhered.a111573

Reis, R. E., Malabarba, L. R., and Pavanelli, C. S. (1992). *Gymnogeophagus setequedas*, a new cichlid species (Teleostei: Labroidei) from middle rio Paraná system. Brazil and Paraguay. *Ichthyol. Explor. Freshw.* 3, 265–272.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* 43, 223–225. doi: 10.1111/j.1558-5646.1989.tb04220.x

Rogers, A. R., and Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9, 552–569.

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes* 4, 137–138. doi: 10.1046/j.1471-8286.2003.00566.x

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* 18, 223–234. doi: 10.1038/72708

Slatkin, M., and Hudson, R. R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129, 555–562.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. F. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4, 535–538. doi: 10.1111/j.1471-8286.2004.00684.x

Wilson, G. A., and Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163, 1177–1191.

Wofford, J. E. B., Gresswell, R. E., and Banks, M. A. (2005). Influence of barriers to movement on within-watershed genetic variation of coastal cutthroat trout. *Ecol. Appl.* 15, 628–637. doi: 10.1890/04-0095

Wu, F. Q., Shen, S. K., Zhang, X. J., Wang, Y. H., and Sun, W. B. (2015). Genetic diversity and population structure of an extremely endangered species: the world's largest Rhododendron. *AoB Plants* 7:plu082. doi: 10.1093/aobpla/plu082

Yeh, F. C., Yang, R., Boyle, T. J., and Xiyan, J. M. (2000). *Pop Gene 32: Microsoft Window-Based Freeware for Population Genetic Analysis v.1.32*. Edmonton, AB: University of Alberta.

Zawadzki, C. H., Renesto, E., and Bini, L. M. (1999). Genetic and morphometric analysis of three species of the genus *Hypostomus* Lacépède, 1803 (Osteichthyes: Loricariidae) from the Rio Iguaçu basin (Brazil). *Rev. Suisse Zool.* 106, 91–105. doi: 10.5962/bhl.part.80072

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 05 February 2018
doi: 10.3389/fgene.2018.00017

Identification of a New Mullet Species Complex Based on an Integrative Molecular and Cytogenetic Investigation of *Mugil hospes* (Mugilidae: Mugiliformes)

*Mauro Nirchio[1], Fabilene G. Paim[2], Valentina Milana[3], Anna R. Rossi[3] and Claudio Oliveira[2]**

[1] Facultad de Ciencias Agropecuarias, Universidad Técnica de Machala, Machala, Ecuador, [2] Departamento de Morfologia, Instituto de Biociências, Universidade Estadual Paulista "Júlio de Mesquita Filho", São Paulo, Brazil, [3] Dipartimento di Biologia e Biotecnologie "C. Darwin", Sapienza Università di Roma, Rome, Italy

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni,
Ponta Grossa State University, Brazil

Reviewed by:
Marcelo R. S. Briones,
Federal University of São Paulo, Brazil
Marcelo De Bello Cioffi,
Federal University of São Carlos, Brazil

***Correspondence:**
Claudio Oliveira
claudio@ibb.unesp.br

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

Received: 08 July 2017
Accepted: 15 January 2018
Published: 05 February 2018

Citation:
Nirchio M, Paim FG, Milana V, Rossi AR and Oliveira C (2018) Identification of a New Mullet Species Complex Based on an Integrative Molecular and Cytogenetic Investigation of Mugil hospes (Mugilidae: Mugiliformes). Front. Genet. 9:17. doi: 10.3389/fgene.2018.00017

Mullets are very common fishes included in the family Mugilidae, (Mugiliformes), which are characterized by both a remarkably uniform external morphology and internal anatomy. Recently, within this family, different species complexes were molecularly identified within *Mugil,* a genus which is characterized by lineages that sometimes show very different karyotypes. Here we report the results of cytogenetic and molecular analyses conducted on *Mugil hospes*, commonly known as the hospe mullet, from Ecuador. The study aims to verify whether the original described species from the Pacific Ocean corresponds to that identified in the Atlantic Ocean, and to identify species-specific chromosome markers that can add new comparative data about Mugilidae karyotype evolution. The karyotype of *M. hospes* from Ecuador is composed of 48 acrocentric chromosomes and shows two active nucleolar organizer regions (NORs). *In situ* hybridization, using different types of repetitive sequences (rDNAs, U1 snDNA, telomeric repeats) as probes, identified species-specific chromosome markers that have been compared with those of other species of the genus *Mugil*. Cytochrome c oxidase subunit I (COI) sequence analysis shows only 92–93% similarity with sequences previously deposited under this species name in GenBank, all of which were from the Atlantic Ocean. Phylogenetic reconstructions indicate the presence of three well-supported hospe mullet lineages whose molecular divergence is compatible with the presence of distinct species. Indeed, the first lineage includes samples from Ecuador, whereas the other two lineages include the Atlantic samples and correspond to *M. brevirostris* from Brazil and *Mugil* sp. R from Belize/Venezuela. Results here provided reiterate the pivotal importance of an integrative molecular and cytogenetic approach in the reconstruction of the relationships within Mugilidae.

Keywords: fish cytogenetics, fish molecular phylogeny, COI, chromosomal evolution, FISH, Mugilidae

INTRODUCTION

Mullets is the popular name of fishes included in the Mugilidae, a species rich family that is the only representative of the order Mugiliformes. These fishes are distributed in several coastal aquatic habitats in tropical, subtropical and temperate regions of the world, where they are ecologically, recreationally and commercially important (Thomson, 1966). According to different authors (see González-Castro and Ghasemzadeh, 2016 and references herein), the family has approximately 26 genera, but Eschmeyer and Fong (2017) ascribe to Mugilidae 20 genera and 75 valid species.

In Mugilidae, most of the classical morphological characters used in species identification and/or systematics have poor diagnostic power and morphometric variability is limited (Schultz, 1946; Thomson, 1997; González-Castro, 2007; González-Castro and Ghasemzadeh, 2016). These characteristics are associated with the wide distribution of most of the species, which raises questions about their actual taxonomic status. Cytogenetic and molecular studies have provided important data for understanding the systematic relationships and evolutionary pathways among mullet species (Harrison et al., 2007; Sola et al., 2007; Durand et al., 2012; Durand and Borsa, 2015). These studies have also shown that it is necessary to use integrative approaches to study mugilids. Indeed, the use of repetitive sequences such as ribosomal genes (18S rDNA and 5S rDNA) as probes in FISH mapping has been shown to be a very informative cytotaxonomic tool in revealing different lineages/species within Mugilidae (Nirchio et al., 2007, 2017; Sola et al., 2007). On the other hand, the utility of molecular markers in this family to identify the species, better define the genera, and reconstruct their phylogenetic relationships, is well-represented by the huge amount of literature on this topic published in the last 15 years (see Rossi et al., 2016 for a review and Durand et al., 2017). In addition, molecular phylogenetic analyses have been used successfully in the investigation of chromosome evolution in some fish groups as those of the genus *Characidium* (Pansonato-Alves et al., 2014) and *Triportheus* (Yano et al., 2014), and in *Geophagus brasiliensis* (Alves-Silva and Dergam, 2015).

Mugil, which presently includes 16 valid species (Eschmeyer and Fong, 2017), is the most cytogenetically studied genus among the Mugilidae. Nine species have been investigated to date (see section "Discussion" and **Figure 6**). Nonetheless, the number of species is probably underestimated currently, as recent molecular data have indicated that there are different species complexes within this genus. For example, the cosmopolitan *M. cephalus* was found to be composed of 15 well supported mitochondrial lineages (Durand and Borsa, 2015), including the one sampled in the type-locality (Mediterranean Sea); six of these lineages have already been cytogenetically analyzed (Rossi et al., 1996, 2016). However, these lineages lack formal descriptions and species name attribution.

Very recently, Durand et al. (2017) reported the presence of two well-supported mitochondrial lineages in the hospe mullet "*Mugil hospes,*" a species that, according to Barletta and Dantas (2016), is distributed in the western Atlantic from Belize to Brazil and in the eastern Pacific from Mexico to Ecuador. The first molecular lineage includes sequences from Brazil and corresponds to the resurrected species *Mugil brevirostris*, which is distributed from the northern Brazilian coast (Amapá) to the southern Brazilian coast (Rio Grande do Sul) (Menezes et al., 2015); the second lineage is represented by haplotypes collected in the Gulf of Mexico (Belize/Venezuela) and was named *Mugil* sp. R. Samples from the eastern Pacific were not included in these analyses or in any other molecular study. The karyotype of the species remains undescribed.

In this research, specimens of *M. hospes* from Ecuador have been collected and their morphological characters accurately analyzed to make sure of the correct species identification. Cytogenetic and mitochondrial cytochrome c oxidase subunit I (COI) sequence analyses were performed aiming to (a) verify whether the original described *M. hospes* from the Pacific Ocean corresponds to one of the two lineages identified in the Atlantic Ocean or represents a third lineage, (b) estimate if the divergence among lineages is sufficient to attribute them to different species, (c) identify species-specific chromosome markers and add new comparative data that allow cytotaxonomic inferences on Mugilidae karyotype evolution.

MATERIALS AND METHODS

Fourteen specimens of *Mugil hospes* (four males, four females, six immature), were collected with a cast net from a reservoir that provides water to a shrimp pool located at Barbones, El Oro Province, Ecuador (3°09′14.0″ S 79°53′53.1″ W). Fishes were transported to the laboratory in sealed plastic bags (32′) containing two gallons of water, and the air in the bags was replaced with pure oxygen. All 14 individuals were used to prepare cell suspensions. A subsample of eight individuals was used for molecular and morphological analyses. Voucher specimens were deposited in the fish collection of the Laboratório de Biologia e Genética de Peixes (LBP), UNESP, Botucatu (São Paulo State, Brazil) (collection numbers LBP 23325) and Universidad Técnica de Machala (UTMACH-174-UTMACH-182; UTMACH-187; UTMACH-191-UTMACH-194). All experiments were conducted according to the Ethical Committee of Instituto de Biociências/UNESP/Botucatu, under protocol number 1057.

Morphological Analysis

Each fish was measured. Measurements and counts were taken as described by Menezes et al. (2010). Mouth width and mouth depth were measured as described by Thomson (1997). Twenty morphometric characters (Supplementary Table 1) and nine meristic characters (Supplementary Table 2) were recorded for each fish.

Molecular Analysis

Genomic DNA was extracted from muscle tissue that was preserved in 95% ethanol. DNA samples were obtained for eight specimens (one male, two female, five immature), according to procedures described by Aljanabi and Martínez (1997). A 655 bp fragment of the mitochondrial COI was amplified by PCR and

sequenced using primers and protocols reported by Nirchio et al. (2017). DNA sequences were aligned using the software Clustal X (Thompson et al., 1997) and deposited in GenBank (Accession numbers: KY964500-KY964504). The basic local alignment search tool (BLAST[1]) was used to search for similar sequences to confirm species assignment.

For phylogenetic tree reconstruction, a subset of the COI sequences of *Mugil,* previously analyzed by Durand et al. (2017), was considered. Those sequences that showed greater than 90% similarity (i.e., the six sequences of *M. brevirostris* and the seven sequences of *Mugil* sp. R) were also included; *Agonostomus monticola* (Bancroft, 1834) (JQ060401) was used as an outgroup.

Three types of phylogenetic reconstructions were conducted: neighbour-joining (NJ), maximum-likelihood (ML) and Bayesian inference (BI) analyses. NJ and ML analyses (1000 bootstrap pseudoreplicates) were performed using MEGA7 (Kumar et al., 2016) and PhyML 3.0 (Guindon et al., 2010), respectively. Bayesian analyses were carried out as implemented in MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001). Two independent runs of four Markov chains, each for 1,000,000 generations were performed. ModelTest 3.7 (Posada and Crandall, 1998) and MrModelTest 2.3 (Nylander, 2008) were used to select, according to the Akaike information criterion, the evolutionary models that best fit the data set for the ML (GTR + I + G, with nst = 6, gamma shape = 4.682, and proportion of invariant sites = 0.637) and the BI (GTR + I + G) analyses, respectively. Genetic distances were calculated with MEGA7 using the Kimura-2-parameters substitution model (Kimura, 1980).

Cytogenetic Analysis

Each fish received an intra-abdominal injection of 0.0125% colchicine (1.0 ml/100 g body weight) 50 min before being sacrificed by administering a numbing overdose of benzocaine (250 mg/L) as recommended by the Guidelines for the Euthanasia of Animals of the American Veterinary Medical Association (AVMA, 2013). Kidney cells were suspended, and chromosomes were prepared by following the conventional air-drying method, as described by Nirchio and Oliveira (2006). Classical staining techniques (Giemsa, Ag-staining, C-banding) and fluorescence *in situ* hybridization (FISH) were used to map ribosomal gene clusters (5S rDNA and 18S rDNA) and U1 snRNA gene clusters (U1 snRNA is a non-coding RNA that forms part of the spliceosome) (Nilsen, 2003). Telomeric probes were also applied. For the conventional karyotype, slides were stained for 20 min with 10% Giemsa in phosphate buffer at pH 6.88. Active nucleolus organizer regions (NORs) were revealed by silver (Ag) staining as described by Howell and Black (1980); this was performed after Giemsa staining (Rábová et al., 2015). C-banding was performed following the method of Sumner (1972).

The 5S rDNA, 18S rDNA, U1 snRNA genes and telomeric repeats were mapped onto chromosomes by FISH using the method described by Pinkel et al. (1986). Sequences of 5S rDNA, 18S rDNA, U1 snDNA and telomeric repeats were obtained by polymerase chain reaction (PCR) from the genome of *Hypsolebias flagellatus* and used as probes. The primers used for amplification were 5SA (5′-TCA ACC AAC CAC AAA GAC ATT GGC AC-3′) and 5SB (5′-TAG ACT TCT GGG TGG CCA AAG GAA TCA-3′) (Pendás et al., 1995), 18S6F (5′-CTC TTT CGA GGC CCT GTA AT-3′) and 18S6R (5′-CAG CTT TGC AAC CAT ACT CC-3′) (Utsunomia et al., 2016), U1F (5′-GCA GTC GAG ATT CCC ACA TT-3′) and U1R (5′-CTT ACC TGG CAG GGG AGA TA-3′) (Silva et al., 2015) and (TTAGGG)5 and (CCCTAA)5 (Ijdo et al., 1991). The 5S rDNA and telomeric probes were labeled with biotin-16-dUTP (2′-deoxyuridine 5′-triphosphate), and the 18S rDNA and U1 snRNA gene probes were labeled by including digoxigenin-11-dUTP in the PCR. Hybridization was detected with fluorescein-conjugated avidin (FITC, Sigma–Aldrich[2]) and anti-digoxigenin-rhodamine conjugate (Roche Applied Science[3]), respectively. Chromosomes were counterstained with 4,6-diamidino-2-phenylindole (DAPI), which was included in the Vectashield mounting medium (Vector Laboratories[4]).

Conventionally stained metaphase cells were photographed using a Motic B400, equipped with a Moticam 5000C digital camera using Motic Images Plus 2.0 ML software. FISH images were captured with an Olympus BX61 photomicroscope equipped with a DP70 digital camera using Image-Pro plus 6.0 software (Media Cybernetics). Images were merged and edited for optimization of brightness and contrast using Photoshop (Adobe Systems, Inc.) Version 2015.0.0.

RESULTS

Meristic and Morphometric Characters

The fresh specimens were gray on the dorsal side and white/silver on the ventral side. The pelvic fins had a yellowish tone, and the base of each pectoral fin had a visible dark spot. The dorsal fins and caudal fins were dusky. The distal tips of the anterior rays of the second dorsal fin were slightly darker. The pelvic and anal fins were pale. The body was elongated, with a slightly pointed snout (see **Figure 1**). The origin of the first dorsal fin was midway between the tip of the snout and the base of the caudal fin. The second dorsal fin and anal fin were profusely covered with scales. One row of small teeth was visible on the upper and lower lips (viewed under the microscope). There were adipose eyelids and widely separated spiny-rayed dorsal fins with four spines in the first dorsal fin and one spine plus eight soft rays in the second dorsal fin (small specimen with nine soft rays). Pelvic fins were sub-abdominal with one spine and 5–6 branched soft rays (commonly I+5). Pectoral fins were long, reaching the level of the origin of the first dorsal fin or extending just beyond, with two spines (the first spine very small) and 11–13 soft rays (commonly 12 rays). The anal fin had three spines and nine soft rays (first spine very short, and hidden by overlying scales). There was a large pectoral axillary scale, with 37–38 scales in longitudinal series (commonly 38), 11–14 scales in an oblique row extending to the origin of the pelvic fin (commonly 13) and 13 scales in a

[1]https://blast.ncbi.nlm.nih.gov/Blast.cgi

[2]www.sigma-aldrich.com

[3]https://lifescience.roche.com/

[4]https://vectorlabs.com/

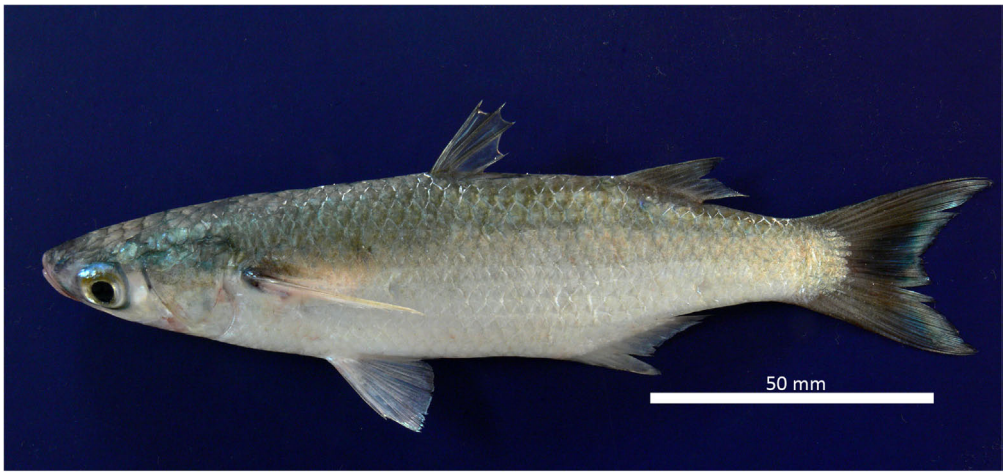


FIGURE 1 | Specimen of *Mugil hospes* from Ecuador.

transversal series, as well as 17–22 scales in a circum-peduncular series (commonly 19) (Supplementary Table 2).

Molecular Analysis

BLAST was used to show that COI nucleotide sequences from GenBank have 92–93% similarity with specimens originally identified as *Mugil hospes*, and with *M. trichodon* and *Mugil* sp., which were all collected in the Atlantic Ocean (from Brazil and Belize). Similarity values with other *Mugil* species were all below 90%.

The phylogenetic tree obtained by NJ, ML (lnL = −3276.75382), and BI (lnL = −3986.503923) analyses (**Figure 2**) shows three well-supported lineages of "*M. hospes*." The first two correspond to the *M. brevirostris* (Brazil) and *Mugil* sp. R (Belize and Venezuela) lineages identified by Durand et al. (2017), whereas the third, referred to hereafter as *Mugil hospes* (see Discussion), includes all the sequences from Ecuador obtained in this study (**Figure 2**). The genetic distance is 0.077 between *M. hospes*/*M. brevirostris* and between *M. hospes*/*Mugil* sp. R, and 0.073 between *M. brevirostris*/*Mugil* sp. R.

Cytogenetic Analysis

All individuals showed a diploid number of 2n = 48 and a karyotype composed entirely of uniformly decreasing acrocentric chromosomes. Thus, the fundamental arm number (NF) was 48. Only two pairs of homologous chromosomes can be identified with certainty: pair 5, due to a clear interstitial secondary constriction, and pair 24, which is distinctly small (**Figure 3A**). Sequential Giemsa-silver (Ag) nitrate staining enabled the identification of two actively transcribing NORs, interstitially located on the secondary constriction of chromosome pair 5 (**Figures 3B,C**). C-banding showed that constitutive heterochromatin is restricted to the centromeric regions of all chromosomes, and there is a pericentromeric heterochromatin block on the secondary constriction of chromosome pair 5 (**Figure 3D**).

Double FISH experiments using 5S and 18S rDNA as probes revealed two positive sites detected for each probe (**Figure 4A**), located on different chromosome pairs. The 18S rDNA positive sites correspond to the $AgNO_3$ sites on the secondary constriction of chromosome pair 5. The 5S rDNA

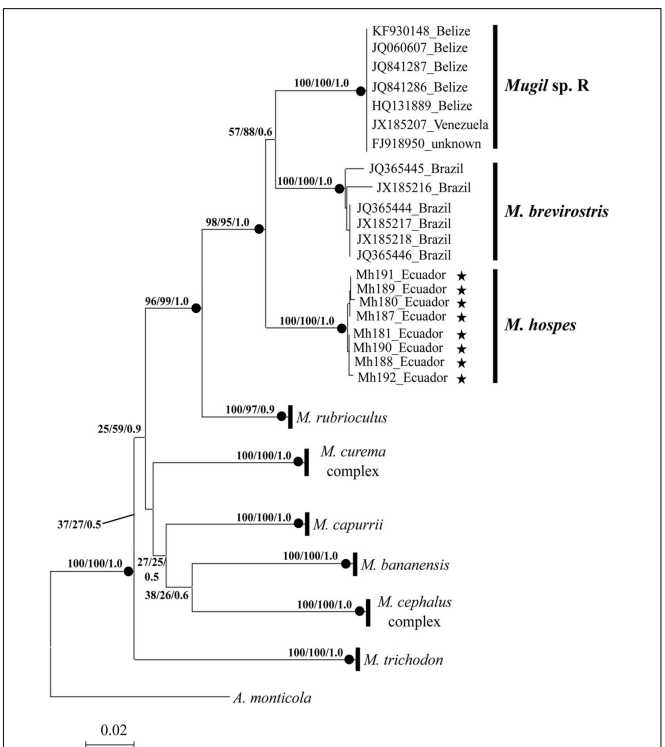


FIGURE 2 | Neighbor-joining tree based on COI sequences. At each node bootstrap values (NJ and ML) and posterior probabilities (BI) are shown. Black dots at the nodes indicate bootstrap values > 70% (NJ and ML) and posterior probabilities > 0.9 (BI). Stars indicate sequences obtained in this study; the remaining sequences are from Durand et al. (2017).

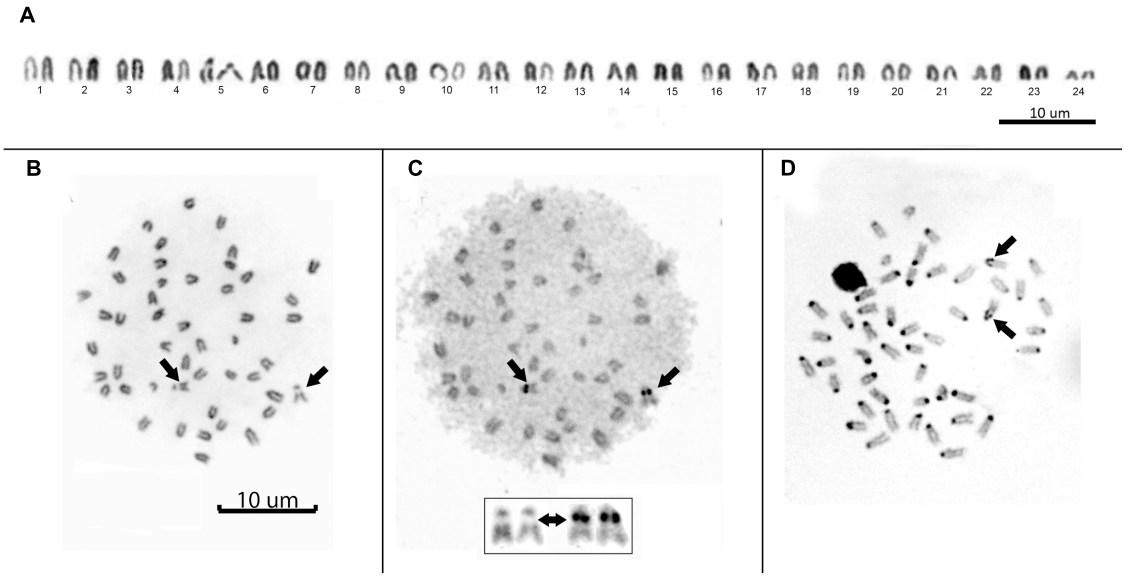


FIGURE 3 | (A) Conventional Giemsa-stained karyotype of *Mugil hospes.* Sequential Giemsa **(B)** and $AgNO_3$-staining **(C)** of the metaphase plate (inset shows NOR-bearing chromosomes); arrows indicate the NOR-bearing chromosome. **(D)** C-banded chromosomes; arrows indicate positive heterochromatic blocks on the interstitial secondary constriction of chromosome pair 5.

probes hybridized interstitially on a pair of medium-sized chromosomes. Double FISH using the U1 snDNA and 5S rDNA probes revealed positive U1 snDNA signals on the telomeric region of a pair of medium-sized chromosomes distinct from the 5S-rDNA-bearing chromosome pair (**Figure 4B**). When chromosomes treated by double FISH were sorted by decreasing size, it was possible to assign the 5S rDNA sites to pair 15 and the U1 snDNA signals to pair 10 (**Figure 5**). Telomeric repeats were located at both ends of each chromosome, although signal intensities varied between chromosomes (**Figure 4C**).

DISCUSSION

Meristic and morphometric data of samples from Ecuador agree with the original description of *M. hospes* (Jordan and Culver 1895 in Jordan, 1895): this species has pectoral fins whose tips reach and extend slightly past the vertical line passing through the origin of the first dorsal fin (with four spines). This morphological character is shared with *M. brevirostris*, which inhabits the opposite side of the Americas (i.e., the Atlantic coast).

Sequence analysis showed that samples of the hospe mullet from Ecuador are genetically very different from those collected in the Atlantic Ocean, all of which were originally identified with the same name *M. hospes*. Thus, in addition to the two lineages identified by Durand et al. (2017) in the Atlantic Ocean, *M. brevirostris* and *Mugil* sp. R, a third lineage is present in the Pacific Ocean. The genetic distances between the Pacific and the two Atlantic lineages are higher than the COI 2% threshold value that discriminates different species (Ward, 2009), and in the phylogenetic reconstruction, the three species form a monophyletic and well-supported clade. Cryptic species are defined as distinct evolutionary lineages not detectable with traditional taxonomic approaches, due to the absence of morphological differences (Avise, 2000; Mallet, 2010). In the last decade, barcoding methods based on COI sequences have made possible their identification in several marine and freshwater fish species (Ward et al., 2008; Lara et al., 2010; Puckridge et al., 2013; Mateussi et al., 2017; Okamoto et al., 2017; Ramirez et al., 2017; Shimabukuro-Dias et al., 2017). In Mugilidae evidences of cryptic species were inferred from mitochondrial tree topology, independent data from nuclear markers, and on the base of the geographic distribution of sister lineages (Durand and Borsa, 2015). Our results indicate that the genetic distances between the different hospe mullet are comparable to those reported among species within both the *M. cephalus* and *M. curema* species complexes (Durand et al., 2017), and the three lineages inhabits different geographic areas. Thus, we hypothesize that besides the *Mugil cephalus* and *Mugil curema* species complexes, there is an additional putative one, which should be identified as the *M. hospes* species complex; the name *M. hospes* should be kept by the Pacific samples, being Mazatlán (in the eastern Pacific) the species type-locality.

Cytogenetic analysis shows that the 48 acrocentric chromosome karyotype detected in *M. hospes* is consistent with the generally available data on diploid chromosome number and karyotype structure in Mugilidae (Sola et al., 2007; Rossi et al., 2016). This confirms that the only exception is represented by the mullets belonging to the *Mugil curema* species complex (Nirchio et al., 2017).

Apart from the number of chromosomes, many microstructural changes are evident looking at the variability in the locations of ribosomal genes in *Mugil*. For example, 5S rDNA cistrons are always localized to an interstitial position,

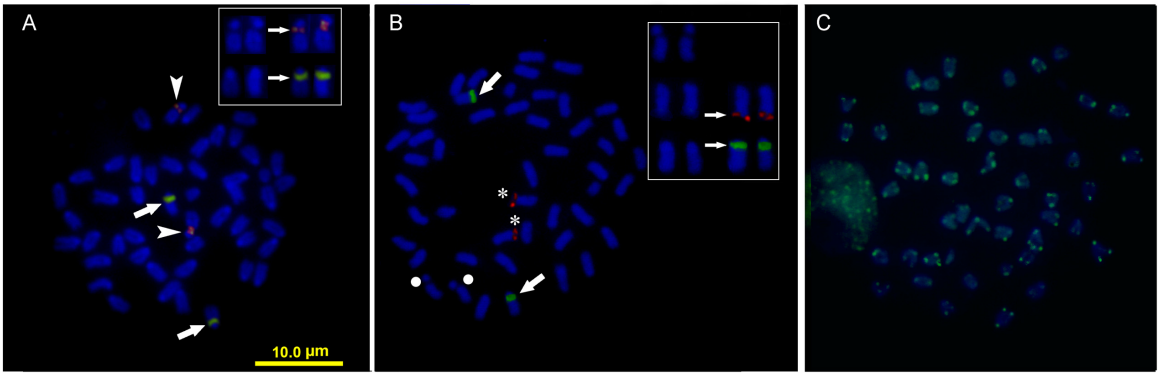


FIGURE 4 | Somatic metaphase chromosomes of *Mugil hospes* assayed by FISH and counterstained with DAPI: **(A)** 5S rDNA (arrows) and 18S rDNA (arrowheads); **(B)** U1 snRNA (asterisks), 5S rDNA (arrows), chromosome pair 5 (circle), and **(C)** telomeric repeats. Enlargement of selected samples of chromosome pairs after DAPI staining (left) and FISH (right) are shown in the insets: **(A)** chromosome pairs 5 and 15, with probes showing 18S rDNA (above) and 5S rDNA (below) positive sites; **(B)** chromosome pair 5 (above), chromosome pair 10 (center) showing U1 snDNA positive sites and chromosome pair 15 (below) showing 5S rDNA.

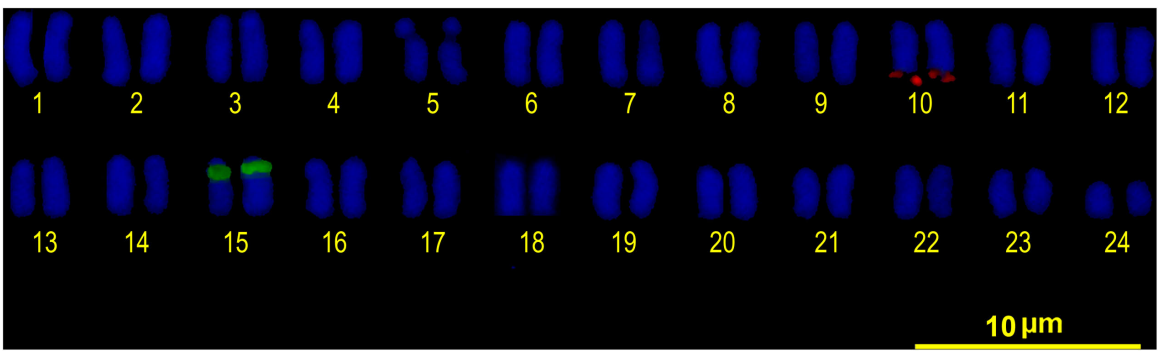


FIGURE 5 | FISH karyotype. Interstitial secondary constriction corresponding to NOR (chromosome pair 5); 5S rDNA (chromosome pair 15) and U1 snDNA (chromosome pair 10) positive sites are evident.

although they are on different chromosomes in different species (**Figure 6**). The 18S rDNA cistrons seems to be more variable and can be found in the telomeric or interstitial regions of a long chromosome arm, or even on the short arms of different chromosomes. The variability in the localization of the major ribosomal genes could be attributable to their association with heterochromatinas that is observed in *Mugil cephalus*, *M. margaritae* (formerly *M. curema*), *M. rubrioculus*, *M. curema*, *M. liza, M. trichodon*, *M. incilis*, *Mugil* sp. O (Rossi et al., 1996, 2005; Nirchio et al., 2005a,b, 2007, 2017; Hett et al., 2011), and *M. hospes* (present study). Heterochromatin is known to evolve rapidly, and its composition, that includes highly repetitive simple sequences like satellite DNA and transposable elements, is often different even between closely related species. This characteristic might promote rearrangements of the associated genes and might play an important role in reproductive isolation between sister species (Hughes and Hawley, 2009).

Cytogenetic mapping of U1 snDNA probes in *M. hospes* showed the presence of a single U1 gene cluster, located in the terminal position of a chromosomal pair different from the 18S rDNA and 5S bearing chromosomes (**Figures 4B**, **5**). There are no data available on the chromosome mapping of these sequences in other Mugilidae; thus, it is not possible, at this stage, to compare our results with those of other species in the family. However, the mapping of these sequences, combined with other repetitive sequences in other *Mugil*, might allow the identification of other chromosome re-arrangement. The analysis of chromosome localization of these sequences is restricted to a few other fish species. In *Merluccius merluccius* (Merlucciidae), multiple interstitial U1 sites are present (García-Souto et al., 2015). In 19 species of cichlids (Cabral-de-Mello et al., 2012), these sites could be either interstitial or terminal on a single st/a chromosome pair, and represent good chromosomal markers that allow the detection of many microstructural chromosomal rearrangements. On the contrary, in five species of *Astyanax* (Characidae), there is a conserved pattern in the number of U1 sites per genome, and these sequences are frequently associated with 5S rDNA sequences (Silva et al., 2015).

Telomeric DNA repeat sequences were found at the very ends of chromosomes, as observed in 15 different orders of teleosts (Ocalewicz, 2013). In mugilids, telomeric repeats have been mapped in nine species (Gornung et al., 2004; Rossi

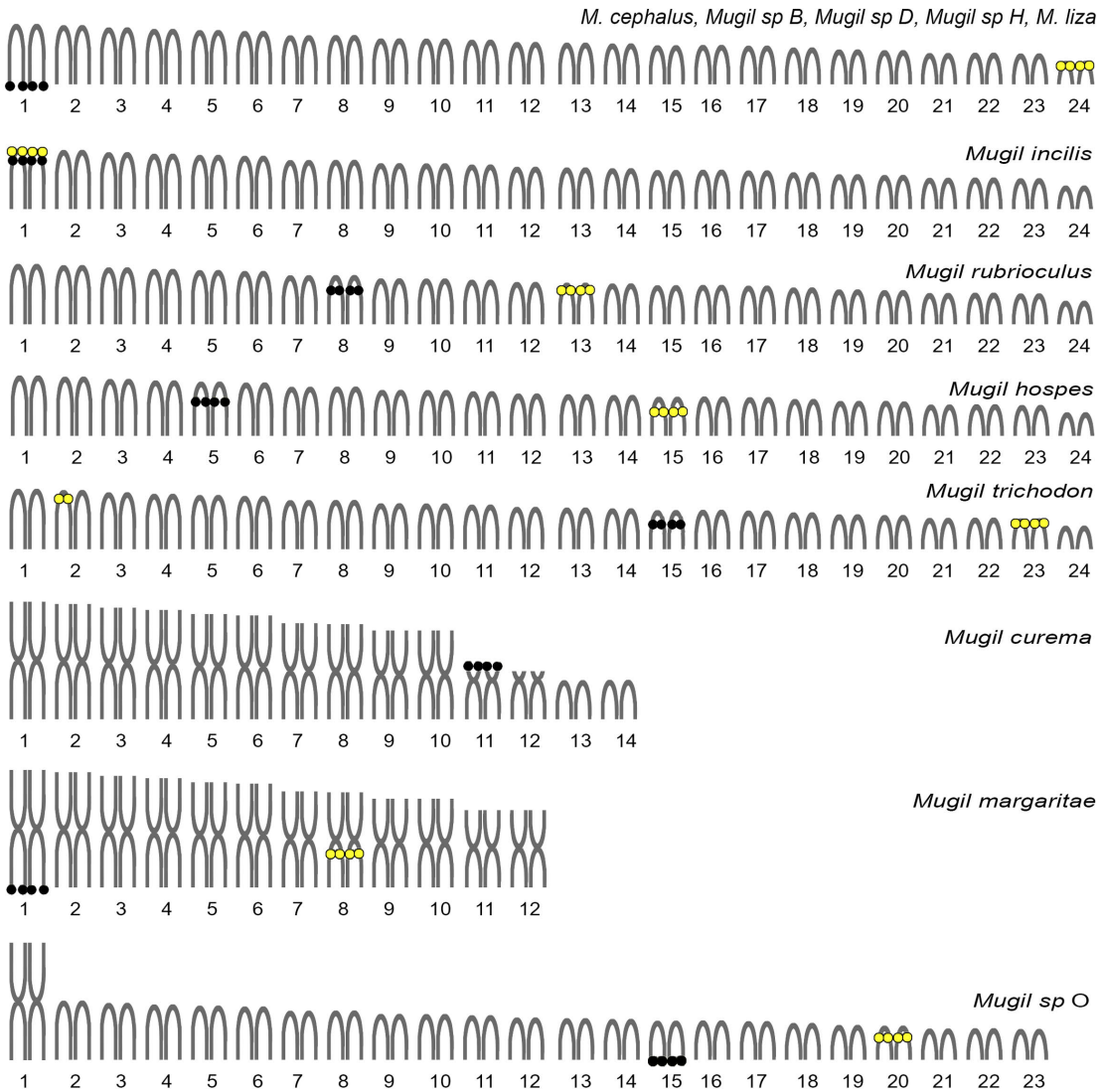


FIGURE 6 | Idiograms of the karyotypes observed in *Mugil* species. Re-drawn and updated from Rossi et al. (2016). Black circles: major ribosomal gene locations. Yellow circles: minor ribosomal gene locations.

et al., 2005; Nirchio et al., 2017) and were found also to be interspersed in NORs. Signal intensity variability between chromosomes, as observed in *M. hospes*, has been previously reported in other fishes (Rocco et al., 2002; Ocalewicz and Dobosz, 2009; Pomianowski et al., 2012), including *Mugil* species such as *M. cephalus* (Gornung et al., 2004), *M. liza* and *M. margaritae* (Rossi et al., 2005), and *Mugil* sp. O (Nirchio et al., 2017). This variability is probably due to differing copy numbers of these sequences in the different sites (Lansdorp et al., 1996).

CONCLUSION

The data presented here confirm that a complex dynamic has played in the karyotype evolution of *Mugil,* and they reiterate the usefulness of cytogenetic and molecular data in the reconstruction of relationships among taxa within Mugilidae. Species of this family are usually characterized by morphological features that are "insufficient to describe its actual species diversity" (Durand and Borsa, 2015). The combined use of morphological, molecular and cytogenetic analysis is necessary in these fishes to avoid species misidentification and to reconcile the confused picture obtained by morphology-based taxonomy with molecular-based taxonomy. In the case of the *M. hospes* species complex, *Mugil* sp. R, which is distributed in the Caribbean Sea, still deserves a formal morphological description and specific name attribution. This species, along with the Brazilian species *M. brevirostris*, also requires a karyotype description. Thus, at this stage, it is not possible to determine whether this complex is characterized by karyotypes that differ in the total number and morphology of chromosomes, like the *M. curema* species

complex (Nirchio et al., 2017), or whether it is characterized by karyotype homogeneity, like *M. cephalus* species complex (Rossi et al., 1996).

AUTHOR CONTRIBUTIONS

MN, CO, and AR designed the study. FP and VM conducted the lab work. MN, AR, and VM designed and conducted the analyses. All authors analyzed the results and wrote the manuscript.

FUNDING

Funding was provided by: Centro de Investigación of Universidad Técnica de Machala, Ecuador; Fundação de Amparo à Pesquisa do Estado de São Paulo -FAPESP-, Brazil; Conselho Nacional de Desenvolvimento Científico e Tecnológico -CNPq-, Brazil; Sapienza University, Rome (Progetto Ricerca Università 2016).

ACKNOWLEDGMENTS

The authors are grateful to Dr. Naercio Menezes and Dr. Ricardo Britzke, who helped with the identification of *M. hospes*. All experiments comply with the current laws of Ecuador and Italy.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00017/full#supplementary-material

REFERENCES

Aljanabi, S. M., and Martínez, L. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* 25, 4692–4693. doi: 10.1093/nar/25.22.4692

Alves-Silva, A. P., and Dergam, J. A. (2015). Cryptic speciation within the Neotropical Cichlid *Geophagus brasiliensis* (Quoy & Gaimard, 1824) (Teleostei Cichlidae): a new paradigm in karyotypical and molecular evolution. *Zebrafish* 12, 91–101. doi: 10.1089/zeb.2014.0999

Avise, J. C. (2000). *Phylogeography*. Boston, MA: Harvard University Press.

AVMA (2013). *Guidelines for the Euthanasia of Animals. Version 2013.0.1.* Available at: https://www.avma.org/KB/Policies/Documents/euthanasia.pdf (accessed May 27, 2016).

Bancroft, E. N. (1834). "The class Pisces, arranged by the Baron Cuvier, with supplementary additions," in *The Animal Kingdom*, eds E. Griffith and C. H. Smith (London: Whittaker & Co.), 680.

Barletta, M., and Dantas, D. V. (2016). "Biogeography and distribution of Mugilidae in the Americas," in *Biology, Ecology and Culture of Grey Mullet (Mugilidae)*, eds D. Crosetti and S. J. M. Blaber (Boca Raton, FL: CRC Press), 42–62. doi: 10.1201/b19927-4

Cabral-de-Mello, D. C., Valente, G. T., Nakajima, R. T., and Martins, C. (2012). Genomic organization and comparative chromosome mapping of the U1 snRNA gene in cichlid fish, with an emphasis in *Oreochromis niloticus*. *Chromosome Res.* 20, 279–292. doi: 10.1007/s10577-011-9271-y

Durand, J. D., and Borsa, P. (2015). Mitochondrial phylogeny of grey mullets (Acanthopterygii: Mugilidae) suggests high proportion of cryptic species. *C. R. Biol.* 338, 266–277. doi: 10.1016/j.crvi.2015.01.007

Durand, J. D., Chen, W. J., Shen, K. N., Jamandre, B. W., Blel, H., Diop, K., et al. (2012). Systematics of the grey mullets (Teleostei: Mugiliformes: Mugilidae): molecular phylogenetic evidence challenges two centuries of morphology-based taxonomy. *Mol. Phylogenet. Evol.* 64, 73–92. doi: 10.1016/j.ympev.2012.03.006

Durand, J. D., Hubert, N., Shen, K. N., and Borsa, P. (2017). DNA barcoding grey mullets. *Rev. Fish Biol. Fish.* 27, 233–243. doi: 10.1007/s11160-016-9457-7

Eschmeyer, W. N., and Fong, J. D. (2017). *Catalog of Fishes, Species by Family/Subfamily. Electronic Version.* Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp (accessed May 30, 2017).

García-Souto, D., Troncoso, T., Pérez, M., and Pasantes, J. J. (2015). Molecular cytogenetic analysis of the European hake *Merluccius merluccius* (Merlucciidae, Gadiformes): U1 and U2 snRNA Gene Clusters Map to the Same Location. *PLOS ONE* 10:e0146150. doi: 10.1371/journal.pone.0146150

González-Castro, M. (2007). *Los Peces Representantes de la Familia Mugilidae en Argentina*. Ph.D. thesis, Universidad Nacional de Mar del Plata, Argentina.

González-Castro, M., and Ghasemzadeh, J. (2016). "Morphology and morphometry based taxonomy of Mugilidae," in *Biology, Ecology and Culture of Grey Mullet (Mugilidae)*, eds D. Crosetti and S. J. M. Blaber (Boca Raton, FL: CRC Press), 1–20.

Gornung, E., Mannarelli, M. E., Rossi, A. R., and Sola, L. (2004). Chromosomal evolution in Mugilidae (Pisces, Mugiliformes): FISH mapping of the (TTAGGG)n telomeric repeat in the six Mediterranean mullets. *Hereditas* 140, 158–159. doi: 10.1111/j.1601-5223.2004.01784.x

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Harrison, I. J., Nirchio, M., Oliveira, C., Ron, E., and Gaviria, J. (2007). A new species of mullet (Teleostei: Mugilidae) from Venezuela, with a discussion on the taxonomy of *Mugil gaimardianus*. *J. Fish Biol.* 71, 76–97. doi: 10.1111/j.1095-8649.2007.01520.x

Hett, A. K., Nirchio, M., Oliveira, C., Sicchia, Z. R., Rossi, A. R., and Sola, L. (2011). Karyotype characterization of Mugil incilis Hancock, 1830 (Mugiliformes: Mugilidae), including a description of an unusual co-localization of major and minor ribosomal genes in the family. *Neotrop. Ichthyol.* 9, 107–112. doi: 10.1590/S1679-62252011005000005

Howell, W. M., and Black, D. A. (1980). Controlled silver staining of nucleolus organizer regions with a protective colloidal developer: a 1-step method. *Experientia* 36, 1014–1015. doi: 10.1007/BF01953855

Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755. doi: 10.1093/bioinformatics/17.8.754

Hughes, S. E., and Hawley, R. S. (2009). Heterochromatin: a rapidly evolving species barrier. *PLOS Biol.* 7:e1000233. doi: 10.1371/journal.pbio.1000233

Ijdo, J. W., Wells, R. A., Baldini, A., and Reeders, S. T. (1991). Improved telomere detection using a telomere repeat probe (TTAGGG)n generated by PCR. *Nucleic Acids Res.* 19:4780. doi: 10.1093/nar/19.17.4780

Jordan, D. S. (1895). The fishes of Sinaloa. *Proc. California Acad. Sci.* 5, 377–514.

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Biol.* 16, 111–120. doi: 10.1007/BF01731581

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lansdorp, P. P. M., Verwoerd, N. P., Van den Rijke, F. M., Dragowska, V., Little, M. T., Dirks, R. W., et al. (1996). Heterogeneity in telomere length of human chromosomes. *Hum. Mol. Genet.* 5, 685–691. doi: 10.1093/hmg/5.5.685

Lara, A., Ponce de León, J. L., Rodríguez, R., Casane, D., Côté, G., Bernatchez, L., et al. (2010). DNA barcoding of Cuban freshwater fishes: evidence for cryptic species and taxonomic conflicts. *Mol. Ecol. Resour.* 10, 421–430. doi: 10.1111/j.1755-0998.2009.02785.x

Mallet, J. (2010). Why was Darwin's view of species rejected by twentieth century biologists? *Biol. Philos.* 25, 497–527. doi: 10.1007/s10539-010-9213-7

Mateussi, N. T. B., Pavanelli, C. S., and Oliveira, C. (2017). Molecular identification of cryptic diversity in species of cis-Andean *Mylossoma* (Characiformes: Serrasalmidae). *Mitochondrial DNA A DNA Mapp. Seq. Anal.* 28, 778–780. doi: 10.1080/24701394.2016.1180515

Menezes, N. A., Nirchio, M., Oliveira, C., and Siccha Ramirez, R. (2015). Taxonomic review of the species of Mugil (Teleostei: Perciformes: Mugilidae) from the Atlantic South Caribbean and South America, with integration of morphological, cytogenetic and molecular data. *Zootaxa* 3918, 1–38. doi: 10.11646/zootaxa.3918.1.1

Menezes, N. A., Oliveira, C., and Nirchio, M. (2010). An old taxonomic dilemma: the identity of the western south Atlantic lebranche mullet (Teleostei: Perciformes: Mugilidae). *Zootaxa* 2519, 59–68. doi: 10.5281/zenodo.196200

Nilsen, T. W. (2003). The spliceosome: the most complex macromolecular machine in the cell? *Bioessays* 25, 1147–1149.

Nirchio, M., Cipriano, R. R., Cestari, M. M., and Fenocchio, A. S. (2005a). Cytogenetical and morphological features reveal significant differences among Venezuelan and Brazilian samples of *Mugil curema*. *Neotrop. Ichthyol.* 3, 99–102. doi: 10.1590/S1679-62252005000100006

Nirchio, M., Ron, E., and Rossi, A. R. (2005b). Karyological characterization of *Mugil trichodon* Poey, 1876 (Pisces: Mugilidae). *Sci. Mar.* 69, 525–530. doi: 10.3989/scimar.2005.69n4525

Nirchio, M., and Oliveira, C. (2006). *Citogenética de Peces*. Cumaná: Rectorado de la Universidad de Oriente.

Nirchio, M., Oliveira, C., Ferreira, I. A., Pérez, J. E., Gaviria, J. I., Harrison, I., et al. (2007). Comparative cytogenetic and allozyme analysis of *Mugil rubrioculus* and *M. curema* (Teleostei: Mugilidae) from Venezuela. *Interciencia* 32, 757–762.

Nirchio, M., Oliveira, C., Siccha Ramirez, Z. R., de Sene, V. F., Sola, L., Milana, V., et al. (2017). The *Mugil curema* species complex: a new karyotype for the Pacific white mullet mitochondrial lineage. *Comp. Cytogenet.* 11, 225–237. doi: 10.3897/CompCytogen.v11i2.11579

Nylander, J. A. A. (2008). *MrModeltest v2.3. Program Distributed by the Author*. Uppsala: Uppsala University.

Ocalewicz, K. (2013). Telomeres in Fishes. *Cytogenet. Genome Res.* 141, 114–125. doi: 10.1159/000354278

Ocalewicz, K., and Dobosz, S. (2009). Karyotype variation in the albino rainbow trout (*Oncorhynchus mykiss* Walbaum). *Genome* 52, 347–352. doi: 10.1139/g09-009

Okamoto, N., Maung-Saw-Htoo-Thaw, Sein-Thaung, Koike, K., Shibata, J., and Tomiyama, T. (2017). Species composition of hairtails (Trichiuridae) in Myanmar. *Reg. Stud. Mar. Sci.* 17, 73–77. doi: 10.1016/j.rsma.2017.11.007

Pansonato-Alves, J. C., Serrano, É. A., Utsunomia, R., Camacho, J. P. M., Costa Silva, G. J. D., Vicari, M. R., et al. (2014). Single origin of sex chromosomes and multiple origins of B chromosomes in fish genus Characidium. *PLOS ONE* 9:e107169. doi: 10.1371/journal.pone.0107169

Pendás, A. M., Morán, P., Martínez, J. L., and Garcia-Vásquez, E. (1995). Applications of 5S rDNA in Atlantic salmon, brown trout, and in Atlantic salmon x brown trout hybrid identification. *Mol. Ecol.* 4, 275–276. doi: 10.1111/j.1365-294X.1995.tb00220.x

Pinkel, D., Straume, T., and Gray, J. W. (1986). Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc. Natl. Acad. Sci. U.S.A.* 83, 2934–2938. doi: 10.1073/pnas.83.9.2934

Pomianowski, L., Jankun, M., and Ocalewicz, K. (2012). Detection of interstitial telomeric sequences in the Arctic charr (*Salvelinus alpinus*, Linnaeus 1758) (Teleostei, Salmonidae). *Genome* 55, 26–32. doi: 10.1139/G11-073

Posada, D., and Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818. doi: 10.1093/bioinformatics/14.9.817

Puckridge, M., Andreakis, N., Appleyard, S. A., and Ward, R. D. (2013). Cryptic diversity in flathead fishes (Scorpaeniformes: Platycephalidae) across the Indo-West Pacific uncovered by DNA barcoding. *Mol. Ecol. Resour.* 13, 32–42. doi: 10.1111/1755-0998.12022

Rábová, M., Völker, M., Pelikánová, S., and Ráb, P. (2015). "Sequential chromosome banding in fishes," in *Fish Cytogenetic Techniques – Ray-fin fishes and Chondrichthyans*, eds C. Ozouf-Costaz, E. Pisano, F. Foresti, and L. F. Almeida-Toledo (Boca Raton, FL: CRC Press), 92–102. doi: 10.1201/b18534-15

Ramirez, J. L., Birindelli, J. L., Carvalho, D. C., Affonso, P. R. A. M., Venere, P. C., Ortega, H., et al. (2017). Revealing hidden diversity of the underestimated neotropical ichthyofauna: DNA barcoding in the recently described genus Megaleporinus (Characiformes: Anostomidae). *Front. Genet.* 8:149. doi: 10.3389/fgene.2017.00149

Rocco, L., Morescalchi, M. A., Costagliola, D., and Stingo, V. (2002). Karyotype and genome characterization in four cartilaginous fishes. *Gene* 295, 289–298. doi: 10.1016/S0378-1119(02)00730-8

Rossi, A. R., Crosetti, D., Gornung, E., and Sola, L. (1996). Cytogenetic analysis of global populations of *Mugil cephalus* (striped mullet) by different staining techniques and fluorescent in situ hybridization. *Heredity* 76, 77–82. doi: 10.1038/hdy.1996.10

Rossi, A. R., Crosetti, D., and Livi, S. (2016). "Genetics of Mugilidae," in *Biology, Ecology and Culture of Grey Mullet (Mugilidae)*, eds D. Crosetti and S. J. M. Blaber (Boca Raton, FL: CRC Press), 349–397. doi: 10.1201/b19927-16

Rossi, A. R., Gornung, E., Sola, L., and Nirchio, M. (2005). Comparative molecular cytogenetic analysis of two congeneric species, *Mugil curema* and M. liza, characterized by significant karyotype diversity. *Genetica* 125, 27–32. doi: 10.1007/s10709-005-2668-9

Schultz, L. P. (1946). A revision of the genera of mullets, fishes of the genera Mugilidae, with description of three new genera. *Proc. U. S. Natl. Mus.* 96, 377–395. doi: 10.5479/si.00963801.96-3204.377

Shimabukuro-Dias, C. K., da Costa Silva, G. J., Ashikaga, F. Y., Foresti, F., and Oliveira, C. (2017). Molecular identification of the fish fauna from the pantanal flood plain area in Brazil. *Mitochondrial DNA A DNA Mapp. Seq. Anal.* 28, 588–592. doi: 10.3109/24701394.2016.1149826

Silva, D. M., Utsunomia, R., Pansonato-Alves, J. C., Oliveira, C., and Foresti, F. (2015). Chromosomal mapping of repetitive DNA sequences in five species of *Astyanax* (Characiformes, Characidae) reveals independent location of U1 and U2 snRNA sites and association of U1 snRNA and 5S rDNA. *Cytogenet. Genome Res.* 146, 144–152. doi: 10.1159/000438813

Sola, L., Gornung, E., Mannarelli, M. E., and Rossi, A. R. (2007). "Chromosomal evolution in Mugilidae, Mugilomorpha: an overview," in *Fish Cytogenetics*, eds E. Pisano, C. Ozouf-Costaz, F. Foresti, and B. G. Kapoor (Enfield, NH: Science Publishers), 165–194.

Sumner, A. T. (1972). A simple technique for demonstrating centromeric heterochromatin. *Exp. Cell Res.* 75, 304–306. doi: 10.1016/0014-4827(72)90558-7

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882. doi: 10.1093/nar/25.24.4876

Thomson, J. M. (1966). The grey mullets. *Oceanogr. Mar. Biol.* 4, 301–335.

Thomson, J. M. (1997). The Mugilidae of the world. *Mem. Queensl. Mus.* 43, 457–562.

Utsunomia, R., Silva, D. M. Z. A., Ruiz-Ruano, F. J., Araya-Jaime, C., Pansonato-Alves, J. C., Scacchetti, P. C., et al. (2016). Uncovering the ancestry of B chromosomes in *Moenkhausia sanctaefilomenae* (Teleostei, Characidae). *PLOS ONE* 11:e0150573. doi: 10.1371/journal.pone.0150573

Ward, R. D. (2009). DNA barcode divergence among species and genera of birds and fishes. *Mol. Ecol. Res.* 9, 1077–1085. doi: 10.1111/j.1755-0998.2009.02541.x

Ward, R. D., Costa, F. O., Holmes, B. H., and Steinke, D. (2008). DNA barcoding of shared fish species from the North Atlantic and Australasia: minimal divergence for most taxa, but Zeus faber and *Lepidopus caudatus* each probably constitute two species. *Aquat. Biol.* 3, 71–78. doi: 10.3354/ab00068

Yano, C. F., Poltronieri, J., Bertollo, L. A. C., Artoni, R. F., Liehr, T., and Cioffi, M. B. (2014). Chromosomal mapping of repetitive DNAs in *Triportheus trifurcatus* (Characidae, Characiformes): insights into the differentiation of the Z and W chromosomes. *PLOS ONE* 9:e90946. doi: 10.1371/journal.pone.0090946

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 06 February 2018
doi: 10.3389/fgene.2018.00028

Gene and Blood Analysis Reveal That Transfer from Brackish Water to Freshwater Is More Stressful to the Silverside *Odontesthes humensis*

*Tony L. R. Silveira[1†], Gabriel B. Martins[2†], William B. Domingues[1], Mariana H. Remião[3], Bruna F. Barreto[1], Ingrid M. Lessa[1], Lucas Santos[1], Danillo Pinhal[4], Odir A. Dellagostin[5], Fabiana K. Seixas[3], Tiago Collares[3], Ricardo B. Robaldo[2] and Vinicius F. Campos[1]**

[1] Laboratory of Structural Genomics, Technological Development Center, Federal University of Pelotas, Pelotas, Brazil, [2] Laboratory of Physiology, Institute of Biology, Federal University of Pelotas, Pelotas, Brazil, [3] Laboratory of Cancer Biotechnology, Technological Development Center, Federal University of Pelotas, Pelotas, Brazil, [4] Genomics and Molecular Evolution Laboratory, Department of Genetics, Institute of Biosciences of Botucatu, São Paulo State University, Botucatu, Brazil, [5] Laboratory of Vaccinology, Technological Development Center, Federal University of Pelotas, Pelotas, Brazil

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Daniele Aparecida Matoso, Federal University of Amazonas, Brazil
Ricardo Shohei Hattori, Agência Paulista de Tecnologia dos Agronegócios, Brazil

***Correspondence:**
Vinicius F. Campos
fariascampos@gmail.com

† These authors have contributed equally to this work.

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

Received: *26 October 2017*
Accepted: *22 January 2018*
Published: *06 February 2018*

Citation:
Silveira TLR, Martins GB, Domingues WB, Remião MH, Barreto BF, Lessa IM, Santos L, Pinhal D, Dellagostin OA, Seixas FK, Collares T, Robaldo RB and Campos VF (2018) Gene and Blood Analysis Reveal That Transfer from Brackish Water to Freshwater Is More Stressful to the Silverside Odontesthes humensis. Front. Genet. 9:28. doi: 10.3389/fgene.2018.00028

Silversides are fish that inhabit marine coastal waters, coastal lagoons, and estuarine regions in southern South America. The freshwater (FW) silversides have the ability to tolerate salinity variations. *Odontesthes humensis* have similar habitats and biological characteristics of congeneric *O. bonariensis*, the most studied silverside species and with great economic importance. Studies revealed that *O. bonariensis* is not fully adapted to FW, despite inhabiting hyposmotic environments in nature. However, there is little information about stressful environments for cultivation of silverside *O. humensis*. Thus, the aim of this study was to evaluate the stress and osmoregulation responses triggered by the osmotic transfers on silverside *O. humensis*. Silversides were acclimated to FW (0 ppt) and to brackish water (BW, 10 ppt) and then they were exposed to opposite salinity treatment. Silverside gills and blood were sampled on pre-transfer (D0) and 1, 7, and 15 days (D1, D7, and D15) after changes in environmental salinity, the expression levels of genes *atp1a3a*, *slc12a2b*, *kcnh1*, and *hspa1a* were determined by quantitative reverse transcription-PCR for evaluation of osmoregulatory and stress responses. Furthermore, glycemia, hematocrit, and osmolality were also evaluated. The expression of *atp1a3a* was up- and down-regulated at D1 after the FW–BW and BW–FW transfers, respectively. *Slc12a2b* was up-regulated after FW–BW transfer. Similarly, *kcnh1* and *hspa1a* were up-regulated at D1 after the BW–FW transfer. *O. humensis* blood osmolality decreased after the exposure to FW. It remained stable after exposure to BW, indicating an efficient hyposmoregulation. The glycemia had a peak at D1 after BW–FW transfer. No changes were observed in hematocrit. The return to the pre-transfer levels at D7 after the significant increases in responses of almost all evaluated molecular and blood parameters indicated that this period is enough for acclimation to the experimental conditions. In conclusion, our results suggest that BW–FW transfer is more stressful to *O. humensis* than FW–BW transfer and the physiology of *O. humensis* is only partially adapted to FW.

Keywords: acclimation, blood, brackish water, fish, freshwater, genes, salt, transfer

INTRODUCTION

Odontesthes spp., popularly known as pejerreyes or silversides, is a fish genus naturally endemic from South America waters (Bemvenuti, 2006). This genus comprises the biggest number of species of Atherinopsidae and the broader area of distribution, with species inhabiting marine coastal waters, coastal lagoons, and estuarine regions in southern South America (Dyer, 2006). Currently, some fishes from these species occupy FW environments, despite all *Odontesthes* spp. have a recent common marine origin (Heras and Roldán, 2011; Campanella et al., 2015). Thus, the FW silversides have an interesting ability to tolerate salinity variations. This euryhaline characteristic has drawn attention of researchers to its application in aquaculture in estuarine regions, where there are continuous alterations in water salinity (Piedras et al., 2009).

The most studied species of silverside, with economic importance and favorable projections for production, is *Odontesthes bonariensis*. In Argentina, it has an importance for flesh commercialization and sport fishing (Menone et al., 2000; Somoza et al., 2008). In Japan, *O. bonariensis* cultivation has been deployed, due to the efforts to farming conditions improvements in the last few years (Tsuzuki et al., 2000b, 2001, 2007). This species has also been introduced in Chile, Bolivia, and Italy (Dyer, 2000, 2006). The species is also exotic and has economic importance in Peru, where *O. bonariensis* was introduced in the Titicaca Lake in 1946 and until nowadays continues to be caught (Loubens and Osorio, 1991).

On the other hand, the congeneric *Odontesthes humensis* receives little attention from researchers and fish farmers. Only studies involving fry growth (Pouey et al., 2012), embryos salinity tolerance (Piedras et al., 2009), growth in intensive system (Tavares et al., 2014), and toxicology (Zebral et al., 2017) are developed in captivity specimens. Both *O. bonariensis* and *O. humensis* species have similar habitats, biological characteristics (including morphology), and are capable of interbreeding. Due to this, both species have similar potential for captivity cultivation (Dyer, 2006; Piedras et al., 2009). Main differences between both species are: *O. bonariensis* is the largest species in Atheriniforms and it has 32–38 gill rakers on lower branch, vomerine teeth present, and the prognathism can be present. Furthermore, this species is native to Brazil and Argentina, but not to Uruguay. On the other hand, *O. humensis* is smaller, has 13–19 gill rakers on lower branch, vomerine teeth is absent, prognathism is never present, and the species is native to Brazil, Uruguay, and Argentina (Dyer, 2006). Although the commercial production of *O. humensis* has not increased in recent years, this species has potential to be a bioindicator in toxicological studies (Zebral et al., 2017) due to its demands for water quality, surviving only in a narrow range of water parameters.

Thus, information about basic biology of *O. humensis* is necessary to the adequate maintenance of this silverside species in farm and laboratory conditions. Some studies revealed that the silverside *O. bonariensis* is not fully adapted to FW yet, despite inhabiting hyposmotic environments in nature (Tsuzuki et al., 2000b, 2001). However, there is no information about the less stressful environment, if with or without salt, to the cultivation of silverside *O. humensis*. We hypothesized that *O. humensis* shared, besides the habitats and reproductive physiology, the same pattern of response of *O. bonariensis* to environmental salinity. The corroboration of this hypothesis will help in maintenance of laboratorial or commercial cultivation, in addition to reinforcing the theory of invasion of FW habitats by silversides marine ancestors. Thus, the aim of this study was to evaluate, using cellular, biochemical, and molecular analysis, the stress and osmoregulation responses triggered by the osmotic transfers in silverside *O. humensis*.

Abbreviations: BW, brackish water; FW, freshwater; SW, seawater.

MATERIALS AND METHODS

Animals and Conditions

The silversides *O. humensis* used in this study came from eggs collected in nature (Arroio Grande, Brazil – 32°14′15″S/53°05′13″W) and hatched in tanks. The eggs were collected from wild breeders and incubated at laboratory conditions in FW (0‰) and pH values near 7.5, the same conditions applied to larviculture. The fish were 1.5 years old and have mean weight of 27.3 ± 9.5 g at experimental time. The silversides were maintained in 1,000 L cylindrical plastic tanks within an experimental room, fed three times a day with commercial feed (Supra, 38% crude protein) until satiety, under natural photoperiod of autumn (11 h light/13 h dark), with water pH 7.5 ± 0.3, and 19 ± 0.3°C for acclimation. The tank sides were opaque to reduce visual stress and incidence of luminosity through the top according to the daily variance. The acclimation period was 4 weeks for silversides in FW (salinity of 0 ppt) and in BW (salinity of 10 ppt), under the same previous conditions. The salinity levels were achieved through the dissolution of non-iodized sea salt in the water.

The fish continued to be fed three times a day during the experimental period. Once a week the water of the tanks was totally renewed with dissolved oxygen maintained in 7.8 ± 0.3 mg L^{-1}; the total ammonia levels lower than 0.6 ± 0.2 mg L^{-1}; temperature 20.01 ± 0.34°C; and pH 7.78 ± 0.03.

Experimental Design and Protocol

The experimental setup consisted in two treatments performed in quadruplicate, totaling eight tanks with 15 fish each. Four tanks contained silversides acclimated to FW and other four tanks contained silversides acclimated to BW. Both FW and BW groups were shortly transferred to opposite salt treatment. Fish were sampled at four different time points getting three fish per tank: a control before the osmotic transfer (D0) and on day 1 (D1), on day 7 (D7), and on day 15 (D15) after the transfer. Hence, 96 silverside fish were analyzed in different times of salt exposure.

Sampling

When captured, at all time points, the silversides were anesthetized with benzocaine (50 mg L^{-1}). Blood samples

were collected from caudal puncture using a heparinized syringe for hematocrit, glycemia, and osmolality evaluation. While anesthetized, the fish were euthanized by medullary section and brain excision to perform the post-mortem gills collection. Gill samples from three fish of each tank, at the four time points, were collected using sterilized material. Left second branchial arches were collected and immediately stored in liquid nitrogen (N_2) until posterior RNA preparation. All experimental procedures were previously approved by the Ethics Committee on Animal Experimentation of Federal University of Pelotas (Process no. 23110.007018/2015-85).

Analyzed Genes

The Na^+/K^+-ATPase (NKA) enzyme is responsible to active transports three Na^+ out of and two K^+ into animal cells. In fishes, NKA is directly involved in the osmoregulation process (Hiroi and McCormick, 2012). Among NKAα1∼α4, we selected NKAα3 (encoded by *atp1a3*) which has similar responses (Stefansson et al., 2007; Whitehead et al., 2012; Armesto et al., 2014; Li et al., 2014; Taugbøl et al., 2014) and is expressed in similar tissues (Guynn et al., 2002) as NKAα1. NKAα1 is the most studied isoform of NKA (Havird et al., 2013), however, its mRNA expression can be influenced by the stress (Kiilerich et al., 2011) and this could disrupt the interpretation of results.

The $Na^+/K^+/2Cl^-$ cotransporter 1 (NKCC1) mediates the intake of one Na^+, one K^+, and two Cl^- ions into the ionocytes down the electrochemical gradient generated by NKA (Hwang et al., 2011). In teleosts there are two isoforms of NKCC1 and NKCC2: NKCC1a/b and NKCC2a/b (Kakumura et al., 2015). NKCC1 (encoded by *slc12a2*) is more associated with ion secretion in teleost gills (Tipsmark et al., 2002; Nilsen et al., 2007), whereas NKCC2 mRNA is not present in large amount in fish gills (Hiroi and McCormick, 2012).

The voltage-gated potassium channels (KCNH) are gated by changes in membrane potential and play an important role in the repolarization of cells (Morais-Cabral and Robertson, 2015). KCNH1 has already been correlated to myoblast fusion (Occhiodoro et al., 1998), to embryonic development (Stengel et al., 2012), and has been used as target to cancer researches (Pardo and Stühmer, 2013) but the precise role of KCNH1 in normal tissues is poorly known. The *kcnh1* gene has two copies in fishes, *kcnh1a*/*b* that are expressed differently in the organism (Stengel et al., 2012). KCNH4 has already been proposed as an osmoregulation marker (Taugbøl et al., 2014).

The heat shock 70kDa proteins (HSP70) are mainly responsible to protein folding, degradation of misfolded proteins, membrane translocation, and activation of immune system. HSP70 are expressed under normal conditions but also in response to stress conditions (Morimoto, 1998; Yamashita et al., 2010; McCallister et al., 2016). The genes encoding HSP70 in fishes are *hspa1a* and *hspa1b* (He et al., 2013). Thus, both *hspa1* genes and the HSP70 protein are considered stress indicators in fish (Basu et al., 2001; Larsen et al., 2012; Taugbøl et al., 2014).

Molecular Cloning and Sequencing

Total RNA was isolated from gill samples of each fish separately, using RNeasy® Mini Kit (Qiagen, United States) according to the manufacturer's protocol. DNase treatment of RNA samples was conducted with a DNA-Free® Kit (Ambion, United States) following the manufacturer's protocol. The RNA concentration and quality were verified by spectrophotometry using NanoVue™ Plus (GE Healthcare Life Sciences, United states), and only samples with absorption A_{260}/A_{280} ratio ≥2.0 were used in reverse transcription (RT) reactions.

First-strand cDNA was performed with 2 μg of RNA using random primers and SuperScript® III Reverse Transcriptase (Invitrogen, United States) according to the manufacturer's protocol. Primer sets (**Table 1**) were designed to clone partial cDNA sequences of *atp1a3*, *slc12a2*, *kcnh1*, and *hspa1* genes from gills of silverside fish based on the alignment of sequences of other fish species. The PCR parameters were: an initial denaturation for 1 min at 94°C, followed by 35 cycles of 94°C for 30 s; 57°C (for *atp1a3*); 62°C (for *slc12a2*); 64.6°C (for *kcnh1*); and 61°C (for *hspa1*) for 30 s; and 72°C for 1 min, with a final extension of 5 min at 72°C. PCR products were inserted into pCR™4-TOPO® TA cloning vector and transformed in the electrocompetent *Escherichia coli* strain DH5α. These procedures were performed according to the manufacturer specifications

TABLE 1 | Sequences of primers for *Odontesthes humensis* used in this study.

Gene symbol	Sequence 5′→3′	Use
atp1a3a		
	F: AACCCCAGAGATGCCAA	Cloning
	R: AAGGCACAGAACCACCA	
	F: ATACCGGGGCAGTAGGAGAG	qPCR
	R: CAGTATCGTGGTGGTGCAGT	
slc12a2b		
	F: GGRATGGARTTGGARGCHAARGC	Cloning
	R: CCDGABACCAGGCTCATBACCTG	
	F: GCCATCCTCATCACAGGATT	qPCR
	R: ACTGTGTCGCTCATCGTGTC	
kcnh1		
	F: CAGCACCTGCAGCTTCATG	Cloning
	R: CCACRATGCTGTCYACCACCAGCC	
	F: TTGTCGGGGTACCATAGAGC	qPCR
	R: AGTCGCACTTTTTCCGTTGT	
hspa1a		
	F: TCTGCAGCTAAAGGTGTAGC	Cloning
	R: TTGAAGGGCCAGTGCTTCATG	
	F: TCCAGCACGGAAAAGTAGAGA	qPCR
	R: ACAGTGTTGCTGGGGTTCAT	
actb		
	F: CTCTGGTCGTACCACTGGTATCG	qPCR reference gene
	R: GCAGAGCGTAGCCTTCATAGATG	
h3a		
	F: CGTGGCTCTGAGAGAGATCC	qPCR reference gene
	R: TCGGTCTTGAAGTCCTGAGC	

(Invitrogen, United States). The positive clones were sequenced using a Big Dye® V3.1 Terminator Kit and M13 Primers (Applied Biosystems, United States). The cloned fragments were sequenced using an Applied Biosystems 3500 Genetic Analyzer® automatic sequencer (Life Technologies, United States).

In Silico and Phylogenetic Analysis

The identity of all sequenced cDNA was confirmed by Blast tool of the NCBI database[1]. The translation of sequenced nucleotides to amino acid sequences as well as the open-reading frame (ORF) identification was made using ExPASy bioinformatics resource portal[2]. Conserved domains and sites were mapped from UniProt database[3]. Amino acid sequences of NKA, NKCC, KCNH, and HSP70 from various vertebrate species were obtained from GenBank and aligned to the new deduced amino acid sequences of *O. humensis* using ClustalW. Phylogenetic analyses were carried out using MEGA version 6 (Tamura et al., 2013). Phylogenetic trees were constructed based on alignments of amino acid sequences using Neighbor-Joining method. Bootstrap analyses were done with 10,000 replications to test the tree reliability. The outgroups were *Caenorhabditis elegans* (for NKA and HSP70 analysis), *Strongylocentrotus purpuratus* (for NKCC analysis), and *Toxocara canis* (for KCNH analysis).

Evaluation of Gene Expression after Salinity Changes

Quantitative RT-PCR (qRT-PCR) was run on CFX96™ Real-Time PCR Detection System (Bio-Rad Laboratories, United States) using SYBR® Green PCR Master Mix (Applied Biosystems, United States). Primers (**Table 1**) for the identified genes *atp1a3*, *slc12a2*, *kcnh1*, and *hspa1* and for the used reference genes β-actin (*actb*, GenBank Accession No. EF044319) and histone h3a (*h3a*, GenBank Accession no. KX060037) were designed using Primer3 online software[4]. Initial validation experiments were conducted to ensure that all primer pairs had equivalent PCR efficiencies calculated by standard dilution curve. Amplification was carried out at the standard cycling conditions of 95°C for 10 min, followed by 40 cycles at 95°C for 15 s, 60°C for 60 s followed by conditions to calculate the melting curve. All PCR runs for each cDNA sample were performed in triplicate. The qRT-PCR data were analyzed using the $2^{-\Delta\Delta C_t}$ method considering primer amplification efficiencies (Livak and Schmittgen, 2001).

Blood Analysis

After blood collection, with fasted animals, the glycemia was measured using Accu Chek Glucometer (Roche Diagnostics, United Kingdom). The samples from each fish were immediately transferred to microcapillaries and centrifuged at 12,000 × *g* for 5 min to hematocrit analysis. The blood was transferred to sterilized microtubes and centrifuged at 1,500 × *g* at 4°C for 10 min to plasma separation for osmolality measurements in Vapro 5520® Vapor Pressure Osmometer (Wescor, United States).

Statistical Analysis

Quantitative data were expressed as means ± standard error of mean. Significant differences among means were evaluated by two-way ANOVA followed by Tukey's test, setting the significance level at 95% ($P < 0.05$). The exception was the expression of reference genes, which was expressed as means of cycle threshold (Ct) values ± standard deviation of mean and the differences were analyzed by one-way ANOVA followed by Tukey's test.

RESULTS

cDNA Cloning and Characterization

The *atp1a3*, *slc12a2*, *kcnh1*, and *hspa1* cloned fragments from *O. humensis* were, respectively, 975, 423, 560, and 286 bp (base pairs) in length. They were sequenced and deposited under GenBank accession numbers KR920364, KT001464, KX035016, and KU639716, respectively. The cloned fragments of *atp1a3* and *kcnh1* belong, respectively, to the middle of the ORF +2 and to the initial part of the ORF +3. The fragments of *slc12a2* and *hspa1* are part of the center of ORF +1. The *atp1a3* fragment codes to 324 amino acids belonging to the P-type ATPase family. The *slc12a2* codes to 141 amino acid residues belonging to the Na/K/Cl co-transporter 1 family. The cloned fragment of *kcnh1* codes to 186 amino acids belonging to the potassium channel, voltage-dependent family. The cloned fragment of *hspa1* codes to 95 amino acids belonging to the heat shock protein 70 family.

The percentage identity values between NKAα3, NKCC1, KCNH, and HSP70 putative sequences of silverside and of the other analyzed species were, respectively, 77–98, 55–92, 26–99, and 37–97% (**Supplementary Figures S1**–**S4**). The phylogenetic tree constructed based on alignment of NKAα-subunits amino acid sequences (**Supplementary Figure S1**) reveals that α1, α2, α3, and α4 form different clusters with mammal and fish sequences in each, except in α4 group, which was devoid of NKA sequences of fishes. Furthermore, the silverside NKAα3 was grouped in the monophyletic α3 cluster, showing that it was more closely related to NKAα3a than NKAα3b and NKAα2. Based on these evidences, the sequence found in the gills of *O. humensis* was identified as *atp1a3*α. NKCC1 and NKCC2 were grouped in different clusters (**Supplementary Figure S2**).

The paralogous NKCC1a and NKCC1b were grouped together in NKCC1 group. Furthermore, the tree has shown more evidence that silverside NKCC is closer to fish NKCC1b than to NKCC1a or NKCC2. Thus, the cloned and sequenced cDNA from silverside gills was identified as *slc12a2b*. The KCNH phylogenetic tree (**Supplementary Figure S3**) forms two monophyletic main groups: ELK family cluster, composed by vertebrates KCNH3, KCNH4, and KCNH8, and EAG family cluster, composed by vertebrates KCNH5 and KCNH1, in which the new sequence of silverside KCNH1 was grouped. Based on this analysis, the cloned KCNH sequence from silverside gills was confirmed as *kcnh1*.

[1]http://blast.ncbi.nlm.nih.gov/

[2]https://www.expasy.org/

[3]http://www.uniprot.org/

[4]http://primer3.ut.ee

The HSP70 phylogenetic analysis (**Supplementary Figure S4**) separated HSPA1 and HSPA2 sequences in different clusters. HSPA2 was composed only by sequences of mammals while HSPA1 was composed by mammals and teleosts. The new sequence of HSPA1 from silverside was clustered within fish HSPA1a. Based on these results, the HSPA sequence from silverside gills was identified as *hspa1a*.

Gene Expression after Salinity Changes

Primer sets for *atp1a3a*, *slc12a2b*, *kcnh1*, *hspa1a*, *actb*, and *h3a* had amplification efficiencies in qRT-PCR of 1.14, 1.07, 1.06, 1.02, 0.98, and 1.05, respectively. No difference was observed in the expression of reference genes between four different time points or between the treatments (**Supplementary Table S1**). The relative expression of *atp1a3a* (**Figure 1A**) was not different ($P > 0.05$) between FW- and BW-acclimated fishes in control D0. However, after the transfer of FW-acclimated fish to BW (FW–BW) and vice versa (BW–FW), a change in *atp1a3a* expression was observed. In FW–BW-transferred fish, there was an increase ($P < 0.05$) in *atp1a3a* expression, while in BW–FW there was a decrease at D1 ($P < 0.05$). The levels of *atp1a3a* expression returned close to the initial conditions in both FW–BW and BW–FW groups, without difference between groups ($P > 0.05$), at D7 and D15.

The relative expression of *slc12a2b* (**Figure 1B**) was not different between FW- and BW-acclimated fish at D0 ($P > 0.05$). However, after the FW–BW and BW–FW transfers, a change in *slc12a2b* mRNA expression was observed. In FW–BW-transferred fish, there was an increase ($P < 0.05$) in *slc12a2b* mRNA levels at D1. The gene expression decreased ($P < 0.05$) to intermediary levels at D7 and remained without difference until D15 ($P > 0.05$). The *slc12a2b* expression levels did not change ($P > 0.05$) after the transfer in the BW–FW group in any experimental time. Difference in *slc12a2b* relative expression between the BW–FW and FW–BW groups ($P < 0.05$) was observed at D1 and D7.

None significant difference was observed between the relative expression of *kcnh1* (**Figure 1C**) at D0 from FW- or BW-acclimated silversides ($P > 0.05$). After the transfers, at D1, it was observed a significant increase ($P < 0.05$) in the *kcnh1* expression of the BW–FW group. The expression levels decreased ($P < 0.05$), returning to the initial patterns at D7 and D15. The *kcnh1* expression levels did not change significantly ($P > 0.05$) after the transfer in the FW–BW group in any experimental time. Difference in *kcnh1* relative expression between the BW–FW and FW–BW groups ($P < 0.05$) was observed only at D1.

The relative expression of *hspa1a* (**Figure 1D**) was not different between FW- and BW-acclimated fish at D0 ($P > 0.05$). After the transfers, at D1, it was observed a significant increase ($P < 0.05$) in the *hspa1a* relative expression of the BW–FW-transferred fish. The expression levels decreased ($P < 0.05$) to initial pattern at D7 and remained unchanged until D15 ($P > 0.05$). The *hspa1a* expression levels did not change ($P > 0.05$) after the transfer in the FW–BW group in any experimental time. Difference in *hspa1a* relative expression between the BW–FW and FW–BW groups ($P < 0.05$) was observed only at D1.

Glycemia

No difference in blood glucose concentration ($P > 0.05$) was observed between FW- or BW-acclimated fish at D0 (**Figure 2A**). The BW–FW-transferred fish presented an increase ($P < 0.05$) in glycemia values at D1, a decrease ($P < 0.05$) to pre-transfer levels at D7, and remained without difference ($P > 0.05$) until D15. None difference ($P > 0.05$) between time-points was observed in FW–BW transfer group. Significant difference between both groups was observed only at D1.

Hematocrit

No difference ($P > 0.05$) between the hematocrit (**Figure 2B**) in none group of analyzed silverside of experimental time was observed. Only trends of increase in hematocrit soon after BW–FW transfer and decrease after FW–BW were verified.

Osmolality

No difference ($P > 0.05$) between the blood osmotic concentration (**Figure 2C**) of FW- and BW-acclimated fishes was observed at D0. After transfer, there was a reduction ($P < 0.05$) in osmolality of BW–FW-transferred fish that was kept the same until D15. Differences were not observed in blood osmolality of FW–BW-transferred fish. Differences ($P < 0.05$) between groups were found right after the transfer (D1) and at the final measurements (D15).

DISCUSSION

Salinity is a frequent abiotic stressor that restrains fish growth and development by inducing osmotic stress responses. When fishes are subjected to salinity stress, related genes are activated to induce salinity stress tolerance. This study was successful in the identification and characterization of *atp1a3a*, *slc12a2b*, *kcnh1*, and *hspa1a* mRNA sequences in the silverside *O. humensis*. The molecular cloning made possible to adequately monitor changes in the expression of these genes after salinity challenges. In addition, it was demonstrated that alterations in blood glucose and osmolality after environmental salinity change.

In FW–BW-transferred fish, *atp1a3a* expression increased soon upon salinity stress, whereas in BW–FW group it was downregulated. Similar effects in *atp1a3a* expression were also observed in gills of Mozambique tilapia challenged by salinity stress, which presented higher expression levels in SW and lower in FW (Li et al., 2014). In threespine stickleback, the SW transfer also generated an increase in *atp1a3a* mRNA expression (Taugbøl et al., 2014). The increase of mRNA expression and protein activity of NKAα-subunit in gills, in response to the SW transfer, occurred in Atlantic salmon (D'Cotta et al., 2000), in brown trout (Tipsmark et al., 2002), and in killifish (Scott and Schulte, 2005; Scott et al., 2008). Yet, the transfer to lower salinity generates a decrease in mRNA expression and protein activity of NKAα-subunit in gills of Mozambique tilapia (Lin et al., 2004). In the present study, the downregulation in *atp1a3a* expression

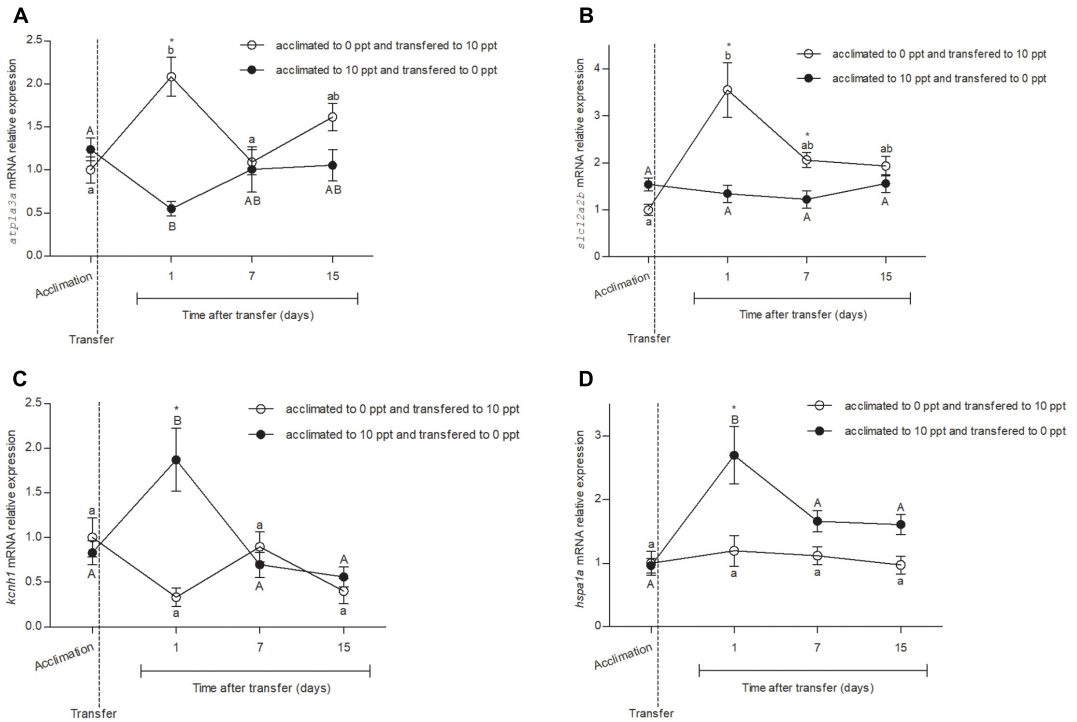


FIGURE 1 | Relative expression of cloned genes *atp1a3a*, *slc12a2b*, *kcnh1*, and *hspa1a* from *Odontesthes humensis* submitted to salinity transfer. **(A)** *Atp1a3a* mRNA relative expression before and after the salinity change. **(B)** *Slc12a2b* mRNA relative expression before and after the salinity change. **(C)** *Kcnh1* mRNA relative expression before and after the salinity change. **(D)** *Hspa1a* mRNA relative expression before and after the salinity change. Mean ± SEM (n = 8–10). Different letters in each group (lowercase to FW–BW transferred group and uppercase to BW–FW transferred group) represent significant difference between the measurements in a same group. The asterisk represents the significant difference between groups in a same experimental time. The vertical broken line indicates the transfer from BW to FW and from FW to BW moments.

indicates a reduction in osmoregulatory activity due the lack of salt in water.

In silverside, *slc12a2b* expression in gills and in response to increase of salinity was well detectable, evidencing the high importance of NKCC1b to hyposmoregulation of *O. humensis*. While NKA can have activity in ion absorption and secretion, NKCC1 is more associated with ion secretion in teleost gills. Here, the augmented water salt concentration leads *O. humensis* to increase its *slc12a2b* mRNA expression when FW–BW transferred. This fact, in association with the blood osmolality stabilization even when transferred to BW, indicates an efficient ion secretion activity. Furthermore, the transfer to lower salinity media did not affect the *slc12a2b* mRNA expression. Despite the opposite responses of NKCC in hypo- or hyperosmotic environment are the most common reported in fishes, none variations after salinity changes is also possible, especially in transfers for low salinities. In gills of brown trout (Tipsmark et al., 2002) and striped bass (Tipsmark et al., 2004), both gene and protein expression of NKCC1 were positively correlated to environmental salinity, with increases after FW–SW transfer and decreases after SW–FW transfer. In brackish medaka (Kang et al., 2010) and climbing perch (Loong et al., 2012), the *slc12a2b* mRNA expression and NKCC protein quantity followed the saline concentration of water inhabited by FW- and SW-acclimated fish. However, no response was reported in *slc12a2a* mRNA expression after the transfer of Mozambique tilapia from BW–FW (Breves et al., 2010).

The branchial tissue of silversides of this study had an increase in *kcnh1* mRNA expression soon after transfer to FW. A similar result was obtained in threespine stickleback, in which the *kcnh4* mRNA expression was higher in FW- than in SW-acclimated fish (Taugbøl et al., 2014). Furthermore, the return of mRNA expression to pre-transfer levels indicates an acclimation to media. Although *kcnh1* functions in organisms are not well known, our results reveal that this gene may play a role in osmoregulation.

The HSP70 overexpression is induced by some environmental stressors, such as temperature changes, UV and γ-irradiation, and chemical exposure (Place and Hofmann, 2005; Yamashita et al., 2010; Rajeshkumar et al., 2013). In *O. humensis*, the salinity change was able to induce a response of *hspa1a* gene. The same was observed in threespine stickleback (Taugbøl et al., 2014), in North Sea cod and Baltic Sea cod (Larsen et al., 2012), and in Kaluga (Peng et al., 2016). The up-regulation in *hspa1a* mRNA after BW–FW transfer indicates an increase in stress. The significant down-regulation observed after 7 days in FW indicates a reduction in stress levels, indicating acclimation. On the other hand, the unchanged *hspa1a* mRNA levels after the FW–BW transfer reveal that this salinity challenge is less stressful to *O. humensis*.

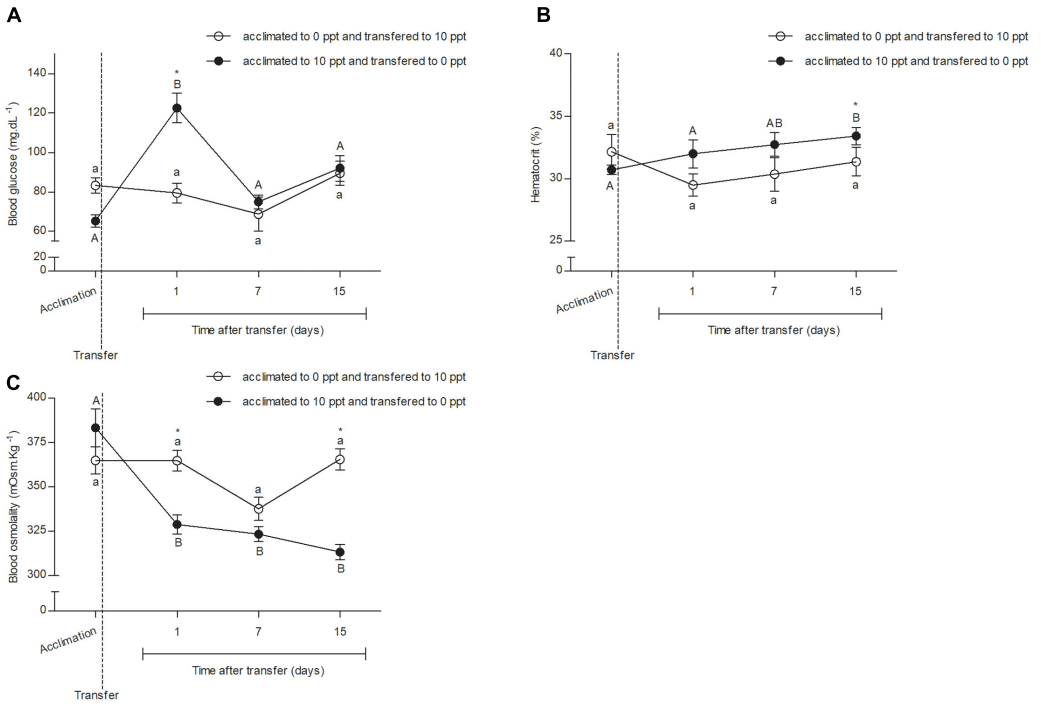


FIGURE 2 | Blood responses from *O. humensis* submitted to salinity transfer. **(A)** Glycemia before and after the salinity change. **(B)** Hematocrit before and after the salinity change. **(C)** Blood osmolality before and after the salinity change. Mean ± SEM (n = 12). Different letters in each group (lowercase to FW–SW transferred group and uppercase to SW–FW transferred group) represent significant difference between the measurements in a same group. The asterisk represents the significant difference between groups in a same experimental time. The vertical broken line indicates the transfer from SW to FW and from FW to SW moments.

Blood glucose has been used to estimate acute stress conditions in fishes (Cataldi et al., 2005), as reported to silverside *O. bonariensis* (Tsuzuki et al., 2001). Furthermore, the hematocrit levels also can serve as stress indicator in fishes. The increment in hematocrit may indicate the presence of an active stress factor (Hudson et al., 2008). The quick increase in glycemia and the trend of increase in hematocrit, allied to the increase in *hspa1a* expression after BW–FW transfer, reveal that this "salt-free" environment is significantly more stressful to *O. humensis* than the BW medium. In *O. bonariensis*, the transfer from FW to 20 ppt BW triggered a decrease in stress followed by a decrease in glycemia and hematocrit, indicating that FW is more stressful than BW also in this species (Tsuzuki et al., 2001). Moreover, the unchanged blood glucose and hematocrit corroborates to *hspa1a* expression results, which indicates that FW–BW transfer leads to lower levels of stress in *O. humensis*.

Reduction in *atp1a3a* expression and blood osmolality in BW–FW-transferred silversides indicates a decrease in osmoregulation activity, which is energetically expensive to fishes. Even so, the stress in BW–FW group was higher, suggesting that the physiology of *O. humensis* is only partially adapted to FW. In the FW–BW transfer group, the blood osmolality maintains its levels. In *O. bonariensis*, the transfer to 20 ppt BW generates a significant increment in osmolality until 24 h post-transfer (Tsuzuki et al., 2001, 2007). The 5 ppt BW transfer also generated an increase in blood osmolality, but not as significant as 20 ppt (Tsuzuki et al., 2007). Here, *O. humensis* osmolality remained without significant variation in FW–BW. Maybe the 10 ppt salt concentration medium was insufficient to increase significantly the blood osmolality of *O. humensis*, despite the half salt concentration is able to increase this parameter of *O. bonariensis*. However, we hypothesized that FW–BW transference causes increased recruitment of both *atp1a3a* and *slc12a2b*, which would work together in the control of ion secretion and stabilization of blood osmolality. This may indicate that the silverside *O. humensis* is more efficient than *O. bonariensis* in the hyposmoregulation process.

The return of osmoregulatory genes expression and blood glucose close to pre-transfer levels after significant increase indicates an efficient acclimation (Grutter and Pankhurst, 2000; Havird et al., 2013). The blood glucose and *atp1a3a*, *slc12a2b*, *kcnh1*, and *hspa1a* mRNA stabilized in 7 days. These markers of osmoregulation and stress corroborated acclimation after 1 week despite the high stress occurred after the saline challenge.

Our results suggest that the BW–FW transfer is more stressful to *O. humensis* than the FW–BW transfer. Even though *O. humensis* is not widely cultivated in aquaculture today, the information that this species follows the responses of *O. bonariensis* and remains less stressed when in brackish environments is very important. Salinity levels were shown to modulate the energy supply available for growth and reproduction in farmed fishes (Altinok and Grizzle, 2001; Chand et al., 2015), and its optimal adjustment can benefit *O. humensis* production in captivity for aquaculture purposes.

The brackish medium has potential for decrease in economic losses by mortality due to handling, transport, crowding, and poor water quality (Strüssmann et al., 1996; Tsuzuki et al., 2000b, 2001) and increase the survival rate of embryos in farm production (Piedras et al., 2009). Keeping *O. humensis* in this near-isosmotic environment can potentially allow for better growth rates, food conversion ratio, energy absorption efficiency, among other parameters of interest. Furthermore, the cultivation of *O. humensis* in conditions closer to the ideal decreases the stress and the physiological variations due to it and favors the use of this species as a model in scientific research.

Even regarded as FW, these species are commonly in contact with salt and BW on estuary and coastal lagoons of South America and have better development and survival in saline environments (Tsuzuki et al., 2000a, 2001; Piedras et al., 2009). The coastal plain of southern Brazil system was originated from successive transgressions and regressions moves since the upper Pleistocene (Bemvenuti, 2006). The formation and radiation of *Odontesthes* spp. is also recent and occurs during Pleistocene–Holocene (Beheregaray et al., 2002; Lovejoy et al., 2006; Heras and Roldán, 2011). Furthermore, there are strong evidences that various genus of Atherinopsidae order, including *Odontesthes*, have an evolutionary tendency to invade the continental FWs (Bloom et al., 2013; Campanella et al., 2015). Thus, the same events that originated the Southeastern South America continental lakes and lagoons may have triggered the speciation process of *Odontesthes* spp.

The previous and the present study corroborate to the theory that the FW is not the ideal environment for *O. humensis* and give more arguments to the theory of repeated invasion of FW habitats by silversides marine ancestors. The conquest of the FW environment is very recent, so there has not been enough time for the selection of adaptive mechanisms for life in this condition without the elimination of some basal stress levels.

AUTHOR CONTRIBUTIONS

TS, GM, and VC were responsible for experimental design, data analysis, and manuscript writing. TS, GM, MR, and RR were responsible for fish acclimation and maintenance. TS, GM, BB, IL, LS, MR, RR, and WD were responsible for the biological collections. TS, BB, IL, LS, WD, and VC were responsible for the molecular biology, from RNA extraction to sequencing, and qRT-PCR analysis. OD, TC, and FS were also responsible for qRT-PCR analysis. DP and RR were also responsible for data analysis and language review.

FUNDING

This study was supported by the Ministério da Ciência (422292/2016-8), Tecnologia e Inovação/Conselho Nacional de Desenvolvimento Científico e Tecnológico (Edital Universal No. 472210/2013-0) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (AUXPE 2900/2014). TS, GM, WD, and MR are individually supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. DP, OD, TC, FS, RR, and VC are also individually supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico.

ACKNOWLEDGMENTS

We are greatly thankful to Mr. Nilton Link and Mr. Vítor Colvara for the help in the maintenance of animals and Janaína Pedron for the help in blood osmolality analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00028/full#supplementary-material

FIGURE S1 | Phylogenetic analysis of *NKAα*-subunit amino acid sequences of mammals and teleosts. The tree was generated by MEGA v6 software using the Neighbor-Joining method. The bootstrap values from 10,000 replicates are showed at each node. Values between parentheses represent the identity (in %) of the sequence with *NKAα3a* of *Odontesthes humensis*. Scale bar units represent the number of amino acid substitutions per site. *NKAα1* GenBank ID: *Homo sapiens*, CAA27840.1. *NKAα1a* GenBank ID: *Danio rerio*, NP_571762.1; *Solea senegalensis*, BAN17690.1; *Fundulus heteroclitus*, AAL18002.1; *Oncorhynchus masou*, BAJ13363.1; *Oncorhynchus mykiss*, NP_001117933.1. *NKAα1b* GenBank ID: *O. mykiss*, NP_001117932.1; *O. masou*, BAJ13362.1; *S. senegalensis*, BAN17691.1; *D. rerio*, NP_571765.1. *NKAα2* GenBank ID: *Rattus norvegicus*, NP_036637.1; *O. mykiss*, NP_001117930.1; *D. rerio*, AAF98359.1; *S. senegalensis*, BAO02373.1; *F. heteroclitus*, AAL18003.1. *NKAα3* GenBank ID: *H. sapiens*, NP_689509.1; *R. norvegicus*, NP_036638.1; *Oreochromis mossambicus*, AAF75108.1; *O. mykiss*, NP_001118102.1. *NKAα3a* GenBank ID: *D. rerio*, NP_571759.2; *S. senegalensis*, BAN17693.1; *Monopterus albus*, AGV06213.1. *NKAα3b* GenBank ID: *M. albus*, AGV06214.1; *S. senegalensis*, BAN17692.1. *NKAα4* GenBank ID: *H. sapiens*, Q13733.3; *R. norvegicus*, NP_074039.1. *NKAα1* from *Caenorhabditis elegans*, NP_506269.1, was used as outgroup.

FIGURE S2 | Phylogenetic analysis of *NKCC* amino acid sequences of mammals and teleosts. The tree was generated by MEGA v6 software using the Neighbor-Joining method. The bootstrap values from 10,000 replicates are showed at each node. Values between parentheses represent the identity (in %) of the sequence with *NKCC1b* of *O. humensis*. Scale bar units represent the number of amino acid substitutions per site. *NKCC1* GenBank ID: *H. sapiens*, P55011.1; *Mus musculus*, NP_033220.2. *NKCC1a* GenBank ID: *Anguilla Anguilla*, CAD31111.1; *Salmo salar*, NP_001117155.1; *Oryzias dancena*, ADN18710.1; *O. mossambicus*, AAR97731.1; *Anabas testudineus* AFK29496.1. *NKCC1b* GenBank ID: *O. dancena*, ADK47392.1; *M. albus*, AGX01628.1; *O. mossambicus*, AAR97732.1; *A. anguilla*, CAD31112.1. *NKCC2* GenBank ID: *H. sapiens*, Q13621.2; *M. musculus*, P55014.2; *O. mossambicus*, AAR97733.1; *Takifugu obscurus*, BAH20440.1. *NKCC1* from *Strongylocentrotus purpuratus*, NP_001106707.1, was used as outgroup.

FIGURE S3 | Phylogenetic analysis of *KCNH* amino acid sequences from EAG and ELK subfamilies of mammals and teleosts. Third existent ERG subfamily was discard in order to maintain an abbreviated analysis. The tree was generated by MEGA v6 software using the Neighbor-Joining method. The bootstrap values from 10,000 replicates are showed at each node. Values between parentheses represent the identity (in %) of the sequence with *KCNH1* of *O. humensis*. Scale bar units represent the number of amino acid substitutions per site. *KCNH1* GenBank ID: *H. sapiens*, NP_758872.1; *M. musculus*, NP_034730.1; *D. rerio*, XP_009291371.1; *Oreochromis niloticus*, XP_005474450.1. *KCNH3* GenBank ID: *H. sapiens*, NP_001300959.1; *D. rerio*, XP_001919436.3; *Nothobranchius furzeri*, SBP57798.1. *KCNH4* GenBank ID: *H. sapiens*, NP_036417.1; *P. reticulata*, XP_008415281.2; *O. niloticus*, XP_019213276.1; *Maylandia zebra*, XP_012774834.1. *KCNH5* GenBank ID: *H. sapiens*, NP_647479.2; *M. musculus*,

NP_766393.2; *D. rerio*, NP_001263209.1; *O. niloticus*, XP_003451242.1. *KCNH8* GenBank ID: *H. sapiens*, NP_653234.2; *M. musculus*, NP_001026981.2; *O. niloticus*, XP_003448945.1; *M. zebra*, XP_004547937.1; *Poecilia reticulata*, XP_008419899.1. *KCNH1* from *Toxocara canis*, KHN74999.1, was used as outgroup.

FIGURE S4 | Phylogenetic analysis of *HSP70* amino acid sequences of mammals and teleosts. The tree was generated by MEGA v6 software using the Neighbor-Joining method. The bootstrap values from 10,000 replicates are showed at each node. Values between parentheses represent the identity (in %) of the sequence with *HSP70* of *Odontesthes humensis*. Scale bar units represent the number of amino acid substitutions per site. *HSP70-1A* GenBank ID: *H. sapiens*, NP_005336.3; *M. musculus*, NP_034609.2; *Notothenia coriiceps*, XP_010769991.1; *M. albus*, AGO01980.1; *Nothobranchius korthausae*, SBQ68887.1. *HSP70-1B* GenBank ID: *M. albus*, AGO01981.1; *D. rerio*, NP_001093532.1. *HSP70-2* GenBank ID: *H. sapiens*, AAH36107.1; *M. musculus*, EDL36460.1. *HSP70-1* from *C. elegans*, NP_503068.1, was used as outgroup.

TABLE S1 | Expression of reference genes of *Odontesthes humensis* before (D0) and after (D1, D7, and D15) hypo- and hyperosmotic shock. Equal letters represent no difference in a same column. Abbreviations: Ct, cycle threshold values; D0, day zero; D1, day one; D7, day seven; D15, day fifteen; FW–BW, transfer from freshwater to brackish water; BW–FW, transfer from brackish water to freshwater.

REFERENCES

Altinok, I., and Grizzle, J. M. (2001). Effects of brackish water on growth, feed conversion and energy absorption efficiency by juvenile euryhaline and freshwater stenohaline fishes. *J. Fish Biol.* 59, 1142–1152. doi: 10.1111/j.1095-8649.2001.tb00181.x

Armesto, P., Campinho, M. A., Rodríguez-Rúa, A., Cousin, X., Power, D. M., Manchado, M., et al. (2014). Molecular characterization and transcriptional regulation of the Na^+/K^+ ATPase α subunit isoforms during development and salinity challenge in a teleost fish, the Senegalese sole (*Solea senegalensis*). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 175, 23–38. doi: 10.1016/j.cbpb.2014.06.004

Basu, N., Nakano, T., Grau, E. G., and Iwama, G. K. (2001). The effects of cortisol on Heat Shock Protein 70 levels in two fish species. *Gen. Comp. Endocrinol.* 124, 97–105. doi: 10.1006/gcen.2001.7688

Beheregaray, L. B., Sunnucks, P., and Briscoe, D. A. (2002). A rapid fish radiation associated with the last sea-level changes in southern Brazil: the silverside *Odontesthes perugiae* complex. *Proc. R. Soc. B Biol. Sci.* 269, 65–73. doi: 10.1098/rspb.2001.1838

Bemvenuti, M. A. (2006). Silversides in South Brazil: morphological and ecological aspects. *Biocell* 30, 111–118.

Bloom, D. D., Weir, J. T., Piller, K. R., and Lovejoy, N. R. (2013). Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of New World silversides (Atherinopsidae). *Evolution* 67, 2040–2057. doi: 10.1111/evo.12074

Breves, J. P., Hasegawa, S., Yoshioka, M., Fox, B. K., Davis, L. K., Lerner, D. T., et al. (2010). Acute salinity challenges in Mozambique and *Nile tilapia*: differential responses of plasma prolactin, growth hormone and branchial expression of ion transporters. *Gen. Comp. Endocrinol.* 167, 135–142. doi: 10.1016/j.ygcen.2010.01.022

Campanella, D., Hughes, L. C., Unmack, P. J., Bloom, D. D., Piller, K. R., and Ortí, G. (2015). Multi-locus fossil-calibrated phylogeny of Atheriniformes (Teleostei, Ovalentaria). *Mol. Phylogenet. Evol.* 86, 8–23. doi: 10.1016/j.ympev.2015.03.001

Cataldi, E., Mandich, A., Ozzimo, A., and Cataudella, S. (2005). The interrelationships between stress and osmoregulation in a euryhaline fish, *Oreochromis mossambicus*. *J. Appl. Ichthyol.* 21, 229–231. doi: 10.1111/j.1439-0426.2005.00634.x

Chand, B. K., Trivedi, R. K., Dubey, S. K., Rout, S. K., Beg, M. M., and Das, U. K. (2015). Effect of salinity on survival and growth of giant freshwater prawn *Macrobrachium rosenbergii* (de Man). *Aquacult. Rep.* 2, 26–33. doi: 10.1016/j.aqrep.2015.05.002

D'Cotta, H., Valotaire, C., Gac, F., Le, and Prunet, P. (2000). Synthesis of gill Na^+-K^+-ATPase in Atlantic salmon smolts: differences in α-mRNA and α-protein levels. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 278, R101–R110. doi: 10.1152/ajpregu.2000.278.1.R101

Dyer, B. S. (2000). Systematic review and biogeography of the freshwater fishes of Chile. *Estud. Oceanol.* 19, 77–98. doi: 10.1002/dvdy.24533

Dyer, B. S. (2006). Systematic revision of the South American silversides (Teleostei, Atheriniformes). *Biocell* 30, 69–88.

Grutter, A. S., and Pankhurst, N. W. (2000). The effects of capture, handling, confinement and ectoparasite load on plasma levels of cortisol, glucose and lactate in the coral reef fish *Hemigymnus melapterus*. *J. Fish Biol.* 57, 391–401. doi: 10.1006/jfbi.2000.1312

Guynn, S. R., Scofield, M. A., and Petzel, D. H. (2002). Identification of mRNA and protein expression of the Na/K-ATPase α1-, α2- and α3-subunit isoforms in Antarctic and New Zealand notothenid fishes. *J. Exp. Mar. Biol. Ecol.* 273, 15–32. doi: 10.1016/S0022-0981(02)00136-3

Havird, J. C., Henry, R. P., and Wilson, A. E. (2013). Altered expression of Na^+/K^+–ATPase and other osmoregulatory genes in the gills of euryhaline animals in response to salinity transfer: a meta-analysis of 59 quantitative PCR studies over 10 years. *Comp. Biochem. Physiol. D Genomics Proteomics* 8, 131–140. doi: 10.1016/j.cbd.2013.01.003

He, Y., Luo, M., Yi, M., Sheng, Y., Cheng, Y., Zhou, R., et al. (2013). Identification of a testis-enriched Heat Shock Protein and fourteen members of Hsp70 family in the swamp eel. *PLOS ONE* 8:e65269. doi: 10.1371/journal.pone.0065269

Heras, S., and Roldán, M. I. (2011). Phylogenetic inference in *Odontesthes* and *Atherina* (Teleostei: Atheriniformes) with insights into ecological adaptation. *C. R. Biol.* 334, 273–281. doi: 10.1016/j.crvi.2011.03.004

Hiroi, J., and McCormick, S. D. (2012). New insights into gill ionocyte and ion transporter function in euryhaline and diadromous fish. *Respir. Physiol. Neurobiol.* 184, 257–268. doi: 10.1016/j.resp.2012.07.019

Hudson, H. A., Brauer, P. R., Scofield, M. A., and Petzel, D. H. (2008). Effects of warm acclimation on serum osmolality, cortisol and hematocrit levels in the Antarctic fish, *Trematomus bernacchii*. *Polar Biol.* 31, 991–997. doi: 10.1007/s00300-008-0438-8

Hwang, P., Lee, T., and Lin, L. (2011). Ion regulation in fish gills: recent progress in the cellular and molecular mechanisms. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 301, R28–R47. doi: 10.1152/ajpregu.00047.2011

Kakumura, K., Takabe, S., Takagi, W., Hasegawa, K., Konno, N., Bell, J. D., et al. (2015). Morphological and molecular investigations of the holocephalan elephant fish nephron: the existence of a countercurrent-like configuration and two separate diluting segments in the distal tubule. *Cell Tissue Res.* 362, 677–688. doi: 10.1007/s00441-015-2234-4

Kang, C. K., Tsai, H. J., Liu, C. C., Lee, T. H., and Hwang, P. P. (2010). Salinity-dependent expression of a Na^+, K^+, $2Cl^-$ cotransporter in gills of the brackish medaka *Oryzias dancena*: a molecular correlate for hyposmoregulatory endurance. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 157, 7–18. doi: 10.1016/j.cbpa.2010.05.013

Kiilerich, P., Tipsmark, C. K., Borski, R. J., and Madsen, S. S. (2011). Differential effects of cortisol and 11-deoxycorticosterone on ion transport protein mRNA levels in gills of two euryhaline teleosts, Mozambique tilapia (*Oreochromis mossambicus*) and striped bass (*Morone saxatilis*). *J. Endocrinol.* 209, 115–126. doi: 10.1530/JOE-10-0326

Larsen, P. F., Nielsen, E. E., Meier, K., Olsvik, P. A., Hansen, M. M., and Loeschcke, V. (2012). Differences in salinity tolerance and gene expression between two populations of Atlantic cod (*Gadus morhua*) in response to salinity stress. *Biochem. Genet.* 50, 454–466. doi: 10.1007/s10528-011-9490-0

Li, Z., Lui, E. Y., Wilson, J. M., Ip, Y. K., Lin, Q., and Lam, T. J. (2014). Expression of key ion transporters in the gill and esophageal-gastrointestinal tract of euryhaline Mozambique tilapia *Oreochromis mossambicus* acclimated to fresh water, seawater and hypersaline water. *PLOS ONE* 9:e87591. doi: 10.1371/journal.pone.0087591

Lin, C.-H., Huang, C.-L., Yang, C.-H., Lee, T.-H., and Hwang, P.-P. (2004). Time-course changes in the expression of Na, K-ATPase and the morphometry of mitochondrion-rich cells in gills of euryhaline tilapia (*Oreochromis mossambicus*) during freshwater acclimation. *J. Exp. Zool. A Comp. Exp. Biol.* 301A, 85–96. doi: 10.1002/jez.a.20007

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Loong, A. M., Chew, S. F., Wong, W. P., Lam, S. H., and Ip, Y. K. (2012). Both seawater acclimation and environmental ammonia exposure lead to increases in mRNA expression and protein abundance of Na^+:K^+:2Cl^- cotransporter in the gills of the climbing perch, *Anabas testudineus*. *J. Comp. Physiol. B Biochem. Syst. Environ. Physiol.* 182, 491–506. doi: 10.1007/s00360-011-0634-7

Loubens, G., and Osorio, F. (1991). "Especies introducidas: *Basilichthyes bonariesis*," in *El Lago Titicaca: Síntesis del Conocimiento Limnológico Actual*, eds C. Dejoux and A. Iltis (La Paz, BO: ORSTOM/HISBOL), 431–449.

Lovejoy, N. R., Albert, J. S., and Crampton, W. G. R. (2006). Miocene marine incursions and marine/freshwater transitions: evidence from Neotropical fishes. *J. South Am. Earth Sci.* 21, 5–13. doi: 10.1016/j.jsames.2005.07.009

McCallister, C., Kdeiss, B., Oliverio, R., and Nikolaidis, N. (2016). Characterization of the binding between a 70-kDa heat shock protein, HspA1A, and phosphoinositides. *Biochem. Biophys. Res. Commun.* 472, 270–275. doi: 10.1016/j.bbrc.2016.02.103

Menone, M. L., Aizpún de Moreno, J. E., Moreno, V. J., Lanfranchi, A. L., Metcalfe, T. L., and Metcalfe, C. D. (2000). Environmental contamination and toxicology PCBs and organochlorines in tissues of silverside (*Odontesthes bonariensis*) from a coastal lagoon in Argentina. *Arch. Environ. Contam. Toxicol.* 38, 202–208. doi: 10.1007/s002449910027

Morais-Cabral, J. H., and Robertson, G. A. (2015). The enigmatic cytoplasmic regions of KCNH channels. *J. Mol. Biol.* 427, 67–76. doi: 10.1016/j.jmb.2014.08.008

Morimoto, R. I. (1998). Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. *Genes Dev.* 12, 3788–3796. doi: 10.1101/gad.12.24.3788

Nilsen, T. O., Ebbesson, L. O. E., Madsen, S. S., McCormick, S. D., Andersson, E., Björnsson, B. T., et al. (2007). Differential expression of gill Na^+,K^+-ATPase alpha- and beta-subunits, Na^+,K^+,2Cl^- cotransporter and CFTR anion channel in juvenile anadromous and landlocked Atlantic salmon *Salmo salar*. *J. Exp. Biol.* 210, 2885–2896. doi: 10.1242/jeb.002873

Occhiodoro, T., Bernheim, L., Liu, J. H., Bijlenga, P., Sinnreich, M., Bader, C. R., et al. (1998). Cloning of a human ether-a-go-go potassium channel expressed in myoblasts at the onset of fusion. *FEBS Lett.* 434, 177–182. doi: 10.1016/S0014-5793(98)00973-9

Pardo, L. A., and Stühmer, W. (2013). The roles of K^+ channels in cancer. *Nat. Rev. Cancer* 14, 39–48. doi: 10.1038/nrc3635

Peng, G., Zhao, W., Shi, Z., Chen, H., Liu, Y., Wei, J., et al. (2016). Cloning HSP70 and HSP90 genes of kaluga (*Huso dauricus*) and the effects of temperature and salinity stress on their gene expression. *Cell Stress Chaperones* 21, 349–359. doi: 10.1007/s12192-015-0665-1

Piedras, S. R. N., Fernandes, J. L. O., Motoyama, I. S., and Martins, G. B. (2009). Efeito de diferentes concentrações de salinas (NaCl) na sobrevivência de embriões de peixe – rei *Odontesthes bonariensis* e *Odontesthes humensis*. *Biotemas* 22, 235–238. doi: 10.5007/2175-7925.2009v22n3p235

Place, S. P., and Hofmann, G. E. (2005). Constitutive expression of a stress-inducible heat shock protein gene, *hsp70*, in phylogenetically distant Antarctic fish. *Polar Biol.* 28, 261–267. doi: 10.1007/s00300-004-0697-y

Pouey, J. L. O. F., Rocha, C. B., Tavares, R. A., Portelinha, M. K., and Piedras, S. R. N. (2012). Frequência alimentar no crescimento de alevinos de peixe-rei *Odontesthes humensis*. *Semin. Agrar.* 33, 2423–2427. doi: 10.5433/1679-0359.2012v33n6p2423

Rajeshkumar, S., Mini, J., and Munuswamy, N. (2013). Effects of heavy metals on antioxidants and expression of HSP70 in different tissues of Milk fish (*Chanos chanos*) of Kaattuppalli Island, Chennai, India. *Ecotoxicol. Environ. Saf.* 98, 8–18. doi: 10.1016/j.ecoenv.2013.07.029

Scott, G. R., Baker, D. W., Schulte, P. M., and Wood, C. M. (2008). Physiological and molecular mechanisms of osmoregulatory plasticity in killifish after seawater transfer. *J. Exp. Biol.* 211, 2450–2459. doi: 10.1242/jeb.017947

Scott, G. R., and Schulte, P. M. (2005). Intraspecific variation in gene expression after seawater transfer in gills of the euryhaline killifish *Fundulus heteroclitus*. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 141, 176–182. doi: 10.1016/j.cbpb.2005.05.002

Somoza, G. M., Miranda, L. A., Berasain, G. E., Colautti, D., Remes Lenicov, M., and Strüssmann, C. A. (2008). Historical aspects, current status and prospects of pejerrey aquaculture in South America. *Aquacult. Res.* 39, 784–793. doi: 10.1111/j.1365-2109.2008.01930.x

Stefansson, S. O., Nilsen, T. O., Ebbesson, L. O. E., Wargelius, A., Madsen, S. S., Björnsson, B. T., et al. (2007). Molecular mechanisms of continuous light inhibition of Atlantic salmon parr–smolt transformation. *Aquaculture* 273, 235–245. doi: 10.1016/j.aquaculture.2007.10.005

Stengel, R., Rivera-Milla, E., Sahoo, N., Ebert, C., Bollig, F., Heinemann, S. H., et al. (2012). Kcnh1 voltage-gated potassium channels are essential for early zebrafish development. *J. Biol. Chem.* 287, 35565–35575. doi: 10.1074/jbc.M112.363978

Strüssmann, C. A., Moriyama, S., Hanke, E. F., Cota, J. C. C., and Takashima, F. (1996). Evidence of thermolabile sex determination in pejerrey. *J. Fish Biol.* 48, 643–651. doi: 10.1111/j.1095-8649.1996.tb01459.x

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Taugbøl, A., Arntsen, T., Østbye, K., and Vøllestad, L. A. (2014). Small changes in gene expression of targeted osmoregulatory genes when exposing marine and freshwater threespine stickleback (*Gasterosteus aculeatus*) to abrupt salinity transfers. *PLOS ONE* 9:e106894. doi: 10.1371/journal.pone.0106894

Tavares, R. A., Fernandes, J. M., Garcia, V. H., Piedras, S. R. N., Pouey, J. L. O. F., Dionello, N. J. L., et al. (2014). Growth performance of three pejerrey genetic groups in intensive culture system. *Semin. Cienc. Agrar.* 35, 2749–2758. doi: 10.5433/1679-0359.2014v35n5p2749

Tipsmark, C. K., Madsen, S. S., and Borski, R. J. (2004). Effect of salinity on expression of branchial ion transporters in striped bass (*Morone saxatilis*). *J. Exp. Zool. Comp. Exp. Biol.* 301A, 979–991. doi: 10.1002/jez.a.119

Tipsmark, C. K., Madsen, S. S., Seidelin, M., Christensen, A. S., Cutler, C. P., and Cramb, G. (2002). Dynamics of Na^+,K^+,2Cl^- cotransporter and Na^+,K^+-ATPase expression in the branchial epithelium of brown trout (*Salmo trutta*) and Atlantic salmon (*Salmo salar*). *J. Exp. Zool.* 293, 106–118. doi: 10.1002/jez.10118

Tsuzuki, M. Y., Aikawa, H., Strüssmann, C. A., and Takashima, F. (2000a). Comparative survival and growth of embryos, larvae, and juveniles of pejerrey *Odontesthes bonariensis* and *O. hatcheri* at different salinities. *J. Appl. Ichthyol.* 16, 126–130. doi: 10.1046/j.1439-0426.2000.00227.x

Tsuzuki, M. Y., Aikawa, H., Strüssmann, C. A., and Takashima, F. (2000b). Physiological responses to salinity increases in the freshwater silversides *Odontesthes bonariensis* and *O. hatcheri* (Pisces, Atherinidae). *Rev. Bras. Oceanogr.* 48, 81–85. doi: 10.1590/S1413-77392000000100007

Tsuzuki, M. Y., Ogawa, K., Strüssmann, C. A., Maita, M., and Takashima, F. (2001). Physiological responses during stress and subsequent recovery at different salinities in adult pejerrey *Odontesthes bonariensis*. *Aquaculture* 200, 349–362. doi: 10.1016/S0044-8486(00)00573-1

Tsuzuki, M. Y., Ogawa, K., Strüssmann, C. A., Maita, M., Takashima, F., and Melo, C. M. R. (2007). The significance of cortisol on acclimation to salinity in pejerrey *Odontesthes bonariensis*. *Arq. Bras. Med. Vet. Zootec.* 59, 1301–1307. doi: 10.1590/S0102-09352007000500030

Whitehead, A., Roach, J. L., Zhang, S., and Galvez, F. (2012). Salinity- and population-dependent genome regulatory response during osmotic acclimation in the killifish (*Fundulus heteroclitus*) gill. *J. Exp. Biol.* 215(Pt 8), 1293–1305. doi: 10.1242/jeb.062075

Yamashita, M., Yabu, T., and Ojima, N. (2010). Stress protein HSP70 in fish. *Aqua Biosci. Monogr.* 3, 111–141. doi: 10.5047/absm.2010.00304.0111

Zebral, Y. D., Costa, P. G., Knopp, B. C., Lansini, L. R., Zafalon-Silva, B., Bianchini, A., et al. (2017). Effects of a glyphosate-based herbicide in pejerrey *Odontesthes humensis* embryonic development. *Chemosphere* 185, 860–867. doi: 10.1016/j.chemosphere.2017.07.069

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 15 February 2018
doi: 10.3389/fgene.2018.00047

Molecular Evidences of a Hidden Complex Scenario in *Leporinus* cf. *friderici*

Rosane Silva-Santos, Jorge L. Ramirez, Pedro M. Galetti Jr. and Patrícia D. Freitas*

Laboratório de Biodiversidade Molecular e Conservação, Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Paulo, Brazil

The megadiversity of the neotropical ichthyofauna has been associated to recent diversification processes, reflecting in subtle or lacking morphological differentiation between species, challenging the classical taxonomic identification. *Leporinus friderici* occurs in several river basins of South America, and its nominal taxonomic validity has been questioned. Its wide distribution within the Brazilian Shield suggests that this species could be genetically structured among the hydrographic basins, despite a sharp morphological similarity. In this study, we performed phylogenetic analyses, based on three nuclear (recombination activating gene 1, RAG1, recombination activating gene 2, RAG2, and myosin heavy chain 6 cardiac muscle alpha gene, Myh6) and two mitochondrial (COI and Cytochrome b, Cytb) markers, in specimens morphologically similar to *L. friderici* and related species from different hydrographic basins in South America. Our phylogenetic tree identified four well-supported clades, which point out to the existence of taxonomic inconsistencies within this fish group. A clade named *L.* cf. *friderici sensu stricto* included eight Molecular Operational Taxonomic Units recently diversified in the Brazilian Shield basins. These results were also confirmed by a single-gene species delimitation analysis. It is suggested that this clade includes a species complex, characterizing taxonomic uncertainties. Another clade recovered only *L. friderici* from the Suriname rivers, validating this nominal species in its type locality. A third no-named clade, characterized by deeper species divergence, recovered five different nominal species interleaved with other undescribed forms previously also recognized as *L.* cf. *friderici*, indicating taxonomic errors. The fourth clade only included *L. taeniatus*. Our results showed a complex scenario involving the morphotype *L.* cf. *friderici* and allowed us to address aspects related to evolutionary diversification of this fish group and historical processes involved with, highlighting the importance of revealing hidden biodiversity for the taxonomy and conservationist action plans of these fish.

Keywords: neotropical fish, MOTUs, fish phylogeny, taxonomic uncertainties, cryptic species

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

Reviewed by:
Izeni Pires Farias, Federal University of Amazonas, Brazil
Marcelo Ricardo Vicari, Ponta Grossa State University, Brazil

****Correspondence:***
Rosane Silva-Santos
rosanesantos.gen@gmail.com

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

Received: *30 October 2017*
Accepted: *31 January 2018*
Published: *15 February 2018*

Citation:
Silva-Santos R, Ramirez JL, Galetti PM Jr and Freitas PD (2018) Molecular Evidences of a Hidden Complex Scenario in Leporinus cf. friderici. Front. Genet. 9:47. doi: 10.3389/fgene.2018.00047

INTRODUCTION

South America freshwater fish represent one-third of the world continental ichthyofauna (Reis et al., 2016). However, this huge biodiversity is relatively recent, mostly due to extensive speciation events during the last 10 Ma (Hubert et al., 2007; Albert and Reis, 2011). Several taxa diverged <1 Ma during the late Pleistocene because of the Quaternary activity that

led to topographic changes such as recent drainage rearrangements (Ribeiro, 2006). These events probably reflected in subtle or lacking morphological differentiation between the emerged species, challenging the classical taxonomic identification.

Molecular analyses have been largely used to aid species identification and delimitation within neotropical fish (e.g., Carvalho et al., 2011; Pereira et al., 2011, 2013; Ramirez and Galetti, 2015; Machado et al., 2016), contributing in revealing hidden biodiversity (Pires et al., 2017; Ramirez et al., 2017a). Both DNA barcoding and phylogeny studies have provided important contributions for better understand the phylogenetic relationships within taxa, and historical and evolutionary processes involved in the species diversification (e.g., Hebert et al., 2003; Carvalho-Costa et al., 2011; Ramirez et al., 2017b). In this way, mitochondrial and nuclear DNA data can be used to define discrete genetic lineages, detecting reciprocal monophyly, and characterizing a Molecular Operational Taxonomic Unit (MOTU, cluster of orthologous sequences generated by an explicit algorithm), representing a monophyletic lineage that could or not correspond to a taxa (Blaxter et al., 2005; Jones et al., 2011).

Leporinus, from the Anostomidae family, is considered one of the richest genus within Characiformes, a predominant freshwater fish order in South America (Garavello and Britski, 2003). Recent morphological and molecular studies have shown *Leporinus* as a non-monophyletic genus (Sidlauskas and Vari, 2008; Ramirez et al., 2016), highlighting the need of a deep taxonomic revision on this group. Freshly, an integrated morphological, chromosomal, and molecular approach described the new genus *Megaleporinus*, gathering the largest body sized *Leporinus* species in a monophyletic clade (Ramirez et al., 2017b).

Among the remaining *Leporinus* species, *Leporinus friderici* Block (1794) has a wide geographic distribution occurring in most rivers of South America (Garavello et al., 1992). This large distribution has been investigated by several authors, who have found morphological (Géry et al., 1987; Renno et al., 1990, 1991; Garavello et al., 1992; Sidlauskas and Vari, 2012) and genetic (Renno et al., 1990, 1991) variations among populations, suggesting that this *L. friderici* morphotype may contain a species complex (Sidlauskas and Vari, 2012). While *L. friderici* from Suriname and French Guiana rivers has been recognized as the type species (Sidlauskas and Vari, 2012), the provisional nomenclature *Leporinus* cf. *friderici* has been used to refer to the remaining individuals of this morphotype. However, because of its wide distribution within the Brazilian Shield, one could expect some levels of genetic differentiation within the *L.* cf. *friderici* among hydrographic basins. Morphological variations in the eyes, body length, and color patterns have been already reported between *L.* cf. *friderici* from the Amazon and Paraná-Paraguay basins (Géry et al., 1987; Garavello et al., 1992).

In this study we tested the hypothesis that *L.* cf. *friderici* consists of a monophyletic group formed by different MOTUs currently separated in distinct basins within the Brazilian Shield. We performed a phylogenetic analysis using mitochondrial and nuclear markers to confirm *L.* cf. *friderici* as a monophyletic group, and a single-gene species delimitation analysis to characterize MOTUs. Finally, we infer on the evolutionary historical processes possibly responsible for the diversification within this group.

MATERIALS AND METHODS

Ethics Statements

This research was conducted within the protocols approved by the Ethics Committee on Animal Experimentation (CEUA, Federal University of São Carlos, São Carlos, São Paulo, Brazil) and SISBIO-ICMBio (Authorization System and Biodiversity Information, Chico Mendes Institute for Biodiversity Conservation, Ministry of Environment, Brazil). The biological samples were obtained under CMBIO/MMA No. 32215 and CEUA No. 3893250615 permits following all legal requirements. Samples from Colombia and Peru were obtained by Jose Ariel Rodriguez and Hernan Ortega, respectively, who provided DNA aliquots for this study.

Biological Specimens and Data Sampling

Samples of fin, muscle, or liver tissues were collected of 53 *L.* cf. *friderici* specimens from eight hydrographic basins from the Brazilian Shield. **Figure 1** illustrates the studied river basins. Samples of *Leporinus agassizii* Steindachner (1876), *L. boehlkei* Garavello (1988), and *L. piau* Fowler (1941), from Jaguaribe and São Francisco rivers and *L.* cf. *parae* Eigenmann and Ogle (1907), were also obtained. In addition, DNA sequences of *Leporinus desmotes* Fowler (1941), *Leporinus fasciatus* Block (1794), *Leporinus lacustris* Campos (1945), *Leporinus octomaculatus* Britski and Garavello (1993), *Leporinus taeniatus* Lütken (1875), *Leporinus venerei* Britski and Birindelli (2008), and *Hypomasticus pachycheilus* Britski (1976), were downloaded from GenBank (Ramirez et al., 2016). All these species were included in our study because they have been considered as closely related to *L. friderici* (Garavello, 1988; Britski and Birindelli, 2008; Ramirez et al., 2016). Sequences of *L. friderici* from Suriname, that is considered the type locality, were obtained from GenBank (Melo et al., 2016) as well. Information related to the specimens, vouchers, IDs, and site localities are recorded in the Supplementary Material (**Supplementary Tables S1**, **S2**).

DNA Extraction, Gene Amplification, and Sequencing

Total DNA was extracted using the conventional phenol–chloroform/proteinase K protocol (Sambrook et al., 1989). For the phylogenetic analysis at least one individual per each major sampled hydrographic basin was used. Cytochrome b (Cytb) and cytochrome oxidase subunit 1 (COI) mitochondrial regions were amplified according to Ramirez and Galetti (2015). The nuclear recombination activating gene 1 (RAG1), recombination activating gene 2 (RAG2), and myosin heavy chain 6 cardiac muscle alpha gene (Myh6) were amplified following Oliveira et al. (2011). PCR products obtained for both DNA strands were sequenced on an ABI 373xl sequencer (Applied Biosystems, Little Chalfont, United Kingdom).

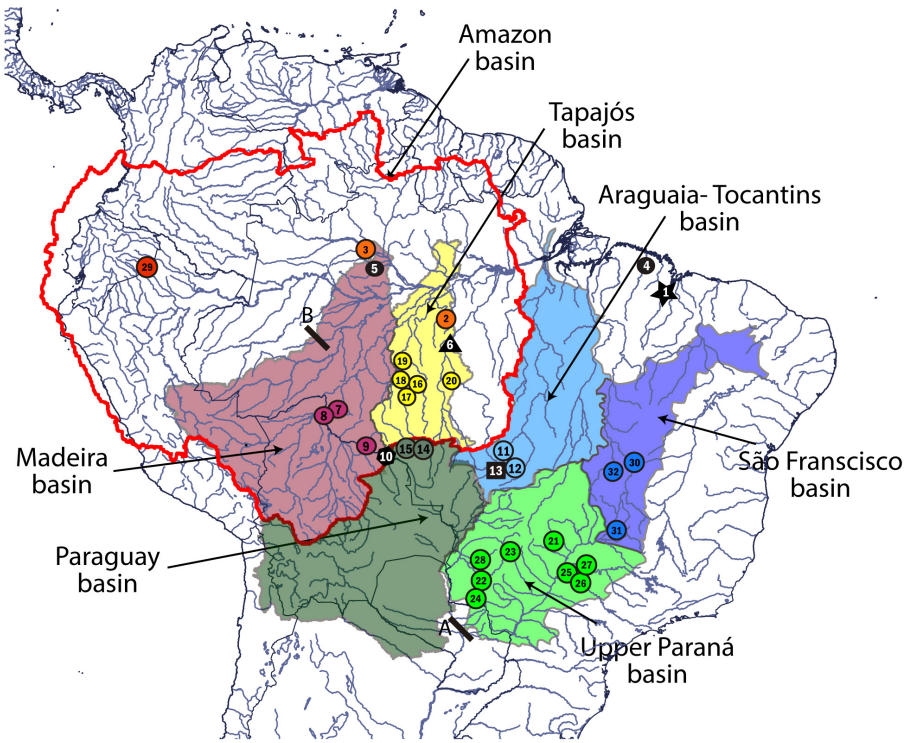


FIGURE 1 | Collection sites and hydrographic basin of occurrence of *Leporinus* cf. *friderici*. Colored circles represent the MOTUs found in this study: *L. agassizii* (red), *L.* cf. *friderici* Amazon 1 (orange), *L.* cf. *friderici* Madeira 1 (purple), *L.* cf. *friderici* Paraná (light green), *L.* cf. *friderici* Paraguay (dark green), *L. piau* São Francisco (dark blue), *L.* cf. *friderici* Upper Tapajós (yellow), and *L.* cf. *friderici* Tocantins 1 (light blue). Black dot represents individuals collected as *L.* cf. *friderici* Amazon 2; black star represents *L.* cf. *friderici* Mearim; black square represents *L.* cf. *friderici* Tocantins 2; black polygon represents *L.* cf. *friderici* Madeira 2; black triangle represents *L.* cf. *friderici* Xingú. A, Prior Sete Quedas waterfalls region and Itaipú dam. B, Teotônio falls in the Madeira River. Map created by Q-GIS (http://www.qgis.org/). The collection site numbers are according to **Supplementary Table S1**.

Data Analysis

The obtained sequences were manually edited and aligned using Bioedit (Hall, 1999) and Clustal W (Thompson et al., 1994), respectively. All sequences were checked for indels and stop codons. The haplotypes of the nuclear genes were combined into a consensus sequence by coding polymorphic sites with the IUPAC ambiguity codes (IUPAC, 1974).

The phylogenetic analyses were conducted for concatenated sequences of all mitochondrial and nuclear genes using maximum parsimony (MP), implemented in PAUP*4.0 (Swofford, 2003), with 1000 bootstrap replicates. We also performed maximum-likelihood (ML) analyses, using RAxML in XSEDE (Stamatakis, 2006; Stamatakis et al., 2008), through the web server CIPRES Science Gateway (Miller et al., 2010), and a partitioned model determined by PartitionFinder (Lanfear et al., 2012), under a GTR+G model, and 1000 bootstrap replicates.

A multilocus Bayesian species tree (BST) was estimated by *BEAST (Star-BEAST) (Heled and Drummond, 2010) using 150 million generations, sampled every 5000, and a burn-in of 300. A nucleotide substitution model selected was based on the Bayesian criterion, using JModeltest 2 (Darriba et al., 2012). The models chosen were HKY+I+G, HKY+G, K80+I, K80+I, and K80+I for COI, Cytb, Myh6, RAG1, and RAG2, respectively. The two mitochondrial gene tree topologies were linked and set to have an effective population size of one-quarter from that of nuclear genes. A lognormal relaxed clock was used for all partitions. Bayesian trees using all sequences were established for each gene, separately. These trees were estimated with 10 million of generations, sampling every 5000, and a burn-in of 10%. Yule speciation model was used and nucleotide substitution models followed the same criteria above cited (**Supplementary Figures S1–S5**). The convergence (sample size >200) and stationarity of the values were checked in TRACER v1.6 (Rambaut et al., 2014).

Two different analyses, using a single-gene species delimitation approach based on COI sequences, were performed to determine the number of MOTUs within the clade widely distributed in the Brazilian Shield (*L.* cf. *friderici sensu stricto*, see below). First, the General Mixed Yule-Coalescent (GMYC) model (Pons et al., 2006) was used to determine the clustering of the COI haplotypes. This analysis was performed with a single threshold that was implemented in the *SPLITS* package using R 3.3.3 statistical software (R Core Team, 2017). For this analysis an ultrametric tree was generated using BEAST 2.3.2 (Bouckaert et al., 2014), with a lognormal relaxed clock, a birth and death model, and a HKY substitution model chosen by jModeltest 2 (Darriba et al., 2012). A total of 50 million MCMC generations and a burn-in of 10% were used. In second, a Bayesian analysis

of genetic structure was implemented using BAPS software (Corander et al., 2008). The maximum number of genetically diverged groups (K) was firstly set up for 10 replicates, 10 times. The obtained groups containing samples from different basins were submitted to a second layer of analysis in BAPS using K = 1–3, replicated also 10 times. This hierarchical approach to DNA sequence clustering provides a useful way to increase statistical power and detect separated haplogroups that are assigned to conservative clusters (Cheng et al., 2013). The most likely K was chosen based on log (likelihood) and posterior probability values. Next, we consider the concordant groups between the two different analyses, and the groups presenting allopatry and reciprocal monophyly as the final number of MOTUs.

The genetic distances between MOTUS were calculated based on K2P model using MEGA 6.06 (Tamura et al., 2013). Finally, a haplotype network was generated using Median Joining (Bandelt et al., 1999) in the POPART software (Leigh and Bryant, 2015).

RESULTS

We successful obtained a total of 127 sequences for the specimens studied herein [63 for the COI gene (557 bp) and 16 for each remaining amplified marker – Cytb (1005 bp), Myh6 (754 bp), RAG1 (1477 bp), and RAG2 (1023 bp)]. The GenBank accession numbers are shown in the Supplementary Material section (**Supplementary Table S2**).

The phylogenetic trees, generated by MP, ML, and BST analyses (**Figure 2**), strongly supported four clades with maximum support values. A clade recovered *L. friderici* from the Suriname basin alone. A monophyletic clade characterized by a recent diversification within the Brazilian Shield, named herein as *L.* cf. *friderici sensu stricto*, included specimens of *L.* cf. *friderici* from Amazonas (main channel), Madeira, Upper Tapajós, Tocantins, Paraguay, and Paraná basins, *L. agassizii*, and specimens of *L. piau* from São Francisco basin. A no-named clade, showing older diversification, joined specimens of *L.* cf. *friderici*, from Mearim, Tocantins, Turiaçu, Xingu, and Madeira basins, interleaved with *L. boehlkei*, *L. lacustris*, *L.* cf. *parae*, *L. piau* from Jaguaribe basin, and *L. venerei*. Lastly, a fourth clade recovered only *L. taeniatus*.

The GMYC analysis, considering 23 parsimony-informative sites and no insertions or deletions within the COI sequences, identified seven MOTUs (CI: 6–7), with a significant likelihood ratio of 10.97 ($P < 0.005$) within *L.* cf. *friderici sensu stricto*. From these, six MOTUs corresponded to *L. agassizii*, *L.* cf. *friderici* Amazon 1, *L.* cf. *friderici* Madeira 1, *L.* cf. *friderici* Paraná, *L.* cf. *friderici* Paraguay, and *L.* cf. *friderici* Upper Tapajós. The seventh MOTU joined *L. piau* São Francisco and *L.* cf. *friderici* Tocantins 1 (**Figure 3**). The results of the first BAPS layer presented four as the most likely K with log (ml) = −554.0008 and 0.93 posterior probability values. In addition, the results of the second BAPS layer were similar to the GMYC MOTUs, but recovered *L. piau* São Francisco and *L.* cf. *friderici* Tocantins as two different MOTUs. The hierarchical analysis of BAPS could separate these populations, since information related to sample locations were given. Differently, GMYC method considers no prior information. Despite this little divergence in detecting MOTUs, both these lineages presenting recent divergence are reciprocally monophyletic and geographically isolated in nature.

The mean genetic distance COI K2P values among the MOTUs ranged from 0.4 to 2.4%. The maximum intra-MOTU distance (0.5%) was observed in *L. agassizii*, while the minimum inter-MOTU distance (0.4%) was between *L.* cf. *friderici* Tocantins and *L. piau* São Francisco (**Supplementary Table S3**).

A total of 27 haplotypes was obtained within *L.* cf. *friderici sensu stricto*, in which each MOTU was represented by a haplogroup, except for the MOTUs from Paraná and Paraguay that shared one haplotype. *L. piau* São Francisco and *L.* cf. *friderici* Upper Tapajós were separated by only one mutational step, while the other haplogroups were connected by at least two mutational steps (**Figure 4**).

DISCUSSION

Our phylogenetic analyses showed that specimens morphologically identified as *L. friderici* constitute a polyphyletic group, widely distributed along the South America (**Figure 2**). The individuals collected as *L. friderici* across the Brazilian Shield basins are not conspecific with *L. friderici* from the type locality, representing a different species. *L.* cf. *friderici sensu stricto* constitutes a monophyletic species complex distributed in Amazon, Madeira, Upper Tapajós, Tocantins, São Francisco, Paraná, and Paraguay river basins. This finding can represent a typical situation of recent diversification forming a strictly related group composed of potential cryptic species, revealing typical taxonomic uncertainties (Ramirez et al., 2017a). On the other hand, in the no-named clade the five nominal species *L. boehlkei*, *L.* cf. *parae*, *L. lacustris*, *L. piau* Jaguaribe, and *L. venerei* were interleaved with individuals that morphologically fit with the description of *L.* cf. *friderici*. In this clade, characterized by an older diversification, the use of the term *L.* cf. *friderici* hides undescribed cryptic species. All species from this clade share discriminative morphological general pattern with *L. friderici* (i.e., one–three spots on the body along the lateral-line, and dental formulae 4/4, except *L. venerei* that has 4/3, unique in Anostomidae), hindering the identification of these cryptic species. Moreover, three (*L. lacustris*, *L. parae* – *L.* cf. *parae* –, and *L. venerei*) of the five nominal species of this clade have been already considered as very similar in morphology due to their deep body, terminal mouth, anal fin long and dark, and three blotches on the lateral line (Britski and Birindelli, 2008).

Overall, the results obtained for *L.* cf. *friderici sensu stricto* confirm our hypothesis that there are different MOTUs within *L.* cf. *friderici*, currently separated in distinct river basins, but not all provisionally recognized as *L.* cf. *friderici*, that can be joined in a single monophyletic group.

Within the clade *L.* cf. *friderici sensu stricto*, *L. agassizii* is clearly recognized as a valid species different from *L. friderici*, mainly due to the presence of a longitudinal stripe, extending from dorsal fin to just before the caudal fin (Birindelli et al.,

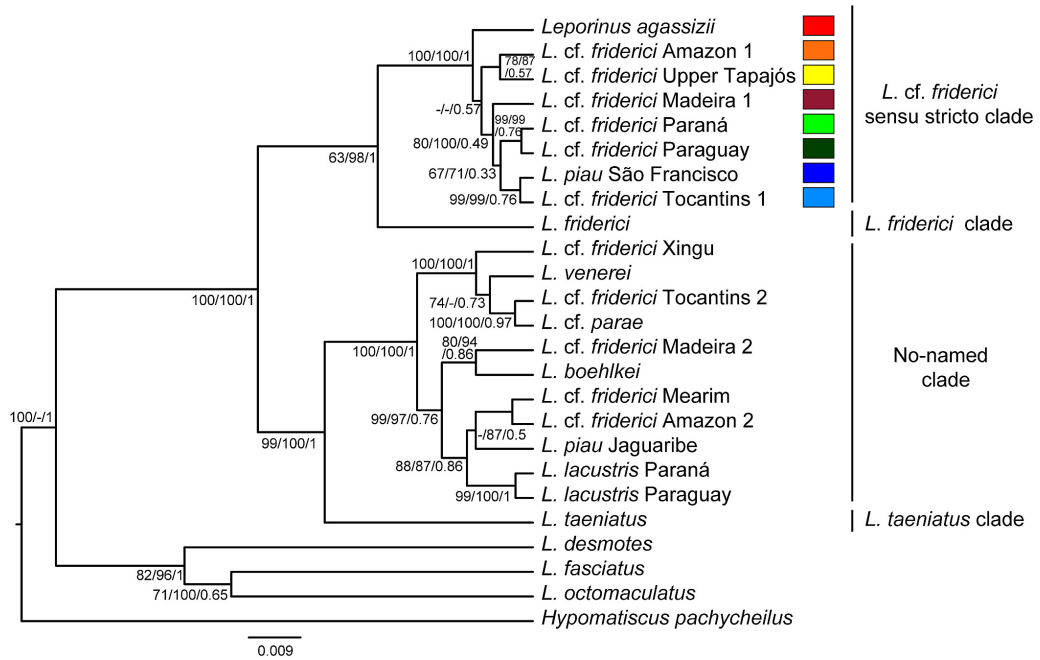


FIGURE 2 | Species tree showing phylogenetic relationships of *Leporinus* cf. *friderici*. Trees were generated using five molecular markers (sequences with approximately 4900 bp). The topology corresponds to Bayesian tree. The numbers on the branches are bootstrap values for maximum parsimony and maximum likelihood, and posterior probability for Bayesian species tree. The scale bar indicates nucleotide substitutions per site.

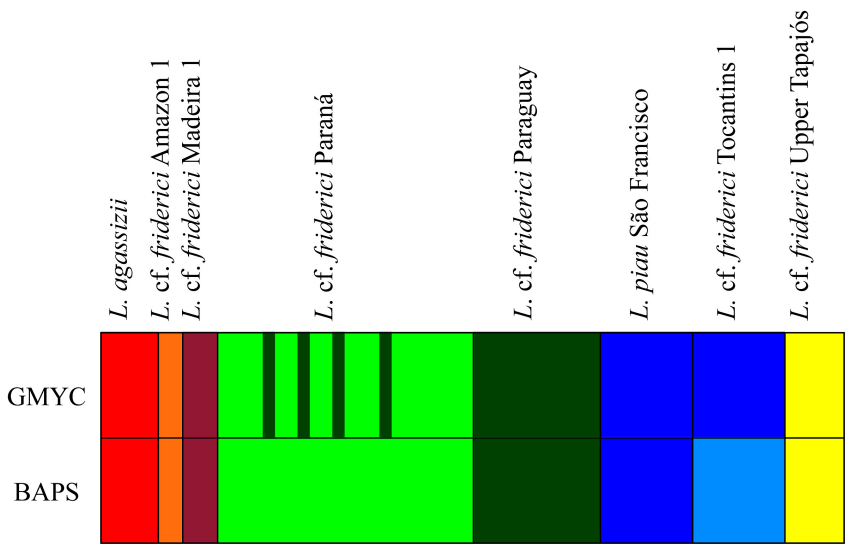


FIGURE 3 | MOTUs identified based on the GMYC and BAPS analyses for *Leporinus* cf. *friderici sensu stricto* clade.

2013). *L. agassizii* was firstly described for the Iça river, Upper Amazon basin (Steindachner, 1876), and posteriorly it was also found in the Tefe lake, Nanay, Negro, and Branco rivers (Birindelli et al., 2013), from the same basin. This species has been described as restricted to the Upper Amazon basin. A parsimony analysis of endemism in the South America reported the Upper Amazon as a separate clade from the Amazon drainages (Hubert and Renno, 2006), suggesting that uplift of the paleoarches has promoted allopatric divergence in the ichthyofauna of this region, which was enhanced by marine incursions.

In our study, all individuals collected in rivers above Madeira river falls were joined in the single MOTU *L.* cf. *friderici* Madeira 1, while individuals downstream the Amazon basin were recovered in the MOTU *L.* cf. *friderici* Amazon 1 (**Figure 3**), with the exception of individuals from Upper Tapajós (see below),

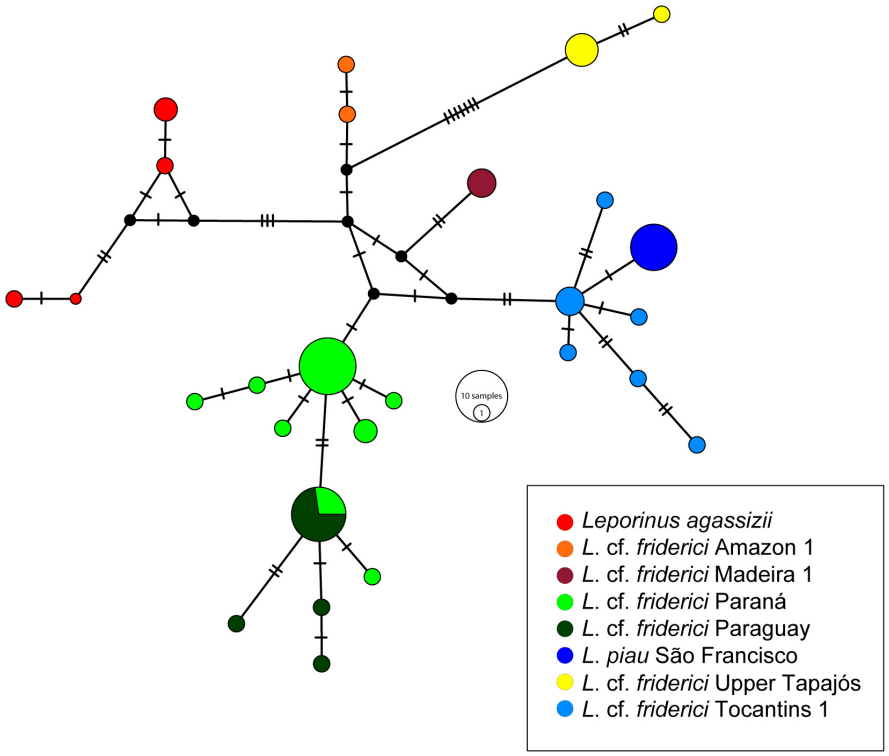


FIGURE 4 | The Median Joining haplotype network of *Leporinus* cf. *friderici sensu stricto* clade using COI gene data. The small dark dots indicate the missing or not sampled haplotypes. The haplotypes are shown in different colors according to each MOTU in legend.

indicating these falls as possible barrier that limits the fish species distribution in the region. Previous studies had already reported evidences of structuring along the Amazon basin (Goulding, 1979; de Queiroz et al., 2013). This subdivision was attributed to geomorphological agents that allowed allopatric fragmentation and diversification events (Albert and Reis, 2011). In the Madeira river, the Teotônio fall seems to play a relevant role on the ichthyofauna diversification of the Amazon basin. This fall, besides other rapids, has been keeping apart the rivers from Upper and Lower Madeira, and has been considered a geographic barrier by limiting the fish species distribution in the region (Zanata and Toledo-Piza, 2004; Hubert et al., 2007; Torrente-Vilara et al., 2011).

Barrier effects can also explain the presence of *L.* cf. *friderici* Upper Tapajós joining individuals caught in the Juruena – Teles Pires sub-basin, in the Upper Tapajós (**Supplementary Tables S1**, **S2**). Waterfalls and rapids along the Tapajós river, and in its tributaries seem acting as barriers to fish dispersal (Britski and Garavello, 2005; Britski and Lima, 2008; Dagosta and de Pinna, 2017). The region above the Juruena – Teles Pires confluence river has been characterized by an endemic ichthyofauna different from other Amazon rivers (Carvalho and Bertaco, 2006; Britski and Lima, 2008), which could account to the separation of *L.* cf. *friderici* Upper Tapajós.

In turn, the Tocantins basin is considered an independent system from the Amazon basin, since its waters flow directly into the Atlantic Ocean (Albert and Reis, 2011). This fact was reflected in our analyses, in which the Tocantins individuals corresponded to a different genetic group (*L.* cf. *friderici* Tocantins 1). The final establishment of modern course of Tocantins (1.8 Ma) separated definitively this basin from Amazon (Rossetti and Valeriano, 2007), and the differentiation of the Tocantins ichthyofauna has been often associated to the rise of Gurupá arch, the Tucurui rapids, or the limited connectivity (Hubert et al., 2007; Hrbek et al., 2014).

The relationship between *L. piau* São Francisco and *L.* cf. *friderici* Tocantins 1 observed here can be accounted for a biogeographic history between the Tocantins and São Francisco basins, and the low genetic divergence (0.4%) between them likely represents a recent diversification. These two hydrographic basins share an extensive watershed, where the Sapão river (São Francisco basin) shares headwaters with the Galheiros river, Tocantins basin (Lima and Caires, 2011). The existence of these common headwaters can allow a fauna exchange between these basins. Geological evidence shows that the western border of Serra Geral from the Goiás plateau has been gradually eroded and could have potentially promoted headwater capture events between the São Francisco and Tocantins rivers (Lima and Caires, 2011). Geodispersal events (i.e., headwater capture) from Amazon river to eastern basins of the Brazilian Shield (as São Francisco river) have been already claimed in studies using molecular approaches (Hubert et al., 2007; Ramirez et al., 2017b).

While the *L. piau* specimens from São Francisco was linked to the *L.* cf. *friderici sensu strito* clade, *L. piau* from Jaguaribe

was grouped in the no-named clade (**Figure 2**), revealing a clear taxonomic inconsistency. Fowler (1941) claimed the Salgado river (Jaguaribe basin) from Ceará state as type locality of *L. piau*, and included one paratype from Jatobá river (São Francisco basin). Consequently, specimens from São Francisco river have been usually cited as *L. piau* (Garavello and Britski, 2003; Carvalho et al., 2011). Our results pointed that specimens from São Francisco basin indeed constitute a different species from the *L. piau* from the Jaguaribe river, the type basin.

Still within *L.* cf. *friderici sensu stricto*, a well-supported differentiation between specimens from the Upper Paraná and Paraguay basins was also observed, although some individuals from the Upper Paraná showed haplotypes from the Paraguay basin (**Figures 3**, **4**). It is possible that both *L.* cf. *friderici* Paraná and *L.* cf. *friderici* Paraguay reached their current distribution through ancient geodispersal events, as headwater captures between Amazon rivers and the Paraná and Paraguay basins. In the modern river basin landscape, the Paraguay basin has a watershed with the Guaporé, Tapajós, and Xingu rivers, while the Upper Paraná shared a watershed at the headwaters of the Tocantins basin (Albert and Reis, 2011). These hydrographic systems have experienced a long history of major capture events and formation of semipermeable barriers (Lundberg et al., 1998) that can support this hypothesis. The Upper Paraná ichthyofauna was separated from the Lower Paraná by the Sete Quedas Falls, a natural geographic barrier which no longer occurs. In the past this barrier isolated the Upper Paraná, where the ichthyofauna has been diverging, as already reported in *Megaleporinus* (Ramirez et al., 2017b) and *Salminus* (Machado et al., 2016). The shared haplotypes between *L.* cf. *friderici* Paraná and *L.* cf. *friderici* Paraguay are probably resulting of the removal of the natural barrier when the Itaipu hydroelectric was built. The resulted dam flooded an extensive area, including the no longer existent Sete Quedas Falls, allowing the connection between both ichthyofauna from Lower and Upper Paraná facilitating contact between mitochondrial lineages since the formation of Sete Quedas Falls (Júlio et al., 2009; Prioli et al., 2012).

CONCLUSION

Our study showed that *L.* cf. *friderici* as provisionally used hides at least two major situations. First, *L.* cf. *friderici sensu stricto*, a monophyletic clade joining eight MOTUs, potentially includes a true species complex, characterized by recent diversification across the Brazilian Shield basins. According to our initial expectations, *L.* cf. *friderici sensu stricto* is genetically structured along the Brazilian shield basins, and this structure appears to be related to geomorphological agents, determining the current hydrographic structure. Its taxonomic significance is an open question, requesting complementary studies for resolving this typical situation of taxonomic uncertainties. Second, a no-named clade, characterized by relatively older diversification, in our opinion, hides undescribed cryptic species under *L.* cf. *friderici* denomination, likely due to the morphology similarities that characterize the clades here studied (except *L. taeniatus* clade). However, this new MOTUs show deep phylogenetic divergence and they are interleaved with other nominal valid species (*L. venerei*, *L. boehlkei*, *L. lacustris*, *L. piau*, and *L.* cf. *parae*), supporting them as potential new species.

Overall, our results have important significance for the taxonomy and evolutionary knowledge of this fish group as well as for its conservation. Moreover, this scenario indicates that *L.* cf. *friderici sensu stricto* can constitute an excellent phylogeographic model in studying evolutionary and speciation processes acting in the South America basins. Despite its migratory behavior, *L.* cf. *friderici* cannot be considered as a single genetic stock even within the same basin (i.e., Amazon basin) and needs to be well known for having its whole diversity considered in any conservation effort. For a more complete understanding, the taxonomic status of each MOTU that was revealed herein needs to be evaluated using preferentially morphological and molecular data in an integrative approach.

AUTHOR CONTRIBUTIONS

RS-S and JR collected the data, reviewed the literature, and achieved the bioinformatic analyses. All authors contributed to design the research, article writing and discussion, and approved the final version of the manuscript.

FUNDING

This study was supported by Conselho Nacional de Desenvolvimento Cientifíco e Tecnológico (CNPq, 304440/2009-4 and 473474/2011-5), SISBIOTA-Brazil Program (CNPq, 563299/2010-0; FAPESP, 10/52315-7), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, número to RS-S), and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, 2011/21836-4 to JR).

ACKNOWLEDGMENTS

We are grateful to C. Cramer, C. Doria, D. Carvalho, H. Ortega, J. A. Rodriguez, J. C. Riofrio, P. Venere, W. Troy, and U. Lopes for help to obtain part of the tissue or DNA samples and ICMBIO/MMA for sampling fish authorization (32215-1). We also thank H. Britski, P. Venere, J. Zuanon, N. Priorski, and H. Ortega for the specimen identification and the two reviewers for their comments and suggestions, which have improved the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00047/full#supplementary-material

FIGURE S1 | Bayesian tree for the cytochrome oxidase subunit 1 (COI) gene. Values on nodes represent the posterior probability.

FIGURE S2 | Bayesian tree for the cytochrome b (Cytb) gene. Values on nodes represent the posterior probability.

FIGURE S3 | Bayesian tree for the myosin heavy chain 6 cardiac muscle alpha (Myh6) gene. Values on nodes represent the posterior probability.

FiGURE S4 | Bayesian tree for the recombination activating gene 1 (RAG1) gene. Values on nodes represent the posterior probability.

FiGURE S5 | Bayesian tree for the recombination activating gene 2 (RAG2) gene. Values on nodes represent the posterior probability.

TABLE S1 | Vouchers and collection sites for all analyzed *Leporinus* cf. *friderici*.

TABLE S2 | GenBank accession numbers for all analyzed species.

TABLE S3 | Pairwise mean genetic distances values inter and intra-MOTU (in bold) *Leporinus* cf. *friderici sensu stricto* clade using K2p model. Values as percentage.

REFERENCES

Albert, J. S., and Reis, R. E. (eds). (2011). "Major biogeographic and phylogenetic patterns," in *Historical Biogeography of Neotropical Freshwater Fishes*, (Los Angeles, CA: University of California Press). doi: 10.1525/california/9780520268685.001.0001

Bandelt, H.-J. J., Forster, P., Röhl, A., and Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48. doi: 10.1093/oxfordjournals.molbev.a026036

Birindelli, J. L. O., Britski, H. A., and Lima, F. C. T. (2013). New species of *Leporinus* from the Rio Tapajós Basin, Brazil, and redescription of L. moralesi (Characiformes: Anostomidae). *Copeia* 2013, 238–247. doi: 10.1643/CI-12-081

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., et al. (2005). Defining operational taxonomic units using DNA barcode data. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1935–1943. doi: 10.1098/rstb.2005.1725

Block, M. E. (1794). Naturgeschichte der ausländischen Fische. *Berlin* 8, 361–396.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., et al. (2014). BEAST 2: a software platform for bayesian evolutionary analysis. *PLOS Comput. Biol.* 10:e1003537. doi: 10.1371/journal.pcbi.1003537

Britski, H. A. (1976). Sobre uma nova espécie Leporinus da Amazônia. *Acta Amaz.* 6, 87–89. doi: 10.1590/1809-43921976064s087

Britski, H. A., and Birindelli, J. L. O. (2008). Description of a new species of the genus *Leporinus* Spix (Characiformes: Anostomidae) from the rio Araguaia, Brazil, with comments on the taxonomy and distribution of *L. parae* and *L. lacustris*. *Neotrop. Ichthyol.* 6, 45–51. doi: 10.1590/S1679-62252008000100005

Britski, H. A., and Garavello, J. C. (1993). Descrição de duas espécies novas de Leporinus da bacia do Tapajós (Pisces, Characiformes). *Comun. Mus. Ciênc. Tecnol. PUCRS Sér. Zool.* 6, 29-40.

Britski, H. A., and Garavello, J. C. (2005). Uma nova espécie de *Leporinus* Agassiz, 1829, da bacia Amazônica (Ostariophysi: Characiformes: Anostomidae). *Comun. do Mus. Ciências e Tecnol. da PUCRS Série Zool.* 18, 75–83.

Britski, H. A., and Lima, F. C. T. (2008). A new species of *Hemigrammus* from the upper Rio Tapajós Basin in Brazil (Teleostei: Characiformes: Characidae). *Copeia* 2008, 565–569. doi: 10.1643/CI-07-134

Campos, A. A. (1945). Contribuição ao estudo das espécies brasileiras do gênero *Leporinus*. *Pap. Avulsos Dept. Zool.* 5, 141–158.

Carvalho, D. C., Oliveira, D. A. A., Pompeu, P. S., Leal, C. G., Oliveira, C., and Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River basin. *Mitochondrial DNA* 22, 80–86. doi: 10.3109/19401736.2011.588214

Carvalho, T. P., and Bertaco, V. A. (2006). Two new species of *Hyphessobrycon* (Teleostei: Characidae) from upper rio Tapajós basin on Chapada dos Parecis, central Brazil. *Neotrop. Ichthyol.* 4, 301–308. doi: 10.1590/S1679-62252006000300001

Carvalho-Costa, L. F., Piorski, N. M., Willis, S. C., Galetti, P. M. Jr., and Ortí, G. (2011). Molecular systematics of the neotropical shovelnose catfish genus *Pseudoplatystoma* Bleeker 1862 based on nuclear and mtDNA markers. *Mol. Phylogenet. Evol.* 59, 177–194. doi: 10.1016/j.ympev.2011.02.005

Cheng, L., Connor, T. R., Siren, J., Aanensen, D. M., and Corander, J. (2013). Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Mol. Biol. Evol.* 30, 1224–1228. doi: 10.1093/molbev/mst028

Corander, J., Marttinen, P., Sirén, J., and Tang, J. (2008). Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9:539. doi: 10.1186/1471-2105-9-539

Dagosta, F. C. P., and de Pinna, M. C. C. (2017). Biogeography of Amazonian fishes: deconstructing river basins as biogeographical units. *Neotrop. Ichthyol.* 15, 1–24. doi: 10.1590/1982-0224-20170034

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9:772. doi: 10.1038/nmeth.2109

de Queiroz, L. J., Torrente-Vilara, G., Ohara, W. M., Pires, T., Zuanon, J. A., and Doria, C. R. (eds). (2013). *Peixes do Rio Madeira*. São Paulo: Dialeto Latin American Documentary.

Eigenmann, C. H., and Ogle, F. (1907). An Annotated list of characin fishes in the United States National Museum and the Museum of Indiana University, with descriptions of new species. *Proc. U.S. Natl. Mus.* 33, 1–36.

Fowler, H. W. (1941). Academy of natural sciences A collection of fresh-water fishes obtained in Eastern Brazil by Dr. Rodolpho Von Ihering. *Proc. Acad. Nat. Sci. Phila.* 93, 123–199.

Garavello, J. C. (1988). A new species of the genus *Leporinus* Spix from the Rio Meta, Colombia, South America (Pisces, Ostariophysi, Anostomidae). *Proc. Acad. Nat. Sci. Phila.* 140, 143–149.

Garavello, J. C., and Britski, H. A. (2003). "Family anostomidae," in *Check List of the Freshwater Fishes of South and Central America*, eds R. E. Reis, S. O. Kullander, and C. J. Ferraris Jr. (Porto Alegre: EDIPUCRS), 71–84.

Garavello, J. C., Dos Reis, S. F., and Strauss, R. E. (1992). Geographic variation in *Leporinus friderici* (Bloch) (Pisces: Ostariophysi: Anostomidae) from the Paraná-Paraguay and Amazon River basins. *Zool. Scr.* 21, 197–200. doi: 10.1111/j.1463-6409.1992.tb00320.x

Géry, J., Mahnert, V., and Dlouhy, C. (1987). Poisson characoïdes non characidae du paraguay (Pisces, Ostariophysi). *Rev. Suisse Zool.* 94, 357–464. doi: 10.5281/zenodo.46287

Goulding, M. (1979). *Ecologia da Pesca do rio Madeira*. Manaus: Instituto Nacional de Pesquisas da Amazônia.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Hebert, P. D. N., Cywinska, A., Ball, S. L., and deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc. Biol. Sci.* 270, 313–321. doi: 10.1098/rspb.2002.2218

Heled, J., and Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27, 570–580. doi: 10.1093/molbev/msp274

Hrbek, T., da Silva, V. M. F., Dutra, N., Gravena, W., Martin, A. R., and Farias, I. P. (2014). A new species of river dolphin from Brazil or: how little do we know our biodiversity. *PLOS ONE* 9:e83623. doi: 10.1371/journal.pone.0083623

Hubert, N., Duponchelle, F., Nuñez, J., Garcia-Davila, C., Paugy, D., and Renno, J.-F. (2007). Phylogeography of the piranha genera *Serrasalmus* and *Pygocentrus*: implications for the diversification of the Neotropical ichthyofauna. *Mol. Ecol.* 16, 2115–2136. doi: 10.1111/j.1365-294X.2007.03267.x

Hubert, N., and Renno, J.-F. (2006). Historical biogeography of South American freshwater fishes. *J. Biogeogr.* 33, 1414–1436. doi: 10.1111/j.1365-2699.2006.01518.x

IUPAC (1974). Abbreviations and symbols for Nucleic Acids, Polynucleotides and their constituents. *Int. Union Pure Appl. Chem. Int. Union Biochem.* 40, 277–331.

Jones, M., Ghoorah, A., and Blaxter, M. (2011). JMOTU and taxonerator: turning DNA barcode sequences into annotated operational taxonomic units. *PLOS ONE* 6:e19259. doi: 10.1371/journal.pone.0019259

Júlio, H. F. Jr., Tós, C. D., Agostinho, Ă. A., and Pavanelli, C. S. (2009). A massive invasion of fish species after eliminating a natural barrier in the upper rio Paraná basin. *Neotrop. Ichthyol.* 7, 709–718. doi: 10.1590/S1679-62252009000400021

Lanfear, R., Calcott, B., Ho, S. Y. W., and Guindon, S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29, 1695–1701. doi: 10.1093/molbev/mss020

Leigh, J. W., and Bryant, D. (2015). POPART: full-feature software for haplotype network construction. *Methods Ecol. Evol.* 6, 1110–1116. doi: 10.1111/2041-210X.12410

Lima, F. C. T., and Caires, R. A. (2011). Peixes da Estação Ecológica Serra Geral do Tocantins, bacias dos Rios Tocantins e São Francisco, com observações sobre as implicações biogeográficas das "águas emendadas" dos Rios Sapão e Galheiros. *Biota Neotrop.* 11, 231–250. doi: 10.1590/S1676-06032011000100024

Lundberg, J. G., Marshall, L. G., Guerrero, J., Horton, B., Malabarba, M. C. S. L., and Wesselingh, F. P. (1998). "The stage for Neotropical fish diversification: a history of tropical South American rivers," in *Phylogeny and Classification of Neotropical Fishes*, eds L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M. S. Lucena, and C. A. S. Lucena (Porto Alegre: EDIPUCRS), 13–48.

Lütken, C. F. (1875). Characinae novae Brasiliae centralis a clarissimo J. Reinhardt in provincia Minas-Geraes circa oppidulum Lagoa Santa in lacu ejusdem nominis, flumine Rio das Velhas et rivulis affluentibus collectae, secundum characteres essentiales breviter descriptae. *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger og dets Medlemmers Arbeider (Kjøbenhavn)* 1874, 127–143.

Machado, C. D. B., Ishizuka, T. K., Freitas, P. D. D., Valiati, V. H., and Galetti, P. M. (2016). DNA barcoding reveals taxonomic uncertainty in *Salminus* (Characiformes). *Syst. Biodivers* 15, 372–382. doi: 10.1080/14772000.2016.1254390

Melo, B. F., Sidlauskas, B. L., Hoekzema, K., Frable, B. W., Vari, R. P., and Oliveira, C. (2016). Molecular phylogenetics of the Neotropical fish family Prochilodontidae (Teleostei: Characiformes). *Mol. Phylogenet. Evol.* 102, 189–201. doi: 10.1016/j.ympev.2016.05.037

Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). "Creating the CIPRES science gateway for inference of large phylogenetic trees," in *Proceedings of the 2010 Gateway Computing Environments Workshop (GCE)*, (New Orleans, LA: IEEE), 1–8. doi: 10.1109/GCE.2010.5676129

Oliveira, C., Avelino, G. S., Abe, K. T., Mariguela, T. C., Benine, R. C., Ortí, G., et al. (2011). Phylogenetic relationships within the speciose family characidae (Teleostei: Ostariophysi: Characiformes) based on multilocus analysis and extensive ingroup sampling. *BMC Evol. Biol.* 11:275. doi: 10.1186/1471-2148-11-275

Pereira, L. H. G., Hanner, R., Foresti, F., and Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genet.* 14:20. doi: 10.1186/1471-2156-14-20

Pereira, L. H. G., Pazian, M. F., Hanner, R., Foresti, F., and Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (Characiformes: Characidae) from the Upper Paraná Basin of Brazil. *Mitochondrial DNA* 22(Suppl. 1), 87–96. doi: 10.3109/19401736.2011.588213

Pires, A. A., Ramirez, J. L., Galetti, P. M., Troy, W. P., and Freitas, P. D. (2017). Molecular analysis reveals hidden diversity in *Zungaro* (Siluriformes: Pimelodidade): a genus of giant South American catfish. *Genetica* 145, 335–340. doi: 10.1007/s10709-017-9968-8

Pons, J., Barraclough, T., Gomez-Zurita, J., Cardoso, A., Duran, D., Hazell, S., et al. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55, 595–609. doi: 10.1080/10635150600852011

Prioli, A. J., Carlo, V. A., Soria, T. V., Prioli, S. M., Pavanelli, C. S., Prioli, R. A., et al. (2012). Mitochondrial D-loop nucleotide diversity in Astyanax (Osteichthyes, Characidae) from the upper Paraná and upper Paraguay River basins. *Genet. Mol. Res* 11, 1064–1074. doi: 10.4238/2012.April.27.5

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Available at: https://www.r-project.org/

Rambaut, A., Suchard, M. A., Xie, D., and Drummond, A. J. (2014). *Tracer v1.6*. Available at: http://tree.bio.ed.ac.uk/software/tracer/

Ramirez, J. L., Birindelli, J. L., Carvalho, D. C., Affonso, P. R. A. M., Venere, P. C., Ortega, H., et al. (2017a). Revealing hidden diversity of the underestimated neotropical ichthyofauna: DNA barcoding in the recently described genus *Megaleporinus* (Characiformes: Anostomidae). *Front. Genet.* 8:149. doi: 10.3389/fgene.2017.00149

Ramirez, J. L., Birindelli, J. L. O., and Galetti, P. M. (2017b). A new genus of anostomidae (Ostariophysi: Characiformes): diversity, phylogeny and biogeography based on cytogenetic, molecular and morphological data. *Mol. Phylogenet. Evol.* 107, 308–323. doi: 10.1016/j.ympev.2016.11.012

Ramirez, J. L., Carvalho-Costa, L. F., Venere, P. C., Carvalho, D. C., Troy, W. P., and Galetti, P. M. (2016). Testing monophyly of the freshwater fish *Leporinus* (Characiformes. Anostomidae) through molecular analysis. *J. Fish Biol.* 88, 1204–1214. doi: 10.1111/jfb.12906

Ramirez, J. L., and Galetti, P. M. Jr. (2015). DNA barcode and evolutionary relationship within *Laemolyta* Cope 1872 (Characiformes: Anostomidae) through molecular analyses. *Mol. Phylogenet. Evol* 93, 77–82. doi: 10.1016/j.ympev.2015.07.021

Reis, R. E., Albert, J. S., Di Dario, F., Mincarone, M. M. M., Petry, P. L., and Rocha, L. R. (2016). Fish biodiversity and conservation in South America. *J. Fish Biol.* 89, 12–47. doi: 10.1111/jfb.13016

Renno, J.-F., Berrebi, P., Boujard, T., and Guyomard, R. (1990). Intraspecific genetica differentiation of *Leporinus friderici* (Anostomidae, Pisces) in French Guiana and Brazil: a genetic approach to the refuge theory. *J. Fish Biol.* 36, 85–95. doi: 10.1111/j.1095-8649.1990.tb03522.x

Renno, J.-F., Machordom, A., Blanquer, A., and Boursot, P. (1991). Polymorphism of mitochondrial genes in populations of *Leporinus friderici* (Block, 1794): intraspecific structure and zoogeography of the Neotropical fish. *Genetica* 84, 137–142. doi: 10.1007/BF00116554

Ribeiro, A. C. (2006). Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: an example of faunal evolution associated with a divergent continental margin. *Neotrop. Ichthyol.* 4, 225–246. doi: 10.1590/S1679-62252006000200009

Rossetti, D. F., and Valeriano, M. M. (2007). Evolution of the lowest amazon basin modeled from the integration of geological and SRTM topographic data. *Catena* 70, 253–265. doi: 10.1016/j.catena.2006.08.009

Sambrook, J., Fritish, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. NewYork, NY: Cold Spring Harbor Laboratory Press.

Sidlauskas, B. L., and Vari, R. P. (2008). Phylogenetic relationships within the South American fish family Anostomidae (Teleostei, Ostariophysi,Characiformes). *Zool. J. Linn. Soc.* 154, 70–210. doi: 10.1111/j.1096-3642.2008.00407.x

Sidlauskas, B. L., and Vari, R. P. (2012). Diversity and distribution of anostomoid fishes (Teleostei: Characiformes) throughout the Guianas. *Cybium* 36, 71–103.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446

Stamatakis, A., Hoover, P., and Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML Web servers. *Syst. Biol.* 57, 758–771. doi: 10.1080/10635150802429642

Steindachner, F. (1876). *Ichthyologische Beiträge (V). Sitzungsberichte der Kais*. Vienna: Gerold.

Swofford, D. L. (2003). *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version* 4. Sunderland, MA: Sinauer Associates.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Torrente-Vilara, G., Zuanon, J., Leprieur, F., Oberdorff, T., and Tedesco, P. A. (2011). Effects of natural rapids and waterfalls on fish assemblage structure in the Madeira River (Amazon Basin). *Ecol. Freshw. Fish* 20, 588–597. doi: 10.1111/j.1600-0633.2011.00508.x

Zanata, A. M., and Toledo-Piza, M. (2004). Taxonomic revision of the South American fish genus *Chalceus* Cuvier (Teleostei: Ostariophysi: Characiformes) with the description of three new species. *Zool. J. Linn. Soc.* 140, 103–135. doi: 10.1111/j.1096-3642.2004.00090.x

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 19 February 2018
doi: 10.3389/fgene.2018.00048

Genetic Variation of the Endangered Neotropical Catfish *Steindachneridion scriptum* (Siluriformes: Pimelodidae)

Rômulo V. Paixão[1,2], Josiane Ribolli[1,2] and Evoy Zaniboni-Filho[1,2]*

[1] *Laboratório de Biologia e Cultivo de Peixes de Água Doce, Departamento de Aquicultura, Universidade Federal de Santa Catarina, Florianópolis, Brazil,* [2] *Programa de Pós-Graduação em Aquicultura, Centro de Ciências Agrárias, Universidade Federal de Santa Catarina, Florianópolis, Brazil*

Steindachneridion scriptum is an important species as a resource for fisheries and aquaculture; it is currently threatened and has a reduced occurrence in South America. The damming of rivers, overfishing, and contamination of freshwater environments are the main impacts on the maintenance of this species. We accessed the genetic diversity and structure of *S. scriptum* using the DNA barcode and control region (D-loop) sequences of 43 individuals from the Upper Uruguay River Basin (UUR) and 10 sequences from the Upper Paraná River Basin (UPR), which were obtained from GenBank. *S. scriptum* from the UUR and the UPR were assigned in two distinct molecular operational taxonomic units (MOTUs) with higher inter-specific K2P distance than the optimum threshold (OT = 0.0079). The COI Intra-MOTU distances of *S. scriptum* specimens from the UUR ranged from 0.0000 to 0.0100. The control region indicated a high number of haplotypes and low nucleotide diversity, compatible with a new population in recent expansion process. Genetic structure was observed, with high differentiation between UUR and UPR basins, identified by BAPS, haplotype network, AMOVA (F_{ST} = 0.78, $p < 0.05$) and Mantel test. *S. scriptum* from the UUR showed a slight differentiation (F_{ST} = 0.068, $p < 0.05$), but not isolation-by-distance. Negative values of Tajima's *D* and Fu's *Fs* suggest recent demographic oscillations. The Bayesian skyline plot analysis indicated possible population expansion from beginning 2,500 years ago and a recent reduction in the population size. Low nucleotide diversity, spatial population structure, and the reduction of effective population size should be considered for the planning of strategies aimed at the conservation and rehabilitation of this important fisheries resource.

Keywords: control region mitochondrial DNA, conservation of natural resources, DNA barcode, endangered species, freshwater fishes

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

Reviewed by:
Jorge Luis Ramirez, Federal University of São Carlos, Brazil
Jorge Abdala Dergam, Universidade Federal de Viçosa, Brazil

****Correspondence:***
Rômulo V. Paixão
romulo.veiga.paixao@gmail.com

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

Received: *17 October 2017*
Accepted: *31 January 2018*
Published: *19 February 2018*

Citation:
Paixão RV, Ribolli J and Zaniboni-Filho E (2018) Genetic Variation of the Endangered Neotropical Catfish Steindachneridion scriptum (Siluriformes: Pimelodidae). Front. Genet. 9:48. doi: 10.3389/fgene.2018.00048

INTRODUCTION

Freshwater ecosystems are among the most endangered ecosystems (Dudgeon et al., 2006). Habitat degradation, hydrologic alterations, habitat fragmentation, sediment deposition, and overfishing are the principal causes of declines and extinctions of freshwater fishes (Dudgeon et al., 2006; Agostinho et al., 2008; Helfman, 2008; Hoeinghaus et al., 2009). In addition, species with

geographically restricted distribution are more susceptible to erosion of the genetic diversity due to habitat fragmentation (Vrijenhoek et al., 1985). Fishery resources are an integral part of most societies and make important contributions to economic and social health and well-being in many countries and areas (FAO, 2002). The understanding of the genetic diversity and structure of wild populations of fish species are important to the regulation of fisheries and conservation management strategies (Carvalho and Hauser, 1994; Iervolino et al., 2010).

Steindachneridion scriptum (Miranda Ribeiro, 1918) is a large catfish belonging to the family Pimelodidae. This potamodromous fish species presents restricted distribution in the Upper Uruguay River (UUR) and Upper Paraná River (UPR) basins (Marques et al., 2002). *S. scriptum* is an important fishing resource to the riverine fishermen (Schork et al., 2013); nonetheless, it was recently classified as an endangered species according to the Chico Mendes Institute for Biodiversity Conservation (ICMBIO, 2014). Human activities (e.g., damming of rivers, illegal fishing, and industrial waste) are the main threats to *S. scriptum* in the UUR (Fontana et al., 2003) and are the principle reason for population reduction in this basin (Beux and Zaniboni-Filho, 2008).

The correct management of fish stocks depends on the precise identification of the target species, which may present very similar morphological characteristics to other species. The latest revision of the genera *Steindachneridion* recognizes six valid species (Garavello, 2005). Using museum specimen, Garavello (2005) suggests the presence of *S. punctatum* in the UUR and UPR with few characteristics that differ from *S. scriptum*. Despite the description of the two species of the genus *Steindachneridion*, ichthyofauna studies never report the presence of *S. punctatum* in the UUR (Schork et al., 2012, 2013). In addition to the correct ichthyofauna and management of fish stocks, uncertainties of taxonomic identification may be a problem for stock management (e.g., formation of *in vivo* and *in vitro* banks, restocking programs) and adequate fisheries control.

Taxonomic uncertainties are common in fishes (Pereira et al., 2013) and can be investigated using DNA barcode methodology, which permits the unambiguous identification of the majority of fish species (Ward, 2009). Particularly for endangered species, prior knowledge of the distribution of genetic variability within and among natural populations as well as the implementation of an efficient management plan based on genetic features are important measures for its maintenance and recovery (Cross, 2000). The distribution of the genetic variation within and between populations can be assessed using the mitochondrial control region (Sivasundar et al., 2001; Iervolino et al., 2010; Ochoa et al., 2015). Given these findings, we tested the null hypothesis that *S. scriptum* from UUR represent a single molecular operational taxonomic unit (MOTU). Posteriorly, we investigate the genetic diversity and population structure of *S. scriptum* from the UUR and UPR basins using the mitochondrial control region.

MATERIALS AND METHODS

Study Area and Sampling

The Uruguay River (Uru) originates in Brazilian territory (in the Serra Geral Mountains) in Southern Brazil, together with the Paraná and Paraguay Rivers form the La Plata Basin (Zaniboni-Filho and Schulz, 2003). Samples of *S. scriptum* were collected by scientific fishing and local fishermen between 2006 and 2015, with authorization of the Brazilian Institute of the Environment and Renewable Natural Resources (IBAMA; protocol number: 02026.005762/2004-71). A total of 19 individuals from the Uru River and 24 from the Canoas River (Can), UUR basin, were sampled (**Figure 1**). Tissues were preserved in 95% ethanol until extraction. This research was conducted under Animal Care Protocol PP00788 of the Federal University of Santa Catarina (UFSC).

DNA Extraction and Amplification of the Mitochondrial Fragments

Total DNA was obtained with fin clips following a salt extraction method (Aljanabi and Martinez, 1997). For the DNA barcode analyses, a fragment of 652 bp of Cytochrome Oxidase subunit I (COI) was amplified through polymerase chain reaction (PCR), using primers FishF1/FishR1 (Ward et al., 2005) and following Bellafronte et al. (2013). The partial amplification of the mitochondrial control region (D-loop) was performed using primers FTTP-L and DLR1-H according to Huergo et al. (2011). The PCR products were checked for amplification using gel electrophoresis with 1% agarose gels purified using PEG 20% (Lis, 1980). Sequencing reactions were performed using BigDye TM Terminator v 3.1 (Applied Biosystems), and the PCR products were sequenced for both strands in ABI 3500XL (Applied Biosystems).

Data Analysis

DNA sequences from each individual of both genes were edited using Geneious 5.4.4 (Wu and Drummond, 2011) to generate a consensus sequence. For the DNA barcode analysis, we combined the COI reference sequences of *S. scriptum* (access FUPR686-09, PDCAP027-14, and PDCAP028-14) from the UPR basin, *S. parahybae* (access FPSR293-10–FPSR297-10) from the Paraíba do Sul Basin and one specimen of *Pseudoplatystoma corruscans* and one specimen of *Zungaro jahu* to root our phylogenetic analyses. All sequences available in Barcode of Life Data System (BOLD). Intra- and inter-specific genetic distances based on the Kimura 2-parameter (K2P) evolution model were calculated using Mega 6 (Tamura et al., 2013).

We used the phylogenetic General Mixed Yule Coalescent (GMYC) approach based on single-locus data that is a relative robust tool for species delimitation (Pons et al., 2006; Fujisawa and Barraclough, 2013). The ultrametric tree was generated in BEAST v.2.2.1 (Bouckaert et al., 2014), with the substitution model calculated in the JModelTest 2.1.4 (HKY+G; Darriba et al., 2012), using relaxed molecular clock with a lognormal distribution and birth–death model. Three independent runs were carried out with 20 million generations each. Posteriorly, the

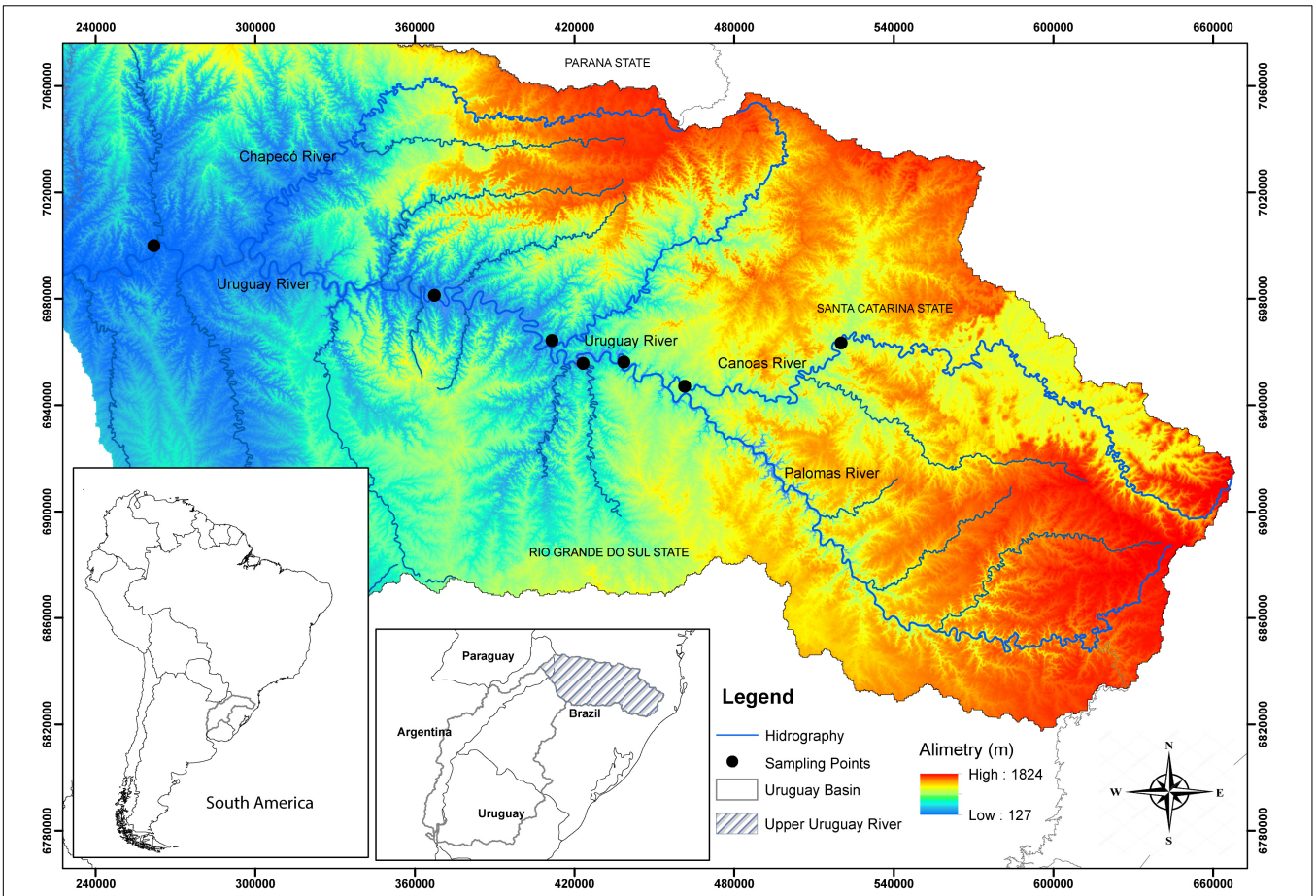


FIGURE 1 | Map showing the locations of sampling sites of the *Steindachneridion scriptum* in the Upper Uruguay River Basin. Black dots indicate sample points.

runs were combined using the LogCombiner v.1.8.3 (Drummond et al., 2012), with a burn-in of 25. Data mixing and effective sample size (ESS) were verified in Tracer v1.5. GMYC was carried out in Species Limits by Threshold Statistics (SPLITs; Monaghan et al., 2009) with RStudio[1], using the unique threshold method to detect the transition point between intra- and inter-specific relationships.

In addition to the standard threshold adopted to Neotropical fishes (Pereira et al., 2013), we calculated an optimum threshold (OT; Collins et al., 2012) directly from all dataset, using the *local minima* function in the R package SPIDER (SPecies IDentity and Evolution in R; Brown et al., 2012). The OT value was used to define the MOTUs using the software jMOTU (Jones et al., 2011). The graphical representation of the MOTUs was performed by a neighbor-joining analysis (NJ) using the K2P model with Mega 6.6 (Tamura et al., 2013). The support of the clades was tested by the bootstrap method with 10,000 pseudo-replicates.

DNA D-loop sequences were combined with 16 *S. scriptum* sequences downloaded from GenBank, of specimens collected between 1995 and 2002 that corresponded to the Uru River (access EU930029.1–EU930038.1) and specimens from Tibaji River (UPR) (access EU930039.1–EU930044.1). The overall genetic diversity was estimated using the following DnaSP software (Rozas et al., 2003) parameters: nucleotide diversity (π) (Nei, 1987), haplotype diversity (Hd) (Nei and Tajima, 1981), and number of polymorphic sites (S). Genetic diversity within and between sample sites was hierarchically tested by Analysis of Molecular Variance (AMOVA) (Excoffier et al., 1992) with 10,000 permutations to test the pairwise population comparison (F_{ST}) using Arlequin 3.1 (Excoffier and Lischer, 2010). Spatial genetic structure was inferred using Bayesian Analysis of Population Structure 6.0 (BAPS) software (Corander et al., 2013). First, BAPS was run with 10 replicates for every level of k (1–6) without origin information ("clustering of individuals") and subsequently using "clustering of groups of individuals." We used Mantel tests as implemented in the Alleles In Space (AIS) 1.0 (Miller, 2005) to test for a correlation of geographic stream distance and genetic distance (isolation-by-distance; IBD) expressed as F_{ST}, with 10,000 permutations to assess significance. Tajima's (1989) D, Fu's (1997) Fs, and mismatch distributions were estimated with *DnaSP* (Rozas et al., 2003). A median-joining haplotype network was generated through PopART (Bandelt et al., 1999). Demographic history was investigated using Bayesian Skyline Plot (BSP) in BEAST 2.1.3 (Drummond et al., 2012) with the

[1]http://r-forge.r-project.org/projects/splits

evolutionary model obtained in the Jmodeltest program. The graphic was generated in Tracer 1.66 (Rambaut et al., 2014).

The COI sequences were deposited in BOLD systems (accession UUR001-17–UUR052-17), and DNA D-loop sequences were uploaded in GenBank (accession MF045370–MF045412). The voucher corresponding to *S. scriptum* from the UUR basin was deposited in the Zoology Museum of the Universidade Estadual de Londrina MZUEL 15569 (Type locality: Itaqui, RS – Uru Basin).

TABLE 1 | Intra-MOTU distances (in bold) and inter-MOTU genetic distances using the COI gene and K2p model.

	1	2	3
1 *S. scriptum* UUR	**0.001 (0.000–0.010)**		
2 *S. scriptum* UPR	0.012 (0.010–0.017)	**0.000 (0.000)**	
3 *S. parahybae*	0.102 (0.100–0.106)	0.090 (0.090)	**0.003 (0.000–0.003)**

Distance range in parentheses. UUR = Upper Uruguay River; UPR = Upper Paraná River.

RESULTS

Mitochondrial DNA Barcoding

The consensus alignment of 43 COI sequences were obtained from samples identified morphologically as *S. scriptum*, resulting in a total length of 611 bp, with 7 polymorphic sites and 6 haplotypes defined. The inter-specific nucleotide frequencies were 26.02% of Cytosine, 28.97% of Thymine, 27.19% of Adenine, and 17.82% of Guanine. No stop codons, insertions, or deletions were observed in the COI sequences, indicating that they represent fragments of functional mitochondrial genes and not nuclear mitochondrial pseudo-genes (Numts). Considering *P. corruscans* and *Z. jahu* as the outgroup, the maximum likelihood for the GMYC model was significantly superior ($L = 628.3242$) to the likelihood of the null model ($Lo = 614.7156$, $p < 0.0001$). The single-threshold GMYC model suggested the presence two clusters (confidence interval 2–13) of four ML entities (GMYC 'species,' named 'MOTUs' herein) with a confidence interval of 4–17 (*S. scriptum*, *S. parahybae*, and two outgroups).

The optimal threshold calculated for all *S. scriptum* sequences used in this study was OT = 0.0079 (0.79%) of divergence. From the set of the genus *Steindachneridion* sequences available on BOLD, three MOTUs were identified using the OT value and software jMOTU: *S. scriptum* from the UUR, *S. scriptum* from the UPR, and *S. parahybae*. The COI inter-MOTU between *S. scriptum* from the UUR and *S. scriptum* from the UPR showed values larger than the OT (mean 0.012, minimum 0.010), while intra-MOTU values for fishes from both basins were lower than the average OT (UUR = 0.000 to 0.010; UPR = 0.000) (**Table 1**). The minimum inter-specific distances between *S. scriptum* MOTUs (UUR and UPR) and *S. parahybae* were 0.100 (10%) and 0.090 (9%), respectively.

The NJ-K2P (Supplementary Figure SI 1) and Bayesian Inference topologies (**Figure 2**) were clustered in two principal clusters, corresponding to *S. scriptum* and *S. parahybae* species. The *S. scriptum* clade appeared divided into two well-supported sub-clades formed by UUR and UPR individuals, in both NJ and BI phylogenetic trees. Based on these results, all specimens from Uru were composed of a single MOTU, named here as *S. scriptum*.

Mitochondrial DNA D-Loop

The final alignment size of the 59 D-loop consensus sequences of *S. scriptum* specimens from the UUR and UPR (43 newly sequenced, 16 downloaded from GenBank) was 865 bp. A total of 56 variable sites were found in the region defining a total of 36 haplotypes. Overall, 30 haplotypes identified in the analyses were unique and exclusive. Of these, 20 were from the Uru, 4 from Can, and 6 from UPR basin (Tibagi River). The most common haplotype was H8, which was recorded 12 times and was shared by samples from the Uru and Can rivers, and haplotypes H4 and H27, both with three records, were exclusive to the Uru and Can rivers, respectively. The haplotype network (**Figure 3**) revealed a high degree of similarity between the specimens from the Uru and Can rivers, even though with slight differences in the haplotype frequencies and exclusive haplotypes from each river. These specimens were also differentiated from the samples of UPR by 12 mutations.

The average nucleotide frequencies were 34.38% of Adenine, 33.17% of Thymine, 12.19% of Guanine, and 20.26% of Cytosine. Genetic variability, expressed as Hd) and nucleotide diversity (π), was higher in *S. scriptum* from the Uru (Hd = 0.959/π = 0.007) in comparison with the Can River (Hd = 0.837/π = 0.004) (**Table 2**). In addition, the samples from UPR have high diversity indices, in comparison with samples from UUR ($N = 6$; $Nh = 6$; Hd = 1.0000, $\pi = 0.00698$; $D = -0.88901$, $p = 0.24002$; $Fs = -1.81313$, $p = 0.07768$).

The patterns of genetic variability found within and between populations in the AMOVA were based on the two principal clusters: Uruguay vs. Paraná Basins and Uru vs. Can rivers (**Table 3**). When the populations were considered as two basin groups, the AMOVA among groups was 78.17% and F_{ST} value was highly significant ($F_{ST} = 0.781$; $p = 0.000$). Genetic divergence between individuals from the Uruguay Basin (Uru vs. Can rivers) was low but significant ($F_{ST} = 0.0682$, $p = 0.00475$). The population groupings generated by the BAPS, without origin information of the samples, revealed the existence of three clusters ($K = 3$, Supplementary Figure SI 2), with slight differentiation between Can and Uru individuals. On the other hand, analysis with the individuals identified by sample group indicated two clusters ($K = 2$, **Figure 4**), corresponding to Uruguay and Paraná Basins. The IBD analysis showed a significant positive correlation ($r^2 = 0.85$, $p < 0.001$) between the geographical distance and corresponding F_{ST} for *S. scriptum* from UUR and UPR (Supplementary Figure SI 3). On the other hand, F_{ST} values plotted over distance no reveal patter of isolation by distance for S. scriptum from UUR ($r = -0.052$, $p = 0.682$).

Tajima's (1989) *D*-neutrality tests, applied to detect evidence of strong selective pressures, and Fu's (1997) *F*s-tests, used

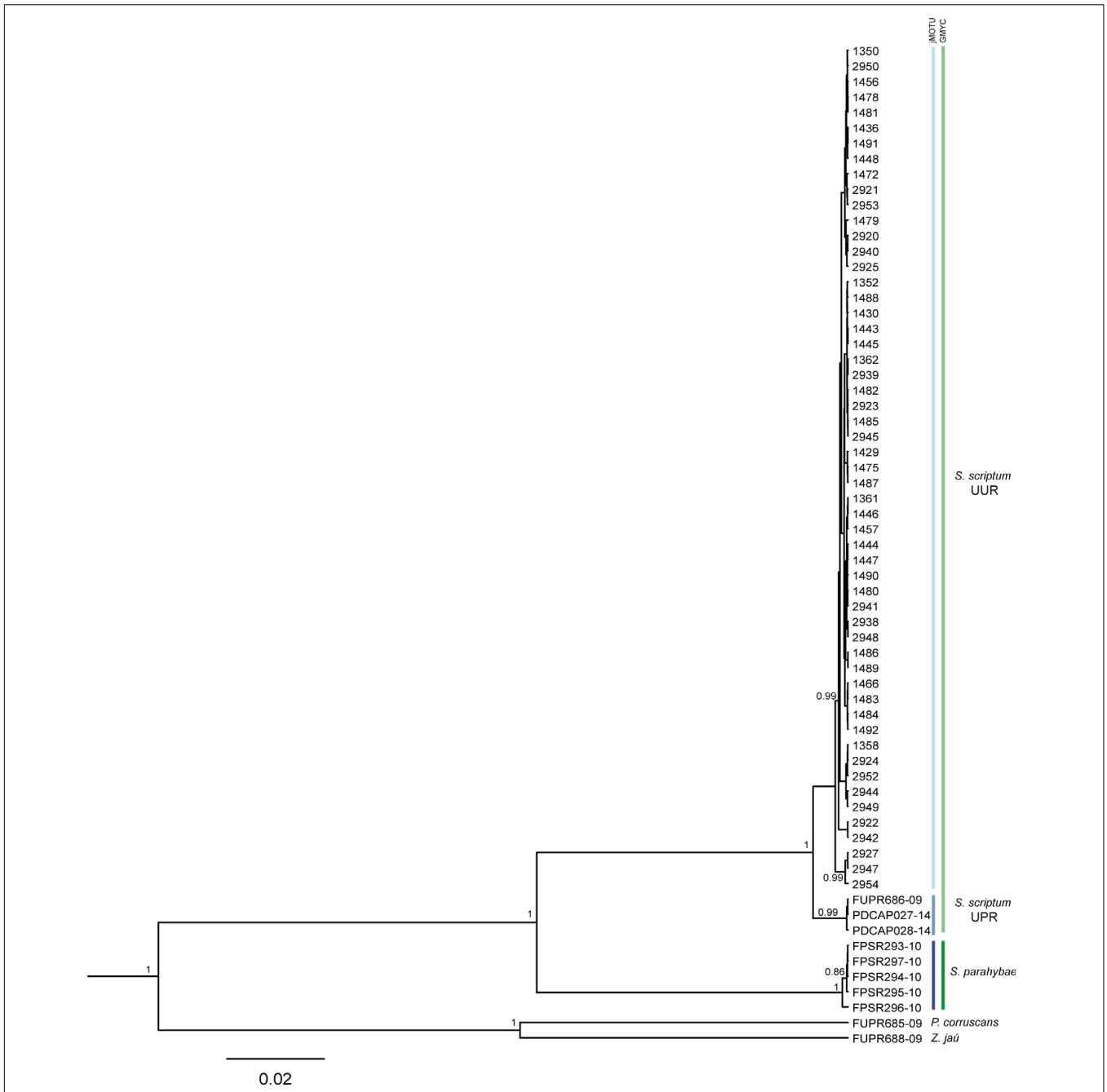


FIGURE 2 | Phylogenetic tree of the genus *Steindachneridion* based on the mitochondrial gene COI inferred with BEAST. Values before nodes are Bayesian posterior probabilities. Codes FUPR, PDCAP, and FPSR were downloaded from the BOLD system. *Pseudoplatystoma corruscans* and *Z. jahu* are outgroups. Vertical lines in Blue = jMOTU MOTUs, and vertical lines in Green indicate GMYC MOTUs.

specifically to detect population expansion, revealed significant negative values for all individuals from the UUR ($D = -1.907$, $p < 0.05$; $Fs = -19.246$, $p < 0.01$; **Table 3**). Non-significant negative values were estimated for D and Fs indexes for specimens from the Can River and UPR, whereas fish sampled in the Uru showed significant negative values estimated for D and Fs indexes. The BSP analysis (Supplementary Figure SI 4), used to explore previous demographic signals of *S. scriptum,* indicated early demographic expansion approximately 2,500 years ago as well as a fairly recent population reduction.

DISCUSSION

The DNA barcode confirmed the identification of all the individuals of *S. scriptum* from UUR as a single MOTU.

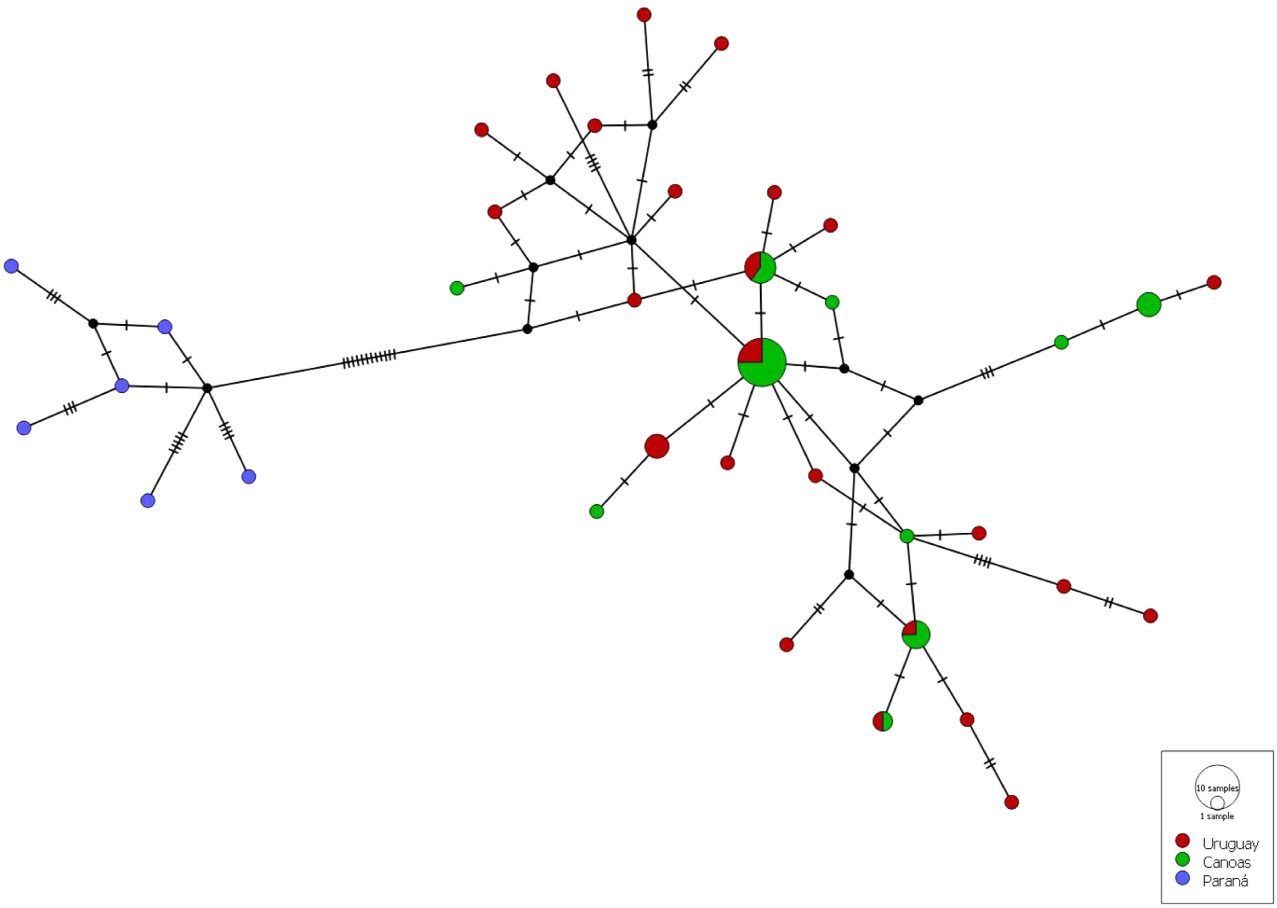


FIGURE 3 | Median-joining network of *Steindachneridion scriptum*, based on haplotypes of mtDNA control region. The colors indicate locality according to the legend, size of the circles illustrates the number of identical haplotypes, and small black circles, hypothetical ancestors or unsampled haplotypes. Hatch marks represent the number of mutations by which haplotypes differ.

TABLE 2 | Genetic diversity of *Steindachneridion scriptum* from the Uruguay and Canoas rivers (Upper Uruguay River Basin) estimated using D-loop control region.

Sampling Site	*N*	*Nh*	Hd	π	*D*	*Fs*
Canoas River	24	12	0.837	0.004	−0.049	−2.223
Uruguay River	29	23	0.959	0.007	−2.016*	−16.346*
Total	53	32	0.932	0.005	−1.907*	−19.246**

*N = number of sequenced individuals; Nh = number of haplotypes; Hd = haplotype diversity; π = nucleotide diversity; D = Tajima's D-test; Fs = Fu's Fs-test. Significant values to D-test and Fs-test: *p < 0.05, **p < 0.01.*

The different methods, jMOTU and GMYC, were congruent in delimiting *S. scriptum* and *S. parahybae*. However, within *S. scriptum*, jMOTU methodology identified two distinct MOTUs between UUR basin and UPR basin, while GMYC only one, despite strongly supported clade. Inter-MOTU divergence between *S. scriptum* from the UUR and the UPR was higher than the OT and the mean intra-specific divergence found for freshwater fish (0.3%) (Ward et al., 2005; Lara et al., 2010; Pereira et al., 2013). The MOTUs do not necessarily represent species (Blaxter et al., 2005) but can indicate molecular entities (Casiraghi et al., 2010). The two MOTUs estimated between *S. scriptum* from the UUR and *S. scriptum* from the UPR could be explained by geographic isolation between watersheds that occurred during the Miocene epoch (between 5 and 24 million years ago) when the UUR and the UPR became isolated (Albert and Reis, 2011). Based on the results, *S. scriptum* from these two hydrographic systems are most likely in the process of incipient allopatric speciation since the genetic structuring in fish is in fact often evidenced and influenced principally by geological, ecological, and behavioral factors (Allan and Flecker, 1993). Congeners *S. scriptum* and *S. parahybae* species showed mean inter-MOTUs 10 times greater than the OT, indicating the existence of the barcode gap (Hebert et al., 2003, 2004) that allows us to

TABLE 3 | Hierarchical AMOVA analysis and F_{ST} values for *Steindachneridion scriptum* according to their geographical location, estimated using D-loop control region.

Type of variation	Component of variation	% of variation	F_{ST} (*p*)
UUR × UPR			
Between Basins	9.22585	78.17	0.78167*(0.0000)
Within Basins	2.57696	21.83	
Upper Uruguay Basin			
Between populations	0.17916	6.83	0.06825* (0.0047)
Within populations	2.44580	93.17	

*UUR = Upper Uruguay River; UPR = Upper Parana River. *Significant value: p-value (F_{ST}) = < 0.05.*

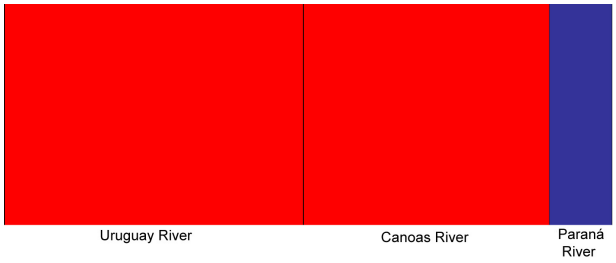


FIGURE 4 | Estimate of the probable groups of populations produced by the BAPS between *Steindachneridion scriptum* from Upper Uruguay and Upper Parana Basins, assigned to two clusters (*K* = 2).

assign an unknown *Steindachneridion* specimen to its species using a genetic distance criterion with an insignificant error rate.

Although they belong to the same watershed, individuals from the Can River and the Uru River showed a slight genetic differentiation probably due to the topography of the region and the interaction between the species' biology and environmental characteristics (e.g., the Can River is located at a higher altitude with a lower water temperature than the Uru River). The haplotype network indicated a greater genetic similarity between the specimens from the Can and Uru, whereas is possible to observe differences in the haplotype frequencies, and exclusive haplotypes for each river. Recent studies with potamodromous fish species reported genetic structure in hydrographic systems without apparent physical barriers, resulting from behaviors related to IBD (Hardy and Vekemans, 1999; Primmer et al., 2006; Han et al., 2010), *homing* (Windle and Rose, 2005; Batista and Alves-Gomes, 2006; Neville et al., 2006), and isolation-by-time (IBT) (Hendry and Day, 2005; Braga-Silva and Galetti, 2016; Ribolli et al., 2017).

Nucleotide diversity of *S. scriptum* from the UUR was low in comparison with the values found in neotropical freshwater fish (π = 1.5%) (Batista and Alves-Gomes, 2006; Iervolino et al., 2010; Ashikaga et al., 2015). The high haplotype diversity and low nucleotide diversity seem quite compatible with a new population in recent expansion process, similar to what is shown in BSP analyses. This pattern may be a signature of such expansion that is long enough to examine a change in the haplotypes resulted from the mutation, but is not long enough to accumulate large differences between sequences (Avise, 2000).

Individuals from the Uru River were genetically more diverse than fishes from the Can River, indicating that the main channel of the Uru River allows the meeting of individuals of different areas (or tributaries), favoring the maintenance of the highest level of genetic diversity. Although it is a relevant fishing resource, useful molecular markers such as microsatellites, extensively employed in fish genetic studies, are still not developed for any *Steindachneridion* species, and the knowledge about genetic characteristics of the genus is incipient. RAPD markers indicated low genetic diversity for *S. scriptum* from the UUR, as reported by Ramella et al. (2006), as well as for *S. melanodermatum* from the Iguaçu River Basin (Matoso et al., 2011). Low genetic diversity may indicate recent or historic reduction of this diversity; however, some endangered populations may have a historic maintenance of small effective population sizes (Matocq and Villablanca, 2001). In this way, the low diversity detected in *S. scriptum* from the UUR can be attributed to the following: (1) population history of this species: a Bayesian skyline plot analysis revealed a subtle increase in effective population size over time and demographic swings in the recent past, with a notable increase in effective population size between 1,000 and 2,500 years ago and a subsequent reduction in the effective size of females; and (2) evolutionary history of the species: according to Garavello (2005) and Swarça et al. (2005), *S. scriptum* notably maintain conserved morphological and cytogenetic patterns.

In general, concerns and actions of conservation are more related to the perception of the disappearance of a given species than of genetic diversity reduction (Frankham, 2005). Therefore, given the low genetic diversity associated with the current scenario of fragmentation of the UUR basin and the population reduction of *S. scriptum* in some stretches of the Itá and Machadinho reservoirs (Schork et al., 2012, 2013), this study highlights the necessity of mitigation measures and more intense monitoring of illegal fishing to avoid the collapse of this important fishing resource. In addition, our results were congruent, identifying great differentiation between individuals from the UPR and UUR Basins. Further studies with a larger number of samples morphological analyze to better define the taxonomic status of endangered *S. scriptum*.

AUTHOR CONTRIBUTIONS

RP made substantial contributions to the design of the work and acquisition, analysis, and interpretation of data for the work, and drafted it until the approval for publication of the content. JR made substantial contributions to the conception and design of the work, analysis and interpretation of the data, and drafted and revised it critically for important intellectual content. EZ-F made substantial contributions to the conception of the work,

revised it critically for important intellectual content, and provided approval for publication of the content. The authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

FUNDING

This work was supported by the Engie Tractebel Energia, Consórcio Machadinho e Consórcio Itá. RP thanks Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); JR acknowledges the Programa Nacional de Pós-Doutorado (PNPD/CAPES); EZ-F thanks Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Grant No. 302860/2014-2).

ACKNOWLEDGMENTS

We are grateful to Laboratório de Biologia e Cultivo de Peixes de Água Doce (LAPAD) of Universidade Federal de Santa Catarina (UFSC), Pedro Iaczinki, and local fishermen for help with fish collections. We thank Dra. Carolina Machado, Leonardo Porto Ferreira, and Ueslei Lopes for their contributions. We also thank the two reviewers for their valuable contributions on improving the manuscript and for the suggestions. Research was conducted under Animal Care Protocol PP00788.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00048/full#supplementary-material

REFERENCES

Agostinho, A. A., Pelicice, F. M., and Gomes, L. C. (2008). Dams and the fish fauna of the Neotropical region: impacts and management related to diversity and fisheries. *Braz. J. Biol.* 68, 1119–1132. doi: 10.1590/S1519-69842008000500019

Albert, J. S., and Reis, R. (2011). *Historical Biogeography of Neotropical Freshwater Fishes*. Berkeley, CA: University of California Press, 162–165. doi: 10.1093/sysbio/syr113

Aljanabi, S. M., and Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* 25, 4692–4693. doi: 10.1093/nar/25.22.4692

Allan, J. D., and Flecker, A. S. (1993). Biodiversity conservation in running waters. *Bioscience* 43, 32–43. doi: 10.2307/1312104

Ashikaga, F. Y., Orsi, M. L., Oliveira, C., and Senhorini, J. A. (2015). The endangered species *Brycon orbignyanus*: genetic analysis and definition of priority areas for conservation. *Environ. Biol. Fishes* 98, 1845–1855. doi: 10.1007/s10641-015-0402-8

Avise, J. C. (2000). *Phylogeography: the History and Formation of Species*. Cambridge: Harvard University Press, 447.

Batista, J. S., and Alves-Gomes, J. A. (2006). Phylogeography of *Brachyplatystoma rousseauxii* (Siluriformes - Pimelodidae) in the Amazon Basin offers preliminary evidence for the first case of "homing" for an Amazonian migratory catfish. *Genet. Mol. Res.* 5, 723–740.

Bandelt, H. J., Forster, P., and Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48. doi: 10.1093/oxfordjournals.molbev.a026036

Bellafronte, E., Mariguela, T. C., Henrique, L., Pereira, L. H. G., Oliveira, C., and Moreira-Filho, O. (2013). DNA barcode of Parodontidae species from the La Plata river basin - applying new data to clarify taxonomic problems. *Neotrop. Ichthyol.* 11, 497–506. doi: 10.1590/S1679-62252013000300003

Beux, L., and Zaniboni-Filho, E. (2008). "Produção pesqueira do reservatório de Itá. *Alto Rio Uruguai*," in *Reservatório de Itá: Estudos Ambientais. Desenvolvimento de Tecnologias de Cultivo e Conservação da Ictiofauna*, ed. E. Z. Filho (Florianópolis: UFSC), 87–108.

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., et al. (2005). Defining operational taxonomic units using DNA barcode data. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1935–1943. doi: 10.1098/rstb.2005.1725

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., et al. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10:e1003537. doi: 10.1371/journal.pcbi.1003537

Braga-Silva, A., and Galetti, P. M. (2016). Evidence of isolation by time in freshwater migratory fish *Prochilodus costatus* (Characiformes, Prochilodontidae). *Hydrobiologia* 765, 159–167. doi: 10.1007/s10750-015-2409-8

Brown, S. D. J., Collins, R. A., Boyer, S., Lefort, M. C., Malumbres-Olarte, J., Vink, C. J., et al. (2012). Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. *Mol. Ecol. Resour.* 12, 562–565. doi: 10.1111/j.1755-0998.2011.03108.x

Carvalho, G., and Hauser, L. (1994). Molecular genetics and the stock concept in fisheries. *Rev. Fish Biol. Fish.* 4, 326–350. doi: 10.1007/BF00042908

Casiraghi, M., Labra, M., Ferri, E., Galimberti, A., and de Mattia, F. (2010). DNA barcoding: a six-question tour to improve users' awareness about the method. *Brief. Bioinform.* 11, 440–453. doi: 10.1093/bib/bbq003

Collins, R. A., Boykin, L. M., Cruickshank, R. H., and Armstrong, K. F. (2012). Barcoding's next top model: an evaluation of nucleotide substitution models for specimen identification. *Methods Ecol. Evol.* 3, 457–465. doi: 10.1111/j.2041-210X.2011.00176.x

Corander, J., Marttinen, P., Sirén, J., and Tang, J. (2013). *BAPS: Bayesian Analysis of Population Structure*. Helsinki: University of Helsinki, 28.

Cross, T. F. (2000). Genetic implications of translocation and stocking of fish species, with particular reference to Western Australia. *Aquac. Res.* 31, 83–94. doi: 10.1046/j.1365-2109.2000.00439.x

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 30:772. doi: 10.1038/nmeth.2109

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. doi: 10.1093/molbev/mss075

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z. I., Knowler, D. J., Lévêque, C., et al. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biol. Rev. Camb. Philos. Soc.* 81, 163–182. doi: 10.1017/S1464793105006950

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Excoffier, L., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.

FAO (2002). *A Fishery Manager's Guidebook. Management Measures and Their Application*, 2nd Edn. Rome: FAO.

Fontana, C. S., Bencke, G. A., and Reis, R. E. (2003). *Livro Vermelho da Fauna Ameaçada de Extinção no Rio Grande do Sul*. Porto Alegre: Edipucrs.

Frankham, R. (2005). Genetics and extinction. *Biol. Conserv.* 126, 131–140. doi: 10.1016/j.biocon.2005.05.002

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.

Fujisawa, T., and Barraclough, T. G. (2013). Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method

and evaluation on simulated data sets. *Syst. Biol.* 62, 707–724. doi: 10.1093/sysbio/syt033

Garavello, J. C. (2005). Revision of genus Steindachneridion (Siluriformes: Pimelodidae). *Neotrop. Ichthyol.* 3, 607–623. doi: 10.1590/S1679-62252005000400018

Han, Y., Hung, C., Liao, Y., and Tzeng, W. (2010). Population genetic structure of the Japanese eel *Anguilla japonica*: panmixia at spatial and temporal scales. *Mar. Ecol. Prog. Ser.* 401, 221–232. doi: 10.3354/meps08422

Hardy, O. J., and Vekemans, X. (1999). Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity* 83, 145–154. doi: 10.1046/j.1365-2540.1999.00558.x

Hebert, P. D. N., Ratnasingham, S., and Jeremy, R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. Biol. Sci.* 270, 96–99. doi: 10.1098/rsbl.2003.0025

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S., and Francis, C. M. (2004). Identification of Birds through DNA Barcodes. *PLoS Biol.* 2:e312. doi: 10.1371/journal.pbio.0020312

Helfman, G. S. (2008). *Fish Conservation: A Guide to Understanding and Restoring Global Aquatic Biodiversity and Fishery Resources*. Washington, DC: Island Press.

Hendry, A. P., and Day, T. (2005). Population structure attributable to reproductive time: isolation by time and adaptation by time. *Mol. Ecol.* 14, 901–916. doi: 10.1111/j.1365-294X.2005.02480.x

Hoeinghaus, D. J., Agostinho, A. A., Gomes, L. C., Pelicice, F. M., Okada, E. K., Latini, J. D., et al. (2009). Effects of river impoundment on ecosystem services of large tropical rivers: embodied energy and market value of artisanal fisheries. *Conserv. Biol.* 23, 1222–1231. doi: 10.1111/j.1523-1739.2009.01248.x

Huergo, G. M., Filgueiras-Souza, R. J., Batista, J. D. S., Formiga-Aquino, K., and Alves-Gomes, J. A. (2011). Molecular genetics as a tool for fisheries management in the Brazilian Amazon: Piraíba (*Brachyplatystoma filamentosum* and *Brachyplatystoma capapretum*) (Siluriformes: Pimelodidae) in white-water rivers. *Pan Am. J. Aquat. Sci.* 6, 280–289.

ICMBIO (2014). *Lista de Especies Ameaçadas. Portaria MMA no 445*. Brasília: ICMBIO.

Iervolino, F., Kawakami, E., Resende, D., Wagner, A., and Hilsdorf, S. (2010). The lack of genetic differentiation of pacu (*Piaractus mesopotamicus*) populations in the Upper-Paraguay Basin revealed by the mitochondrial DNA D-loop region: implications for fishery management. *Fish. Res.* 101, 27–31. doi: 10.1016/j.fishres.2009.09.003

Jones, M., Ghoorah, A., and Blaxter, M. (2011). jMOTU and Taxonerator: turning DNA barcode sequences into annotated operational taxonomic units. *PLoS One* 6:e19259. doi: 10.1371/journal.pone.0019259

Lara, A., Ponce de León, J. L., Rodríguez, R., Casane, D., Côté, G., Bernatchez, L., et al. (2010). DNA barcoding of Cuban freshwater fishes: evidence for cryptic species and taxonomic conflicts. *Mol. Ecol. Resour.* 10, 421–430. doi: 10.1111/j.1755-0998.2009.02785.x

Lis, J. T. (1980). Fractionation of DNA fragments by polyethylene glycol induced precipitation. *Methods Enzymol.* 65, 347–353. doi: 10.1016/S0076-6879(80)65044-7

Marques, A. A. B., Fontana, C. S., Vélez, E., Bencke, G. A., Schneider, M., and Reis, R. E. (2002). *Lista das Espécies da Fauna Ameaçadas de Extinção no Rio Grande do Sul*. Porto Alegre: Fundação Zoobotânica do Rio Grande do Sul.

Matocq, M. D., and Villablanca, F. X. (2001). Low genetic diversity in an endangered species: recent or historic pattern? *Biol. Conserv.* 98, 61–68. doi: 10.1016/S0006-3207(00)00142-7

Matoso, D. A., da Silva, M., Cortinhas, M. C. S., Cestari, M. M., de Almeida, M. C., Vicari, M. R., et al. (2011). Two genetic stocks of *Steindachneridion melanodermatum* living in sympatry in nature and genetic variability of wild parents and F1 generation. *Genet. Mol. Res.* 10, 2606–2612. doi: 10.4238/2011.October.25.8

Miller, M. P. (2005). Alleles In Space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. *J. Hered.* 96, 722–724. doi: 10.1093/jhered/esi119

Miranda Ribeiro, A. (1918). Tres generos e dezesete especies novas de peixes Brasileiros. *Rev. Mus. Paulista* 10, 629–646.

Monaghan, M. T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D. J. G., et al. (2009). Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Syst. Biol.* 58, 298–311. doi: 10.1093/sysbio/syp027

Nei, M. (1987). *Molecular Evolutionary Genetics*. New York, NY: Columbia University Press.

Nei, M., and Tajima, F. (1981). DNA polymorphism detectable by restriction endonucleases. *Genetics* 97, 145–163.

Neville, H. M., Isaak, D. J., Dunham, J. B., Thurow, R. F., and Rieman, B. E. (2006). Fine-scale natal homing and localized movement as shaped by sex and spawning habitat in Chinook salmon: insights from spatial autocorrelation analysis of individual genotypes. *Mol. Ecol.* 15, 4589–4602. doi: 10.1111/j.1365-294X.2006.03082.x

Ochoa, L. E., Pereira, L. H. G., Costa-Silva, G. J., Roxo, F. F., Batista, J. S., Formiga, K., et al. (2015). Genetic structure and historical diversification of catfish *Brachyplatystoma platynemum* (Siluriformes: Pimelodidae) in the Amazon basin with implications for its conservation. *Ecol. Evol.* 5, 2005–2020. doi: 10.1002/ece3.1486

Pereira, L. H. G., Hanner, R., Foresti, F., and Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genet.* 14:20. doi: 10.1186/1471-2156-14-20

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., et al. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55, 595–609. doi: 10.1080/10635150600852011

Primmer, C. R., Veselov, A. J., Zubchenko, A., Poututkin, A., Bakhmet, I., and Koskinen, M. T. (2006). Isolation by distance within a river system: genetic population structuring of Atlantic salmon, Salmo salar, in tributaries of the Varzuga River in northwest Russia. *Mol. Ecol.* 15, 653–666. doi: 10.1111/j.1365-294X.2005.02844.x

Rambaut, A., Suchard, M., Xie, D., and Drummond, A. (2014). *Tracer v1.6 | BEAST*. Available at: http://beast.bio.ed.ac.uk/tracer [accessed March 27, 2017].

Ramella, M. S., Kroth, M. A., Meurer, S., De, A. P., Nuñer, O., Filho, E. Z., et al. (2006). Genetic variability in four fish species (*Pimelodus maculatus*, *Prochilodus lineatus*, *Salminus brasiliensis* and *Steindachneridion scripta*) from Uruguay River Basin. *Braz. Arch. Biol. Technol.* 49, 589–598. doi: 10.1590/S1516-89132006000500008

Ribolli, J., Hoeinghaus, D. J., Johnson, J. A., Zaniboni-Filho, E., de Freitas, P. D., and Galetti, P. M. (2017). Isolation-by-time population structure in potamodromous Dourado *Salminus brasiliensis* in southern Brazil. *Conserv. Genet.* 18, 67–76. doi: 10.1007/s10592-016-0882-x

Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X., and Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496–2497. doi: 10.1093/bioinformatics/btg359

Schork, G., Hermes-Silva, S., Beux, L. F., Zaniboni-Filho, E., Nuñer, A. P., and de, O. (2012). Diagnóstico da pesca artesanal na usina hidrelétrica de machadinho, alto Rio Uruguai - Brasil. *Bol. Inst. Pesca* 38, 97–108.

Schork, G., Hermes-Silva, S., and Zaniboni-Filho, E. (2013). Analysis of fishing activity in the Itá reservoir, upper uruguay river, in the period 2004-2009. *Braz. J. Biol.* 73, 559–571. doi: 10.1590/S1519-69842013000300014

Sivasundar, A., Bermingham, E., and Ortí, G. (2001). Population structure and biogeography of migratory freshwater fishes (Prochilodus: Characiformes) in major South American rivers. *Mol. Ecol.* 10, 407–417. doi: 10.1046/j.1365-294X.2001.01194.x

Swarça, A. C., Fenocchio, A. S., Cestari, M. M., and Dias, A. L. (2005). First chromosome data on *Steindachneridion scripta* (Pisces, Siluriformes, Pimelodidae) from Brazilian rivers: Giemsa, CBG, G-, and RE banding. *Genet. Mol. Res.* 4, 734–741.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Vrijenhoek, R., Douglas, M., and Meffe, G. (1985). Conservation genetics of endangered fish populations in Arizona. *Science* 229, 400–402. doi: 10.1126/science.229.4711.400

Ward, R. D. (2009). DNA barcode divergence among species and genera of birds and fishes. *Mol. Ecol. Resour.* 9, 1077–1085. doi: 10.1111/j.1755-0998.2009.02541.x

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., and Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 360, 1847–1857. doi: 10.1098/rstb.2005.1716

Windle, M., and Rose, G. (2005). Migration route familiarity and homing of transplanted Atlantic cod (*Gadus morhua*). *Fish. Res.* 75, 193–199. doi: 10.1016/j.fishres.2005.05.006

Wu, C. H., and Drummond, A. J. (2011). Joint inference of microsatellite mutation models, population history and genealogies using transdimensional Markov Chain Monte Carlo. *Genetics* 188, 151–164. doi: 10.1534/genetics.110.125260

Zaniboni-Filho, E., and Schulz, U. H. (2003). "Migratory fishes of the Uruguay River," in *Migratory Fishes of South America. Biology, Fisheries and Conservation Status*, ed. J. Carolsfeld (Washington, DC: World Bank), 157–194.

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 22 February 2018
doi: 10.3389/fgene.2018.00046

Genetic Characterization of the Fish *Piaractus brachypomus* by Microsatellites Derived from Transcriptome Sequencing

*Paulo H. Jorge[1], Vito A. Mastrochirico-Filho[1], Milene E. Hata[1], Natália J. Mendes[1], Raquel B. Ariede[1], Milena Vieira de Freitas[1], Manuel Vera[2], Fábio Porto-Foresti[3] and Diogo T. Hashimoto[1]**

[1] Aquaculture Center of Universidade Estadual Paulista Júlio de Mesquita Filho, São Paulo State University, Jaboticabal, Brazil, [2] Veterinary Faculty, University of Santiago de Compostela, Lugo, Spain, [3] School of Sciences, São Paulo State University, Bauru, Brazil

The pirapitinga, *Piaractus brachypomus* (Characiformes, Serrasalmidae), is a fish from the Amazon basin and is considered to be one of the main native species used in aquaculture production in South America. The objectives of this study were: (1) to perform liver transcriptome sequencing of pirapitinga through NGS and then validate a set of microsatellite markers for this species; and (2) to use polymorphic microsatellites for analysis of genetic variability in farmed stocks. The transcriptome sequencing was carried out through the Roche/454 technology, which resulted in 3,696 non-redundant contigs. Of this total, 2,568 contigs had similarity in the non-redundant (nr) protein database (Genbank) and 2,075 sequences were characterized in the categories of Gene Ontology (GO). After the validation process of 30 microsatellite loci, eight markers showed polymorphism. The analysis of these polymorphic markers in farmed stocks revealed that fish farms from North Brazil had a higher genetic diversity than fish farms from Southeast Brazil. AMOVA demonstrated that the highest proportion of variation was presented within the populations. However, when comparing different groups (1: Wild; 2: North fish farms; 3: Southeast fish farms), a considerable variation between the groups was observed. The F_{ST} values showed the occurrence of genetic structure among the broodstocks from different regions of Brazil. The transcriptome sequencing in pirapitinga provided important genetic resources for biological studies in this non-model species, and microsatellite data can be used as the framework for the genetic management of breeding stocks in Brazil, which might provide a basis for a genetic pre-breeding programme.

Keywords: aquaculture, genetic structure, NGS, Pirapitinga, Serrasalmidae

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

Reviewed by:
Pedro Manoel Galetti Jr, Federal University of São Carlos, Brazil
Maria Raquel Moura Coimbra, Federal Rural University of Pernambuco, Brazil

***Correspondence:**
Diogo T. Hashimoto
diogo@caunesp.unesp.br

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

Received: *01 September 2017*
Accepted: *31 January 2018*
Published: *22 February 2018*

Citation:
Jorge PH, Mastrochirico-Filho VA, Hata ME, Mendes NJ, Ariede RB, Freitas MV, Vera M, Porto-Foresti F and Hashimoto DT (2018) Genetic Characterization of the Fish Piaractus brachypomus by Microsatellites Derived from Transcriptome Sequencing. Front. Genet. 9:46. doi: 10.3389/fgene.2018.00046

INTRODUCTION

The pirapitinga (*Piaractus brachypomus*) is a native fish from the Amazon and Orinoco Rivers and can reach up to 20 kg of weight (Alcântara et al., 1990). This species is used for fish farming, is valued for its meat and has fast growth performance (Fresneda et al., 2004). In Brazil, pirapitinga farming represents the third largest fish production operation (about 10,000 tons) among the native

fish species (MPA, 2013a). Furthermore, this species has been widely used for the production of interspecific hybrids, particularly the tambatinga (female tambaqui *Colossoma macropomum* × male pirapitinga *P. brachypomus*), and patinga (female pacu *Piaractus mesopotamicus* × male pirapitinga *P. brachypomus;* IBGE, 2016). The aquaculture production of pirapitinga in Brazil is concentrated mainly in the Midwest and North (87%), followed by the Southeast (9%), Northeast (3%) and South (1%) (MPA, 2013b). This species also has economic importance for aquaculture in other countries in South America (Colombia, Peru, and Venezuela) and in Asia (China, Myanmar, Thailand, and Vietnam; Flores Nava, 2007; Honglang, 2007; Lin et al., 2015).

However, despite this representation of aquaculture production, few scientific studies have focused on understanding the biology of pirapitinga, especially of genetic traits. So, the generation of genetic resources for this species is fundamental to advancing studies of breeding and genetic management, as occurred in model species used in aquaculture, such as salmon, catfish, carp, and tilapia (Lien et al., 2011; Liu et al., 2011; Guyon et al., 2012; Ji et al., 2012).

Model fish species, such as zebrafish *Danio rerio*, have been described with more than 26,000 genes (Howe et al., 2013). However, few genes and their metabolic pathways have been characterized for non-model species without reference genomes, as is the case for pirapitinga. In the field of genetics and molecular biology, Next-Generation Sequencing (NGS) technologies are causing a revolution, allowing the sequencing of genome and transcriptome of any organism, quickly and at low cost (Seeb et al., 2011). RNA-seq (transcriptome sequencing) is considered one of the most used strategies of NGS technology for the transcripts analysis (Qian et al., 2014), wherein all the messenger RNA (mRNA) of a specific tissue or set of tissues are used as a source for sequencing. Moreover, RNA-seq is an effective tool for discovery of molecular markers, particularly for the prospection of gene-associated microsatellites (Teacher et al., 2012; Xu et al., 2013).

Due to the usefulness of revealing the genetic variation among individuals (Liu and Cordes, 2004), microsatellite markers have proven to be efficient for genetic characterization of wild populations and breeding stocks of farmed fish (Koljonen et al., 2002; Lehoczky et al., 2005), such as to prevent inbreeding (Ponzoni et al., 2008), to identify and preserve live gene banks (Machado-Schiaffino et al., 2007), to detect genetic structure (Do Prado et al., 2018), to direct matings during the formation of the population base of breeding programmes (Fernández et al., 2014), and to perform marker assisted selection (MAS) for economic traits (Houston et al., 2010). However, these markers are not available for pirapitinga, one of most important species for the aquaculture in South America.

For the aquaculture of pirapitinga, analysis of genetic variability in farmed stocks still needs to be performed, which will allow three hypotheses to be tested: (1) farmed stocks of pirapitinga have lower genetic diversity in relation to wild stocks; (2) farmed stocks of pirapitinga in Brazil are genetically structured; and (3) gene-linked microsatellites can be associated to economic traits of pirapitinga, such as growth and disease resistance. These analyses will support the creation of a breeding programme to increase the productivity of pirapitinga, by directed matings which lead to the formation of families, avoiding the problems of bottlenecks and inbreeding in the base population (Fernández et al., 2014), and by the identification of quantitative trait loci (QTL), which will assist the selection of superior genotypes by MAS (Houston et al., 2010).

Thus, the objective of the present study was to characterize genetic resources for the proper management of this non-model species in aquaculture, through transcriptome characterization and genetic variability analysis of stocks using microsatellite markers.

MATERIALS AND METHODS

Ethics Statement

This study was carried out in strict accordance with the animal welfare guidelines of the National Council for Control of Animal Experimentation (Brazilian Ministry for Science, Technology, and Innovation). The present study was performed under authorization N° 33435-1, issued through ICMBio (Chico Mendes Institute for the Conservation of Biodiversity, Brazilian Ministry for Environment). No animal was housed or cared for in the laboratory. Fish were euthanized by benzocaine anesthetic overdose for collection of liver tissue for transcriptome sequencing. For microsatellite validation and genetic variability analysis, fin fragments were collected from each fish under benzocaine anesthesia and all efforts were made to minimize suffering.

Samples for Transcriptome Sequencing

To perform the transcriptome sequencing, samples of liver tissue were taken from 10 individual fish from three different Brazilian fish farms and one wild population: Aquaculture Center of São Paulo State University, CAUNESP, Jaboticabal, SP ($n = 3$); Projeto Peixe fish farm, Sales Oliveira, SP ($n = 1$); Fazenda São Paulo fish farm, Brejinho de Nazaré, TO ($n = 5$); and Tocantins River, Lajeado, TO ($n = 1$). Individuals from different origins were used in order to achieve the highest genetic variability in microsatellite discovery analysis. Liver samples were selected for transcriptome studies because the liver plays a critical role in coordinating various physiological processes, including digestion, metabolism, detoxification, and endocrine system immune response (Martin et al., 2010).

Samples for Genetic Variability Analysis

Analyses of microsatellite validation were performed in 22 individual pirapitinga collected from the Tocantins River (TO) from Lajeado City, Tocantins State, Brazil. We then used the microsatellite markers to study the genetic variability in samples collected from four commercial fish farms: TO1 ($n = 25$) and TO2 ($n = 26$), from Tocantins State (North Brazil); and SP1 ($n = 36$) and SP2 ($n = 20$), from São Paulo State (Southeast Brazil). To maintain the confidentiality of these fish farms, the names of and information on the fish farms have been preserved.

Analysis of Genetic Purity in Pirapitinga Individuals

According Hashimoto et al. (2014), interspecific hybrids have been detected in broodstocks of Brazilian fish farms. The pirapitinga can be crossed with tambaqui *C. macropomum* or pacu *P. mesopotamicus*, resulting in viable and fertile hybrids (Hashimoto et al., 2012, 2014). Therefore, in the present study, special attention was given to analyze pure pirapitinga, and not interspecific hybrids. The analysis of genetic purity in all animals herein studied was performed using the mitochondrial genes, Cytochrome C Oxidase subunit I (*mt-co1*) and Cytochrome b (*mt-cyb*); and the nuclear genes, α-Tropomyosin (*tpm1*) and Recombination Activating Gene 2 (*rag2*), according to the protocols and methods of Hashimoto et al. (2011). Fish identified as interspecific hybrids were excluded from further analysis in this study.

cDNA Library Construction and Roche 454 Platform Sequencing

Samples of ∼100 mg of liver fixed in RNAlater were extracted with Rneasy Mini Kit (Qiagen). Each sample was quantified by spectrophotometry using NanoDrop ND-1000 equipment and the quality (integrity) was checked by 2100 Bioanalyzer equipment. It succeeded the preparation of an equimolar pool of total RNA samples (from 10 individuals) to mRNA enrichment with μMACS mRNA Isolation Kit (Miltenyi Biotech).

A non-normalized cDNA library was prepared using cDNA Synthesis System Kit with random primer GS Rapid Library Prep Kit and GS Rapid Library MID Adaptors Kit (Roche). The High Sensitivity DNA LabChip Kit (Agilent Technologies) with 2100 Bioanalyzer was used for quality analysis of the cDNA library. The concentration of sample (molecules/μL) was obtained by QuantiFluor™—ST fluorimeter (Promega). Titration of emPCR (emulsion PCR) was performed with the GS FLX Titanium SV em PCR Kit (Lib-L) (Roche), according to the emPCR Amplification Method Manual—Libl SV, GS FLX+ Series, to identify the optimal number of DNA molecules per bead (cpb = copies per bead). After emPCR titration, the emPCR was performed with GS FLX Titanium LV emPCR Kit (Lib-L) (Roche), according to the emPCR Amplification Method Manual—LibL LV, GS FLX+ Series. The transcriptome sequencing was conducted using the Roche/454 technology (GS FLX Titanium Sequencing Kit XL +) from HELIXXA company (Campinas, SP, Brazil), which has been used for transcriptome analysis of non-model fish species (Renaut et al., 2010).

Bioinformatic Analysis

Filtering of the initial quality of the 454 sequences in sff format was performed using the Roche Newbler programme. Sequence analysis was performed using the high-throughput sequencing module of CLC Genomics Workbench (version 7.5.1; CLC bio, Aarhus, Denmark). The raw reads were cleaned by trimming low quality sequences with quality scores of <20. Terminal nucleotides (five nucleotides at each extremity 5′ and 3′), ambiguous nucleotides, adapter sequences and reads <15 base pairs (bp) were discarded. For *de novo* assembly, contigs <200 bp were also discarded and the default local alignment settings were used to rank potential matches (mismatch cost of 2, insertion cost of 3, deletion cost of 3). The highest scoring matches that shared ≥50% of their length with ≥80% of similarity were included in the alignment. The assembled transcripts were subjected to cd-hit-est programme with an identity threshold of 90% to remove redundancy (Li and Godzik, 2006; Duan et al., 2012). In order to remove any mitochondrial and ribosomal contamination, sequences were compared against pacu mitochondrial genome and zebrafish ribosomal RNA RefSeqs (NCBI database) using CLC Genomic Workbench (version 8.0.3; CLC Bio, Aarhus, Denmark).

Functional annotation of the unique consensus sequences was performed by homology searches against the *National Center for Biotechnology Information* (NCBI) non-redundant protein database (nr) (cutoff E-value of 1E-3) using BLAST2GO software (Conesa et al., 2005) to obtain the putative gene identity. All BLASTx hits were filtered for redundancy in protein accessions. The gene ontology (GO) terms were assigned to each unique gene based on the GO terms annotated to the corresponding homologs in the NCBI database (e-value cutoff 1e-6). The transcripts were further annotated in InterPro, Enzyme code (EC), and Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathways analysis through the Bi-directional Best Hits (BBH) method.

Microsatellites were identified in the contigs using msatcommander software (Faircloth, 2008). Primers flanking the microsatellite loci were designed with Primer3plus software (Rozen and Skaletsky, 2000). The six possible reading frames of the consensus sequence of each functionally annotated contig containing microsatellite were compared against the NCBI protein database using BLASTx (e-value 1e-10) in order to find *Open Reading Frame* (ORF) regions. These approaches allowed us to locate microsatellites in coding sequences (CDS) or untranslated regions (5′UTR and 3′UTR) through graphical sequence viewer Tablet (Milne et al., 2013).

Microsatellite Genotyping and Validation

DNA was extracted from fin fragments using the Wizard Genomic DNA Purification Kit (Promega), according to the manufacturer's protocol. Microsatellite validation was performed in 30 loci, selected according to the motif and functional annotation of the contigs. Amplifications were performed by polymerase chain reaction (PCR) in a total volume of 25 μl containing 100 μM of each dNTP (dATP, dTTP, dGTP, and dCTP), 1.5 mM $MgCl_2$, 1X *Taq* DNA buffer (20 mM Tris-HCl, pH 8.4, and 50 mM KCl), 0.1 μM of each primer, 0.5 units of *Taq* Polymerase (Invitrogen) and 10-50 ng of genomic DNA. The reactions were performed in a thermocycler (ProFlex™ PCR System, Life Technologies) following initial denaturing for 10 min at 95°C; 35 cycles of 30 s at 95°C, 30 s at 55–60°C (adjusted for each primer set), 20 s at 72°C; and a final extension at 72°C for 20 min.

Microsatellites that showed polymorphism in 6% polyacrylamide gels were analyzed in a 3130xl sequencer (Life Technologies) to get better accuracy of allele determination. The sequencing strategy adopted in this study was according to protocols described by Schuelke (2000), using the CAGtag primer (5′-CAGTCGGGCGTCATCA-3′; Shirk et al., 2013) labeled with the fluorochromes HEX or FAM. The genotyping

PCR was performed with the following reagents: 100 μM of each dNTP, 1.5 mM $MgCl_2$, 1X *Taq* DNA buffer, 0.1 μM of each primer (F and R), 0.01 μM of the CAGtag primer, 0.5 units of *Taq* Polymerase (Invitrogen), and 10–50 ng of genomic DNA. The cycling programme for amplification consisted of: nine cycles at 95°C for 30 s, 55–60°C for 30 s (adjusted for each primer set), 72°C for 20 s; then, 30 cycles at 95°C for 30 s, 50°C for 30 s, and 72°C for 20 s. During the first nine cycles, the annealing temperature of 55–60°C allows incorporation of the primers (F and R) from the microsatellite loci. Then, in the following 30 cycles, the temperature of 50°C facilitates the annealing of the fluorescent dye-labeled CAGtag primer. PCR products were analyzed by capillary electrophoresis with a 3130xl genetic analyzer, using the DS-30 matrix, with the GeneScan 500 ROX dye Size Standard (Thermo). The programme GeneMapper 3.7 (Applied Biosystems) was used to determine the allele sizes.

Microsatellite Diversity and Population Analysis

For statistical analysis, we initially used GenAlex analysis 6.1 software (Peakall and Smouse, 2012) to convert the arrays into specific formats for each programme. The observed (H_o) and expected (H_e) heterozygosity, Hardy-Weinberg Equilibrium (HWE) and Analysis of Molecular Variance (AMOVA) (Excoffier et al., 1992) were calculated using the Arlequim 3.5 programme (Excoffier and Lischer, 2010). The levels of significance for the HWE test were adjusted with the Bonferroni correction (Rice, 1989). The inbreeding coefficient (F_{IS}) was performed using Genepop 4.0.11 (Rousset, 2008), based on Weir and Cockerham (1984) estimates. The fixation index (F_{ST}) was calculated using FSTAT 9.3.2 software (Goudet, 1995). Wright (1965) threshold values were adopted, F_{ST} = little genetic differentiation (0–0.05); moderate genetic differentiation (0.05–0.25); high level of genetic differentiation (> 0.25). The programme Cervus v.3.0.7 (Marshall et al., 1998) was applied to verify the presence of null alleles. Linkage disequilibrium (LD) was estimated using Arlequin v.3.5.2.2. The levels of significance were adjusted to multiple tests using the Bonferroni correction.

After LD analysis, level of admixture among population samples was inferred by estimating the optimum number of clusters (K), as suggested by Evanno et al. (2005), using the programme STRUCTURE version 2.3.4 (Pritchard et al., 2000) without prior information about population. Primarily, we determined the distribution of ΔK, an *ad hoc* statistic based on the rate of change in the log probability of data between successive K values. The range of clusters (K) was predefined from 1 to 5. The analysis was performed in 25 replicated runs using 200,000 iterations after a burn-in period of 50,000 runs. The K value most likely to explain the population structure is the modal value of this ΔK. The outputs of STRUCTURE analysis were visualized through the STRUCTURE HARVESTER programme (Earl, 2012).

Analysis for population bottlenecks was tested using BOTTLENECK (Cornuet and Luikart, 1996; Piry et al., 1999), by using the mutation–drift equilibrium assuming the two-phase model (TPM) with 70% stepwise mutation model (SMM) and 30% infinite allele model (IAM). Deviations between the observed and expected frequency distributions were tested using the Wilcoxon's signed rank test. BOTTLENECK was run for 10,000 iterations.

RESULTS

Transcriptome Sequencing

The results of liver transcriptome sequencing in pirapitinga yielded a total of 192,373 reads, which were deposited in the Short Read Archive (SRA) of NCBI under the accession number SRR6303971. The raw reads presented an average length of 395.5 bp, comprising a total of ∼76 Mbp. After the trimming process, the average length of the reads was of 362.1 bp, resulting in a total of ∼69 Mbp (192,077 *reads;* **Table 1**). As *P. brachypomus* is considered a non-model organism, and therefore without reference genome, *de novo* assembly strategy was performed for transcriptome analysis, which yielded 3,696 non-redundant contigs as a result of 174,272 overlapping reads (63,460,229 bp). The size characteristics of the contigs are presented in **Table 1**. A total of 17,805 remaining reads (6,084,530 bp) was considered as singletons, and therefore they were not used for subsequent analysis.

Non-redundant sequences were annotated by BLASTx algorithm against the NCBI databases: non-redundant protein (*nr*), protein RefSeq of *zebrafish* and *fugu*. A total of 2,568 unique protein accessions (69.4% of transcripts) had significant similarity in the *nr* database. In relation to the protein RefSeq of *zebrafish* and *fugu*, we found similar numbers of annotated genes, which were of 2,498 (67.6%) and 2,419 (65.4%), respectively. No sequence showed homology with known pirapitinga protein sequences deposited in NCBI database, because the available sequences database is still limited to mostly mitochondrial sequences.

Of the 2,568 contigs with correspondence in the *nr* database, 2,075 (80.8%) were annotated in the categories of Gene Ontology (GO). A total of 1,831 assignments to Biological Process (88.2%) were found, followed by 1,757 to Molecular Function (84.6%) and 1,378 to Cellular Component (66.4%). In relation to the GO subcategories, the most abundant terms were related to: metabolic process, cellular process, and single-organism process of the Biological Process category; binding, catalytic activity, and

TABLE 1 | Data of *de novo* assembly from liver transcriptome of pirapitinga *Piaractus brachypomus*.

Matched reads for assembly	174,272
Total nucleotides of matched reads	63,460,229
Number of contigs	3,797
Total of contig nucleotides	2,999,680
Minimum contig length (bp)	202
Maximum contig length (bp)	7,812
Average contig length (bp)	790
N75 (bp)	593
N50 (bp)	861
N25 (bp)	1,383

transporter activity of the Molecular Function; cell, organelle, and membrane of the Cellular Component (**Figure 1**). In the present study, genes assigned to the immune system, growth and reproduction were found, and therefore these data will serve as support for future studies on the aquaculture of pirapitinga.

The transcripts characterization in the KEGG database demonstrated that 1,122 sequences were identified in 106 metabolic pathways. Genes involved in the biosynthesis of antibiotics, purine metabolism and glycolysis/gluconeogenesis were able to be highlighted (**Figure 2**).

Microsatellite Diversity and Population Analysis

The search for short sequence repeats (SSR) in the 3,696 contigs resulted in the discovery of 130 microsatellite markers distributed in 95 contigs. In total, 75 pairs of primers were designed adjacent to the microsatellite loci, including the following sequence repeats: 56 di, 13 tri, 4 tetra, and 2 pentanucleotide. Among the dinucleotide motifs, the main repeats were the types AC (48.28%), AG (39.65%), AT (10.35%), and CG (1.72%). In relation to the trinucleotide motifs, we identified seven types (AGC, AGG, ATC, AAT, ACG, CCG, and AAG). The tetranucleotide (ATCT, AAAG, AATG, and AAAT) and pentanucleotide (ACTAT and ATAGT) sequences were described with the presence of four and two types of motifs. In relation to the gene position, 26.76% of the microsatellite markers were found in the 3′UTR (untranslated region), 19.71% in the 5′UTR, and 29.58% in the cds (coding sequence).

In the process of microsatellite validation, 30 markers were evaluated in 22 samples of pirapitinga collected from the wild. Of these markers, eight microsatellite loci showed polymorphism (GenBank accession numbers MG595996—MG596003), revealed by the presence of different fragment sizes (**Table 2**). The number of alleles was low, which ranged from 2 (loci C25, C64, C410, and C1832) to 5 (C1376) and mean of 2.750 ± 0.366. The expected (H_e) and observed (H_o) heterozygosity in the wild population had an average of 0.466 ± 0.061 and 0.355 ± 0.076, respectively. Most of the loci showed positive values for F_{IS}, except the locus C64. Three microsatellite loci (C13, C25, and C1716) showed significant deviation from the Hardy–Weinberg Equilibrium (HWE) after Bonferroni correction (adjusted $p = 0.00625$).

The results of genetic variability in farmed stocks revealed that North fish farms TO1 and TO2 had higher diversity than the wild population, demonstrated by number of alleles (mean of 4.500 ± 0.423 and 3.375 ± 0.596, respectively) and average values of H_e (0.589 ± 0.033 and 0.488 ± 0.044, respectively) and H_o (0.520 ± 0.060 and 0.447 ± 0.040, respectively; **Table 3**). The Southeast fish farms SP1 and SP2 showed the lowest genetic variability when compared to other populations, with lower allele number (mean of 2.250 ± 0.313 and 3.125 ± 0.581, respectively), and average of H_e(0.226 ± 0.077 and 0.278 ± 0.085, respectively; $p < 0.05$) and H_o (0.259 ± 0.103 and 0.251 ± 0.079, respectively; **Table 3**). Most of the microsatellite loci were characterized with positive values of F_{IS}, except for SP1 and SP2. The mean value of F_{IS} and null alleles was positive in most populations, with the exception of SP1 (−0.071 ± 0.076 and −0.011 ± 0.080). The majority of the markers were in concordance to HWE, after Bonferroni correction, with the exception of C25 (TO1, SP1, and SP2), C64 (SP1 and SP2) and C1376 (SP1) (**Table 3**). Linkage disequilibrium was found between the microsatellites C410 and C1005 ($p < 0.00625$). Although molecular markers on linkage disequilibrium were not applied in genetic variability studies, this information can be useful in future analysis of genetic mapping.

In bottleneck analyses, evidence for recent reductions in population size (bottleneck) using TPM was not found, except for the wild population of Tocantins River ($p = 0.027$).

In the evaluation of the level of admixture among stocks by STRUCTURE, the model-based clustering analyses detected K = 3, allowing the identification of 3 main clusters between the populations: Group 1 (SP1 and SP2), Group 2 (TO1 and TO2), and Group 3 (wild) (**Figure 3**).

The global F_{ST} was 0.379, which showed high genetic differentiation among the populations ($p < 0.05$). Pairwise F_{ST} detected a higher genetic differentiation between the wild population and all farmed stocks, particularly when compared to SP1 ($F_{ST} = 0.538$, $p < 0.05$) and SP2 ($F_{ST} = 0.537$, $p < 0.05$). Additionally, high genetic structure was found between the populations from North and Southeast Brazil, as observed between TO2 with SP1 ($F_{ST} = 0.463$, $p < 0.05$) and TO1 with SP2 ($F_{ST} = 0.380$, $p < 0.05$; **Table 4**). Moreover, values of pairwise F_{ST} after stock clustering detected a higher genetic differentiation when comparing Group 1/Group 2 ($F_{ST} = 0.379$, $p < 0.05$), Group 1/Group 3 ($F_{ST} = 0.549$, $p < 0.05$), and Group 2/Group 3 ($F_{ST} = 0.144$, $p < 0.05$).

The results of AMOVA showed that the majority of genetic variation (29.11%, $F_{CT} = 0.291$, $p < 0.001$) occurred between groups (according to STRUCTURE clustering), while the variation among individuals within populations was only 8.21% ($F_{IS} = 0.126$, $p < 0.001$) and among populations within groups presented 6.06% of genetic variation ($F_{SC} = 0.085$, $p < 0.001$).

DISCUSSION

Transcriptome Sequencing

Currently, genetic resources for pirapitinga *P. brachypomus* are limited only to sequences of the mitochondrial genome (Chen et al., 2016). Thus, one of the main results of this study was the data generated through transcriptome sequencing, because little knowledge was available about the genes of this species. The efficiency of the Roche/454 sequencing system in the functional genomics analysis of pirapitinga can be observed because of the 3,696 transcripts that were generated in this study. According to Seeb et al. (2011), genome reduction strategies for NGS sequencing (e.g., transcriptome sequencing) are more viable when the objective is to prospect molecular markers and genetic information for use in aquaculture, in a low cost and fast way. Roche/454 sequencing technology is one of the main methods used in NGS transcriptome of non-model fish (Renaut et al., 2010;

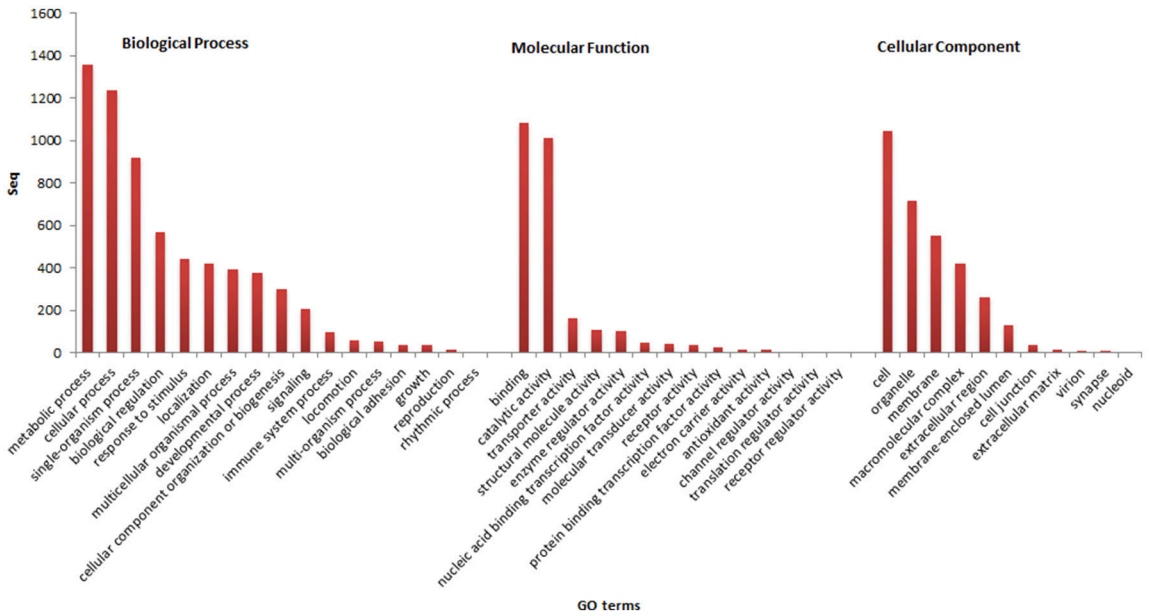


FIGURE 1 | Results of functional annotation and the assignment of genes in the GO categories and subcategories.

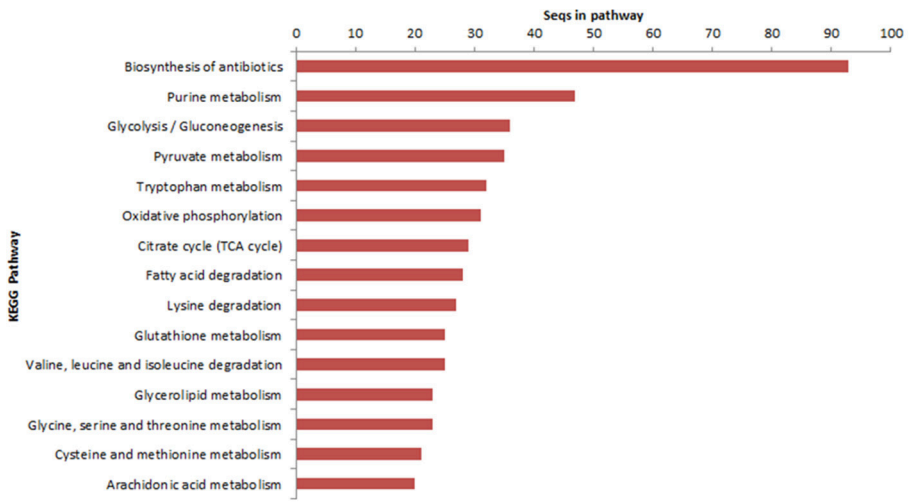


FIGURE 2 | Transcripts characterized in metabolic pathways database of KEGG enzymes *(Kyoto Encyclopedia of Genes and Genomes)*.

Shin et al., 2012; Calduch-Giner et al., 2013; Mutz et al., 2013).

The results of functional annotation showed that the sequences of pirapitinga had a high proportion of annotated genes when compared to the database of *zebrafish* and *fugu* proteins. The gene annotation allowed identification of genomic regions responsible for ontogenetic development processes, biological regulation, the immune system, and regions involved in processes of growth and reproduction. Consequently, the present data can be used as the basis of further biological studies of other areas of aquaculture or for future breeding programmes. In addition, through transcriptome sequencing, the discovery of gene-associated microsatellites can be considered to be the main result which can be applied to pirapitinga aquaculture, as already demonstrated in previous studies of fish (Renaut et al., 2010; Helyar et al., 2012; Shin et al., 2012). The use of gene-associated markers becomes even more important in the construction of genetic maps (Shin et al., 2012) because, by comparative genomics using fish genome references already sequenced, it is possible to presume the location of each studied locus.

TABLE 2 | Characterization of the genetic diversity of eight polymorphic microsatellites in the wild population of pirapitinga (*Piaractus brachypomus*).

Locus	Sequence description	Gene position	Motifs	Primers 5′–3′	T_A°C	Size range	Na	H_O	H_e	P(HWE)	F_{IS}	F(Null)
C13	dihydroxyvitamin d24-like	3′UTR	$(AGC)_6$	F: TCTCTTCAAGCCTCCTCTGC R: ATGCTGCAGCTCCTCCTGT	60°C	143–149	3	0.545	0.669	0.000	0.188	0.075
C25	cytosolic 5-nucleotidase 3a-like isoform x 1	3′UTR	$(AT)_{11}$	F: CTTTGTCTGCTTTGGGTCGT R: CTTAGAAGAATGTGCAAATTGAAA	60°C	117–120	2	0.000	0.169	0.001	1.000	0.887
C64	sodium-coupled neutral amino acid transporter	5′UTR	$(AAAG)_7$	F: CAAAGCAAACTCAAAAAGGAAAA R: TGGGAACGTTTAGCATCTCA	55°C	143–151	2	0.545	0.474	0.650	−0.156	−0.082
C410	apolipoprotein e	3′UTR	$(AG)_8$	F: CGCACAGGTCTAAAGGCACT R: CTCCCACACAGTGAAAAGCA	60°C	125−127	2	0.273	0.359	0.271	0.245	0.125
C1005	apolipoprotein e	3′UTR	$(AG)_7$	F: AGTTGTTGCACCAAATGCAG R: CTTGTTCCCTCCCACACAGT	60°C	137−141	3	0.318	0.369	0.225	0.140	0.090
C1376	frutose-biphosphatase 1-like	5′UTR	$(AC)_{10}$	F: GTGTTACATGGCAGGCGTTT R: CAAGTGAGACCAAATCCAAGG	60°C	157–175	5	0.500	0.608	0.185	0.180	0.073
C1716	vacuolar atp synthase 16 kda proteolipid subunit	3′UTR	$(GT)_7$	F: AACCGAAGAGAGGGGAGTGT R: GCATTTACAAGGGGACGCAC	60°C	155–171	3	0.545	0.659	0.000	0.175	0.059
C1832	–	–	$(AC)_6$	F: GGTGCTATGTCGTAGAGGCC R: AGGAAGGCATGACCAGTGTG	60°C	159–169	2	0.111	0.529	0.036	0.800	ND

The wild population analyzed corresponds to 22 individuals collected on the Tocantins River. T_A, annealing temperature (°C); Na, number of alleles per locus; H_O, observed heterozygosity; H_e, expected heterozygosity; P (HWE), p-value from Hardy-Weinberg equilibrium and F_{IS}, inbreeding coefficient. F(Null), Null alleles; ND, not performed.

Moreover, some examples have demonstrated that gene-linked microsatellite markers can be correlated with interesting productive traits, especially for growth performance. In the fish *Sparus aurata*, a dinucleotide microsatellite in the 5′ UTR of the growth hormone gene (GH) is linked with faster growth rate, especially the alleles 250 and 254, which can be used for breeding management and genetic selection for this trait (Almuly et al., 2005). In other fish species, such as *Oreochromis niloticus* and *Lates calcarifer* (Yue et al., 2001; Yue and Orban, 2002), microsatellites have also been reported for genes of interest (prolactin, GH and igf2) and, therefore, they can be used in marker-assisted selection (MAS) programmes. In the present study, eight polymorphic microsatellite loci were validated, some of them located in gene regions that may be useful for productive characteristics in aquaculture. In this case, a microsatellite *locus* was found in the gene *Tetraspanin−3 isoform x1* (C1832), which plays a role in viral infection pathology (Martin et al., 2005; Shoshana and Shoham, 2005). There is another microsatellite in the gene *Cytosolic 5 – nucleotidase 3 a-like* (NTC5C3) (C25), which contributes in the production of red blood cells and its mutation can cause hemolytic anemia and influence on the immune system (Aksoy et al., 2009). Thus, the microsatellites described in this study will be also important in future analysis of (QTL) linked to traits of disease resistance, which has received special attention in aquaculture species, such as turbot (*Scophthalmus maximus*), rainbow trout (*Oncorhynchus mykiss*), salmon (*Salmo salar*), Nile tilapia (*O. niloticus*), and cod (*Gadus morhua*), investigating the resistance to pathogens (Pardo et al., 2008; Ødegård et al., 2010, 2011; Yáñez et al., 2014; Evenhuis et al., 2015). Furthermore, one microsatellite locus was also detected in the gene *apolipoprotein e* (C410), which is associated with the central nervous system and the senescence process (Wang et al., 2014). These markers can provide useful information for studies of the biology of the pirapitinga, besides serving as a framework for other native species.

Population Analysis

The validation of eight microsatellites showed a low level of genetic diversity in these loci, both in wild and farmed stocks. In the wild, the observed heterozygosity (H_o) ranged from 0.000 to 0.545 and an average of 2.750 alleles per locus. These values confirm the low genetic variability when compared with related species, such as pacu *P. mesopotamicus* (H_o range from 0.068 to 0.911 and average of 8.5 alleles per locus), and tambaqui *C. macropomum* (H_o range from 0.430 to 0.880 and average of 12.8 alleles per locus; Calcagnotto and DeSalle, 2009; Fazzi-Gomes et al., 2017). In contrast to neutral markers (microsatellites in noncoding regions), gene-associated microsatellites might be more susceptible to selection pressure and, therefore, they have low values of gene diversity.

Analysis of the genetic diversity in pirapitinga farmed stocks showed significant differences between fish farms in different regions of Brazil, two from the Southeast (São Paulo State: SP1 and SP2) and two from the North (Tocantins State: TO1 and TO2). In general, farmed stocks were expected to have low genetic variability as a result of genetic decline, genetic

TABLE 3 | Values of genetic diversity of eight microsatellite loci of *Piaractus brachypomus*.

Stocks		Loci							
		C13	**C25**	**C64**	**C410**	**C1005**	**C1376**	**C1716**	**C1832**
Wild	N	22	22	22	22	22	22	22	9
	Na	3	2	2	2	3	5	3	2
	H_o	0.545	0.000	0.545	0.273	0.318	0.500	0.545	0.111
	H_e	0.669	0.169	0.474	0.359	0.369	0.608	0.659	0.529
	P(HWE)	0.000	0.001	0.650	0.271	0.225	0.185	0.000	0.036
	F_{IS}	0.188	1.000	−0.156	0.245	0.140	0.180	0.175	0.800
	F(Null)	0.075	0.887	−0.082	0.125	0.090	0.073	0.059	ND
TO1	N	25	25	25	25	25	25	25	25
	Na	4	5	3	3	4	6	6	5
	H_o	0.600	0.160	0.560	0.520	0.560	0.440	0.760	0.560
	H_e	0.536	0.704	0.495	0.537	0.562	0.609	0.782	0.580
	P(HWE)	0.650	0.001	0.365	0.267	0.421	0.010	0.545	0.388
	F_{IS}	−0.121	0.776	−0.133	0.034	0.004	0.282	0.028	0.035
	F(Null)	−0.067	0.620	−0.065	−0.015	−0.034	0.164	0.011	−0.001
TO2	N	26	21	26	26	26	26	26	26
	Na	3	2	2	2	2	6	5	5
	H_o	0.500	0.380	0.346	0.384	0.384	0.538	0.461	0.653
	H_e	0.528	0.315	0.382	0.506	0.506	0.632	0.515	0.638
	P(HWE)	0.840	1.000	0.626	0.256	0.256	0.020	0.030	0.670
	F_{IS}	0.055	−0.212	0.096	0.244	0.244	0.151	0.107	−0.253
	F(Null)	0.018	−0.103	0.040	0.127	0.127	0.055	0.067	−0.027
SP1	N	14	14	14	14	14	14	14	14
	Na	2	1	1	3	2	3	3	3
	H_o	0.143	–	–	0.287	0.500	0.857	0.143	0.143
	H_e	0.137	–	–	0.264	0.494	0.634	0.140	0.203
	P(HWE)	1.000	–	–	1.000	1.000	0.000	1.000	0.109
	F_{IS}	−0.040	–	–	−0.083	−0.011	−0.368	−0.019	0.306
	F(Null)	−0.028	–	–	−0.069	−0.023	−0.215	−0.027	0.272
SP2	N	19	19	19	19	19	16	18	19
	Na	2	2	1	3	3	5	3	6
	H_o	0.368	0.000	–	0.157	0.157	0.625	0.222	0.473
	H_e	0.308	0.193	–	0.152	0.152	0.790	0.207	0.486
	P(HWE)	1.000	0.002	–	1.000	1.000	0.112	1.000	0.387
	F_{IS}	−0.200	1.000	–	−0.038	−0.385	0.214	−0.070	0.027
	F(Null)	−0.099	0.916	–	−0.032	−0.032	0.093	−0.051	0.024

Wild, population from the Tocantins River; TO1 and TO2, fish farms from Tocantins; SP1 and SP2, fish farms from São Paulo, Na, number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity; P (HWE), P-value from Hardy-Weinberg equilibrium; F_{IS}, inbreeding coefficient; F(Null), Null alleles; ND, not performed.

drift, selection and inbreeding (Theodorou and Couvet, 2015). However, the results of this study showed higher genetic variability in breeding stocks from North fish farms in relation to the wild stocks ($p < 0.05$; higher values of allelic frequency and heterozygosity), which was also observed in studies with other related species (Barroso et al., 2005; Panarari-Antunes et al., 2011). The basis of this result could be considered from three different perspectives: (1) North fish farms had originated from different wild stocks resulting in high level of genetic variability; (2) problems of sample size bias, such as few microsatellite loci and individuals analyzed; (3) evidence for recent genetic bottlenecks in the wild population. Some studies of fish have reported bottlenecks in natural populations, particularly due to habitat loss and fragmentation by human disturbance (Brauer et al., 2016). In the case of pirapitinga, the fragmentation of the Tocantins River by hydroelectric dams in the 80′s and 90′s (e.g., Tucuruí and Luiz Eduardo Magalhães dams, where wild fish were collected for this study) could be responsible for

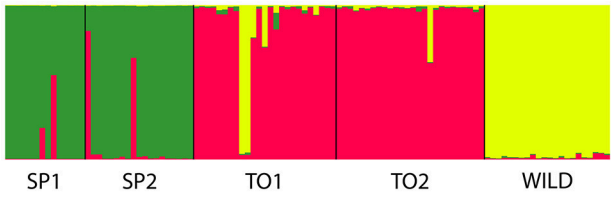


FIGURE 3 | Evaluation of the level of admixture among stocks by STRUCTURE, showing three main clusters between the populations: Group 1 (SP1 and SP2) in green, Group 2 (TO1 and TO2) in red, and Group 3 (wild) in yellow.

TABLE 4 | Analysis of pairwise F_{ST} based on eight microsatellite loci between populations of *Piaractus brachypomus*.

	SP1	SP2	TO1	TO2	wild
SP1	–				
SP2	0.18438	–			
TO1	0.39201	0.38080	–		
TO2	0.46373	0.42999	0.08345	–	
wild	0.53856	0.53782	0.16086	0.17915	–

Wild, population from the Tocantins River; TO1 and TO2, fish farms from Tocantins; SP1 and SP2, fish farms from São Paulo. All results of F_{ST} were significant statistically $p < 0.05$.

a population reduction and subsequent genetic variation loss detected by our microsatellite analysis. There are considerable numbers of hydropower dams in the basin, which can affect the reproduction, migratory routes, and egg and larvae drift of fish (Agostinho et al., 2008). Alteration of the migratory flow consequently leads to a decrease in or interruption of the gene flow, reducing the population size, which makes the fish more susceptible to the effects of genetic drift (Hatanaka and Galetti, 2003), which results in genetic structure for some fish species (Calcagnotto and DeSalle, 2009; Do Prado et al., 2018).

STRUCTURE and pairwise F_{ST} analyses suggested a high genetic structure between the stocks herein analyzed, particularly as result of the fixation of specific alleles in some loci, which resulted in three clusters (**Figure 3**). There are three hypothetical explanations for these genetic patterns: (1) differentiation of wild population in relation to farmed stocks, which could be due to the selection of the fittest individuals for farming systems or low number of founders for the establishment of the farmed broodstocks; (2) lower genetic structure in North/wild than Southeast/wild, which suggests that North fish farms had frequent broodstock renovation from the wild; (3) fish farms were genetically clustered due to the geographic distribution, i.e., the degree of genetic similarity is higher when one fish farm is closer to the other, indicating interchange of individuals between nearby fish farms, common origin of the farmed broodstocks, or fixation/selection of specific alleles for different climatic conditions that are found in Brazil (North and South). However, these genetic patterns should be also evaluated using neutral markers (microsatellites in noncoding regions) and through techniques of higher genome coverage (SNP, single-nucleotide polymorphism).

Through AMOVA analysis, the main genetic variation was found to be present within populations (64.8%). This genetic pattern has also been reported in studies carried out with pacu (Calcagnotto and DeSalle, 2009; Iervolino et al., 2010) and tambaqui (Aguiar et al., 2013). Moreover, highly significant genetic variation was associated with differences between groups (Wild, SP, and TO), which represented 29.11% of genetic variation, in contrast to low differences among populations within groups (6.06%).

In general, our study of genetic characterization in piratininga farmed stocks provides important insights which can lead to better management of this species in aquaculture. Our results are fundamental to beginning a breeding programme, since the genetic structure should be taken into consideration when composing an initial base population, where matings between farmed individuals from North and Southeast Brazil are shown to result in higher genetic variability in the families. Moreover, the data suggested levels of genetic diversity which were higher in farmed stocks than in wild fish, discarding the occurrence of inbreeding. In general, lack of knowledge on genetic variability of stocks can result in inbreeding and fixation of deleterious genes, reduced growth rates, disease resistance problems and reduced ability to adapt to new environments (Arkush et al., 2002; Gallardo et al., 2004; Neira et al., 2006; Hillen et al., 2017). Therefore, besides the identification of QTL to assist in the selection of superior genotypes by MAS, studies of microsatellites are important for genetic monitoring, supporting pirapitinga aquaculture and increasing its productivity.

FINAL CONSIDERATIONS

The prospection of genetic data for pirapitinga is one of the priority issues for aquaculture, since this species is of high economic importance in national and global fish farming. The identification of gene-associated microsatellites by NGS is fundamental to understanding the genetic structure of wild and farmed populations, providing support for further management programmes and genetic pre-breeding programmes. Moreover, the microsatellites described herein are interesting targets used to find QTL markers, specifically related to the immune system of pirapitinga.

AUTHOR CONTRIBUTIONS

PJ: Acquisition, analysis and interpretation of data, draft of the work, final approval of the version; VM-F, RA, and MdF: Draft of the work, development of intellectual content, final approval of the version; MH: Analysis and interpretation of data, draft of the work, final approval of the version; NM: Analysis and interpretation of data, draft of the work, final approval of the version; MV: Analysis and interpretation of data, draft

of the work, final approval of the version; FP-F: Acquisition, Interpretation of data, development of intellectual content, final approval of the version; DH: Draft of the work, development of intellectual content, writing of the manuscript, final approval of the version.

ACKNOWLEDGMENTS

This work was supported by grants from FAPESP (2014/03772-7 and 2014/05732-2), CNPq (446779/2014-8 and 305916/2015-7), and PROPE/UNESP.

REFERENCES

Agostinho, A. A., Pelicice, F. M., and Gomes, L. C. (2008). Dams and the fish fauna of the Neotropical region: impacts and management related to diversity and fisheries. *Braz. J. Biol.* 68, 1119–1132. doi: 10.1590/S1519-69842008000500019

Aguiar, J., Schneider, H., Gomes, F., Carneiro, J., Santos, S., Rodrigues, L. R., et al. (2013). Genetic variation in native and farmed populations of Tambaqui (*Colossoma macropomum*) in the Brazilian Amazon: regional discrepancies in farming systems. *An. Acad. Bras. Ciênc.* 85, 1439–1447. doi: 10.1590/0001-376520130007

Aksoy, P., Zhu, M. J., Kalari, K. R., Moon, I., Pelleymounter, L. L., Eckloff, B. W., et al. (2009). Cytosolic 5′-nucleotidase III (NT5C3): gene sequence variation e functional genomics. *Pharmacogenet. Genomics* 19, 567–576. doi: 10.1097/FPC.0b013e32832c14b8

Alcântara, P. F., Oliveira, A. A., and Nobre, M. I. S. N. (1990). Considerações sobre a amostragem da pirapitinga, *Colossoma brachypomum*, Cuvier, no estado do Ceará (Brasil). *Ciênc. Agron.* 21, 43–49.

Almuly, R., Poleg-Danin, Y., Gorshkov, S., Gorshkova, G., Rapoport, B., Soller, M., et al. (2005). Characterization of the 5′ flanking region of the growth hormone gene of the marine teleost, gilthead sea bream *Sparus aurata*: analysis of a polymorphic microsatellite in the proximal promoter. *Fish. Sci.* 71, 479–490. doi: 10.1111/j.1444-2906.2005.00991.x

Arkush, K. D., Giese, A. R., Mendonca, H. L., McBride, A. M., Marty, G. D., and Hedrick, P. W. (2002). Resistance to three pathogens in the endangered winter-run chinook salmon (*Oncorhynchus tshawytscha*): effects of inbreeding and major histocompatibility complex genotypes. *Can. J. Fish. Aquat. Sci.* 59, 966–975. doi: 10.1139/f02-066

Barroso, R. M., Hilsdorf, A. W. S., Moreira, H. L. M., Cabello, P. H., and Traub-Cseko, Y. M. (2005). Genetic diversity of wild and cultured populations of *Brycon opalinus* (Cuvier, 1819) (*Characiforme*, *Characidae*, *Bryconiae*) using microssatellites. *Aquaculture* 247, 51–65. doi: 10.1016/j.aquaculture.2005.02.004

Brauer, C. J., Hammer, M. P., and Beheregaray, L. B. (2016). Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin. *Mol. Ecol.* 25, 5093–5113. doi: 10.1111/mec.13830

Calcagnotto, D., and DeSalle, R. (2009). Population genetic structuring in pacu (*Piaractus mesopotamicus*) across the Paraná-Paraguay basin: evidence from microsatellites. *Neotrop. Ichthyol.* 7, 607–616. doi: 10.1590/S1679-62252009000400008

Calduch-Giner, J. A., Bermejo-Nogales, A., Benedito-Palos, L., Estensoro, I., Ballester-Lozano, G., Sitjà-Bobadilla, A., et al. (2013). Deep sequencing for de novo construction of a marine fish (*Sparus aurata*) transcriptome database with a large coverage of protein-coding transcripts. *BMC Genomics*. 14:178. doi: 10.1186/1471-2164-14-178

Chen, H. P., Li, S. S., Xie, Z. Z., Zhang, Y., Zhu, C. H., Deng, S. P., et al. (2016). The complete mitochondrial genome of the *Piaractus brachypomus* (Characiformes: Characidae). *Mitochondrial DNA*. 27, 1289–1290. doi: 10.3109/19401736.2014.945560

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

Cornuet, J. M., and Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144, 2001–2014.

Do Prado, F. D., Fernandez-Cebrián, R., Foresti, F., Oliveira, C., Martínez, P., and Porto-Foresti, F. (2018). Genetic structure and evidence of anthropogenic effects on wild populations of two Neotropical catfishes: baselines for conservation. *J. Fish Biol.* 92, 55–72. doi: 10.1111/jfb.13486

Duan, J., Xia, C., Zhao, G., Jia, J., and Kong, X. (2012). Optimizing de novo common wheat transcriptome assembly using short-read RNA-Seq data. *BMC Genomics* 13:392. doi: 10.1186/1471-2164-13-392

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Evenhuis, J. P., Leeds, T. D., Marancik, D. P., LaPatra, S. E., and Wiens, G. D. (2015). Rainbow trout (*Oncorhynchus mykiss*) resistance to columnaris disease is heritable and favorably correlated with bacterial cold water disease resistance. *J. Anim. Sci.* 93, 1546–1554. doi: 10.2527/jas.2014-8566

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Excoffier, L., Smouse, P. E., and Quatrro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics.* 131, 479 491.

Faircloth, B. C. (2008). Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol. Ecol. Res.* 8, 92–94. doi: 10.1111/j.1471-8286.2007.01884.x

Fazzi-Gomes, P., Guerreiro, S., Palheta, G. D. A., Melo, N. F. A. C., Santos, S., and Hamoy, I. (2017). High genetic diversity and connectivity in *Colossoma macropomum* in the Amazon basin revealed by microsatellite markers. *Genet. Mol. Biol.* 40, 142–146. doi: 10.1590/1678-4685-gmb-2015-0222

Fernández, J., Toro, M. Á., Sonesson, A. K., and Villanueva, B. (2014). Optimizing the creation of base populations for aquaculture breeding programs using phenotypic and genomic data and its consequences on genetic progress. *Front. Genet.* 5:414. doi: 10.3389/fgene.2014.00414

Flores Nava, A. (2007). "Aquaculture seed resources in Latin America: a regional synthesis," in *Assessment of Freshwater Fish Seed Resources for Sustainable Aquaculture*, ed M.G. Bondad-Reantaso (Rome: FAO Fisheries Technical Paper, No. 501, FAO), 91–102.

Fresneda, A., Lenis, G., Agudelo, E., and Olivera Ángel, M. (2004). Espermiación inducida y crioconservación de semen de cachama blanca (*Piaractus brachypomus*). *Rev. Colomb. Cienc. Pec.* 17, 46–52.

Gallardo, J. A., Garcia, X., Lhorente, J. P., and Neira, R. (2004). Inbreeding and inbreeding depression of female reproductive traits in two populations of Coho salmon selected using BLUP predictors of breeding values. *Aquaculture* 234, 111–122. doi: 10.1016/j.aquaculture.2004.01.009

Goudet, J. (1995). FSTAT (Version 1.2): a computer program to calculate f-statistics. *J. Hered.* 86, 485–486. doi: 10.1093/oxfordjournals.jhered.a111627

Guyon, R., Rakotomanga, M., Azzouzi, N., Coutanceau, J. P., Bonillo, C., D'Cotta, H., et al. (2012). A high-resolution map of the Nile tilapia genome: a resource for studying cichlids and other percomorphs. *BMC Genomics* 13:222. doi: 10.1186/1471-2164-13-222

Hashimoto, D. T., Mendonça, F. F., Senhorini, J. A., Oliveira, C., Foresti, F., and Porto-Foresti, F. (2011). Molecular diagnostic methods for identifying Serrasalmid fish (*Pacu, Pirapitinga*, and *Tambaqui*) and their hybrids in the Brazilian aquaculture industry. *Aquaculture* 321, 49–53. doi: 10.1016/j.aquaculture.2011.08.018

Hashimoto, D. T., Senhorini, J. A., Foresti, F., Martínez, P., and Porto-Foresti, F. (2014). Genetic identification of F1 and post-F1 Serrasalmid juvenile hybrids in Brazilian aquaculture. *PLoS ONE* 9:e89902. doi: 10.1371/journal.pone.0089902

Hashimoto, D. T., Senhorini, J. A., Foresti, F., and Porto-Foresti, F. (2012). Interspecific fish hybrids in Brazil: management of genetic resources for sustainable use. *Rev. Aquacult.* 4, 108–118. doi: 10.1111/j.1753-5131.2012.01067.x

Hatanaka, T., and Galetti, P. M. Jr. (2003). RAPD markers indicate the occurrence of structured populations in a migratory freshwater fish species. *Genet. Mol. Biol.* 26, 19–25. doi: 10.1590/S1415-47572003000100004

Helyar, S. J., Limborg, M. T., Bekkevold, D., Babbucci, M., van Houdt, J., Maes, G. E., et al. (2012). SNP discovery using next generation transcriptomic sequencing in atlantic herring (*Clupea harengus*). *PLoS ONE* 7:e42089. doi: 10.1371/journal.pone.0042089

Hillen, J. E. J., Coscia, I., Vandeputte, M., Herten, K., Hellemans, B., Maroso, F., et al. (2017). Estimates of genetic variability and inbreeding in experimentally selected populations of European sea bass. *Aquaculture* 479, 742–749. doi: 10.1016/j.aquaculture.2017.07.012

Honglang, H. (2007). "Freshwater fish seed resources in China," in *Assessment of Freshwater Fish Seed Resources for Sustainable Aquaculture*, ed. M.G. Bondad-Reantaso (FAO Fisheries Technical Paper No. 501. Rome: FAO), 185–199.

Houston, R. D., Haley, C. S., Hamilton, A., Guy, D. R., Mota-Velasco, J. C., Gheyas, A. A., et al. (2010). The susceptibility of Atlantic salmon fry to freshwater infectious pancreatic necrosis is largely explained by a major QTL. *Heredity* 105, 318–327. doi: 10.1038/hdy.2009.171

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., et al. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496, 498–503. doi: 10.1038/nature12111

IBGE (2016). Produção da Pecuária Municipal 2016. *Rio de Janeiro* 44, 1–51.

Iervolino, F., Resende, E. K., and Hilsdorf, A. W. S. (2010). The lack of genetic differentiation of pacu (*Piaractus mesopotamicus*) populations in the Upper-Paraguay Basin revealed by the mitochondrial DNA D-loop region: implications for fishery management. *Fish. Res.* 101, 27–31. doi: 10.1016/j.fishres.2009.09.003

Ji, P., Liu, G., Xu, J., Wang, X., Li, J., Zhao, Z., et al. (2012). Characterization of common carp transcriptome: sequencing, de novo assembly, annotation and comparative genomics. *PLoS ONE* 7:e35152. doi: 10.1371/journal.pone.0035152

Koljonen, M. L., Tahtinen, J., Saisa, M., and Koskiniemi, J. (2002). Maintenance of genetic diversity of Atlantic salmon (*Salmo salar*) by captive breeding programmes and the geographic distribution of microsatellite variation. *Aquaculture* 212, 69–92. doi: 10.1016/S0044-8486(01)00808-0

Lehoczky, I., Magyary, I., Hancz, C., and Weiss, S. (2005). Preliminary studies on the genetic variability of six Hungarian common carp strains using microsatellite DNA markers. *Hydrobiologia* 533, 223–228. doi: 10.1007/s10750-004-2490-x

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and compare large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158

Lien, S., Gidskehaug, L., Moen, T., Hayes, B. J., Berg, P. R., Davidson, W. S., et al. (2011). A dense SNP-based linkage map for Atlantic salmon (*Salmo salar*) reveals extended chromosome homeologies and striking differences in sex-specific recombination patterns. *BMC Genomics*. 12:615. doi: 10.1186/1471-2164-12-615

Lin, Y., Gao, Z., and Zhan, A. (2015). Introduction and use of non-native species for aquaculture in China status, risk and management solutions. *Rev. Aquacult.* 7, 28–58. doi: 10.1111/raq.12052

Liu, S., Zhou, Z., Lu, J., Sun, F., Wang, S., Liu, H., et al. (2011). Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array. *BMC Genomics*. 12:53. doi: 10.1186/1471-2164-12-53

Liu, Z. J., and Cordes, J. F. (2004). DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* 238, 1–37. doi: 10.1016/j.aquaculture.2004.05.027

Machado-Schiaffino, G., Dopico, E., and Garcia-Vazquez, E. (2007). Genetic variation losses in Atlantic salmon stocks created for supportive breeding. *Aquaculture* 264, 59–65. doi: 10.1016/j.aquaculture.2006.12.026

Marshall, T., Slate, J., Kruuk, L., and Pemberton, J. (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* 7, 639–655. doi: 10.1046/j.1365-294x.1998.00374.x

Martin, F., Roth, D. M., Jans, D. A., Pouton, C. W., Partridge, L. J., Monk, P. N., et al. (2005). Tetraspanins in viral infections: a fundamental role in viral biology? *J. Virol.* 79, 10839–10851. doi: 10.1128/JVI.79.17.10839-10851.2005

Martin, S. A. M., Douglas, A., Houlihan, D. F., and Secombes, C. J. (2010). Starvation alters the liver transcriptome of the innate immune response in Atlantic salmon (*Salmo salar*). *BMC Genomics* 11:418. doi: 10.1186/1471-2164-11-418

Milne, I., Stephen, G., Bayer, M., Cock, P. J. A., Pritchard, L., Cardle, L., et al. (2013). Using tablet for visual exploration of second-generation sequencing data. *Brief. Bioinform.* 14, 193–202. doi: 10.1093/bib/bbs012

MPA (2013a). *Boletim Estatístico da Pesca e Aquicultura—Brasil 2011*. Ministério da Pesca e Aquicultura.

MPA (2013b). *Censo Aquícola Nacional—Ano 2008*. Ministério da Pesca e Aquicultura.

Mutz, K. O., Heilkenbrinker, A., Lönne, M., Walter, J. G., and Stahl, F. (2013). Transcriptome analysis using next-generation sequencing. *Curr. Opin. Biotechnol.* 24, 22–30. doi: 10.1016/j.copbio.2012.09.004

Neira, R., Díaz, N. F., Gall, G. A., Gallardo, J. A., Lhorente, J. P., and Manterola, R. (2006). Genetic improvement in Coho salmon (*Oncorhynchus kisutch*). I: selection response and inbreeding depression on harvest weight. *Aquaculture* 257, 9–17. doi: 10.1016/j.aquaculture.2006.03.002

Ødegård, J., Baranski, M., Gjerde, B., and Gjedrem, T. (2011). Methodology for genetic evaluation of disease resistence in aquaculture species: challenges and future prospects. *Aquacult. Res.* 42, 103–114. doi: 10.1111/j.1365-2109.2010.02669.x

Ødegård, J., Sommer, A. I., and Praebel, A. K. (2010). Heritability of resistance to viral nervous necrosis in Atlantic cod (*Gadus morhua*). *Aquaculture* 300, 59–64. doi: 10.1016/j.aquaculture.2010.01.006

Panarari-Antunes, R. S., Prioli, A. J., Prioli, S. M. A. P., Galdino, A. S., Junior, J., Ferreira, H., et al. (2011). Genetic variability of brycon orbignyanus (Valenciennes,1850) (*Characiformes*: *Characidae*) in cultivated and natural populations of the Upper Paraná River, and implications for the conservation of the species. *Braz. Arch. Biol. Technol.* 54, 839–848. doi: 10.1590/S1516-89132011000400025

Pardo, B. G., Fernández, C., Millán, A., Bouza, C., Vázquez-López, A., Vera, M., et al. (2008). Expressed sequence tags (ESTs) from immune tissues of turbot (*Scophthalmus maximus*) challenged with pathogens. *BMC Vet. Res.* 4:37. doi: 10.1186/1746-6148-4-37

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28, 2537–2539. doi: 10.1093/bioinformatics/bts460

Piry, S., Luikart, G., and Cornuet, J. M. (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J. Hered.* 90, 502–503. doi: 10.1093/jhered/90.4.502

Ponzoni, R. W., Nguyen, N. H., Khaw, H. L., and Ninh, N. H. (2008). Accounting for genotype by environment interaction in economic appraisal of genetic improvement programs in common carp *Cyprinus carpio*. *Aquaculture* 285, 47–55. doi: 10.1016/j.aquaculture.2008.08.012

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Qian, X., Ba, Y., Zhuang, Q., and Zhong, G. (2014). RNA-Seq technology and its application in fish transcriptomics. *OMICS* 18, 98–110. doi: 10.1089/omi.2013.0110

Renaut, S., Nolte, A. W., and Bernatchez, L. (2010). Mining transcriptoma sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus* spp. *Salmonidae*). *Mol. Ecol. Res.* 19, 115–131. doi: 10.1111/j.1365-294X.2009.04477.x

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* 43, 223–225. doi: 10.1111/j.1558-5646.1989.tb04220.x

Rousset, F. (2008). GENEPOP'007: a complete re-implementation of the GENEPOP software for windows and linux. *Mol. Ecol. Resour.* 8, 103–106. doi: 10.1111/j.1471-8286.2007.01931.x

Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132, 365–386. doi: 10.1385/1-59259-192-2:365

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nat. Biotechnol.* 18, 233–234. doi: 10.1038/72708

Seeb, J. E., Carvalho, G., Hauser, L., Naish, K., Roberts, S., and Seeb, L. W. (2011). Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Mol. Ecol. Res.* 11, 1–8. doi: 10.1111/j.1755-0998.2010.02979.x

Shin, S. C., Kim, S. J., Lee, J. K., Ahn, D. H., Kim, M. G., Lee, H., et al. (2012). Transcriptomics and comparative analysis of three Antarctic notothenioid fishes. *PLoS ONE* 7:e43762. doi: 10.1371/journal.pone.0043762

Shirk, R. Y., Glenn, T. C., Chang, S., and Hamrick, J. L. (2013). Development and characterization of microsatellite Primers in *Geranium carolinianum* (Geraniaceae) with 454 sequencing. *Appl. Plant Sci.* 1:1300006. doi: 10.3732/apps.1300006

Shoshana, L., and Shoham, T. (2005). The tetraspanin web modulates immune – signaling complexes. *Nat. Rev. Immunol.* 5, 136–148. doi: 10.1038/nri1548

Teacher, A., Kähkönen, K., and Merilä, J. (2012). Development of 61 new transcriptome-derived microsatellites for the Atlantic herring (*Clupea harengus*). *Conserv. Genet. Resour.* 4, 71–74. doi: 10.1007/s12686-011-9477-5

Theodorou, K., and Couvet, D. (2015). The efficiency of close inbreeding to reduce genetic adaptation to captivity. *Heredity* 114:38. doi: 10.1038/hdy.2014.63

Wang, X., Shang, X., Luan, J., and Zhang, S. (2014). Identification, expression and function of apolipoprotein e in animal fish *Nothobranchis guentheri*: implication for an aging marker. *Biogerontology* 15, 233–243. doi: 10.1007/s10522-014-9493-4

Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19, 395–420. doi: 10.1111/j.1558-5646.1965.tb01731.x

Xu, J., Liu, L., Xu, Y., Chen, C., Rong, T., Ali, F., et al. (2013). Development and characterization of simple sequence repeat markers providing genome-wide coverage and high resolution in maize. *DNA Res.* 20, 497–509. doi: 10.1093/dnares/dst026

Yáñez, J. M., Houston, R. D., and Newman, S. (2014). Genetics and genomics of disease resistance in salmonid species. *Front. Genet*. 5:415. doi: 10.3389/fgene.2014.00415

Yue, G. H., and Orban, L. (2002). Microsatellites from genes show polymorphism in two related *Oreochromis* species. *Mol. Ecol. Notes* 2, 99–100. doi: 10.1046/j.1471-8286.2002.00159.x

Yue, G., Li, Y., and Orban, L. (2001). Characterization of microsatellites in the *IGF-2* and *GH* genes of Asian seabass (*Lates calcarifer*). *Mar. Biotechnol.* 3, 1–3. doi: 10.1007/s101260000043

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 02 March 2018
doi: 10.3389/fgene.2018.00069

Genetic and Morphological Analyses Demonstrate That *Schizolecis guntheri* (Siluriformes: Loricariidae) Is Likely to Be a Species Complex

Camila S. Souza[1], Guilherme J. Costa-Silva[1,2], Fábio F. Roxo[1], Fausto Foresti[1] and Claudio Oliveira[1]*

[1] Departamento de Morfologia, Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu, Botucatu, Brazil, [2] Departamento de Biologia, Universidade Santo Amaro, São Paulo, Brazil

Schizolecis is a monotypic genus of Siluriformes widely distributed throughout isolated coastal drainages of southeastern Brazil. Previous studies have shown that fish groups found in isolated river basins tend to differentiate over time in the absence of gene flow, resulting in allopatric speciation. In this study, we used partial sequences of the mitochondrial gene COI with the analysis of the General Mixed Yule Coalescent model (GMYC) and the Automatic Barcode Gap Discovery (ABGD) for single locus species delimitation, and a Principal Component Analysis (PCA) of external morphology to test the hypothesis that *Schizolecis guntheri* is a complex of species. We analyzed 94 samples of *S. guntheri* for GMYC and ABGD, and 82 samples for PCA from 22 coastal rivers draining to the Atlantic in southeastern Brazil from the Paraná State to the north of the Rio de Janeiro State. As a result, the GMYC model and the ABGD delimited five operational taxonomy units (OTUs – a nomenclature referred to in the present study of the possible new species delimited for the genetic analysis), a much higher number compared to the traditional alfa taxonomy that only recognizes *S. guntheri* across the isolated coastal rivers of Brazil. Furthermore, the PCA analysis suggests that *S. guntheri* is highly variable in aspects of external body proportions, including dorsal-fin spine length, pectoral-fin spine length, pelvic-fin spine length, lower caudal-fin spine length, caudal peduncle depth, anal width and mandibular ramus length. However, no exclusive character was found among the isolated populations that could be used to describe a new species of *Schizolecis*. Therefore, we can conclude, based on our results of PCA contrasting with the results of GMYC and ABGD, that *S. guntheri* represents a complex of species.

Keywords: coastal drainages, catfish, molecular identification, COI gene, GMYC model

OPEN ACCESS

Edited by:
Samuel A. Cushman, United States Forest Service (USDA), United States

Reviewed by:
Gonzalo Gajardo, University of Los Lagos, Chile
Chandra Shekhar Prabhakar, Bihar Agricultural University, India

***Correspondence:**
Fábio F. Roxo
roxoff@hotmail.com.br

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

Received: *25 August 2017*
Accepted: *15 February 2018*
Published: *02 March 2018*

Citation:
Souza CS, Costa-Silva GJ, Roxo FF, Foresti F and Oliveira C (2018) Genetic and Morphological Analyses Demonstrate That Schizolecis guntheri (Siluriformes: Loricariidae) Is Likely to Be a Species Complex. Front. Genet. 9:69. doi: 10.3389/fgene.2018.00069

INTRODUCTION

The distribution pattern of single fish species throughout independent hydrographic systems (i.e., current not connected rivers) is unusual among fishes of the Atlantic rainforest rivers (Menezes et al., 2007) as well as among members of Otothyrinae (Reis et al., 2003). Recently, several genetic studies focusing on freshwater fishes, such as *Rineloricaria* (Costa-Silva et al., 2015), *Curimatopsis*

(Melo et al., 2016), *Piabina* (Pereira et al., 2011), and *Astyanax* (Ornelas-Garcia et al., 2008), have shown that species of these groups may present large discontinuities in their distribution patterns with high genetic divergences, but with low morphological variability among geographically isolated populations. These results suggest that these groups may represent a complex of species –, i.e., *they are constituted by two or more morphological variable species that are erroneously classified (and hidden) under one species name* (Brown et al., 1995). Usually, studies focused on morphology alone are inadequate to recognize species complex. Integrative studies using molecular markers (e.g., DNA sequencing and allozymes), in addition to morphological comparison, are more powerful for the recognition of possible new species in such complex groups (Sytsma and Schaal, 1985).

The fast development of DNA sequencing and advances in molecular techniques in the last few years have been effective in recognizing species in several organism groups – birds (e.g., Tavares et al., 2011; Saitoh et al., 2015), fishes (e.g., Ward et al., 2005; Pereira et al., 2011, 2013; Roxo et al., 2012, 2015; Shimabukuro-Dias et al., 2016), insects (e.g., Hebert et al., 2004; Versteirt et al., 2015; Batovska et al., 2016), mammals (e.g., Borisenko et al., 2008; Li et al., 2015), plants (Kress et al., 2005; Lahaye et al., 2008; Braukmann et al., 2017), fungi (Begerow et al., 2010; Kelly et al., 2011), archaea (Bates et al., 2011), and bacteria (Sogin et al., 2006). The use of DNA sequences combined with several analytical methods, such as General Mixed Yule Coalescent (GMYC; Pons et al., 2006) and Automatic Barcode Gap Discovery (ABGD; Puillandre et al., 2012), support species delineation with single-locus data. The GMYC method is based on a likelihood method that seeks to determine the threshold between speciation and coalescent events from an ultrametric gene tree, whereas ABGD methods use the gap among organisms belonging to the same species and organisms from different species as a limit to species delimitation.

Schizolecis was described by Britski and Garavello (1984), being the species-type *Microlepidogaster guntheri* Miranda Ribeiro (1918). Currently *Schizolecis guntheri* is the only species of *Schizolecis*. This species is a descendant of a very ancient lineage that arose during the Middle Eocene approximately 42 Mya (Roxo et al., 2014), and inhabits small to median size streams with rocky and sandy bottoms, mostly in shallows and backwaters up to 30 cm deep, with slow water flow (Burgess, 1989). The work conducted by Britski and Garavello (1984) detected morphological differences only related to orbits of the eyes, body depth and head depth among populations, but without enough evidence to support the hypothesis that some of the analyzed populations could represent a new species. Therefore, despite *Schizolecis guntheri* being widely distributed across adjacent and not connected Atlantic Coastal rivers from the north of Santa Catarina to the north of Rio de Janeiro States (Menezes et al., 2007), and present small morphological variations among isolated populations; the doubt of whether *S. guntheri* represents a complex of species still remains.

In the present study, we used genetic data of 94 samples of *Schizolecis guntheri* from 22 coastal rivers draining directly to the Atlantic in southeastern Brazil and employed analytical methods to support species delineation with single-locus data (GMYC and ABGD), and analyzed the body shape variation among isolated populations using a PCA to test whether this species represents a complex of species.

MATERIALS AND METHODS

Morphological Analysis

Principal Component Analysis (PCA) (Jolliffe, 2002) was used to check the external morphology variation among 82 samples of *Schizolecis guntheri* among regions of genetic groups (Supplementary Table S3) using the program Past version v1.28 (Hammer et al., 2004). Landmarks distances followed those originally proposed by Carvalho and Reis (2009), and they were measured for adult specimens (>26.2 mm SL). Prior to the PCA analysis, we followed the method of Dryden and Mardia (1998) to minimize body size influence on morphometric data. We normalized the first two coordinate dimensions, divided all coordinate values by the centroid size for each specimen, and conducted a Procrustes superimposition of the left half to a mirrored version of the right half. After that we also log

TABLE 1 | Variable loadings in the first and second axes of the size-free principal component analysis (Axis 1 and Axis 2) of samples of *Schizolecis guntheri*.

	Axis 1	Axis 2
Standard length	0.03469	0.01922
Predorsal length	0.05589	0.05622
Preanal length	−0.0378	−0.08091
Head length	0.1271	0.0697
Cleithral width	0.1482	0.1679
Dorsal-fin spine length	**0.3355**	−0.2164
Base of dorsal-fin length	0.1576	0.002704
Thorax length	0.1218	0.01957
Pectoral-fin spine length	**0.3305**	−0.1625
Abdomen length	0.1143	0.1715
Pelvic-fin spine length	**0.3478**	0.1035
Anal-fin spine length	0.2795	−0.2604
Lower caudal-fin spine length	**0.5282**	−0.3743
Caudal peduncle depth	0.1384	**0.3224**
Caudal peduncle length	−0.04707	0.05234
Anal width	**0.3044**	0.287
Snout-opercle length	0.07479	0.1523
Head width	0.1181	0.14
Head depth	0.1298	0.2021
Snout length	0.07969	0.153
Interorbital width	0.08071	0.1185
Orbital diameter	0.01212	0.2747
Suborbital depth	0.1547	0.1987
Mandibular ramus length	0.127	**0.4584**

Bold values represent the character with the highest variations.

transformed the data for base 10. The PCA loadings are presented in **Table 1**.

Taxon Sampling for Genetic Analysis

We analyzed 94 *Schizolecis guntheri* specimens from 22 coastal rivers draining directly to the Atlantic, from the south of Paraná to the north of Rio de Janeiro States (**Figure 1**), almost comprising the entire distribution of this species. Information about the sample used in the present study is available in Bold Systems with the accession numbers for individuals present in Supplementary Table S1. The vouchers and tissues are deposited in the fish collection of the LBP – Laboratório de Biologia e Genética de Peixes, Universidade Estadual Paulista, Botucatu, São Paulo, Brazil. A sequence from one additional species of Loricariidae (*Hypostomus ancistroides*) was used as an outgroup to root our tree.

The abbreviation OTU (operational taxonomical unit) was used in the text to refer to the possible new species of the molecular analysis of the GMYC model and ABGD method (*sensu* Blaxter et al., 2005). This last author defined the term OTU to refer to genetic clusters of unknown organisms grouped by DNA sequences.

Ethics Statement

All fishes collected for this study were collected in accordance with Brazilian laws, under a scientific collection license in the name of Dr. Claudio Oliveira (SISBIO). Furthermore, our laboratory has special federal permission to keep animals and tissues from a public collection under our care. To work with the animals, we followed all the ethical prescriptions stated by our internal committee of ethics (protocol number 388), called the "Comissão de Ética na Experimentação Animal" (CEEA), involving animal experiments that were *approved* for this study. After collection, animals were anesthetized with benzocaine, a piece of muscle tissue was extracted from the right side of the body and preserved in 95% ethanol. Specimens were fixed in 10% formalin for 2 weeks, and then transferred to 70% ethanol for permanent storage.

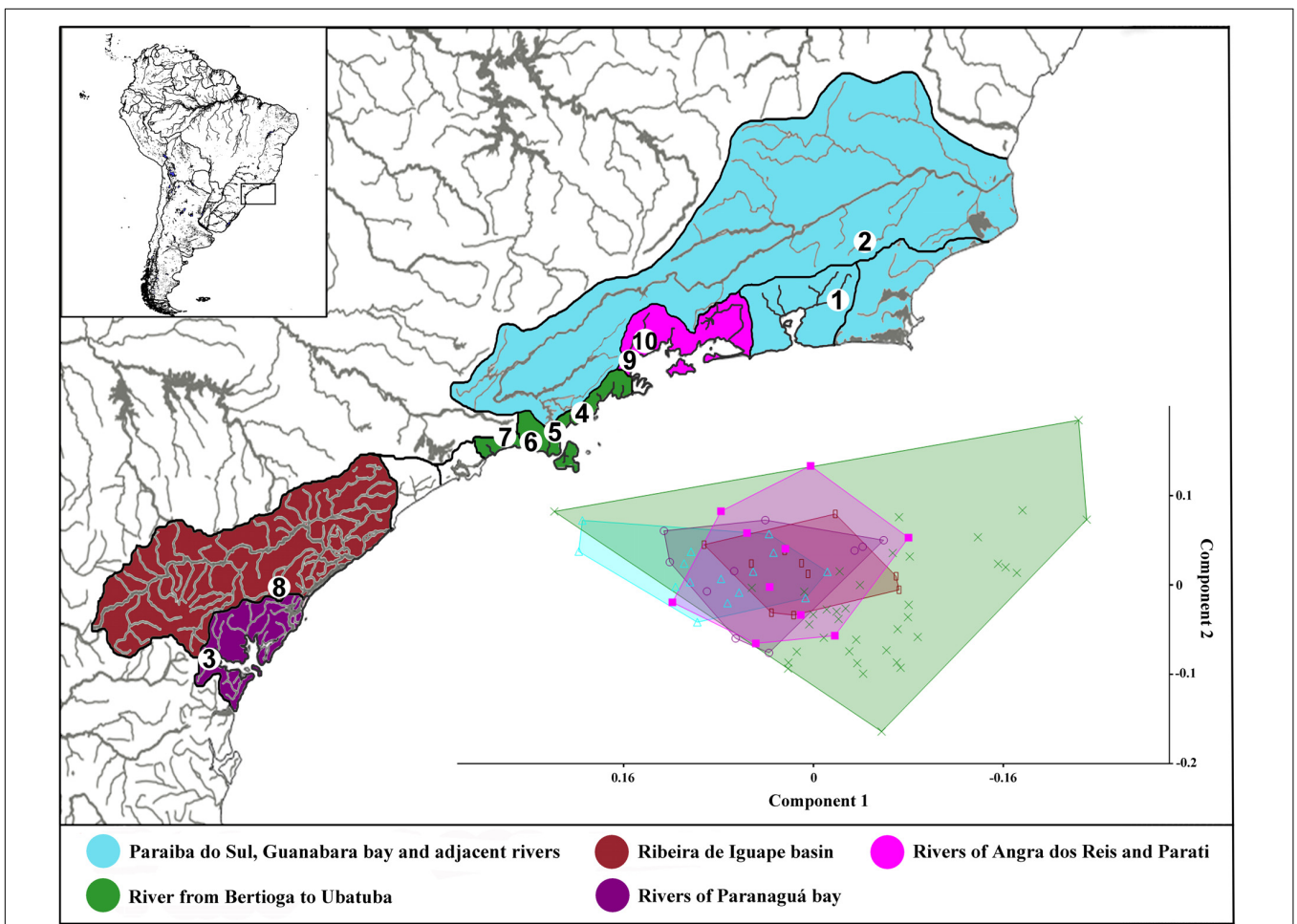


FIGURE 1 | Image showing the map of coastal rivers of southeastern of Brazil following the paleodrainages inference of Thomaz et al. (2015), and adding the rivers of Guanabara Bay, the Paraiba do Sul basin and adjacent rivers. Numbers in the maps represents collection sites: **1** – Itaboraí-RJ; **2** – Bom Jardim-RJ; **3** – Morretes-PR; **4** – Ubatuba-SP; **5** – Caraguatatuba-SP; **6** – São Sebastião-SP; **7** – Bertioga-SP; **8** – Cajati-SP; **9** – Parati-RJ; and **10** – Angra dos Reis-RJ. Each number can represent more than one collection site.

DNA Extraction, Amplification, and Sequencing

We conducted the total genomic DNA extraction using the protocol described by Ivanova et al. (2006). Partial sequences of the cytochrome oxidase C subunit I (COI) gene were amplified (approximately 655 bp) using the primers Fish F1 and Fish R1 (Ward et al., 2005). Amplifications were performed in a total reaction mixture volume of 12.5 µl. Each reaction includes 1.25 µl of 10 X Buffer, 0.25 µl of $MgCl_2$ (50 mM), 0.2 µl dNTPs (2 mM), 0.5 µl of each primer (5 mM), 0.1 µl of Pht Taq DNA polymerase (Phoneutria Biotecnologia e Serviços Ltda, Belo Horizonte, Brazil), 1 µl of genomic DNA (200 ng) and 8.7 ml ddH_2O. The conditions for each PCR reaction consisted of an initial denaturation (5 min at 94°C), followed by 30 cycles of chain denaturation (40 s at 94°C), primer hybridization (30 s at 50–54°C), nucleotide extension (1 min at 68°C, considering the optimum temperature of the Pht Tap DNA polymerase) and final extension (8 min at 72°C, to stabilize the reaction). The amplified products were checked on 1% agarose gels and then purified using ExoSap-IT (USB Corporation, Cleveland, OH, United States) following the manufacturer's instructions. We accomplished the sequencing reactions using the BigDye TM Terminator v 3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Austin, TX, United States) and purified again by ethanol precipitation. DNA sequencing was conducted in an ABI 3130 DNA Analyzer automatic sequencer (Applied Biosystems, Foster City, CA, United States).

Genetic Analysis

The consensus sequences were obtained using the program Geneious 7.1.4[1] (Kearse et al., 2012) and the alignment was generated with the algorithm Muscle (Edgar, 2004) under default parameters. To evaluate the occurrence of substitution saturation in our molecular data, we estimated whether the Iss (index of substitution saturation) is significantly lower than Iss.cAsym (assuming asymmetrical topology) using the method described by Xia et al. (2003) with the software DAMBE 5.3.38 (Xia, 2013). After the identification of the OTUs by the GMYC model and ABGD analysis, we calculated the genetic variation within and among the OTUs delimited by each method using the Kimura-2-parameter (K2P) model in the MEGA v.6.06 software (Tamura et al., 2013).

GMYC Model

The GMYC method requires an ultrametric tree as input for the analysis. To estimate the ultrametric tree, we used Beast v1.8.2 (Drummond et al., 2012), employing an uncorrelated lognormal relaxed clock and birth-death speciation process, and the General Time Reversible (GTR) model (Lanave et al., 1984; Tavaré, 1986). The Bayesian topology reconstruction started with a UPGMA tree and the Markov Chain Monte Carlo (MCMC) method was performed for 100 million generations; a tree was sampled for every 20,000 generations. We used the software Tracer v1.6 (Rambaut et al., 2014) to check the convergence of the values. All sampled topologies beneath the asymptote (20,000,000 generations) were discarded as part of a burn-in procedure, and the remaining trees were used to construct a 90% majority-rule consensus tree using Tree Annotator v1.8.2 (Drummond et al., 2012). The GMYC analysis was performed with the package Species Limits by Threshold Statistics ("splits") (Fujisawa and Barraclough, 2013) using R v3.0.0 (R Development Core Team, 2014) that only includes the ingroup (the outgroup *Hypostomus ancistroides* was excluded).

ABGD Analysis

Automatic Barcode Gap Discovery analysis (Puillandre et al., 2012), was processed using the "graphic" web version available at http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html, under the default parameters of P_{min} = 0.001 to P_{max} = 0.1, steps = 10, X (relative gap width) = 1.5, Nb bins (for distance distribution) = 20, and the Kimura (K80) molecular model. For this analysis, the external group (*H. ancistroides*) was excluded from the input data.

RESULTS

Morphological Analysis

The first (PC1) and second (PC2) principal component axis of our analysis explained 28.1% and 12.4%, respectively, of variation in body shape for all analyzed *Schizolecis guntheri* specimens. The variation is partly distributed within populations and partly between populations, and apparently, it represents a continuous distribution of external morphology, as we can observe in the PCA scatter plot (**Figure 1**). Our results also showed that the measures with greater variations were, respectively: dorsal-fin spine length, pectoral-fin spine length, pelvic-fin spine length, lower caudal-fin spine length, caudal peduncle depth, anal width and mandibular ramus length, as we can observe in the PCA loading values (**Table 1**).

Statistics of the DNA Matrix

We obtained a matrix with 533 characters, 230 of which were variable. The nucleotide frequencies were A (25.4%), G (16.7%), T (29.3%), and C (28.7%). No insertions, deletions, stop codons or contamination in the sequences were detected. The data were not saturated considering that the Iss.cAsym values are higher than the Iss for the different numbers of NumOTU analyzed (4 OTUs: Iss = 0.187, Iss.cAsym = 0.763, P = 0; 8 OTUs: Iss = 0. 177, Iss.cAsym = 0.643, P = 0; 16 OTUs: Iss = 0.179, Iss.cAsym = 0.516, P = 0; and 32 OTUs: Iss = 0.180, Iss.cAsym = 0.383, P = 0) in the total matrix (all molecular characters including gaps).

GMYC Model and ABGD Analysis

The phylogenetic reconstruction resulted in a tree with high values of posterior probability among the clades (>95%), highlighting the existence of five lineages within *Schizolecis guntheri* (**Figure 2**). The analysis of species delimitation using the GMYC model under a phylogenetic tree estimated using a Birth–Death model prior of branching rates showed that the found threshold time was −0.0055687, indicating the time

[1] http://www.geneious.com

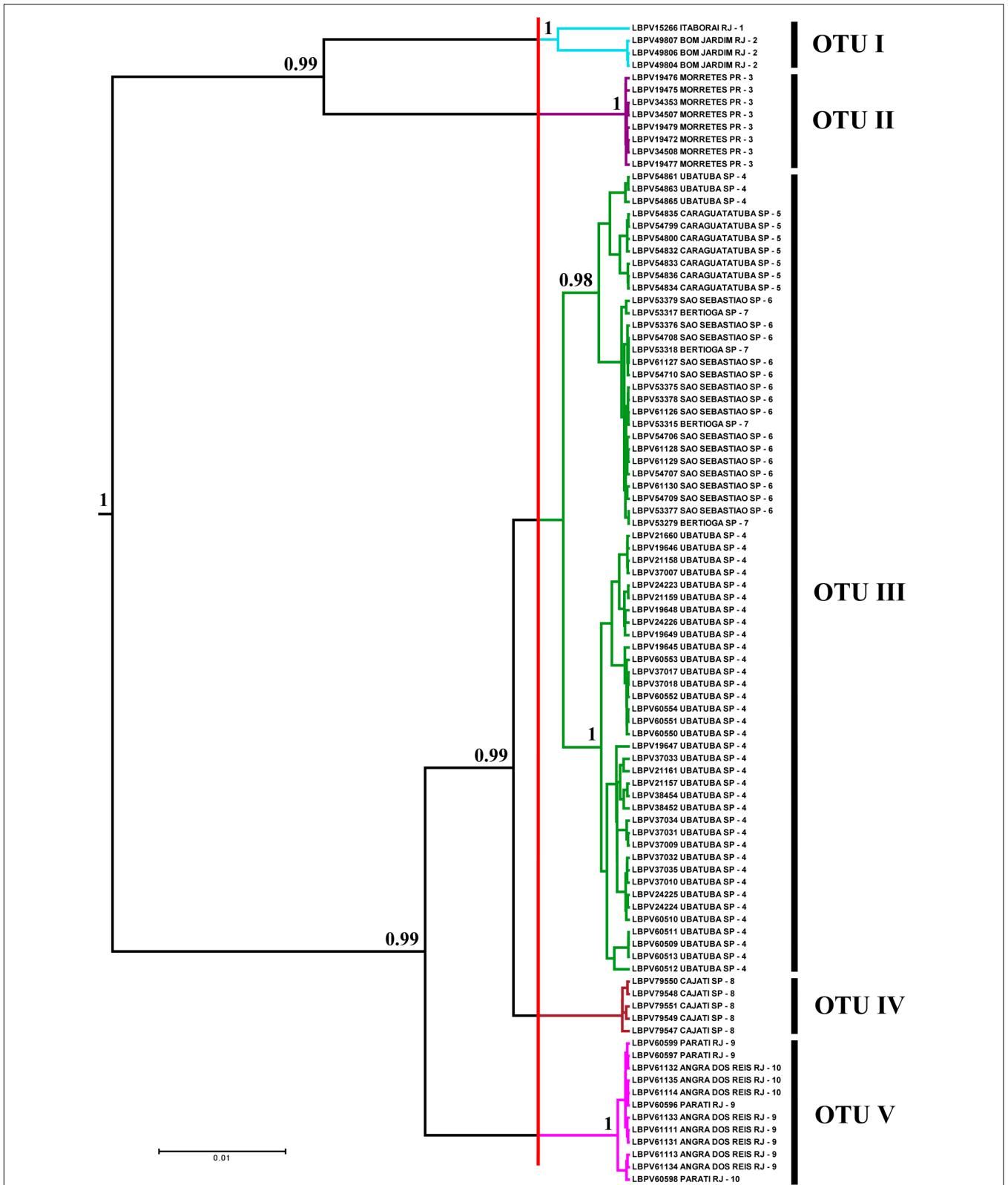


FIGURE 2 | Bayesian tree showing the distribution of specimens by OTUs selected through GMYC analysis. Numbers after branches are posterior probabilities. Values below 0.95 are not shown. The vertical red line represents the threshold time (–0.0055687) – the limit between species and population according to the GMYC model in the tree. The colors of the branches to the right of the red line correspond to the colors of the paleodrainages in **Figure 1**, and numbers after are local to collection sites also in **Figure 1**. LBPV is the code for samples of the fish collection of the *Laboratório de Biologia e Genética de Peixes* on the Bold systems site.

before which all nodes reflect diversification events and after which all nodes in the tree reflect coalescent events. The likelihood of the null model was 958.2078 and the maximum likelihood of the GMYC model was 962.0088. The standard log-likelihood ratio test is used to assess whether the alternative model (GMYC model) provides a better fit than the null model in the branching process (see Goldman, 1993 and Pons et al., 2006 for a better explanation about how the GMYC model assesses species delimitation).

The analysis of ABGD partitioned the data from 19 groups (P_{min} = 0.0010003) to 3 groups (P_{max} = 0.021544). The value of P = 0.004642 delimited five groups, the same number of groups delimited by the GMYC model (see Supplementary Table S1). Therefore, the OTUs for *Schizolecis guntheri* were divided following the genetic delimitation of five clusters (OTUs) according to the results of both genetic methods, and named according to their localities: OTU I – Itaboraí-RJ and Bom Jardim-RJ; OTU II – Morretes-PR; OTU III – Ubatuba-SP, Caraguatatuba-SP, São Sebastião-SP and Bertioga-SP; OTU IV – Cajati-SP; and OTU V – Angra dos Reis and Parati (**Figures 1**, **2** and Supplementary Table S1). The values of genetic distance among the five OTUs ranged from 1.1% (OTU III and OTU IV) to 8.7% (OTU II and OTU V) (**Table 2**).

DISCUSSION

The results of the Bayesian and GMYC analyses (**Figure 2**) of the samples from 22 coastal rivers localities draining directly to the Atlantic of southeastern Brazil (**Figure 1**) highlighted the existence of five monophyletic and highly statistically supported (>95%) lineages (OTUs *sensu* Blaxter et al., 2005) within *Schizolecis guntheri* (**Figure 2**), the same result found by the ABGD method (Supplementary Table S2) that also divided the data into five clusters (or five OTUs). Several authors (e.g., Hänflig and Brandl, 1998; Paggi et al., 1998; Baric and Sturmbauer, 1999; Ferguson, 2002) have summarized the arguments for using genetic divergence for identifying separate species. According to these authors, sufficient genetic distance indicates reproductive isolation between probable/possible species (*sensu* Mayr, 1963) that gradually accumulates genetic differences between lineages, and after long-time periods, accumulates sufficient genetic divergence so that a speciation event could be detected.

TABLE 2 | Genetic divergences based on the Kimura-2-parameter (K2P) nucleotide model for OTUs delimitated by GMYC analysis.

	OTU I	OTU II	OTU III	OTU IV	OTU V
OTU I	**1.1 ± 0.3**				
OTU II	6.4 ± 1.2	**0**			
OTU III	8.1 ± 1.3	8.6 ± 1.4	**0.8 ± 0.3**		
OTU IV	8.0 ± 1.4	8.4 ± 1.5	1.1 ± 0.4	**0**	
OTU V	8.0 ± 1.4	8.7 ± 1.5	2.3 ± 0.7	1.9 ± 0.6	**0**

In the main diagonal are the values of intragroup genetic divergences highlighted in bold. Below the main diagonal are the values of intergroup (OTUs) genetic divergences. The values are shown in percentages.

However, despite the similarity of the results of the GMYC and ABGD analysis methods in the present study, if we apply the 2% threshold of interspecific genetic divergence as a limit among population and species (as proposed by Smith et al., 2005) the OTU III and OTU IV with 1.1% and OTU IV and OTU V with 1.9% (**Table 2**) should be considered members of the same cluster. The different number of species delimited for different molecular methods is also a problematic question in species delimitation using single locus genes (Pons et al., 2006). However, the two analytical molecular methods used in the present study (e.g., GMYC and ABGD) resulted in the same numbers of OTUs (i.e., five OTUs) and the application of these methods has been encouraging, highlighting a hidden genetic diversity in several neotropical fish species (e.g., Costa-Silva et al., 2015; Roxo et al., 2015; Melo et al., 2016). Furthermore, molecular studies in several Neotropical fish groups (e.g., Pereira et al., 2011; Costa-Silva et al., 2015; Melo et al., 2016) have shown that widely distributed nominal species isolated in independent hydrographic systems can sometimes represent species complex (i.e., species with continuous morphological variation, but discontinuous variation in genetic analysis) and the combined usage of DNA barcoding and morphological data has provided support to recognize and describe possible new species (Melo et al., 2011; Amaral et al., 2013; Silva et al., 2013).

The results of *Schizolecis guntheri* morphological analysis exhibits high external morphological variation across its

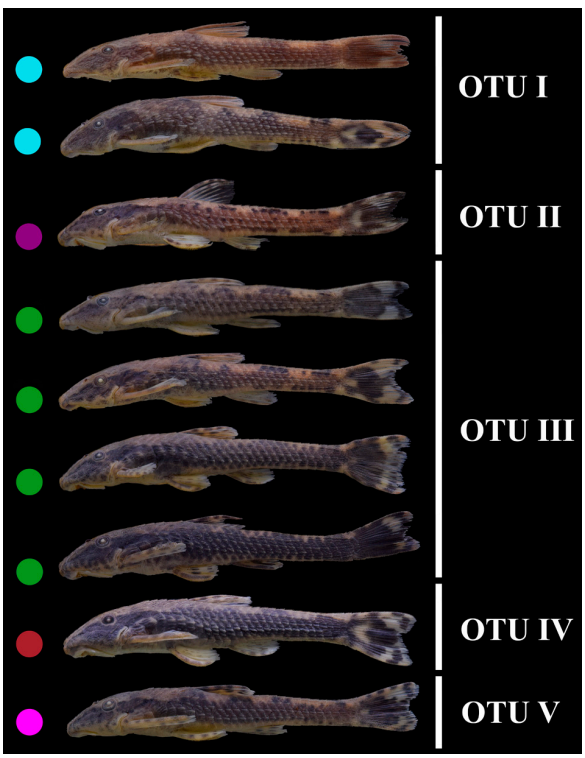


FIGURE 3 | Pictures showing the variation in external morphology among *Schizolecis guntheri* populations. Colored circles represent the paleodrainages proposed by Thomaz et al. (2015) shown in **Figure 1**.

distribution, including color pattern (**Figures 1**, **3**), but especially in morphometric characteristics, such as dorsal-fin spine length, pectoral-fin spine length, pelvic-fin spine length, lower caudal-fin spine length, caudal peduncle depth, anal width and mandibular ramus length, as we can observe in the PCA loading values (**Table 1**). However, this diversity is partly distributed within populations and among populations, and represents a continuous variation in external morphology, as we can observe in the PCA scatter plot (**Figure 1**). This continuous morphological variation is typically found in a species complex (Brown et al., 1995). Therefore, we do not find any exclusive character to support a possible new species for any of the OTUs distributed across the isolated coastal drainages of the present study. A similar result was previously found by Britski and Garavello (1984). These authors found differences in the morphometric characters of eyes orbits, body depth and head depth, but no exclusive character that could support a new species of the genus *Schizolecis* among the analyzed samples.

De Queiroz (2007) argued that the confusion among species delimitation is associated with the fact that different methods, including different properties (molecular or morphological), are focused on different stages of the speciation process. Our results suggested different numbers of OTUs of *S. guntheri* delimited based on the GMYC model and ABGD (Genotypic Cluster) that recognized five monophyletic groups and the morphological analysis (Phenetic Cluster) that only recognized *S. guntheri* with a continuous variation in the external morphology, as shown by the PCA (typical results of a species complex). With the passing of time, two independent lineages develop and increasingly acquire different properties relative to each other –, i.e., they become phenetically distinct, reach the reciprocal monophyly, became ecologically distinct or reproductively incompatible. Before the recovery of the first estate, everybody will recognize that there is only one species, and after the acquisition of several estates, everyone will recognize two species. Otherwise, in between, there will be confusion. De Queiroz (2007) called the area where two groups of species come into conflict and the boundaries among species are unclear as the "gray zone." On either side of the gray zone, there will be consistent agreement about the species number, but the "gray zone" has conflict. Therefore, the conflict among different numbers of recognized species between genetic methods (GMYC and ABGD) and the morphological analysis in *S. guntheri* could be associated with the fact that this species is in the gray zone suggested by De Queiroz (2007). Furthermore, a species complex could be interpreted as species that are in the gray zone (e.g., molecularly distinguished but morphologically undistinguished).

Therefore, considering the lack of phenotypic discontinuities and the presence of relatively high levels of genetic divergence among some local populations of *Schizolecis guntheri*, we hypothesize that this species may represent a species complex, like suggested for other freshwater fish species (Pereira et al., 2011; Bellafronte et al., 2013; Marques et al., 2013).

AUTHOR CONTRIBUTIONS

CS, CO, GC-S, and FR designed the ideas of the research. CS collected data. CS, GC-S, and FR performed the analyses. CS, CO, GC-S, and FR contributed to the writing of the manuscript. FF provided physical structure of the laboratory to develop the work.

FUNDING

This research was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico proc. 150415/2015–0 to GC-S), FAPESP (Fundo de Amparo à Pesquisa do Estado de São Paulo, proc. 2014/05051–5 and 2015/00691–9 to FR), and MCT/CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) (Edital Universal, proc. N. 441347/2014–2 coord. FR).

ACKNOWLEDGMENTS

The authors are grateful to Priscila Camelier and Renato Devidé for their help during the collection expeditions. The authors also wish to thank Gabriel S. C. Silva for his help with the identification of the specimens.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00069/full#supplementary-material

REFERENCES

Amaral, C. R. L., Brito, P. M., Silva, D. A., and Carvalho, E. F. (2013). A new cryptic species of South American freshwater pufferfish of the genus *Colomesus* (Tetraodontidae), based on both morphology and DNA data. *PLoS One* 8:e74397. doi: 10.1371/journal.pone.0074397

Baric, S., and Sturmbauer, C. (1999). Ecological parallelism and cryptic species in the genus *Ophiothrix* derived from mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* 11, 157–162. doi: 10.1006/mpev.1998.0551

Bates, S. T., Berg-Lyons, D., Caporaso, J. G., Walters, W. A., Knight, R., and Fierer, N. (2011). Examining the global distribution of dominant archaeal populations in soil. *ISME J.* 5, 908–917. doi: 10.1038/ismej.2010.171

Batovska, J., Blacket, M. J., Brown, K., and Lynch, S. E. (2016). Molecular identification of mosquitoes (Diptera: Culicidae) in southeastern Australia. *Ecol. Evol.* 6, 3001–3011. doi: 10.1002/ece3.2095

Begerow, D., Nilsson, H., Unterseher, M., and Maier, W. (2010). Current state and perspectives of fungal DNA barcoding and rapid identification procedures. *Appl. Microbiol. Biotechnol.* 87, 99–108. doi: 10.1007/s00253-010-2585-4

Bellafronte, E., Mariguela, T., Garcia-Pereira, L., Oliveira, C., and Moreira-Filho, O. (2013). DNA barcode of Parodontidae species from the La Plata river basin-applying new data to clarify taxonomic problems. *Neotrop. Ichthyol.* 11, 497–506. doi: 10.1590/S1679-62252013000300003

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., et al. (2005). Defining operational taxonomic units using DNA barcode data. *Philos.*

Trans. R. Soc. Lond. B Biol. Sci. 360, 1935–1943. doi: 10.1098/rstb.2005.1725

Borisenko, A. V., Lim, B. K., Ivanova, N. V., Hanner, R. H., and Hebert, P. D. N. (2008). DNA barcoding in surveys of small mammal communities: a field study in Suriname. *Mol. Ecol. Res.* 8, 471–479. doi: 10.1111/j.1471-8286.2007.01998.x

Braukmann, T. W., Kuzmina, M. L., Sills, J., Zakharov, E. V., and Hebert, P. D. (2017). Testing the efficacy of DNA barcodes for identifying the vascular plants of Canada. *PLoS One* 12:e0169515. doi: 10.1371/journal.pone.0169515

Britski, H., and Garavello, J. (1984). Two new Southeastern Brazilian genera of hypoptopomatinae and a redescription of *Pseudotocinclus* Nichols, 1919 (Ostariophysi, Loricariidae). *Pap. Avulsos Zool.* 35, 225–241.

Brown, J. K., Frohlich, D. R., and Rosell, R. C. (1995). The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or a species complex? *Annu. Rev. Entomol.* 40, 511–534. doi: 10.1146/annurev.en.40.010195.002455

Burgess, W. E. (1989). *An Atlas of Freshwater and Marine Catfishes in A Preliminary Survey of the Siluriformes*. Neptune City, NJ: T.F.H. Publications, Inc., 784.

Carvalho, T. P., and Reis, R. E. (2009). Four new species of *Hisonotus* (Siluriformes: Loricariidae) from the upper rio Uruguay, southeastern South America, with a review of the genus in the rio Uruguay basin. *Zootaxa* 2113, 1–40.

Costa-Silva, G. J., Rodriguez, M. S., Roxo, F. F., Foresti, F., and Oliveira, C. (2015). Using different methods to access the difficult task of delimiting species in a complex neotropical hyperdiverse group. *PLoS One* 10:e0135075. doi: 10.1371/journal.pone.0135075

De Queiroz, K. (2007). Species concepts and species delimitation. *Syst. Biol.* 56, 879–886. doi: 10.1080/10635150701701083

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. doi: 10.1093/molbev/mss075

Dryden, I. L., and Mardia, K. V. (1998). *Statistical Shape Analysis*. New York, NY: John Wiley & Sons, 347.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Ferguson, J. W. H. (2002). On the use of genetic divergence for identifying species. *Biol. J. Linn. Soc.* 75, 509–516. doi: 10.1046/j.1095-8312.2002.00042.x

Fujisawa, T., and Barraclough, T. G. (2013). Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent (GMYC) approach: a revised method and evaluation on simulated datasets. *Syst. Biol.* 62, 707–724. doi: 10.1093/sysbio/syt033

Goldman, N. (1993). Statistical tests of models of DNA substitution. *J. Mol. Evol.* 36, 182–198. doi: 10.1007/BF00166252

Hammer, O., Harper, D. A. T., and Ryan, P. D. (2004). *Past - Palaeontological Statistics, ver. 1.32*. Oslo: University of Oslo.

Hänflig, B., and Brandl, R. (1998). Genetic differentiation of the bullhead *Cottos gobio* L. across watersheds in central Europe – evidence for two taxa. *Heredity* 80, 110–117. doi: 10.1046/j.1365-2540.1998.00279.x

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., and Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 14812–14817. doi: 10.1073/pnas.0406166101

Ivanova, N. V., Dewaard, J. R., and Hebert, P. D. N. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Mol. Ecol. Notes* 6, 998–1002. doi: 10.1111/j.1471-8286.2006.01428.x

Jolliffe, I. T. (2002). *Principal Component Analysis. Springer Series in Statistics*, 2nd Edn. New York, NY: Springer, 487.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kelly, L. J., Hollingsworth, P. M., Coppins, B. J., Ellis, C. J., Harrold, P., Tosh, J., et al. (2011). DNA barcoding of lichenized fungi demonstrates high identification success in a floristic context. *New Phytol.* 191, 288–300. doi: 10.1111/j.1469-8137.2011.03677.x

Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A., and Janzen, D. H. (2005). Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8369–8374. doi: 10.1073/pnas.0503123102

Lahaye, R., Van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., et al. (2008). DNA barcoding the floras of biodiversity hotspots. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2923–2928. doi: 10.1073/pnas.0709936105

Lanave, C., Preparata, G., Saccone, C., and Serio, G. (1984). A new method for calculating evolutionary substitution rates. *J. Mol. Evol.* 20, 86–93. doi: 10.1007/BF02101990

Li, J., Zheng, X., Cai, Y., Zhang, X., Yang, M., Yue, B., et al. (2015). DNA barcoding of murinae (Rodentia: Muridae) and arvicolinae (Rodentia: Cricetidae) distributed in China. *Mol. Ecol. Resour.* 15, 153–167. doi: 10.1111/1755-0998.12279

Marques, D., Santos, F., and Silva, S. (2013). Cytogenetic and DNA barcoding reveals high divergence within the trahira, *Hoplias malabaricus* (Characiformes:Erythrinidae) from the lower Amazon River. *Neotrop. Ichthyol.* 11, 459–466. doi: 10.1590/S1679-62252013000200015

Mayr, E. (1963). *Animal Species and Evolution*. Cambridge, MA: Belknap Press of Harvard University Press.

Melo, B. F., Benine, R. C., Mariguela, T. C., and Oliveira, C. (2011). A new species of *Tetragonopterus* cuvier, 1816 (Characiformes: Characidae: Tetragonopterinae) from the rio Jari, Amapá, northern Brazil. *Neotrop. Ichthyol.* 9, 49–56. doi: 10.1590/S1679-62252011000100002

Melo, B. F., Ochoa, L. E., Vari, R. P., and Oliveira, C. (2016). Cryptic species in the Neotropical fish genus (Teleostei, Characiformes). *Zool. Scripta* 45, 650–658. doi: 10.1111/zsc.12178

Menezes, N., Weitzman, S., Oyakawa, O., Lima, F., Castro, R., and Weitzman, M. (2007). *Peixes de Água Doce da Mata Atlântica: Lista Preliminar das Espécies e Comentários sobre Conservação de Peixes de água doce Neotropicais*. São Paulo: Museu de Zoologia da Universidade.

Miranda Ribeiro, A. (1918). Três gêneros e dezessete espécies novas de peixes Brasileiros. *Rev. Mus. Paul.* 10, 631–646.

Ornelas-Garcia, C. P., Dominguez-Dominguez, O., and Doadrio, I. (2008). Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actynopterigii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evol. Biol.* 8:340. doi: 10.1186/1471-2148-8-340

Paggi, L., Nascetti, G., Webb, S. C., Mattiucci, S., Cianchi, R., and Bullini, L. (1998). A new species of *Anisakis* Dujardin, 1845 (Nematoda, Anisakidae) from beaked whales (Ziphiidae) – allozyme and morphological evidence. *Sytemat. Parasitol.* 40, 161–174. doi: 10.1023/A:1006093201920

Pereira, L. H. G., Hanner, R., Foresti, F., and Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genetics* 14:20. doi: 10.1186/1471-2156-14-20

Pereira, L. H. G., Pazian, M. F., Hanner, R., Foresti, F., and Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (Characiformes: Characidae) from the Upper Paraná basin of Brazil. *Mitochondrial DNA* 22(Suppl. 1), 87–96. doi: 10.3109/19401736.2011.588213

Pons, J., Barraclough, T., Gomez-Zurita, J., Cardoso, A., Duran, D., Hazell, S., et al. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55, 595–609. doi: 10.1080/10635150600852011

Puillandre, N., Lambert, A., Brouillet, S., and Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Mol. Ecol.* 21, 1864–1877. doi: 10.1111/j.1365-294X.2011.05239.x

R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rambaut, A., Suchard, M. A., Xie, D., and Drummond, A. J. (2014). *Tracer v1.6*. Available at: http://tree.bio.ed.ac.uk/software/tracer/ [accessed November 2014].

Reis, R. E., Kullander, O. S., and Ferraris, C. J. J. (2003). *Check List of the Freshwater Fishes of South and Central America*. Porto Alegre: *EDIPUCRS*, 742.

Roxo, F. F., Albert, J. S., Silva, G. S. C., Zawadzki, C. H., Foresti, F., and Oliveira, C. (2014). Molecular phylogeny and biogeographic history of the armored neotropical catfish subfamilies hypoptopomatinae, neoplecostominae and Otothyrinae (Siluriformes: Loricariidae). *PLoS One* 9:e105564. doi: 10.1371/journal.pone.0105564

Roxo, F. F., Ochoa, L. E., Costa-Silva, G. J., and Oliveira, C. (2015). Species delimitation in *Neoplecostomus* (Siluriformes: Loricariidae) using morphologic and genetic approaches. *DNA Barcodes* 3, 110–117. doi: 10.1515/dna-2015-0014

Roxo, F. F., Zawadzki, C. H., Alexandrou, M. A., Silva, G. J. C., Chiachio, M. C., Foresti, F., et al. (2012). Evolutionary and biogeographic history of the subfamily Neoplecostominae (Siluriformes: Loricariidae). *Ecol. Evol.* 2, 2438–2449. doi: 10.1002/ece3.368

Saitoh, T., Sugita, N., Someya, S., Iwami, Y., Kobayashi, S., Kamigaichi, H., et al. (2015). DNA barcoding reveals 24 distinct lineages as cryptic bird species candidates in and around the Japanese Archipelago. *Mol. Ecol. Res.* 15, 177–186. doi: 10.1111/1755-0998.12282

Shimabukuro-Dias, C. K., Costa-Silva, G. J., Ashikaga, F. Y., Foresti, F., and Oliveira, C. (2016). Molecular identification of the fish fauna from the pantanal flood plain area in Brazil. *Mitochondrial DNA A DNA Mapp. Seq. Anal.* 28, 588–592. doi: 10.3109/24701394.2016.1149826

Silva, G. S. C., Melo, B. F., Oliveira, C., and Benine, R. C. (2013). Morphological and molecular evidence for two new species of *Tetragonopterus* (Characiformes: Characidae) from central Brazil. *J. Fish Biol.* 82, 1613–1631. doi: 10.1111/jfb.12094

Smith, M. A., Fisher, B. L., and Hebert, P. D. N. (2005). DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Philos. Trans. R. Soc. B.* 360, 1825–1834. doi: 10.1098/rstb.2005.1714

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., et al. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci. U.S.A.* 103, 12115–12120. doi: 10.1073/pnas.0605127103

Sytsma, K. J., and Schaal, B. A. (1985). Genetic variation, differentiation, and evolution in a species complex of tropical shrubs based on isozymic data. *Evolution* 39, 582–593. doi: 10.1111/j.1558-5646.1985.tb00396.x

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Lect. Math. Life Sci.* 17, 57–86.

Tavares, E. S., Gonçalves, P., Miyaki, C. Y., and Baker, A. J. (2011). DNA barcode detects high genetic structure within neotropical bird species. *PLoS One* 6:e28543. doi: 10.1371/journal.pone.0028543

Thomaz, A. T., Malabarba, L. R., Bonatto, S. L., and Knowles, L. L. (2015). 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. *J. Biogeogr.* 42, 2389–2401. doi: 10.1111/jbi.12597

Versteirt, V., Nagy, Z. T., Roelants, P., Denis, L., Breman, F. C., Damiens, D., et al. (2015). Identification of Belgian mosquito species (Diptera: Culicidae) by DNA barcoding. *Mol. Ecol. Res.* 15, 449–457. doi: 10.1111/1755-0998.12318

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., and Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1847–1857. doi: 10.1098/rstb.2005.1716

Xia, X. (2013). DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Mol. Biol. Evol.* 30, 1720–1728. doi: 10.1093/molbev/mst064

Xia, X., Xie, Z., Salemi, M., Chen, L., and Wang, Y. (2003). An index of substitution saturation and its application. *Mol. Phylogenet. Evol.* 26, 1–7. doi: 10.1016/S1055-7903(02)00326-3

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 05 March 2018
doi: 10.3389/fgene.2018.00071

Sex Chromosome Evolution and Genomic Divergence in the Fish *Hoplias malabaricus* (Characiformes, Erythrinidae)

*Alexandr Sember[1], Luiz A. C. Bertollo[2], Petr Ráb[1], Cassia F. Yano[2], Terumi Hatanaka[2], Ezequiel A. de Oliveira[2,3] and Marcelo de Bello Cioffi[2]**

[1] Laboratory of Fish Genetics, Institute of Animal Physiology and Genetics, Czech Academy of Sciences, Liběchov, Czechia, [2] Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, Brazil, [3] Secretaria de Estado de Educação de Mato Grosso (SEDUC-MT), Cuiabá, Brazil

The Erythrinidae family (Teleostei: Characiformes) is a small Neotropical fish group with a wide distribution throughout South America, where *Hoplias malabaricus* corresponds to the most widespread and cytogenetically studied taxon. This species possesses significant genetic variation, as well as huge karyotype diversity among populations, as reflected by its seven major karyotype forms (i.e., karyomorphs A-G) identified up to now. Although morphological differences in their bodies are not outstanding, *H. malabaricus* karyomorphs are easily identified by differences in 2*n*, morphology and size of chromosomes, as well as by distinct evolutionary steps of sex chromosomes development. Here, we performed comparative genomic hybridization (CGH) to analyse both the intra- and inter-genomic status in terms of repetitive DNA divergence among all but one (E) *H. malabaricus* karyomorphs. Our results indicated that they have close relationships, but with evolutionary divergences among their genomes, yielding a range of non-overlapping karyomorph-specific signals. Besides, male-specific regions were uncovered on the sex chromosomes, confirming their differential evolutionary trajectories. In conclusion, the hypothesis that *H. malabaricus* karyomorphs are result of speciation events was strengthened.

Keywords: fish cytogenetics, multiple sex chromosomes, sex-determining region, sex chromosome turnover, CGH, intraspecific variability, species complex, speciation

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Julio Cesar Pieczarka, Universidade Federal do Pará, Brazil
Daniel Pacheco Bruschi, Universidade Federal do Paraná, Brazil

****Correspondence:***
Marcelo de Bello Cioffi
mbcioffi@ufscar.br

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

Received: *29 August 2017*
Accepted: *16 February 2018*
Published: *05 March 2018*

Citation:
Sember A, Bertollo LAC, Ráb P, Yano CF, Hatanaka T, de Oliveira EA and Cioffi MdB (2018) Sex Chromosome Evolution and Genomic Divergence in the Fish Hoplias malabaricus (Characiformes, Erythrinidae). Front. Genet. 9:71. doi: 10.3389/fgene.2018.00071

INTRODUCTION

The Erythrinidae family (Teleostei: Characiformes) is a small group of Neotropical fishes with a wide distribution throughout South America. This family currently consists of three well-recognized genera—*Erythrinus* (Scopoli, 1777), *Hoplias* (Gill, 1903), and *Hoplerythrinus* (Gill, 1895) with at least 15 until now recognized species (Oyakawa, 2003; Oyakawa and Mattox, 2009). Erythrinids live in diverse habitats, from small lakes and lagoons to large rivers (Oyakawa, 2003). However, unlike the large migratory Neotropical fishes, they are usually not able to overcome obstacles such as waterfalls and large rapids, due to their sedentary lifestyle (Oyakawa, 2003). This situation may have contributed to reduced gene flow between sub-populations in the same hydrographic basin. Consequently, great genome diversity has been documented within the

Erythrinidae family, reflected in the noticeable diversity of karyotypes—particularly in the diploid number (2*n*), karyotype structure and sex chromosome systems (reviewed in Bertollo, 2007).

Hoplias malabaricus is the most widespread as well as cytogenetically investigated taxon, with analyzed populations from north to south of Brazil, Uruguay, Argentina, and Suriname. Despite its wide distribution, this taxon is characterized by low vagility, tending to constitute small populations, being able to survive under low oxygen conditions and to adapt to new environments (Rantin et al., 1992, 1993; Rios et al., 2002). Such characteristics are probably associated with the significant genetic variation and enormous karyotype diversity evidenced by the seven major karyotype forms (i.e., karyomorphs A-G Bertollo et al., 2000; Cioffi et al., 2012). However, despite rather morphological uniformity of body plan, *H. malabaricus* karyomorphs are easily distinguished by 2*n*, morphology and size of chromosomes, as well as by different evolutionary stages of distinct sex chromosome systems, indicating the occurrence of an unrecognized species diversity (Bertollo et al., 2000; Bertollo, 2007). Among the seven karyomorphs examined to date, only those of A and E do not show heteromorphic sex chromosomes. Indeed, a well-differentiated XY sex chromosome system occurs in the karyomorph B, while karyomorphs C and F possess such system in an early state of differentiation and finally, karyomorphs D and G harbor a X_1X_2Y and XY_1Y_2 multiple sex system, respectively (reviewed in Freitas et al., 2017). These findings indicate independent origins of sex chromosome systems as evidenced by whole chromosome painting (WCP) data (Cioffi et al., 2013). Altogether, *H. malabaricus* provides a suited model for evolutionary, cytotaxonomic and biodiversity analyses (for review see Cioffi et al., 2012).

The development of recent molecular methologies has allowed a qualitative improvement on chromosome researches of different biological taxa. Among them, the genomic *in situ* hybridization (GISH) and comparative genomic hybridization (CGH), originally developed for clinical studies (Kallioniemi et al., 1992), are now successfully applied for several other purposes, such as the identification of parental genomes in hybrids/allopolyploids (Bi and Bogart, 2006; Knytl et al., 2013; Symonová et al., 2013a, 2015; Doležálková et al., 2016; Majtánová et al., 2016), the detection of sex-specific content on homomorphic sex chromosomes (Ezaz et al., 2006; Altmanová et al., 2016; Rovatsos et al., 2016; Freitas et al., 2017) and the genome comparisons among related species (Valente et al., 2009; Symonová et al., 2013b; Majka et al., 2016; Carvalho et al., 2017; Moraes et al., 2017). All these (and many other) studies proved that GISH and CGH technologies, despite representing rather "rough" molecular tools, may be successful in providing clues about the genome evolution, with their resolution being based on the differential distribution of already divergent genome-specific repetitive DNA classes (Kato et al., 2005; Chester et al., 2010), as these sequences are generally highly abundant in eukaryotic genomes and display faster evolutionary rates than the single-copy ones (Charlesworth et al., 1994; Cioffi and Bertollo, 2012; López-Flores and Garrido-Ramos, 2012). Hence, in the present study, we performed analyses including CGH procedures to explore both inter- and intra-genomic divergences (within the range defined above) among *H. malabaricus* karyomorphs. Our results provided new insigths for better understanding of the ongoing processes of the karyomorph differentiations, as well as their sex chromosome systems.

Abbreviations: 2*n*, iploid chromosome number; CGH, comparative genomic hybridization; DAPI, 4′,6-diamidino-2-phenylindole; dUTP, 2′-Deoxyuridine-5′-Triphosphate; FISH, fluorescence *in situ* hybridization; gDNA, total genomic DNA; m, metacentric chromosome; NFDM, non-fat dried milk; NOR, nucleolar organizer region, PCR, polymerase chain reaction; rDNA, ribosomal DNA; SD, sex determination; sm, submetacentric chromosome; TEs, transposable elements; WCP, whole chromosomal painting.

MATERIALS AND METHODS

Individuals and Mitotic Chromosome Preparations

Analyzed representatives of *H. malabaricus* karyomorphs are given in **Table 1**. The individuals were collected under appropriate authorization of the Brazilian environmental agency ICMBIO/SISBIO (License numbers 48628-2 and 10538-1) and deposited in the fish collection of the Cytogenetic Laboratory, Departamento de Genética e Evolução, Universidade Federal de São Carlos. Mitotic chromosomes were obtained by protocols described in Bertollo et al. (2015). The experiments followed ethical and anesthesia conducts, in accordance with the Ethics Committee on Animal Experimentation of the Universidade Federal de São Carlos (Process number CEUA 1853260315).

Preparation of Probes for Comparative Genomic Hybridization (CGH)

The total genomic DNAs (gDNAs) from male and female specimens of all karyomorphs listed in **Table 1** were extracted from liver tissue by the standard phenol-chloroform-isoamylalkohol method (Sambrook and Russell, 2001). Two different experimental designs were used for this study, as

TABLE 1 | Collection sites of *Hoplias malabaricus* karyomorphs, with the respective localities and sample sizes.

***Hoplias malabaricus* karyomorphs**	**Locality**	***N***	**References**
Karyomorph A	Pântano River (SP)	08♂ 06♀	Cioffi et al., 2009
Karyomorph B	Doce River (MG)	07♂ 08♀	Cioffi et al., 2010
Karyomorph C	Bento Gomes River (MT)	11♂ 09♀	Cioffi and Bertollo, 2010
Karyomorph D	UFSCar reservoir: Monjolinho Stream (SP)	10♂ 07♀	Cioffi and Bertollo, 2010
Karyomorph F	São Francisco River (MG)	12♂ 11♀	Freitas et al., 2017
Karyomorph G	Aripuanã (MT)—Aripuanã River	12♂ 11♀	Oliveira et al., 2018

AM, Amazonas; SP, São Paulo; MT, Mato Grosso; MG, Minas Gerais States.

outlined in **Figure 1**. In the first set of experiments (inter-karyomorph genomic comparisons), the gDNA of karyomorph B male specimens was chosen as a reference due to its well-differentiated XY sex chromosomes and used for hybridization against metaphase chromosomes of the other karyomorphs (**Figure 1A**). For this purpose, male-derived gDNAs of each karyomorph A, C, D, F and G were labeled with digoxigenin-11-dUTP using DIG-nick-translation Mix (Roche, Mannheim, Germany), while male-derived gDNA of karyomorph B was labeled with biotin-16-dUTP using BIO-nick-translation Mix (Roche). For blocking the repetitive sequences in all experiments, we used unlabeled C_0t-1 DNA (i.e., fraction of genomic DNA enriched for highly and moderately repetitive sequences), prepared according to Zwick et al. (1997). Hence, the final probe cocktail for each slide was composed of 500 ng of male-derived gDNA of karyomorph B, 500 ng of male-derived DNA corresponding to one of the comparative karyomorphs, 15 μg of female-derived C_0t-1 DNA of karyomorph B and 15 μg of female-derived C_0t-1 DNA from the respective comparative karyomorph. The probe was ethanol-precipitated and the dry pellets were resuspended in hybridization buffer containing 50% formamide, 2 × SSC, 10% SDS, 10% dextran sulfate and Denhardt's buffer, pH 7.0. In the second set of experiments (**Figure 1B**) we focused on intra-karyomorph comparisons, with special emphasis on molecular composition of putative, nascent or well-differentiated sex chromosomes. In this case, male-derived gDNAs of all karyomorphs were labeled with biotin-16-dUTP and female gDNAs with digoxigenin-11-dUTP by means of nick translation as described above. The final hybridization mixture for each slide (20 μl) was composed of male- and female-derived gDNAs (500 ng each), 25 μg of female-derived C_0t-1 DNA from the respective karyomorph and the hybridization buffer described above. The chosen ratio of probe vs. C_0t-1 DNA amount was based on the experiments performed in our previous studies in fishes including erythrinids (Symonová et al., 2013a,b, 2015; Carvalho et al., 2017; Freitas et al., 2017; Moraes et al., 2017; Yano et al., 2017; Oliveira et al., 2018) and corroborated the ratio used in other related fish studies (e.g., Valente et al., 2009). According to our experiences, this ratio reflects high stringency towards repetitive DNA blocking and yet avoids the probability of improper probe dissolution in the hybridization buffer, which would otherwise cause artifacts.

Fish Used for CGH

The CGH experiments followed the methodology described in Symonová et al. (2015), with modifications. Briefly, prior to hybridization, slides were aged at 37°C for 2h, followed by an RNAse A (90 min, 37°C) and then pepsin (50 μg/ml in 10 mM HCl, 3 min, 37°C) treatments. Chromosomes were subsequently denatured in 75% formamide (pH 7.0) in 2 × SSC (74°C, 3 min), and then immediately cooled and dehydrated through 70% (cold), 85%, and 100% (RT) ethanol series. The hybridization mixture was denatured at 86°C for 6 min, cooled at 4°C (10 min) and then applied on the slides. The hybridization was performed at 37°C for 72 h. Post-hybridization washes were carried out once in 50% formamide in 2 × SSC (pH 7.0) (44°C, 10 min each) and three times in 1× SSC (44°C, 7 min each). Prior to probe detection, the slides were incubated with 3% non-fat dried milk (NFDM) in order to avoid the non-specific binding of antibodies. The hybridization signal was detected using Anti-Digoxigenin-Rhodamin (Roche) and Avidin-FITC (Sigma). Chromosomes were counterstained and mounted in antifade containing 1.5 μg/ml DAPI (Vector, Burlingame, CA, USA).

Microscopic Analyses and Image Processing

At least 30 metaphase spreads per individual were analyzed to confirm the 2*n*, karyotype structure and CGH results. Images were captured using an Olympus BX50 microscope (Olympus Corporation, Ishikawa, Japan), with CoolSNAP and the images were processed using Image Pro Plus 4.1 software (Media Cybernetics, Silver Spring, MD, USA). Chromosome morphology was classified according to Levan et al. (1964).

RESULTS

Inter-Karyomorph Genomic Relationships

In each experiment, both genome-derived probes showed rather equal binding to all chromosomes, with preferential localization in the centromeric and pericentromeric regions of most chromosomes and in terminal parts of some of them (yellow signals, i.e., combination of green and red), indicating the shared repetitive DNA content in such regions. The hybridization pattern in karyomorph A displayed stronger binding of the A-derived probe to the centomeric or telomeric regions of several chromosomes, while the B-derived probe that co-hybridized to these regions, produced signals of less intensity. Moreover, several exclusive A-specific markings appeared mostly in distal chromosomal regions (**Figures 2A–D**). Similar situation was observed also in karyomorph C (**Figures 2E–H**), where the majority of the accumulated blocks was shared by both probes, including those in the pericentromeric regions of the XY chromosomes, but several signals and especially those located in the terminal region of the long arms of the largest m pair were found to be accumulated with C-specific probe only. In karyomorph D, stronger binding of the D-derived probe was highlighted in many centromeres, in addition to some telomeric segments (**Figures 2I–L**). Remarkably, both genomic probes equally stained the heterochromatic block displayed by the neo-Y chromosome. In the karyomorph F, besides the shared binding pattern to majority of heterochromatic blocks, the F-derived probe yielded specific signals on the chromosomal pair bearing NOR-like regions (**Figures 2M–P**). In addition, the male-specific region on the nascent Y chromosome of this karyomorph displayed some affinity to B-derived male probe, despite being preferentially labeled with the F-male specific one (**Figures 2N,O**). In comparison to that, the male-specific region of karyomorph G, covering the entire short arms of the Y_1 chromosome, was stained almost exclusively by the G-derived probe, while the B-derived probe produced only faint and dispersed signals in this region (**Figures 2Q–T**). In a similar way, the G-specific probe showed predominant binding to terminally located repetitive blocks on a few other chromosomes.

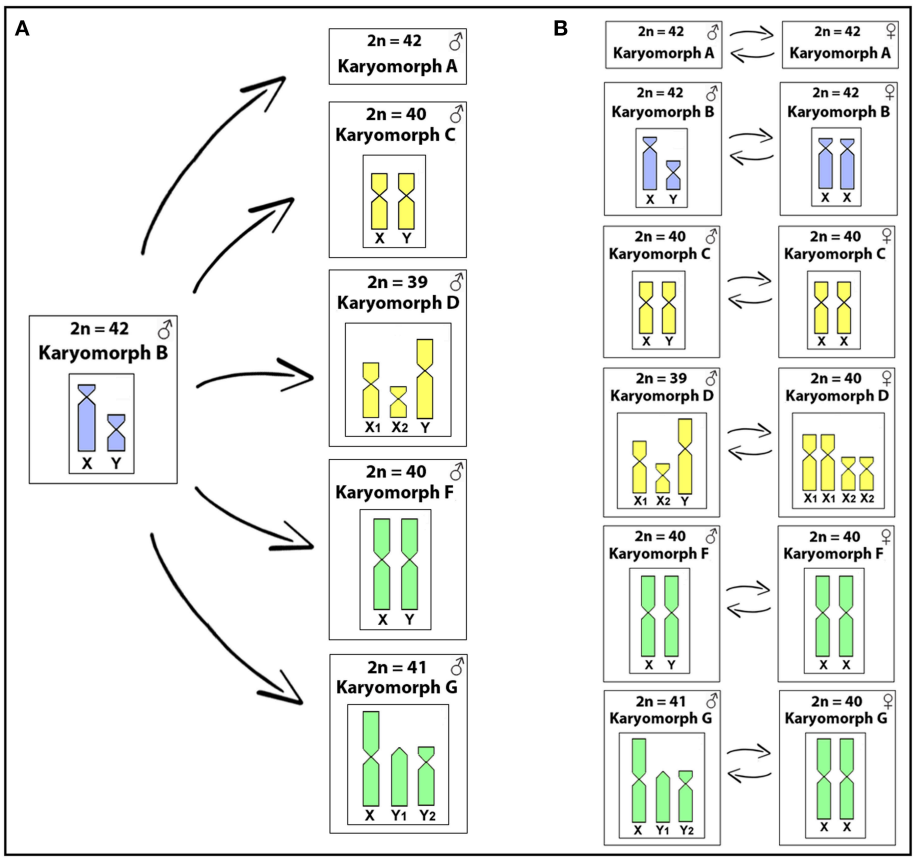


FIGURE 1 | The experimental designs used in this study. In the first one **(A)**, gDNA of karyomorph B male specimens was used for hybridization against chromosomal background of the other karyomorphs (except for karyomorph E), focusing on inter-karyomorph comparisons. In the second set of experiments **(B)** male- and female-derived gDNAs of each karyomorph under study were hybridized together, focusing on the intra-karyomorph comparisons, with special emphasis on the molecular composition of putative, nascent, or well-differentiated sex chromosomes.

Intra-Karyomorph Genomic Relationships: Detecting Male-Specific Regions

Experiments performed on female chromosome spreads of karyomorphs B, C, D, F, and G showed the absence of identifiable sex-specific segments.

Regarding karyomorph A, no exclusive male-specific regions were identified on male chromosome complement (**Figures 3A–D**). In male chromosome spreads of karyomorph B, CGH enabled to recognize male-specific region located terminally on the long arms of the Y chromosome (**Figures 3E–H**). In karyomorph C, unlike the biased accumulation of repetitive DNAs in the X pericentromeric region (Cioffi and Bertollo, 2010), a slight binding preference for the male-derived probe to the pericentromeric region of Y chromosome was evidenced (**Figures 3I–L**). Female-derived probe produced only a faint hybridization signal in such region, while both probes matched equally the large heterochromatic segment located in the pericentromeric part of the X chromosome. Accordingly, CGH procedure failed to detect any sex-specific region on male chromosomes of karyomorph D (**Figures 3M–P**). In karyomorph F, a prominent interstitial band on the metacentric Y chromosome was also enriched with male-specific sequences, although a concurrent faint hybridization signal produced by the female-derived probe was also apparent (**Figures 3Q–T**). CGH on male preparations from karyomorph G unmasked a clear male-specific region covering the short arms of Y_1 chromosome (**Figures 3U–X**). The summary of observed intra-karyomorph CGH patterns is given in **Figure 4**.

DISCUSSION

Genomic Diversity Among Karyomorphs

Sex chromosome systems in *H. malabaricus* display only male heterogamety and therefore inter-karyomorph genomic comparisons between males were supposed to be informative in indicating their interrelationships. Hence, the male-derived gDNA of reference karyomorph B was probed on chromosomes of karyomorphs A, C, D, F, and G. Even though the B-derived probe showed lower affinity to chromosomes of other karyomorphs, the overall pattern of these experiments was relatively similar. In all experiments, both genome-derived probes showed preferred accumulation to chromosome regions

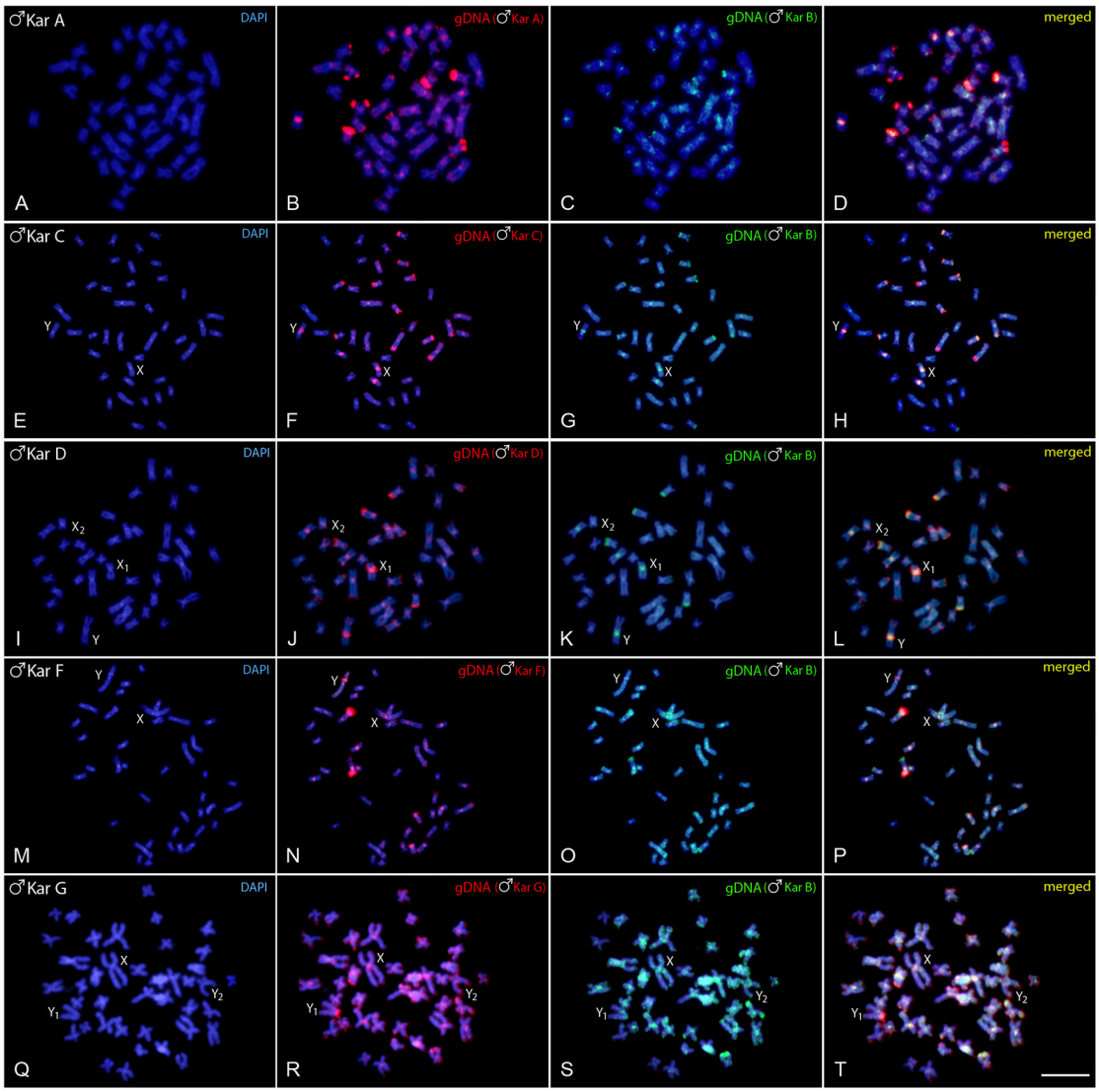


FIGURE 2 | Mitotic chromosome spreads of *Hoplias malabaricus* males after CGH—interkaryomorph comparison. Male-derived genomic probe from karyomorph B mapped against male chromosomes of karyomorph A **(A–D)**, karyomorph C **(E–H)**, karyomorph D **(I–L)**, karyomorph F **(M–P),** and karyomorph G **(Q–T)**. First column **(A,E,I,M,Q,I)**: DAPI images (blue); Second column **(B,F,J,N,R)**: hybridization pattern using the male-derived probe (red) of each analyzed karyomorph; Third column **(C,G,K,O,S)**: hybridization pattern using the male-derived probe of karyomorph B (green). Fourth column **(D,H,L,P,T)**: merged images of both genomic probes and DAPI staining. The common genomic regions of both compared karyomorphs are depicted in yellow. Bar = 10 μm.

previously identified as C-bands and C_0t-1 DNA hybridization sites (Born and Bertollo, 2000; Cioffi and Bertollo, 2010; Cioffi et al., 2010, 2011a), documenting their repetitive DNA content. However, despite less intense, hybridization signals were also apparent along the rest of chromosomal material. Our findings are in line with the general patterns observed in previous GISH/CGH-based reports (e.g., Traut and Winking, 2001; Valente et al., 2009; Koubová et al., 2014; Altmanová et al., 2016) in the sense of biased hybridization in heterochromatic regions and point to the fact that even high amount of C_0t-1 DNA is insufficient to outcompete highly repetitive (heterochromatic) regions. Given that the resolution of the CGH procedure predominantly relies on the presence of species-specific (or sex-specific) repetitive sequences, together with the evolutionary distance of the compared genomes (Kato et al., 2005; Chester et al., 2010), our overall results indicate that karyomorphs of *H. malabaricus* are closely related, but with divergences among their genomes, yielding a range of non-overlapping karyomorph-specific signals. Remarkably, the B-derived probe displayed the lowest degree of hybridization correspondence with karyomorphs D and F, suggesting the ongoing processes of sequence divergence. These findings are indicative of ongoing evolutionary processes driving the divergence and possibly also speciation within *H. malabaricus* populations, facilitated by the

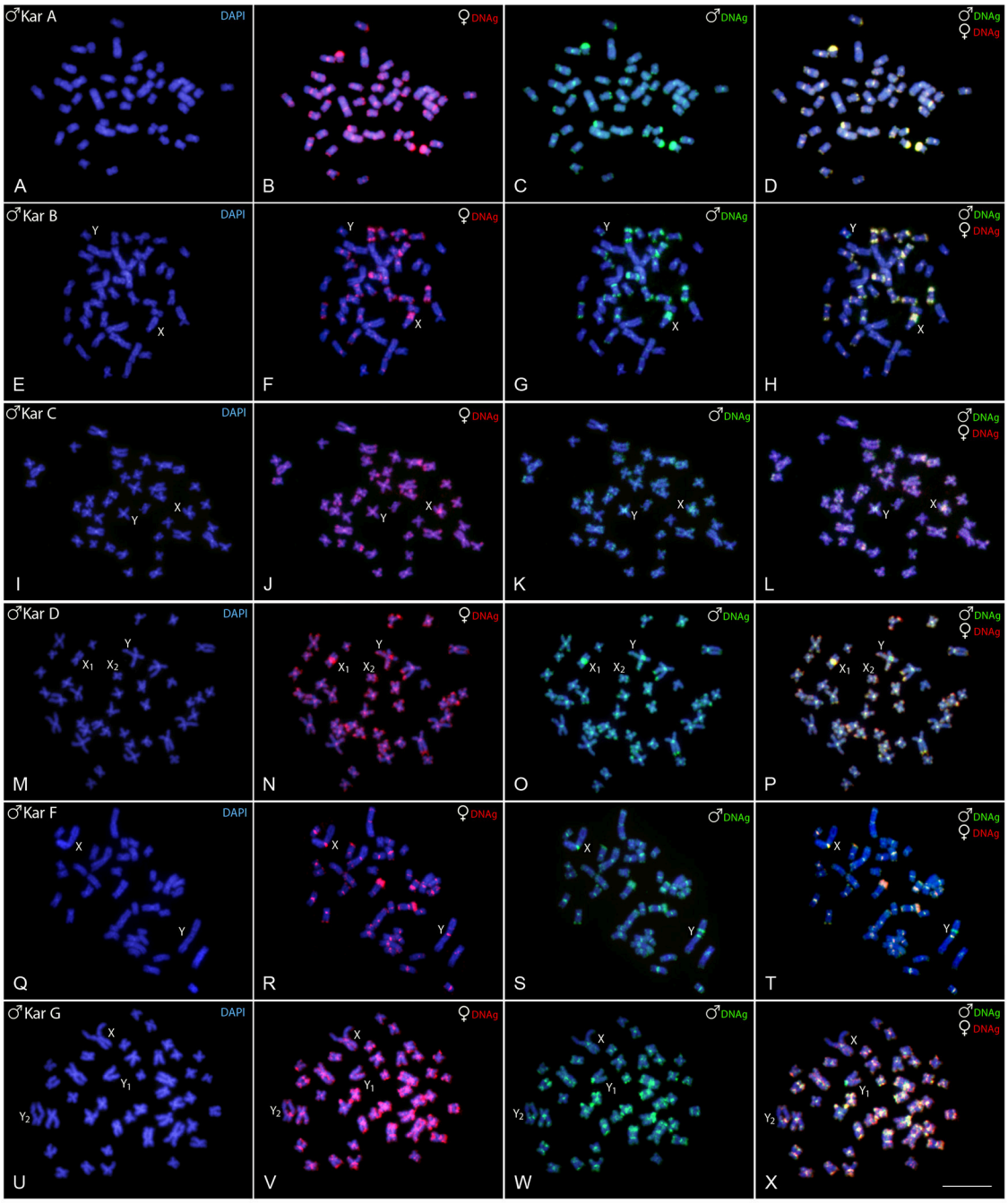


FIGURE 3 | Mitotic chromosome spreads of *Hoplias malabaricus* males after CGH—intra-karyomorph hybridizations. Male- and female-derived genomic probes hybridized together for each karyomorph. First column **(A,E,I,M,Q,U)**: DAPI images (blue); Second column **(B,F,J,N,R,V)**: hybridization pattern of the female-derived probe (red) of each analyzed karyomorph; Third column **(C,G,K,O,S,W)**: hybridization pattern of the male-derived probe (green) of the respective karyomorph. Fourth column **(D,H,L,P,T,X)**: merged images of both genomic probes and DAPI staining. The common genomic regions for male and female are depicted in yellow. Bar = 10 μm.

sedentary lifestyle of these fishes (as discussed in detail further in the text).

An array of molecular and cytogenetic methods, including DNA barcoding and phylogeographic approaches, have already led to the hypothesis that *H. malabaricus* likely represents a "species complex," with several undescribed species (Bertollo et al., 2000; Santos et al., 2009; Cioffi et al., 2012; Marques et al., 2013). Furthermore, different karyomorphs with a sympatric or syntopic occurrence were found to lack any hybrid forms, as proven by cytogenetic and RAPD (random amplified

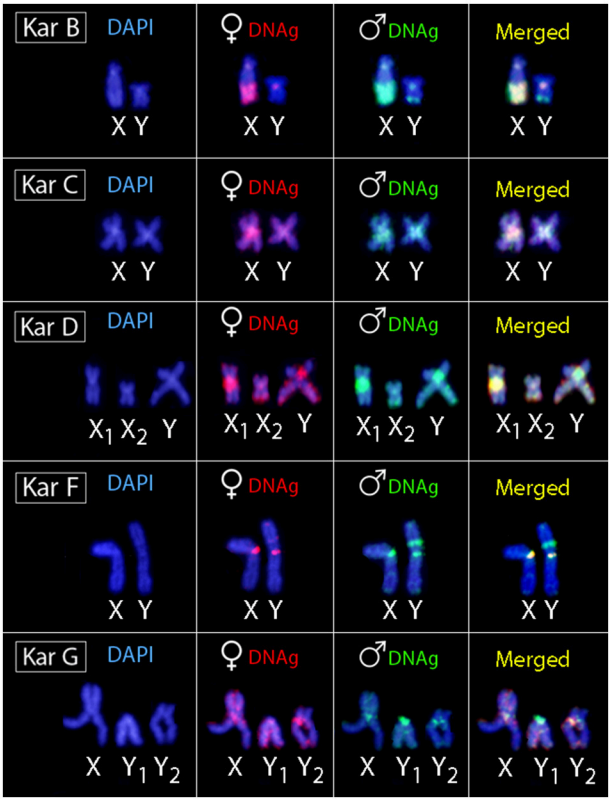


FIGURE 4 | Sex chromosomes of *H. malabaricus* karyomorphs B, C, D, F and G after CGH procedures. First column: DAPI images; Second column: hybridization pattern using the female-derived probe (red); Third column: hybridization pattern using the male-derived probe (green); Fourth column: merged images of both genomic probes and DAPI staining. Common genomic regions for male and female are depicted in yellow.

polymorphic DNA) analyses (Dergam et al., 1998, 2002; Bertollo et al., 2000), yet a sporadic case of hybridization followed by an elevation of the ploidy level was already reported (Utsunomia et al., 2014).

When compared to other members of the Erythrinidae family, similar degree of cytotaxonomic diversity can be found in *Erythrinus erythrinus* and *Hoplerythrinus unitaeniatus* groups (Bertollo, 2007; Cioffi et al., 2012; Rosa et al., 2012; Martinez et al., 2015, 2016), while—in stark contrast—the *Hoplias lacerdae* species complex exhibits highly conserved karyotype structure (Blanco et al., 2011; de Oliveira et al., 2015). It is likely that the chromosomal diversity inside some Erythrinidae species might be associated with their species-specific lifestyles. In this sense, because *H. malabaricus*, *E. erythrinus,* and *H. unitaeniatus* appear to constitute small and restricted populations, with low vagility (Blanco et al., 2011), they experience a higher rate of stochastic fixation of chromosome rearrangements and, consequently, an elevated evolutionary genome dynamism that might contribute to speciation and/or local adaptation processes (see King, 1993; Faria and Navarro, 2010 for an exhaustive discussion). It is therefore remarkable that the three *H. malabaricus* karyomorphs with more restricted geographic distribution, i.e., the B, D and G ones, possess morphogically recognizable sex chromosomes (Bertollo et al., 2000; Cioffi et al., 2012).

Intra-Karyomorph Genomic Diversity and Male-Specific Sequences

Based on data available from the last comprehensive fish karyotype overview (Arai, 2011), so far only 5% of karyologically analyzed actinopterygian fish species possess heteromorphic sex chromosomes. However, this is very likely an underestimation, with many other well-differentiated or even nascent sex chromosome systems still awaiting their discovery, especially when taking into account that homomorphic (i.e., cytologically indistinguishable or hardly detectable) sex chromosomes are thought to be frequent in fishes (Mank and Avise, 2009; Schartl et al., 2016). Within the Erythrinidae family, three different simple or multiple sex chromosome systems in advanced and/or nascent evolutionary stages have been reported among both *E. erythrinus* and *H. malabaricus* karyomorphs (reviewed in Cioffi et al., 2012), where males possess always the heterogametic sex.

Within *H. malabaricus* group, karyomorph A is characterized by $2n = 42$ for both males and females, without an apparent sex chromosome system, while karyomorph B, though also with $2n = 42$ for both sexes, exhibits a well-differentiated ♀XX/♂XY sex chromosome system. In this case, the subtelocentric large X chromosome is clearly distinguished from the small-sized metacentric Y, in addition to presence of a conspicuous heterochromatic block distally located on its long arms (Born and Bertollo, 2000; Cioffi et al., 2010). Previous repetitive DNA mapping and WCP data indicate that such sex chromosome system is likely derived from a proto-sex chromosome (the 21st pair of karyomorph A) due enrichment in several types of DNA repeats confined to only one of the homologs, namely the X chromosome in karyomorph B (Cioffi et al., 2010, 2011a,c, 2013). However, CGH procedures did not reveal any sex-specific region in karyomorph A, probably due to low level of sex-specific repetitive DNA divergence or due the small size of the sex-determining region, remaining below the detection limit of the CGH method (that ranges approximately between 2 and 3 Mbp; Schoumans et al., 2004). Theoretically, alternate mechanisms of sex determining region creation such as, e.g., epimutation-coupled recombination suppression (as eloquently discussed in Ezaz and Deakin, 2014) cannot be entirely ruled out and these, again, may have gone undetected through CGH. Finally, the possibility that the sex-determining region in this karyomorph is completely absent seems equally likely. However, the frequent changes of master sex determining genes repeatedly observed in fishes, especially in XX/XY systems (Heule et al., 2014; Martínez et al., 2014), favor the former view as the new non-recombining region is needed to be established again. In support of this view, in several another fish species, the sex determining regions might be very tiny (reviewed in Schartl et al., 2016), with the extreme case of fugu genome, *Takifugu rubripes*, where the Y-specific sex-determining gene differs from the homologous region on the X chromosome by a single non-synonymous substitution (Kamiya et al., 2012). Moreover, even in the platyfish, *Xiphophorus maculatus*, with genetically defined sex chromosomes, no visible

differences between X and Y were evidenced after CGH (Traut and Winking, 2001), similarly to what had been occasionally observed in some other animals (Koubová et al., 2014; Altmanová et al., 2016; Green et al., 2016; Gazoni et al., 2018). In yet another case, however, CGH proved to be resolute even in sex chromosomes of a very young age (Montiel et al., 2017). Finally, we cannot entirely exclude the possibility that bright fluorescence from major rDNA loci located on several chromosomes of the karyomorph A complement, most probably including the pair no. 21, could disable the detection of the hypothetical sex-determining region in its vicinity. Maybe a finer-scale approach such as BAC FISH mapping of specific candidate genes identified based on utilization of recent genome sequencing approaches and corresponding bioinformatic tools, togehter with other related "state-of-the-art" technologies may shed more light on this issue (for examples in fishes, see: Reichwald et al., 2015; Portela-Bens et al., 2017; Sutherland et al., 2017; Liu et al., 2018).

A male-specific region confined to a distal part of the long arms of Y chromosome was identified in the karyomorph B corresponding to the location of a constitutive heterochromatin block previously described (Born and Bertollo, 2000). The position of this region is noteworthy since a large NOR is found in the corresponding homologous region on the X chromosome, leading to the considerable size difference between both sex chromosomes (Born and Bertollo, 2000). Therefore, we are dealing here with an unusual situation, resembling the findings in weakly electric fish *Eigenmannia virescens* (de Almeida-Toledo and Foresti, 2001; Henning et al., 2008) or in the snake eel *Ophisurus serpens* (Salvadori et al., 2018), where the accumulation of rDNA and other repetitive DNAs occurs also on the X instead of Y chromosome. In our specific case, the sex-specific region is present on the corresponding C-positive but NOR-negative (Born and Bertollo, 2000; Cioffi et al., 2010) region on the Y chromosome. It is likely that the differential accumulation of repetitive DNA sequences might have decreased the recombination rate between the sex pair due to their delayed pairing during meiosis (Griffin et al., 2002). Alternatively, the co-amplification of the NOR region with other repetitive DNA sequences on the X chromosome can be viewed as a consequence of the whole differentiation process of the sex pair, helping to buffer the absence of functional rDNA copies on the Y chromosome. In fact, it is noteworthy that the NOR on the X chromosome is always genetically active (Born and Bertollo, 2000). The growing number of reports pointing to sex chromosome-specific NORs (see Kawai et al., 2007; Badenhorst et al., 2013; Yano et al., 2017 for references) possibly indicates that such regions might have played a more relevant role in nascent sex chromosome evolution than currently known.

In karyomorph C ($2n = 40$, for both sexes), the nascent and morphologically undifferentiated XY sex chromosomes were formerly evidenced by a small accumulation of repetitive DNAs occurring exclusively on the X chromosome (Cioffi and Bertollo, 2010). Here, it is likely that these newly emerging sex-related elements have not had the necessary evolutionary time to evolve and hence accumulated low proportion of tentatively Y-specific sequences Despite that, we cannot rule out that some differentiation in the hybridization pattern of both genomic probes in the pericentromeric region of the nascent Y chromosome is caused by a copy number variation of interspersed repetitive sequences between the sex chromosomes. Finally, it is worth mentioning that B-derived gDNA probe displayed strong binding to this region (on both X and Y) in the inter-karyomorph experiment (see **Figures 2F–H**), suggesting certain degree of shared sequences, but the extent of overlap with C-derived probe was not absolute.

Karyomorph D ($2n = 40$ in females/39 in males) is characterized by a X_1X_2Y multiple sex chromosomes system, where the neo-Y originated via a tandem fusion between the nascent Y chromosome and one autosomal homolog corresponding to the pair No. 20 of karyomorph C (Bertollo et al., 1997; Cioffi and Bertollo, 2010). Indeed, such origin was also confirmed by additional data from inter-karyomorph chromosome painting and mapping of several repetitive DNA classes (Cioffi et al., 2009, 2011b,c). Previous studies on male meiosis showed stabilized pachytene sex trivalents, as well as asynapsis in the region of presumed sequence divergences (Bertollo and Mestriner, 1998), thus pointing to a putative sex-specific region. In favor of this scenario, Rosa et al. (2009) reported noticeable alterations in location of constitutive heterochromatin and 18S rDNA sites on the neo-Y chromosome, indicating that pericentric inversions probably have also taken place in the early process of the sex-specific chromosome differentiation. However, although in karyomorph C, a slight binding preference for the male-derived probe to the pericentromeric region of Y chromosome was observed, our CGH data did not reveal any conspicuous Y-specific region in the neo-sex chromosome system of karyomorph D. In this sense, while in karyomorph D the recombination arrest and the establishment of the stable multiple sex chromosomes was most likely achieved by chromosomal rearrangements, in karyomorph C the accumulation of repetitive DNA sequences seems to have a central role in triggering the differentiation of the nascent XY sex system (Bertollo et al., 1997; Cioffi and Bertollo, 2010).

Karyomorphs E, F, and G were supposed to be closely related (Bertollo et al., 2000). Although the karyomorph E ($2n = 42$) was not sampled in this study, our results confirmed previous findings in karyomorphs F (Freitas et al., 2017) and G (Oliveira et al., 2018). More specifically, karyomorph F ($2n = 40$, for both sexes) was found to exhibit a nascent XY sex chromosome system, where the male-specific content was highlighted as a prominent interstitial heterochromatic block on the large metacentric Y chromosome, coincident with several microsatellite motifs and retrotransposons (RTEs) (Freitas et al., 2017, present study). Importantly, as the faint hybridization signal produced by the female-specific probe was allocated also within this region, we witness here a bit similar situation to that found in karyomorph C. If we admit that slightly preferred accumulation of male-exclusive sequences in pericentromeric region of the Y chromosome in males of karyomorph C might be related to the early stage of sex-determining region formation, the observed pattern in karyomorph F may reflect a later phase of similar process. At this stage, the accumulation of repetitive DNA in the Y-specific region in karyomorph F probably involves also the portion of sequences that are common for both sexes.

In contrast to karyomorph F, the sex chromosome system found in karyomorph G ($2n = 40$ in females/41 in males) is characterized by presence of XY_1Y_2 chromosomes, where the unusual acrocentric Y_1 element carries the male-specific region, enriched with several different types of repetitive DNAs including 5S rDNA (Oliveira et al., 2018), hence strengthening the view discussed above. As initially proposed by Bertollo et al. (2000) and confirmed by the recent findings (Oliveira et al., 2018), the emergence of sex chromosomes in karyomorph G proceeded through a tandem fusion involving chromosomes from two different pairs that might be tentatively assigned to specific pairs in karyomorph E—a hypothetical ancestral karyotype to both F and G karyomorphs. Importantly, while the tandem fusion was fixed in heterozygous condition in karyomorph G as only one homolog from each pair underwent this rearrangement (hence resulting in unpaired large-sized metacentric X chromosome complemented with the remaining unfused Y_1 and Y_2 chromosomes in males), in karyomorph F both homologs from the mentioned pairs gave rise to two large-sized metacentric chromosomes, the X and Y ones (Freitas et al., 2017; Oliveira et al., 2018, this study). Noteworthy, the XY_1Y_2 system of the karyomorph G differs from the X_1X_2Y neo-sex chromosomes of karyomorph D in the way that a sex-determining region is clearly detectable by CGH only in the former case, indicating a different evolutionary stage between such sex systems. All these findings in karyomorphs F and G, i.e., (i) shared homology between their sex chromosomes, pointing to a common origin through tandem fusion and (ii) lack of homology between multiple sex chromosomes found in karyomorphs D and G, are supported also by recent Zoo-FISH experiments (Oliveira et al., 2018).

In summary, our findings support trends in teleost fishes concerning the independent and repeated evolution of sex chromosomes regardless their phylogenetic relationships (Devlin and Nagahama, 2002; Woram et al., 2003; Schartl, 2004; Mank et al., 2006; Mank and Avise, 2009; Cioffi et al., 2013). It has been shown that the sex chromosome systems and/or the stage of their differentiation may differ evidently not only among closely-related species, but also among different populations of the same species (Takehana et al., 2007; Ross et al., 2009; Zhou et al., 2010; Cioffi et al., 2012; Cnaani, 2013). Such an exceptional sex chromosome variability could be possibly associated with the high plasticity and dynamics of teleost genomes (Ravi and Venkatesh, 2008), a feature usually assigned to a specific whole-genome duplication (TSGD) that occurred at the base of teleostean radiation (Hurley et al., 2007). As a consequence, duplicated redundant copies of different genes might have evolved into master sex-determining genes (Matsuda et al., 2002; Nanda et al., 2002), thus leading to emergence of distinct sex chromosomes in different evolutionary lineages (Schartl, 2004; Mank and Avise, 2009). The outstanding pace of sex chromosome turnover is, however, commonly observable also in other cold-blooded vertebrates such as amphibians and reptiles, hence a number of alternative hypotheses about the evolutionary forces standing behind this phenomenon have already been proposed (for recent reviews and in depth discussion, see Mank and Avise, 2009; Kitano and Peichel, 2012; Kikuchi and Hamaguchi, 2013; Bachtrog et al., 2014; Brykov, 2014; Pokorná et al., 2014; Pennell et al., 2015; Schartl et al., 2016). In a broader context, handful of studies have provided direct evidence that the emergence of sex chromosomes, or even the sex chromosome turnover itself, might play a major role in reproductive isolation promoting evolutionary divergences and eventually speciation (e.g., Kitano et al., 2009; Nguyen et al., 2013), which is evidently the case for *H. malabaricus*.

CONCLUSION

Our data provided additional layer of evidence about the status of the taxon *H. malabaricus* and corroborated previous studies in the conclusion that it includes taxonomically distinct species. The CGH procedures proved to be very useful in detecting the hidden biodiversity in this fish group, as they have opened novel views and widen our understanding of the ongoing processes of inter-karyomorph genome differentiation, as well as the amazing variety of sex chromosome systems inside this fish group. Besides, our approach not only uncovered the male-specific regions on the sex chromosomes, but also confirmed different trajectories of the sex chromosome evolution. Future studies using high throughput sequencing will be applied in microdissected sex chromosomes for furthering our understanding of sex determination in this species complex and its possible link with the speciation process.

AUTHOR CONTRIBUTIONS

AS: Designed the study, performed the experiments, and drafted the manuscript. CY, TH, and EdO: Performed the experiments and drafted the manuscript; PR and LB: Drafted and revised the manuscript; MC: Designed the study drafted and revised the manuscript. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

This study was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico—CNPq (Proc. nos 304992/2015-1, 401575/2016-0 and 152105/2016-6), Fundação de Amparo à Pesquisa do Estado de São Paulo- FAPESP (Proc. No 2016/21411-7) and the project EXCELLENCE CZ.02.1.01/0.0/0.0/15_003/0000460 OP RDE and RVO: 67985904.

REFERENCES

de Almeida-Toledo, L. F., and Foresti, F. (2001). Morphologically differentiated sex chromosomes in neotropical freshwater fish. *Genetica* 111, 91–100. doi: 10.1023/A:1013768104422

Altmanová, M., Rovatsos, M., Kratochvíl, L., and Johnson Pokorná, M. (2016). Minute Y chromosomes and karyotype evolution in Madagascan iguanas (Squamata: Iguania: Opluridae). *Biol. J. Linn. Soc.* 118, 618–633. doi: 10.1111/bij.12751

Arai, R. (2011). *Fish Karyotypes: A Check List, 1st Edn*. Tokyo: Springer.

Bachtrog, D., Mank, J. E., Peichel, C. L., Kirkpatrick, M., Otto, S. P., Ashman, T. L., et al. (2014). Sex determination: why so many ways *of* doing it? *PLoS Biol.* 12:e1001899. doi: 10.1371/journal.pbio.1001899

Badenhorst, D., Stanyon, R., Engstrom, T., and Valenzuela, N. (2013). A ZZ/ZW microchromosome system in the spiny softshell turtle, *Apalone spinifera*, reveals an intriguing sex chromosome conservation in Trionychidae. *Chromosome Res.* 21, 137–147. doi: 10.1007/s10577-013-9343-2

Bertollo, L. A. C. (2007). "Chromosome evolution in the Neotropical Erythrinidae fish family: an overview," in *Fish Cytogenetics*, eds E. Pisano, C. Ozouf-Costaz, F. Foresti, and B. G. Kapoor (Enfield, NH: Science Publishers), 195–211.

Bertollo, L. A. C., and Mestriner, C. A. (1998). The X_1X_2Y sex chromosome system in the fish *Hoplias malabaricus* II Meiotic analyses. *Chromosome Res.* 6, 141–147. doi: 10.1023/A:1009243114124

Bertollo, L. A., Born, G. G., Dergam, J. A., Fenocchio, A. S., and Moreira-Filho, O. (2000). A biodiversity approach in the neotropical Erythrinidae fish, *Hoplias malabaricus*. Karyotypic survey, geographic distribution of cytotypes and cytotaxonomic considerations. *Chromosome Res.* 8, 603–613. doi: 10.1023/A:1009233907558

Bertollo, L. A. C., Fontes, M. S., Fenocchio, A. S., and Cano, J. (1997). The X_1X_2Y sex chromosome system in the fish *Hoplias malabaricus*. I. G-, C- and chromosome replication banding. *Chromosome Res.* 5, 493–499. doi: 10.1023/A:1018477232354

Bertollo, L. A. C., Moreira-Filho, O., and Cioffi, M. B. (2015). "Direct chromosome preparations from freshwater teleost fishes," in *Fish Techniques, Ray-Fin Fishes and Chondrichthyans*, eds C. Ozouf-Costaz, E. Pisano, F. Foresti, and L. F. de Almeida Toledo (Boca Ranton, FL: CRC Press), 21–26.

Bi, K., and Bogart, J. P. (2006). Identification of intergenomic recombinations in unisexual salamanders of the genus *Ambystoma* by genomic *in situ* hybridization (GISH). *Cytogenet. Genome Res.* 112, 307–312. doi: 10.1159/000089885

Blanco, D. R., Lui, R. L., Vicari, M. R., Bertollo, L. A., and Moreira-Filho, O. (2011). Comparative cytogenetics of giant trahiras *Hoplias aimara* and *H. intermedius (Characiformes, Erythrinidae)*: chromosomal characteristics of minor and major ribosomal DNA and cross-species repetitive centromeric sequences mapping differ among morphologically identical karyotypes. *Cytogenet. Genome Res.* 132, 71–78. doi: 10.1159/000320923

Born, G. G., and Bertollo, L. A.(2000). An XX/XY sex chromosome system in a fish species, *Hoplias malabaricus*, with a polymorphic NOR-bearing X chromosome. *Chromosome Res.* 8, 111–118. doi: 10.1023/A:1009238402051

Brykov, V. A. (2014). Mechanisms of sex determination in fish: evolutionary and practical aspects. *Russ. J. Mar. Biol.* 40, 407–417. doi: 10.1134/S1063074014060145

Carvalho, P. C., de Oliveira, E. A., Bertollo, L. A. C., Yano, C. F., Oliveira, C., Decru, E., et al. (2017). First chromosomal analysis in Hepsetidae (Actinopterygii, Characiformes): insights into relationship between African and Neotropical fish groups. *Front. Genet.* 8:203. doi: 10.3389/fgene.2017.00203

Charlesworth, B., Sniegowski, P., and Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371, 215–220. doi: 10.1038/371215a0

Chester, M., Leitch, A. R., Soltis, P. S., and Soltis, D. E. (2010). Review of the application of modern cytogenetic methods (FISH/GISH) to the study of reticulation (polyploidy/hybridisation). *Genes* 1, 166–192. doi: 10.3390/genes1020166

Cioffi, M. B., and Bertollo, L. A., (2010). Initial steps in XY chromosome differentiation in *Hoplias malabaricus* and the origin of an X_1X_2Y sex chromosome system in this fish group. *Heredity* 105, 554–561. doi: 10.1038/hdy.2010.18

Cioffi, M. B., and Bertollo, L. A. (2012). Chromosomal distribution and evolution of repetitive DNAs in fish. *Genome Dyn.* 7, 197–221. doi: 10.1159/000337950

Cioffi, M. B., Liehr, T., Trifonov, V., Molina, W. F., and Bertollo, L. A. C. (2013). Independent sex chromosome evolution in lower vertebrates: a molecular cytogenetic overview in the erythrinidae fish family. *Cytogenet. Genome Res.* 141, 186–194. doi: 10.1159/000354039

Cioffi, M. B., Martins, C., and Bertollo, L. A. C. (2009). Comparative chromosome mapping of repetitive sequences. Implications for genomic evolution in the fish, *Hoplias malabaricus*. *BMC Genet.* 10:34. doi: 10.1186/1471-2156-10-34

Cioffi, M. B., Martins, C., Vicari, M. R., Rebordinos, L., and Bertollo, L. A. (2010). Differentiation of the XY sex chromosomes in the fish *Hoplias malabaricus* (Characiformes, Erythrinidae). Unusual accumulation of repetitive sequences on the X chromosome. *Sex. Dev.* 4, 176–185. doi: 10.1159/000309726

Cioffi, M. B., Molina, W. F., Artoni, R. F., and Bertollo, L. A. (2012). "Chromosomes as tools for discovering biodiversity – the case of Erythrinidae fish family," in *Recent Trends in Cytogenetic Studies – Methodologies and Applications*, ed P. Tirunilai (Rijeka: InTech Publisher), 125–146.

Cioffi, M. B., Molina, W. F., Moreira-Filho, O., and Bertollo, L. A. (2011a). Chromosomal distribution of repetitive DNA sequences highlights the independent differentiation of multiple sex chromosomes in two closely related fish species. *Cytogenet. Genome Res.* 134, 295–302. doi: 10.1159/000329481

Cioffi, M. B., Sánchez, A., Marchal, J. A., Kosyakova, N., Liehr, T., Trifonov, V., et al. (2011b). Cross-species chromosome painting tracks the independent origin of multiple sex chromosomes in two cofamiliar Erythrinidae fishes. *BMC Evol. Biol.* 11:186. doi: 10.1186/1471-2148-11-186

Cioffi, M. B., Sánchez, A., Marchal, J. A., Kosyakova, N., Liehr, T., Trifonov, V., et al. (2011c). Whole chromosome painting reveals independent origin of sex chromosomes in closely related forms of a fish species. *Genetica* 139, 1065–1072. doi: 10.1007/s10709-011-9610-0

Cnaani, A. (2013). The tilapias' chromosomes influencing sex determination. *Cytogenet. Genome Res.* 141, 195–205. doi: 10.1159/000355304

Dergam, J. A., Paiva, S. R., Schaefer, C. E., Godinho, A. L., and Vieira, F. (2002). Phylogeography and RAPD-PCR variation in *Hoplias malabaricus* (Bloch, 1974) (Pisces, Teleostei) in southeastern Brazil. *Genet. Mol. Biol.* 25, 379–387. doi: 10.1590/S1415-47572002000400005

Dergam, J. A., Suzuki, H. I., Shibatta, O. A., Duboc, L. F., Júlio, H. F. Jr., Giuliano-Caetano, L. et al. (1998). Molecular biogeography of the neotropical fish *Hoplias malabaricus* (Erythrinidae, Characiformes) in the Iguaçu, Tibagi, and Paraná rivers. *Genet. Mol. Biol.* 21, 493–496. doi: 10.1590/S1415-47571998000400015

Devlin, R. H., and Nagahama, Y. (2002). Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture* 208, 191–364. doi: 10.1016/S0044-8486(02)00057-1

Doležálková, M., Sember, A., Marec, F., Ráb, P., Plötner, J., and Choleva, L. (2016). Is premeiotic genome elimination an exclusive mechanism for hemiclonal reproduction in hybrid males of the genus *Pelophylax*? *BMC Genet.* 17:100. doi: 10.1186/s12863-016-0408-z

Ezaz, T., and Deakin, J. E. (2014). Repetitive sequence and sex chromosome evolution in vertebrates. *Adv. Evol. Biol.* 2014, 1–9. doi: 10.1155/2014/104683

Ezaz, T., Valenzuela, N., Grützner, F., Miura, I., Georges, A., Burke, R. L., et al. (2006). An XX/XY sex microchromosome system in a freshwater turtle, *Chelodina longicollis* (Testudines: Chelidae) with genetic sex determination. *Chromosome Res.* 14, 139–150. doi: 10.1007/s10577-006-1029-6

Faria, R., and Navarro, A. (2010). Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends Ecol. Evol.* 25, 660–669. doi: 10.1016/j.tree.2010.07.008

Freitas, N. L., Al-Rikabi, A. B. H., Bertollo, L. A. C., Ezaz, T., Yano, C. F., Oliveira, E. A., et al. (2017). Early stages of XY sex chromosomes differentiation in the fish *Hoplias malabaricus* (Characiformes, Erythrinidae) revealed by DNA repeats accumulation. *Curr. Genomics* 19, 216–226. doi: 10.2174/1389202918666170711160528

Gazoni, T., Haddad, C. F. B., Narimatsu, H., Cabral-de-Mello, D. C., Lyra, M. L., and Parise-Maltempi P. P. (2018). More sex chromosomes than autosomes in the Amazonian frog *Leptodactylus pentadactylus*. *Chromosoma*. doi: 10.1007/s00412-018-0663-z. [Epub ahead of print].

Green, J. E., Dalíková, M., Sahara, K., Marec, F., and Akam, M. (2016). XX/XY system of sex determination in the geophilomorph centipede *Strigamia maritima*. *PLoS ONE* 11:e0150292. doi: 10.1371/journal.pone.0150292

Griffin, D. K., Harvey, S. C., Campos-Ramos, R., Ayling, L.-J., Bromage, N. R., Masabanda, J. S., et al. (2002). Early origins of the X and Y chromosome: lessons from tilapia. *Cytogenet. Genome Res.* 99, 157–163. doi: 10.1159/000071588

Henning, F., Trifonov, V., Ferguson-Smith, M. A., and de Almeida-Toledo, L. F. (2008). Non-homologous sex chromosomes in two species of the genus *Eigenmannia* (Teleostei: Gymnotiformes). *Cytogenet Genome Res.* 121, 55–58. doi: 10.1159/000124382

Heule, C., Salzburger, W., and Böhne, A. (2014). Genetics of sexual development: an evolutionary playground for fish. *Genetics* 196, 579–591. doi: 10.1534/genetics.114.161158

Hurley, I. A., Mueller, R. L., Dunn, K. A., Schmidt, E. J., Friedman, M., Ho, R. K., et al. (2007). A new time-scale for ray-finned fish evolution. *Proc. R. Soc. B Biol. Sci*. 274, 489–498. doi: 10.1098/rspb.2006.3749

Kallioniemi, A., Kallioniemi, O. P., Sudar, D., Rutovitz, D., Gray, J. W., Waldman, F., et al. (1992). Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258, 818–821. doi: 10.1126/science.1359641

Kamiya, T., Kai, W., Tasumi, S., Oka, A., Matsunaga, T., Mizuno, N., et al. (2012). A trans-species missense SNP in *Amhr2* is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (Fugu). *PLoS Genet*. 8:e1002798. doi: 10.1371/journal.pgen.1002798

Kato, A., Vega, J. M., Han, F., Lamb, J. C., and Bircher, J. A. (2005). Advances in plant chromosome identification and cytogenetic techniques. *Curr. Opin. Plant. Biol*. 8, 148–154. doi: 10.1016/j.pbi.2005.01.014

Kawai, A., Nishida-Umehara, C., Ishijima, J., Tsuda, Y., Ota, H., and Matsuda, Y. (2007). Different origins of bird and reptile sex chromosomes inferred from comparative mapping of chicken Z-linked genes. *Cytogenet. Genome Res*. 117, 92–102. doi: 10.1159/000103169

Kikuchi, K., and Hamaguchi, S. (2013). Novel sex-determining genes in fish and sex chromosome evolution. *Dev. Dyn*. 242, 339–353. doi: 10.1002/dvdy.23927

King, M. (1993). *Species Evolution: The Role of Chromosome Change*. Cambridge, GB: University Press.

Kitano, J., and Peichel, C. L. (2012). Turnover of sex chromosomes and speciation in fishes. *Environ. Biol. Fishes*. 94, 549–558. doi: 10.1007/s10641-011-9853-8

Kitano, J., Ross, J. A., Mori, S., Kume, M., Jones, F. C., Chan, Y. F., et al. (2009). A role for a neo-sex chromosome in stickleback speciation. *Nature* 461, 1079–1083. doi: 10.1038/nature08441

Knytl, M., Kalous, L., Symonová, R., Rylková, K., and Ráb, P. (2013). Chromosome studies of European cyprinid fishes: cross-species painting reveals natural allotetraploid origin of a *Carassius* female with 206 chromosomes. *Cytogenet. Genome Res*. 139, 276–283. doi: 10.1159/000350689

Koubová, M., Pokorná, M. J., Rovatsos, M., Farkacová, K., Altmanová, M., and Kratochvíl, L. (2014). Sex determination in Madagascar geckos of the genus *Paroedura* (Squamata: Gekkonidae): are differentiated sex chromosomes indeed so evolutionary stable? *Chromosome Res*. 22, 441–452. doi: 10.1007/s10577-014-9430-z

Levan, A., Fredga, K., and Sandberg, A. A. (1964). Nomenclature for centromeric position on chromosomes. *Hereditas* 52, 201–220. doi: 10.1111/j.1601-5223.1964.tb01953.x

Liu, H., Pang, M., Yu, X., Zhou, Y., Tong, J., and Fu, B. (2018). Sex-specific markers developed by next-generation sequencing confirmed an XX/XY sex determination system in bighead carp (*Hypophthalmichehys nobilis*) and silver carp (*Hypophthalmichthys molitrix*). *DNA Res.* doi: 10.1093/dnares/dsx054. [Epub ahead of print].

López-Flores, I., and Garrido-Ramos, M. A. (2012). The repetitive DNA content of eukaryotic genomes. *Genome Dyn.* 7, 1–28. doi: 10.1159/000337118

Majka, J., Majka, M., Kwiatek, M., and Wiśniewska, H. (2016). Similarities and differences in the nuclear genome organization within Pooideae species revealed by comparative genomic *in situ* hybridization (GISH). *J. Appl. Genet*. 58, 151–161. doi: 10.1007/s13353-016-0369-y

Majtánová, Z., Choleva, L., Symonová, R., Ráb, P., Kotusz, J., Pekárik, L., et al. (2016). Asexual reproduction does not apparently increase the rate of chromosomal evolution: karyotype stability in diploid and triploid clonal hybrid fish (*Cobitis*, Cypriniformes, Teleostei). *PLoS ONE* 11:e0146872. doi: 10.1371/journal.pone.0146872

Mank, J. E., and Avise, J. C. (2009). Evolutionary diversity and turn-over of sex determination in teleost fishes. *Sex Dev*. 3, 60–67. doi: 10.1159/000223071

Mank, J. E., Promislow, D. E. L., and Avise, J. C. (2006). Evolution of alternative sex-determining mechanisms in teleost fishes. *Biol. J. Linn. Soc*. 87, 83–93. doi: 10.1111/j.1095-8312.2006.00558.x

Marques, D. F., Santos, F. A., da Silva, S. S., Sampaio, I., and Rodrigues, L. R. R. (2013). Cytogenetic and DNA barcoding reveals high divergence within the trahira, *Hoplias malabaricus* (Characiformes: Erythrinidae) from the lower Amazon River. *Neotrop. Ichthyol.* 11, 459–466. doi: 10.1590/S1679-62252013000200015

Martinez, J. F., Lui, R. L., Traldi, J. B., Blanco, D. R., and Moreira-Filho, O. (2016). Comparative cytogenetics of *Hoplerythrinus unitaeniatus* (Agassiz, 1829) (Characiformes, Erythrinidae) species complex from different brazilian hydrographic basins. *Cytogenet. Genome Res*. 149, 191–200. doi: 10.1159/000448153

Martinez, J. F., Lui, R. L., Traldi, J. B., Blanco, D. R., and Moreira-Filho, O. (2015). Occurrence of natural hybrids among sympatric karyomorphs in *Hoplerythrinus unitaeniatus* (Characiformes, Erythrinidae). *Zebrafish* 12, 281–287. doi: 10.1089/zeb.2015.1083

Martínez, P., Viñas, A. M., Sánchez, L., Díaz, N., Ribas, L., and Piferrer, F. (2014). Genetic architecture of sex determination in fish: applications to sex ratio control in aquaculture. *Front. Genet*. 5:340. doi: 10.3389/fgene.2014.00340

Matsuda, M., Nagahama, Y., Shinomiya, A., Sato, T., Matsuda, C., Kobayashi, T., et al. (2002). *DMY* is a Y specific DM-domain gene required for male development in the medaka fish. *Nature* 417, 559–563. doi: 10.1038/nature751

Montiel, E. E., Badenhorst, D., Tamplin, J., Burke, R. L., and Velanzuela, N. (2017). Discovery of the youngest sex chromosomes reveals first case of convergent co-option of ancestral autosomes in turtles. *Chromosoma* 126, 105–113. doi: 10.1007/s00412-016-0576-7

Moraes, R. L. R., Bertollo, L. A. C., Marinho, M. M. F., Yano, C. F., Hatanaka, T., Barby, F. F., et al. (2017). Evolutionary relationships and cytotaxonomy considerations in the genus *Pyrrhulina* (Characiformes, Lebiasinidae). *Zebrafish* 14, 536–546. doi: 10.1089/zeb.2017.1465

Nanda, I., Kondo, M., Hornung, U., Asakawa, S., Winkler, C., Shimizu, A., et al. (2002). A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. *Proc. Natl. Acad. Sci. U.S.A*. 99, 11778–11783. doi: 10.1073/pnas.182314699

Nguyen, P., Sýkorová, M., Šíchová, J., Kuta, V., and Dalíková, M., Capková, Frydrychová, R., et al. (2013). Neo-sex chromosomes and adaptive potential in tortricid pests. *Proc. Natl. Acad. Sci. U.S.A*. 110, 6931–6936. doi: 10.1073/pnas.1220372110

de Oliveira, E. A., Bertollo, L. A. C., Yano, C. F., Liehr, T., and Cioffi, M. B. (2015). Comparative cytogenetics in the genus *Hoplias* (Characiformes, Erythrinidae) highlights contrasting karyotype evolution among congeneric species. *Mol. Cytogenet*. 8:56. doi: 10.1186/s13039-015-0161-4

Oliveira, E. A., Sember, A., Bertollo, L. A. C., Yano, C. F., Ezaz, T., Moreira-Filho, O., et al. (2018). Tracking the evolutionary pathway of sex chromosomes among fishes: characterizing the unique XX/XY_1Y_2 system in *Hoplias malabaricus* (Teleostei, Characiformes). *Chromosoma* 127, 115–128. doi: 10.1007/s00412-017-0648-3

Oyakawa, O. T. (2003). "Family Erythrinidae," in *Check List of the freshwater fishes of South and Central America,* eds R. E. Reis, S. O. Kullander, and C. J., Ferraris Jr (Porto Alegre: Edipucrs), 238–240.

Oyakawa, O. T., and Mattox, G. M. T. (2009). Revision of the Neotropical trahiras of the *Hoplias lacerdae* species-group (Ostariophysi: Characiformes: Erythrinidae) with descriptions of two new species. *Neotrop. Ichthyol*. 7, 117–140. doi: 10.1590/S1679-62252009000200001

Pennell, M. W., Kirkpatrick, M., Otto, S. P., Vamosi, J. C., Peichel, C. L., Valenzuela, N., et al. (2015). Y fuse? Sex chromosome fusions in fishes and reptiles. *PLoS Genet*. 11:e1005237. doi: 10.1371/journal.pgen.1005237

Pokorná, M., Altmanová, M., and Kratochvíl, L. (2014). Multiple sex chromosomes in the light of female meiotic drive in amniote vertebrates. *Chromosome Res*. 22, 35–44. doi: 10.1007/s10577-014-9403-2

Portela-Bens, S., Merlo, M. A., Rodríguez, M. E., Cross, I., Manchado, M., Kosyakova, N., et al. (2017). Integrated gene mapping and synteny studies give insights into the evolution of a sex protochromosome in *Solea senegalensis*. *Chromosoma* 126, 261–277. doi: 10.1007/s00412-016-0589-2

Rantin, F. T., Glass, M. L., Kalinin, A. L., Verzola, R. M. M., and Fernandes, M. N. (1993). Cardio-respiratory responses in two ecologically distinct erythrinids (*Hoplias malabaricus* and *Hoplias lacerdae*) exposed to graded environmental hypoxia. *Environ. Biol. Fish*. 36, 93–97. doi: 10.1007/BF00005983

Rantin, F. T., Kalinin, A. L., Glass, M. L., and Fernandes, M. N. (1992). Respiratory responses to hypoxia in relation to mode of life of two erythrinid species (*Hoplias malabaricus* and *Hoplias lacerdae*). *J. Fish. Biol*. 41, 805–812. doi: 10.1111/j.1095-8649.1992.tb02708.x

Ravi, V., and Venkatesh, B. (2008). Rapidly evolving fish genomes and teleost diversity. *Curr. Opin. Genet. Dev*. 18:544–550. doi: 10.1016/j.gde.2008.11.001

Reichwald, K., Petzold, A., Koch, P., Downie, B. R., Hartmann, N., Pietsch, S., et al. (2015). Insights into sex chromosome evolution and aging from the genome of a short-lived fish. *Cell* 163, 1527–1538. doi: 10.1016/j.cell.2015.10.071

Rios, F. S., Kalinin, A. L., and Rantin, F. T. (2002). The effects of long-term food deprivation on respiration and hematology of the neotropical fish *Hoplias malabaricus*. *J. Fish. Biol.* 61, 85–95. doi: 10.1111/j.1095-8649.2002.tb01738.x

Rosa, R., Laforga Vanzela, A. L., Rubert, M., Martins-Santos, I. C., and Giuliano-Caetano, L. (2009). Differentiation of Y chromosome in the $X_1X_1X_2X_2/X_1X_2Y$ sex chromosome system of *Hoplias malabaricus* (Characiformes, Erythrinidae). *Cytogenet. Genome Res.* 127, 54–60. doi: 10.1159/000269736

Rosa, R., Rubert, M., Martins-Santos, I., and Giuliano-Caetano, L. (2012). Evolutionary trends in *Hoplerythrinus unitaeniatus* (Agassiz 1829) (Characiformes, Erythrinidae). *Rev. Fish Biol. Fish.* 22, 467–475. doi: 10.1007/s11160-011-9237-3

Ross, J. A., Urton, J. R., Boland, J., Shapiro, M. D., and Peichel, C. L. (2009). Turnover of sex chromosomes in the stickleback fishes (Gasterosteidae). *PLoS Genet.* 5:e1000391. doi: 10.1371/journal.pgen.1000391

Rovatsos, M., Johnson Pokorná, M., Altmanová, M., and Kratochvíl, L. (2016). Mixed-up sex chromosomes: identification of sex chromosomes in the $X_1X_1X_2X_2/X_1X_2Y$ system of the legless lizards of the genus *Lialis* (Squamata: Gekkota: Pygopodidae). *Cytogenet. Genome Res.* 149, 282–289. doi: 10.1159/000450734

Salvadori, S., Deiana, A. M., Deidda, F., Lobina, C., Mulas, A., and Coluccia, E. (2018). XX/XY sex chromosome system and chromosome markers in the snake eel *Ophisurus serpens* (Anguilliformes: Ophichtidae). *Mar. Biol. Res.* 14, 158–164. doi: 10.1080/17451000.2017.1406665

Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual, 3rd Edn*. New York, NY: Cold Spring Harbor Laboratory Press.

Santos, U., Völcker, C. M., Belei, F. A., Cioffi, M. B., Bertollo, L. A., Paiva, S. R., et al. (2009). Molecular and karyotypic phylogeography in the Neotropical *Hoplias malabaricus* (Erythrinidae) fish in eastern. *Brazil. J. Fish Biol.* 75, 2326–2343. doi: 10.1111/j.1095-8649.2009.02489.x

Schartl, M. (2004). Sex chromosome evolution in non-mammalian vertebrates. *Curr. Opin. Genet. Dev.* 14, 634–641. doi: 10.1016/j.gde.2004.09.005

Schartl, M., Schmid, M., and Nanda, I. (2016). Dynamics of vertebrate sex chromosome evolution: from equal size to giants and dwarfs. *Chromosoma* 125, 553–571. doi: 10.1007/s00412-015-0569-y

Schoumans, J., Nielsen, K., Jeppesen, I., Anderlid, B. M., Blennow, E., Brøndum-Nielsen, K., et al. (2004). A comparison of different metaphase CGH methods for the detection of cryptic chromosome aberrations of defined size. *Eur. J. Hum. Genet.* 12, 447–454. doi: 10.1038/sj.ejhg.5201175

Sutherland, B. J. G., Rico, C., Audet, C., and Bernatchez, L. (2017). Sex chromosome evolution, heterochiasmy, and physiological QTL in the salmonid brook charr *Salvelinus fontinalis*. *G3* 7, 2749–2762. doi: 10.1534/g3.117.040915

Symonová, R., Flajšhans, M., Sember, A., Havelka, M., Gela, D., Korínková, T., et al. (2013a). Molecular cytogenetics in artificial hybrid and highly polyploid sturgeons: an evolutionary story narrated by repetitive sequences. *Cytogenet. Genome Res.* 141, 153–162. doi: 10.1159/000354882

Symonová, R., Majtánová, Z., Sember, A., Staaks, G. B., Bohlen, J., Freyhof, J., et al. (2013b). Genome differentiation in a species pair of coregonine fishes: an extremely rapid speciation driven by stress-activated retrotransposons mediating extensive ribosomal DNA multiplications. *BMC Evol. Biol.* 13:42. doi: 10.1186/1471-2148-13-42

Symonová, R., Sember, A., Majtánová, Z., and Ráb, P. (2015). "Characterization of fish genomes by GISH and CGH," in *Fish techniques, Ray-Fin Fishes and Chondrichthyans*, eds C. Ozouf-Costaz, E. Pisano, F. Foresti, and L. F. de Almeida Toledo (Boca Ranton, FL: CRC Press), 118–131.

Takehana, Y., Naruse, K., Hamaguchi, S., and Sakaizumi, M. (2007). Evolution of ZZ/ZW and XX/XY sex-determination systems in the closely related medaka species, *Oryzias hubbsi* and *O. dancena*. *Chromosoma* 116, 463–470. doi: 10.1007/s00412-007-0110-z

Traut, W., and Winking, H. (2001). Meiotic chromosomes and stages of sex chromosome evolution in fish: zebrafish, platyfish and guppy. *Chromosome Res.* 9, 659–672. doi: 10.1023/A:1012956324417

Utsunomia, R., Pansonato-Alves, J. C., Paiva, L. R. S., Costa Silva, G. J., Oliveira, C., Bertollo, L. A. C., et al. (2014). Genetic differentiation among distinct karyomorphs of the wolf fish *Hoplias malabaricus* species complex (Characiformes, Erythrinidae) and report of unusual hybridization with natural triploidy. *J. Fish Biol.* 85, 1682–1692. doi: 10.1111/jfb.12526

Valente, T. G., Schneider, C. H., Gross, M. C., Feldberg, E., and Martins, C. (2009). Comparative cytogenetics of cichlid fishes through genomic *in-situ* hybridization (GISH) with emphasis on *Oreochromis niloticus*. *Chromosome Res.* 17, 791–799. doi: 10.1007/s10577-009-9067-5

Woram, R. A., Gharbi, K., Sakamoto, T., Hoyheim, B., Holm, L. E.;, Naish, K., et al. (2003). Comparative genome analysis of the primary sex-determining locus in salmonid fishes. *Genome Res.* 13, 272–280. doi: 10.1101/gr.578503

Yano, C. F., Bertollo, L. A., Ezaz, T., Trifonov, V., Sember, A., Liehr, T., et al. (2017). Highly conserved Z and molecularly diverged W chromosomes in the fish genus *Triportheus* (Characiformes, Triportheidae). *Heredity* 118, 276–283. doi: 10.1038/hdy.2016.83

Zhou, Q., Braasch, I., Froschauer, A., and Böhne, A. (2010). A novel marker for the platyfish (*Xiphophorus maculatus*) W chromosome is derived from a *Polinton* transposon. *J. Genet. Genomics* 37, 181–188. doi: 10.1016/S1673-8527(09)60036-9

Zwick, M. S., Hanson, R. E., McKnight, T. D., Islam-Faridi, M. N., Stelly, D. M., Wing, R. A., et al. (1997). A rapid procedure for the isolation of C_0t-1 DNA from plants. *Genome* 40, 138–142. doi: 10.1139/g97-020

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

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

frontiers in Genetics

METHODS
published: 09 March 2018
doi: 10.3389/fgene.2018.00073

High-Throughput Sequencing Strategy for Microsatellite Genotyping Using Neotropical Fish as a Model

*Juliana S. M. Pimentel[1], Anderson O. Carmo[1], Izinara C. Rosse[1], Ana P. V. Martins[1], Sandra Ludwig[2], Susanne Facchin[1], Adriana H. Pereira[1], Pedro F. P. Brandão-Dias[1], Nazaré L. Abreu[1] and Evanguedes Kalapothakis[1]**

[1] *Laboratory of Biotechnology and Molecular Markers, Department of General Biology, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Brazil,* [2] *Department of Zoology, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Brazil*

Genetic diversity and population studies are essential for conservation and wildlife management programs. However, monitoring requires the analysis of multiple *loci* from many samples. These processes can be laborious and expensive. The choice of microsatellites and PCR calibration for genotyping are particularly daunting. Here we optimized a low-cost genotyping method using multiple microsatellite *loci* for simultaneous genotyping of up to 384 samples using next-generation sequencing (NGS). We designed primers with adapters to the combinatorial barcoding amplicon library and sequenced samples by MiSeq. Next, we adapted a bioinformatics pipeline for genotyping microsatellites based on read-length and sequence content. Using primer pairs for eight microsatellite *loci* from the fish *Prochilodus costatus*, we amplified, sequenced, and analyzed the DNA of 96, 288, or 384 individuals for allele detection. The most cost-effective methodology was a pseudo-multiplex reaction using a low-throughput kit of 1 M reads (Nano) for 384 DNA samples. We observed an average of 325 reads per individual per *locus* when genotyping eight *loci*. Assuming a minimum requirement of 10 reads per *loci*, two to four times more *loci* could be tested in each run, depending on the quality of the PCR reaction of each *locus*. In conclusion, we present a novel method for microsatellite genotyping using Illumina combinatorial barcoding that dispenses exhaustive PCR calibrations, since non-specific amplicons can be eliminated by bioinformatics analyses. This methodology rapidly provides genotyping data and is therefore a promising development for large-scale conservation-genetics studies.

Keywords: microsatellite, fish, genotyping, next-generation sequencing, conservation genetics

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Alexandre Wagner Silva Hilsdorf, University of Mogi das Cruzes, Brazil
Masafumi Nozawa, Tokyo Metropolitan University, Japan

****Correspondence:***
Evanguedes Kalapothakis
kalapothakis@gmail.com

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

Received: *14 September 2017*
Accepted: *19 February 2018*
Published: *09 March 2018*

Citation:
Pimentel JSM, Carmo AO, Rosse IC, Martins APV, Ludwig S, Facchin S, Pereira AH, Brandão-Dias PFP, Abreu NL and Kalapothakis E (2018) High-Throughput Sequencing Strategy for Microsatellite Genotyping Using Neotropical Fish as a Model. Front. Genet. 9:73. doi: 10.3389/fgene.2018.00073

INTRODUCTION

Innovative technological applications in the field of conservation genetics can contribute to wildlife monitoring and management programs. Technological advances in genomics have greatly expanded the use of genetic markers in biological research, resulting in more extensive and efficient generation and analysis of population genetics data (Putman and Carbone, 2014).

Microsatellites are DNA units composed of repeating motifs in tandem, also known as simple sequence repeats (SSRs) or short tandem repeats. Due to their high degree of polymorphism, microsatellites are used as molecular markers in genetic structure, kinship identification, genetic mapping, and others population genetics studies (Buschiazzo and Gemmell, 2006; Chistiakov et al., 2006; Yazbeck and Kalapothakis, 2007; Bhargava and Fuentes, 2010). The high statistical power per *locus* obtained through microsatellites genotyping makes this a powerful tool in population studies (Guichoux et al., 2011; Putman and Carbone, 2014). Moreover, microsatellites are preferred in forensic and kinship analyses due to their high mutation rates and multiallelic nature (Clayton et al., 1998), and are the markers most frequently used in human paternity tests (Guichoux et al., 2011).

The early difficulties in isolating sequences from microsatellites were circumvented with the use of next-generation sequencing (NGS). Indeed, thousands of microsatellite *loci* can now be identified from a single NGS run (Tang et al., 2008; Boomer and Stow, 2010; Castoe et al., 2010; Rosazlina et al., 2015). However, the current techniques present difficulties regarding PCR calibration and the choice of informative microsatellites with high specificity. Thus, a balanced reaction that maintains great sensitivity and specificity to target DNA remains a challenge. In addition, studies involving microsatellites frequently use electrophoresis, which makes PCR calibration and fragment size identification laborious, and compromises the analysis of homoplasy cases (Delmotte et al., 2001; Pasqualotto et al., 2007).

NGS has enabled high-throughput genotyping, with a range of protocols for partial sequencing of genomes such as the use of restriction enzymes, e.g., RAD (Baird et al., 2008), ddRAD (Peterson et al., 2012), and 2bRAD (Wang et al., 2012). This is collectively known as 'genotyping-by-sequencing' (GBS) (Narum et al., 2013). Sequencing of a target region (PCR amplicon) can also be used as a GBS method (Vartia et al., 2015).GBS enables analyses of multiple target sequences in several different samples simultaneously, thereby saving time and resources. Moreover, target region sequencing can be applied to DNA samples with a certain degree of degradation such as forensic and cancer biopsy samples (Kerick et al., 2011; Van Neste et al., 2012; König et al., 2015). New approaches with microsatellite genotyping using individual combinatorial barcoding have been used in conservation genetics (Scheible et al., 2011; Vartia et al., 2015).

Here, we optimized a high-throughput genotyping methodology for genetic studies in conservation programs. We genotyped microsatellite *loci* employing NGS technology using *Prochilodus costatus*, a migratory fish species found in Brazil, as a model. Microsatellite markers with a maximum size of 200 base pairs (bp) and tri- or tetra-nucleotide motifs were used to facilitate data analysis (de Valk et al., 2005). We tested distinct NGS reagent kits and varying numbers of individuals per analysis. We believe that the use of high-throughput technology in conservation studies will advance the field, allowing large amounts of data to be generated quickly and efficiently.

MATERIALS AND METHODS

Ethics Statement

We collected *P. costatus* samples in the region under the influence of the Três Marias dam (located in the state of Minas Gerais, Southeastern Brazil). To collect the samples needed for the study we obtained a Permanent Field Permit for Collecting from the Instituto Chico Mendes de Conservação da Biodiversidade (protocol number 57204-1) and also from the Instituto Estadual de Florestas (protocol number 014.007 2017).

DNA Extraction

Using a sterile metallic mold, we removed 12.5-mm^2 fragments of fish fin from samples previously collected and stored in 70% ethanol (v/v). Fragments were washed with ultrapure water and individually placed in 96-well plates. We added 50 μL of NaOH (50 mM) to the wells, sealed the plates, vortexed for 10 s, incubated at 95°C for 10 min, and vortexed again for 10 s. Then, 7.5 μL of Tris-HCl (0.5 M, pH 8.0) were added to each well and the plates vortexed for 15 s. Supernatants were transferred to new 96-well plates. We purified DNA from the supernatants using Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA, United States), according to the manufacturer's protocol. DNA was quantified using Qubit 2.0 (Invitrogen, Carlsbad, CA, United States) and its purity was evaluated using NanoDrop 2000 (ThermoScientific, Waltham, MA, United States).

Analysis and Identification of Microsatellites

Contigs previously generated by shotgun genome sequencing of *P. costatus* and deposited in the database of the Laboratory of Biotechnology and Molecular Markers were used for *in silico* analysis. The software msatcommander (Faircloth, 2008) was used to identify microsatellite regions. Over 100,000 contigs containing microsatellite regions were identified, and about 1,000 primer pairs suggested. We screened the primer pairs for the generated fragment size (amplicons of 130–200 bp), and selected microsatellites containing at least seven repeats of tri- or tetra-nucleotide motifs. de Valk et al. (2005) reported that microsatellites with tri- and tetra-nucleotide motifs can be easily discerned, avoiding issues such as stutter patter.

We selected 50 primer pairs that matched the requirements and tested them in *P. costatus* samples. PCR optimization was carried out by testing buffers with distinct concentrations of Mg^{2+} and KCl (Phoneutria Biotecnologia e Serviços LTDA, Brazil). PCR was performed using different annealing temperatures (50 – 65°C), dimethyl sulfoxide concentrations (0 – 7%), and numbers of amplification cycles (20 – 35). Amplicons were analyzed in 8% (w/v) polyacrylamide gel stained with silver nitrate. Twenty-four sets of primers were selected based on the presence of a single band (homozygous) or two adjacent bands (heterozygous). For genotyping, we used Illumina's metagenomics protocol for amplicon resequencing. An 'overhang adapter' complementary to the 'index' in Illumina's Nextera kits was added to the primers. The 'overhang adapter' sequences were as follows: Forward: 5′-TCGTCGGCAGC

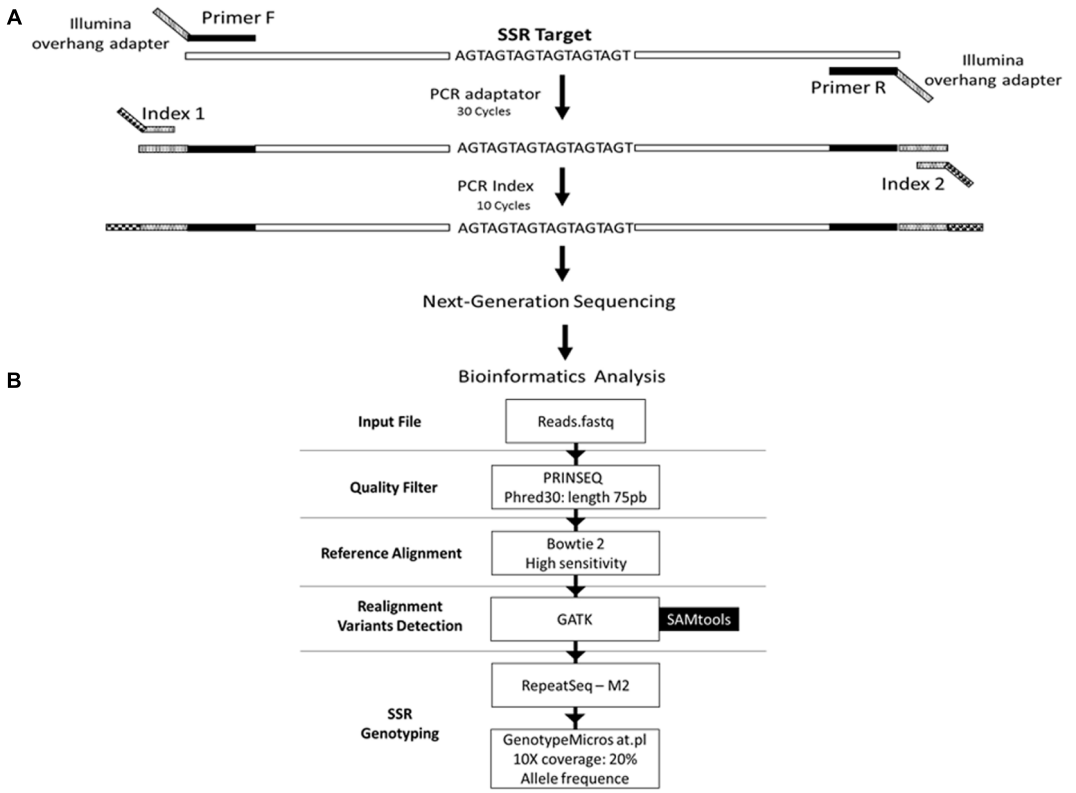


FIGURE 1 | Schematic representation of the genotyping strategy. **(A)** Amplification protocol for generation of the amplicons for sequencing. **(B)** Data analysis pipeline. The reads from the sequencing are filtered based on the quality of the bases and the length of the read, using the PRINSEQ tool. Filtered reads are aligned against the reference sequence using the software Bowtie2 with high sensitivity parameter. Then, the reads are realigned against the reference, using tools from the GATK package, to increase the confidence in genotyping. Subsequently, alleles are detected and quantified using the RepeatSeq tool. Results are filtered to decrease the false negative rate using the GenotypeMicrosat.pl script. Finally, the result of the genotyping is converted into the format required by population genetics analysis software using the same script.

GTCAGATGTGTATAAGAGACAG + *locus*-specific forward primer sequence-3′, and Reverse: 5′-GTCTCGTGGGCTCG GAGATG TGTATAAGAGACAG + *locus*-specific reverse primer sequence-3′. The primers with the overhang adapters were incorporated into the target DNA through 30 PCR cycles. A specific primer (index) containing the MiSeq adapter (which individualizes samples in the NGS procedures) was attached to amplicons through 10 PCR amplification cycles (**Figure 1A**). A preliminary genotyping test (NGS sequencing) was performed using *P. costatus* DNA isolated from five individuals. From the 24 sets of primers, eight were selected for further analyses (**Table 1**). These were selected based on performance and PCR robustness (i.e., satisfactory amplification in non-ideal conditions, such as little or degraded DNA, the presence of inhibitors, etc.) and the presence of homozygous or heterozygous alleles with low amounts of unspecific DNA. GenBank accession number: Proc10 MG456705; Proc18 MG456707; Proc22 MG456708; Proc36 MG456709; Proc37 MG456710; Proc44 MG456712; Proc48 MG456715; Proc49 MG456716.

We tested several NGS genotyping strategies: (a) multiplex, in which all primer pairs were used to amplify a single DNA sample in a single reaction (30 cycles) followed by the incorporation of the index (10 cycles) and NGS sequencing; (b) pseudo-multiplex reaction, in which each primer pair was individually used to amplify a DNA target (30 cycles), followed by pooling the amplicons of one individual and a second PCR (10 cycles) for index incorporation and MiSeq sequencing; (c) monoplex, in which DNA target amplification (30 cycles) and index incorporation (10 cycles) were performed individually for each DNA target and followed by MiSeq sequencing. **Table 2** indicates the strategies, the types of cartridges, and the number of target DNA samples used (minimum of five and maximum of 384).

To optimize the NGS genotyping procedures, we performed multiplex PCR tests in an initial MiSeq run (run 1) using a low-throughput kit (1 M reads, MiSeq Reagent Kit Nano, 300 cycles; Illumina). We mixed eight, four, and two primer pairs for the first, second, and third tests, respectively. Additionally, a monoplex test was carried out for each of the eight primer pairs selected. We used DNA from five individuals in all tests (**Table 2**) and different index sequences for each test.

The second NGS run (run 2) was carried out using a multiplex reaction with the eight selected primer pairs, 192 individuals, and a high-throughput kit (15 M cartridge, MiSeq Reagent Kit

Standard, 300 cycles; Illumina). Pseudo-multiplex reactions with eight primer pairs and 1M cartridges were performed in the runs 3, 4, 5, and 6 with 192, 96, 288, and 384 target DNAs, respectively (**Table 2**).

We used Nextera XT Index Kit v2 Sets A, B, C, and D (Illumina), which allows the sequencing of up to 384 individuals in a single MiSeq run and, therefore, accelerates the process and increases the benefit-cost ratios of the analysis. We further increased the number of individuals genotyped per run using a random nucleotide sequence (such as AAA or TTT) between the adapter and the *locus*-specific sequences. We used the following primers: Forward: 5′-TCGTCGGCAG CGTCAGATGT GTATAAGAGA CAG + **AAA** or **TTT** + *locus*-specific forward primer sequence-3′, and Reverse: 5′-GTCTCGTGGGCTCGGAGATG TGTATAAGAG ACAG + **AAA** or **TTT** + *locus*-specific reverse primer sequence-3′ (see **Table 1** for primer sequences). In these runs, only one set of adapters was used (Nextera XT Index kit v2 set A) and 288 individuals were genotyped. Noteworthy, other sequences can be used including CCC or sequences with four or more nucleotides. However, shorter sequences (one or two nucleotides) may hinder bioinformatics analysis, whereas longer sequences may reduce the efficiency of amplification. Additional tests should be performed to evaluate each case.

Amplicons of the eight *loci* were polled for each individual, quantified using Qubit, and diluted to 10 ng/μl. Then, a new pool (the combinatorial barcoding amplicon library) was prepared with all the quantified material. This library was then quantified by qPCR using the KAPA Library Quant Illumina/Universal kit (KAPA Biosciences) following the manufacturers' instructions. The library was used as input in a MiSeq run with a final concentration of 15 pM. A ready-to-use control library from Illumina (PhiX Control v3) was used in each sequencing run.

Genotyping of Microsatellites

We developed a bioinformatics pipeline for microsatellite genotyping (described below and in **Figure 1B**). To improve reliability, we trimmed and filtered the obtained reads using the PRINSEQ tool (Schmieder and Edwards, 2011). Bases with Phred scores lower than 30 and/or read lengths shorter than 75 bp were removed. Filtered reads were aligned against a FASTA file containing reference sequences for the eight microsatellite *loci* using the software Bowtie 2 (Langmead and Salzberg, 2012) with the high sensitivity option.

Alignment against a reference region that contains insertions or deletions of nucleotides, such as the microsatellite variants requires careful curation because variations in the edge of the repeat can lead to error in alignment and consequent misidentification of the alleles. One way to increase the

TABLE 1 | Primers used for amplification of the microsatellite markers selected for genotyping of *Prochilodus costatus*.

Name	Primer F	Primer R	SSR Motif	Amplicon size (bp)	Buffer*	AT (°C)	DMSO (%)	Number of Cycles
ProC10	ATTCCTGTCAATTTCCGGCC	AGGCCCAAACAGAAGGTAGG	ATT	130	4B	58	3	30
ProC18	GATCAGACCTCAGACGGGAC	GTTGTACGGAGATGCACTGC	CCGT	197	1C	60	5	25
ProC22	AGAGCTGGGATAGGCTCCAGC	CCTGGACAGGCTCCCAGTCC	AATG	130	4B	64	7	25
ProC36	GACGGAACGTCTTTAGAACC	TCTGCACATGCACGAGCGCGG	AAG	167	4B	62	5	30
ProC37	TGAAGGTGCACAGGGATAGT	TGTGAAGGTCCTGGAACCCAC	TTTA	130	4B	62	–	25
ProC44	CTTAGTGAACTGGAGCACG	GGTCCAGATTGGGCATATACAC	AAC	183	1C/4B	58	–	25
ProC48	AGCTTAGATGTGTACTAAC	GTTGAGCAGTGGTGGGGTAC	AAAG	153	1C	55	–	25
ProC49	GGTGTTTGGTTAATCACCCC	AGAGATGTGCTTATGCACGC	AAAG	146	4B	62	7	25

*SSR, simple sequence repeat; bp, base pairs; AT, annealing temperature; DMSO, dimethyl sulfoxide. *Buffer classification according to Phoneutria Biotecnologia e Seviços LTDA, Brazil.*

TABLE 2 | Strategies used for optimization of next generation sequencing (NGS) microsatellite genotyping of *Prochilodus costatus*.

NGS Run	Strategy[#]	N° of primer pairs	N° of DNA samples	Cartridge[¥]	N° of filtered reads	**N° of alleles
1	Multiplex	8, 4 and 2	5	1 M	13,691	*
	Monoplex	8			829,631	*
2	Multiplex	8	192	15 M	608,390	30
3	Pseudo-multiplex	8	192	1 M	486,130	49
4	Pseudo-multiplex	8	96	1 M	646,223	48
5	Pseudo-multiplex	8	288	1 M	1,051,375	52
6	Pseudo-multiplex	8	384	1 M	991,337	53

*Not applicate due to the low number of individuals. **Total number of alleles obtained for eight loci and variable numbers of individuals. [#]Reactions used for amplification. Multiplex, in which all primer pairs were used to amplify a single DNA sample in a single reaction followed by the incorporation of the index and NGS sequencing; Pseudo-multiplex, in which each primer pair was individually used to amplify a DNA target, followed by pooling of the amplicons of one individual (eight loci per individual) and a second PCR for index incorporation and MiSeq sequencing; Monoplex, in which DNA target amplification and index incorporation were performed individually for each DNA target and followed by MiSeq sequencing. [¥]Type of cartridge used. 1 M is a MiSeq Reagent Nano Kit v2 (1 M reads) and 15 M is a MiSeq Reagent Kit v2 (15 M reads).*

confidence and reduce error is to realign the reads taking into account the nucleotide sequence of the edges of the repeat and possible variants within the region. As no information on the variants of the microsatellite region studied here was available in public databases, we identified the possible variants at the edges of the repetition of each of the eight *loci*. We used the SAMtools package (Li et al., 2009) to detect possible variants in the mapping file (BAM format) of each of the 384 sequenced individuals. As a result, we obtained a Variant Call Format (VCF) file containing all the variants in the regions around the microsatellites repeat motifs. Next, we realigned the reads that mapped to the reference repeating regions using the tools RealignTargetCreator and IndelRealigner from the GATK package (McKenna et al., 2010). These tools perform a local realignment to the regions of the repeat motifs taking into account only high-quality reads that completely cover the repeat region and the variants described for the region (VCF file). Reads that did not match these criteria were removed. A realignment file (BAM format) containing only reads that realigned to each *locus* (tested twice) was obtained for each individual, thus increasing the confidence in the identification of the alleles.

We used the RepeatSeq tool (Highnam et al., 2012), with parameter −M 2 (minimum sequencing quality required value) to identify and quantify the alleles from the realignment files. This tool requires a file containing the chromosome coordinates and the repeat region motif sequence. Since *P. costatus* genomic information was unavailable, we used the information obtained through the microsatellites amplicon sequencing as an independent chromosome. We created an input file containing the name, the starting and ending positions of the repeat sequence in the amplicon, and the sequence of repeat motif for each *locus*. The information of the chromosomal regions was replaced by the information of each of the amplicons. There is no limit regarding the size or number of amplicons. However, it is important to enter the correct location and base sequence of the repeat region. The RepeatSeq tool uses the coordinate file to search for repeat regions in the realignment files and calculates the repeat length, which determines the alleles. A repeat ATTATTATTATT, for example, would be defined as allele 12. After identification of the repeat motif, the reads that aligned to that region are selected and quantified, according to their number of repeats. The resulting file contains the full read annotation of the reference microsatellite, including the total number of alleles detected, total number of reads, total number of reads per allele, and mapping quality score.

To avoid false negatives and to convert the results into the input format required by the software commonly used in population genetics, we developed a Perl script, named GenotypeMicrosat.pl. This script performs a detailed analysis of the RepeatSeq output file. We determined the individual's genotype for each *locus* using the following filter criteria: (1) maximum of two alleles per individual per *locus*; (2) at least 10 reads per *locus* in the entire repeat sequence, including eight bases in the 5′ and eight bases in the 3′ flanking regions; and (3) at least 20% of reads corresponding to a second allele for an individual to be considered heterozygous in a given *locus*. Individuals with a second allele coverage of less than 20% were considered homozygous. Following application of these filter criteria, we generated a spreadsheet containing the genotypes of each individual for each of the eight *loci*. *Loci* that did not attain the filter requirements were identified as 'NA.' The generated spreadsheet can be easily adapted for other population genetics analysis programs.

RESULTS

In comparison with monoplex, multiplex reactions generated 10- to 100-times fewer reads (run 1) that could be used in genotyping (**Figure 2**). For some *loci*, no reads were detected in the multiplex tests, which suggests intense primer competition. On the other hand, consistent results were observed with monoplex-amplified samples (**Figure 2**).

We compared the genotyping results of 192 individuals by multiplex (run 2) or pseudo-multiplex (run 3) reactions. The 15 M cartridge tested in run 2 yielded about 4.7 times more reads than the 1 M cartridge (Nano) kit tested in run 3. Nevertheless, the number of reads generated with the Nano kit (run 3) allowed genotyping of all individuals. In run 2 (multiplex, **Table 3**), reads were obtained for five *loci* only. Amplification efficiency, determined as the fraction of individuals successfully genotyped for a given *locus*, was superior in the pseudo-multiplex reaction (run 3) for all *loci*, except proC10 and proC37 (**Table 3**).

As expected, the standard curve generated from runs 4, 5, and 6 revealed a strong negative correlation between the number of individuals tested and the number of reads generated (**Figure 3**). The obtained distribution of reads per *locus* in these three runs (**Table 4**) was uneven. When analyzing 384 individuals (run 6), we encountered an average of 325 reads per individual per *locus* for all *loci*. The yield differed depending on the *locus* being evaluated and for the same *locus* in different runs. For example, marker proC36 yielded an average of 1,018 reads per individual in run 5 and 198 reads per individual in run 6. On

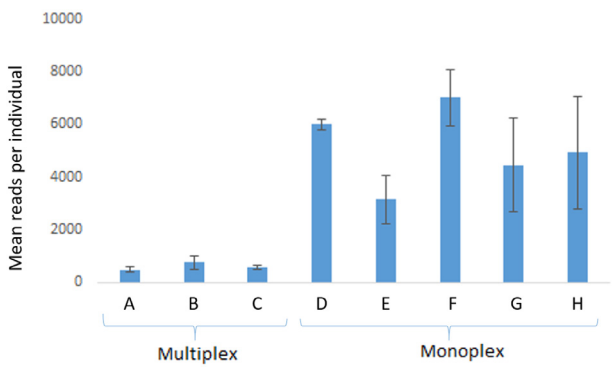


FIGURE 2 | Average (mean) read yield in run 1 (n = 5 individuals). Multiplex reactions were performed in A with eight *loci* (proC10, proC18, proC22, proC36, proC37, proC44, proC48, proC49), in B with four *loci* (proC18, proC36, proC37, proC44), and in C with two *loci* (proC48 and proC49). Monoplex reactions were performed in D (proC10), E (proC18), F (proC36), G (proC48), and H (proC49). Error bars are standard deviation. MiSeq Reagent Kit v2 1 M reads was used.

TABLE 3 | Amplification and genotyping efficiency test for multiplex (run 2) and pseudo-multiplex (run 3) systems.

Loci	Multiplex (%)	Pseudo-multiplex (%)
ProC10	97.86	86.09
ProC18	X	70.05
ProC22	44.92	80.21
ProC36	50.27	93.05
ProC37	50.80	16.04
ProC44	19.79	65.77
ProC48	X	63.64
ProC49	X	97.86

Values correspond to the genotyping efficiency (%) for 192 individuals (Prochilodus costatus). x, no date obtained.

the other hand, proC37 yielded only 65 and 42 reads in runs 5 and 6, respectively. This variation did not compromise the genotyping analysis because of the bioinformatics parameters used.

From the 3,771,786 reads yielded in run 6, 1,998,885 (53%) passed quality filtering. Repeatseq identified 991,337 high-quality reads (**Table 2**), which were subsequently used for allele detection. Of the 3,072 microsatellites genotyped in this run (eight *loci* from 384 individuals), 1,179 (39%) remained with undetermined genotype due to the low number of reads per *locus* (<10). The total number of alleles obtained in run 6 (384 individuals and 1 M cartridge) was 53, with the *locus* proC49 showing the highest number of alleles (10) while proC10 showed the lowest (4). Genotyping using the primers marked with base triads AAA and TTT generated results similar to those obtained in run 6.

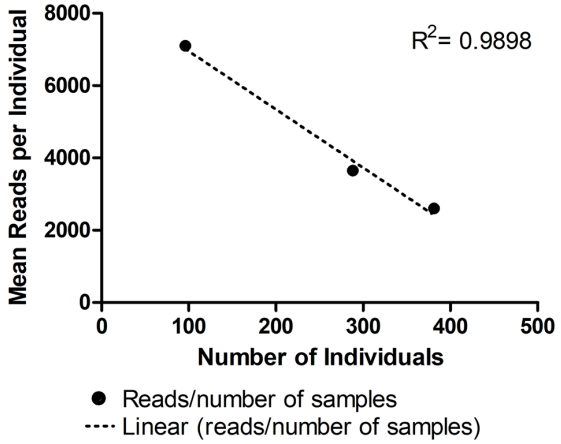


FIGURE 3 | Average number of reads generated per individual in three distinct runs using MiSeq Reagent Kit v2 1 M reads: run 4 (96 individuals), run 5 (288 individuals), and run 6 (384 individuals). • Represents the mean number of reads for each number of samples tested. Linear regression is represented by the dashed line (·····). For more details, see **Table 2**.

TABLE 4 | Percentage* of reads generated for each *locus* tested in three distinct sequencing runs (runs 4, 5, and 6).

Loci	% run 4	% run 5	% run 6
ProC10	5.79	10.36	5.44
ProC18	5.50	10.66	3.01
ProC22	12.69	8.08	14.52
ProC36	6.39	27.89	19.02
ProC37	0.26	1.78	0.21
ProC44	4.49	6.26	2.49
ProC48	7.32	3.99	3.81
ProC49	57.56	30.99	51.50

**Percentage was calculated from 646,223 reads generated in run 4 with 96 individuals, 1,051,375 reads generated in run 5 with 288 individuals, and 991,337 reads generated in run 6 with 384 individuals. One MiSeq Reagent Nano Kit v2 (1 M reads) was used for each run.*

DISCUSSION

In the present paper we showed the potential of microsatellite genotyping by NGS as a fast and cost-effective methodology to be implemented in large-scale population genetic studies. We used a combination of commercially available indexes and genotyped 384 individuals per run. The efficiency observed with the Nano (1 M reads) kit represents a substantial cost reduction over the NGS runs with the 15 M reads kit.

The flowchart presented herein was developed to ensure high accuracy in microsatellite genotyping. We excluded low-quality reads from the analysis and aligned the reads against the reference using Bowtie 2, the best-suited software for INDEL-rich *loci* (Highnam et al., 2012). This realignment step increases the confidence in allele detection, since our analysis considers only reads that cover the whole repeat region, including eight bases in the 5′ and eight bases in the 3′ flanking regions. Additionally, the pipeline proposed verifies all neighboring variations (5′ and 3′ regions flanking the microsatellite motif), allowing the identification of homoplastic motif repeat numbers and fragment length. The microsatellite genotyping tool RepeatSeq uses a Bayesian approach, which considers characteristics of the read and the sequence under analysis (Highnam et al., 2012). The developed GenotypeMicrosat.pl script further increases the confidence of the genotyping by establishing a minimum of 10 reads to confirm an allele.

Read realignment with the GATK package requires a file containing all the variants described for the analyzed species. However, information about variations in our model species is scarce in public databases. As an alternative, we used single nucleotide polymorphisms (SNPs) and INDELs detected through sequenced amplicons as input for GATK. To circumvent the lack of data on the *P. costatus* microsatellite genomic localization required by RepeatSeq, we generated a file containing the genomic coordinates and repeat sequences to use as input. Our successful attempts to overcome the lack of genomic information for our species of interest highlight the potential application of the pipeline proposed for microsatellite genotyping of species for which genomic data are not available, and further support its use in genetic monitoring programs.

The maximum number of individuals genotyped per run is limited by the number of commercially available indexes

(currently 384). However, the amplification success of primers containing adenine or thymine trios shows the potential of this tool to increase the number of individuals genotyped in a single run. We genotyped eight *loci* per individual, with an average coverage of 325 reads per *locus* per individual in the runs with the Nano kit using 384 samples. As the genotyping pipeline considers a minimum of 10 reads per allele per individual, none of the runs reached the maximum capacity of the cartridge. In theory, the number of individuals or the number of *loci* could be increased up to four times per run (32 *loci* or 1536 individuals), based on the regression analysis. However, this possibility must be weighted carefully since the number of alleles considerably varies among the *loci*. For instance, ProC49 showed an average of 700 reads, while ProC37 showed 20 reads per individual in run 6. The number of *loci* may, therefore, be increased or decreased depending on the quality of the *loci* tested. Previous knowledge of the quality of a given *locus* also allows for the use of greater amounts of amplicon for *loci* with low yield. These findings open the prospect of using *loci* for which PCR reactions are not 100% efficient and represent an advantage in the genetic analysis of understudied species with limited availability of microsatellites.

Traditional methods employing microsatellite molecular markers have disadvantages such as the long optimization process and the elevated costs, especially in the development of multiplex systems. Additionally, automation limitations and data management requirements can prevent technology transfer among different laboratories (Guichoux et al., 2011). Previous studies have shown that allele sizes generated by capillary electrophoresis may vary depending on the equipment and running conditions (Delmotte et al., 2001; Foulet et al., 2005; Pasqualotto et al., 2007), and the number of *loci* that can be multiplexed with this technique is limited by the number of commercially available fluorophores. On the other hand, many *loci* can be simultaneously genotyped in a single NGS run (Scheible et al., 2011). The direct sequencing of *loci* is a more reliable approach as it allows for the analysis of all the variations in the fragment, thus ensuring greater reliability. Furthermore, technology transfer and detection of technical errors are facilitated by NGS. Here we tested eight *loci* per individual. Nevertheless, our pipeline has the potential to provide analysis of a significantly larger number of *loci* and recent publications with neotropical migratory fish revealed a minimum of seven and a maximum of 13 *loci* for up to 30 individuals (Rueda et al., 2011; Berdugo and Narvaez Barandica, 2014; Coimbra et al., 2017).

Despite the poor results of direct multiplex reactions, we successfully optimized a 'pseudo-multiplex strategy,' in which previous monoplex reactions were performed for each sample and the amplicons mixed in the indexing reaction. This strategy reduced the cost and duration of the analysis and may be used in the genotyping of other markers, such as SNPs, and in metagenomics studies.

CONCLUSION

We present a novel method for microsatellite genotyping based on the Illumina combinatorial barcoding using a Nano kit. This approach is faster and more efficient than those currently available and offers large amounts of high-quality data for conservation genetics and population studies.

AUTHOR CONTRIBUTIONS

JP and EK designed and coordinated the work. SL, SF, AP, PB-D, and NA acquired the data. AC, IR, and AM developed the pipeline. JP, AC, IR, and SL analyzed the data. All authors contributed to data interpretation and provided substantial contributions to manuscript writing. All authors approved the final version prior to submission.

FUNDING

This work was funded by Companhia Energética de Minas Gerais (CEMIG) (project P&D GT 455), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (Edital Ciências Forenses no. 25/2014), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

REFERENCES

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3:e3376. doi: 10.1371/journal.pone.0003376

Berdugo, G. O., and Narvaez Barandica, J. C. (2014). Genetic diversity and population structure of bocachico *Prochilodus magdalenae* (Pisces, Prochilodontidae) in the Magdalena River basin and its tributaries, Colombia. *Genet. Mol. Biol.* 37, 37–45.

Bhargava, A., and Fuentes, F. F. (2010). Mutational dynamics of microsatellites. *Mol. Biotechnol.* 44, 250–266. doi: 10.1007/s12033-009-9230-4

Boomer, J. J., and Stow, A. J. (2010). Rapid isolation of the first set of polymorphic microsatellite loci from the Australian gummy shark, *Mustelus antarcticus* and their utility across divergent shark taxa. *Conserv. Genet. Resour.* 2, 393–395. doi: 10.1007/s12686-010-9274-6

Buschiazzo, E., and Gemmell, N. J. (2006). The rise, fall and renaissance of microsatellites in eukaryotic genomes. *Bioessays* 28, 1040–1050. doi: 10.1002/bies.20470

Castoe, T. A., Poole, A. W., Gu, W., Jason de Koning, A. P., Daza, J. M., Smith, E. N., et al. (2010). Rapid identification of thousands of copperhead snake (*Agkistrodon contortrix*) microsatellite loci from modest amounts of 454 shotgun genome sequence. *Mol. Ecol. Resour.* 10, 341–347. doi: 10.1111/j.1755-0998.2009.02750.x

Chistiakov, D. A., Hellemans, B., and Volckaert, F. A. (2006). Microsatellites and their genomic distribution, evolution, function and applications: a review with special reference to fish genetics. *Aquaculture* 255, 1–29. doi: 10.1016/j.aquaculture.2005.11.031

Clayton, T. M., Whitaker, J. P., Sparkes, R., and Gill, P. (1998). Analysis and interpretation of mixed forensic stains using DNA STR profiling. *Forensic Sci. Int.* 91, 55–70. doi: 10.1016/S0379-0738(97)00175-8

Coimbra, M. R. M., Lima, A. P. S., Oliveira, K. K. C., and Severi, W. (2017). Microsatellite assessment of the genetic diversity in indigenous populations of curimba (*Prochilodus argenteus*) in the São Francisco river (Brazil). *Conserv. Genet.* 18, 965–975. doi: 10.1007/s10592-017-0947-5

de Valk, H. A., Meis, J. F., Curfs, I. M., Muehlethaler, K., and Mouton, J. W. (2005). Use of a novel panel of nine short tandem repeats for exact and high-287

resolution fingerprinting of *Aspergillus fumigatus* isolates. *J. Clin. Microbiol.* 43, 4112–4120. doi: 10.1128/JCM.43.8.4112-4120.2005

Delmotte, F., Leterme, N., and Simon, J. C. (2001). Microsatellite allele sizing: difference between automated capillary electrophoresis and manual technique. *Biotechniques* 31, 810, 814–816, 818.

Faircloth, B. C. (2008). Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol. Ecol. Resour.* 8, 92–94. doi: 10.1111/j.1471-8286.2007.01884.x

Foulet, F., Nicolas, N., Eloy, O., Botterel, F., Gantier, J. C., Costa, J. M., et al. (2005). Microsatellite marker analysis as a typing system for *Candida glabrata*. *J. Clin. Microbiol.* 43, 4574–4579. doi: 10.1128/JCM.43.9.4574-4579.2005

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., et al. (2011). Current trends in microsatellite genotyping. *Mol. Ecol. Resour.* 11, 591–611. doi: 10.1111/j.1755-0998.2011.03014.x

Highnam, G., Franck, C., Martin, A., Stephens, C., Puthige, A., and Mittelman, D. (2012). Accurate human microsatellite genotypes from high-throughput resequencing data using informed error profiles. *Nucleic Acids Res.* 41:e32. doi: 10.1093/nar/gks981

Kerick, M., Isau, M., Timmermann, B., Sültmann, H., Herwig, R., Krobitsch, S., et al. (2011). Targeted high throughput sequencing in clinical cancer settings: formaldehyde fixed-paraffin embedded (FFPE) tumor tissues, input amount and tumor heterogeneity. *BMC Med. Genomics* 4:68. doi: 10.1186/1755-8794-4-68

König, K., Peifer, M., Fassunke, J., Ihle, M. A., Künstlinger, H., Heydt, C., et al. (2015). Implementation of amplicon parallel sequencing leads to improvement of diagnosis and therapy of lung cancer patients. *J. Thorac. Oncol.* 10, 1049–1057. doi: 10.1097/JTO.0000000000000570

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., and Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Mol. Ecol.* 22, 2841–2847. doi: 10.1111/mec.12350

Pasqualotto, A. C., Denning, D. W., and Anderson, M. J. (2007). A cautionary tale: lack of consistency in allele sizes between two laboratories for a published multilocus microsatellite typing system. *J. Clin. Microbiol.* 45, 522–528. doi: 10.1128/JCM.02136-06

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., and Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS One* 7:e0037135. doi: 10.1371/journal.pone.0037135

Putman, A. L., and Carbone, I. (2014). Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecol. Evol.* 4, 4399–4428. doi: 10.1002/ece3.1305

Rosazlina, R., Jacobsen, N., Ørgaard, M., and Othman, A. S. (2015). Utilizing next generation sequencing to characterize microsatellite loci in a tropical aquatic plant species *Cryptocoryne cordata* var. *cordata (Araceae)*. *Biochem. Syst. Ecol.* 61, 385–389. doi: 10.1016/j.bse.2015.06.033

Rueda, E., Sommer, J., Scarabotti, P., Markariani, R., and Ortí, G. (2011). Isolation and characterization of polymorphic microsatellite loci in the migratory freshwater fish *Prochilodus lineatus* (*Characiformes. Prochilodontidae*). *Conserv. Genet. Resour.* 3, 681–684. doi: 10.1111/jfb.12632

Scheible, M., Loreille, O., Just, R., and Irwin, J. (2011). Short tandem repeat sequencing on the 454 platform. *Forensic Sci. Int.* 3, e357–e358. doi: 10.1016/j.fsigen.2014.04.010

Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27, 863–864. doi: 10.1093/bioinformatics/btr026

Tang, J., Baldwin, S. J., Jacobs, J. M., van der Linden, C. G., Voorrips, R. E., Leunissen, J. A., et al. (2008). Large-scale identification of polymorphic microsatellites using an in silico approach. *BMC Bioinformatics* 9:374. doi: 10.1186/1471-2105-9-374

Van Neste, C., Van Nieuwerburgh, F., Van Hoofstat, D., and Deforce, D. (2012). Forensic STR analysis using massive parallel sequencing. *Forensic Sci. Int. Genet.* 6, 810–818. doi: 10.1016/j.fsigen.2012.03.004

Vartia, S., Villanueva-Cañas, J. L., Finarelli, J., Farrell, E. D., Collins, P. C., Hughes, G. M., et al. (2015). A novel method of microsatellite genotyping-by-sequencing using individual combinatorial barcoding. *R Soc. Open Sci.* 3:15056. doi: 10.1098/rsos.150565

Wang, S., Meyer, E., McKay, J. K., and Matz, M. V. (2012). 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nat. Methods* 9, 808–810. doi: 10.1038/nmeth.2023

Yazbeck, G. M., and Kalapothakis, E. (2007). Isolation and characterization of microsatellite DNA in the piracema fish *Prochilodus lineatus* (Characiformes). *Genet. Mol. Res.* 5, 1026–1034.

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 14 March 2018
doi: 10.3389/fgene.2018.00075

Evaluation of Reference Genes to Analyze Gene Expression in Silverside *Odontesthes humensis* Under Different Environmental Conditions

*Tony L. R. Silveira[1], William B. Domingues[1], Mariana H. Remião[1], Lucas Santos[1], Bruna Barreto[1], Ingrid M. Lessa[1], Antonio Sergio Varela Junior[2], Diego Martins Pires[3], Carine Corcini[3], Tiago Collares[4], Fabiana K. Seixas[4], Ricardo B. Robaldo[5] and Vinicius F. Campos[1]**

[1] Laboratory of Structural Genomics, Biotechnology Graduate Program, Federal University of Pelotas, Pelotas, Brazil, [2] Institute of Biological Sciences, Federal University of Rio Grande, Rio Grande, Brazil, [3] Veterinary Faculty, Federal University of Pelotas, Pelotas, Brazil, [4] Laboratory of Cancer Biotechnology, Biotechnology Graduate Program, Federal University of Pelotas, Pelotas, Brazil, [5] Laboratory of Physiology, Institute of Biology, Federal University of Pelotas, Pelotas, Brazil

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

Reviewed by:
Bruno Cavalheiro Araújo, University of Mogi das Cruzes, Brazil
Diogo Teruo Hashimoto, Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), Brazil

***Correspondence:**
Vinicius F. Campos
fariascampos@gmail.com

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

Received: *23 October 2017*
Accepted: *19 February 2018*
Published: *14 March 2018*

Citation:
Silveira TLR, Domingues WB, Remião MH, Santos L, Barreto B, Lessa IM, Varela Junior AS, Martins Pires D, Corcini C, Collares T, Seixas FK, Robaldo RB and Campos VF (2018) Evaluation of Reference Genes to Analyze Gene Expression in Silverside Odontesthes humensis Under Different Environmental Conditions. Front. Genet. 9:75. doi: 10.3389/fgene.2018.00075

Some mammalian reference genes, which are widely used to normalize the qRT-PCR, could not be used for this purpose due to its high expression variation. The normalization with false reference genes leads to misinterpretation of results. The silversides (*Odontesthes* spp.) has been used as models for evolutionary, osmoregulatory and environmental pollution studies but, up to now, there are no studies about reference genes in any *Odontesthes* species. Furthermore, many studies on silversides have used reference genes without previous validations. Thus, present study aimed to was to clone and sequence potential reference genes, thereby identifying the best ones in *Odontesthes humensis* considering different tissues, ages and conditions. For this purpose, animals belonging to three ages (adults, juveniles, and immature) were exposed to control, Roundup®, and seawater treatments for 24 h. Blood samples were subjected to flow-cytometry and other collected tissues to RNA extraction; cDNA synthesis; molecular cloning; DNA sequencing; and qRT-PCR. The candidate genes tested included *18s*, *actb*, *ef1a*, *eif3g*, *gapdh*, *h3a*, *atp1a*, and *tuba*. Gene expression results were analyzed using five algorithms that ranked the candidate genes. The flow-cytometry data showed that the environmental challenges could trigger a systemic response in the treated fish. Even during this systemic physiological disorder, the consensus analysis of gene expression revealed *h3a* to be the most stable gene expression when only the treatments were considered. On the other hand, *tuba* was the least stable gene in the control and *gapdh* was the least stable in both Roundup® and seawater groups. In conclusion, the consensus analyses of different tissues, ages, and treatments groups revealed that *h3a* is the most stable gene whereas *gapdh* and *tuba* are the least stable genes, even being considered two constitutive genes.

Keywords: algorithm, expression, fish, gene, normalization, real time PCR, sequencing, validation

INTRODUCTION

Northern-blotting, ribonuclease protection assay, microarray, RNA-Seq, semi-quantitative RT-PCR, and quantitative RT-PCR (qRT-PCR) are among the most currently used techniques for evaluation of gene expression levels. The latter stands out for fast readout, high throughput and high automation potential (Tang et al., 2012). In absolute quantification of qRT-PCR, the increasing level of fluorescent signal emitted by the amplified products is compared to a standard curve. The relative quantification describes the change in the expression pattern of genes of interest when subjected to certain controlled situations compared to the constant expression of reference genes (RGs), also known as housekeeping genes under same conditions. Therefore, it is extremely important to have a stable and constant expression pattern of RGs regardless of gender, age, organ or tissue, and even after environmental changes. Thus, it is necessary to validate diverse RGs for various species and their stages of development in different tissues, as well as for each type of controlled environmental condition (Livak and Schmittgen, 2001; Zheng and Sun, 2011).

However, most studies related to the quantification of gene expression in various species of teleosts are not based on RGs displaying unchanged expression. Most studies use RGs properly confirmed in mammalian species to normalize the relative quantification using qRT-PCR of non-mammalian species, such as fishes. The main RGs from mammals that have been empirically used to normalize experiments in teleost species are the *GAPDH* (Lin et al., 2004); *ACTB* (Choi and An, 2008; Nawata et al., 2010); *EF1α* (Scott et al., 2004; Nilsen et al., 2007; Patterson et al., 2012; Sinha et al., 2015); *RPL* (Hsu et al., 2014) and *18S* (Sinha et al., 2015). Studies also report the distribution of RGs in the tissues; however, these genes are not objectively evaluated in relation to the maintenance of the constant expression during development stages of individuals, in both genders or in different controlled environmental conditions (Tomy et al., 2009; Hsu et al., 2014).

Numerous studies have reported that some mammalian RGs, which are widely used to normalize the qRT-PCR reactions of different species (i.e., *ACTB* and *GAPDH*), could not be used owing to their high variation in expression levels (Bustin, 2002; Dheda et al., 2004; Silver et al., 2006). This could lead to misinterpretation of results. It is impossible to find suitable RGs that exhibit constant expression pattern for all species and under all experimental conditions (Zheng and Sun, 2011; Tang et al., 2012). Due to these reasons, the identification of RGs for each species and each experimental condition is justified.

The silversides (*Odontesthes* spp.) comprise a genus of fish that are naturally endemic to the South American waters (Bemvenuti, 2006). This genus consists of the biggest number of species of the South American atherinopsids (Dyer, 2006). In nature, the silversides are usually found in southern Brazil, Uruguay, and Argentina (Bemvenuti, 2006). In these regions, some species have economic importance for production, flesh commercialization and sport fishing (Menone et al., 2000; Somoza et al., 2008). Furthermore, some silversides have been used as experimental models for evolutionary, osmoregulatory, and environmental pollution studies due to the radiation after continental waters invasion from ocean, euryhaline biology and inhabitation of environments contaminated by glyphosate coming from rice and soybean monocultures in fields near to the habitat of silversides (Menone et al., 2000; Carriquiriborde and Ronco, 2006; Tsuzuki et al., 2007; Piedras et al., 2009; Bloom et al., 2013; Hughes et al., 2017; Zebral et al., 2017). Moreover, the silversides have already been utilized to address various aspects of gene expression (Strobl-Mazzulla et al., 2005; Karube et al., 2007; Fernandino et al., 2008, 2011; Majhi et al., 2009; Miranda et al., 2009; Blasco et al., 2010; Gómez-Requeni et al., 2012; Pérez et al., 2012; González et al., 2015). The expression pattern of a gene can be used to assess physiological responses of fishes to environmental changes, such as those to which some silversides are exposed in their habitat (salinity variations and contamination by glyphosate) (Wang et al., 2014; Velasques et al., 2016).

Similarly, *Odontesthes humensis* de Buen, 1953, a native species to coastal lagoons from southern Brazil, Uruguay and Argentina, has similar potential to be used as a biological model. However, a previous basic knowledge about the RGs is necessary to employ modern techniques such as qRT-PCR to study the gene expression of different genes of interest, independent from the biological model. Up to now, no studies have focused on the sequence determination and validation of RGs in any species of *Odontesthes* spp. Due to this, the present study aimed to clone and sequence eight potential RGs, thereby identifying and validating the best RGs in *O. humensis* using different tissues, stages of development and a natural and an anthropized environmental conditions, to standardize the qRT-PCR technique.

Abbreviations: RG, reference gene.

MATERIALS AND METHODS

Animals and Sample Collection

The silversides *Odontesthes humensis* used in this study was procreated from the eggs previously collected in nature (Arroio Grande, Brazil: 32°14′15″S/53°05′13″W) and hatched in tanks. The silversides were kept within an experimental room in nine 1,000 L circular plastic tanks with water at a temperature of 13.3 ± 1.8°C, pH 7.0 ± 0.7, salinity 3.2 ± 0.67 ppt, dissolved oxygen 9.5 ± 0.5 mg/L, ammonia levels lower than 0.6 mg/L and a natural photoperiod of winter. The tank sides were opaque to reduce visual stress, and once a week 2/3 of the water in the tanks were renewed. Three tanks were occupied by 3-year-old fish (5 animals/tank), classified as adults considering the complete development of gonads. Other three tanks were occupied by 1.5-year-old fish (5 animals/tank), classified as juveniles considering the partial development of gonads. In addition, three tanks were occupied by 6-month-old silversides (5 animals/tank), classified as immature due to the undifferentiated gonads. The animals were fed thrice a day on a commercial diet (Supra, 38% crude protein) and zooplankton *ad libitum*. The acclimation period was 4 weeks under these conditions.

The experimental design consisted of three treatment groups: the "control group," exposed to water in previously mentioned conditions; the "Roundup® group," exposed to Roundup® Transorb (a glyphosate based herbicide) diluted in water at a concentration of 10 mg.L^{-1} (acid equivalent, a.e.) of glyphosate; and the "seawater group," exposed to seawater at 30 ppt (dissolution of non-iodized sea salt in the medium). Three tanks were prepared for each treatment, the first one received the adults, the second one received the juveniles, and the third one received the immature individuals. After the acclimation period, the animals were exposed to treatments for 24 h. The water quality parameters did not differ statistically from those observed in the acclimatization period, except for the salinity of the group exposed to seawater. During the experimental period, the fish continued to be fed as during the acclimation period.

Each animal was hooked from the tank and anesthetized by submersion in benzocaine at 50 ppm. Three fish from each tank were collected, leading to a total of 27 animals. The animals were weighed, measured, and euthanized by cranial spinal section and excision of the brain. Adults, juveniles and immature silversides had mean lengths of 20.2 ± 2.3 cm, 15.2 ± 1.2 cm, and 11.7 ± 0.7 cm and mean weight of 55.2 ± 11.7 g, 24.6 ± 6.5 g, and 9.6 ± 1.9 g, respectively. Brain, hepatopancreas, gills, and kidney (the main organs affected by glyphosate and/or hyperosmotic environments in fishes; Tipsmark et al., 2004; Carriquiriborde and Ronco, 2006; Hiroi and McCormick, 2012; Harayashiki et al., 2013; Velasques et al., 2016) from adults, juveniles and immature silversides in the three treated groups were collected and preserved in liquid nitrogen (N_2). Furthermore, a total of 20 μL of peripheral blood was collected by puncturing the caudal fin. Each blood sample was added to 1 mL of fetal bovine serum (FBS), stored at 4°C, and protected from the light until used for flow-cytometry analysis. The animal use and all handling practices were approved by the Ethics Committee on Animal Experimentation of the Federal University of Pelotas (process no. 23110.007018/2015-85). The Roundup® residual from experimental tanks was sent to university chemical waste facility to to be treated prior to disposal.

Flow-Cytometric Analysis

The flow-cytometry analysis was performed using the Attune® Acoustic Focusing Flow Cytometer (Applied Biosystems, USA) to evaluate the efficacy of treatments to affect the physiology of silversides. Each blood sample was exposed to 2 mM of Hoechst 33342 (H33342; Sigma-Aldrich, USA) for 5 min before each read, except in the DNA damage analysis. Events were detected by fluorochrome with violet laser (405 nm) and photomultiplier (PMT) VL1 (filter 450/40 nm). The green, orange, and red fluorescences were analyzed by blue laser (488 nm) and PMTs BL1 (filter 530/30 nm), BL2 (filter 575/24 nm), and BL3 (filter > 640 nm) filters, respectively. Cytometry fluorescence stability was tested daily using standard beads (Invitrogen, USA). The acquisition rate was 200 events/s totaling 20,000 events per sample. All assays were performed in triplicate. The results were analyzed using the Attune Cytometric Software v2.1. The non-cellular events were eliminated from the analysis by scatter plots of forward scatter (FSC) × side scatter (SSC) (Petrunkina et al., 2005) and negative fluorescence of H33342 events (debris).

For evaluation of reactive oxygen species (ROS) produced by the erythrocytes, 10 μL of blood sample previously collected and stored was added to 20 μL of saline solution with 2 μM of 2′,7′-dichlorofluorescein diacetate (DCFH-DA) and 5 μM of propidium iodide (PI) fluorescent probes (Sigma-Aldrich Co., USA). The samples were analyzed in triplicate after incubation for 60 min at 22°C in the dark. The ROS production was measured by the median of the intensity of the green fluorescence. Only intact cells (PI negative) were evaluated.

To evaluate DNA damage in the erythrocytes, 10 μL of blood sample previously collected and stored was added to 5 μL of TNE (0.01 M Tris-HCl. 0.15 M NaCl, 0.001 M EDTA, pH 7.2) and 30 s later to 10 μL of Triton 1X (Triton X-100, 1%, v/v). Then, 50 μL acridine orange dye (2 mg/mL, #A6014, Sigma-Aldrich, USA) was added to the sample, followed by incubation from 30 s up to 2 min before each reading. The samples were analyzed in triplicate. The DNA of the erythrocytes was classified as integrated (green fluorescence emission) and damaged (orange/red fluorescence emission). The percentage of DNA fragmentation index (DFI) was obtained by the median of the red fluorescence intensity/(median of the red + green fluorescence intensities) × 100.

RNA Extraction and cDNA Synthesis

The total RNA was extracted from tissue samples using the commercial kit RNeasy® Mini Kit (Qiagen, USA) following the manufacturer's instructions. The RNA was treated with DNA-free® Kit (Ambion, USA) to remove genomic DNA contamination. Subsequently, the RNA concentration and purity were measured using a spectrophotometer NanoVue™ Plus (GE Healthcare Life Science, USA), and only the samples presenting high purity ($OD_{260/280} \geq 2.0$ nm) were used. In addition, RNA quality was analyzed on the 4,200 TapeStation (Agilent Technologies, USA) system using the TapeStation analysis application. The mean RIN number for each experimental group is listed in Supplementary Table 1. First-strand cDNA synthesis was performed with 2 μg of RNA using the commercial kit High Capacity cDNA Reverse Transcription (Applied Biosystems, USA) according to manufacturer's recommendation. The reactions were run in SimpliAmp™ thermal cycler (Applied Biosystems, USA). Finally, the cDNA was stored at −20°C until its use.

Amplification and Cloning of Candidate RGs

The tested candidates RGs were: 18S ribosomal RNA (*18s*), β-actin (*actb*), elongation factor 1 α (*ef1a*), eukaryotic translation initiation factor 3g (*eif3g*), glyceraldehyde-3-phosphate dehydrogenase (*gapdh*), histone h3a (*h3a*), Na^+/K^+-ATPase α (*atp1a*), and tubulin-α (*tuba*) (**Table 1**). The *atp1a* gene represented the control, which is a constitutive gene that is reported to be affected by glyphosate-based herbicides and salt treatments (Shiogiri et al., 2012; Armesto et al., 2014).

TABLE 1 | Summary of candidate reference genes evaluated in the present study.

Gene symbol	Gene name	Function	GenBank accession No.
18s	18S ribosomal RNA	Ribosomal subunit	KU639718
actb	β-Actin	Cytoskeletal protein	KX060039
ef1a	Elongation factor 1-α	Factor for protein translation	KU639717
eif3g	Eukaryotic translation initiation factor 3g	Initiation of protein synthesis	KX060035
gapdh	Glyceraldehyde-3-phosphate dehydrogenase	Glycolytic enzyme	KX060038
h3a	Histone h3a	Component of nucleosomes	KX060037
atp1a	Na^+/K^+-ATPase-α	Active transport of Na^+ and K^+	KR920364
tuba	Tubulin-α	Cytoskeletal protein	KX035015

The gene fragments were amplified by PCR using primers designed in the PriFi online tool (https://services.birc.au.dk/prifi/) after alignments (Supplementary Table 2) of known sequences deposited in the GenBank (**Table 2**). The PCR was performed using the mix buffer GoTaq® G2 Flexi DNA Polymerase (Promega, USA) following manufacturer's instructions. The reactions were run in the SimpliAmp™ thermal cycler (Applied Biosystems, USA). The PCR parameters were: an initial denaturation step for 1 min at 94°C, followed by 35 cycles at 94°C for 30 s, 59.9 to 65.5°C (depending on the primer sequence, **Table 2**) for 30 s; and 72°C for 1 min, with a final extension of 5 min at 72°C. To confirm the amplification of the fragments, the reactions were analyzed using 1.5% agarose gel electrophoresis. The PCR products were inserted into the cloning vector pCR™4-TOPO® TA (Invitrogen, USA). The transformed vector was used to transform electrocompetent *Escherichia coli* DH5α.

Purification and Sequencing of the Fragments of Interest

The transformed *E. coli* colonies were selected and cultured in 3 mL of Luria broth (LB) medium containing kanamycin antibiotic. The resistant bacteria were used to isolate the vector and the fragment was purified using the commercial kit Illustra plasmidPrep Mini Spin (GE Healthcare, USA) following the manufacturer's instructions. The purified samples of ten cloning fragments were submitted to sequencing by BigDye® Terminator v3.1 Cycle Sequencing (Applied Biosystems, USA) following the manufacturer's instructions. The reaction mixtures were incubated in a SimpliAmp™ thermal cycler (Applied Biosystems, USA) under the following parameters: 96°C for 1 min, 35 cycles of 96°C for 1 min, 50°C for 5 s, 60°C for 4 min.

One more purification was carried out using the BigDye® XTerminator™ Purification (Applied Biosystems, USA) according to the manufacturer's instructions. Following the purification, the samples were submitted to sequencing using an Applied Biosystems 3,500 Genetic Analyzer® automatic sequencer (Life Technologies, USA). The sequences were analyzed using the Vector NTI and deposited in GenBank®.

Gene Expression Analysis by qRT-PCR

The primers used for qRT-PCR were designed using the previous sequence of the cloned fragments and Primer3 online tool (www.bioinfo.ut.ee/primer3-0.4.0/). The primers used in this study are presented in **Table 2**. The qRT-PCR was run in a Stratagene® Mx3005P™ Real-Time PCR System (Agilent Technologies, USA) using SYBR® Green PCR Master Mix (Applied Biosystems, USA). The amplification conditions were 95°C for 10 min, 40 cycles at 95°C for 15 s, and 60°C for 60 s followed by the conditions needed to calculate the melting curve. All the reactions were performed in triplicate.

Data Analysis

The normality distribution of flow-cytometry data was evaluated by the Shapiro-Wilk test. None of the evaluated parameters presented normal distribution, thus, the data were submitted to the Kruskal-Wallis non-parametric test followed by Dunn's all-pairwise multiple comparisons, with a significance level of 5%. The results of ROS production and DFI were expressed as the mean fluorescent intensity and mean percentage, respectively, ± standard error of the means (SEM).

The cycle threshold (Ct) values were exported from the detection software MxPro™ v.4.1 (Stratagene, Waldbronn, Germany) to Excel files to facilitate the descriptive analysis and the values transformations required by some software. The specific analyses of stability of gene expression were performed using the comparative delta-Ct (dCt) method (Silver et al., 2006) and geNorm (Vandesompele et al., 2002), NormFinder (Andersen et al., 2004), BestKeeper (Pfaffl et al., 2004), and RefFinder (http://150.216.56.64/referencegene.php?type=reference) statistical approaches as previously reported (Yang et al., 2013; Taki et al., 2014). The results were expressed as average of standard deviation (SD) to dCt method; average expression stability value to geNorm and NormFinder; Pearson's correlation coefficient ([r]) to BestKeeper; and geometric mean of ranking values to RefFinder.

RESULTS

Cloning and Sequencing of RGs

Except for *atp1a* gene, which was previously known (Silveira et al., 2018), all the other seven silverside *O. humensis* candidate RGs analyzed in this study were cloned and partially sequenced for the first time. Furthermore, the new sequences were deposited in the GenBank® (**Table 1**). The sequence length of *18s* was 640 base pairs (bp). The sequence lengths of *actb*; *ef1a*; *eif3g*; *gapdh*; *h3a*; and *tuba* cloned fragments from *O. humensis* were 350, 638, 435, 325, 222, and 527 bp, respectively, that encoded 116, 212, 144, 108, 73, and 176 amino acids, respectively. The cloned fragments of *tuba* belongs to the open reading frame (ORF) +1, the fragments of *actb* and *eif3g* are part of the ORF +2, and the fragments of *ef1a*, *gapdh*, and *h3a* compose the ORF +3.

TABLE 2 | Primers information of the candidate reference genes in *Odontesthes humensis*.

Gene symbol	Sequence 5′ → 3′	Average T_m (°C)	Product size (bp)	E (%)	R^2	Use
18s	GCCGGTACAGTGAAACTGCGAATG	62	716	–	–	Cloning
	TTTCAAAGTAAACGCTTCGGACCCCG					
	AAACGGCTACCACATCCAAG	60	112	100.1	0.998	qRT-PCR
	CAATTACAGGGCCTCGAAAG					
actb	AGGAGCACCCWGTCCTGCT	64.3	448	–	–	Cloning
	ATGACCTGTCCGTCRGGCAG					
	AGGCTGTGCTGTCCCTGTAT	60	104	108.2	0.993	qRT-PCR
	CAGGGCGTAACCCTCATAGA					
ef1a	GAGCGTGAGCGTGGTATCACC	61.5	680	–	–	Cloning
	ACAGACTTCACCTCAGTGGTCAGGTTG					
	CGTTTCGAGGAAATCCAAAA	60	141	104.5	0.999	qRT-PCR
	CTTGAACCAGCCCATCTTGT					
eif3g	AAAGAACTGGAAGAARTTTGGCAACTC	59.9	501	–	–	Cloning
	GGYCTGAAGAGCTCCTGCA					
	GTATTTGCAAAGGCGACCAT	60	118	93.9	0.986	qRT-PCR
	GCAGGCTTGTCTTTGTCTCC					
gapdh	GTATGACTCCACCCACGGMCG	65.5	400	–	–	Cloning
	GTAGGCGTGRACKGTGGTCATGAG					
	GGTGGTGCCAAGAGAGTCAT	60	124	90.8	0.989	qRT-PCR
	TAGTTGTGCAGGAGGCATTG					
h3a	AGGAAGCAGCTGGCCACC	60	285	–	–	Cloning
	GGCGCACAGGTTGGTGTCCTC					
	CGTGGCTCTGAGAGAGATCC	60	108	105.3	0.995	qRT-PCR
	TCGGTCTTGAAGTCCTGAGC					
atp1a	AACCCCAGAGATGCCAA	63	1010	–	–	Cloning
	AAGGCACAGAACCACCA					
	ATACCGGGGCAGTAGGAGAG	60	150	114	0.993	qRT-PCR
	CAGTATCGTGGTGGTGCAGT					
tuba	ACTCCATCCTGACCACCCACACCACC	69	589	–	–	Cloning
	GTGGTGTTGCTCAGCATGCACAC					
	ATGAGCAGCTTTCGGTGTCT	60	119	101.8	0.983	qRT-PCR
	ATCACCACGGAACAGTAGGC					

E, efficience; bp, base pair; R^2, Pearson coefficient of determination; Tm, melting temperature.

Flow-Cytometry Analysis

The exposure of silversides for 24 h to both Roundup® and seawater caused an increase in the ROS production in erythrocytes ($P < 0.05$) when compared to the control group (**Figure 1A**). Similarly, the Roundup® and seawater treatments increased ($P < 0.05$) the DFI of the erythrocytes in the exposed silversides (**Figure 1B**). The increase in ROS production and DNA damage in the treated fish confirmed that both Roundup® and seawater treatments negatively affected the physiology of silversides, thereby generating a systemic effect on fish.

The Comparative dCt Method Analysis

The comparative dCt method was used to select the most stable RG. A low average of SD value represented a low expression variance, or high stability. The analysis revealed the following order of stability of evaluated genes, from higher to lower, in water with controlled quality parameters (control group): *h3a* (2.57) > *ef1a* (2.82) = *actb* (2.82) > *eif3g* (3.00) > *18s* (3.21) > *gapdh* (3.66) > *tuba* (3.77) > *atp1a* (4.98) (**Figure 2A**). The ranking order of stability of the evaluated genes upon exposure of silversides to Roundup® treatment was *h3a* (2.74) > *actb* (2.79) > *ef1a* (2.95) > *eif3g* (3.03) > *18s* (3.51) > *tuba* (3.82) > *gapdh* (3.95) > *atp1a* (6.18) (**Figure 2B**). In saline condition, the seawater treated group presented the following stability order: *h3a* (2.62) > *eif3g* (2.91) > *ef1a* (3.00) > *actb* (3.19) > *18s* (3.72) > *tuba* (3.84) > *gapdh* (3.90) > *atp1a* (5.05) (**Figure 2C**).

GeNorm Analysis

The geNorm analysis was carried out for the result data set after transformation of the Ct values into relative quantities through the $2^{(\text{minimum Ct value in a set sample} - \text{Ct value of a sample})}$ formula and compare pairwise variation (SD values) for each gene pair. Then, the geometric mean of SD values was used to calculate the *M*-value. The generation of low average expression stability

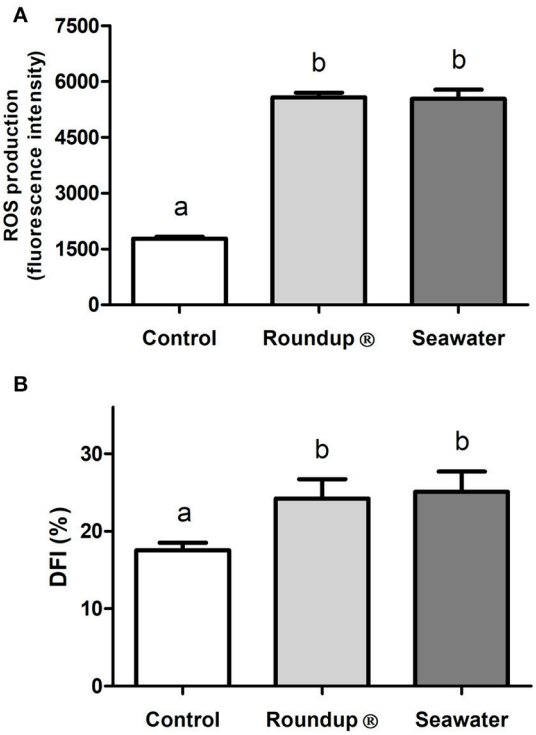


FIGURE 1 | Effects in erythrocytes of *Odontesthes humensis* of exposures to Roundup® and to seawater. Effects of exposures to Roundup® Transorb [10 mg.L^{-1} (acid equivalent) of glyphosate] and to seawater (30 ppt) for 24 h on ROS (Reactive Oxygen Species) production **(A)** and in DFI (DNA Fragmentation Index, %) **(B)**. Mean ± SEM, $n = 9$. Different letters represents significant statistical difference between the measurements of the treatment groups ($P < 0.05$).

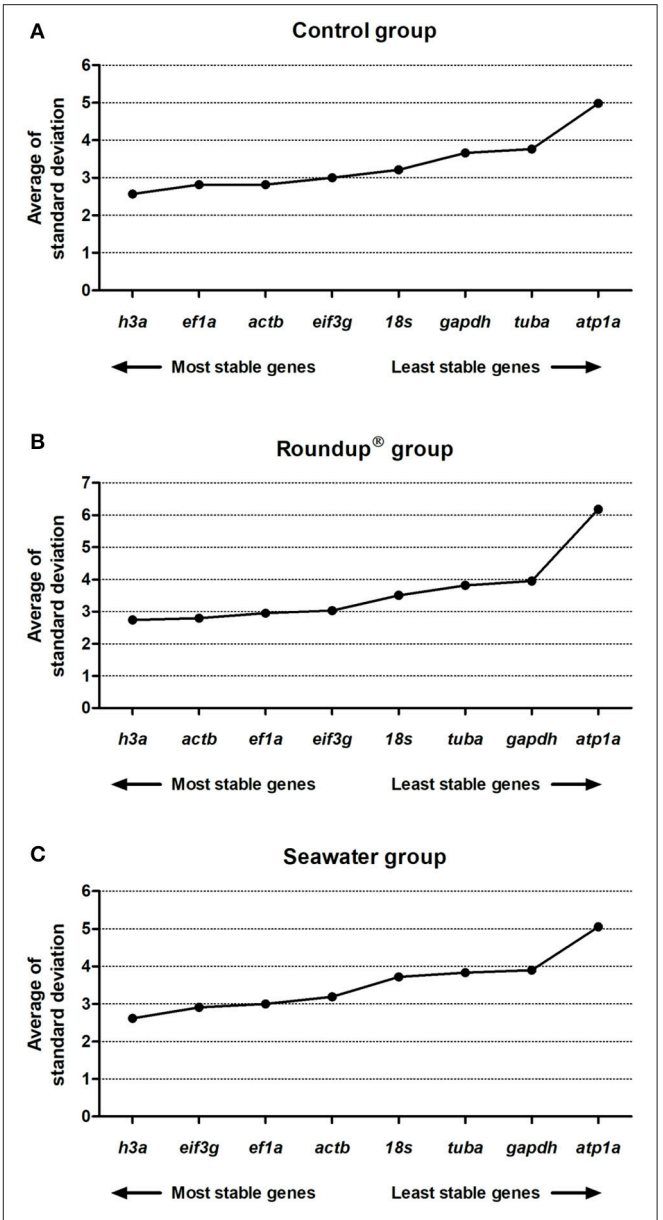


FIGURE 2 | Stability analysis of the candidates reference genes in *Odontesthes humensis* calculated by dCt method. Gene expression stability in silversides exposed to water with quality parameters controlled **(A)**; to Roundup® Transorb [10 mg.L^{-1} (acid equivalent) of glyphosate] **(B)**; and to seawater (30 ppt) **(C)** across the different treatments for 24 h. The most stable genes are displayed on the left, and the least stable genes are displayed on the right of the x-axis.

represents a low variance. The geNorm analysis revealed the following ranking order of stability of genes in the control group: *h3a*/*actb* (1.14) > *eif3g* (1.62) > *ef1a* (1.85) > *18s* (2.19) > *tuba* (2.53) > *gapdh* (2.81) > *atp1a* (3.35) (**Figure 3A**). In the Roundup® group, the sequence of stability of the evaluated genes from the highest to the lowest was *h3a*/*actb* (1.19) > *eif3g* (1.43) > *ef1a* (1.61) > *tuba* (2.09) > *18s* (2.44) > *gapdh* (2.77) > *atp1a* (3.62) (**Figure 3B**). When exposed to seawater, the silversides genes exhibited the following stability ranking order: *h3a*/*eif3g* (1.48) > *actb* (1.68) > *ef1a* (1.85) > *tuba* (2.36) > *18s* (2.74) > *gapdh* (3.02) > *atp1a* (3.53) (**Figure 3C**).

NormFinder Analysis

The NormFinder was applied to analyze the most stable evaluated genes. A low average of expression stability represents a low variance. The NormFinder analysis revealed the ranking order of stability in the control group to be *h3a* (0.79) > *ef1a* (1.26) > *actb* (1.63) > *eif3g* (1.95) > *18s* (2.06) > *gapdh* (2.79) > *tuba* (3.05) > *atp1a* (4.55) (**Figure 4A**). For the Roundup® group, the ranking of expression stability was *h3a* (0.89) > *actb* (1.23) > *ef1a* (1.43) > *eif3g* (1.79) > *18s* (2.15) > *tuba* (2.89) > *gapdh* (3.01) > *atp1a* (5.83) (**Figure 4B**). When exposed to the seawater, the ranking order of RGs was *h3a* (0.74) > *ef1a* (1.53) > *eif3g* (1.61) > *actb* (2.19) > *18s* (2.70) > *tuba* (2.97) > *gapdh* (3.04) > *atp1a* (4.54) (**Figure 4C**). In addition, the best combination of genes for the experimental conditions was evaluated. A comparison of the gene expression in silversides of the control and the Roundup® groups, and of the control and the seawater groups revealed the best combination of genes to be *h3a* and *ef1a* with combined stability values of 0.28 and 0.29, respectively. The best suitable combination of two genes for all

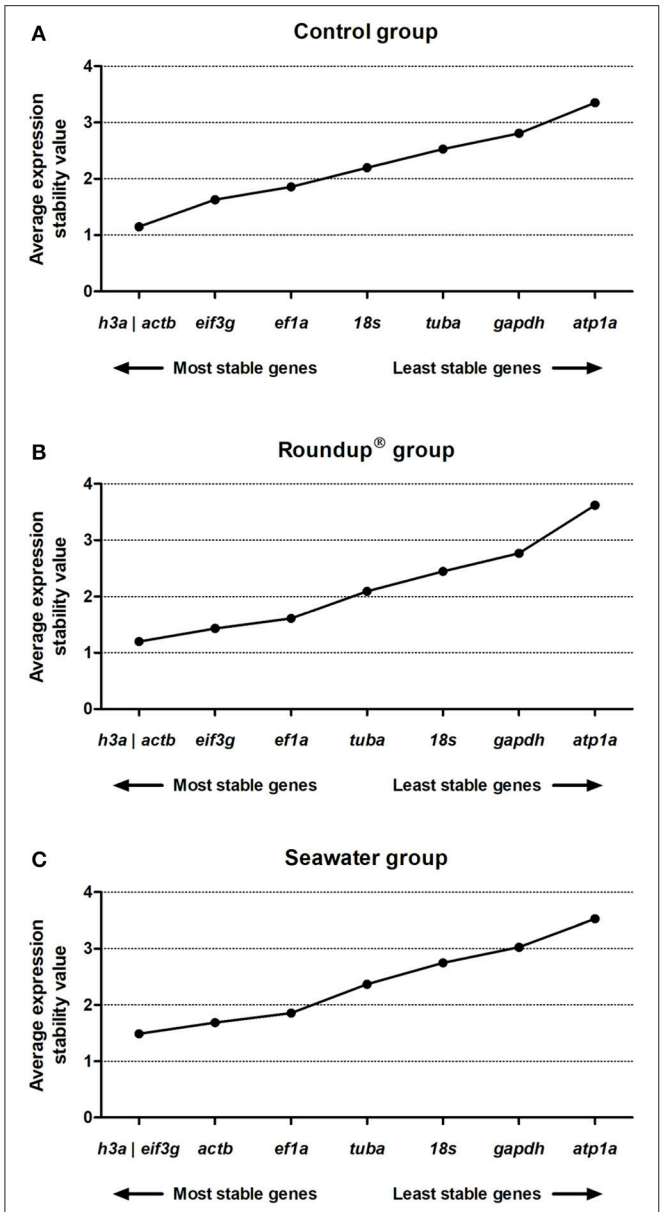


FIGURE 3 | Stability analysis of the candidates reference genes in *Odontesthes humensis* calculated by geNorm algorithm. Gene expression stability in silversides exposed to water with quality parameters controlled **(A)**; to Roundup® Transorb [10 mg.L^{-1} (acid equivalent) of glyphosate] **(B)**; and to seawater (30 ppt) **(C)** across the different treatments for 24 h. The most stable genes are displayed on the left, and the least stable genes are displayed on the right of the x-axis.

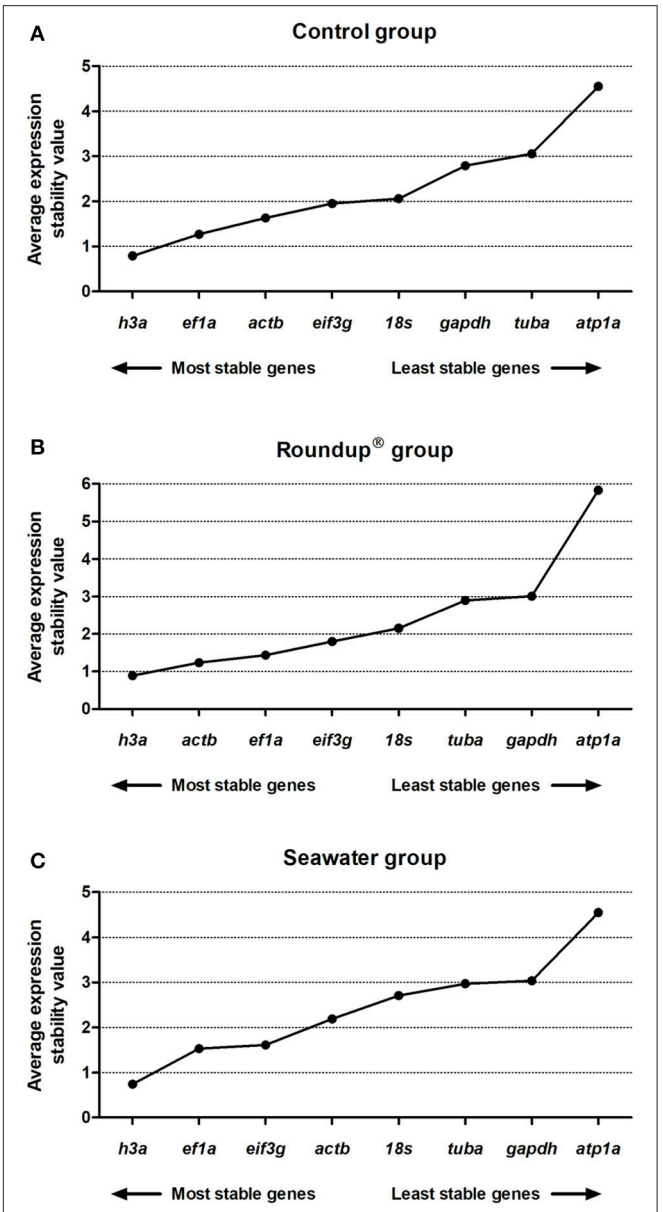


FIGURE 4 | Stability analysis of the candidates reference genes in *Odontesthes humensis* calculated by NormFinder algorithm. Gene expression stability in silversides exposed to water with quality parameters controlled **(A)**; to Roundup® Transorb [10 mg.L^{-1} (acid equivalent) of glyphosate] **(B)**; and to seawater (30 ppt) **(C)** across the different treatments for 24 h. The most stable genes are displayed on the left, and the least stable genes are displayed on the right of the x-axis.

three tested conditions was also *h3a* and *ef1a* with a combined stability value of 0.26.

BestKeeper Analysis

The BestKeeper analysis provided two-interpretation-ways to rank the gene stability: one based on the samples SD values of Ct and other based on the Pearson's correlation of expression among the genes. Thus, the genes with low SDs and high correlation with the BestKeeper index (indicating high similarity among the expression levels of the RGs) are ranked as the most stable genes. To construct the consensus comprehensive ranking, the SD values were used, which is a more conservative approach. However, to construct the rankings of the most stable genes with BestKeeper, [r] and *P*-values of Pearson's correlation were employed. This way is more sophisticated and statistically robust, as it results in the rankings more similar to the ones obtained using other algorithms.

The analyses based on Pearson's correlation revealed the following ranking order of stability in the control group: *h3a* (0.93) > *actb* (0.90) > *ef1a* (0.80) > *eif3g* (0.79) > *tuba* (0.76) > *18s* (0.62) > *gapdh* (0.46) > *atp1a* (0.24) (**Figure 5A**). The genes presented a positive correlation with the BestKeeper index with $P = 0.001$ except atp1a with $P = 0.05$. In the Roundup® group the ranking of the best RG was *h3a* (0.83) = *actb* (0.83) > *ef1a* (0.79) > *eif3g* (0.78) > *tuba* (0.70) > *gapdh* (0.49) > *18s* (0.41) > *atp1a* (0.26) (**Figure 5B**). All correlation analysis presented $P = 0.001$ except *atp1a* with $P = 0.08$. In the seawater group the ranking order from the most to the least stable genes was *h3a* (0.91) > *ef1a* (0.89) > *actb* (0.87) > *eif3g* (0.83) > *tuba* (0.76) > *gapdh* (0.49) > *18s* (0.23) > *atp1a* (0.22) (**Figure 5C**). The genes *atp1a* and *18s* exhibited $P = 0.1$ and $P = 0.08$, respectively. All the other genes displayed positive correlations with $P = 0.001$.

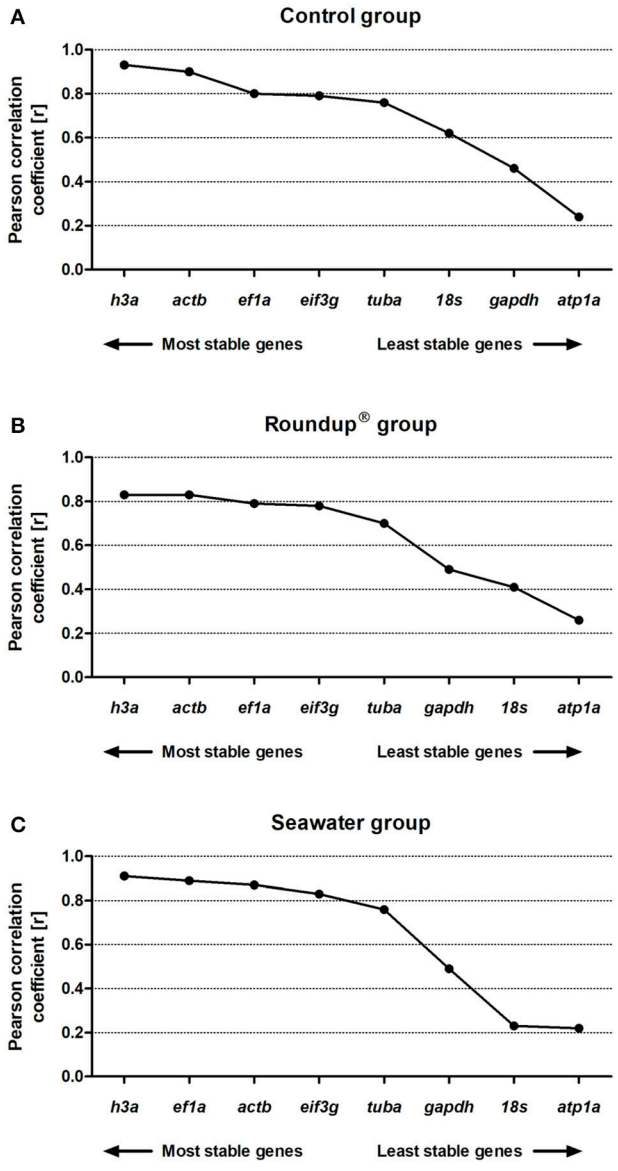


FIGURE 5 | Stability analysis of the candidates reference genes in *Odontesthes humensis* calculated by BestKeeper algorithm. Gene expression stability in silversides exposed to water with quality parameters controlled **(A)**; to Roundup® Transorb [10 mg.L^{-1} (acid equivalent) of glyphosate] **(B)**; and to seawater (30 ppt) **(C)** across the different treatments for 24 h. The most stable genes are displayed on the left, and the least stable genes are displayed on the right of the x-axis.

RefFinder Analysis

The RefFinder is an online tool used to construct a consensus comprehensive ranking of stability of RGs among all the other methods using the calculation of geometric mean for the ranks calculated by each of other algorithms. Candidate genes with the lowest geometric mean are most stable.

Besides the graphical rankings of RefFinder, it presented the consensus results discriminated by treatment group (control, Roundup®, and seawater) by developmental stages (adults, juveniles, and immature) and tissue types (brain, gills, hepatopancreas, and kidney).

In the control group, the most stable gene across the different ages and tissues was *h3a*. The only exception was the brain samples, where the most stable gene was *tuba*. The least stable gene across the developmental stages and tissues was *atp1a* (**Table 3**).

In the Roundup® group the most stable gene across different ages was *h3a* in adults and immature and *ef1a* for juveniles. In different tissues, the most stable gene was also *h3a* in gills and hepatopancreas, but *actb* was most stable in brain and kidney. As expected, the gene with the lowest stability was *atp1a* across different ages and tissues, except in hepatopancreas, where *tuba* exhibited the lowest stability (**Table 4**).

The *h3a* gene displayed the highest stability in different developmental stages in the seawater group. It was also the most stably expressed gene in the gills and hepatopancreas, while *actb*

TABLE 3 | Consensus stability ranking by RefFinder of the candidate reference genes in *Odontesthes humensis* under normal conditions.

Control group		Ranking							
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Ages	Adults	*h3a*	*actb*	*18s*	*ef1a*	*eif3g*	*gapdh*	*tuba*	*atp1a*
		1.32	2.00	3.16	3.41	4.68	4.74	6.96	7.74
	Juveniles	*h3a*	*eif3g*	*ef1a*	*18s*	*actb*	*gapdh*	*tuba*	*atp1a*
		1.41	2.45	3.13	3.34	3.66	4.56	7.24	7.74
	Immatures	*h3a*	*18s*	*ef1a*	*actb*	*eif3g*	*gapdh*	*tuba*	*atp1a*
		1.41	2.28	2.51	2.99	4.16	6.48	6.48	8.00
Tissues	Brain	*tuba*	*h3a*	*eif3g*	*18s*	*actb*	*ef1a*	*gapdh*	*atp1a*
		1.32	1.68	2.45	4.61	4.68	5.18	7.00	8.00
	Gills	*h3a*	*actb*	*eif3g*	*ef1a*	*18s*	*gapdh*	*tuba*	*atp1a*
		1.00	2.91	3.16	3.46	3.50	5.42	7.00	8.00
	Hepatopancreas	*h3a*	*actb*	*gapdh*	*18s*	*eif3g*	*ef1a*	*tuba*	*atp1a*
		1.50	2.21	2.71	3.46	4.40	4.43	7.00	8.00
	Kidney	*h3a*	*actb*	*eif3g*	*ef1a*	*18s*	*gapdh*	*tuba*	*atp1a*
		1.57	2.45	2.63	3.16	3.60	5.48	6.74	8.00

TABLE 4 | Consensus stability ranking by RefFinder of the candidate reference genes in *Odontesthes humensis* exposed to Roundup®.

Roundup® exposure (10 mg.L^{-1}, a.e.)		Ranking							
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Ages	Adults	*h3a*	*eif3g*	*actb*	*18s*	*ef1a*	*tuba*	*gapdh*	*atp1a*
		1.57	2.00	2.28	3.46	4.68	5.23	7.00	8.00
	Juveniles	*ef1a*	*actb*	*18s*	*h3a*	*eif3g*	*gapdh*	*tuba*	*atp1a*
		1.41	1.57	2.78	3.66	4.95	5.86	6.48	8.00
	Immatures	*h3a*	*actb*	*ef1a*	*18s*	*eif3g*	*tuba*	*gapdh*	*atp1a*
		1.19	2.21	2.71	3.34	4.23	6.24	6.74	8.00
Tissues	Brain	*actb*	*eif3g*	*18s*	*h3a*	*tuba*	*ef1a*	*gapdh*	*atp1a*
		1.32	1.68	3.98	4.16	4.28	4.61	5.86	8.00
	Gills	*h3a*	*tuba*	*eif3g*	*actb*	*18s*	*ef1a*	*gapdh*	*atp1a*
		1.41	2.66	2.78	3.46	4.30	5.18	5.24	8.00
	Hepatopancreas	*h3a*	*actb*	*eif3g*	*18s*	*gapdh*	*ef1a*	*atp1a*	*tuba*
		1.32	2.21	2.82	3.13	4.95	5.48	6.73	7.24
	Kidney	*actb*	*h3a*	*eif3g*	*gapdh*	*ef1a*	*18s*	*tuba*	*atp1a*
		1.68	1.86	2.45	3.50	4.16	6.09	6.24	8.00

TABLE 5 | Consensus stability ranking by RefFinder of the candidate reference genes in *Odontesthes humensis* exposed to seawater.

Seawater exposure (30 ppt)		Ranking							
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Ages	Adults	*h3a*	*eif3g*	*18s*	*actb*	*ef1a*	*gapdh*	*tuba*	*atp1a*
		1.57	2.00	2.59	3.64	4.43	5.66	6.45	7.11
	Juveniles	*h3a*	*ef1a*	*eif3g*	*18s*	*actb*	*gapdh*	*tuba*	*atp1a*
		1.19	2.00	3.00	4.30	4.43	5.23	6.16	7.74
	Immatures	*h3a*	*eif3g*	*ef1a*	*18s*	*actb*	*tuba*	*gapdh*	*atp1a*
		1.57	2.28	2.83	3.16	3.16	6.00	7.00	8.00
Tissues	Brain	*actb*	*h3a*	*eif3g*	*gapdh*	*ef1a*	*18s*	*tuba*	*atp1a*
		1.19	1.41	3.22	4.23	5.23	5.66	6.24	8.00
	Gills	*h3a*	*ef1a*	*actb*	*eif3g*	*18s*	*tuba*	*gapdh*	*atp1a*
		1.32	2.63	2.91	2.99	4.30	5.23	6.24	8.00
	Hepatopancreas	*h3a*	*actb*	*eif3g*	*gapdh*	*18s*	*ef1a*	*atp1a*	*tuba*
		1.68	2.00	2.59	4.12	4.30	5.44	6.26	6.45
	Kidney	*h3a*	*eif3g*	*actb*	*18s*	*gapdh*	*ef1a*	*tuba*	*atp1a*
		1.32	1.68	3.22	3.34	4.23	6.24	6.74	8.00

was the most stable in brain and kidney. The gene with the lowest stability across the ages and tissues was *atp1a*, except in hepatopancreas, where *tuba* exhibited the lowest stability in expression (**Table 5**).

In the control group, the consensus comprehensive ranking constructed by RefFinder was *h3a* (1.32) > *ef1a* (2.38) > *actb* (2.59) > *18s* (3.34) > *eif3g* (4.12) > *gapdh* (5.63) > *tuba* (6.96) > *atp1a* (7.74) (**Figure 6A**). In the Roundup® group, the ranking order of stability was *h3a* (1.19) > *actb* (1.86) > *18s* (3.50) > *ef1a* (3.66) > *eif3g* (3.72) > *tuba* (5.96) > *gapdh* (6.74) > *atp1a* (8.00) (**Figure 6B**). The consensus ranking of stability constructed based on the expression data of the seawater group was: *h3a* (1.19) > *eif3g* (2.06) > *ef1a* (3.13) > *18s* (3.50) > *actb* (4.12) > *tuba* (5.96) > *gapdh* (6.44) > *atp1a* (8.00) (**Figure 6C**). The final consensus stability analysis, considering all methods and grouping the different tissues, ages, and treatments groups, revealed the following ranking order: *h3a* (1.19) > *ef1a* (2.63) > *actb* (2.99) > *eif3g* (3.22) > *18s* (3.34) > *gapdh* (6.48) > *tuba* (6.48) > *atp1a* (8.00) (**Figure 6D**).

DISCUSSION

In the present study, it was used individual methods (comparative dCt, geNorm, NormFinder, and BestKeeper) to obtain independent rankings of gene expression stability of candidate RGs in *O. humensis* and an algorithm that deduces a consensus ranking among all methods (RefFinder) such as presented in recent studies about RGs (Taki et al., 2014; Xu et al., 2014; Huang et al., 2015). Different algorithms can give different results depending of the set of analyzed genes and experimental variables (Yang et al., 2013; Zhang et al., 2013; Liu et al., 2014; Purohit et al., 2016) and no one statistical approach can cover all variables (Taki et al., 2014). For these reasons, the use of more than one algorithm is indicated. The different results presented by each algorithm are natural and expected due to their distinct statistical approach to construct the rankings (Liu et al., 2014). For example, the dCt method compares the relative expression of gene pairs with their SD values to get the most stable. GeNorm goes beyond, it calculates the *M*-value from the SD values and the geometric mean, and compares all gene pairs results. BestKeeper also starts from the SD values of gene expression and a geometric mean of the genes with the most stable Ct value are used to calculate the BestKeeper index. Then, it is calculated the Pearson's correlation coefficient ([r]) and the *P*-value to determine the similarity between the candidate RGs expression. The NormFinder is based in analysis of variance (ANOVA) and it is the unique of the used algorithms that evaluates the intragroup variation (i.e., replicates of a same treatment group) besides the intergroup variation (i.e., control vs. Roundup® or seawater groups), which also used by the other pairwise comparisons approaches, and combine both variations into a stability value. In this study, some differences were obtained in the ranking orders occupied by the genes in result of each algorithm analysis mainly from the second to the penultimate positions (**Figures 2**–**5**). Due to this, the use of the analysis by the ReFinder, which calculates the geometric mean for the previously obtained ranks, was important to a consensual unification of the discrepant results.

In the current study, the most stable gene across the treatment groups, independent of the method of analysis, including the final consensus RefFinder analysis, was *h3a*. This gene was also the most stable across different ages, except in the juveniles from the Roundup® group. However, in the consensus analysis of the tissues samples revealed *h3a* to be not as stable as in

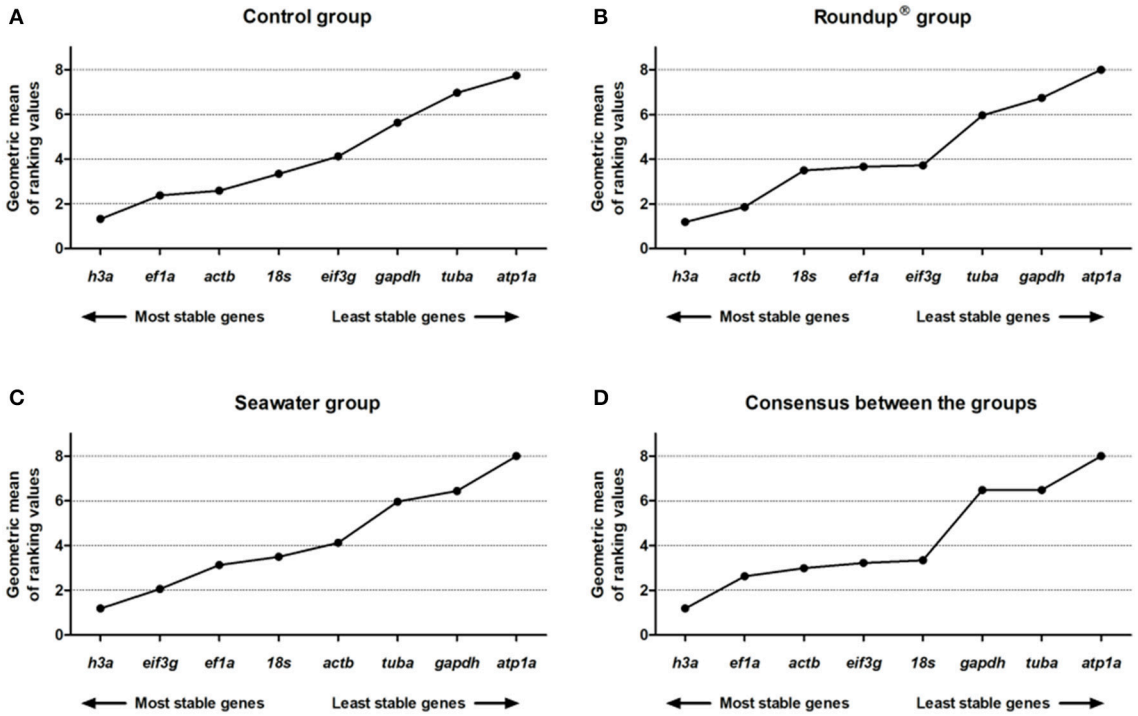


FIGURE 6 | Consensus stability analysis of the candidates reference genes in *Odontesthes humensis* calculated by RefFinder algorithm. Gene expression stability in silversides exposed to water with quality parameters controlled **(A)**; to Roundup® Transorb [10 mg.L^{-1} (acid equivalent) of glyphosate] **(B)**; and to seawater (30 ppt) **(C)** across the different treatments for 24 h; and final consensus stability analysis between all methods and grouping the three treatments groups **(D)**. The most stable genes are displayed on the left, and the least stable genes are displayed on the right of the x-axis.

the age analysis. For example, the results of brain analysis demonstrated this gene not to be the most stable in all treatments. Moreover, *h3a* did not exhibit the highest stability in the kidney of Roundup® group. The *h3a* gene has already been evaluated in studies for validation of RGs showing high stability (Taylor et al., 2007; Koenigstein et al., 2013) and can be used to normalize qRT-PCR reactions (Ramachandra et al., 2008). However, up to now, no study has reported this gene as the candidate RG in a fish species. In the final consensus ranking, the second most stable gene was *ef1a*. All performed analysis revealed this gene to have high stability, occupying at least the third position of the constructed rankings, except in consensus ranking of the Roundup® group constructed by RefFinder, where it occupied the fourth position. A deviation in the ranking order of stability is natural owing to the use of different algorithms. Studies on stability of candidates RGs in different experimental protocols have revealed that *ef1a* is usually the most stable gene of fishes (Jorgensen et al., 2006; Tang et al., 2007, 2012; Infante et al., 2008; Liman et al., 2013; Hu et al., 2014).

The studies that evaluated the gene expression in the South American silversides have used *actb* as the internal control for qRT-PCR (Strobl-Mazzulla et al., 2005; Karube et al., 2007; Fernandino et al., 2008, 2011; Majhi et al., 2009; Miranda et al., 2009; Blasco et al., 2010; Gómez-Requeni et al., 2012; Pérez et al., 2012; González et al., 2015). In the present study, RefFinder analysis focused on the individual tissues revealing *actb* to be the most stable gene in brain and kidney from the Roundup® group and in the brain from the seawater group. Only the geNorm analysis of stability, based on control and Roundup® groups, revealed *actb* as the most stable gene. However, our results showed that *actb* was not the most stable gene. The RefFinder final consensus comprehensive ranking pointed this gene as the third most stable in *O. humensis*. There is still much debate on the use of *actb* to normalize qRT-PCR reactions in fishes. Some authors have reported that this gene is the least stable in some fish species (Jorgensen et al., 2006; Filby and Tyler, 2007; Yang et al., 2013; Hu et al., 2014; Chapman and Waldenström, 2015). However, the studies that present *actb* as a good RG have also reported that its high stability in fishes is tissue dependent within a species (Zheng and Sun, 2011; Zhang et al., 2013; Sun and Hu, 2015), such as in *O. humensis*. This does not mean that the studies using *actb* as the RG should not be considered. However, future studies should pay more attention to the choice of RG.

The *18s* gene has been a subject of discussions on its use as the RG in fish models. This gene has been shown as the most or to be among the most stable candidates in some fish species (Jorgensen et al., 2006; Filby and Tyler, 2007; Tang et al., 2007; Small et al., 2008; Yang et al., 2013; Liu et al., 2014). However, in other species, this gene is considered unsuitable for use as an internal control for qRT-PCR (Infante et al., 2008; Hu et al., 2014; Purohit et al., 2016). In *O. humensis* from control, Roundup®, and seawater groups, and across different ages and tissues, *18s* generally

held the intermediary position in the rankings, attracting little attention for its use in normalization methods. The *gapdh* gene, as well as *actb* and *18s*, has presented discrepant results in the literature that reports its expression stability. The *gapdh* has already been reported as one of the most stable genes in half-smooth tongue sole (Liu et al., 2014). Furthermore, its expression is usually stable depending on the evaluated tissue, being able to occupy both extremities of rankings of stability (Yang et al., 2013; Zhang et al., 2013). So, generally, *gapdh* is refuted as a suitable gene for an efficient normalization (Jorgensen et al., 2006; Filby and Tyler, 2007; Tang et al., 2007, 2012; Infante et al., 2008; Small et al., 2008; Ahi et al., 2013; Liman et al., 2013; Hu et al., 2014; Chapman and Waldenström, 2015). In the present study, *gapdh* appeared to be the second most unstable gene, only after *atp1a* gene.

The *tuba* gene usually presents high stability in studies on validation of candidates RGs. This gene was the second most stable after analysis by dCt and NormFinder in Nile tilapia under normal conditions (Yang et al., 2013). This gene also exhibited high stability in intestine and liver from Japanese flounder at 24 and 72 h after viral infection, respectively, as well as in muscle of turbot after 72 h post-infection (Zhang et al., 2013). Even with *tuba* gene presenting higher stability in the brain of silversides from the control group, this gene did not maintain this expression pattern. In the brain of silversides from other treatment groups, *tuba* had low stability levels. Across other tissues and different developmental stages, the *tuba* gene frequently occupied the last places of the final ranking, even behind *atp1a*, as well as across treatment groups and different algorithms used. The final consensus comprehensive ranking constructed by RefFinder presented *tuba* as the most unstable gene in *O. humensis*, disregarding *atp1a*. Finally, the control gene *atp1a*, despite being a constitutive gene, was the most unstable, as expected. It occupied the last position in the rankings of stability constructed by all algorithms used to compare the three treatment groups. This gene also exhibited lower stability across different ages and tissues from silversides from all treatments, except in the hepatopancreas of the Roundup® and seawater groups.

In the present study, a set of candidates of RGs was analyzed for the first time in *O. humensis*. The genes *atp1a*, *tuba*, *gapdh*, and *18s* were the least stable and highly unsuitable to be used for qRT-PCR normalization in studies in silverside *O. humensis*. The three most stable genes were *h3a*, *ef1a*, and *actb*. The discordance observed among the results of analyses of the candidate RGs in different fish species, tissues, developmental stages, and experimental conditions, it is evident the importance of identification of efficient RGs for each experimental design instead of the indiscriminate use of traditional genes. Furthermore, the gene *h3a* exhibited a greater potential to be used as a RG in *O. humensis* in comparison with other traditional genes, such as *actb* and *ef1a*. Perhaps the *h3a* gene had the same potential to be used as the best RG in other fish species, since the stability of this gene has never been evaluated in other fishes. Therefore, future studies warrant the evaluation of *h3a* as candidate RG in other fish species, besides *O. humensis* or silversides.

AUTHOR CONTRIBUTIONS

TS and VC are responsible for experimental design, data analysis, manuscript writing. TS, MR, and RR were responsible for fish acclimation and maintenance. TS, BB, IL, LS, MR, RR, and WD were responsible for the biological collections. TS, BB, IL, LS, WD, and VC were responsible for the molecular biology, from RNA extraction to sequencing and qRT-PCR analysis. TC and FS were also responsible for qRT-PCR analysis. AV, CC, and DM were responsible for the flow-cytometry analysis.

FUNDING

This study was supported by the Ministério da Ciência, Tecnologia e Inovação/Conselho Nacional de Desenvolvimento Científico e Tecnológico (Edital Universal #472210/2013-0 and #422292/2016-8) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (AUXPE 2900/2014). TS, WD, MR, and DM are individually supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. AV, CC, TC, FS, RR, and VC are also individually supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico.

ACKNOWLEDGMENTS

We are greatly thankful to Mr. Nilton Link and Mr. Vítor Colvara for the help in the maintenance of animals.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00075/full#supplementary-material

REFERENCES

Ahi, E. P., Guðbrandsson, J., Kapralova, K. H., Franzdóttir, S. R., Snorrason, S. S., Maier, V. H., et al. (2013). Validation of reference genes for expression studies during craniofacial development in arctic charr. *PLoS ONE* 8:e66389. doi: 10.1371/journal.pone.0066389

Andersen, C. L., Jensen, J. L., and Ørntoft, T. F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64, 5245–5250. doi: 10.1158/0008-5472.CAN-04-0496

Armesto, P., Campinho, M. A., Rodríguez-Rúa, A., Cousin, X., Power, D. M., Manchado, M., et al. (2014). Molecular characterization and transcriptional regulation of the Na^+/K^+ ATPase α subunit isoforms during development and salinity challenge in a teleost fish, the Senegalese sole (*Solea senegalensis*). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 175, 23–38. doi: 10.1016/j.cbpb.2014.06.004

Bemvenuti, M. A. (2006). Silversides in South Brazil: morphological and ecological aspects. *Biocell* 30, 111–118.

Blasco, M., Fernandino, J. I., Guilgur, L. G., Strüssmann, C. A., Somoza, G. M., and Vizziano-Cantonnet, D. (2010). Molecular characterization of cyp11a1 and cyp11b1 and their gene expression profile in pejerrey (*Odontesthes bonariensis*) during early gonadal development. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 156, 110–118. doi: 10.1016/j.cbpa.2010.01.006

Bloom, D. D., Weir, J. T., Piller, K. R., and Lovejoy, N. R. (2013). Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of New World silversides (Atherinopsidae). *Evolution* 67, 2040–2057. doi: 10.1111/evo.12074

Bustin, S. A. (2002). Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J. Mol. Endocrinol.* 29, 23–39. doi: 10.1677/jme.0.0290023

Carriquiriborde, P., and Ronco, A. (2006). Ecotoxicological studies on the pejerrey (*Odontesthes bonariensis*, Pisces Atherinopsidae). *Biocell* 30, 97–109.

Chapman, J. R., and Waldenström, J. (2015). With reference to reference genes: a systematic review of endogenous controls in gene expression studies. *PLoS ONE* 10:e0141853. doi: 10.1371/journal.pone.0141853

Choi, C. Y., and An, K. W. (2008). Cloning and expression of Na^+/K^+-ATPase and osmotic stress transcription factor 1 mRNA in black porgy, *Acanthopagrus schlegeli* during osmotic stress. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 149, 91–100. doi: 10.1016/j.cbpb.2007.08.009

Dheda, K., Huggett, J. F., Bustin, S. A., Johnson, M. A., Rook, G., and Zumla, A. (2004). Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* 37, 112–119.

Dyer, B. S. (2006). Systematic revision of the South American silversides (Teleostei, Atheriniformes). *Biocell* 30, 69–88.

Fernandino, J. I., Hattori, R. S., Kimura, H., Strüssmann, C. A., and Somoza, G. M. (2008). Expression profile and estrogenic regulation of anti-müllerian hormone during gonadal development in pejerrey *Odontesthes bonariensis*, a teleost fish with strong temperature-dependent sex determination. *Dev. Dyn.* 237, 3192–3199. doi: 10.1002/dvdy.21731

Fernandino, J. I., Popesku, J. T., Paul-Prasanth, B., Xiong, H., Hattori, R. S., Oura, M., et al. (2011). Analysis of sexually dimorphic expression of genes at early gonadogenesis of pejerrey *Odontesthes bonariensis* using a heterologous microarray. *Sex Dev.* 5, 89–101. doi: 10.1159/000324423

Filby, A. L., and Tyler, C. R. (2007). Appropriate "housekeeping" genes for use in expression profiling the effects of environmental estrogens in fish. *BMC Mol. Biol.* 8:10. doi: 10.1186/1471-2199-8-10

Gómez-Requeni, P., Kraemer, M. N., and Canosa, L. F. (2012). Regulation of somatic growth and gene expression of the GH-IGF system and PRP-PACAP by dietary lipid level in early juveniles of a teleost fish, the pejerrey (*Odontesthes bonariensis*). *J. Comp. Physiol. B Biochem. Syst. Environ. Physiol.* 182, 517–530. doi: 10.1007/s00360-011-0640-9

González, A., Fernandino, J. I., and Somoza, G. M. (2015). Effects of 5α-dihydrotestosterone on expression of genes related to steroidogenesis and spermatogenesis during the sex determination and differentiation periods of the pejerrey, *Odontesthes bonariensis*. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 182, 1–7. doi: 10.1016/j.cbpa.2014.12.003

Harayashiki, C. A., Varela Junior, A. S., Machado, A. A., Cabrera, L. C., Primel, E. G., Bianchini, A., et al. (2013). Toxic effects of the herbicide Roundup in the guppy *Poecilia vivipara* acclimated to fresh water. *Aquat. Toxicol.* 142–143, 176–184. doi: 10.1016/j.aquatox.2013.08.006

Hiroi, J., and McCormick, S. D. (2012). New insights into gill ionocyte and ion transporter function in euryhaline and diadromous fish. *Respir. Physiol. Neurobiol.* 184, 257–268. doi: 10.1016/j.resp.2012.07.019

Hsu, H. H., Lin, L. Y., Tseng, Y. C., Horng, J. L., and Hwang, P. P. (2014). A new model for fish ion regulation: identification of ionocytes in freshwater- and seawater-acclimated medaka (*Oryzias latipes*). *Cell Tissue Res.* 357, 225–243. doi: 10.1007/s00441-014-1883-z

Hu, Q., Guo, W., Gao, Y., Tang, R., and Li, D. (2014). Reference gene selection for real-time RT-PCR normalization in rice field eel (*Monopterus albus*) during gonad development. *Fish Physiol. Biochem.* 40, 1721–1730. doi: 10.1007/s10695-014-9962-3

Huang, X., Gao, Y., Jiang, B., Zhou, Z., and Zhan, A. (2015). Reference gene selection for quantitative gene expression studies during biological invasions: a test on multiple genes and tissues in a model ascidian *Ciona savignyi*. *Gene* 576, 79–87. doi: 10.1016/j.gene.2015.09.066

Hughes, L. C., Somoza, G. M., Nguyen, B. N., Bernot, J. P., González-Castro, M., Díaz de Astarloa, J. M., et al. (2017). Transcriptomic differentiation underlying marine-to-freshwater transitions in the South American silversides *Odontesthes argentinensis* and *O. bonariensis* (Atheriniformes). *Ecol. Evol.* 7, 1–11. doi: 10.1002/ece3.3133

Infante, C., Matsuoka, M. P., Asensio, E., Cañavate, J. P., Reith, M., and Manchado, M. (2008). Selection of housekeeping genes for gene expression studies in larvae from flatfish using real-time PCR. *BMC Mol. Biol.* 9:28. doi: 10.1186/1471-2199-9-28

Jorgensen, S. M., Kleveland, E. J., Grimholt, U., and Gjoen, T. (2006). Validation of reference genes for real-time polymerase chain reaction studies in Atlantic salmon. *Mar. Biotechnol.* 8, 398–408. doi: 10.1007/s10126-005-5164-4

Karube, M., Fernandino, J. I., Strobl-Mazzulla, P., Strüssmann, C. A., Yoshzaki, G., Somoza, G. M., et al. (2007). Characterization and expression profile of the ovarian cytochrome P-450 aromatase (cyp19A1) gene during thermolabile sex determination in pejerrey, *Odontesthes bonariensis*. *J. Exp. Zool. A Ecol. Genet. Physiol.* 307, 625–636. doi: 10.1002/jez.416

Koenigstein, S., Pöhlmann, K., Held, C., and Abele, D. (2013). Ecological comparison of cellular stress responses among populations – normalizing RT-qPCR values to investigate differential environmental adaptations. *BMC Ecol.* 13:21. doi: 10.1186/1472-6785-13-21

Liman, M., Wenji, W., Conghui, L., Haiyang, Y., Zhigang, W., Xubo, W., et al. (2013). Selection of reference genes for reverse transcription quantitative real-time PCR normalization in black rockfish (*Sebastes schlegeli*). *Mar. Genomics* 11, 67–73. doi: 10.1016/j.margen.2013.08.002

Lin, C. H., Huang, C. L., Yang, C. H., Lee, T. H., and Hwang, P. P. (2004). Time-course changes in the expression of Na, K-ATPase and the morphometry of mitochondrion-rich cells in gills of euryhaline tilapia (*Oreochromis mossambicus*) during freshwater acclimation. *J. Exp. Zool. A Comp. Exp. Biol.* 301, 85–96. doi: 10.1002/jez.a.20007

Liu, C., Xin, N., Zhai, Y., Jiang, L., Zhai, J., Zhang, Q., et al. (2014). Reference gene selection for quantitative real-time RT-PCR normalization in the half-smooth tongue sole (*Cynoglossus semilaevis*) at different developmental stages, in various tissue types and on exposure to chemicals. *PLoS ONE* 9:e91715. doi: 10.1371/journal.pone.0091715

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Majhi, S. K., Hattori, R. S., Rahman, S. M., Suzuki, T., and Strüssmann, C. A. (2009). Experimentally induced depletion of germ cells in sub-adult Patagonian pejerrey (*Odontesthes hatcheri*). *Theriogenology* 71, 1162–1172. doi: 10.1016/j.theriogenology.2008.12.008

Menone, M. L., Aizpún de Moreno, J. E., Moreno, V. J., Lanfranchi, A. L., Metcalfe, T. L., and Metcalfe, C. D. (2000). Environmental contamination and toxicology PCBs and organochlorines in tissues of silverside (*Odontesthes bonariensis*) from a coastal lagoon in Argentina. *Arch. Environ. Contam. Toxicol.* 38, 202–208. doi: 10.1007/s002449910027

Miranda, L. A., Strüssmann, C. A., and Somoza, G. M. (2009). Effects of light and temperature conditions on the expression of GnRH and GtH genes and levels of plasma steroids in *Odontesthes bonariensis* females. *Fish Physiol. Biochem.* 35, 101–108. doi: 10.1007/s10695-008-9232-3

Nawata, C. M., Hirose, S., Nakada, T., Wood, C. M., and Kato, A. (2010). Rh glycoprotein expression is modulated in pufferfish (*Takifugu rubripes*) during high environmental ammonia exposure. *J. Exp. Biol.* 213, 3150–3160. doi: 10.1242/jeb.044719

Nilsen, T. O., Ebbesson, L. O., Madsen, S. S., McCormick, S. D., Andersson, E., Björnsson, B. T., et al. (2007). Differential expression of gill Na^+, K^+-ATPase alpha- and beta-subunits, Na^+, K^+, $2Cl^-$ cotransporter and CFTR anion channel in juvenile anadromous and landlocked Atlantic salmon *Salmo salar*. *J. Exp. Biol.* 210, 2885–2896. doi: 10.1242/jeb.002873

Patterson, J., Bodinier, C., and Green, C. (2012). Effects of low salinity media on growth, condition, and gill ion transporter expression in juvenile Gulf killifish, *Fundulus grandis*. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 161, 415–421. doi: 10.1016/j.cbpa.2011.12.019

Pérez, M. R., Fernandino, J. I., Carriquiriborde, P., and Somoza, G. M. (2012). Feminization and altered gonadal gene expression profile by ethinylestradiol

exposure to pejerrey, *Odontesthes bonariensis*, a South American teleost fish. *Environ. Toxicol. Chem.* 31, 941–946. doi: 10.1002/etc.1789

Petrunkina, A. M., Volker, G., Weitze, K. F., Beyerbach, M., Töpfer-Petersen, E., and Waberski, D. (2005). Detection of cooling-induced membrane changes in the response of boar sperm to capacitating conditions. *Theriogenology* 63, 2278–2299. doi: 10.1016/j.theriogenology.2004.10.008

Pfaffl, M. W., Tichopad, A., Prgomet, C., and Neuvians, T. P. (2004). Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: bestKeeper – excel-based tool using pair-wise correlations. *Biotechnol. Lett.* 26, 509–515. doi: 10.1023/B:BILE.0000019559.84305.47

Piedras, S. R. N., Fernandes, J. L. O., Motoyama, I. S., and Martins, G. B. (2009). Efeito de diferentes concentrações de salinas (NaCl) na sobrevivência de embriões de peixe – rei *Odontesthes bonariensis* e *Odontesthes humensis*. *Biotemas* 22, 235–238. doi: 10.5007/2175-7925.2009v22n3p235

Purohit, G. K., Mahanty, A., Mohanty, B. P., and Mohanty, S. (2016). Evaluation of housekeeping genes as references for quantitative real-time PCR analysis of gene expression in the murrel *Channa striatus* under high-temperature stress. *Fish Physiol. Biochem.* 42, 125–135. doi: 10.1007/s10695-015-0123-0

Ramachandra, R. K., Salem, M., Gahr, S., Rexroad, C. E. III., and Yao, J. (2008). Cloning and characterization of microRNAs from rainbow trout (*Oncorhynchus mykiss*): Their expression during early embryonic development. *BMC Dev. Biol.* 8:41. doi: 10.1186/1471-213X-8-41

Scott, G. R., Richards, J. G., Forbush, B., Isenring, P., and Schulte, P. M. (2004). Changes in gene expression in gills of the euryhaline killifish *Fundulus heteroclitus* after abrupt salinity transfer. *Am. J. Cell Physiol.* 287, C300–C309. doi: 10.1152/ajpcell.00054.2004

Shiogiri, N. S., Paulino, M. G., Carraschi, S. P., Baraldi, F. G., da Cruz, C., and Fernandes, M. N. (2012). Acute exposure of a glyphosate-based herbicide affects the gills and liver of the Neotropical fish, *Piaractus mesopotamicus*. *Environ. Toxicol. Pharmacol.* 34, 388–396. doi: 10.1016/j.etap.2012.05.007

Silver, N., Best, S., Jiang, J., and Thein, S. L. (2006). Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol. Biol.* 7:33. doi: 10.1186/1471-2199-7-33

Sinha, A. K., Dasan, A. F., Rasoloniriana, R., Pipralia, N., Blust, R., and De Boeck, G. (2015). Hypo-osmotic stress-induced physiological and ion-osmoregulatory responses in European sea bass (*Dicentrarchus labrax*) are modulated differentially by nutritional status. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 181, 87–99. doi: 10.1016/j.cbpa.2014.11.024

Silveira, T. L. R., Martins, G. B., Domingues, W. B., Remião, M. H., Barreto, B. F., Lessa, I. M., et al. (2018). Gene and blood analysis reveal that transfer from brackish water to freshwater is less stressful to the silverside *Odontesthes humensis*. *Front. Genet.* 9:28. doi: 10.3389/fgene.2018.00028

Small, B. C., Murdock, C. A., Bilodeau-Bourgeois, A. L., Peterson, B. C., and Waldbieser, G. C. (2008). Stability of reference genes for real-time PCR analyses in channel catfish (*Ictalurus punctatus*) tissues under varying physiological conditions. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 151, 296–304. doi: 10.1016/j.cbpb.2008.07.010

Somoza, G. M., Miranda, L. A., Berasain, G. E., Colautti, D., Remes Lenicov, M., and Strüssmann, C. A. (2008). Historical aspects, current status and prospects of pejerrey aquaculture in South America. *Aquac. Res.* 39, 784–793. doi: 10.1111/j.1365-2109.2008.01930.x

Strobl-Mazzulla, P. H., Moncaut, N. P., López, G. C., Miranda, L. A., Canario, A. V., and Somoza, G. M. (2005). Brain aromatase from pejerrey fish (*Odontesthes bonariensis*): cDNA cloning, tissue expression, and immunohistochemical localization. *Gen. Comp. Endocrinol.* 143, 21–32. doi: 10.1016/j.ygcen.2005.02.026

Sun, B. G., and Hu, Y. H. (2015). Evaluation of potential internal references for quantitative real-time RT-PCR normalization of gene expression in red drum (*Sciaenops ocellatus*). *Fish Physiol. Biochem.* 41, 695–704. doi: 10.1007/s10695-015-0039-8

Taki, F. A., Abdel-Rahman, A. A., and Zhang, B. (2014). A comprehensive approach to identify reliable reference gene candidates to investigate the link between alcoholism and endocrinology in Sprague-Dawley rats. *PLoS ONE* 9:e94311. doi: 10.1371/journal.pone.0094311

Tang, R., Dodd, A., Lai, D., McNabb, W. C., and Love, D. R. (2007). Validation of zebrafish (*Danio rerio*) reference genes for quantitative real-time RT-PCR normalization. *Acta Biochim. Biophys. Sin. (Shanghai)* 39, 384–390. doi: 10.1111/j.1745-7270.2007.00283.x

Tang, Y. K., Yu, J. H., Xu, P., Li, J. L., Li, H. X., and Ren, H. T. (2012). Identification of housekeeping genes suitable for gene expression analysis in Jian carp (*Cyprinus carpio* var. jian*). Fish Shellfish Immunol.* 33, 775–779. doi: 10.1016/j.fsi.2012.06.027

Taylor, D. L., Thomson, P. C., de Silva, K., and Whittington, R. J. (2007). Validation of endogenous reference genes for expression profiling of RAW264.7 cells infected with *Mycobacterium avium* subsp. *paratuberculosis* by quantitative PCR. *Vet. Immunol. Immunopathol.* 115, 43–55. doi: 10.1016/j.vetimm.2006.10.007

Tipsmark, C. K., Madsen, S. S., and Borski, R. J. (2004). Effect of salinity on expression of branchial ion transporters in striped bass (*Morone saxatilis*). *J. Exp. Zool. A Comp. Exp. Biol.* 301, 979–991. doi: 10.1002/jez.a.119

Tomy, S., Chang, Y. M., Chen, Y. H., Cao, J. C., Wang, T. P., and Chang, C. F. (2009). Salinity effects on the expression of osmoregulatory genes in the euryhaline black porgy *Acanthopagrus schlegeli*. *Gen. Comp. Endocrinol.* 161, 123–132. doi: 10.1016/j.ygcen.2008.12.003

Tsuzuki, M. Y., Ogawa, K., Strüssmann, C. A., Maita, M., Takashima, F., and Melo, C. M. R. (2007). The significance of cortisol on acclimation to salinity in pejerrey *Odontesthes bonariensis*. *Arq. Bras. Med. Vet. Zootec.* 59, 1301–1307. doi: 10.1590/S0102-09352007000500030

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., et al. (2002). Accurate normalization of real-time quantitative RT -PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3, 1–12. doi: 10.1186/gb-2002-3-7-research0034

Velasques, R. R., Sandrini, J. Z., and da Rosa, C. E. (2016). Roundup in zebrafish: effects on oxidative status and gene expression. *Zebrafish* 13, 432–441. doi: 10.1089/zeb.2016.1259

Wang, G., Yang, E., Smith, K. J., Zeng, Y., Ji, G., Connon, R., et al. (2014). Gene expression responses of threespine stickleback to salinity: implications for salt-sensitive hypertension. *Front. Genet.* 5:312. doi: 10.3389/fgene.2014.00312

Xu, X. Y., Shen, Y. B., Fu, J. J., Lu, L. U., and Li, J. L. (2014). Determination of reference microRNAs for relative quantification in grass carp (*Ctenopharyngodon idella*). *Fish Shellfish Immunol.* 36, 374–382. doi: 10.1016/j.fsi.2013.12.007

Yang, C. G., Wang, X. L., Tian, J., Liu, W., Wu, F., Jiang, M., et al. (2013). Evaluation of reference genes for quantitative real-time RT-PCR analysis of gene expression in Nile tilapia (*Oreochromis niloticus*). *Gene* 527, 183–192. doi: 10.1016/j.gene.2013.06.013

Zebral, Y. D., Costa, P. G., de Castro Knopp, B., Lansini, L. R., Zafalon-Silva, B., Bianchini, A., et al. (2017). Effects of a glyphosate-based herbicide in pejerrey *Odontesthes humensis* embryonic development. *Chemosphere* 185, 860–867. doi: 10.1016/j.chemosphere.2017.07.069

Zhang, J., Hu, Y., Sun, B., Xiao, Z., and Sun, L. (2013). Selection of normalization factors for quantitative real time RT-PCR studies in Japanese flounder (*Paralichthys olivaceus*) and turbot (*Scophthalmus maximus*) under conditions of viral infection. *Vet. Immunol. Immunopathol.* 152, 303–316. doi: 10.1016/j.vetimm.2012.12.018

Zheng, W. J., and Sun, L. (2011). Evaluation of housekeeping genes as references for quantitative real time RT-PCR analysis of gene expression in Japanese flounder (*Paralichthys olivaceus*). *Fish Shellfish Immunol.* 30, 638–645. doi: 10.1016/j.fsi.2010.12.014

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

Copyright © 2018 Silveira, Domingues, Remião, Santos, Barreto, Lessa, Varela Junior, Martins Pires, Corcini, Collares, Seixas, Robaldo and Campos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 15 March 2018
doi: 10.3389/fevo.2018.00022

Unrevealing Parasitic Trophic Interactions—A Molecular Approach for Fluid-Feeding Fishes

*Karine O. Bonato, Priscilla C. Silva and Luiz R. Malabarba**

Laboratório de Ictiologia, Programa de Pós-Graduação em Biologia Animal, Departamento de Zoologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

Fish diets have been traditionally studied through the direct visual identification of food items found in their stomachs. Stomach contents of Vandeliinae and Stegophilinae (family Trichomycteridae) parasite catfishes, however, cannot be identified by usual optical methods due to their mucophagic, lepidophagic, or hematophagic diets, in such a way that the trophic interactions and the dynamics of food webs in aquatic systems involving these catfishes are mostly unknown. The knowledge about trophic interactions, including difficult relation between parasites and hosts, are crucial to understand the whole working of food webs. In this way, molecular markers can be useful to determine the truly hosts of these catfishes, proving a preference in their feeding behavior for specific organisms and not a generalist. Sequences of *cytochrome oxidase subunit 1* (*COI*) were successfully extracted and amplified from mucus or scales found in the stomach contents of two species of stegophilines, *Homodiaetus anisitsi,* and *Pseudostegophilus maculatus*, to identify the host species. The two species were found to be obligatory mucus-feeders and occasionally lepidophagic. Selection of host species is associated to host behavior, being constituted mainly by substrate-sifting benthivores. Characiformes are preferred hosts, but host choice depends on what characiform species are available in their environments, usually corresponding to the most abundant species. This is the first time that host species of parasitic fishes bearing mucophagous habits are identified, and demonstrates the effectiveness of the extraction and amplification of mitochondrial DNA from the ingested mucus in gut contents. The molecular markers effectively allowed determine parasite preferences and helps in better understanding the food web and trophic interaction on which fish species are involved. Despite, the methodology applied here can be used for an infinitive of organisms improving ecological trophic studies.

Keywords: annealing blocking primer, DNA barcode, food webs, parasite-host interaction, stegophilinae, vandeliinae

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Gonzalo Gajardo, University of Los Lagos, Chile
Igor De Paiva Affonso, Universidade Tecnológica Federal do Paraná, Brazil

****Correspondence:***
Luiz R. Malabarba
malabarb@ufrgs.br

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

Received: *19 August 2017*
Accepted: *26 February 2018*
Published: *15 March 2018*

Citation:
Bonato KO, Silva PC and Malabarba LR (2018) Unrevealing Parasitic Trophic Interactions—A Molecular Approach for Fluid-Feeding Fishes. Front. Ecol. Evol. 6:22. doi: 10.3389/fevo.2018.00022

INTRODUCTION

The role of parasites in food webs has been largely disregarded (Sukhdeo, 2012). As an example, Winemiller and Polis (1996) was the most important contribution of food web studies but not approached the parasite-host interaction as part of the trophic webs. Since Elton insights (Elton et al., 1931) it is known that parasites are very important links in the food webs and are capable of externing major effects on ecological interactions. Today, there is no longer the need to argue that

parasites must be included in all models of ecosystem function (Sukhdeo, 2012). However, many questions remain as how parasites might fit in food webs, if it should be included or excluded in the food webs, what is the role of parasites in host population regulation, and what are the evolutionary and ecological implications of parasite mediation in trophic interactions (Sukhdeo and Hernandez, 2005). Nevertheless, small alteration in the position of parasites and host can change the food chain length, connectance, and the establishment of food patterns (Huxham et al., 1995; Leaper and Huxham, 2002).

Indeed, only dietary studies allow the comprehension of the trophic interactions and of the dynamics of food webs specific for fishes in aquatic systems (Winemiller and Polis, 1996; Carreon-Martinez and Heath, 2010; Murray et al., 2011). For these purposes, food items are traditionally identified through visual analysis of stomach, gut, or fecal contents under microscope (Hyslop, 1980; Taguchi et al., 2014) and through stable isotope analysis (DeNiro and Epstein, 1978, 1981; Post, 2002; Mercado-Silva et al., 2015). None of these methods, however, allow the identification of items that are easily digested or of specific parasitic-host (or predator-prey) interactions, which may cause a bias in data interpretation (Sheppard and Harwood, 2005; Paquin et al., 2014). This may be especially problematic when dealing with small species (Sheppard and Harwood, 2005; Jo et al., 2014; Paquin et al., 2014).

Since the 1940's many methods of diet analysis have been developed with a recent use of molecular tools (Symondson, 2002; King et al., 2008; Hardy et al., 2010; Pompanon et al., 2012; Taguchi et al., 2014). Molecular techniques are efficient for the identification of small preys or digested items that cannot be identified through traditional methodologies. It is also presumed to allow a higher taxonomic resolution (Carreon-Martinez et al., 2011) through the use of sequences of specific regions of the mitochondrial genome to identify any species (DNA Barcode). In most animal groups, the *cytochrome oxidase subunit 1* (*COI*) is the reference for the DNA barcoding system (Hebert et al., 2003). *COI* sequences libraries are available through on-line systems (such as GenBank, Bold), enabling its use in the identifications of host or prey species (Valentini et al., 2009; Corse et al., 2010; Leray et al., 2013a; Jo et al., 2014). DNA barcode techniques have been used to determine host-parasitoids webs in arthropods (Hrček and Godfray, 2014), as in Lepidoptera (Janzen et al., 2009) and Hemiptera (Gordon and Weirauch, 2015). However, in fish species the establishment of parasitic interaction has been reported as difficult (Sazima, 1983; Lima et al., 2012). According to Paine (1980) "*The central significance of webs is derived from the fact that the links between species are often easily identified and the resultant trophic scaffolding provides a tempting descriptor of community structure*." Thus, the no comprehension of the parasite/host interaction leads to the impossibility of understanding food webs and the trophic ecology at the ecosystems.

The Neotropical catfish family Trichomycteridae is composed of eight subfamilies, two of which consist exclusively of fishes referred to as parasites, the Stegophilinae ,and Vandeliinae (de Pinna, 1998, 2016; Datovo and Bockmann, 2010; Ferrer and Malabarba, 2013). The vandeliines feed exclusively on blood from the gills of other fishes (Kelley and Atz, 1964; Machado and Sazima, 1983) while the stegophilines have been reported as scale and mucus-feeders (Eigenmann and Allen, 1942; Roberts, 1972; Baskin et al., 1980; Machado and Sazima, 1983; Winemiller and Yan, 1989; Neto and de Pinna, 2016). The feeding habits of the representatives of a third subfamily, the Tridentinae, remains uncertain, but there are circumstantial records of semi-parasitic (scale-eating) or predation of small invertebrates among its species. Certainly, their biology is key to understanding the evolution of parasitic feeding behavior of stegophilines and vandeliines (de Pinna, 2016).

The species of the stegophiline genus *Homodiaetus* are small (maximum 42.0 mm of standard length) and translucent when alive, except the head and abdominal region (Koch, 2002). *Homodiaetus anisitsi* Eigenmann and Ward, 1907 (**Figure 1**) is found in lakes and rivers in the lower Paraná-Paraguay System and coastal river drainages of the Rio Grande do Sul State, Brazil (Koch, 2002). The stegophiline *Pseudostegophilus maculatus* (Steindachner, 1879; **Figure 1**) reaches a maximum size of 60.0 mm of standard length and is found in the lower Paraná and Uruguay river basins, South America (de Pinna and Wosiacki, 2003). Specimens of both species inhabit fine grain sandy bottom environments. In aquarium observations, specimens of *H. anisitsi* remain buried in sandy bed except for the eyes. They quickly move out from the substrate and attach to the side of fishes passing by for mucus and occasionally scale ingestion (LRM pers. obs.).

Host identification of mucophagous, lepidophagous, and hematophagous trichomycterids cannot be made by conventional analysis of the stomach contents due to the nature of the ingested items. This is notably exemplified by Winemiller and Yan (1989) that examined 245 specimens of the stegophiline *Ochmacanthus alternus* Myers, 1927. They were able to identify the presence of ingested mucus in 95% of the stomachs, but were not able to identify a single hosts species among the 88 fish species found syntopic with this catfish. Likewise, there is no information available about the host species exploited by *H. anisitsi*, *P. maculatus* or other stegophilines in natural environments. The same is true for most vandeliines, except for

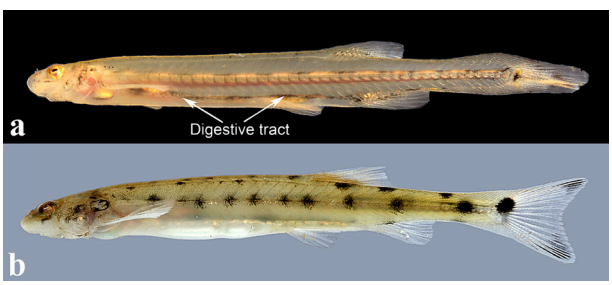


FIGURE 1 | Studied species of Stegophilinae. **(a)** *Homodiaetus anisitsi*, 32 mm standard length, specimen from Lagoa dos Quadros, rio Tramandaí basin, Rio Grande do Sul State, Brazil. **(b)** *Pseudostegophilus maculatus*, 38 mm standard length, specimen from rio Ibicuí, rio Uruguay basin, Rio Grande do Sul State, Brazil. Arrows indicate the digestive tract, visible by transparency as a relatively straight tube from the mouth to the anus.

a few visual records of specimens observed attached to gill arches of some large fishes (Zuanon and Sazima, 2005) or experiments in aquaria or observation on large fishes tied to river banks (Machado and Sazima, 1983; Zuanon and Sazima, 2004).

Considering mucophagous or hematophagous catfishes, molecular techniques provide an alternative to determine the origin of the food items ingested by mucus-, scale- and blood-feeders, being considered as a powerful tool for studying feeding ecology (King et al., 2008). However, this tool has not been used yet to identify the species to which belong these items found in the stegophiline stomachs. Therefore, our hypothesis is that molecular markers can be useful to determine the truly hosts of these catfishes, proving a preference in their feeding behavior for specific organisms and not a generalist. Notwithstanding, the aims of this study are: (1) to verify the viability of DNA extraction and amplification of the *cytochrome oxidase subunit 1* gene (*COI*) from the mucus and occasionally scales ingested by *H. anisitsi* and *P. maculatus*; (2) the identification of their hosts, and (3) establishment of specific parasitic-host interactions.

METHODS AND MATERIALS

Ethics Statement

This study was approved by the Animal Ethics Committee of the Universidade Federal do Rio Grande do Sul (Permit Number: 24434) and was conducted in accordance with protocols in their ethical and methodological aspects for the use of fish.

Analysis

Sixty three specimens of *H. anisitsi* and 18 specimens of *P. maculatus* (**Tables 1**, **2** and Supplementary Table 1) were analyzed and vouchers cataloged in the fish collection of the Departamento de Zoologia, Universidade Federal do Rio Grande do Sul (UFRGS; CGEN Process # 02000.002101/2007-25) (Supplementary Table 2).

The digestive tract of *H. anisitsi* and *P. maculatus* is a relatively straight tube from the mouth to the anus (**Figure 1**) with a very thin wall. The entire digestive tract was opened and all the content was examined under a stereomicroscope to identify the food items. Immediately after that, the ingested items and a tissue sample of the dissected specimens were moved to separate and empty sterilized tubes of 0.2 μl to proceed with the DNA extraction. DNA was extracted from all stomach contents using the "Phire Animal Tissue Direct PCR Kit" by Thermo Scientific, following manufacturer instructions. Extracted DNA samples were then submitted to two different protocols for gene amplification.

The PCR reactions were carried out in a reaction volume of 20 μL [12.9 μL of H20, 2 μL of 10 × reaction buffer (Platinum®Taq), 0.6 μL of MgCl2 (50 mM), 2 μL of dNTPs (2 mM), 0.2 μL of each primer (10 μM), 2.0 μL of the blocking primer, 0.1 μL (5 U) of Platinum® Taq (Invitrogen), and 100 ng of template DNA. A versatile PCR primer (mlCOIintF/jgHCO2198 Geller et al., 2013; Leray et al., 2013b) was used to amplify a 313 bp region of the mitochondrial COI region. In *H. anisitsi* samples, this primer was associated to a parasite-specific annealing blocking primer (blocking primer sequence: 5′ CGAARAATCARAAYARRTGTTG3SpC3 3′). Because the predator DNA co-amplification is known to inhibit prey DNA detection in stomach contents (Vestheim and Jarman, 2008), the blocking primer was included at ten times the concentration of versatile primers (Leray et al., 2013b). The PCR cocktail and touchdown temperature profile used in this study follow Leray et al. (2013b).

The quantification of the amplified gene was carried out in agarose gel using a Low Mass Ladder 100 pb by Ludwig Biokee as comparison parameter. The PCR products were purified by enzymatic methods (ExoSap) and sequencing was performed on a sequencing platform of Actgene (Porto Alegre, Brazil) and Laboratory of analytical biology at National Museum of Natural History, Smithsonian Institution (Washington DC). The chromatogram qualities of the generated sequences were checked in MEGA 6.0 (Tamura et al., 2013), and were aligned using Clustal W (Higgins et al., 1994).

The sequences obtained from stomach contents were compared to the GenBank on-line platform; and to local Barcode inventories (sequences deposited in Genbank; access numbers in Supplementary Table 2). The genetic distance between the stomach contents and the local Barcode inventories were estimated using p-distance in MEGA 6.0 (Tamura et al., 2013) software using as molecular evolution model Kimura-2 parameters.

RESULTS

Stomachs of three specimens of *H. anisitsi* contained sand only, and 15 stomachs of *H. anisitsi* and eight of *P. maculatus* were empty, being not submitted to DNA extraction. The stomachs of 42 specimens of *H. anisitsi* and 10 specimens of *P. maculatus* contained ingested items identified as mucus and occasionally scales, sometimes associated with sand (**Tables 1**, **2** and Supplementary Table 1). DNA was successfully extracted from gut contents of all these specimens, but amplification gave positive results only for 10 specimens of *H. anisitsi* and 7 specimens of *P. maculatus*.

The BLAST search (GenBank) based on the sequences obtained from stomach contents generated identifications with values of identity ranging from as low as 80% and up to 100%. Values of BLAST identity lower than 98% were not considered (Supplementary Table 1), since it generates spurious identifications, including fish species absent in the hydrographic regions sampled (e.g., *Astyanax altiparanae*) or even marine species of Balistidae and Apogonidae. Further comparison to local barcode data allowed a second and independent identification, allowing the confirmation of the identity of the species spoiled by these stegophilines, or identification of species whose sequences are not available in GenBank (**Tables 1**, **2**).

The *COI* sequences amplified from stomach contents resulted always in the identification of a single species per stomach. Six host species were identified in the contents of eight stomachs in *H. anisitsi* (**Table 1**), five freshwater species belonging to

TABLE 1 | Results of the identification of the collected samples from stomach contents of *Homodiaetus anisitsi*.

Voucher UFRGS	SL (mm)	Drainage	Stomach Contents	BLAST identity (>95%)	*p* distance to local barcode
13611A	29.4	Patos Lagoon	Scales	*Bryconamericus iheringii* (98%)	*Bryconamericus iheringii* ($p = 0.01$)
13611B	30.9	Patos Lagoon	Scales	*Bryconamericus iheringii* (98%)	*Bryconamericus iheringii* ($p = 0.01$)
13937	34.9	Patos Lagoon	Mucus	–	*Homodiaetus anisitsi* ($p = 0.01$)
16715	28.7	Tramandaí River	Sand and mucus	*Mugil liza* (100%)	*Mugil liza* ($p = 0$)
19372	33.7	Tramandaí River	Mucus and scales	*Cyphocharax voga* (99%)	*Cyphocharax voga* ($p = 0$)
19373	28.4	Tramandaí River	Mucus	*Mugil liza* (100%)	*Mugil liza* ($p = 0$)
13807	36.3	Uruguay River	Mucus	–	*Astyanax fasciatus* ($p = 0.03$)
14658	36.1	Uruguay River	Mucus	–	*Acestrorhynchus pantaneiro* ($p = 0.12$)
17927A	34.8	Uruguay River	Detritus and mucus	*Astyanax altiparanae* (98%)	*Astyanax jacuhiensis* ($p = 0.00$
17927B	31.4	Uruguay River	Mucus and scales		*Homodiaetus anisitsi* ($p = 000$)
17927C	42.3	Uruguay River	Mucus	Failed	Failed

Sequences were compared to GenBank (BLAST identity), to local Barcode data (p distance) and to the parasitic catfish COI sequence. UFRGS, catalog number of vouchers in the Fish Collection, Universidade Federal do Rio Grande do Sul; SL, standard length in mm.

TABLE 2 | Results of the identification of the collected samples from stomach contents of *Pseudostegophilus maculatus*.

Voucher UFRGS	SL (mm)	Stomach contents	BLAST identity	*p* distance to local barcode
20127-5	41.8	Mucus	*Pseudostegophilus maculatus* (98%)	*Pseudostegophilus maculatus* ($p = 0.00$)
20127-6	40.4	Mucus	*Pseudostegophilus maculatus* (100%)	*Pseudostegophilus maculatus* ($p = 0.00$)
20127-8	38.3	Mucus	*Prochilodus lineatus* (100%)	
20127-11	35.0	Mucus	*Pseudostegophilus maculatus* (100%)	*Pseudostegophilus maculatus* ($p = 0.00$)
20127-11	35.0	Scale	*Bryconamericus turiuba* (98%)	*Piabarchus stramineus* ($p= 0.01$)
20127-13	33.8	Mucus	*Prochilodus lineatus* (100%)	
20127-14	33.1	Mucus	*Prochilodus lineatus* (100%)	

Sequences were compared to GenBank (BLAST identity), to local Barcode data (p distance) and to the parasitic catfish COI sequence. UFRGS, catalog number of vouchers in the Fish Collection, Universidade Federal do Rio Grande do Sul; SL, standard length in mm.

the order Characiformes (six stomachs) and one species of the diadromous mugilid *Mugil liza* (two stomachs). Two host species were identified in the contents of eight stomachs in *P. maculatus* (**Table 2**), both belonging to the freshwater order Characiformes (four stomachs).

Both species were mucophagous and lepidophagous, and parasitized organisms that live close to the bottom or that feed on the bottom.

DISCUSSION

The analysis of guts contents of *H. anisitsi* and *P. maculatus* suggests they are obligatory mucus feeders, as also observed in the stegophiline *O. alternus* by Winemiller and Yan (1989). This diet is possible by the morphological specializations shared by the members of the subfamily Stegophilinae that possess a wide and ventral mouth with numerous teeth in both jaws, arranged

in several rows (Baskin et al., 1980). This allows mucus scraping of their hosts and may function as a sucker (Winemiller and Yan, 1989).

A scale-feeding behavior has been also described for the stegophilines (Eigenmann and Allen, 1942; Roberts, 1972; Baskin et al., 1980; Machado and Sazima, 1983; Winemiller and Yan, 1989), but we found lepidophagy in the two studied species to be non-obligatory and related to the body size and consequently scale size of the host species (Supplementary Table 1). Even though we do not have information on the body size of the hosts whose mucus were found in the stomachs (adult size or juveniles), none of the stomachs containing mucus of the largest host species (*M. liza* and *Prochilodus lineatus*) showed the presence of scales (**Figures 2**, **3**). Instead, scales were found associated with DNA of *Bryconamericus* and *Cyphocharax*, both with smaller body sizes (**Figures 2**, **3**).

It is remarkable the preference for characiforms as hosts in the two stegophiline species. In a recent inventory of the freshwater fish species from the rio Uruguay, rio Tramandaí and laguna dos Patos drainages, that embraces the areas of collection of the samples studied herein, Siluriformes was found to be the dominant group, corresponding to 42% of the species found in these drainages (Bertaco et al., 2016), but none correspond to the stomach contents. Instead, except for two stomachs containing one species of mullets (Mugilidae), all stomachs contained species of Characiformes, that corresponds to nearly 1/4 (28%) of the freshwater species found in these drainages (Bertaco et al., 2016). In addition and further supporting the preference for characiforms, several species of Siluriformes and Cichlidae (Cichliformes) were collected along with the specimens of *H. anisitsi* or *P. maculatus* analyzed here (Supplementary Table 3), but were not found in their stomachs.

Host species of *H. anisitsi* varied according to the locality of collection, and seems to be related to hosts that live close to the bottom or that feed on the bottom, and to the species available. *M. liza* is very abundant in coastal lagoons of southern Brazil (Reis and D'Incao, 2000), including the locality sampled in the rio Tramandaí drainage, being found in two stomachs of *H. anisitsi*. The second species found in stomach contents in this drainage is *Cyphocharax voga*, one of the five most abundant species in these coastal lagoons (Schifino et al., 2004). Both host species are detritivorous and edentulous (Dualiby, 1988; Oliveira and Soares, 1996; Corrêa and Piedras, 2008). In the laguna dos Patos drainage the two identified stomach contents included scales of *Bryconamericus iheringii*, a characid species with inferior mouth that feeds in the substrate (Orciolli and Bennemann, 2006) and is very abundant in creeks of this drainage (Corrêa et al., 2015). In the rio Uruguay drainage, three host species were found (*Astyanax jacuhiensis*, *Astyanax fasciatus,* and *Acestrorhynchus pantaneiro*), but in this case none have feeding habits associated to bottom feeding. Instead, *A. lacustris* and *A. fasciatus* normally inhabit the column of water and surface and can inspect the bottom for feeding (Casatti et al., 2001).

A similar result was obtained from *P. maculatus*, being host species related to host behavior and to the species availability. *P. lineatus*, the most common host found in three stomachs, feeds on algae and detritus grasping the substrate (Fugi et al., 2001). *Piabarchus stramineus* found in one stomach inhabits the water column and can inspect the bottom for feeding (Casatti et al., 2003; Brandão-Gonçalves et al., 2009). Likewise, most hosts of *H. anisitsi*, *P. lineatus* is very abundant in the Rio de la Plata basin,

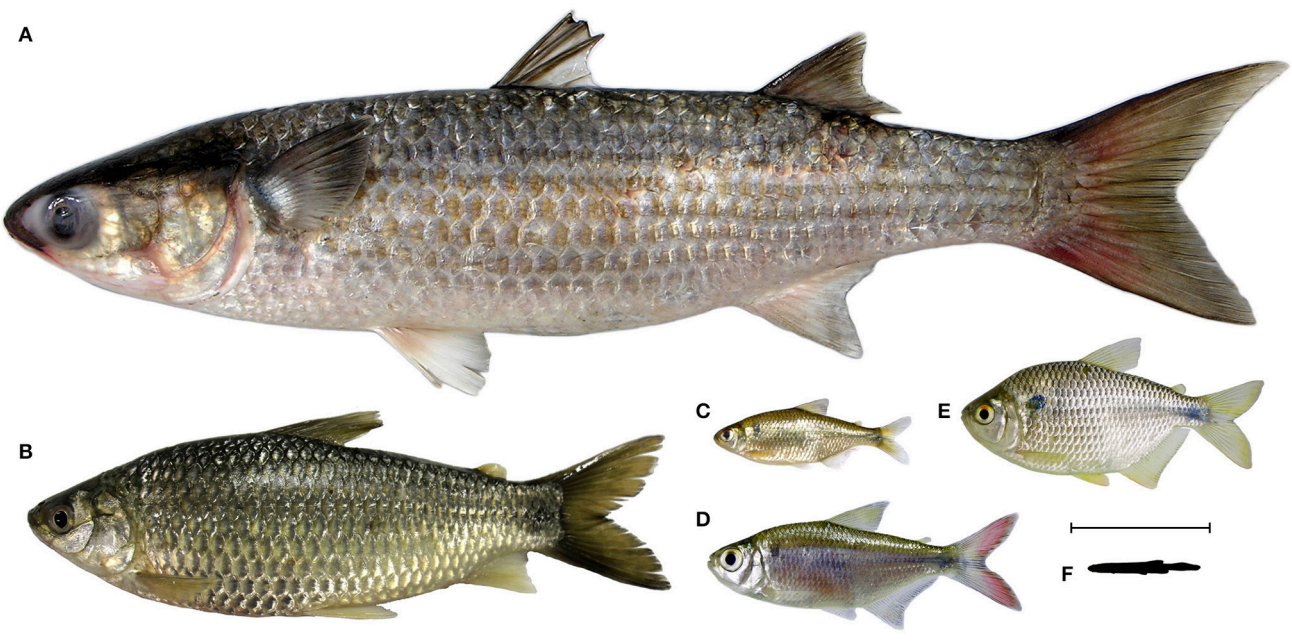


FIGURE 2 | Host species of *Homodiaetus anisitsi* represented in scale. **(A)** *Mugil liza*; **(B)** *Cyphocharax voga*; **(C)** *Bryconamericus iheringii*; **(D)** *Astyanax fasciatus*; **(E)** *Astyanax jacuhiensis*; **(F)** *Homodiaetus anisitsi* represented in black. Scale bar 50 mm.



FIGURE 3 | Host species of *Pseudostegophilus maculatus* represented in scale. **(A)** *Prochilodus lineatus*; **(B)** *Piabarchus stramineus*; **(C)** *Pseudostegophilus maculatus* represented in black. Scale bar 50 mm.

including the rio Uruguay, and dominates the biomass, being the target of the principal freshwater fishery and is the main prey item for large predatory fish (Speranza et al., 2013).

When studying trophic ecology, one of the most difficult concepts is the classification of organisms as specialists or generalists. Parasites are usually highly host-specific (Sukhdeo and Hernandez, 2005), but the problem is why some species appear to not specialize as consistently on a single host species (Thompson, 1982). In the best revision about this approach Sukhdeo and Hernandez (2005) indicated that phylogenetic history as morphologic characters and feeding behaviors of both players are able to promote parasite-host specificity. Authors argue that if you base your definition of specificity on the number of higher taxa on which a parasite feeds it allows inferences about the extent to which a parasite is tracking it host taxon phylogenetically. Our results suggest that these two species of stegophilines show host-specificity, because they boot feed mainly on species belonging to the same order. The study of the feeding behavior of other stegophilines may test if the evolution of mucous and scale feeding behavior of these catfishes has been associated to a coevolutionary interaction with the order Characiformes.

This is the first time the mucous found in the stomachs of mucophagic species were identified through DNA extraction and sequencing, and this approach may be used to any parasite/host interactions in any group of organisms. It can also test hypotheses based on empiric observations that lack testing support. For example, Sazima (1983) reports the characid fish *Probolodus heterostomus* as lepidophagous and associates the scales found in its stomach with the shoal that this species is mimetic. Since then, no study has been done to determine the origin of scales commonly found in the stomachs of *P. heterostomus*: are them from its shoal or can be originated from other species, preferentially not belonging to the shoal?

In other examples, Lima et al. (2012) reports a behavior of "mutilating predation" for *Odontostilbe pequira* preferentially attacking *Leporinus friderici*, but do not show any support as molecular identification of food items. Based only in field observations of the putative attack, *O. pequira* is currently classified as omnivore. Other problem is related to the determination of species in piscivorous fish diets. In most cases it is possible to determine only the genus or family of prey by bone remains of semi digested preys in the stomachs due the similarity of structures or the stage of fragment digestibility (Hansel et al., 1988). In such cases, the DNA identification of hosts or preys is the best alternative to solve these questions, and can be applied to any parasite/host or predator/prey interaction.

Thus, in our case, regardless the relative small samples from which stomach contents were positively identified through DNA extraction, some patterns are clearly discernible in the two stegophiline species: (1) they are truly parasites once they are obligatory mucophagous and lepidophagous; (2) Characiformes are preferred hosts event thought they are not the most abundant order in their environments; (3) host choice is related to the habits of hosts, choosing those that feed and live in or near the bottom. The unanswered question now is: why the preferential choice for Characiformes?

Finally, the accurate analysis of parasite/host interactions of the stegophilines allowed us to better understand their function on the freshwater ecosystem where these are inserted. From this point, we conclude that molecular helps in reveal any trophic interaction, including the identification of food items not allowed by usual methods, permitting a whole comprehension of the interactions inside of the food webs and any trophic ecology of the system.

AUTHOR CONTRIBUTIONS

KB, PS, and LM: Conceived and designed the experiments; KB and PS: Performed the experiments; KB, PS, and LM: Analyzed the data; PS and LM: Contributed reagents, materials, analysis tools; PS: Drawn the blocked primer; KB, PS, and LM: Wrote the paper; KB, PS, and LM: Reviewing the manuscript.

ACKNOWLEDGMENTS

Authors are grateful to MC Malabarba and two reviewers for critically reading the manuscript and providing helpful suggestions; to UFRGS laboratory team that collected the analyzed samples; to Lee Weigt and Jeff Hunt for all support to PS at LAB, NMNH, Smithsonian Institution; to Mathieu Leray that helped designing the blocking primer. This research is supported by CNPq (process # 307890/2016-3 and 401204/2016-2 to LM; 205678/2014-9 and 150956/2017-7 to PS).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2018.00022/full#supplementary-material

REFERENCES

Baskin, J. N., Zaret, T. M., and Mago-Leccia, F. (1980). Feeding of reportedly parasitic catfishes (Trichomycteridae and Cetopsidae) in the Río Portuguesa basin, Venezuela. *Biotropica* 12, 182–186. doi: 10.2307/2387969

Bertaco, V. A., Ferrer, J., Carvalho, F. R., and Malabarba, L. R. (2016). Inventory of the freshwater fishes from a densely collected area in South America – a case study of the current knowledge of Neotropical fish diversity. *Zootaxa* 4138, 401–440. doi: 10.11646/zootaxa.4138.3.1

Brandão-Gonçalves, L., Lima-Junior, S. E., and Suarez, Y. R. (2009). Hábitos alimentares de *Bryconamericus stramineus* Eigenmann, 1908 (Characidae), em diferentes riachos da sub-bacia do Rio Guiraí, Mato Grosso do Sul, Brasil. *Biota Neotrop.* 9, 135–143. doi: 10.1590/S1676-06032009000100016

Carreon-Martinez, L., and Heath, D. D. (2010). Revolution in food web analysis and trophic ecology diet analysis by DNA and stable isotope analysis. *Mol. Ecol.* 19, 25–27. doi: 10.1111/j.1365-294X.2009.04412.x

Carreon-Martinez, L., Johnson, T. B., Ludsin, S. A., and Heath, D. D. (2011). Utilization of stomach content DNA to determine diet diversity in piscivorous fishes. *J. Fish Biol.* 70, 1170–1182. doi: 10.1111/j.1095-8649.2011.02925.x

Casatti, L., Langeani, F., and Castro, M. R. C. (2001). Peixes de riacho do Parque Estadual Morro do Diabo, bacia do Alto Rio Paraná, SP. *Biota Neotrop.* 1, 1–15. doi: 10.1590/S1676-06032001000100005

Casatti, L., Mendes, H. F., and Ferreira, K. M. (2003). Aquatic macrophytes as feeding site for small fishes in the Rosana reservoir, Paranapanema river, southeastern Brazil. *Braz. J. Biol.* 63, 213–222. doi: 10.1590/S1519-69842003000200006

Corrêa, F., Oliveira, F. E., Tuchtenhagen, T., Pouey, J., and Piedras, S. (2015). Ichthyofauna of the hydrographic basin of the Chasqueiro Stream (Mirim Lagoon system, Southern Brazil): generating subsidies for conservation and management. *Biota Neotrop.* 15:e0006. doi: 10.1590/1676-0611-BN-2015-0006

Corrêa, F., and Piedras, S. R. N. (2008). Alimentação de *Cyphocharax voga* (Hensel, 1869) (Characiformes, Curimatidae) no arroio Corrientes, Pelotas, Rio Grande do Sul, Brasil. *Biotemas.* 21, 117–122. doi: 10.5007/2175-7925.2008v21n4p117

Corse, E., Costedoat, C., Chappaz, R., Pech, N., Martin, J. F., and Gilles, A. (2010). A PCR-based method for diet analysis in freshwater organisms using 18S rDNA barcoding on faeces. *Mol. Ecol. Resour.* 10, 96–108. doi: 10.1111/j.1755-0998.2009.02795.x

Datovo, A., and Bockmann, F. A. (2010). Dorsolateral head muscles of the catfish families Nematogenyidae and Trichomycteridae (Siluriformes: Loricarioidei): comparative anatomy and phylogenetic analysis. *Neotrop. Ichthyol.* 8, 193–246. doi: 10.1590/S1679-62252010000200001

DeNiro, M. J., and Epstein, S. (1978). Influence of diet on the distribution of carbon isotopes in animals. *Geochim. Cosmochim. Acta.* 42, 495–506. doi: 10.1016/0016-7037(78)90199-0

DeNiro, M. J., and Epstein, S. (1981). Influence of diet on the distribution of nitrogen isotopes in animals. *Geochim. Cosmochim. Acta.* 45, 341–351. doi: 10.1016/0016-7037(81)90244-1

de Pinna, M. C. C. (1998). "Phylogenetic relationships of Neotropical Siluriformes (Teleostei: Ostariophysi): historical overview and synthesis of hypotheses," in *Phylogeny and Classification of Neotropical Fishes*, eds L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M. Lucena, and C. A. S. Lucena (Porto Alegre, RS: EDIPUCRS), 279–330.

de Pinna, M. C. C. (2016). The dawn of phylogenetic research on Neotropical fishes: a commentary and introduction to Baskin (1973), with an overview of past progress on trichomycterid phylogenetics. *Neotrop. Ichthyol.* 14:e150127. doi: 10.1590/1982-0224-20150127

de Pinna, M. C. C., and Wosiacki, W. (2003). "Family Trichomycteridae (Pencil or parasitic catfishes)," in *Check List of the Freshwater Fishes of South and Central America*, eds R. E. Reis, S. O. Kullander, and C. J. Ferraris (Porto Alegre, RS: EDIPUCRS), 270–290.

Dualiby, D. O. (1988). Ecología trófica de *Mugil curema*, *M. incilis* y *M. liza* (Pisces: Mugilidae) en la Ciénaga Grande de Santa Maria, Caribe colombiano. I. Análisis cualitativo y cuantitativo. *Anal. Instit. Investig. Marinas de Punta Betín.* 18, 113–126.

Eigenmann, C. H., and Allen, W. R. (1942). *Fishes of Western South America*. Lexington: University of Kentucky.

Elton, C., Ford, E. B., Baker, J. R., and Gardner, A. D. (1931). The health and parasites of a wild mouse population. *J. Zool.* 101, 657–721. doi: 10.1111/j.1096-3642.1931.tb01037.x

Ferrer, J., and Malabarba, L. R. (2013). Taxonomic review of genus *Trichomycterus* Valenciennes (Siluriformes: Trichomycteridae) from the laguna dos Patos system, Southern Brazil. *Neotrop. Ichthyol.* 11, 217–246. doi: 10.1590/S1679-62252013000200001

Fugi, R., Agostinho, A. A., and Hahn, N. S. (2001). Trophic morphology of five benthic-feeding fish species of a tropical floodplain. *Rev. Bras. Biol.* 61, 27–33. doi: 10.1590/S0034-71082001000100005

Geller, J. B., Meyer, C. P., Parker, M., and Hawk, H. (2013). Redesign of PCR primers for mitochondrial Cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Mol. Ecol. Resour.* 13, 851–861. doi: 10.1111/1755-0998.12138

Gordon, R. L., and Weirauch, C. (2015). Efficient capture of natural history data reveals prey conservatism of cryptic termite predators. *Mol. Phyl. Evol.* 94, 65–73. doi: 10.1016/j.ympev.2015.08.015

Hansel, H. C., Duke, S. D., Lofy, P. T., and Gray, G. A. (1988). Use of Diagnostic Bones to Identify and Estimate Original Lengths of Ingested Prey Fishes. *Transac. Amer. Fish. Soc.* 17, 55–62. doi: 10.1577/1548-8659(1988)117<0055:UODBTI>2.3.CO;2

Hardy, C. M., Krull, E. S., Hartley, D. M., and Oliver, R. L. (2010). Carbon source accounting for fish using combined DNA and stable isotope analyses in a regulated lowland river weir pool. *Mol. Ecol.* 19, 197–212. doi: 10.1111/j.1365-294X.2009.04411.x

Hebert, P. D., Ratnasingham, S., and Waard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B.* 270, 596–599. doi: 10.1098/rsbl.2003.0025

Higgins, D., Thompson, J. D., Gibson, T., Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Hrček, J., and Godfray, H. C. J. (2014). What do molecular methods bring to host–parasitoid food webs? *Tre. Parasitol.* 31, 30-35. doi: 10.1016/j.pt.2014.10.008

Huxham, M. D., Raffaelli, D. G., and Pike, A. (1995). Parasites and food web patterns. *J. Anim. Ecol.* 64, 168–176. doi: 10.2307/5752

Hyslop, E. J. (1980). Stomach contents analysis – a review of methods and their application. *J. Fish Biol.* 17, 411–429. doi: 10.1111/j.1095-8649.1980.tb02775.x

Janzen, D. H., Hallwachs, W., Blandin, P., Burns, J. M., Cadiou, J.-M., Chancon, I., et al. (2009). Integration of DNA barcoding into an ongoing

inventory of complex tropical biodiversity. *Mol. Ecol. Resour.* 9, 1–26. doi: 10.1111/j.1755-0998.2009.02628.x

Jo, H., Gim, J. A., Jeong, K. S., Kim, H. S., and Joo, G. J. (2014). Application of DNA barcoding for identification of freshwater carnivorous fish diets: Is number of prey items dependent on size class for *Micropterus salmoides*? *Ecol. Evol.* 4, 219–229. doi: 10.1002/ece3.921

Kelley, W. E., and Atz, J. (1964). A pygidiid catfish that can suck blood from goldfish. *Copeia* 1964, 702–704. doi: 10.2307/1441447

King, R. A., Read, D. S., Traugott, M., and Symondson, W. O. C. (2008). Molecular analysis of predation: a review of best practice for DNA-based approaches. *Mol. Ecol.* 17, 947–963. doi: 10.1111/j.1365-294X.2007.03613.x

Koch, W. R. (2002). Revisão taxonômica do gênero *Homodiaetus* (Teleostei, Siluriformes, Trichomycteridae). *Iheringia Ser. Zool.* 92, 33–46. doi: 10.1590/S0073-47212002000300004

Leaper, R., and Huxham, M. (2002). Size constraints in a real food web: predator, parasite and prey body-size relationships. *Oikos* 99, 443–456. doi: 10.1034/j.1600-0706.2002.10888.x

Leray, M., Agudelo, N., Mills, S. C., and Meyer, C. P. (2013a). Effectiveness of annealing blocking primers versus restriction enzymes for characterization of generalist diets: unexpected prey revealed in the gut contents of two coral reef fish species. *PLoS ONE* 8:e58076. doi: 10.1371/journal.pone.0058076

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., et al. (2013b). A new versatile primer set targeting a shot fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* 10:34. doi: 10.1186/1742-9994-10-34

Lima, M. R. L., Bessa, E., Krinski, D., and Carvalho, L. N. (2012). Mutilating predation in the Cheirodontinae *Odontostilbe pequira* (Characiformes: Characidae). *Neotrop. Ichthyol.* 10, 361–368. doi: 10.1590/S1679-62252012000200011

Machado, F. A., and Sazima, I. (1983). Comportamento alimentar do peixe hematófago *Branchioica bertonii* (Siluriformes, Trichomycteridae). *Cienc. Cult.* 35, 344–348.

Mercado-Silva, N., Lyons, J., Moncayo-Estrada, R., Gesundheit, P., Krabbenhoft, T. J., Powell, D. L., et al. (2015). Stable isotope evidence for trophic overlap of sympatric Mexican Lake Chapala silversides (Teleostei: Atherinopsidae: *Chirostoma* spp.). *Neotrop. Ichthyol.* 13, 389–400. doi: 10.1590/1982-0224-20140079

Murray, D. C., Bunce, M., Cannell, B. L., Oliver, R., Houston, J., White, N. E., et al. (2011). DNA-based faecal dietary analysis: a comparison of qPCR and high throughput sequencing approaches. *PLoS ONE* 6:e25776. doi: 10.1371/journal.pone.0025776

Neto, C. S., and de Pinna, M. (2016). Redescription of *Ochmacanthus batrachostoma* (Miranda-Ribeiro, 1912) (Siluriformes: Trichomycteridae): a possible case of incipient paedomorphism. *Neotrop. Ichthyol.* 14:e150030. doi: 10.1590/1982-0224-20150030

Oliveira, I. R., and Soares, L. S. H. (1996). Alimentação da tainha *Mugil platanus* Günther, 1880 (Pisces: Mugilidae), da região estuarino-lagunar de Cananéia, São Paulo, Brasil. *Bol. Instit. Pesca.* 23, 95–104.

Orciolli, M. C. G., and Bennemann, S. T. (2006). Dieta de *Bryconamericus iheringii* (Ostariophysi: Characidae) em riachos da bacia do rio Tibagi, Estado do Estado do Paraná. *Acta Sci. Biol. Sci.* 18, 59–63. doi: 10.4025/actascibiolsci.v28i1.1060

Paine, R. T. (1980). Food webs: linkage, interaction strength and community infrastructure. *J. An. Ecol.* 49, 667–685. doi: 10.2307/4220

Paquin, M. M., Buckley, T. W., Hibpshman, R. E., and Canino, M. F. (2014). DNA-based identification methods of prey fish from stomach contents of 12 species of eastern North Pacific groundfish. *Deep Sea Res. Part I: Oceanogr. Res. Pap.* 85, 110–117.doi: 10.1016/j.dsr.2013.12.002

Pompanon, F., Deagle, B. E., Symondson, W. O. C., Brown, D. S., Jarman, S. V., and Taberlet, P. (2012). Who is eating what: diet assessment using next generation sequencing. *Mol. Ecol.* 21, 1931–1950. doi: 10.1111/j.1365-294X.2011.05403.x

Post, D. M. (2002). Using stable isotopes to estimate trophic position: models, methods and assumptions. *Ecology* 83, 703–718. doi: 10.1890/0012-9658(2002)083[0703:USITET2.0.CO;2]

Reis, E. G., and D'Incao, F. (2000). The present status of artisanal fisheries of extreme Southern Brazil: an effort towards community-based management. *Ocean Coast. Manag.* 43, 585–595. doi: 10.1016/S0964-5691(00)00048-X

Roberts, T. R. (1972). Ecology of fishes of the Amazon and Congo Basins. *Bull. Mus. Comp. Zool.* 143, 117–149.

Sazima, I. (1983). Scale-eating in characoids and other fishes. *Environm. Biol Fishes.* 9, 87–101. doi: 10.1007/BF00690855

Schifino, L. C., Fialho, C. B., and Verani, J. R. (2004). Fish community composition, seasonality and abundance in Fortaleza Lagoon, Cidreira. *Braz. Arch. Biol. Technol.* 47, 755–763. doi: 10.1590/S1516-89132004000500011

Sheppard, S. K., and Harwood, J. D. (2005). Advances in molecular ecology: tracking trophic links through predator-prey food-webs. *Funct. Ecol.* 19, 751–762. doi: 10.1111/j.1365-2435.2005.01041.x

Speranza, E. D., Tatone, L. M., Cappelletti, N., and Colombo, J. C. (2013). Cost-benefit of feeding on anthropogenic organic matter: lipid changes in a detritivorous fish (*Prochilodus lineatus*). *Ichthyol. Res.* 60, 334–342. doi: 10.1007/s10228-013-0355-x

Sukhdeo, M. V. K. (2012). Where are the parasites in food webs? *Paras. Vect.* 5:239. doi: 10.1186/1756-3305-5-239

Sukhdeo, M. V. K., and Hernandez, A. D. (2005). "Food web patterns and the parasite's Perspective," in *Parasitism and Ecosystems*, eds F. Thomas, J. F. Guegan, and F. Renaud (Oxford: Oxford University Press), 54–67.

Symondson, W. O. C. (2002). Molecular identification of prey in predator diets. *Mol. Ecol.* 11, 627–641. doi: 10.1046/j.1365-294X.2002.01471.x

Taguchi, T., Miura, Y., Kruegh, D., and Sugiura, S. (2014). Utilizing stomach content and faecal DNA analysis techniques to assess the feeding behavior of largemouth bass *Micropterus salmoides* and bluegill *Lepomis macrochirus*. *J. Fish Biol.* 84, 1271–1288. doi: 10.1111/jfb.12341

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA 6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Thompson, J. N. (1982). *Interaction and Coevolution.* New York, NY: John Wiley & Sons.

Valentini, A., Miquel, C., Nawaz, M. A., Bellemain, E. V. A., Coissac, E., Pompanon, F., et al. (2009). New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach. *Mol. Ecol. Resour.* 9, 51–60. doi: 10.1111/j.1755-0998.2008.02352.x

Vestheim, H., and Jarman, S. N. (2008). Blocking primers to enhance PCR amplification of rare sequences in mixed samples — a case study on prey DNA in Antarctic krill stomachs. *Front. Zool.* 5:12. doi: 10.1186/1742-9994-5-12

Winemiller, K. O., and Polis, G. A. (1996). "Food webs: what can they tell us about the world?," in *Food Webs: Integration of Patterns and Dynamics*, eds G. A. Polis and K. O. Winemiller (New York, NY: Chapman and Hall), 1–23.

Winemiller, K. O., and Yan, H. Y. (1989). Obligate mucus-feeding in a South American trichomycterid catfish (Pisces: Ostariophysi). *Copeia* 1989, 511–514. doi: 10.2307/1445457

Zuanon, J., and Sazima, I. (2004). Vampire catfishes seek the aorta not the jugular: candirus of the genus *Vandelia* (Trichomycteridae) feed on major gill arteries of host fishes. *J. Ichthyol. Aqua. Biol.* 8, 31–36. doi: 10.1080/02705060.2014.967816

Zuanon, J., and Sazima, I. (2005). Free meals on long-distance cruisers: the vampire fish rides giant catfishes in the amazon. *Biota Neotrop.* 5, 109–114. doi: 10.1590/S1676-06032005000100012

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

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

frontiers in Genetics

REVIEW
published: 19 March 2018
doi: 10.3389/fgene.2018.00088

Fishing Into the MicroRNA Transcriptome

*Marcos E. Herkenhoff[1], Arthur C. Oliveira[1], Pedro G. Nachtigall[1], Juliana M. Costa[1], Vinicius F. Campos[2], Alexandre W. S. Hilsdorf[3] and Danillo Pinhal[1]**

[1] Laboratory of Genomics and Molecular Evolution, Department of Genetics, Institute of Biosciences of Botucatu, Sao Paulo State University, Botucatu, Brazil, [2] Laboratory of Structural Genomics (GenEstrut), Graduate Program of Biotechnology, Technology Developmental Center, Federal University of Pelotas, Pelotas, Brazil, [3] Unit of Biotechnology, University of Mogi das Cruzes, Mogi das Cruzes, Brazil

In the last decade, several studies have been focused on revealing the microRNA (miRNA) repertoire and determining their functions in farm animals such as poultry, pigs, cattle, and fish. These small non-protein coding RNA molecules (18–25 nucleotides) are capable of controlling gene expression by binding to messenger RNA (mRNA) targets, thus interfering in the final protein output. MiRNAs have been recognized as the main regulators of biological features of economic interest, including body growth, muscle development, fat deposition, and immunology, among other highly valuable traits, in aquatic livestock. Currently, the miRNA repertoire of some farmed fish species has been identified and characterized, bringing insights about miRNA functions, and novel perspectives for improving health and productivity. In this review, we summarize the current advances in miRNA research by examining available data on Neotropical and other key species exploited by fisheries and in aquaculture worldwide and discuss how future studies on Neotropical fish could benefit from this knowledge. We also make a horizontal comparison of major results and discuss forefront strategies for miRNA manipulation in aquaculture focusing on forward-looking ideas for forthcoming research.

Keywords: aquaculture, farm animals, teleost fish, gene expression, microRNAs

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Diogo Teruo Hashimoto, Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP), Brazil
Adalberto Luis Val, Brazilian National Institute for Research in the Amazon, Brazil

****Correspondence:***
Danillo Pinhal
dlpinhal@ibb.unesp.br

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

Received: *18 October 2017*
Accepted: *02 March 2018*
Published: *19 March 2018*

Citation:
Herkenhoff ME, Oliveira AC, Nachtigall PG, Costa JM, Campos VF, Hilsdorf AWS and Pinhal D (2018) Fishing Into the MicroRNA Transcriptome. Front. Genet. 9:88. doi: 10.3389/fgene.2018.00088

INTRODUCTION

MicroRNAs (miRNAs) are a class of small (17–22 nucleotides), non-coding RNAs that inhibit gene expression post-transcriptionally by pairing with complementary sequences in their target mRNA (**Box 1**). These small regulatory molecules are present in the genome of animals, plants, and even some viruses (Lee et al., 1993; Bartel, 2004; Kim et al., 2009; Xia et al., 2011). The first miRNAs, lin-4 and let-7, were both discovered in *Caenorhabditis elegans* (Lee et al., 1993; Wightman et al., 1993; Reinhart et al., 2000) and have subsequently been found to correspond to a novel and extensive class of small non-coding RNAs (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001).

BOX 1 | miRNA biogenesis and action.

The miRNA genes are firstly transcribed in the nucleus by RNA polymerase II to generate primary miRNAs (pri-miRNAs). In the canonical pathway, the pri-miRNAs bend into hairpins and are processed by the RNase III Drosha to form precursor miRNAs (pre-miRNAs) with around 70 nt, characterized by the stem-loop structure (Lee et al., 2003; Kim, 2005). The pre-miRNA is transported from the nucleus to cytoplasm by Exportin-5 and is recognized and cleaved by the RNase III Dicer to form a double-stranded RNA with 22 nt (Hutvágner et al., 2001; Lund et al., 2004). The double strands unbind and one of the strands enters in the RNA-induced silencing complex (RISC), becoming a mature miRNA that can exercise its gene silencing function, whereas the other strand is generally degraded or, in some cases, can also form an RISC (Khvorova et al., 2003).

In addition to the canonical pathway, there are alternative pathways that can promote miRNA biogenesis. One of them is the "Drosha/DGCR8-independent pathway" (Ha and Kim, 2014). In this process, the pre-miRNA (also called mirtron) comprises the whole intronic region of a gene. Thus, when the spliceosome cleaves the transcript, the pre-miRNA is released, and its biogenesis follows the standard pathways. The "TUTase-dependent pathway" (Ha and Kim, 2014) is another type of alternative processing of pri-miRNAs. In this pathway, TUT family proteins binds to pre-miRNAs with only one nucleotide overhang in the 3' region and adds a uridine in its 3' end, promoting the two-nucleotide overhang, important for the correct miRNA maturation. Another type of non-canonical pathway is the "Dicer-independent pathway" (Ha and Kim, 2014). During this process, AGO2 proteins bind directly to the pre-miRNA, and along with PARN ribonucleases, promote the cleavage and final maturation of the miRNA.

All the aforementioned pathways lead to miRNA maturation and subsequent miRNA activity occurs through miRNA-RISC complex binding mainly to the 3′ UTR region of target mRNA, inhibiting the expression or degrading the mRNA transcript (Lee and Dutta, 2009). Besides the 3' UTR region, non-canonical miRNA-to-target interactions were described to occur in binding sites located inside exons (Reczko et al., 2012; Hausser et al., 2013) orin the 5' UTR region (Devlin et al., 2010; Zhou and Rigoutsos, 2014). It has been estimated that each miRNA may modulate hundreds of messenger RNAs and a single messenger RNA may have its stability or translation controlled by several miRNAs (Doench and Sharp, 2004; Brenneck et al., 2005; Lim et al., 2005).

BOX 2 | Finding miRNA putative targets.

Searching for a true miRNA-mRNA interaction is a difficult task. Despite the availability of several computational target prediction tools, their results are generally inconsistent and often return variable sets of possible miRNA targets. This happens mainly because the rules governing miRNA target recognition are not completely understood and may vary for each miRNA-target interaction (Ritchie and Rasko, 2014). Moreover, several popular tools such as DIANNA-microT (Maragkakis et al., 2009), miRanda-miRSVR (Betel et al., 2010), and miRWalk (Dweep et al., 2011) were designed exclusively for predicting miRNAs in mammals. Other tools, such as TargetScan (Garcia et al., 2011), miRanda (Enright et al., 2003), PITA (Kertesz et al., 2007), and RNA22 (Miranda et al., 2006) algorithms are frequently used as target prediction tools for animal miRNAs (Witkos et al., 2011; Reyes-Herrera and Ficarra, 2012; Dweep et al., 2013; Peterson et al., 2014).

Once there are several options of target prediction tools available in the literature, researchers often look up for which tool or combination of tools allows for the best quality results. Oliveira et al. (2017) shows that the union of the results provided by TargetScan, miRanda and RNA22 provide a better performance (balance of sensitivity and specificity) than any of these tools alone or other combination of them. However, it is important to keep in mind that any target prediction tool or combinatory use may still return false-positive results.

Since the target prediction tools yet lack consistency, once putative targets have been selected, validation experiments must be performed to confirm if the interaction is genuine. The majority of strategies for miRNA-to-target validation rely on the overexpression and/or knockdown of the miRNA of interest and corresponding surveillance of the effect at the mRNA/protein levels of predicted target genes. To evaluate the mRNA and/or protein expression profiles associated with the manipulation of a specific miRNA, several functional analyses have been widely applied (**Table 1**).

Since then, numerous miRNAs have been identified and quantified in several organisms, primarily as a result of numerous enhancements in high throughput sequencing technologies, bioinformatics computer programs and experimental methods (Guerra-Assumpção and Enright, 2012). Together, these approaches have enabled large-scale analyses, thereby facilitating the discovery of species-specific and novel miRNAs, even those with reduced expression, in many organisms belonging to diverse taxa (Berezikov, 2011).

Owing to the key involvement of miRNAs in the regulation of growth, metabolism, and homeostasis, among hundreds of other functions, diverse studies over the last decade have sought to identify genuine miRNA-to-target interactions in farm animals, such as poultry, pigs, cattle (reviewed by Wang et al., 2013), and fish (Rasal et al., 2016), to maximize production (**Box 2**). In addition, several patents have been approved in various countries for the commercial exploration of miRNAs, e.g., a US patent (US 2006/0246491 A1) using miRNAs to regulate muscle cell growth.

Here, we review the available data on miRNA identification and functional characterization in the most important globally farmed fish species, with an emphasis on Neotropical fish, and discuss the current challenges and future directions for the use of miRNAs in aquaculture.

MiRNAs IN ANIMAL BREEDING

Since miRNAs are key elements in gene regulation, they have been the focus of studies on improving the health and productivity of farm animal species. Diverse studies have been conducted in cattle (Townley-Tilson et al., 2010; Kozomara and Griffiths-Jones, 2011; Li H. et al., 2011; Miretti et al., 2011), pigs (Cho et al., 2010; Cirera et al., 2010; Xie et al., 2010; Li G. et al., 2011; Chen et al., 2012), and birds (Hicks et al., 2008, 2009; Li T. et al., 2011; Yao et al., 2011; Wang et al., 2012b). These studies have reported that miRNAs play an important role in a wide range of biological pathways, such as growth, metabolism, immunology, muscle development, and fat deposition (**Figure 1**). Therefore, these small RNAs are of great value in animal breeding, steering research toward the development of new solutions for current and emerging threats to the health and welfare of farm animal species.

Reports aimed at the identification and functional analysis of miRNAs in fish have increased recently, allowing for the characterization of the miRNA repertoire in Neotropical fish and in other prominent fish species in the aquaculture industry worldwide (**Table 1**).

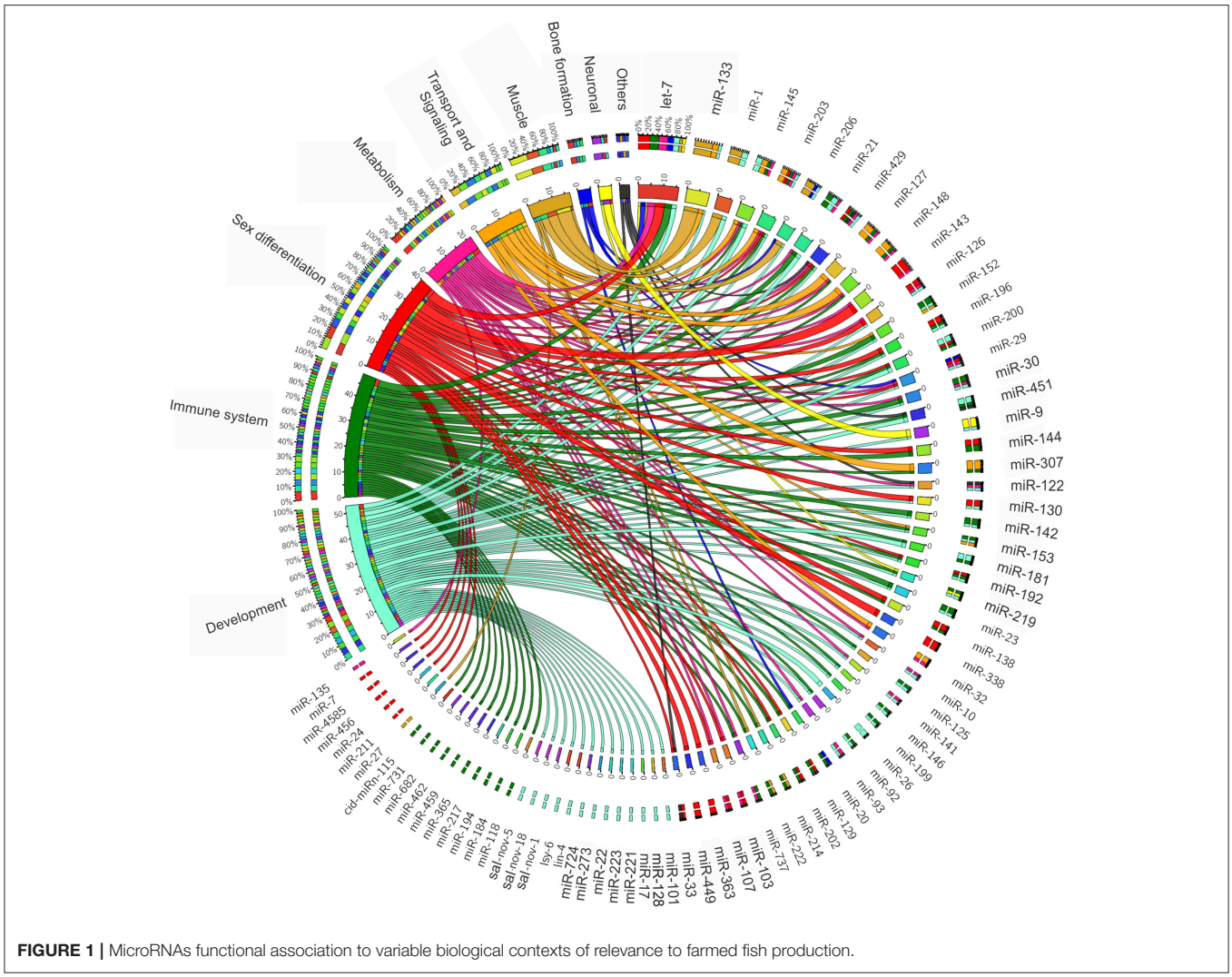


FIGURE 1 | MicroRNAs functional association to variable biological contexts of relevance to farmed fish production.

A REVIEW OF MiRNAs IN FARMED FISHES

Neotropical Fish

Neotropical fish encompass 50% of all freshwater species that are currently identified, which includes ∼13,000 characterized species (Reis et al., 2003; Bertollo et al., 2017). These data highlight the great biodiversity of this group of fishes and, particularly, of those of the Neotropical region. However, despite the growing interest in the study of Neotropical fishes, resources describing miRNA diversity and function are still very limited. Below, we provide a picture of the current knowledge of miRNAs of Neotropical species.

Tambaqui (*Colossoma macropomum*, Characidae)

The black pacu or tambaqui (*Colossoma macropomum*) can reach up to 1 m in length and ∼30 kg in weight and is a main resource of aquaculture and fisheries along the Amazon (Goulding and Carvalho, 1982; IBAMA, 2000). Gomes et al. (2017) characterized the miRNA expression profiles of liver and skin tissues of the Amazon tambaqui. They identified 279 conserved miRNAs in this species, with 257 from the liver and 272 from the skin and with several miRNAs expressed in both tissues. A functional enrichment analysis was performed for the targets of the 10 most highly expressed miRNAs of each tissue, revealing enriched biological pathways associated with metabolism, cell proliferation, calcium/ion transport and carbohydrate kinase activity in the liver as well as regulation of transcription and metabolic processes in the skin.

miR-122, which is highly expressed in the liver of tambaqui (Gomes et al., 2017), is conserved among vertebrate species and is known to be integrated with the regulation of cholesterol metabolism. Moreover, studies in rainbow trout corroborate the influence of miR-122 in liver metabolism (Mennigen et al., 2012, 2013). Thus, miR-122 is an interesting regulatory gene that requires further study regarding its impact on fish metabolism and, consequently, in aquaculture production.

Pacu (*Piaractus mesopotamicus*, Characidae)

The pacu (*Piaractus mesopotamicus*), which is characterized by hardiness, fast growth, adaptation to artificial feeding and

TABLE 1 | Summary of studies on miRNAs of Neotropical and other fish species with relevance on aquaculture.

Group	Species	Tissue	Methodology	References
Neotropical	Tambaqui	Liver and skin	Illumina sequencing technology and functional annotation	Gomes et al., 2017
	Pacu	Skeletal muscle	Real-time PCR	Paula et al., 2017
	Midas cichlids	Not informed by authors	Target prediction and Illumina sequencing technology	Franchini et al., 2016
Others	Nile tilapia	embryo	RNA-seq high-throughput sequencing and miRNA target prediction analysis	Eshel et al., 2014
		Gonads	Solexa sequencing with real-time PCR expression, miRNA target prediction analysis and miRNA pathways analysis	Wang et al., 2016
			Solexa sequencing with real-time PCR expression	Xiao et al., 2014
		Kidney	Real-time PCR and luciferase reporter assay	Yan et al., 2012b,c; Zhao et al., 2016
		Skeletal muscle	Prediction of MyoD-binding miRNAs, real-time PCR and luciferase reporter assay	Khvorova et al., 2003; Yan et al., 2013b
			Microarray and luciferase reporter assay	Kim et al., 2006
			Silencing of miR-206 *in vivo* using the antagomir method	Yan et al., 2013b
	Atlantic Salmon	Blood	RNA-seq high-throughput sequencing	Kure et al., 2013
	Rainbow trout	Larvae	Real-time PCR	Ramachandra et al., 2008
		Liver	Prediction of miRNA targer gene and real-time PCR	Mennigen et al., 2012
		Muscle	RNA-seq high-throughput sequencing and target prediction	Salem et al., 2010
	Grass carp	Blood	Solexa sequencing technology, real-time PCR, directional cloning and induction of expression of three immune factors in vitro.	Xu et al., 2016
		Embryo	Real-time PCR	Xu et al., 2014
		Spleen	Solexa sequencing technology, real-time PCR, target prediction and integration with mRNA expression and luciferase report system	Xu et al., 2015

flavorsome meat (Castagnolli and Cyrino, 1986), is a valuable resource in Brazilian aquaculture (Urbinati and Gonçalves, 2005). Duran et al. (2015) and Paula et al. (2017) studied the role of skeletal muscle miRNAs in this Neotropical fish during development and under food restriction, respectively. Duran et al. (2015) analyzed the impact of the miRNA-target interactions of miR-1/*hdac4*, miR-133-a/b/*srf*, miR-206/*pax7*, and miR-499/*sox6* in fast- and slow-twitch skeletal muscles during growth. They found that miR-1 and miR-206 may promote myoblast differentiation in fast- and slow-twitch muscles in adult individuals, while miR-133a/b acts earlier, promoting myoblast proliferation in juveniles. Finally, miR-499 may promote differentiation of slow-twitch muscle fibers, which corroborates previous findings in Nile tilapia (Nachtigall et al., 2015). Paula et al. (2017) showed that a short period of food restriction significantly increased the expression of miR-1, miR-206, miR-199, and miR-23a in fast muscle and significantly decreased the expression of miR-1 and miR-206 in slow muscle, while their targets (*IGF-1* for miR-1, miR-206, and miR-199; *mTOR* for miR-199; and *MFbx* and *PGC1a* for miR-23a) exhibited negatively correlated expression profiles. Altogether, these data suggest that miRNAs may have a large impact on muscle plasticity and recovery after periods of food restriction.

MiR-1, miR-133a, miR-133b, miR-206, and miR-499 have been shown to be involved in the control of genes related to myoblast proliferation and differentiation. The same miRNAs were also found in the skeletal muscle of other farmed animals and exhibited a highly conserved genomic context (i.e., structural features) among fish species (Nachtigall et al., 2014). These findings underpin the putative applicability of these muscle miRNAs as valuable biomarkers in Neotropical fish breeding. Additional experiments focusing on the characterization of SNPs in miRNAs and target genes from distinct populations exhibiting variable phenotypic traits, such as increased muscle growth and body mass, are required.

Midas cichlids (*Amphilophus* spp., Cichlidae)

The Midas cichlids are Neotropical fishes from Nicaraguan crater lakes. These species are considered to be an excellent model for studying speciation due to their fast parallel sympatric speciation. In contrast to other research that focused on miRNA expression and function, Franchini et al. (2016) studied the role of miRNAs in the diversification of five species of the Midas cichlid lineage (*Amphilophus* spp.).

Interestingly, by examining the putative effects of mutations of 3'UTR miRNA binding sites on the phenotypic diversification and evolution of these cichlids, the authors found that the species from Lake Apoyo (*A. astorquii* and *A. zaliosus*) have undergone less natural selection compared to species from Lake Nicaragua and Lake Managua (*A. citrinellus*) and Lake Xiloà (*A. amarillo* and *A. sagittae*). The impact of variable natural selection in the genome allowed novel mutations in binding sites to accumulate and the regulatory networks to evolve faster in the latter species, thus contributing to novel miRNA-target interactions associated with the phenotypic diversification of Midas cichlids (Franchini et al., 2016). These findings corroborate previous data regarding the "evolutionary race" between miRNAs and targets that have been proposed to influence the evolution of African

TABLE 2 | Summary of miRNA associated functions described in each species reviewed in the manuscript.

Group	Species	miRNA	Associated function
Neotropical	Tambaqui	miR-122	Regulation of cholesterol metabolism
		miR-1; miR-206	Myoblast differentiation
	Pacu	miR-133a/b	Myoblast proliferation
Others	Nile tilapia	miR-499	Differentiation of slow-twitch muscle fibers
		miR-1; miR-206	Myoblast differentiation
		miR-133a/b	Myoblast proliferation
		miR-30c	Regulation of salt tolerance
		miR-429	Regulation of osmotic stress
		miR-21	Regulation of alkaline tolerance
		miR-122	Regulation of stress response
		miR-29; miR-129	Overexpressed in Ovary
		miR-129	Ovary maturation
		miR-29	Oogenesis
		miR-33a; miR-132; miR-135b; miR-212	Overexpressed in Testis
		miR-33a	Testis maturation
		miR-4585	Sex dettermination
		miR-17; miR-20a	Regulation of estrogen production
		miR-138; miR-338; miR-200a	Spermatogonial cell proliferation and spermatogenesis
		miR-456; miR-138	Testis differentiation
		miR-92d	Regulation of C3 complement expression
		miR-310, miR-92, and miR-127	Overexpressed in *Streptococcus agalactiae* infection
		miR-92d, miR-375, miR-146, and miR-694	Downregulated in *Streptococcus agalactiae* infection
	Atlantic Salmon	let-7a; miR-142; miR-144; miR-430; miR-451; miR-1594	Regulation of growth
		miR-29a	Collagen formation
		miR-20a; miR-25; miR-181a; miR-202; let7c/d/a; miR-125b; miR-222a/b; miR337 190a; miR-2188; miR-144; miR-731; miR-8157	Initiation of puberty
		mir-10, mir-21, mir-30, mir-181 and let7	Overexpressed in *Caligus rogercresseyi infection*
		miR-21, miR-181 and miR-30	Overexpressed in *Piscirickettsia salmonis* infection
	Rainbow trout	miR-21, miR-30d, miR-92a, miR-200, and miR-26	Cell differentiation and embryo development
		miR-103, miR-107, and miR-143	Post-prandial regulation of lipid and glucose metabolism
		miR-33; miR-122b	Regulation of lipogenesis and lipolysis
	Grass carp	miR-1a, miR-181a, miR-133a, miR-214, miR-133b, miR-206, miR-146, and miR-26a	Fast-twitch skeletal muscle growth

cichlids (Loh et al., 2011; Brawand et al., 2014), thus indicating the pervasiveness of mutual molecular mechanisms of miRNA functional diversification in fish.

From an aquaculture perspective, such miRNA-target evolutionary mechanisms have selected adaptive traits for a specific environment, which can be tracked to identify economically valuable phenotypes. Moreover, this study reinforces the importance of examining population polymorphisms of both miRNAs and targets while performing fish genetic breeding programs.

Other Key Fish Species

Studies of miRNAs in Neotropical fishes are very recent, still providing a limited source of data regarding the impact of these molecules in Neotropical fish biology. Thus, in this section, we describe advances based on investigations of miRNA repertoires from other four fish species of relevance in aquaculture and fisheries worldwide and discuss the putative transferability of the acquired knowledge to Neotropical fishes.

Nile tilapia (*Oreochromis niloticus*, Cichlidae)

Nile tilapia is one of the most commonly farmed species in freshwater aquaculture worldwide. A fast growth rate and adaptation to a wide range of culture conditions are the attributes responsible for successful fish farming (Yue et al., 2016).

Several studies aimed at miRNA characterization have been carried out in this model species. Using bioinformatics, Loh et al. (2011) analyzed the evolution of miRNAs and their target genes in cichlid fishes from Lake Malawi (East Africa) and detected 100 cichlid miRNAs, including those from tilapia, that are highly

conserved in metazoan genomes. Brawand et al. (2014), who also studied cichlid miRNAs, described diverse loci in tilapia comprising 7 novel miRNAs. Yan et al. (2012a), using small RNA cloning and Sanger sequencing, discovered 25 conserved miRNAs in tilapia skeletal muscle.1'. Huang et al. (2012) employed next-generation sequencing (NGS) in muscle tissue to explore the complete miRNA transcriptome and identified the expression of 184 known mature miRNA sequences.

The functional roles of miRNAs have also evaluated under variable contexts. Several studies have investigated muscle development because muscle constitutes the major edible part of fish and is therefore an economically important trait. Many miRNAs have been proven to participate in the regulation of muscle growth by controlling the genes involved in hyperplasia and hypertrophy (Lagos-Quintana et al., 2002). Studies that focused on the initial embryonic development have also clarified miRNA targets involved in the tightly regulated initial steps of body formation (Giusti et al., 2016).

In tilapia, Yan et al. (2012a) and Nachtigall et al. (2015) showed that miR-1, miR-133a, and miR-206 have similar expression patterns in adult males and females and may assist each other to accurately control the development of skeletal muscles, although they perform distinct biological functions. MiR-1 is responsible for repressing the expression of histone deacetylase 4 (HDAC4), which is a negative regulator of cellular differentiation and thus promotes myocyte differentiation. MiR-1 also blocks a repressor of the MEF2 (myocyte enhancer factor-2) transcription factor. MiR-133a promotes, in part, myocyte proliferation by repressing serum response factor (SRF) (Chen et al., 2006), whereas miR-206 plays an important role in regulating the differentiation of C2C12 myoblasts *in vitro* (Kim et al., 2006). In addition, miR-206 loss of function *in vivo* was shown to significantly improve tilapia growth performance by targeting insulin-like growth factor-1 (IGF-1) (Yan et al., 2013b). IGF-1 is known to play a central role in a complex system that regulates growth, differentiation, and reproduction by selectively promoting mitogenesis and cell differentiation and inhibiting apoptosis (Jones and Clemmons, 1995; Reinecke and Collet, 1998). In many fish species, IGF-1 blood or tissue levels positively correlate with dietary protein levels and body growth rate (Beckman et al., 2004; Carnevali et al., 2006). Therefore, miR-206 could affect tilapia growth by modulating IGF-1 gene expression levels (Yan et al., 2014). Similarly, miR-203b has been shown to promote myogenesis (hyperplastic growth) by targeting MyoD (Yan et al., 2013a), a key protein that initiates the cascade of regulatory events during muscle differentiation. These authors showed that blocking miR-203b results in a significant increase in MyoD expression. In comparison to other approaches, however, using metabolic alterations, Wang D. et al. (2017) showed that miR-223-3p promotes upregulation of pou1f1, a key transcription factor related to somatic growth in Nile tilapia. Pou1f1 binds to the promoter region and transactivates the expression of growth hormone (GH), prolactin (PRL), and somatolactin (SL) in teleost fish. Furthermore, Qiang et al. (2017c) showed that miR-29a antagomir treatment i*n vivo* resulted in stearoyl-CoA desaturase (SCD) upregulation, which plays a role in hepatic lipid metabolism regulation (Dobrzyn and Ntambi, 2004; Ntambi and Miyazaki, 2004).

Other studies have revealed miRNAs regulating metabolic pathways that could indirectly enhance production. For example, there is growing concern about the genetic improvement of salt tolerance in Nile tilapia, which may provide advantages, such as adaptation to certain environmental conditions and higher oxidation. Tilapia are euryhaline fish, and most species can live in a wide range of salinities from freshwater to seawater and therefore are a suitable model organism for studies on ionic and osmotic acclimation in euryhaline teleosts (Deane and Woo, 2004; Wang et al., 2009). For example, miR-30c, a kidney-enriched miRNA, was shown to regulate salt tolerance because its loss of function caused fish to be unable to respond to osmotic stress (Yan et al., 2012b, 2013b). Moreover, osmotic stress transcription factor 1 (OSTF1) was shown to be potentially regulated by miR-429 (Yan et al., 2012c). Recently, Zhao et al. (2016) showed that miR-21 is abundantly expressed and its action modulates alkalinity stress by upregulating VEGFB and VEGFC expression *in vivo* and *in vitro*, which are responsible for regulating alkalinity tolerance. Therefore, miR-21, miR-30c, and miR-429 may be important markers for tilapia and also for other commercial species. Nevertheless, Qiang et al. (2017d) concluded that miR-122 plays an important role in regulating the stress response in the lineage of Nile tilapia liver exposed to cadmium.

Another relevant topic for Nile tilapia production involves sex determination and differentiation, because males grow faster and larger than the female. Therefore, several studies have sought to uncover the molecular mechanisms of sex-determination regulated by miRNAs (Eshel et al., 2014; Xiao et al., 2014).

In the study by Xiao et al. (2014), gonads (testes and ovaries) were screened using high-throughput sequencing, and the data showed distinct miRNA expression signatures. In addition, Nile tilapia testes and ovaries displayed miR-181a, miR-181a-5p, miR-143, and miR-143-3p as the most abundant miRNAs. By contrast, the miR-29 and miR-129 families showed significantly increased expression in ovaries compared to testes (Xiao et al., 2014). In humans, the high expression of miR-129 in ovaries is associated with the control of cell growth and differentiation in the final process of ovary maturation through downregulation of its target mRNAs, whereas miR-29 expression levels, which progressively increase throughout oogenesis, may also be important (Sirotkin et al., 2009). In tilapia testes, the most abundantly expressed miRNA families are miR-33a, miR-132, miR-135b, and miR-212 (Xiao et al., 2014). In mammals, miR-212/132 expression is necessary for the development and function of neurons. Furthermore, both miRNAs are associated with mammary gland development by downregulating the matrix metalloproteinase 9 (MMP-9), which is an activator of TGFβ and is involved in cell proliferation, differentiation, and apoptosis (Kubiczkova et al., 2012). miR-33a controls fatty acid regulation in mammals by repressing insulin receptor substrate 2 (Dávalos et al., 2011), and its high expression in tilapia testes may contribute to testes maturation via regulation of the insulin signaling pathway (Xiao et al., 2014).

Eshel et al. (2014) compared miRNA expression in tilapia gonads and found nine sexually dimorphic expressed miRNAs;

they found a single upregulated miRNA in male embryos (miR-4585) with a perfect inverse correlation in expression pattern with its six target genes, cr/20β-hsd, psmb8, rtn4ip1, casp8, atp5g3, and a non-annotated gene (downregulated in males). cr/20β-hsd is known to be part of the oxidoreductase pathway for oocyte maturation preceding the enzymatic activity of cyp19 (cytochrome P450 aromatase) (Senthilkumaran et al., 2004), and cyp19a1a is proposed to be the major gene for female determination in zebrafish (*Danio rerio*) (Rodríguez-Marí et al., 2010). MiR-4585 is upregulated in male embryos at 2- and 5-days post-fertilization (dpf) and decreases at 9 dpf, which indicates the significance of this miRNA in males soon after fertilization (Eshel et al., 2014). Therefore, miR-4585 could possibly be manipulated for sex reversal in tilapia. Similarly, control of cyp19a1a expression may be relevant for sex reversal of females into males, given the aforementioned differential growth between sexes. Recently, Wang et al. (2016), using ovaries and testes of young Nile tilapia, showed that miR-17-5p and miR-20a were highly expressed in the ovaries and negatively regulated DMRT1 expression, suggesting that these miRNAs could induce estrogen production by inhibiting DMRT1 expression and promoting cyp19a1a expression in Nile tilapia. They found that miR-138, miR-338, and miR-200a negatively regulated cyp17a2 (Wang et al., 2016), which is involved in 20b-dihydroxy-4-pregnen-3-one biosynthesis and thus might be essential for spermatogonial cell proliferation and spermatogenesis (Eshel et al., 2014). Wang et al. (2016) also showed that the miRNAs miR-456 and miR-138 negatively regulate AMH and that lower expression of these miRNAs promotes testis differentiation by allowing AMH to be expressed in testis.

Another aspect to be investigated is immunity. Infected animals cause economic losses and risks to human health. *Streptococcus iniae* causes high mortality and huge economic losses in tilapia cultures in China. This outbreak has seriously hindered the development of the tilapia industry (Qiang et al., 2017a). Qiang et al. (2017a) found that miR-92d regulates the expression of complement C3, which is the central component of the immune system. Furthermore, Qiang et al. (2017b) found three miRNAs (miR-310, miR-92, and miR-127) that were upregulated and four (miR-92d, miR-375, miR-146, and miR-694) that were downregulated by comparing a control group of Nile tilapia with a group infected with *Streptococcus iniae*. In another study of infection by Streptococcus, but using *Streptococcus agalactiae*, Wang et al. (2016) applied high-throughput sequencing followed by functional enrichment and found1121 differentially expressed miRNAs that target 41961 genes during infection. These analyses provide data for future studies of host-pathogen interactions among Nile tilapia and *S agalactiae*. In addition, Wang B. et al. (2017) identified 1981 miRNAs involved in the immune response against meningo encephalitis caused by *Streptococcus agalactiae* in Nile tilapia. Although these data are important, the studies of Wang et al. (2016) and Wang B. et al. (2017) are deficient because of their failure to find miRNA biomarkers that are more effective or that can be manipulated.

Atlantic salmon (*Salmo salar*, Salmonidae)

Atlantic salmon is a domesticated fish of notable economic interest for wild fisheries and aquaculture production. Since 1970, Atlantic salmon have been intensively selected for genetic traits to improve growth performance, considerably benefiting aquaculture (Bentsen and Thodesen, 2005). The increased growth of this species is estimated at ∼14% per generation (Gjedrem, 2010). Genome knowledge of salmonids is advanced in comparison to other farmed fish species (Bekaert et al., 2013); however, there are limited studies describing salmonid miRNAs.

To provide better tools for future analysis of miRNAs, Zavala et al. (2017) identified and validated appropriate endogenous reference miRNA genes. Zavala et al. (2017) showed that the ssa-miR-99-5p gene was the most stable overall and that ssamiR-99-5p and ssa-miR-23a-5p were the best combination.

The first screen for salmonid miRNAs was performed by Barozai (2012) using an *in silico* approach to predict miRNAs in salmon genomes. They detected let-7a-3p as a regulator of zonadhesin-like and growth hormone 2 gene, miR-142-5p as a regulator of heparin-binding growth factor 1, and miR-144 as a regulator of the growth factor receptor-bound protein 2. They also found that miR-430 regulates the transforming growth factor-beta-induced protein ig-h3, miR-451 blocks the anti-dorsalizing morphogenic protein, and miR-1594 activates both titin-cap (telethonin)-like mRNA and growth hormone receptor isoform 2. All of these miRNAs and targets are associated with growth and developmental processes.

Andreassen et al. (2013) characterized miRNA genes in Atlantic salmon by performing deep sequencing analysis of small RNA libraries from nine different tissues and revealed a total of 180 evolutionarily conserved mature miRNAs and 13 distinct novel mature miRNAs. Later, Bekaert et al. (2013) deep sequenced miRNA libraries of young juveniles (4 months old) and identified 547 miRNA transcripts that mapped to 88 miRNA distinct genes.

Johansen and Andreassen (2014) validated miR-25-3p and miR-455-5p as the best-performing two-reference gene combination suitable for quantitative expression analysis in Atlantic salmon. Such findings are relevant for aquaculture because these two miRNA reference genes are now useful for appropriate diagnostics tests, such detection of infection by viral RNA through expression. In addition, Kure et al. (2013) identified the differential expression of miRNAs subjected to an acid environment through RNA-Seq. They found 4 down- and 14 upregulated miRNAs between exposed groups and the control, suggesting alterations in a number of physiological responses that ultimately may interfere in animal growth performance.

The modulatory effect of miRNAs in collagen formation in salmon muscle has also been investigated (Mitchie, 2001). Muscle firmness is appreciated by consumers, and inadequate processing of salmon meat reduces firmness in selected adult salmons (Moreno et al., 2016). miR-29a, which is highly conserved between *Salmo salar* and *Danio rerio* (Andreassen et al., 2013), is the major factor in collagen formation and showed low expression in human fibroblasts with systemic sclerosis (Maurer et al., 2010). Thus, miR-29a may be a future target in expression

studies to improve filet quality with a high potential for applications in aquaculture.

Furthermore, in salmon aquaculture, precocious puberty creates welfare problems and consequently inflicts economic damage. Skaftnesmo et al. (2017) explored which miRNAs regulate mRNAs during initiation of puberty, and several regulated miRNAs in the pubertal stage had earlier been associated (miR-20a, miR-25, miR-181a, miR-202, let7c/d/a, miR-125b, miR-222a/b, miR-190a) or have now been found connected (miR-2188, miR-144, miR-731, miR-8157) to the initiation of puberty.

In addition, there have been some studies on miRNAs associated with immunity. Andreassen et al. (2017) identified miRNAs responding to salmonid alphavirus (SAV) at different time points post-infection. They identified 20 differentially expressed miRNAs that may be important in viral-host interactions in Atlantic salmon. Following the same field of interest, Valenzuela-Miranda et al. (2017) studied the expression of miRNAs during infection with *Piscirickettsia salmonis* in the head, kidney and spleen. *Piscirickettsia salmonis*, a facultative intracellular bacterium, causes salmonid rickettsial septicemia (SRS), which is associated with major mortality in salmonid aquaculture (Mauel and Miller, 2002). The miRNA families miR-181, miR-143, and miR-21 were the most abundant in control groups, while miR-21, miR-181, and miR-30 were the most abundant in animals infected with *P. salmonis* (Valenzuela-Miranda et al., 2017). Sea louse *Caligus rogercresseyi*, which affects Chilean aquaculture, were studied during infestation in Atlantic salmon and the most abundant families were mir-10, mir-21, mir-30, mir-181, and let7 in skin, head and kidney (Valenzuela-Muñoz et al., 2017).

Overall, the collection of experimentally identified Atlantic salmon miRNAs provide an important resource for functional genome research in Neotropical fish and, in particular, helps to determine the actual contribution of miRNAs to phenotypic variation in economic and biologically fundamental characteristics.

Rainbow trout (*Oncorhynchus mykiss*, Salmonidae)

Rainbow trout, the most widely cultivated cold freshwater fish in the world, has been improved in terms of its growth and development by animal breeding programs globally (Sae-Lim et al., 2013). This species is also a model organism for genome-related research on comparative immunology, carcinogenesis, toxicology, disease ecology, physiology, evolutionary genetics, and nutrition (Thorgaard et al., 2002).

Juanchich et al. (2016) provided a repertoire of 2946 miRNA loci in the rainbow trout genome, including 445 already known, in 38 different samples corresponding to 16 different tissues. These data represent the first characterization from a wide variety of tissues and provide a novel resource for research. Subsequently, Mennigen and Zhang (2016) created and reported the microtrout, a comprehensive database, which was developed to implement an algorithm to predict relationships among miRNA-mRNA targets.

Ramachandra et al. (2008) discovered 14 miRNAs in early embryos (5 dpf); among these, miR-21, miR-30d, miR-92a, miR-200, and miR-26 are associated with differentiation and development. Salem et al. (2010) showed that miR-133, which is known to be enriched in mammalian muscle (Shingara et al., 2005; Ason et al., 2006; Chen et al., 2006), is also highly expressed in trout skeletal muscle and therefore is of interest in animal breeding.

Other miRNAs involved in growth and nutritional support metabolic pathways have also been reported. A pioneer study by Mennigen et al. (2012) describes the post-prandial regulation of lipid and glucose metabolism by miRNAs. Among these miRNAs, they predicted conserved targets in fish and humans for miR-103, miR-107, and miR-143 homologs. Research has also shown miR-33 and miR-122b were upregulated, whereas miR-122a was downregulated. MiR-33 and miR-122b act on the hepatic insulin pathway to stimulate lipogenesis and inhibit lipolysis (Fernández-Hernando et al., 2011). MiR-33 and miR-122b may promote lipogenesis and simultaneously inhibit lipolysis. The inhibition of miR-33 expression in mice (*Mus musculus*) resulted in a low level of VLDL triglyceride (Fernández-Hernando et al., 2011). Similarly, inhibition of miR-122 resulted in increased fatty acid oxidation and decreased fatty acid synthesis rates (Krützfeldt et al., 2005; Esau et al., 2006; Elmén et al., 2008). MiR-122b was inhibited alongside an increase in plasma triglycerides, a final product of the lipogenic pathway (Li et al., 2009; Mennigen et al., 2012). Mennigen et al. (2013) analyzed the expression of miRNAs by quantitative real-time RT-PCR in rainbow trout fingerlings while switching from endogenous to exogenous feeding and showed a decrease in miRNA-33 and miRNA-122a/b isomiRs.

Paneru et al. (2017) aimed to investigate the association of miRNA expression with muscle growth and the effects on growth of single nucleotide polymorphisms (SNPs) in miRNA binding sites. They found 90 miRNAs that showed differential expression and were strongly associated with phenotypic variations. Among the 204 SNPs with 3′ UTRs targeted by miRNAs, 78 SNPs may be associated with changes in 5 muscle traits. The expression pattern of 12 miRNAs, including mir-1, mir-133 and mir-206, was validated by real time PCR.

On the basis of these studies, it is possible that down regulation of the above miRNAs could contribute to weight gain in rainbow trout. Therefore, miRNAs could possibly be experimentally manipulated in live animals to improve the production indices in rainbow trout.

Grass carp (*Ctenopharyngodon idella*, Cyprinidae)

Grass carp is extensively cultivated in eastern Asia and is one of the most important freshwater fish globally (Liu et al., 2009). Grass carp was introduced into 115 countries worldwide, and in at least 58 of these (∼50%) it appears to have self-sustaining populations (Fishbase, 2015) and high performance.

Zhu et al. (2014) reported that several miRNAs are likely involved in fast-twitch skeletal muscle growth in grass carp based on their analysis of the response to rapid refeeding following fasting. They recorded changes in the expression levels of eight miRNAs (miR-1a, miR-181a, miR-133a, miR-214, miR-133b, miR-206, miR-146, and miR-26a) shown to be involved in a strong resumption of myogenesis (Zhu et al., 2014).

Xu et al. (2014) studied miRNA expression stability in embryos at distinct developmental stages as well as in several tissues from adults in an effort to establish the best reference genes for quantitative expression analysis. Seven miRNAs (miR-126-3p, miR-101a, miR-451, miR-22a, miR-146, miR-142a-5p, and miR-192) were found to have optimal stability and should be individually prioritized according to the stage and tissue of interest.

Analyzing two grass carp lineages, *Aeromonas* hydrophila-susceptible (SGC) and -resistant (RGC), Xu et al. (2015) found miRNAs related to the immune system. MiR-118 and let-7i are differentially expressed and target both tlr4 and nfil3-6 genes. Let-7i is predominantly expressed in the spleen during bacterial infection and displays a noticeable difference in expression between SGC and RGC (Xu et al., 2015). The let-7 family is known to regulate multiple genes related to the cell cycle and proliferation (Yang et al., 2008), and let-7i was shown to influence innate immunity (Chen et al., 2007). Recently, Xu et al. (2016) studied these two carp lineages using two kidney miRNA transcriptomes from SGC or RGC infected with the highly pathogenic A. hydrophila and identified nine miRNAs differentially expressed between these groups. Furthermore, the spatial and temporal expression of a novel miRNA (cid-miRn-115) and miR-142a-3p suggests that they are potential regulators of anti-bacterial activity (Xu et al., 2016). Overexpression of these miRNAs resulted in a visible change in the immune effector activity in *C. idella* kidney cells, and bioinformatics analysis shows that they directly regulate tlr5 expression (Xu et al., 2016), which is associated with the innate immune response (Yoon et al., 2012).

Therefore, all these miRNAs interfere with grass carp health and should be further investigated for improving fish resistance to diseases that account for great economic losses in aquaculture.

DISCUSSION

Biological Implications of miRNA-Mediated Regulation

In fish, numerous miRNAs have been continuously identified, although only a small fraction have known functions (**Figure 1**; **Table 2**). Regarding Neotropical fish, these numbers are even smaller since very few species have been studied regarding miRNA functions (all species covered in this review). Considering that miRNAs have an important function in shaping both morphological and physiological phenotypes, and its function is highly conserved within vertebrates, analysis of miRNA profiles from other species might potentially be useful to understand and to provide wide applications in Neotropical fish breeding.

Although limited functional studies on farmed fish miRNAs have been conducted, they have provided sufficient evidence for linking miRNA activity to biological aspects of economic interest, such as the control of muscle development and hypertrophy, regulation of oocyte and testis maturation, sex differentiation and physiological resilience to diseases and adaptation to environmental changes. Additionally, deep exploration of classic genetic aspects of miRNA function, such as miRNA mapping, control of expression and genomic context, have contributed to understanding the genetic basis of miRNA expression and have led to new insights into the influence of miRNA on an assortment of complex traits.

Since interactions between miRNAs and their targets can be strongly conserved among species, several interactions verified in previous studies could be extrapolated to Neotropical fish species, thereby remarkably improving fish aquaculture. For instance, the role ofmiR-1, miR-206, and miR-133 during myoblast proliferation and differentiation is recognized to interfere in the hypertrophic growth of skeletal muscle. These interactions has been shown in Nile tilapia (Nachtigall et al., 2015) and may potentially be a widely conserved miRNA-target interaction modulating growth of several Neotropical fishes, as already reported for pacu (Duran et al., 2015; Paula et al., 2017). In the skeletal muscle, the main edible part of the fish, miR-1, miR-133a, and miR-206 have conserved expression patterns in all farmed fish species, which makes them interesting molecules for modulating muscle development and growth. Particularly, in Neotropical fishes, the study of myomiRs (i.e., miRNAs specifically or enriched in this tissue) showed important data that can bring advances to increase their production to its full potential. Moreover, the fact that these miRNAs were found to play the same functional roles in the skeletal muscle of several other animal species corroborates hypotheses about their pivotal importance in gene regulatory networks.

For the large number of non-model fishes that do not have their complete genome sequenced or for those for which little is known about their genome and its interactions, advanced computational tools are becoming sufficiently capable and precise to achieve good and true assessments of the data and genome from model fish species that are evolutionarily close to non-model fish. Consequently, there is less of a gap or challenge between genomic information of non-model and model fishes. Some muscle miRNAs have well-defined target genes, as miR-206 regulates IGF-1; miR-1, miR-122, and miR-462 control IGF-2a; the let-7 family regulates MSTN; miR-103 and miR-107 modulate GHR and FSHR; and miR-138 and miR-211 control LHR. All these genes targeted by muscle miRNAs can be experimentally overexpressed or under expressed through molecular techniques aiming to obtain fishes carrying desired phenotypes. Furthermore, some fish miRNAs have well-defined conserved functions also occurring in humans and mice, and could be controlled to enhance production. For example, miR-29a acts in collagen formation as has conserved targets among fish, mouse and human, implying that this miRNA is a good candidate to be modulated in Tambaqui with the goal of improving filet quality.

Other aspects where miRNAs stand out pertain to the control of fish disease responses that specifically affect confined shoals, owing to the appearance of various stressors, such as poor water quality, high density, and inadequate diets. The megalocytivirus, from the Iridoviridae family, for example, has been a focus of research because it usually causes significant mortality to a wide range of hosts in aquaculture, leading to pronounced economic losses (Subramaniam et al., 2012). For instance, Nile tilapia

(Subramaniam et al., 2016), Atlantic salmon (Crane and Hyatt, 2011), rainbow trout (Ariel and Jensen, 2009; Crane and Hyatt, 2011) and grass carp (Subramaniam et al., 2012) are confirmed hosts of megalocytivirus. Furthermore, grass carp is a potential carrier of megalocytivirus, but infection does not unavoidably cause mortality or clinical changes in this species (Subramaniam et al., 2012). Thus, studies could examine whether miRNAs differentially expressed would play a role in maintaining grass carp resilience to megalocytivirus infection and symptoms, and subsequently to test for a prospective application to Neotropical fish.

Another key aspect in fish biology pertains to the control of sexual differentiation and maturation. Experiments on Nile tilapia have shown that miR-456 and miR-138 inhibit testis differentiation by targeting the Amh gene (Wang et al., 2016), and the high expression of miR-4585 in males interferes with the sexual reversal mechanism. Together, these miRNAs can be good biomarkers for investigating both the mechanisms influenced by sex and mechanisms that influence sexual development, many of which are important parameters for efficient production. For example, some of these findings could be further investigated to produce monosex populations of species that females grows larger than males, such as tambaqui. Overall, this area is still in its infancy, and additional studies focusing on miRNA pathways related to sexual development in fishes will be of great value for improving fish breeding.

Environmental temperature plays a key role in maintaining the life cycle of any fish species. Understanding the molecular mechanisms involved in acclimatization at different temperatures is fundamental in the current context because of global warming. In addition, these molecular mechanisms can be used to genetically select fish in aquaculture programs. For example, the cultivation of tilapias and black pacu is restricted to tropical and subtropical areas, since they prefer temperatures between 27 and 32°C. Handling and transport at low temperatures (<22°C), especially after winter, leads to severe reduction of appetite and increased risk of disease. Furthermore, the culture of these fish at a temperature below 14°C is generally lethal. As a result, tilapia and black pacu breeding companies cannot expand their activities to low-temperature locations. Therefore, miRNAs have been proposed as molecular markers to select cold-tolerant fish. In zebrafish, 25 differentially expressed miRNAs were identified in individuals cultured for 10 days at a temperature of 10°C, which is related to a cellular response of adaptability to the environment (Yang et al., 2011). Thus, miRNAs could be studied as epigenetic markers for the selection of cold tolerant Neotropical fish, allowing the production of several species in subtropical and temperate regions.

Regarding general metabolism, miRNAs that regulate lipid and glucose metabolism are highly conserved between humans and fishes. For instance, miR-103, miR-107, miR-122, and miR-143 have constrained metabolic functions in vertebrates and can be targeted in future genetic therapies. Similarly, salt tolerance is an interesting trait and Nile tilapia has strong tolerance associated with the expression of miR-30c and miR-429, which can be broadly investigated and modulated in less tolerant Neotropical species to increase their productivity.

One of the major challenges, if not the greatest, for the use of miRNAs to increase the production of farm animals is the control of miRNA expression and consequently the control of genes and their products. Such control would allow researchers and farmers to have absolute control over a range of characteristics that would make the animals more productive or exhibit better quality. Efforts have been made to this end, and in the future, it is likely that miRNAs can be fully manipulated. However, in a genomic context, miRNAs can already be very well-exploited with the purpose of improving animal production.

Biological Implications of miRNA Genomic Context and Regulation of Expression of Complex Traits

Characterization of miRNA signatures of expression for quantitative trait loci (named "miR-eQTL") can produce insights into regulatory mechanisms of miRNA transcription and assist in clarifying the role of miRNAs as orchestrators of complex biological traits.

Experiments detecting the quantitative trait locus (QTL) related to important breeding traits have uncovered diverse molecular markers useful for fish production (Sonesson, 2007; Canario et al., 2008). Recently, the relationship between miRNAs and QTLs has been found to interfere in a myriad of processes, as in sexual differentiation and determination in Nile tilapia (Eshel et al., 2014) and in immune defense against viral infection in Atlantic salmon (Lowe et al., 2014). Additionally, colligated assessment of miRNA signatures and SNP within QTLs has contributed to the recognition of several miRNA regulatory loci of interest. Such miR-eQTLs can be used to determine phenotypic information through the integration of QTL polymorphisms with transcriptome data, including genomic loci of miRNA that contribute to the variation of mRNA expression levels. An analysis in European seabass, another economic relevant species, identified 20,779 SNPs over 1,469 gene loci and intergenic spacers (Kuhl et al., 2011), showing that they can be extensively used as genetic markers in population and molecular studies.

The integration of QTL and miRNA allowed the identification of key regulatory pathways involved in human liver diseases (Gamazon et al., 2013). This approach could be used in fish to assess susceptibility to diseases and other biological aspects, such as muscle development. Indeed, miR-eQTLs are promising molecular tools for enhancing aquaculture productivity and/or species conservation. Similarly, global identification and characterization of miRNAs and QTLs could be an important source for further understanding how coding and non-coding DNA work together to generate attractive phenotypic traits for selection programs (Cerdà and Manchado, 2013).

Although miRNAs have the potential to be used as molecular markers in fish genetic breeding programs, some limitations persist in their broad application in aquaculture. For instance, the assessment of miRNA expression levels is usually performed in specific tissues, requiring euthanasia of the fish, which would make it impossible to use the animals for subsequent selection programs. Another drawback comes from the presence of

miRNAs in biofluids, such as blood and urine (Harrill et al., 2016) that were not directly expressed in these fluids, but rather resulted from leakage owing to cell injury, cell death, or active secretion during manipulation. As an alternative to euthanasia for tissue collection, evaluation of circulating miRNA levels in body fluids can be performed to monitor fish metabolism, allowing animals to be further used for genetic selection. In this way, miRNAs could be evaluated in fish blood, since circulating miRNAs are already being proposed as stable epigenetic biomarkers in several mammals, including humans (Harrill et al., 2016; Kasimanickam and Kastelic, 2016).

Another aspect inherent in miRNA modulation pertains to the characteristics of multigenic regulation, that is, a single miRNA can regulate several genes in different biological contexts. So when miRNA expression is altered artificially, the expression of several target genes can be altered. This is one of the challenges of manipulating miRNAs that makes it difficult to obtain the desired final phenotype. This challenge can be overcome or minimized by controlling the modulation of miRNAs specifically or highly expressed in certain tissues during temporal windows of development. For this approach, it is necessary to validate the real interactions between target miRNAs (**Box 2**) in tissues and various cell types. It will then be possible to reduce or eliminate off-target effects and obtain the desired benefits on the stock. The problem of small size and low abundance of some miRNAs has been circumvented through the development and implementation of last generation sequencing techniques and advanced bioinformatics tools.

In addition, to evaluate whether miRNAs with described functions in one species can change the phenotype of other organisms, genetic reverse techniques for gene knock in or knockout can be performed. This approach works by changing gene expression to rigorously test its function, helping to determine the potentially beneficial aspects of down and up regulation in farmed fish species. Additionally, when these effects are confirmed as molecularly conserved in several species, genome-editing analysis can be broadly performed to modulate expression levels of one or a subset of miRNAs in related and non-related fish species. Cutting-edge genome editing techniques have been recently improved with the CRISPR-Cas system and could positively be implemented for aquaculture enhancement. Several papers have already suggested that CRISPR-Cas studies could be widely applied in fish, as first proven by studies using zebrafish models (Bassett et al., 2014).

Based on the aforementioned discussion, we envision that future research on farmed Neotropical fishes will greatly benefit from the ongoing comparative analysis of transcriptomes, which will provide the roadmap for the development of practical applications to expand animal breeding programs. Certainly, studies on farmed fishes that focus on the identification of interactions between SNPs and miRNAs are another important direction. SNPs in miRNA binding sites or in the miRNA precursor sequences may largely impact the desired phenotype and provide the best use of genomics and bioinformatics in animal breeding and aquaculture. Degradome sequencing technology and miRNA-seq can be applied to populations rather than to a few individuals to assess SNPs associated with specific desired phenotypes. This feature could provide an immediate tool for selecting both larval and adult fishes carrying superior traits of economic interest and improve profits.

Finally, gene expression regulation in its various forms is considered one of the fastest and most effective mechanisms underlying adaptive evolution. Thus, several years ago, gene regulation was recognized as a powerful force for diversification and speciation. As was well-discussed in this paper, miRNAs are key molecules involved in the regulation of genes in all complex organisms, including phenotypic characteristics that confer better production or selective advantages in nature. In addition, Franchini et al. (2016) have shown that several species of Neotropical fishes, given their evolutionary history that forces them to form small populations due to physical environmental changes, are and will still serve as good biological models in studies involving molecular mechanisms associated with speciation caused by changes in gene expression. Consequently, miRNAs play a fundamental role in this process. Thus, these species may be at the forefront of studies involving miRNAs and gene regulation as speciation mechanisms.

CONCLUSIONS AND PERSPECTIVES

In conclusion, this review on fish miRNAs shows that these small molecules are great targets for understanding Neotropical fish biology and that miRNAs possess very attractive features for their immediate implementation as biomarkers that can be used to select adaptive traits for a specific environment in worldwide aquaculture. The existence of highly conserved miRNAs among vertebrates is relevant and may lead to the broad application of knowledge acquired from one fish species to another or even allow for genome editing technology transferability within distinct vertebrate groups. Many miRNA-mediated biological characteristics are available for study and implementation into Neotropical fish farming in distinct captivity environments. Genome editing is the best approach to increase or reduce miRNA expression and could contribute to the improvement of important characteristics that can enhance global aquaculture production.

AUTHOR CONTRIBUTIONS

Literature survey: MH, AO, PN, JC. Data discussion: MH, AO, PN, JC, VC, AH, DP. Writing of the first draft of the manuscript: MH, AO, PN, JC. Critically revision and final manuscript writing: MH, VC, AH, DP. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

This work received financial support from Fapesp (Sao Paulo Research Foundation—grants #2014/03062-0, #2012/15589-7, and #2013/06864-7), and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico—grants #460608/2014-2, #312046/2014-6, #158284/2013-5, #131265/2015-6, and #312312/2016-4).

REFERENCES

Andreassen, R., Woldemariam, N. T., Egeland, I. Ø., Agafonov, O., Sindre, H., and Høyheim, B. (2017). Identification of differentially expressed Atlantic salmon miRNAs responding to salmonid alphavirus (SAV) infection. *BMC Genomics* 18:349. doi: 10.1186/s12864-017-3741-3

Andreassen, R., Worren, M. M., and Høyheim, B. (2013). Discovery and characterization of miRNA genes in Atlantic salmon (*Salmo salar*) by use of a deep sequencing approach. *BMC Genomics* 14:482. doi: 10.1186/1471-2164-14-482

Ariel, E., and Jensen, B. B. (2009). Challenge studies of European stocks of redfin perch, *Perca fluviatilis* L., and rainbow trout, *Oncorhynchus mykiss* (Walbaum), with epizootic haematopoietic necrosis virus. *J. Fish Dis.* 32, 1017–1025. doi: 10.1111/j.1365-2761.2009.01088.x

Ason, B., Darnell, D. K., Wittbrodt, B., Berezikov, E., Kloosterman, W. P., Wittbrodt, J., et al. (2006). Differences in vertebrate microRNA expression. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14385–14389. doi: 10.1073/pnas.0603529103

Barozai, M. Y. (2012). Identification and characterization of the microRNAs and their targets in *Salmo salar*. *Gene* 499, 163–168. doi: 10.1016/j.gene.2012.03.006

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297. doi: 10.1016/S0092-8674(04)00045-5

Bassett, A. R., Azzam, G., Wheatley, L., Tibbit, C., Rajakumar, T., McGowan, S., et al. (2014). Understanding functional miRNA-target interactions *in vivo* by site-specific genome engineering. *Nat. Commun.* 5:4640. doi: 10.1038/ncomms5640

Beckman, B. R., Shimizu, M., Gadberry, B. A., and Cooper, K. A. (2004), Response of the somatotropic axis of juvenile coho salmon to alterations in plane of nutrition with an analysis of the relationships among growth rate and circulating IGF-I and 41 kDa IGFBP. *Gen. Comp. Endocrinol.* 135, 334–344. doi: 10.1016/j.ygcen.2003.10.013

Bekaert, M., Lowe, N. R., Bishop, S. C., Bron, J. E., Taggart, J. B., and Houston, R. D. (2013). Sequencing and characterisation of an extensive Atlantic salmon (*Salmo salar* L.) microRNA repertoire. *PLoS ONE* 8:e70136. doi: 10.1371/journal.pone.0070136

Bentsen, H. B., and Thodesen, J. (2005). "Genetic interaction between farmed and wildfish, with examples from the Atlantic salmon case in Norway," in *Selection and Breeding Programs in Aquaculture*, ed T. Gjedrem (Dordrecht: Springer), 319–334.

Berezikov, E. (2011). Evolution of microRNA diversity and regulation in animals. *Nat. Rev. Genet.* 12, 846–860. doi: 10.1038/nrg3079

Bertollo, L. A. C., Cioffi, M. B., Galetti, P. M. Jr., and Filho, O. M. (2017). Contributions to the cytogenetics of the Neotropical fish fauna. *Comp. Cytogenet.* 11, 665–690. doi: 10.3897/compcytogen.v11i4.14713

Betel, D., Koppal, A., Agius, P., Sander, C., and Leslie, C. (2010). Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol.* 11:R90. doi: 10.1186/gb-2010-11-8-r90

Brawand, D., Wagner, C. E., Li, Y. I., Malinsky, M., Keller, I., Fan, S., et al. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature*. 513, 375–381. doi: 10.1038/nature13726

Brenneck, J., Stark, A., Russell, R. B., and Cohen, S. M. (2005). Principles of microRNA-target recognition. *PLoS Biol.* 3:e85. doi: 10.1371/journal.pbio.0030085

Canario, A. V. M., Bargelloni, L., Volckaert, F., Houston, R. D., Massault, C., and Guiguen, Y. (2008). Genomics toolbox for farmed fish. *Rev. Fish. Sci.* 16, 3–15. doi: 10.1080/10641260802319479

Carnevali, O., de Vivo, L., Sulpizio, R., Gioacchini, G., Olivetto, I., Silvi, S., et al. (2006). Growth improvement by probiotic in European sea bass juveniles (*Dicentrarchus labrax*, L.), with particular attention to IGF-1, myostatin and cortisol gene expression. *Aquaculture* 258, 430–438. doi: 10.1016/j.aquaculture.2006.04.025

Castagnolli, N., and Cyrino, J. (1986). *Piscicultura Nos Trópicos*, *1st Edn*. Barueri: Manole.

Cerdà, J., and Manchado, M. (2013). Advances in genomics for flatfish aquaculture. *Genes Nutr.* 8, 5–17. doi: 10.1007/s12263-012-0312-8

Chen, C., Deng, B., Qiao, M., Zheng, R., Chai, J., Ding, Y., et al. (2012). Solexa sequencing identification of conserved and novel microRNAs in backfat of large white and Chinese meishan pigs. *PLoS ONE* 7:e31426. doi: 10.1371/journal.pone.0031426

Chen, J. F., Mandel, E. M., Thomson, J. M., Wu, Q., Callis, T. E., Hammond, S. M., et at. (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat. Genet.* 38, 228–233. doi: 10.1038/ng1725

Chen, X. M., Splinter, P. L., O'Hara, S. P., and LaRusso, N. F. (2007). A cellular micro-RNA, let-7i, regulates Toll-like receptor 4 expression and contributes to cholangiocyte immune responses against *Cryptosporidium parvum* infection. *J. Biol. Chem.* 282, 28929. doi: 10.1074/jbc.m702633200

Cho, I. S., Kim, J., Seo, H. Y., Lim, D. H., Hong, J. S., Park, Y. H., et al. (2010). Cloning andcharacterization of microRNAs from porcine skeletal muscle and adipose tissue. *Mol. Biol. Rep.* 37, 3567–3574. doi: 10.1007/s11033-010-0005-6

Cirera, S., Birck, M., Busk, P. K., and Fredholm, M. (2010). Expression profiles of miRNA-122 and its target CAT1 in minipigs (*Sus scrofa*) fed a high-cholesterol diet. *Comp. Med.* 60, 136–141.

Crane, M., and Hyatt, A. (2011). Viruses of fish: an overview of significant pathogens. *Viruses* 3, 2025–2046. doi: 10.3390/v3112025

Dávalos, A., Goedeke, L., Smibert, P., Ramírez, C. M., Warrier, N. P., Andreo, U., et al. (2011). miR-33a/b contribute to the regulation of fatty acid metabolism and insulin signaling. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9232–9237. doi: 10.1073/pnas.1102281108

Deane, E. E., and Woo, N. Y. S. (2004). Differential gene expression associated with euryhalinity in sea bream (*Sparus sarba*). *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 287, R1054–R1063. doi: 10.1152/ajpregu.00347.2004

Devlin, A. H., Thompson, P., Robson, T., and McKeown, S. R. (2010). Cytochrome P450 1B1 mRNA untranslated regions interact to inhibit protein translation. *Mol. Carcinog.* 49, 190–199.doi: 10.1002/mc.20589

Dobrzyn, A., and Ntambi, J. M. (2004). The role of stearoyl-CoA desaturase in body weight regulation. *Trends Cardiovasc. Med.* 14, 77–81. doi: 10.1016/j.tcm.2003.12.005

Doench, J. G., and Sharp, P. A. (2004). Specificity of microRNA target selection in translational repression. *Genes Dev.* 18, 504–511. doi: 10.1101/gad.1184404

Duran, B. O. S., Fernandez, G. J., Mareco, E. A., Moraes, L. N., Salomão, R. A. S., Gutierrez de Paula, T., et al. (2015). Differential microRNA expression in fast- and slow-twitch skeletal muscle of *Piaractus mesopotamicus* during growth. *PLoS ONE* 10:e0141967. doi: 10.1371/journal.pone.0141967

Dweep, H., Sticht, C., and Gretz, N. (2013). *In silico* algorithms for the screening of possible microRNA binding sites and their interactions. *Curr. Genomics* 14, 127–136. doi: 10.2174/1389202911314020005

Dweep, H., Sticht, C., Pandey, P., and Gretz, N. (2011). miRWalk-database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. *J. Biomed. Inform.* 44, 839–847. doi: 10.1016/j.jbi.2011.05.002

Elmén, J., Lindow, M., Silahtaroglu, A., Bak, M., Christensen, M., Lind-Thomsen, A., et al. (2008). Antagonism of microRNA-122 in mice by systemically administered LNAantimiR leads to up-regulation of a large set of predicted target mRNAs in the liver. *Nucleic Acids Res.* 36, 1153–1162. doi: 10.1093/nar/gkm1113

Enright, A. J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D. S. (2003). MicroRNA targets in drosophila. *Genome Biol.* 5:R1. doi: 10.1186/gb-2003-5-1-r1

Esau, C., Davis, S., Murray, S. F., Yu, X. X., Pandey, S. K., Pear, M., et al. (2006). miR-122 regulation of lipid metabolism revealed by *in vivo* antisense targeting. *Cell Metab.* 3, 87–98. doi: 10.1016/j.cmet.2006.01.005

Eshel, O., Shirak, A., Dor, L., Band, M., Zak, T., Markovich-Gordon, M., et al. (2014). Identification of male-specific amhduplication, sexually differentially expressed genes and microRNAs at early embryonic development of Nile tilapia (*Oreochromis niloticus*). *BMC Genomics* 15:774. doi: 10.1186/1471-2164-15-774

Fernández-Hernando, C., Suárez, Y., Rayner, K. J., and Moore, K. J. (2011). MicroRNAs in lipid metabolism. *Curr. Opin. Lipidol.* 22, 86–92. doi: 10.1097/MOL.0b013e3283428d9d

Fishbase (2015). *Ctenopharyngodon idella (Valenciennes, 1844) Grass carp*. Available online at: http://www.fishbase.org/summary/79 (Accessed August 8, 2015).

Franchini, P., Xiong, P., Fruciano, C., and Meyer, A. (2016). The role of microRNAs in the repeated parallel diversification of lineages of midas cichlid fish from Nicaragua. *Genome Biol. Evol.* 8, 1543–1555. doi: 10.1093/gbe/evw097

Gamazon, E. R., Innocenti, F., Wei, R., Wang, L., Zhang, M., Mirkov, S., et al. (2013). A genome-wide integrative study of microRNAs in human liver. *BMC Genomics* 14:395. doi: 10.1186/1471-2164-14-395

Garcia, D. M., Baek, D., Shin, C., Bell, G. W., Grimson, A., and Bartel, D. P. (2011). Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nat. Struct. Mol. Biol.* 18, 1139–1146. doi: 10.1038/nsmb.2115

Giusti, J., Pinhal, D., Moxon, S., Campos, C. L., Münsterberg, A., and Martins, C. (2016). MicroRNA-10 modulates Hox genes expression during Nile tilapia embryonic development. *Mech. Dev.* 140, 12–18. doi: 10.1016/j.mod.2016.03.002

Gjedrem, T. (2010). The first family-based breeding program in aquaculture. *Rev. Aquacult.* 2, 2–15. doi: 10.1111/j.1753-5131.2010.01011.x

Gomes, F., Watanabe, L., Nozawa, S., Oliveira, L., Cardoso, J., Vianez, J., et al. (2017). Identification and characterization of the expression profile of the microRNAs in the Amazon species Colossoma macropomum by next generation sequencing. *Genomics* 109, 67–74. doi: 10.1016/j.ygeno.2017.02.001

Goulding, M., and Carvalho, M. L. (1982). Live history and management of the tambaqui, (*Colossoma macropomum*, characidae): an important Amazonian food fish. *Rev. Bras. Zootec.* 1, 107–133. doi: 10.1590/S0101-81751982000200001

Guerra-Assumpção, J. A., and Enright, A. J. (2012). Large-scale analysis of microRNA evolution. *BMC Genomics* 13:218. doi: 10.1186/1471-2164-13-218

Ha, M., and Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* 15, 509–524. doi: 10.1038/nrm3838

Harrill, A. H., McCullough, S. D., Wood, C. E., Kahle, J. J., and Chorley, B. N. (2016). MicroRNA biomarkers of toxicity in biological matrices. *Toxicol. Sci.* 152, 264–272. doi: 10.1093/toxsci/kfw090

Hausser, J., Syed, A. P., Bilen, B., and Zavolan, M. (2013). Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. *Genome Res.* 23, 604–615. doi: 10.1101/gr.139758.112

Hicks, J. A., Tembhurne, P., and Liu, H. C. (2008). MicroRNA expression in chicken embryos. *Poult. Sci.* 87, 2335–2343. doi: 10.3382/ps.2008-00114

Hicks, J. A., Tembhurne, P. A., and Liu, H. C. (2009). Identification of microRNA in the developing chick immune organs. *Immunogenetics* 61, 231–240. doi: 10.1007/s00251-009-0355-1

Huang, C. W., Li, Y. H., Hu, S. Y., Chi, J. R., Lin, G. H., Lin, C. C., et al. (2012). Differential expression patterns of growth-related microRNAs in the skeletal muscle of Nile tilapia (*Oreochromis niloticus*). *J. Anim. Sci.* 90, 4266–4279. doi: 10.2527/jas.2012-5142

Hutvágner, G., McLachlan, J., Pasquinelli, A. E., Bálint, E., Tuschl, T., and Zamore, P. D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834–838. doi: 10.1126/science.1062961

IBAMA (2000). "Informe estatístico do desembarque pesqueiro na cidade de Santarém, PA: 1992–1993," in *Recursos Pesqueiros do MédioAmazonas: Biologia e Estatística Pesqueira*, eds V. J. Isaac, M. L. Ruffino, and C. F, Fischer (Brasília: IBAMA/GTZ/GOPA), 225–280.

Johansen, I., and Andreassen, R. (2014). Validation of miRNA genes suitable as reference genes in qPCR analyses of miRNA gene expression in Atlantic salmon (*Salmo salar*). *BMC Res. Notes* 8:945. doi: 10.1186/1756-0500-7-945

Jones, J. I., and Clemmons, D. R. (1995). Insulin-like growth factors and their binding proteins: biological actions. *Endocr. Rev.* 16, 3–34.

Juanchich, A., Bardou, P., Rué, O., Gabillard, J. C., Gaspin, C., Bobe, J., et al. (2016). Characterization of an extensive rainbow trout miRNA transcriptome by next generation sequencing. *BMC Genomics* 17:164. doi: 10.1186/s12864-016-2505-9

Kasimanickam, V., and Kastelic, J. (2016). Circulating cell-free mature microRNAs and their target gene prediction in bovine metritis. *Sci. Rep.* 6:29509. doi: 10.1038/srep29509

Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nat. Genet.* 39, 1278–1284. doi: 10.1038/ng2135

Khvorova, A., Reynolds, A., and Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115, 209–216. doi: 10.1016/S0092-8674(03)00801-8

Kim, H. K., Lee, Y. S., Sivaprasad, U., Malhotra, A., and Dutta, A. (2006). Muscle-specific microRNA miR-206 promotes muscle differentiation. *J. Cell Biol.* 174, 677–687. doi: 10.1083/jcb.200603008

Kim, V. N. (2005). MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol.* 6, 376–385. doi: 10.1038/nrm1644

Kim, V. N., Han, J., and Siomi, M. C. (2009). Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* 10, 126–139. doi: 10.1038/nrm2632

Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* 39, D152–D157. doi: 10.1093/nar/gkq1027

Krützfeldt, J., Rajewsky, N., Braich, R., Rajeev, K. G., Tuschl, T., Manoharan, M., et al. (2005). Silencing of microRNAs *in vivo* with "antagomirs". *Nature* 438, 685–689. doi: 10.1038/nature04303

Kubiczkova, L., Sedlarikova, L., Hajek, R., and Sevcikova, S. (2012). Tgf-beta - an excelente servant but a bad master. *J. Transl. Med.* 10:183. doi: 10.1186/1479-5876-10-183

Kuhl, H., Tine, M., Hecht, J., Knaust, F., and Reinhardt, R. (2011). Analysis of single nucleotide polymorphisms in three chromosomes of European sea bass *Dicentrarchus labrax*. *Comp. Biochem. Physiol. D* 6, 70–75. doi: 10.1016/j.cbd.2010.04.003

Kure, E. H., Sæbø, M., Stangenland, A. M., Hamfjord, J., Hytterød, S., Heggenes, J., et al. (2013). Molecular responses to toxicological stressors: profiling microRNAs in wild Atlantic salmon (*Salmo salar*) exposed to acidic aluminum-rich water. *Aquat. Toxicol.* 138, 98–104. doi: 10.1016/j.aquatox.2013.04.004

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science* 294, 853–858. doi: 10.1126/science.1064921

Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. (2002). Identification of tissue-specific microRNAs from mouse. *Curr. Biol.* 12, 735–739. doi: 10.1016/S0960-9822(02)00809-6

Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858–862. doi: 10.1126/science.1065062

Lee, R. C., and Ambros, V. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294, 862–864. doi: 10.1126/science.1065329

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75, 843–854. doi: 10.1016/0092-8674(93)90529-Y

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., et al. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415–419. doi: 10.1038/nature01957

Lee, Y. S., and Dutta, A. (2009). MicroRNAs in cancer. *Annu. Rev. Pathol.* 4, 199–227. doi: 10.1146/annurev.pathol.4.110807.092222

Li, G., Li, Y., Li, X., Ning, X., Li, M., and Yang, G. (2011). MicroRNA identity and abundance in developing swine adipose tissue as determined by Solexa sequencing. *J. Cell. Biochem.* 112, 1318–1328. doi: 10.1002/jcb.23045

Li, H., Zhang, Z., Zhou, X., Wang, Z., Wang, G., and Han, Z. (2011). Effects of microRNA-143 in the differentiation and proliferation of bovine intramuscular preadipocytes. *Mol. Biol. Rep.* 38, 4273–4280. doi: 10.1007/s11033-010-0550-z

Li, T., Wu, R., Zhang, Y., and Zhu, D. (2011). A systematic analysis of the skeletal muscle miRNA transcriptome of chicken varieties with divergent skeletal muscle growth identifies novel miRNAs and differentially expressed miRNAs. *BMC Genomics* 12:186. doi: 10.1186/1471-2164-12-186

Li, Z., Hassan, M. Q., Jafferji, M., Aqeilan, R. I., Garzon, R., Croce, C. M., et al. (2009). Biological functions of miR-29b contribute to positive regulation of osteoblast differentiation. *J. Biol. Chem.* 284, 15676–15684. doi: 10.1074/jbc.M809787200

Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., et al. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 433, 769–773. doi: 10.1038/nature03315

Liu, J., Luo, X. J., Xiong, A. W., Zhang, Z. D., Yue, S., Zhu, M. S., et al. (2009). MicroRNA-214 promotes myogenic differentiation by facilitating exit from mitosis via downregulation of proto-oncogene N-ras. *J. Biol. Chem.* 285, 26599–26607. doi: 10.1074/jbc.M110.115824

Loh, Y. H. E., Yi, S. V., and Streelman, J. T. (2011). Evolution of MicroRNAs and the diversification of species. *Genome Biol. Evol.* 3, 55–65. doi: 10.1093/gbe/evq085

Lowe, N. R., Bekaert, M., Bishop, S. C., Bron, J. E., Taggart, J. B., and Houston, R. D. (2014). "MicroRNA profiling of Atlantic salmon challenged with Infectious Pancreatic Necrosis virus: comparison between resistant and susceptible fish,"

in *Proceedings, 10th World Congress of Genetics Applied to Livestock Production* (Vancouver, BC).

Lund, E., Güttinger, S., Calado, A., Dahlberg, J. E., and Kutay, U. (2004). Nuclear export of microRNA precursors. *Science* 303, 95–98. doi: 10.1126/science.1090599

Maragkakis, M., Reczko, M., Simossis, V. A., Alexiou, P., Papadopoulos, G. L., and Dalamagas, T. (2009). DIANA-microT webserver: elucidating microRNA functions through target prediction. *Nucleic Acids Res.* 37, 273–276. doi: 10.1093/nar/gkp292

Mauel, M. J., and Miller, D. L. (2002). Piscirickettsiosis and piscirickettsiosis-like infections in fish: a review. *Vet. Microbiol.* 87, 279–289. doi: 10.1016/S0378-1135(02)00085-8

Maurer, B., Stanczyk, J., Jüngel, A., Akhmetshina, A., Trenkmann, M., Brock, M., et al. (2010). microRNA-29, a key regulator of collagen expression in systemic sclerosis. *Arthritis Rheum*. 62, 1733–1743. doi: 10.1002/art.27443

Mennigen, J. A., Panserat, S., Larquier, M., Plagnes-Juan, E., Medale, F., Seiliez, I., et al. (2012). Postprandial regulation of hepatic microRNAs predicted to target the insulin pathway in rainbow trout. *PLoS ONE* 7:e38604. doi: 10.1371/journal.pone.0038604

Mennigen, J. A., Skiba-Cassy, S., and Panserat, S. (2013). Ontogenetic expression of metabolic genes and microRNAs in rainbow trout alevins during the transition from the endogenous to the exogenous feeding period. *J. Exp. Biol.* 16, 1597–1608. doi: 10.1242/jeb.082248

Mennigen, J. A., and Zhang, D. (2016). MicroTrout: a comprehensive, genome-wide miRNA target prediction framework for rainbow trout, *Oncorhynchus mykiss*. *Comp. Biochem. Physiol. D*. 20, 19–26. doi: 10.1016/j.cbd.2016.07.002

Miranda, K. C., Huynh, T., Tay, Y., Ang, Y. S., Tam, W. L., Thomson, A. M., et al. (2006). A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* 126, 1203–1217. doi: 10.1016/j.cell.2006.07.031

Miretti, S., Martignani, E., Taulli, R., Bersani, F., Accornero, P., and Baratta, M. (2011). Differential expression of microRNA-206 in skeletal muscle of female Piedmontese and Friesian cattle. *Vet. J.* 190, 412–413. doi: 10.1016/j.tvjl.2010.12.012

Mitchie, I. (2001). "Causes of downgrading in the salmon farming industry," in *Farmed Fish Quality*, eds S. C. Kestin and P. D. Warris (Oxford: Fishing New Books), 129–136.

Moreno, H. M., Jacq, C., Montero, M. P., Gómez-Guillén, M. C., Borderías, A. J., and Mørkøre, T. (2016). Effect of selective breeding on collagen properties of Atlantic salmon (*Salmo salar* L.). *Food Chem.* 190, 856–863. doi: 10.1016/j.foodchem.2015.06.022

Nachtigall, P. G., Dias, M. C., and Pinhal, D. (2014). Evolution and genomic organization of muscle microRNAs in fish genomes. *BMC Evol. Biol.* 14, 196. doi: 10.1186/s12862-014-0196-x

Nachtigall, P. G., Dias, M. C., Carvalho, R. F., Martins, C., and Pinhal, D. (2015). MicroRNA-499 expression distinctively correlates to target genes sox6 and rod1 profiles to resolve the skeletal muscle phenotype in Nile tilapia. *PLoS ONE* 10:e0119804. doi: 10.1371/journal.pone.0119804

Ntambi, J. M., and Miyazaki, M. (2004). Regulation of stearoyl-CoA desaturases and role in metabolism. *Prog. Lipid Res.* 43, 91–104. doi: 10.1016/S0163-7827(03)00039-0

Oliveira, A. C., Bovolenta, L. A., Nachtigall, P. G., Herkenhoff, M. E., Lemke, N., and Pinhal, D. (2017). Combining results from distinct MicroRNA target prediction tools enhances the performance of analyses. *Front. Genet.* 8:59. doi: 10.3389/fgene.2017.00059

Paneru, B. D., Al-Tobasei, R., Kenney, B., Leeds, T. D., and Salem, M. (2017). RNA-Seq reveals MicroRNA expression signature and genetic polymorphism associated with growth and muscle quality traits in rainbow trout. *Sci. Rep.* 7:9078. doi: 10.1038/s41598-017-09515-4

Paula, T. G., Zanella, B. T. T., Fantinatti, B. E. A., de Moraes, L. N., Duran, B. O. S., de Oliveira, C. B., et al. (2017). Food restriction increase the expression of mTORC1 complex genes in the skeletal muscle of juvenile pacu (*Piaractus mesopotamicus*). ed *PLoS ONE* 12:e0177679. doi: 10.1371/journal.pone.0177679

Peterson, S. M., Thompson, J. A., Ufkin, M. L., Sathyanarayana, P., Liaw, L., and Congdon, C. B. (2014). Common features of microRNA target prediction tools. *Front. Genet.* 5:23. doi: 10.3389/fgene.2014.00023

Qiang, J., Tao, Y. F., He, J., Li, H. X., Xu, P., Bao, J. W., et al. (2017a). Inhibition of miR-92d-3p enhances inflammation responses in genetically improved farmed tilapia (GIFT, *Oreochromis niloticus*) with Streptococcus iniae infection by modulating complement C3. *Fish Shellfish Immunol.* 63, 367–375. doi: 10.1016/j.fsi.2017.02.036

Qiang, J., Tao, Y. F., He, J., Sun, L., Xu, P., and Bao, W. (2017b). Effects of exposure to Streptococcus iniae on microRNA expression in the head kidney of genetically improved farmed tilapia (*Oreochromis niloticus*). *BMC Genomics* 18, 190. doi: 10.1186/s12864-017-3591-z

Qiang, J., Tao, Y. F., He, J., Sun, Y. L., and Xu, P. (2017c). miR-29a modulates SCD expression and is regulated in response to a saturated fatty acid diet in juvenile genetically improved farmed tilapia (*Oreochromis niloticus*). *J. Exp. Biol.* 15, 1481–1489. doi: 10.1242/jeb.151506

Qiang, J., Tao, Y. F., He, J., Xu, P., Bao, J. W., and Sun, Y. L. (2017d). miR-122 promotes hepatic antioxidant defense of genetically improved farmed tilapia (GIFT, *Oreochromis niloticus*) exposed to cadmium by directly targeting a metallothionein gene. *Aquat. Toxicol.* 82, 39–48. doi: 10.1016/j.aquatox.2016.11.009

Ramachandra, R. K., Salem, M., Gahr, S., Rexroad, C. E. III., and Yao, J. (2008). Cloning and characterization of microRNAs from rainbow trout (*Oncorhynchus mykiss*): their expression during early embryonic development. *BMC Dev. Biol.* 8:41. doi: 10.1186/1471-213X-8-41

Rasal, K. D., Nandanpawar, P. C., Swain, P., Badhe, M. R., Sundaray, J. K., and Jayasankar, P. (2016). MicroRNA in aquaculture fishes: a way forward with high-throughput sequencing and a computational approach. *Rev. Fish Biol. Fish.* 26, 199–212. doi: 10.1007/s11160-016-9421-6

Reczko, M., Maragkakis, M., Alexiou, P., Grosse, I., and Hatzigeorgiou, A. G. (2012). Functional microRNA targets in protein coding sequences. *Bioinformatics* 28, 771–776. doi: 10.1093/bioinformatics/bts043

Reinecke, M., and Collet, C. (1998). The phylogeny of the insulin-like growth factors. *Int. Rev. Cytol.* 183, 1–94. doi: 10.1016/S0074-7696(08)60142-4

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., et al. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901–906. doi: 10.1038/35002607

Reis, R. E., Kullander, S. O., and Ferraris, C. J. Jr. (2003). *Check List of Freshwater Fishes of South and Central America* (Porto Alegre: EDIPUCRS), 742.

Reyes-Herrera, P. H., and Ficarra, E. (2012). One decade of development and evolution of microRNA target prediction algorithms. *Genomics Proteomics Bioinformatics* 10, 254–263. doi: 10.1016/j.gpb.2012.10.001

Ritchie, W., and Rasko, J. E. (2014). Refining microRNA target predictions: sorting the wheat from the chaff. *Biochem. Biophys. Res. Commun.* 445, 780–784. doi: 10.1016/j.bbrc.2014.01.181

Rodríguez-Marí, A., Ca-estro, C., Bremiller, R. A., Nguyen-Johnson, A., Asakawa, K., Kawakami, K., et al. (2010). Sex reversal in zebrafish fancl mutants is caused by Tp53-mediated germ cell apoptosis. *PLoS Genet.* 6:e1001034. doi: 10.1371/journal.pgen.1001034

Sae-Lim, P., Kause, A., Mulder, H. A., Martin, K. E., Barfoot, A. J., Parsons, J. E., et al. (2013). Genotype-by-environment interaction of growth traits in rainbow trout (*Oncorhynchus mykiss*): a continental scale study. *J. Anim. Sci.* 91, 5572–5581. doi: 10.2527/jas.2012-5949

Salem, M., Xiao, C., Womack, J., Rexroad, C. E. III., and Yao, J. (2010). A microRNA repertoire for functional genome research in rainbow trout (*Oncorhynchus mykiss*). *Mar. Biotechnol.* 12, 410–429. doi: 10.1007/s10126-009-9232-z

Senthilkumaran, B., Yoshikuni,. M., and Nagahama, Y. (2004). A shift in steroidogenesis occurring in ovarian follicles prior to oocyte maturation. *Mol. Cell. Endocrinol.* 215, 11–18. doi: 10.1016/j.mce.2003.11.012

Shingara, J., Keiger, K., Shelton, J., Laosinchai-Wolf, W., Powers, P., Conrad, R., et al. (2005). An optimized isolation and labeling platform for accurate microRNA expression profiling. *RNA* 11, 1461–1470. doi: 10.1261/rna.2610405

Sirotkin, A. V., Ovcharenko, D., Grossmann, R., Lauková, M., and Mlyncek, M. (2009). Identification of microRNAs controlling human ovarian cell steroidogenesis via a genome-scale screen. *J. Cell. Physiol.* 219, 415–420. doi: 10.1002/jcp.21689

Skaftnesmo, K. O., Edvardsen, R. B., Furmanek, T., Crespo, D., Andersson, E., Kleppe, L., et al. (2017). Integrative testis transcriptome analysis reveals differentially expressed miRNAs and their mRNA targets during early puberty in Atlantic salmon. *BMC Genomics* 18, 801. doi: 10.1186/s12864-017-4205-5

Sonesson, A. K. (2007). Within-family marker-assisted selection for aquaculture species. *Genet. Select. Evol.* 39, 301–317. doi: 10.1186/1297-9686-39-3-301

Subramaniam, K., Gotesman, M., Smith, C. E., Steckler, N. K., Kelley, K. L., Groff, J. M., et al. (2016). Megalocytivirus infection in cultured Nile tilapia *Oreochromis niloticus*. *Dis. Aquat. Org.* 119, 253–258. doi: 10.3354/dao02985

Subramaniam, K., Shariff, M., Omar, A. R., and Hair-Bejo, M. (2012). Megalocytivirus infection in fish. *Rev. Aquac.* 4, 221–233. doi: 10.1111/j.1753-5131.2012.01075.x

Thorgaard, G. H., Bailey, G. S., Williams, D., Buhler, D. R., Kaattari, S. L., Ristow, S. S., et al. (2002). Status and opportunities for genomics research with rainbow trout. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 133, 609–646. doi: 10.1016/S1096-4959(02)00167-7

Townley-Tilson, W., Callis, T. E., and Wang, D. Z. (2010). MicroRNAs 1, 133, and 206: critical factors of skeletal and cardiac muscle development, function, and disease. *Int. J. Biochem. Cell Biol.* 42, 1252–1255. doi: 10.1016/j.biocel.2009.03.002

Urbinati, E., and Gonçalves, D. (2005). *Pacu (Piaractus mesopotamicus)*. UFSM, Santa Maria: Espécies Nativas Para Piscicultura no Brasil.

Valenzuela-Miranda, D., Valenzuela-Muñoz, V., Farlora, R., and Gallardo-Escárate, C. (2017). MicroRNA-based transcriptomic responses of Atlantic salmon during infection by the intracellular bacterium *Piscirickettsia salmonis*. *Dev. Comp. Immunol.* 77, 287–296. doi: 10.1016/j.dci.2017.08.016

Valenzuela-Muñoz, V., Novoa, B., Figueras, A., and Gallardo-Escárate, C. (2017). Modulation of Atlantic salmon miRNome response to sea louse infestation. *Dev. Comp. Immunol.* 6, 380–391. doi: 10.1016/j.dci.2017.07.009

Wang, B., Gan, Z., Wang, Z., Yu, D., Lin, Z., Lu, Y., et al. (2017) Integrated analysis neurimmiRs of tilapia (*Oreochromis niloticus*) involved in immune response to Streptococcus agalactiae, a pathogen causing meningoencephalitis in teleosts. *Fish Shellfish Immunol.* 61:44e60. doi: 10.1016/j.fsi.2016.12.009

Wang, D., Qin, J., Jia, J., Yan, P., and Li, W. (2017). Pou1f1, the key transcription factor related to somatic growth in tilapia (*Orechromis niloticus*), is regulated by two independent posttranscriptional regulation mechanisms. *Biochem. Biophys. Res. Commun.* 483, 559–565. doi: 10.1016/j.bbrc.2016.12.106

Wang, P. J., Lin, C. H., Hwang, L. Y., Huang, C. L., Lee, T. H., and Hwang, P. P. (2009). Differential responses in gills of euryhaline tilapia Oreochromis mossambicus, to various hyperosmotic shocks. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 152, 544–551. doi: 10.1016/j.cbpa.2008.12.012

Wang, W., Liu, W., Liu, Q., Li, B., An, L., Hao, R., et al. (2016). Coordinated microRNA and messenger RNA expression profiles for understanding sexual dimorphism of gonads and the potential roles of microRNA in the steroidogenesis pathway in Nile tilapia (*Oreochromis niloticus*). *Theriogenology* 85, 970–978. doi: 10.1016/j.theriogenology.2015.11.006

Wang, X. G., Yu, J. F., Zhang, Y., Gong, D. Q., and Gu, Z. L. (2012b). Identification a characterization of microRNA from chicken adipose tissue and skeletal muscle. *Poult. Sci.* 91, 139–149. doi: 10.3382/ps.2011-01656

Wang, X., Gu, Z., and Jiang, H. (2013). MicroRNAs in farm animals. *Animal* 7, 1567–1575. doi: 10.1017/S1751731113001183

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* 75, 855–862. doi: 10.1016/0092-8674(93)90530-4

Witkos, T., Koscianska, E., and Krzyzosiak, W. J. (2011). Practical aspects of microRNA target prediction. *Curr. Mol. Med.* 11, 93–109. doi: 10.2174/156652411794859250

Xia, J. H., He, X. P., Bai, Z. Y., and Yue, G. H. (2011). Identification and Characterization of 63 MicroRNAs in the Asian seabass Lates calcarifer. *PLoS ONE* 6:e17537. doi: 10.1371/journal.pone.0017537

Xiao, J., Zhong, H., Zhou, Y., Yu, F., Gao, Y., Luo, Y., et al. (2014). Identification and characterization of microRNAs in ovary and testis of Nile tilapia (*Oreochromis niloticus*) by using Solexa sequencing technology. *PLoS ONE* 9:e86821. doi: 10.1371/journal.pone.0086821

Xie, S. S., Huang, T. H., Shen, Y., Li, X. Y., Zhang, X. X., Zhu, M. J., et al. (2010). Identification and characterization of microRNAs from porcine skeletal muscle. *Anim. Genet.* 41, 179–190. doi: 10.1111/j.1365-2052.2009.01991.x

Xu, X. Y., Shen, Y. B., Fu, J. J., Lu, L. Q., and Li, J. L. (2015). Next-generation sequencing identified microRNAs that associate with motile aeromonad septicemia in grass carp. *Fish Shellfish Immunol.* 45, 94–103. doi: 10.1016/j.fsi.2015.02.008

Xu, X. Y., Shen, Y. B., Fu, J. J., Yu, H. Y., Huang, W. J., Lu, L. Q., et al. (2016). MicroRNA-induced negative regulation of TLR-5 in grass carp, *Ctenopharyngodon idella*. *Sci. Rep.* 6:18595. doi: 10.1038/srep18595

Xu, X., Shen, Y., Fu, J., Lu, L., and Li, J. (2014). De novo assembly of the grass carp Ctenopharyngodon idella transcriptome to identify miRNA targets associated with motile aeromonad septicemia. *PLoS ONE* 9:e112722. doi: 10.1371/journal.pone.0112722

Yan, B., Guo, J. T., Zhao, L. H., and Zhao, J. L. (2012a). microRNA expression signature in skeletal muscle of Nile tilapia. *Aquaculture* 364, 240–246. doi: 10.1016/j.aquaculture.2012.08.018

Yan, B., Guo, J. T., Zhao, L. H., and Zhao, J. L. (2012b). MiR-30c: a novel regulator of salt tolerance in tilapia. *Biochem. Biophys. Res. Commun.* 425, 315–320. doi: 10.1016/j.bbrc.2012.07.088

Yan, B., Guo, J. T., Zhu, C. D., Zhao, L. H., and Zhao, J. L. (2013a). miR-203b: a novel regulator of MyoD expression in tilapia skeletal muscle. *J. Exp. Biol.* 216, 447–451. doi: 10.1242/jeb.076315

Yan, B., Zhao, L. H., Guo, J. T., and Zhao, L. J. (2012c). miR-429 regulation of osmotic stress transcription factor 1 (OSTF1) in tilapia during osmotic stress. *Biochem. Biophys. Res. Commun.* 426, 294–298. doi: 10.1016/j.bbrc.2012.08.029

Yan, B., Zhu, C. D., Guo, J. T., Zhao, L. H., and Zhao, J. L. (2013b). miR-206 regulates the growth of the teleost tilapia (*Oreochromis niloticus*) through the modulation of IGF-1 gene expression. *J. Exp. Biol.* 216, 1265–1269. doi: 10.1242/jeb.079590

Yan, B.,Wang, Z., Zhu, C. D., Guo, J. T., and Zhao, J. L. (2014). MicroRNA repertoire for functional genome research in tilapia identified by deep sequencing. *Mol. Biol. Rep.* 41, 4953–4963. doi: 10.1007/s11033-014-3361-9

Yang, N., Kaur, S., Volinia, S., Greshock, J., Lassus, H., Hasegawa, K., et al. (2008). MicroRNA microarray identifies *Let-7i* as a novel biomarker and therapeutic target in human epithelial ovarian cancer. *Cancer Res.* 68, 10307–10314. doi: 10.1158/0008-5472.CAN-08-1954

Yang, R., Dai, Z., Chen, S., and Chen, L. (2011). MicroRNA-mediated gene regulation plays aminor role in the transcriptomic plasticity of cold-acclimated Zebrafish brain tissue. *BMC Genomics* 12:605. doi: 10.1186/1471-2164-12-605

Yao, J., Wang, Y., Wang, W., Wang, N., and Li, H. (2011). Solexa sequencing analysis of chicken pre-adipocyte microRNAs. *Biosci. Biotechnol. Biochem.* 75, 54–61. doi: 10.1271/bbb.100530

Yoon, S. I., Kurnasov, O., Natarajan, V., Hong, M., Gudkov, A. V., Osterman, A. L., et al. (2012). Structural basis of TLR5-flagellin recognition and signaling. *Science* 335, 859–864. doi: 10.1126/science.1215584

Yue, G. H., Lin, H. R., and Li, J. L. (2016). Tilapia is the fish for next - generation aquaculture. *Int. J. Mar. Sci. Ocean Technol.* 3, 11–13.

Zavala,. E., Reyes, D., Deerenberg, R., and Vidal, R. (2017). Selection of reference genes for microRNA analysis associated to early stress response to handling and confinement in *Salmo salar*. *Sci. Rep.* 7:1756. doi: 10.1038/s41598-017-01970-3

Zhao, Y., Wu, J. W., Wang, Y., and Zhao, J. L. (2016). Role of miR-21 in alkalinity stress tolerance in tilapia. *Biochem. Biophys. Res. Commun.* 471, 26–33. doi: 10.1016/j.bbrc.2016.02.007

Zhou, H., and Rigoutsos, I. (2014). MiR-103a-3p targets the 5' UTR of GPRC5A in pancreatic cells. *RNA* 20, 1431–1439. doi: 10.1261/rna.045757.114

Zhu, X., Chu, W. Y., Wu, P., Yi, T., Chen, T., and Zhang, J. S. (2014). MicroRNA signature in response to nutrient restriction and re-feeding in fast skeletal muscle of grass carp (*Ctenopharyngodon idella*). *Zool. Res.* 35, 404–410. doi: 10.13918/j.issn.2095-8137.2014.5.404

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

The reviewer DTH declared a shared affiliation, with no collaboration, with several of the authors (MH, AO, PN, JC, and DP) and to the handling Editor.

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 04 April 2018
doi: 10.3389/fgene.2018.00107

Little Divergence Among Mitochondrial Lineages of *Prochilodus* (Teleostei, Characiformes)

Bruno F. Melo[1,2]*, *Beatriz F. Dorini*[1], *Fausto Foresti*[1] and *Claudio Oliveira*[1]

[1] Departamento de Morfologia, Instituto de Biociências, Universidade Estadual Paulista, Botucatu, Brazil, [2] Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, United States

Evidence that migration prevents population structure among Neotropical characiform fishes has been reported recently but the effects upon species diversification remain unclear. Migratory species of *Prochilodus* have complex species boundaries and intrincate taxonomy representing a good model to address such questions. Here, we analyzed 147 specimens through barcode sequences covering all species of *Prochilodus* across a broad geographic area of South America. Species delimitation and population genetic methods revealed very little genetic divergence among mitochondrial lineages suggesting that extensive gene flow resulted likely from the highly migratory behavior, natural hybridization or recent radiation prevent accumulation of genetic disparity among lineages. Our results clearly delimit eight genetic lineages in which four of them contain a single species and four contain more than one morphologically problematic taxon including a trans-Andean species pair and species of the *P. nigricans* group. Information about biogeographic distribution of haplotypes presented here might contribute to further research on the population genetics and taxonomy of *Prochilodus*.

Keywords: DNA barcoding, freshwater fishes, gene flow, Neotropics, Prochilodontidae, South America, taxonomy

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, Ponta Grossa State University, Brazil

Reviewed by:
Jorge Abdala Dergam, Universidade Federal de Viçosa, Brazil
Evanguedes Kalapothakis, Universidade Federal de Minas Gerais (UFMG), Brazil

***Correspondence:**
Bruno F. Melo
melo@ibb.unesp.br

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

Received: *18 October 2017*
Accepted: *19 March 2018*
Published: *04 April 2018*

Citation:
Melo BF, Dorini BF, Foresti F and Oliveira C (2018) Little Divergence Among Mitochondrial Lineages of Prochilodus (Teleostei, Characiformes). Front. Genet. 9:107. doi: 10.3389/fgene.2018.00107

INTRODUCTION

Fishes of the characiform family Prochilodontidae are widely distributed across Neotropical freshwaters and represent important fishery resources in South America (Ribeiro and Petrere, 1990; Garcia et al., 2009). The migratory behavior allows them to achieve many hundreds of river kilometers to spawn during rainy seasons (Godinho and Kynard, 2006) and, consequently, permits extensive gene flow among distant populations (Sivasundar et al., 2001; Melo et al., 2013). There is substantial evidence that migration of prochilodontids results in high levels of genetic variability and low levels of population structure (Sivasundar et al., 2001; Rueda et al., 2013; Ferreira et al., 2017; Machado et al., 2017). However, there is no empirical study aimed to detect whether long-distance migrations affect genetic diversification at species level in Neotropical freshwater fishes.

Prochilodontids represent a good model to address such questions because much research on population genetics and phylogeography have provided valuable intraspecific genetic information (Sivasundar et al., 2001; Turner et al., 2004; Hatanaka et al., 2006; Carvalho-Costa et al., 2008; Melo et al., 2013; Rueda et al., 2013; Ferreira et al., 2017; Machado et al., 2017; Sales et al., 2018).

Furthermore, recent barcoding studies in focal regions (i.e., using endemic species) have generated a robust mitochondrial database for *Prochilodus* (e.g., Carvalho et al., 2011; Rosso et al., 2012; Pereira et al., 2013; Chagas et al., 2015; Díaz et al., 2016) that, if combined, might be useful for species-level comparisons.

Prochilodontidae is represented by three genera (*Ichtyoelephas*, *Prochilodus,* and *Semaprochilodus*) spanning 21 species (Castro and Vari, 2004). While *Ichthyoelephas* and *Semaprochilodus* have well-stablished taxonomy, except for questions on species boundaries between *S. kneri* and *S. insignis* (Melo et al., 2016a), the taxonomy of *Prochilodus* remains complex. It has 13 morphologically similar species being two endemic to trans-Andean basins of Río Magdalena (*P. magdalenae*) and Lago Maracaibo (*P. reticulatus*), three from the Amazon basin: the widely distributed *P. nigricans* occupying major tributaries of western Amazon in Colombian, Peruvian, Bolivian, and Brazilian eastern rivers flowing northward such as the Madeira, Tapajós, and Tocantins; *P. rubrotaeniatus*, allopatrically distributed through portions of Rio Negro (i.e., Rio Marauiá) and adjacent Guianese rivers such as the Essequibo, Corantijn, and Marowijne river basins; and the less abundant and endangered *P. britskii* from the Rio Apiacás, a tributary of the upper Rio Tapajós. Remaining species are generally endemic to specific drainages: *P. mariae* (Río Orinoco), *P. lineatus* (La Plata and Rio Paraíba do Sul), *P. argenteus* and *P. costatus* (São Francisco), *P. harttii* and *P. vimboides* (Eastern Brazilian drainages from Rio Pardo to Rio Paraíba do Sul), *P. brevis* (coastal rivers of northeastern Brazil), and *P. lacustris* (Río Parnaíba, Northeastern Brazil). Moreover, species distribution of *Prochilodus* has suffered significant alterations due antropogenic introductions in several rivers of eastern and northeastern Brazil (Castro and Vari, 2004).

Some species groups have very subtle morphological differentiation with species being discriminated by ranges and modal meristic values, and by the biogeographic drainage where they are generally endemic (Castro and Vari, 2004). These are *Prochilodus magdalenae*/*P. reticulatus* from Magdalena-Maracaibo, *P. nigricans*/*P. rubrotaeniatus* from Amazon-Guianas-Orinoco, *P. brevis*/*P. lacustris* from northeastern Brazil, and *P. costatus*/*P. lineatus* from São Francisco-La Plata. Furthermore, a recent molecular phylogeny based on six genes revealed non-monophyly of some species, including *P. magdalenae*, *P. costatus, P. nigricans*, and *P. rubrotaeniatus* (Melo et al., 2016a). This study also revealed a problematic species complex, the *P. nigricans* group that encompasses several specimens of *P. rubrotaeniatus*, *P. brevis,* and *P. lacustris* interspersed within the *P. nigricans sensu lato*. Although Melo et al. (2016a) used various specimens of *P. nigricans* from distinct biogeographic zones across the Amazon basin, which still remain to be complete, they did not use an extensive sampling for those other problematic species.

Haplotypic variation has been applied to study the diversity of Neotropical characiform fishes as well as used to address systematic questions through the expansion of DNA barcoding projects (e.g., Pereira et al., 2011; Bellafronte et al., 2013; Castro Paz et al., 2014; Benzaquem et al., 2015; Melo et al., 2016b; Ramirez et al., 2017). The majority of barcoding studies have demonstrated high levels of interspecific variation (Melo et al., 2016b; Silva et al., 2016a) while others present a more reduced variation pattern (Pereira et al., 2011; Rossini et al., 2016). Despite a substantial number of population genetic studies applied to species of *Prochilodus* and the natural abundance of those fishes in South American rivers, no genetic study aimed to address species diversity within the genus currently exists.

In this context, barcode sequences of a higher number of specimens from distant regions in association with modern species delimitation methods and haplotype variation analysis are applicable to better determine species delineation within problematic taxa (e.g., Castro Paz et al., 2014; Costa-Silva et al., 2015; Melo et al., 2016b), as in the case of *Prochilodus*. Here, we aim to detect the effects of migration in species diversification, to delimit species of *Prochilodus* using a high taxon sampling and to advance the resolution of the problematic species boundaries within the genus.

MATERIALS AND METHODS

Taxon Sampling and DNA Sequencing

Specimens were collected under a permanent permission number 13843-1 from MMA/IBAMA/SISBIO and subsequently preserved in 95% ethanol. We included 146 specimens spanning all 13 species of *Prochilodus* collected across all South America plus *Semaprochilodus taeniurus* to root the trees (total 147 taxa). We sequenced barcodes for 19 specimens and supplemented the matrix with 127 additional barcodes of *Prochilodus* available at the public genetic databases Genbank (www.ncbi.nlm.nih.gov/) and Barcode of Life Database (BOLD; www.boldsystems.org/). **Supplementary Table S1** contains voucher and locality information and accession numbers for databases.

Genomic DNA was extracted from muscle tissues preserved in 95% ethanol with a DNeasy Tissue kit (Qiagen Inc.; http://www.qiagen.com) according to the manufacturer's instructions. We obtained partial sequences of the mitochondrial gene *cytochrome oxidase c subunit I* by amplifying via polymerase chain reaction (PCR) using the primer described in the literature (Melo et al., 2011) and modifying reaction steps as follow: 12.5 μl as a total volume with 9.075 μl of double-distilled water, 1.25 μl 5x buffer, 0.375 μl $MgCl_2$ (50 mM), 0.25 μl dNTP mix, 0.25 μl of each primer at 10 μM, 0.05 μl Platinum Taq DNA polymerase enzyme (5 units/μl, Invitrogen; www.invitrogen.com) and 1.0 μl genomic DNA (10–50 ng). The PCR consisted of an initial denaturation (4 min at 95°C) followed by 28–30 cycles of chain denaturation (30 s at 95°C), primer hybridization (30–60 s at 52–54°C), and nucleotide extension (30–60 s at 72°C). After the visualization of the fragments using 1% agarose gel, we performed the sequencing reaction using dye terminators (BigDye™ Terminator v 3.1 Cycle Sequencing Ready Reaction Kit, Applied Biosystems; http://www.appliedbiosystems.com) purified again through ethanol precipitation. We then loaded the samples onto an automatic sequencer ABI 3130-Genetic Analyzer (Applied Biosystems) at the São Paulo State University, Brazil.

Species Delimitation and Population Genetic Analyses

We assembled and edited the newly generated consensus sequences in Geneious 7.1.9 (Kearse et al., 2012) and aligned the whole matrix with Muscle (Edgar, 2004). This matrix contains 147 taxa (146 *Prochilodus* plus one *Semaprochilodus*) and 648 bp. To evaluate the occurrence of substitution saturation, the index of substitution saturation in asymmetrical (Iss.cAsym) and symmetrical (Iss.cSym) topologies were estimated in Dambe 5.3.38 (Xia, 2013). We used PartitionFinder 1.1.0 (Lanfear et al., 2012) to select the best-fit model of nucleotide evolution for our dataset.

Species were previously identified following the most recent and complete taxonomic revision (Castro and Vari, 2004), and lineages were proposed based on subsequent topologies. Most available sequences are from vouchers already identified by the first author (e.g., Melo et al., 2016b) or from previous studies with endemic species (Carvalho et al., 2011; Rosso et al., 2012; Pereira et al., 2013; Díaz et al., 2016). We then generated overall and pairwise values of genetic distance based on Kimura-2-parameters (K2P)+Gamma using Mega 7.0 (Tamura et al., 2013) and a neighbor-joining tree (NJ) with 1,000 replicates of bootstraps using Geneious 7.1.9. We also performed a maximum likelihood (ML) analysis under RAxML HPC-PTHREADS-SSE3 (Stamatakis, 2006) using five random parsimony trees with the GTRGAMMA model (Stamatakis et al., 2008) without rooting and with other parameters at default. We used the autoMRE function to generate pseudoreplicates through MRE-based stopping criteria (Pattengale et al., 2009) that ran a total of 650 replicates. Stopping criteria determine when enough replicates have been generated so that robust bootstraps under ML analysis become computationally practical (Pattengale et al., 2009).

An ultrametric gene tree was generated in a Bayesian inference with Beast 1.8.0 (Drummond et al., 2012) using two independent runs of 50 millions generations sampling trees every 5000th generation. Convergence was indicated by Tracer v1.5 (Rambaut et al., 2014) with estimated sample sizes (ESS) superior to 200. An appropriate number of trees (first 10%) from each run was discarded as burn-in and the MCMC samples was generated using the maximum clade credibility (MCC) topology in TreeAnnotator v1.4.7 (Drummond et al., 2012) and visualized in FigTree v1.4.3.

The general mixed Yule coalescent (GMYC) method (Pons et al., 2006; Fujisawa and Barraclough, 2013) was performed using the ultrametric gene tree estimated with the exponential growth coalescent model (Griffiths and Tavaré, 1994) and the lognormal relaxed clock model (Drummond et al., 2006), which assumes that the rates of molecular evolution are uncorrelated but log-normally distributed among lineages. Species delimitation through GMYC model was conducted using standard parameters [interval = c(0, 10)] and a single threshold that specifies the transition time between to within species branching. Such analysis was conducted with the package *splits* (Species Limits by Threshold Statistics; http://r-forge.r-project.org/projects/splits) in R v.3.0.0 (R Development Core Team, 2013). GMYC appears to be useful for single-locus analysis (Fujisawa and Barraclough, 2013) but depends on the availability of additional data/analyses from independent characters (Esselstyn et al., 2012). Additionally, we used the Bayesian Poisson Tree Processes model (bPTP) (Zhang et al., 2013) in the bPTP webserver (http://species.h-its.org/ptp/) under default parameters. bPTP does not require an ultrametric gene tree and uses, instead, a nexus tree as input file with branch lengths representing the number of nucleotide substitutions (Zhang et al., 2013). We used a nexus MCC tree generated in Beast 1.8.0 (Drummond et al., 2012) as input file and ran 500,000 generations (thinning = 500). We also used a clustering species delimitation analysis through the Automatic Barcode Gap Discovery (ABGD; Puillandre et al., 2012) that automatically defines sequences into hypothetical candidate species based on confidence limits for intraspecific divergence. We used a pairwise distance matrix generated in Mega 7.0 (Kumar et al., 2016) through K2P+G model and 1,000 pseudoreps as input file into the ABGD webserver (wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html) with other parameters left at defaut.

Population genetic analyses were conducted in order to detect levels of genetic variance among haplotypes. We excluded four taxa and excized flanking regions with elevated missing data to properly run those analyses. This reduced matrix contained 143 taxa and 465 bp. Each mitochondrial lineage previously determined by distance and likelihood analyses was treated as a distinct population. We used DnaSP v.5.10.01 (Librado and Rozas, 2009) to obtain the number of polymorphic sites, haplotype number, and nucleotide/haplotype diversity. In Arlequin 3.5.1 (Excoffier and Lischer, 2010), each mitochondrial lineage was set as a single population with the following hypothetic group structuring (group 1 = outgroup; group 2 = lineage 1; group3 = lineage 2; group 4 = lineages 3, 4, and 5; group 5 = lineages 6, 7, and 8) based on the arrangement from ML and Bayesian trees. We ran an analysis of molecular variance (AMOVA; Excoffier et al., 1992) with 1,000 permutations using conventional F-statistics and generated the haplotype network using the median joining analysis (Bandelt et al., 1999) incorporated in PopART 1.7 (Leigh and Bryant, 2015).

RESULTS

The final matrix contained 147 taxa, 648 bp, and 154 variable sites (23.8%). Nucleotide frequencies were 21.1% adenine, 25.0% citosine, 16.1% guanine, and 26.4% tymine. The newly generated sequences of *Prochilodus* are deposited at GenBank with accession numbers MH068824–MH068842 (**Supplementary Table S1**). The Iss indexes indicated no saturation in either transitions and transversions in both asymmetrical (Iss.cAsym) and symmetrical (Iss.cSym) topologies. The overall mean of K2P genetic distances without outgroup was 0.025 ± 0.004. Intraspecific genetic variation ranged from zero within the lineage of *P. magdalenae* and *P. reticulatus* to 0.003 within the lineage of *P. costatus* and *P. lineatus*. The lowest pairwise K2P distance was 0.012 ±

0.004 between *P. harttii* and *P. argenteus*. The highest pairwise K2P distance was 0.103 ± 0.016 between *P. vimboides* and *P. mariae*. Fourteen out of 28 pairwise comparisons received values below 0.03. **Table 1** shows intraspecific and interspecific genetic distances of each lineage.

Species delimitation analysis by GMYC evidenced the presence of eight genetic lineages (interval 3–22) that encompass the 13 valid species of *Prochilodus* (**Figure 1**). The threshold time was −0.004 and indicates the time before which all nodes reflect speciation events and after which all nodes reflect coalescent events. Maximum likelihood for the null model was 1452.484 and maximum likelihood for GMYC model was 1457.71. The bPTP species delimitation analysis through both ML and Bayesian approaches returned a slightly distinct result with a total of 10 lineages of *Prochilodus* plus outgroup. The two additional clusters refer to splits within *P. vimboides* (Lineage 1; low support = 0.569) and within *P. lineatus* (Lineage 8; low support = 0.427). ABGD resulted in eight partitions that ranged from 61 ($P = 0.001$) to one candidade species ($P = 0.03$), with one partition with eight candidate species plus outgroup ($P = 0.002$) that match those obtained in GMYC. The evidence of eight species of *Prochilodus* plus outgroup agrees with NJ and ML topologies showing well-defined branches and reciprocal monophyly. **Supplementary Figure S1** represents NJ tree and **Supplementary Figure S2** represents the best maximum likelihood tree (sum of branch lengths = 0.266).

The reduced matrix for population genetic analyses included 143 sequences (see section Materials and Methods) with 465 bp (364 invariable and 101 polymorphic sites) and a total of 47 haplotypes ($H_D = 0.951$). We found high F_{ST} values among lineages/populations ($F_{ST} = 0.331$, $P < 0.001$) ranging from 0.000 (*Prochilodus mariae* vs. *Semaprochilodus*, *P. harttii* vs. *Semaprochilodus*, *P. harttii* vs. *P. mariae*) to 1.000 (*P. magdalenae/P. reticulatus* vs. *Semaprochilodus*) but without significant values (**Supplementary Table S2**). High F_{ST} values are expected due the fact that we are treating lineages/species as populations. AMOVA results indicated that there is more variation within populations (66.9%) than among populations within groups (21.7%) or among groups (11.4%) (**Supplementary Table S3**). The haplotype network shows the distribution and interrelationships among haplotypes (**Supplementary Figure S3**).

All clusters present strong support for hypothesized lineages in the NJ (bootstrap >74%), ML (bootstrap >76%) and BI (posterior probabilities = 1) analyses. Lineage one includes *Prochilodus vimboides* from eastern Brazil including the Rio Doce, Rio Itaúnas, and Rio Mucuri. Lineage two includes *P. magdalenae* (Río Magdalena in Colombia) and *P. reticulatus* (Lago Maracaibo in Venezuela), the trans-Andean species of *Prochilodus* as a single genetic unit. Third lineage contains three specimens of *P. mariae* from Río Orinoco and lineage four has two specimens of *P. harttii* from Rio Pardo in Eastern Brazil. *Prochilodus argenteus* is represented by lineage five with 14 specimens from the upper, middle and lower Rio São Francisco plus two specimens introduced into Rio Doce and Rio Jequitinhonha. Subsequent lineages (six, seven, and eight) contain more that one species of the *P. nigricans* group (*sensu* Melo et al., 2016a) plus *P. lineatus* and *P. costatus*. The lineage six incorporates the haplotypic group composed by *P. nigricans* from uplands of the Eastern Amazon (Rio Araguaia, upper and middle Rio Tapajós), *P. britskii* from the upper Rio Tapajós (Rio Apiacás), *P. brevis* from northeastern Brazil (states of Ceará and Rio Grande do Norte), *P. lacustris* from Rio Parnaíba, and *P. rubrotaeniatus* from the upper Río Orinoco in Venezuela and the upper Essequibo river basin in Guyana. Lineage seven contains three specimens of *P. rubrotaeniatus* from Corantijn, Coppename, and Marowijne river basins in Suriname plus 26 specimens of *P. nigricans* from lowlands of the Western Amazon, including mainstream Rio Amazonas in Manaus (Brazil), the Río Itaya at the Iquitos region (Peru), Rio Madeira, and Rio Purus. Finally, the eighth lineage contains the species pair composed by *P. costatus* from distinct regions of the Rio São Francisco together with 42 specimens of *P. lineatus* from Rio Paraíba do Sul, upper Rio Paraná, upper Rio Paraguai (all in Brazil) and the lower Rio Paraná (Argentina). Analyses of NJ, ML, and BI returned similar results overall, despite some differences in the arrangement of some lineages (**Supplementary Figures S1**, **S2**).

TABLE 1 | Pairwise K2P genetic distance among distinct lineages of *Prochilodus* (below diagonal) and standard error (above diagonal).

N	Lineage	–	1	2	3	4	5	6	7	8
–	*Semaprochilodus taeniurus*	–	0.019	0.018	0.018	0.016	0.015	0.017	0.017	0.016
1	*P. vimboides*	0.149	**0.003**	0.012	0.016	0.013	0.014	0.015	0.014	0.014
2	*P. magdalenae/P. reticulatus*	0.148	0.077	**0.000**	0.013	0.010	0.010	0.012	0.011	0.011
3	*P. mariae*	0.132	0.103	0.075	**0.001**	0.007	0.007	0.007	0.008	0.007
4	*P. harttii*	0.126	0.084	0.052	0.025	**0.002**	0.004	0.006	0.007	0.006
5	*P. argenteus*	0.114	0.087	0.057	0.025	0.012	**0.002**	0.007	0.007	0.006
6	*P. brevis/P. britskii/P. lacustris/P. nigricans EA/P. rubrotaeniatus WG*	0.134	0.096	0.063	0.029	0.022	0.026	**0.002**	0.005	0.004
7	*P. nigricans WA/P. rubrotaeniatus EG*	0.137	0.091	0.064	0.034	0.028	0.026	0.016	**0.003**	0.005
8	*P. costatus/P. lineatus*	0.129	0.090	0.059	0.029	0.021	0.019	0.013	0.014	**0.003**

Intraspecific genetic distance within lineages in bold. Groups were defined based on the GMYC analysis. N, number of lineage matching those in ***Figure 1****; EA, Eastern Amazon; WA, Western Amazon; EG, Eastern Guianas; WG, Western Guianas.*

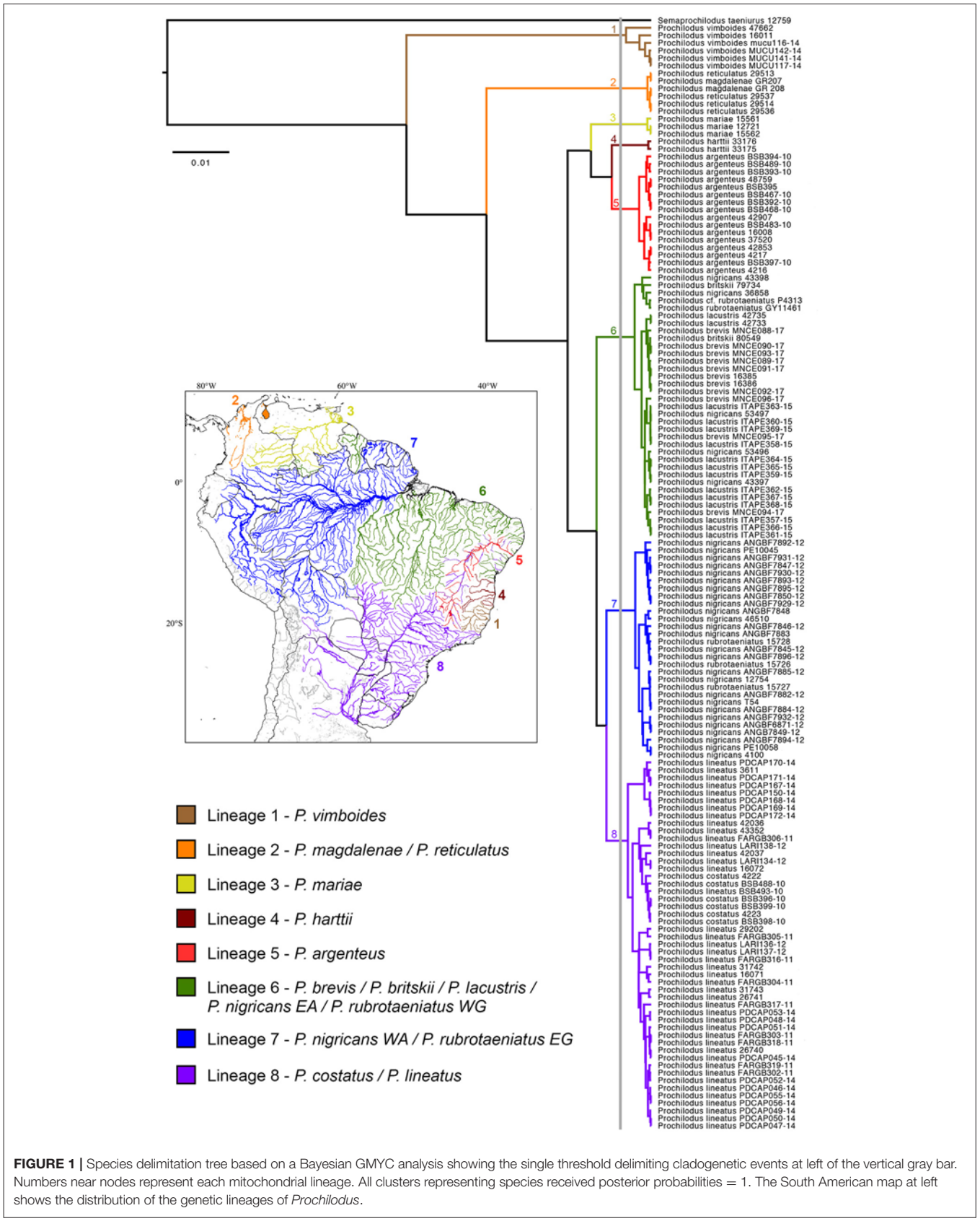


FIGURE 1 | Species delimitation tree based on a Bayesian GMYC analysis showing the single threshold delimiting cladogenetic events at left of the vertical gray bar. Numbers near nodes represent each mitochondrial lineage. All clusters representing species received posterior probabilities = 1. The South American map at left shows the distribution of the genetic lineages of *Prochilodus*.

DISCUSSION

Species Delimitation in *Prochilodus*

Results from the species delimitation analysis revealed the presence of eight genetic lineages covering 13 valid species of *Prochilodus*, in which four lineages (1, 3, 4, and 5) are structured by only one species and the other four lineages (2, 6, 7, and 8) include more than one species. Topologies (**Figure 1**, **Supplementary Figures S1**, **S2**) are quite similar to the molecular phylogeny of Prochilodontidae (Melo et al., 2016a), likely due the locus selection. *Prochilodus vimboides* (lineage 1), for example, splits from the most recent common ancestor of all other *Prochilodus*, although the biogeographic implications for this result still requires a more detailed, time-calibrated analysis of the Prochilodontidae. Other example is the structuring of *P. harttii* (lineage 4) and *P. argenteus* (lineage 5), evidencing distinct genetic lineages even with a recent evidence of hybridization (Sales et al., 2018). In contrast with the molecular phylogeny, results indicate that *P. mariae* is an exclusive cluster, suggesting inconsistencies in the phylogenetic placement of the species (Castro and Vari, 2004; Melo et al., 2016a). A phylogeographic study found that *P. mariae* diverged from *P.* cf. *rubrotaeniatus* in a very recent cladogenesis (Turner et al., 2004), which does not match the Orinoco-Amazon vicariant event resulted from the rise of Vaupes Arch during the Late Miocene (Lujan and Armbruster, 2011).

Our findings suggest the recognition of only one trans-Andean species, in which *Prochilodus magdalenae* from Río Magdalena remains nested within *P. reticulatus* from Lago Maracaibo (lineage 2) as proposed by the molecular phylogeny (Melo et al., 2016a). Species limits among them involve subtle differences in the range and modal values of number of lateral line scales, number of predorsal scales, and number of vertebrae (Castro and Vari, 2004). A further phylogeographic study involving samples from Atrato, Cauca-Magdalena, and Maracaibo might help to elucidate the allopatric distribution of this mitochondrial lineage.

The Amazon basin harbors two distinct mitochondrial lineages of *Prochilodus nigricans* (lineages 6 and 7). A recent study detected population structure in western populations of *P. nigricans* (Madeira and Purus) compared to those from mainstream Rio Amazonas (Machado et al., 2017) despite the lack of samples from eastern tributaries. Interestingly, there is ecological evidence of two distinct migration patterns of *P. nigricans* in the Amazon basin (Araújo-Lima and Ruffino, 2003) that might explain our results. The first involves lateral migrations from floodplain lakes to the mainstream Rio Amazonas with subsequent migration upstream to breeding and spawning (Fernandes, 1997), and the second involves only upstream migrations to upper Rio Tocantins or Araguaia to spawning and downstream migrations to feeding (Carvalho and Mérona, 1986), the latter similar to the well-known pattern observed for *P. argenteus* (Godinho and Kynard, 2006) and *P. lineatus* (Agostinho et al., 2004).

Our results do not support the presence of multiple species within lineage six. *Prochilodus britskii*, a morphologically distinct species, appears for the first time embedded within the lineage, differently from the position as sister to *P. mariae* (Melo et al., 2016a). Both *P. brevis* and *P. lacustris* from northeastern Brazil are distinguished from the species pair *P. nigricans* and *P. rubrotaeniatus* by radial subdivision patterns on body scales (Castro and Vari, 2004). Morphological features diagnosing each of them include overlapped counts of lateral line scales, number of horizontal scale rows below lateral line, and number of circumpeduncular scale rows (Castro and Vari, 2004).

Castro and Vari (2004) redescribed *P. nigricans* by examining almost one thousand Amazonian specimens including type specimens. They designated a neotype from Lago Janauacá at the right margin of Rio Solimões near Manaus in Brazil. Twenty-one individuals of *P. nigricans* from Manaus appear within lineage seven along with specimens from western Amazon. Therefore, this cluster likely constitutes the genetic lineage of the neotype. The position of *P. rubrotaeniatus* still represents a lacuna in our knowledge due the presence of the species in both lineages six and seven. Based on a previous phylogeographic study (Turner et al., 2004), Albert et al. (2011) suggest that *P. rubrotaeniatus* represents an example of paraspecies that gave rise to the endemic *P. mariae*. Paraspecies are paraphyletic, geographically widespread species that originates another peripheral isolated species without becoming extinct (Ackery and Vane-Wright, 1984; Albert et al., 2011). The phylogenetic evidence (Melo et al., 2016a) and the results arrived herein refute such hypothesis and instead, indicate that the present concept of *P. rubrotaeniatus* constitute more than one genetic lineage.

In the Brazilian Shield, *Prochilodus costatus* share the same mitochondrial cluster with *P. lineatus* (lineage 8), which again corroborates the molecular phylogeny (Melo et al., 2016a) and a mitogenome analysis (Chagas et al., 2015). Analyzed specimens of *P. costatus* from Rio Pandeiros/São Francisco are remarkably distant (∼2,500 linear km) from analyzed specimens of *P. lineatus* from Rosario in Argentina. These results agree with previous population genetic and phylogeographic studies that show high genetic diversity and low population divergence (Sivasundar et al., 2001; Carvalho-Costa et al., 2008; Melo et al., 2013; Ferreira et al., 2017). A phylogeographic study of *P. lineatus*, for example, found strong similarity among mitochondrial control regions between samples from the lower Rio Paraná in Argentina and upper Rio Paraná in Brazil (Sivasundar et al., 2001). Overlapped counts of lateral line scales, number of vertebrae and allopatry slightly discriminate the two species (Castro and Vari, 2004), which is clearly not supported herein. In addition, our evidence indicates that future population genetic studies of one or another species should include members of both nominal species.

Little Divergence Among Lineages of *Prochilodus*

Results indicate very little mitochondrial divergence among lineages of *Prochilodus* and provide evidence that distantly sampled specimens, in various instances, correspond to a single mitochondrial lineage. The most plausible hypothesis that might explain such result is that migration affects species diversification. Indeed, migration has been used to explain high levels of gene

flow and low population structure in *Prochilodus* (e.g., Sivasundar et al., 2001; Carvalho-Costa et al., 2008; Melo et al., 2013; Ferreira et al., 2017). This is supported by ecological data from fish tagging that found migratory routes of >120 km for *P. argenteus* along the Rio São Francisco (Godinho and Kynard, 2006) and 250 km of *Prochilodus* sp. in that same basin (Paiva and Bastos, 1981).

Migration and gene flow directly influence morphological stasis (Stanley, 1979). This raises some questions about how migration patterns have influenced population diversification without morphological change in *Prochilodus*. Would distinct environmental settings be responsible for distinct movement behaviors along their evolutionary history? López-Fernández and Albert (2011) suggest that massive prochilodontid migrations evolved during the Oligocene, before the separation of the paleo-Amazon-Orinoco river basin, and that posterior vicariant events allowed their successful colonization throughout major Neotropical basins. The two allopatric lineages of *P. nigricans* and their two migration patterns (Carvalho and Mérona, 1986; Fernandes, 1997) support López-Fernández and Albert (2011)'s conclusion and reinforce the fact that a lineage (or population) once fragmented tends to search for ecological adaptation in distinct environmental conditions. Colonization of Neotropical habitats in upland rivers requires adaptation to a strong selective pressure by acquiring specific morphological innovations (Silva et al., 2016b) or behavior specializations, which appears to be the case of *Prochilodus*.

It is noteworthy, however, that migration is not the exclusive factor aging disfavoring species diversification in *Prochilodus*. There are, at least, two more plausible explanations for the observed low genetic variation. Hybridization between the native *P. harttii* and the introduced *P. argenteus* has been genetically identified in the Rio Jequitinhonha recently (Sales et al., 2018). Although hybridization between native species has not been documented yet, the process might be included as another hypothesis to explain our results, at least for sympatric species. Another plausible hypothesis would be recent episodes of species diversification that did not allow accumulation of haplotype variation. Despite a molecular phylogeny is available (Melo et al., 2016a), the lack of a time-calibrated tree does not allow us to support this hypothesis with better confidence. Testing those three most plausible hypotheses to explain the little divergence among lineages of *Prochilodus* is thus a matter of further research.

AUTHOR CONTRIBUTIONS

BM, FF, and CO designed the project; BM and BD generated, analyzed, and compiled the data; BM wrote most of the text; BM, BD, FF, and CO revised and approved the final version of the manuscript.

ACKNOWLEDGMENTS

The authors thank F. Y. Ashikaga for helping with the population genetic analyses. Research received financial support from FAPESP grants 2011/08374-1, 2013/16436-2, 2016/11313-8 (BM), PIBIC-CNPq (BD), FAPESP grant 2014/26508-3, and CNPq grant 306054/2006-0 (CO).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00107/full#supplementary-material

Supplementary Table S1 | Lineage, taxon, voucher, locality information, and Genbank accession numbers of the analyzed specimens of *Prochilodus*. Lines in bold indicate sequences generated in the present study and asterisks represent BOLD accession numbers.

Supplementary Table S2 | Pairwise F_{ST} values among mitochondrial lineages of *Prochilodus*. $^*P < 0.05$.

Supplementary Table S3 | Analysis of molecular variance (AMOVA) among lineages of *Prochilodus*. Groups were ordered on the basis of previous ML and Bayesian analyses (see section Material and Methods).

Supplementary Figure S1 | Neighbor-joining tree of the species of *Prochilodus* based on partial sequences of the cytochrome oxidase *c* subunit I. Numbers near nodes represent bootstrap support.

Supplementary Figure S2 | Maximum likelihood tree of the *Prochilodus* species based on partial sequences of the cytochrome oxidase *c* subunit I. Numbers near nodes represent bootstrap support. Colors match those in **Figure 1**.

Supplementary Figure S3 | Haplotype network of the eight mitochondrial lineages of *Prochilodus*. Each circle represents a unique haplotype and the size proportional to haplotype frequency. Colors match those in **Figure 1**.

REFERENCES

Ackery, P. R., and Vane-Wright, R. I. (1984). *Milkweed Butterflies their Cladistics and Biology, Being an Account of the Natural History of the Danainae, a Subfamily of the Lepidoptera, Nymphalidae*. London: British Museum (Natural History).

Agostinho, A. A., Gomes, L. C., Suzuki, H. I., and Júlio, H. F. Jr. (2004). "Migratory fishes of the upper Paraná river basin," in *Migratory Fishes of South America: Biology, Fisheries and Conservation Status*, eds J. Carolsfeld, B. Harvey, C. Ross, and A. Baer (Victoria, BC: World Fisheries Trust), 19–98.

Albert, J., Petry, P., and Reis, R. (2011). "Major biogeographic and phylogenetic patterns," in *Historical Biogeography of Neotropical Freshwater Fishes*, eds J. S. Albert and R. E. Reis (Berkeley, CA; Los Angeles, CA: University of California Press), 21–57.

Araújo-Lima, C. A. R. M., and Ruffino, M. L. (2003). "Migratory fishes of the Brazilian Amazon," in *Migratory Fishes of South America: Biology, Fisheries and Conservation Status*, eds J. Carolsfeld, B. Harvey, C. Ross, and A. Baer (Victoria, BC: World Fisheries Trust), 233–301.

Bandelt, H. J., Forster, P., and Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48. doi: 10.1093/oxfordjournals.molbev.a026036

Bellafronte, E., Mariguela, T. C., Pereira, L. H. G., Oliveira, C., and Moreira-Filho, O. (2013). DNA barcode of Parodontidae species from the La Plata river basin - applying new data to clarify taxonomic problems. *Neotrop. Ichthyol.* 11, 497–506. doi: 10.1590/S1679-62252013000300003

Benzaquem, D. C., Oliveira, C., da Silva Batista, J., Zuanon, J., and Porto, J. I. R. (2015). DNA barcoding in pencilfishes (Lebiasinidae: *Nannostomus*) reveals cryptic diversity across the Brazilian amazon. *PLoS ONE* 10:e0112217. doi: 10.1371/journal.pone.0112217

Carvalho, D. C., Oliveira, D. A., Pompeu, P. S., Leal, C. G., Oliveira, C., and Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes:

the case of the São Francisco River basin. *Mitochondrial DNA* 22, 80–86. doi: 10.3109/19401736.2011.588214

Carvalho, J. L., and Mérona, B. (1986). Estudos sobre dois peixes migratorios do baixo Tocantins, antes do fechamento da barragem de Tucurui. *Amazoniana* 4, 595–607.

Carvalho-Costa, L. F., Hatanaka, T., and Galetti, P. Jr. (2008). Evidence of lack of population substructuring in the Brazilian freshwater fish *Prochilodus costatus*. *Genet. Mol. Biol.* 31, 377–380. doi: 10.1590/S1415-47572008000200036

Castro, R. M. C., and Vari, R. P. (2004). Detritivores of the South American fish family Prochilodontidae (Teleostei: Ostariophysi: Characiformes): a phylogenetic and revisionary study. *Smithson. Contrib. Zool.* 622, 1–189. doi: 10.5479/si.00810282.622

Castro Paz, F. P., Batista, J. D. S., and Porto, J. I. R. (2014). DNA barcodes of rosy tetras and allied species (Characiformes: Characidae: *Hyphessobrycon*) from the Brazilian Amazon basin. *PLoS ONE* 9:e98603. doi: 10.1371/journal.pone.0098603

Chagas, A. T., Carmo, A. O., Costa, M. A., Resende, L. C., Brandão Dias, P. F., Martins, A. P., et al. (2015). Description and comparison of two economically important fish species mitogenomes: *Prochilodus argenteus* and *Prochilodus costatus* (Characiformes, Prochilodontidae). *Mitochondrial DNA* 27, 2852–2853. doi: 10.3109/19401736.2015.1053125

Costa-Silva, G. J., Rodriguez, M. S., Roxo, F. F., Foresti, F., and Oliveira, C. (2015). Using different methods to access the difficult task of delimiting species in a complex neotropical hyperdiverse group. *PLoS ONE* 10:e0135075. doi: 10.1371/journal.pone.0135075

Díaz, J., Villanova, G. V., Brancolini, F., Del Pazo, F., Posner, V. M., Grimberg, A., et al. (2016). First DNA barcode reference library for the identification of south American freshwater fish from the lower Paraná river. *PLoS ONE* 11:e0157419. doi: 10.1371/journal.pone.0157419

Drummond, A. J., Ho, S. Y. .w, Phillips, M. J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88. doi: 10.1371/journal.pbio.0040088

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. doi: 10.1093/molbev/mss075

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113. doi: 10.1186/1471-2105-5-113

Esselstyn, J. A., Evans, B. J., Sedlock, J. L., Anwarali Khan, F. A., and Heaney, L. R. (2012). Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. *Proc. R. Soc. B Biol. Sci.* 279, 3678–3686. doi: 10.1098/rspb.2012.0705

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Excoffier, L., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.

Fernandes, C. C. (1997). Lateral migration of fishes in Amazon floodplains. *Ecol. Freshw. Fish* 6, 36–44. doi: 10.1111/j.1600-0633.1997.tb00140.x

Ferreira, D. G., Souza-Shibatta, L., Shibatta, O. A., Sofia, S. H., Carlsson, J., Dias, J. H. P., et al. (2017). Genetic structure and diversity of migratory freshwater fish in a fragmented Neotropical river system. *Rev. Fish Biol. Fish.* 27, 209–231. doi: 10.1007/s11160-016-9441-2

Fujisawa, T., and Barraclough, T. G. (2013). Delimiting species using single-locus data and the generalized mixed yule coalescent approach: a revised method and evaluation on simulated data sets. *Syst. Biol.* 62, 707–724. doi: 10.1093/sysbio/syt033

Garcia, A., Tello, S., Vargas, G., and Duponchelle, F. (2009). Patterns of commercial fish landings in the Loreto region (Peruvian Amazon) between 1984 and 2006. *Fish Physiol. Biochem.* 35, 53–67. doi: 10.1007/s10695-008-9212-7

Godinho, A. L., and Kynard, B. (2006). Migration and Spawning of Radio-Tagged Zulega *Prochilodus argenteus* in a Dammed Brazilian River. *Trans. Am. Fish. Soc.* 135, 811–824. doi: 10.1577/T04-176.1

Griffiths, R. C., and Tavaré, S. (1994). Sampling theory for neutral alleles in a varying environment. *Philos. Trans. Biol. Sci.* 344, 403–410. doi: 10.1098/rstb.1994.0079

Hatanaka, T., Henrique-Silva, F., and Galetti, P. Jr. (2006). Population substructuring in a migratory freshwater fish *Prochilodus argenteus* (Characiformes, Prochilodontidae) from the São Francisco River. *Genetica* 126, 153–159. doi: 10.1007/s10709-005-1445-0

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lanfear, R., Calcott, B., Ho, S. Y. W., and Guindon, S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29, 1695–1701. doi: 10.1093/molbev/mss020

Leigh, J. W., and Bryant, D. (2015). POPART: full-feature software for haplotype network construction. *Methods Ecol. Evol.* 6, 1110–1116. doi: 10.1111/2041-210X.12410

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187

López-Fernández, H., and Albert, J. S. (2011). "Paleogene radiations," in *Historical Biogeography of Neotropical Freshwater Fishes*, eds J. S. Albert and R. E. Reis (Berkeley, CA; Los Angeles, CA: University of California Press), 105–117.

Lujan, N., and Armbruster, J. (2011). "The Guiana Shield," in *Historical Biogeography of Neotropical Freshwater Fishes*, eds J. S. Albert and R. E. Reis (Berkeley, CA; Los Angeles, CA: University of California Press), 211–224.

Machado, V. N., Willis, S. C., Teixeira, A. S., Hrbek, T., and Farias, I. P. (2017). Population genetic structure of the Amazonian black flannelmouth characin (Characiformes, Prochilodontidae: *Prochilodus nigricans* Spix & Agassiz, 1829): contemporary and historical gene flow of a migratory and abundant fishery species. *Environ. Biol. Fishes* 100, 1–16. doi: 10.1007/s10641-016-0547-0

Melo, B. F., Benine, R. C., Mariguela, T. C., and Oliveira, C. (2011). A new species of *Tetragonopterus* Cuvier, 1816 (Characiformes: Characidae: Tetragonopterinae) from the rio Jari, Amapá, northern Brazil. *Neotrop. Ichthyol.* 9, 49–56. doi: 10.1590/S1679-62252011000100002

Melo, B. F., Ochoa, L. E., Vari, R. P., and Oliveira, C. (2016b). Cryptic species in the Neotropical fish genus *Curimatopsis* (Teleostei, Characiformes). *Zool. Scr.* 45, 650–658. doi: 10.1111/zsc.12178

Melo, B. F., Sato, Y., Foresti, F., and Oliveira, C. (2013). The roles of marginal lagoons in the maintenance of genetic diversity in the Brazilian migratory fishes *Prochilodus argenteus* and *P. costatus*. *Neotrop. Ichthyol.* 11, 625–636. doi: 10.1590/S1679-62252013000300016

Melo, B. F., Sidlauskas, B. L., Hoekzema, K., Frable, B. W., Vari, R. P., and Oliveira, C. (2016a). Molecular phylogenetics of the Neotropical fish family Prochilodontidae (Teleostei: Characiformes). *Mol. Phylogenet. Evol.* 102, 189–201. doi: 10.1016/j.ympev.2016.05.037

Paiva, M. P., and Bastos, S. A. (1981). Marcações de peixes nas regiões do alto e médio São Francisco. *Cienc. Cult.* 34, 1362–1365.

Pattengale, N. D., Alipour, M., Bininda-Emonds, O. R. P., Moret, B. M. E., and Stamatakis, A. (2009). How many bootstrap replicates are necessary? *Lect. Notes Comput. Sci.* 5541, 184–200. doi: 10.1007/978-3-642-02008-7_13

Pereira, L. H. G., Hanner, R., Foresti, F., and Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse neotropical freshwater fish fauna? *BMC Genet.* 14:20. doi: 10.1186/1471-2156-14-20

Pereira, L. H. G., Pazian, M. F., Hanner, R., Foresti, F., and Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (Characiformes: Characidae) from the Upper Paraná Basin of Brazil. *Mitochondrial DNA* 22, 87–96. doi: 10.3109/19401736.2011.588213

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., et al. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55, 595–609. doi: 10.1080/10635150600852011

Puillandre, N., Lambert, A., Brouillet, S., and Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* 21, 1864–1877. doi: 10.1111/j.1365-294X.2011.05239.x,

Rambaut, A., Suchard, M. A., Xie, W., and Drummond, A. J. (2014). *Tracer v1.6.1*. Available online at: http://beast.bio.ed.ac.uk

Ramirez, J. L., Birindelli, J. L., Carvalho, D. C., Affonso, P. R. A. M., Venere, P. C., Ortega, H., et al. (2017). Revealing hidden diversity of the underestimated Neotropical ichthyofauna: DNA barcoding in the recently described genus *Megaleporinus* (Characiformes: Anostomidae). *Front. Genet.* 8:149. doi: 10.3389/fgene.2017.00149

R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna.

Ribeiro, M. C. L. B., and Petrere, M. Jr. (1990). Fisheries ecology and management of the jaraqui (*Semaprochilodus taeniurus, S. insignis)* in central Amazonia. *Regul. Rivers Res. Manag.* 5, 195–215. doi: 10.1002/rrr.3450050302

Rossini, B. C., Oliveira, C. A. M., Gonçalves De Melo, F. A., De Araújo Bertaco, V., Díaz De Astarloa, J. M., Rosso, J. J., et al. (2016). Highlighting Astyanax species diversity through DNA barcoding. *PLoS ONE* 11:e0167203. doi: 10.1371/journal.pone.0167203

Rosso, J. J., Mabragaña, E., González Castro, M., and Díaz de Astarloa, J. M. (2012). DNA barcoding Neotropical fishes: recent advances from the Pampa Plain, Argentina. *Mol. Ecol. Resour.* 12, 999–1011. doi: 10.1111/1755-0998.12010

Rueda, E. C., Carriquiriborde, P., Monzón, A. M., Somoza, G. M., and Ortí, G. (2013). Seasonal variation in genetic population structure of sábalo (*Prochilodus lineatus*) in the Lower Uruguay River. *Genetica* 141, 401–407. doi: 10.1007/s10709-013-9739-0

Sales, N. G., Pessali, T. C., Andrade Neto, F. R., and Carvalho, D. C. (2018). Introgression from non-native species unveils a hidden threat to the migratory Neotropical fish *Prochilodus hartii*. *Biol. Invasions*. 20, 555–566. doi: 10.1007/s10530-017-1556-4

Silva, G. S. C., Melo, B. F., Oliveira, C., and Benine, R. C. (2016a). Revision of the South American genus *Tetragonopterus* Cuvier, 1816 (Teleostei: Characidae) with description of four new species. *Zootaxa* 4200, 1–46. doi: 10.11646/zootaxa.4200.1.1

Silva, G. S. C., Roxo, F. F., Lujan, N. K., Tagliacollo, V. A., Zawadzki, C. H., and Oliveira, C. (2016b). Transcontinental dispersal, ecological opportunity and origins of an adaptive radiation in the Neotropical catfish genus *Hypostomus* (Siluriformes: Loricariidae). *Mol. Ecol.* 25, 1511–1529. doi: 10.1111/mec.13583

Sivasundar, A., Bermingham, E., and Ortí, G. (2001). Population structure and biogeography of migratory freshwater fishes (*Prochilodus*: Characiformes) in major South American rivers. *Mol. Ecol.* 10, 407–417. doi: 10.1046/j.1365-294x.2001.01194.x

Stamatakis, A. (2006). RAxML-VI-HPC: maximum-likelihood- based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446

Stamatakis, A., Hoover, P., and Rougemont, J. (2008). A rapid boostrap algorithm for the RAxML Web servers. *Syst. Biol.* 57, 758–771. doi: 10.1080/10635150802429642

Stanley, S. M. (1979). *Macroevolution: Pattern and Process*. Baltimore, MD: The John Hopkins University Press.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Turner, T., McPhee, M. V., Campbell, P., and Winemiller, K. O. (2004). Phylogeography and intraspecific genetic variation of prochilodontid fishes endemic to rivers of northern South America. *J. Fish Biol.* 64, 186–201. doi: 10.1111/j.1095-8649.2004.00299.x

Xia, X. (2013). DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Mol. Biol. Evol.* 30, 1720–1728. doi: 10.1093/molbev/mst064

Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29, 2869–2876. doi: 10.1093/bioinformatics/btt499

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 16 April 2018
doi: 10.3389/fgene.2018.00131

Trends in Karyotype Evolution in *Astyanax* (Teleostei, Characiformes, Characidae): Insights From Molecular Data

Rubens Pazza[1], Jorge A. Dergam[2] and Karine F. Kavalco[1]*

[1] Laboratory of Ecological and Evolutionary Genetics, Institute of Biological and Health Sciences, Federal University of Viçosa, Rio Paranaíba, Brazil, [2] Laboratory of Molecular Systematics "Beagle", Department of Animal Biology, Federal University of Viçosa, Viçosa, Brazil

The study of patterns and evolutionary processes in neotropical fish is not always an easy task due the wide distribution of major fish groups in large and extensive river basins. Thus, it is not always possible to detect or correlate possible effects of chromosome rearrangements in the evolution of biodiversity. In the *Astyanax* genus, chromosome data obtained since the 1970s have shown evidence of cryptic species, karyotypic plasticity, supernumerary chromosomes, triploidies, and minor chromosomal rearrangements. In the present work, we map and discuss the main chromosomal events compatible with the molecular evolution of the genus *Astyanax* (Characiformes, Characidae) using mitochondrial DNA sequence data, in the search for major chromosome evolutionary trends within this taxon.

Keywords: cytotaxonomy, molecular evolution, chromosomal rearrangements, mtDNA, chromosomal symplesiomorphy, chromosomal synapomorphy, chromosomal autapomorphy

OPEN ACCESS

Edited by:
Roberto Ferreira Artoni, State University of Ponta Grossa, Brazil

Reviewed by:
Wenyan Du, Alforex Seeds LLC, United States
Filipe Augusto A. G. Gonçalves De Melo, Universidade Estadual do Piauí, Brazil

****Correspondence:***
Rubens Pazza
rpazza@ufv.br

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

Received: *01 November 2017*
Accepted: *03 April 2018*
Published: *16 April 2018*

Citation:
Pazza R, Dergam JA and Kavalco KF (2018) Trends in Karyotype Evolution in Astyanax (Teleostei, Characiformes, Characidae): Insights From Molecular Data. Front. Genet. 9:131. doi: 10.3389/fgene.2018.00131

INTRODUCTION

The role of chromosomal rearrangements in the evolution of organisms has been a matter of debate for many years. The initial observations that closely related species differ in their karyotypes was later supported by evidence that unbalanced rearrangements can interfere with gametogenesis, decrease gene flow, and reinforce reproductive isolation leading to speciation (Rieseberg, 2001; Navarro and Barton, 2003). On the other hand, some organisms tolerate a certain amount of chromosomal rearrangement, which is often referred to as karyotypic plasticity (Jónsson et al., 2014; Havelka et al., 2016). This complicates the task of explaining the possible role of rearrangements in the evolution of organisms, since the variation may result in speciation or stay as polymorphism within populations.

Since Moreira-Filho and Bertollo (1991), who proposed *Astyanax scabripinnis* as a species complex based on cytogenetic and morphometric data, chromosome variation has been regarded as part of speciation processes in *Astyanax*. Before that, however, it was already acknowledged that some populations could differ in their karyotype formulae and diploid numbers (Morelli et al., 1983). Thus, it is now known that populations currently assigned to *A. scabripinnis* and to the other nominal species are characterized by some degree of inter- or intra-population chromosome variation, with diploid numbers ranging from 46 to 50 chromosomes; such is the case of *Astyanax fasciatus* (Pazza et al., 2006). Molecular cytogenetics (i.e., satellite DNA and ribosomal

genes localization) has also provided further taxonomically informative data (for a review, see Pazza and Kavalco, 2007).

Although little is known about the cause and effect of the rearrangements during the evolutionary process, the correlation between independent data sets such as molecular and chromosomal data suggest that some karyotypic signatures are associated with organismic evolution. Based on *Cytochrome B (CytB)* sequences, Mello et al. (2015) found three genetically distinct clades out of 17 *Astyanax* nominal species from the Iguaçu and adjacent river basins. Likewise, using *Cytochrome Oxydase I* (*COI*) sequences, Rossini et al. (2016) found five clades out of 64 nominal species plus 12 provisionally identified taxa. Based on DNA barcoding criteria, these authors identified only 21 morphological species. These studies also suggest the possibility of horizontal transfer that affects some species within otherwise cogent taxa. Indeed, hybridization may play a fundamental role in the genus speciation, especially when these events involve chromosomal characteristics. The matter is particularly difficult since the claims of possible natural hybrids in the specialized literature are rare (Artoni et al., 2006; Pazza et al., 2006; Sassi et al., 2018).

Thus, we analyzed mitochondrial sequences from individuals with known chromosome characteristics within the five clades proposed by Rossini et al. (2016) to describe the relations between DNA sequences and chromosomal rearrangements. We obtained a phylogenetic tree that depicts chromosomal characteristics that are strongly associated with the proposed clades.

MATERIALS AND METHODS

In the present study we sequenced samples of 195 individuals from 16 nominal species of the genus *Astyanax*, deposited in the tissue collection of the Laboratory of Ecological and Evolutionary Genetics (LaGEEvo), of the Federal University of Viçosa, Rio Paranaíba campus. The species were chosen based on their chromosomal data as analyzed in previous works and because they were included in Rossini et al. (2016) clades. Geographic coordinates, Vouchers, and GenBank sequence access numbers are summarized in Supplementary Table S1. This study was carried out in accordance with the recommendations of the Guide for the Care and Use of Laboratory Animals by the Conselho Nacional de Controle de Experimentação Animal (CONCEA). Tissue samples of fish of the genus *Astyanax* deposited in the tissue bank of the Laboratory of Ecological and Evolutionary Genetics (LaGEEvo). All the specimens used had their chromosomal data analyzed in previous works. No additional animals were sacrificed to this study.

Our hypothesis on chromosome evolution was based on the following chromosome characters: diploid numbers, fundamental number (FN), location of 5S DNA ribosomal sites, presence/absence of a 5S rDNA site in a specific submetacentric, referred to as the "marker" chromosome by Almeida-Toledo et al. (2002) and Kavalco et al. (2005), and finally, the amount and distribution of the repetitive DNA As-51 probe (Mestriner et al., 2000).

Total DNA extraction from muscle or liver samples was carried out using commercial kits (PureLink Genomic DNA minikit, Invitrogen^{TM}), according to the manufacturer's instructions. Amplification of mitochondrial (mtDNA) subunits 6 and 8 of the *ATP synthase* enzyme gene (*ATPase 6/8*) was accomplished using the primers ATP8.2-L8331 (5′-AAAGCRTTRGCCTTTTAAAGC-3′) and CO3.2-H9236 (5′-GTTAGTGGTCAGGGCTTGGRTC-3′) (Sivasundar et al., 2001). PCR was performed in a final volume of 25 μL, with 2.5 μL of 10× Taq buffer, 1 μL of $MgCl_2$, 1 μL of each primer, 0.2 μL Taq DNA polymerase, 12.8 μL ultrapure water, 1.5 μL of dNTP, and 5 μL of DNA. The amplification reactions were performed in a thermocycler at 95°C for initial denaturation (2 min) and 30 cycles of 94°C (30 s), 58°C (30 s), and 72°C (1 min). The PCR product was visualized on 1% agarose gel; purification and sequencing was performed by a third-party company (Macrogen, Korea).

Sequence editing was performed using Chromas Lite v2.01 and sequence identity was checked with BLASTn[1]. Sequences alignment was carried out with ClustalW v1.6 (Thompson et al., 1994) as implemented in MEGA v7 (Kumar et al., 2016). A Maximum Likelihood was obtained using the best model fit with MEGA v7 (Kumar et al., 2016), and chromosome characters were plotted on this ML phylogram. Phylogenetic signal was estimated using bootstrap (Felsenstein, 1985).

RESULTS

A total of 195 mitochondrial DNA sequences were obtained from individuals with known chromosomal characteristics, corresponding to 16 nominal species of *Astyanax* from the Neotropical region. The *ATPase 8* gene yielded a 530 bp partial sequence without insertions, deletions, or stop codons, and the substitution model was Tn93 + G. The maximum likelihood phylogram indicated four main clades with strong bootstrap support. The main events of chromosome differentiation were plotted on a simplified phylogenetic tree (**Figure 1**) according to the trends observed in the analyzed specimens and data from the literature. At its root, and according to the chromosomal characteristics of closely related species, we propose a most recent common ancestor with a karyotype with 2n = 50 chromosomes, low FN, one or two pairs of chromosomes bearing 5S ribosomal rDNA sites at the terminal region, and multiple 18S rDNA sites.

Clade 1

The first clade, composed of 36 individuals, corresponds to exclusively coastal species, here represented here by *A. ribeirae*, *A. intermedius*, *A. giton*, and *A. hastatus* (**Figure 2**), which have been recently proposed as members of the Probolodini (Silva, 2017). In this group, all individuals have 2n = 50 chromosomes; low FN; absence of repetitive DNA As51 and the 5S rDNA marker chromosome; little constitutive heterochromatin, distributed mainly in the pericentromeric region; a greater number of 5S

[1]http://www.ncbi.nlm.nih.gov

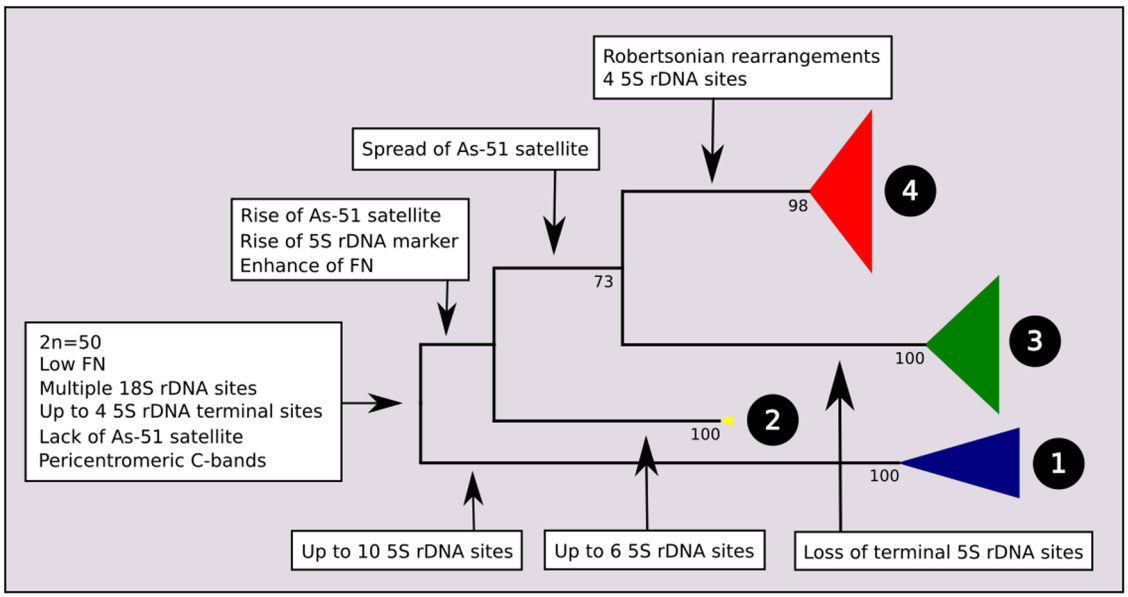


FIGURE 1 | Molecular Phylogenetic analysis by Maximum Likelihood Method (MCL) with synapomorphies of each clade. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura 3-parameter model. The tree with the highest log likelihood (–3974.2120) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the MCL approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites [five categories (+*G*, parameter = 0.4001)]. The condensed tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 196 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 531 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. Value of bootstrap related to 500 replicates is below of branches.

rDNA sites (ranging from 6 to 10 sites) distributed in terminal regions; and a variable number of Ag-NORs and 18S rDNA sites.

Clade 2

This clade is composed of five individuals of *Astyanax mexicanus* (**Figure 3**). This species has a few repetitive As51 DNA sites and 5S rDNA is distributed on six sites, including the submetacentric, marker chromosome pair.

Clade 3

This clade comprises 71 individuals with oval humeral spots distributed in coastal, Upper Paraná, Paraguay, and São Francisco river basins, belonging to the *Astyanax bimaculatus* species complex, such as *A. altiparanae*, *A.* aff. *bimaculatus*, *A. lacustris*, *A. assuncionensis*, and *A. abramis* (**Figure 4**). They all have 50 chromosomes, with a predominance of submetacentric chromosomes, and consequently high FN numbers. The presence of As51 repetitive DNA in a few sites and the presence of 5S rDNA restricted to the submetacentric marker chromosome is remarkable. The distribution of heterochromatin and rDNA 18S/Ag-NORs is variable.

Clade 4

This clade encompasses species widely distributed in the Upper Paraná, São Francisco, and coastal river basins. Eighty four individuals of *A. paranae*, *A. rivularis*, *A. bifasciatus*, *A. fasciatus*, and *A. bockmanni* are represented in this clade (**Figure 5**). Species of this clade present high intra and interspecific chromosome variability, with 2n = 46 to 2n = 50 chromosomes. Chromosomes carrying As51 satellite DNA sites vary from none to 14 sites. Most species have four 5S rDNA sites, which may include the chromosomal marker pair. As in Clade 3, they have FN numbers, relatively few acrocentric chromosomes, 18S rDNA sites and variable Ag-NORs, and highly variable constitutive heterochromatin distribution.

DISCUSSION

The phylogram obtained in the present work using the mitochondrial DNA sequence (*ATPase* subunit 6) is mostly congruent with that obtained by Mello et al. (2015) using the mitochondrial DNA sequence of the *cytochrome b* gene, as well as that obtained by Rossini et al. (2016) using the mitochondrial *COI* sequence. Karyotypical data allow to give further support to at least some monophyletic groupings within the current genus *Astyanax*, corresponding to the major clades, as it was proposed by Travenzoli et al. (2015) in Bryconidae. Despite this cohesion, some species seem to have haplotypes distributed in different clades, as demonstrated by Rossini et al. (2016). According to the authors, the *COI* sequence is not an appropriate tool to recover phylogenies, but rather to identify species (Hebert and Gregory, 2005). For the *Astyanax* genus, this sequence also seems to be unsuitable for species identification, since of the more than 70 nominal species analyzed by Rossini et al. (2016), only 21 were unequivocally

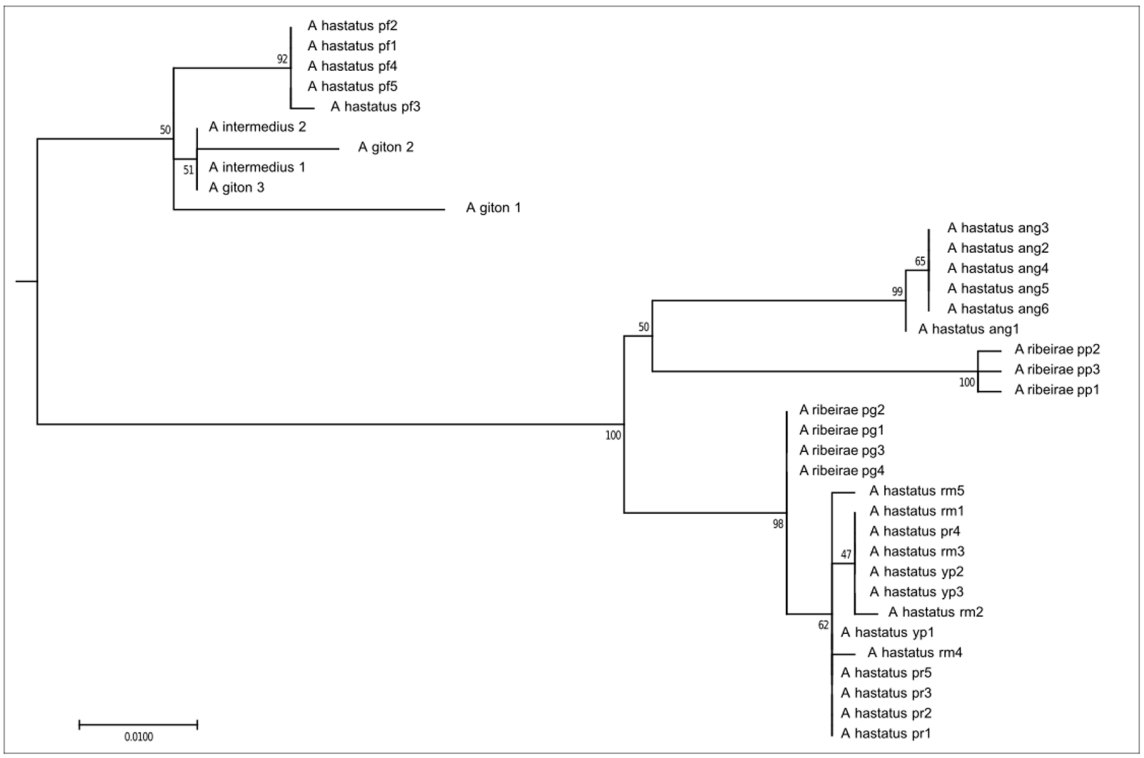


FIGURE 2 | Molecular Phylogenetic of the Clade 1 analysis by Maximum Likelihood Method. Parameters described in **Figure 1**. Clade 1 is formed by species distributed only in coastal drainages as *A. hastatus*, *A. ribeirae*, *A. intermedius*, and *A. giton*.

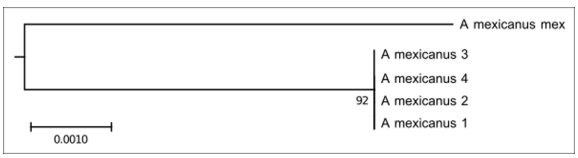


FIGURE 3 | Molecular Phylogenetic of the Clade 2 analysis by Maximum Likelihood Method. Parameters described in **Figure 1**. Clade 2 is formed by *A. mexicanus*.

identified by barcoding. The low genetic distances observed in the mitochondrial analyses among species of the *Astyanax* genus are often explained by the rapid divergence between them (Ornelas-García et al., 2008; Carvalho et al., 2011; Rossini et al., 2016). Therefore, multidisciplinary approaches may be more effective for reconstructing phylogenies within this genus. In lower systematic levels, molecular data associated with the chromosomal features were particularly effective to understand evolutionary patterns in *A.* aff. *bimaculatus* (Kavalco et al., 2011) and *A. fasciatus* (Pansonato-Alves et al., 2013; Kavalco et al., 2016).

Overlapping cytogenetic onto molecular phylograms offers an insight on large-scale chromosome evolutionary *Astyanax*. This analysis allows us to point out as the main specific chromosomal markers to be explored: the distribution patterns of 5S rDNA and As51 repetitive DNA, as they show signatures within the genus *Astyanax*. On the other hand, other characters such as location of the Nuclear Organizing Regions and C-banding shows patterns seem more informative at population level, as demonstrated by Jacobina et al. (2011) in *Hoplias malabaricus* (Erythrinidae). Although other markers have been used sporadically in chromosome studies of the genus *Astyanax* (Barbosa et al., 2015, 2017; Silva et al., 2015; Piscor and Parisi-Maltempi, 2016), their informative value will require intensive sampling in an array of species.

The genus *Astyanax* is currently *incertae sedis* in the Characidae family, and the evolutionary relations within the genus, as well as its phylogenetic relation to other genera of the family are quite challenging both from the morphological point of view (Mirande, 2009, 2010), and from the molecular systematic (Oliveira et al., 2011; Rossini et al., 2016). Despite the incomplete database, others genera related to *Astyanax* usually show more conservative cytogenetic characters, such as steady patterns of 50 chromosomes, multiple NORs and 5S ribosomal genes located in one or more pairs of chromosomes at the terminal region. These suite of characteristics have been reported in the genus *Oligosarcus* (Kavalco et al., 2005; Hattori et al., 2007; Barros et al., 2015), *Deuterodon* (Mendes et al., 2011; Coutinho-Sanches and Dergam, 2015), *Hollandichthys* and *Ctenobrycon* (Carvalho et al., 2002), plus the absence of the homology with the As51 repetitive DNA (Kavalco et al., 2005), suggesting that these characters are close to the evolutionary origins of *Astyanax*, being symplesiomorphies in clade 1.

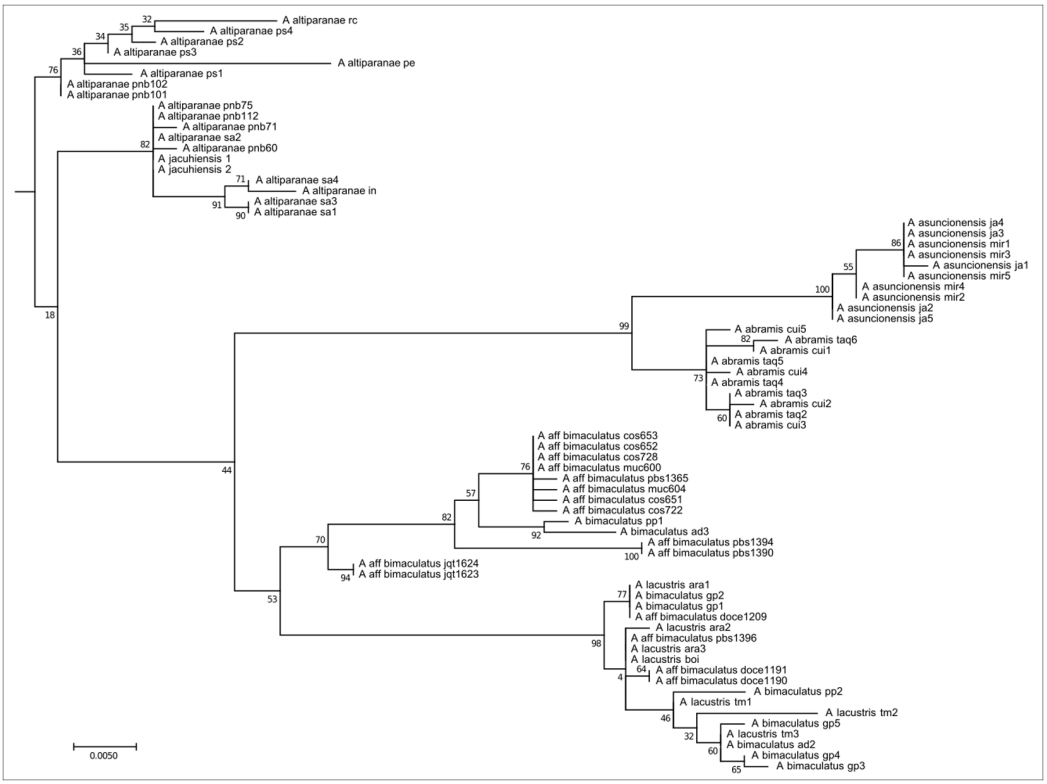


FIGURE 4 | Molecular Phylogenetic of the Clade 3 analysis by Maximum Likelihood Method. Parameters described in **Figure 1**. Clade 3 is formed by species with humeral oval spot like the *A. bimaculatus* complex and the related species.

Clade 1

The clade 1 obtained by our molecular analyses corresponds to clade 5 of Rossini et al. (2016). In the specimens analyzed we can observe the plesiomorphic chromosome characteristics (as a karyotype with 2n = 50 chromosomes and a low FN) and with good structure in *A. ribeirae* (Kavalco et al., 2010), *A. intermedius*, and *A. giton* (Kavalco and Moreira-Filho, 2003; Kavalco et al., 2004). In *A. hastatus*, different cytotypes with 2n = 50 chromosomes have been observed, constituting yet another species complex within the genus *Astyanax* (Kavalco et al., 2009), as corroborated by the present study (**Figure 2**). Also, a combination of molecular and chromosomal data suggests that the current *A. hastatus* may encompass more than one OTU (operational taxonomic unit) and so, more than one ESU (evolutionarily significant unit). The *Astyanax* species distributed in the coastal basins were added to the Probolodini group with species from the genus *Probolodus*, *Deuterodon*, and *Myxiops* and *Hyphessobrycon luetkenii* (Silva, 2017). These other species are less studied by cytogenetic methods, but until now, the chromosome characteristics seems to be shared (Mendes et al., 2011; Coutinho-Sanches and Dergam, 2015), bringing new evidence supporting the group. Unfortunately, As51 repetitive DNA in these coastal distribution species has not been confirmed so far (Kavalco et al., 2007, 2009).

Smaller variations in the number of 5S rDNA sites can be observed among these species, although they are always located in the proximal or distal region of acrocentric chromosomes, reaching up to 10 markings in *A. intermedius* and *A. giton*, which differ by a pericentric inversion (Kavalco et al., 2004). The hypothesis that these characteristics are a symplesiomorphy can be corroborated in independent chromosome data, such as those obtained for *A. taeniatus* (Cunha et al., 2016), which is also included in the same clade 5 by Rossini et al. (2016). The evolutionary dynamics of this gene are related not only to variations in non-transcribed spacers, but also to syntenia with long and short interspersed nuclear elements, non-long terminal repeat retrotransposons, U-snRNA families, and microsatellite polymorphisms (Rebordinos et al., 2013). According to these authors, polymorphisms in non-transcribed regions are observed in fish. Polymorphisms in transcribed regions do not appear to interfere with the cellular activity of 5S rDNA, and the molecular diversity of the 5S rDNA gene families is greater than the chromosome diversity (Rebordinos et al., 2013).

Clade 2

At the root of clades 2, 3, and 4, it is possible to hypothesize two main chromosomal events: the occurrence one of the 5S rDNA sites to the proximal position of a specific submetacentric chromosome pair, which has been considered as a marker (Almeida-Toledo et al., 2002) and the presence of As51 repetitive DNA in the *Astyanax* genome.

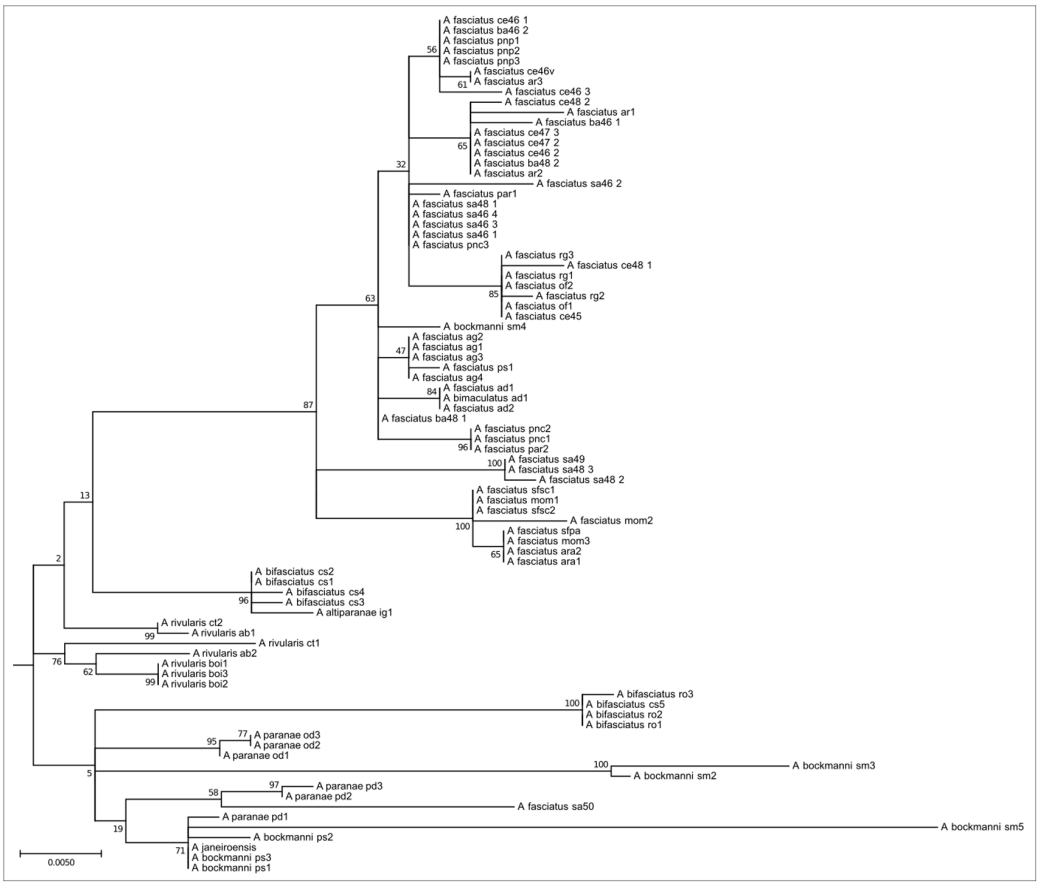


FIGURE 5 | Molecular Phylogenetic of the Clade 4 analysis by Maximum Likelihood Method. Parameters described in **Figure 1**. Clade 4 is formed by red-tailed species and their related group as *A. fasciatus*, the *A. scabripinnis* group, *A. bockmanni*, and *A. bifasciatus*. This group is the most chromosome diverse in the genus and shows little mitochondrial divergence.

The 5S rDNA in clade 2 appears on a pair of two-arm chromosomes in the proximal region of the centromere (Kavalco and Almeida-Toledo, 2007). This marker may have arisen by pericentric inversion from an acrocentric chromosome bearing the 5S rDNA site, which should represent the most basal character state.

In addition to this site, *A. mexicanus* also presents four more 5S rDNA sites on acrocentric chromosomes, being distal markings on one pair and proximal markings on another (Kavalco and Almeida-Toledo, 2007). The isolation of the group represented by the *A. mexicanus* is congruent with some molecular data that relate ichthyofauna invasions in Central America with the genesis of the Panamá Isthmus (around 3.3 Mya; Ornelas-García et al., 2008), as well as with clustering with DNA barcoding (Rossini et al., 2016). Unfortunately, cytogenetic data for the species of this group are scarce, but it would not be surprising to have independent autapomorphies of the karyotype evolution of the cisandine species of *Astyanax*.

As51 satellite DNA is partly repetitive tandem DNA and holds similarities with transposable elements that were isolated from a population of *A. scabripinnis* carrier of chromosome B (Mestriner et al., 2000). In this population, this DNA was located mainly in the B chromosome, besides two to four sites in the distal region of acrocentric chromosomes. Considering the possible origin from transposable elements, it is appropriate to assume that their distribution was initially restricted to a few sites and subsequently spread in the genome by intrinsic mechanisms of the repetitive sequence amplification. In fact, the individuals of *A. mexicanus* that form clade 2 in the present work have few sites that carry this satellite DNA, and that it is shown more diffusely than in other species, which suggests a smaller number of copies or even only partial homology (Kavalco and Almeida-Toledo, 2007). This divergence might be related to independent evolution of *A. mexicanus* relative to the *A. scabripinnis* strain that donated this probe.

Clade 3

The increase of As51 satellite DNA sites characterizes the common ancestral root of clades 3 and 4, in contrast to the loss of 5S rDNA sites on acrocentric chromosomes observed in members of Clade 3. Clade 3 is composed of the species complex "*A. bimaculatus*," represented herein by *A. altiparanae*, *A.* aff. *bimaculatus*, *A. lacustris*, *A. asuncionensis*, and *A. abramis*. Specimens from this clade present 2n = 50 chromosomes,

with a higher FN, multiple Nucleolar Organizing Regions, and consequently multiple sites of 18S rDNA, according to the samples analyzed (Kavalco et al., 2011) and others independent studies (Almeida-Toledo et al., 2002; Fernandes and Martins-Santos, 2004, 2006; Peres et al., 2008; Hashimoto et al., 2011; Giongo et al., 2013).

Clade 3 is characterized by a synapomorphy: only one pair of chromosomes carrying a rDNA 5S site on its pericentromeric region (Kavalco et al., 2011) and in the literature (Almeida-Toledo et al., 2002; Fernandes and Martins-Santos, 2006; Peres et al., 2008, 2012; Hashimoto et al., 2011; Paiz et al., 2015). One exception is *A. abramis*, which presents in addition to the site of the chromosomal marker pair, a pair of extra-acrocentric chromosomes with proximal marking (Paiz et al., 2015; Piscor et al., 2015). This differentiated pattern may be an autapomorphy of *A. abramis*, whose molecular distinctiveness within the group was also evident in our analysis (**Figure 5**). Additionally, the pattern of occurrence of a 5S rDNA carrier pair can be seen in other species related to the *A. bimaculatus* group, as in *A. janeiroensis* (Vicari et al., 2008), *A. goyacensis* (Santos et al., 2013), and *A. elachylepis* (Santos et al., 2016).

In turn, there is little data on the distribution of As51 satellite DNA in this group, besides previous reports (Kavalco et al., 2011). In these species, small sites are observed, suggesting relatively lower number of *in tandem* copies and more restricted distribution through the karyotype, except for *A. bimaculatus* from coastal basins, in which As51 satDNA is absent (Kavalco et al., 2011). On the other hand, in *Astyanax janeiroensis*, a considerable number of very conspicuous sites are observed (Vicari et al., 2008; Kantek et al., 2009b). Although the Catalog of Fishes (Eschmeyer et al., 2017) characterizes the distribution of *A. janeiroensis* as being "in Brazil", Melo (2001) suggests that its distribution is restricted to the basins of the Paraíba do Sul and other coastal drainages. These discrepant and independent occurrences in *A. janeiroensis* and *A. bimaculatus* can be a result of the historical biogeography of coastal drainages, affected by stream capture from continental basins and complex dispersal due to marine regressions and transgressions (Pereira et al., 2013). Fish faunas in these regions are characterized by a high degree of endemism and low species richness (Albert and Reis, 2011).

Clade 4

Clade 4 comprises a group of species with the highest chromosomal variability in the genus, with 2n ranging from 46 to 50 chromosomes and including the species complexes *A. scabripinnis* (encompassing *A. rivularis* and *A. paranae*) and *A. fasciatus*, as well the species *A. bockmanni* and *A. bifasciatus* (formerly referred to as *Astyanax* sp. B). Equally variable among the species of this clade is the number of NOR/rDNA 18S sites and the constitutive heterochromatin distribution patterns. The molecular divergence among these species is apparently quite recent, with low genetic distance indices as observed in the clade 1 and by Rossini et al. (2016) using the *COI* gene. Despite some level of structuring, low bootstrap indexes preclude further hypotheses on the genetic diversification of the group. There is no detailed chromosome information on the other species analyzed by Rossini et al. (2016) belonging to this clade, except for *A. parahybae* (Kavalco and Moreira-Filho, 2003; Kavalco et al., 2004) presenting 2n = 48 chromosomes, and *A. schubarti* (Morelli et al., 1983) presenting 2n = 36 chromosomes. Unfortunately, we were unable to obtain the mitochondrial DNA sequence from *A. parahybae* in the present work.

Among the specimens analyzed, this clade presents some species with a conserved chromosome number, always with 2n = 50 chromosomes, such as *A. bockmanni* and *A. bifasciatus* (Fazoli et al., 2003; Kavalco et al., 2009; Hashimoto and Porto-Foresti, 2010). On the other hand, the others species present high levels of numerical chromosome variation.

This diploid number (2n = 50) was also observed in the *A. paranae* specimens from the Paranaíba and some *A. rivularis* specimens from the São Francisco river basin, although other specimens of the Paranaíba river basin also had 2n = 46 chromosomes (data not shown, in preparation). Both species belong to the historical *A. scabripinnis* species complex, characterized by broad sympatric and allopatric karyotype variation (Moreira-Filho and Bertollo, 1991, among others).

Finally, a diploid number that varies from 2n = 46 or 2n = 48 was observed among the analyzed specimens of *A. fasciatus* (Pazza et al., 2006; Kavalco et al., 2016) that was also observed in other populations (Artoni et al., 2006; Medrado et al., 2008; Pansonato-Alves et al., 2013). This variation is considered common in the species, although 2n = 50 chromosomes have already been reported (Artoni et al., 2006).

In relation to rDNA 5S, most of the species/populations already analyzed have the same following pattern: one site on the marker chromosome plus a pair of acrocentric chromosomes with a rDNA 5S site in the proximal region; this was also observed in the specimens analyzed in the present work (Kavalco et al., 2004, 2007, 2009, 2016; Pazza et al., 2006) and in the literature available for *A. bockmanni*, *A. fasciatus*, and *A. parahybae* (Almeida-Toledo et al., 2002; Hashimoto et al., 2011; Silva et al., 2013; Daniel et al., 2015). Despite this relatively conserved pattern, there are two autapomorphies in the 5S rDNA phenotype in *A. fasciatus*. Medrado et al. (2015) reported the occurrence of a small variation in the distribution of 5S rDNA sites from a population of *Astyanax* aff. *fasciatus* where the marker site of the metacentric chromosome is absent, an evident populational autapomorphy. We also considered the additional occurrence of two 5S rDNA sites of *A.* aff. *fasciatus* from the Paraíba do Sul river basin an autapomorphy. In total, six 5S rDNA-bearing chromosomes are detected, represented always by the marker chromosome pair plus two st/a chromosomes pairs in this sample (Kavalco et al., 2016).

In relation to the *A. scabripinnis* group, the most studied from the cytogenetic point of view (Pazza and Kavalco, 2007), several populations of the Coastal, São Francisco, and Upper Paraná rivers show the same mentioned pattern (Mantovani et al., 2005; Fernandes and Martins-Santos, 2006; Peres et al., 2008; Vicari et al., 2008). However, other studies have already demonstrated, in addition to these standard sites, more distal or proximal sites in other pairs of acrocentric chromosomes in populations of coastal rivers and the Upper Paraná basin (Kavalco et al., 2004).

Although absent in our analyses, *A. schubarti* is a species that has available chromosome data and is found in clade 1 of Rossini et al. (2016), the clade analogous to clade 4 of this work. Its low chromosomal number (2n = 36) and high FN suggest the occurrence of Robertsonian events of chromosome fusion at its origin (Morelli et al., 1983). In fact, this species shows the 5S rDNA sites located in the proximal region of two pairs of metacentric chromosomes (Almeida-Toledo et al., 2002). Interestingly, *A. currentinus*, a species recently described (Mirande et al., 2015), presents 2n = 36 chromosomes and the same distribution pattern of *A. schubarti* 5S rDNA, with an additional odd site (Paiz et al., 2015). The authors suggest that these species may belong to the same morphological group and are phylogenetically related. Unfortunately, there are no available phylogenetic data to support this hypothesis.

Among the specimens used in the present study, populations of the *A. fasciatus* and *A. scabripinnis* species complexes presented sites with homology with the As51 satellite DNA (Kavalco et al., 2007, 2013; Pazza et al., 2008). On the other hand, this probe showed no homology in the chromosomes of *A. bockmanni* (Kavalco et al., 2009). Among the species included in this group and in the clade 1 of Rossini et al. (2016), in only one population of *A. rivularis* (cited as *A. scabripinnis*) this repetitive sequence is absent (Abel et al., 2006). On the other hand, other species belonging to clade 1 of Rossini et al. (2016) have already presented homology with As51 satellite DNA, such as *A. parahybae* (Kavalco et al., 2007) and *A. serratus* (quoted as *Astyanax sp.*; Kantek et al., 2009a). Among these species, the available data are mainly concentrated in the group *A. scabripinnis*/*paranae*/*rivularis* (Abel et al., 2006; Kantek et al., 2009b; Barbosa et al., 2015, 2017), and in *A. fasciatus* (Abel et al., 2006; Kantek et al., 2009b; Medrado et al., 2015). The distribution of this satellite DNA in *A. fasciatus* seems to follow a biogeographic pattern, with an increase in the number of sites in drainage populations in the interior of the continent, and a decrease in coastal populations (Kavalco et al., 2013; Medrado et al., 2015). This pattern does not appear to be the same in the *A. scabripinnis* group and its cryptic species.

The absence of a clear biogeographic pattern of satellite DNA distribution in the *A. scabripinnis* populations should be related to the fact that these fish inhabit isolated headwaters. In addition to facilitating the occurrence of vicariance processes, this ecology keeps low levels of gene flow among, with consequent reduction of population sizes and possible strong effect of genetic drift, resulting in random of As51 satDNA. This could easily generate the biogeographic gaps that occur in the genomic distribution and the existence of populations in which this repetitive DNA is absent, even being a basal characteristic for the group. In fact, it is not yet proven if there is any adaptive role for this DNA, even though we have evidence that it is present in a larger amount in the species with the greatest chromosome diversity, with extensive polymorphisms such as *A. fasciatus*. The occurrence of the As51 satDNA predominantly in supernumerary chromosomes (Mestriner et al., 2000) may imply exogenous origin.

The data of the literature obtained in the present study still cannot answer definitively the questions about the origin and evolution of these sequences within the lineages of *Astyanax*, only to indicate trends within specific clades. Can As51 satDNA promote the chromosomal rearrangements that generate the diversity in diploid numbers and karyotype formulas observed in *Astyanax*; or can the genome of species with chromosomal plasticity, such as those of the genus, facilitate the dispersion of repetitive DNAs? To answer this question, we need more information about the role of this DNA in the karyotype and organismic evolution of the group. Fortunately, the genus *Astyanax*, due to the great concentration of studies, is an excellent model to answer these questions of comparative genomics.

AUTHOR CONTRIBUTIONS

RP and KK designed the research and performed the analyses. All the authors collected data and contributed to the writing and review of the manuscript.

FUNDING

This study was financed by the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG; APQ-01247-08, APQ-00756-09, and APQ-01510-11).

ACKNOWLEDGMENTS

The authors thank the Manejo e Conservação de Ecossistemas Naturais e Agrários (MCENA) post-graduate program, Federal University of Viçosa, Campus Florestal.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00131/full#supplementary-material

REFERENCES

Abel, L. D. S., Mantovani, M., and Moreira-Filho, O. (2006). Chromosomal distribution of the As-51 satellite DNA in two species complexes of the genus *Astyanax* (Pisces, Characidae). *Genet. Mol. Biol.* 29, 448–452. doi: 10.1590/S1415-47572006000300008

Albert, J. S., and Reis, R. E. (eds). (2011). *Historical Biogeography of Neotropical Freshwater Fishes*. Los Angeles: University of California Press, 408.

Almeida-Toledo, L. F., Ozouf-Costaz, C., Foresti, F., Bonillo, C., Porto-Foresti, F., and Daniel-Silva, M. F. Z. (2002). Conservation of the 5S-bearing chromosome pair and co-localization with major rDNA clusters in five species of *Astyanax* (Pisces, Characidae). *Cytogenet. Genome Res.* 97, 229–233. doi: 10.1159/000066609

Artoni, R. F., Shibatta, O. A., Gross, M. C., Schneider, C. H., Almeida, M. C., Vicari, M. R., et al. (2006). *Astyanax* aff. *fasciatus* Cuvier, 1819 (Teleostei; Characidae): evidences of a species complex in the Upper rio Tibagi basin (Paraná,

Brazil). *Neotrop. Ichtyol.* 4, 197–202. doi: 10.1590/S1679-62252006000200005

Barbosa, P., Leal, E. V., Silva, M., Almeida, M. C., Moreira-Filho, O., and Artoni, R. F. (2017). Variability and evolutionary implications of repetitive DNA dynamics in genome of *Astyanax scabripinnis* (Teleostei. Characidae). *Comp. Cytogenet.* 11, 143–162. doi: 10.3897/CompCytogen.v11i1.11149

Barbosa, P., Oliveira, L. A., Pucci, M. B., Santos, M. H., Moreira-Filho, O., Vicari, M. R., et al. (2015). Identification and chromosome mapping of repetitive elements in the *Astyanax scabripinnis* (Teleostei: Characidae) species complex. *Genetica* 143, 55–62. doi: 10.1007/s10709-014-9813-2

Barros, L. C., Santos, U., Cioffi, M. B., and Dergam, J. A. (2015). Evolutionary divergence among *Oligosarcus* spp. (Ostariophysi, Characidae) from the São Francisco and Doce river basins: *Oligosarcus solitarius* Menezes, 1987 shows the highest rates of chromosomal evolution in the Neotropical region. *Zebrafish* 12, 102–110. doi: 10.1089/zeb.2014.1030

Carvalho, D. C., Oliveira, D. A. A., Pompeu, P. S., Leal, C. G., Oliveira, C., and Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River basin. *Mitochondrial DNA* 22, 80–86. doi: 10.3109/19401736.2011.588214

Carvalho, M. L., Oliveira, C., and Foresti, F. (2002). Cytogenetic analysis of five species of the subfamily Tetragonopterinae (Teleostei, Characiformes, Characidae). *Caryologia* 55, 181–188. doi: 10.1080/00087114.2002.10589275

Coutinho-Sanches, N., and Dergam, J. A. (2015). Cytogenetic and molecular data suggest *Deuterodon pedri* Eigenmann, 1907 (Teleostei: Characidae) is a member of an ancient coastal group. *Zebrafish* 12, 357–365. doi: 10.1089/zeb.2014.1068

Cunha, M. S., Reis, V. J. C., and Dergam, J. A. (2016). Closely related syntopic cytotypes of *Astyanax taeniatus* (Jenyns, 1842) from the Upper Piranga River, Upper Doce basin in Southeastern Brazil. *Zebrafish* 13, 112–117. doi: 10.1089/zeb.2015.1163

Daniel, S. N., Penitente, M., Hashimoto, D. T., Silva, D. M. Z. A., Foresti, F., and Porto-Foresti, F. (2015). New insights into karyotypic relationships among populations of *Astyanax bockmanni* (Teleostei, Characiformes) of different watersheds. *Zebrafish* 12, 181–188. doi: 10.1089/zeb.2014.1053

Eschmeyer, W. N., Fricke, R., and van der Laan, R. (2017). *Catalog of Fishes: Genera, Species, References*. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp [accessed October 10, 2017].

Fazoli, L. C., Silva, V. A. B., Portela-Castro, A. L., and Júlio, H. F. Jr. (2003). Chromosome characterization of *Astyanax* sp B (Characidae, Tetragonopterinae), and endemic species of the Iguaçu river, Paraná, Brazil. *Cytologia* 68, 389–394. doi: 10.1508/cytologia.68.389

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791. doi: 10.1111/j.1558-5646.1985.tb00420.x

Fernandes, C. A., and Martins-Santos, I. C. (2004). Cytogenetic studies in two populations of the *Astyanax altiparanae* (Pisces, Characiformes). *Hereditas* 141, 328–332. doi: 10.1111/j.1601-5223.2004.01832.x

Fernandes, C. A., and Martins-Santos, I. C. (2006). Chromosomal location of 5S and 18S rRNA genes in three sympatric cytotypes of *Astyanax scabripinnis* (Characiformes, Characidae) from the Ivaí river basin, state of Paraná, Brazil. *Caryologia* 59, 253–259. doi: 10.1080/00087114.2006.10797923

Giongo, P., Sampaio, W. M. S., Fernandes, A., Dergam, J. A., Kavalco, K. F., and Pazza, R. (2013). Diversidade cariotípica de *Astyanax asuncionensis* (Teleostei, Characiformes) na microbacia do rio Sepotuba (Alto Paraguai). *Evol. Cons. Biodivers.* 3, 53–58. doi: 10.7902/ecb.v3i2.9

Hashimoto, D. T., Ferguson-Smith, M. A., Rens, W., Foresti, F., and Porto-Foresti, F. (2011). Chromosome mapping of H1 histone and 5S rRNA gene clusters in three species of *Astyanax* (Teleostei, Characiformes). *Cytogenet. Genome Res.* 134, 64–71. doi: 10.1159/000323512

Hashimoto, D. T., and Porto-Foresti, F. (2010). Chromosome polymorphism of heterochromatin and nucleolar regions in two populations of the fish *Astyanax bockmanni* (Teleostei: Characiformes). *Neotrop. Ichthyol.* 8, 861–866. doi: 10.1590/S1679-62252010000400016

Hattori, R. S., Daniel-Silva, M. F. Z., and Almeida-Toledo, L. F. (2007). Karyotype characterization and gene mapping of 5S and 18S rDNA in three species of *Oligosarcus* (Teleostei: Characidae). *Caryologia* 60, 372–378. doi: 10.1080/00087114.2007.10797961

Havelka, M., Bytyutskyy, D., Symonová, R., Ráb, P., and Flajshans, M. (2016). The second highest chromosome count among vertebrates is observed in cultured sturgeon and is associated with genome plasticity. *Genet. Sel. Evol.* 48:12. doi: 10.1186/s12711-016-0194-0

Hebert, P. D. N., and Gregory, T. R. (2005). The promise of DNA barcoding for taxonomy. *Mol. Syst. Biol.* 54, 852–859. doi: 10.1080/10635150500354886

Jacobina, U. P., Paiva, E., and Dergam, J. A. (2011). Pleistocene karyotypic divergence in *Hoplias malabaricus* (Bloch, 1794) (Teleostei: Erythrinidae) populations in southeastern Brazil. *Neotrop. Ichthyol.* 9, 325–333. doi: 10.1590/S1679-62252011005000023

Jónsson, H., Schubert, M., Seguin-Orlando, A., Ginolhac, A., Petersen, L., Fumagalli, M., et al. (2014). Speciation with gene flow in equids despite extensive chromosomal plasticity. *Porc. Natl. Acad. Sci. U.S.A.* 111, 18655–18660. doi: 10.1073/pnas.1412627111

Kantek, D. L. Z., Cestari, M. M., Bertollo, L. A. C., and Moreira-Filho, O. (2009a). Characterization of the constitutive heterochromatin of *Astyanax* sp. D (Characiformes: Characidae) from the Upper Iguaçu river (PR). *Folia Biol.* 57, 1–2.

Kantek, D. L. Z., Vicari, M. R., Peres, W. Z. M., Cestari, M. M., Artoni, R. F., Bertollo, L. A. C., et al. (2009b). Chromosomal location and distribution of As51 satellite DNA in five species of the genus *Astyanax* (Teleostei, Characidae, *Incertae sedis*). *J. Fish Biol.* 75, 408–421. doi: 10.1111/j.1095-8649.2009.02333.x

Kavalco, K. F., and Almeida-Toledo, L. F. (2007). Molecular cytogenetics of blind mexican tetra and comments on the karyotypic characteristics of genus *Astyanax* (Teleostei, Characidae). *Zebrafish* 4, 103–111. doi: 10.1089/zeb.2007.0504

Kavalco, K. F., and Moreira-Filho, O. (2003). Cytogenetical analyses in four species of the genus *Astyanax* (Pisces, Characidae) from Paraíba do Sul river basin. *Caryologia* 56, 453–461. doi: 10.1080/00087114.2003.10589358

Kavalco, K. F., Pazza, R., and Almeida-Toledo, L. F. (2009). *Astyanax bockmanni* Vari and Castro, 2007: an ambiguous karyotype in the *Astyanax* genus. *Genetica* 136, 135–139. doi: 10.1007/s10709-008-9325-z

Kavalco, K. F., Pazza, R., and Almeida-Toledo, L. F. (2010). Molecular cytogenetics of *Astyanax ribeirae* (Teleostei, Characidae), an endemic characin of the Atlantic rainforest. *Nucleus* 53, 51–54. doi: 10.1007/s13237-010-0006-9

Kavalco, K. F., Pazza, R., Brandão, K. d. O., and Almeida-Toledo, L. F. (2013). Biogeographic patterns in the chromosomal distribution of a satellite DNA in the banded tetra *Astyanax fasciatus* (Teleostei: Characiformes). *Org. Divers. Evol.* 13, 67–76. doi: 10.1007/s13127-012-0100-8

Kavalco, K. F., Pazza, R., Brandão, K. d. O., Garcia, C., and Almeida-Toledo, L. F. (2011). Comparative cytogenetics and molecular phylogeography in the group *Astyanax altiparanae – Astyanax* aff. *bimaculatus* (Teleostei, Characidae). *Cytogenet. Genome Res.* 134, 108–119. doi: 10.1159/000325539

Kavalco, K. F., Pazza, R., Bertollo, L. A. C., and Moreira-Filho, O. (2004). Gene mapping of 5S rDNA sites in eight fish species from the Paraíba do Sul river basin, Brazil. *Cytogenet. Genome Res.* 106, 107–110. doi: 10.1159/000078567

Kavalco, K. F., Pazza, R., Bertollo, L. A. C., and Moreira-Filho, O. (2005). Molecular cytogenetics of *Oligosarcus hepsetus* (Teleostei, Characiformes) from two Brazilian locations. *Genetica* 124, 85–91. doi: 10.1007/s10709-005-0176-6

Kavalco, K. F., Pazza, R., Bertollo, L. A. C., and Moreira-Filho, O. (2007). Satellite DNA sites in four species of the genus *Astyanax* (Teleostei, Characiformes). *Genet. Mol. Biol.* 30, 529–535. doi: 10.1590/S1415-47572007000400005

Kavalco, K. F., Pazza, R., Brandão, K. O., Garcia, C., Bertollo, L. A. C., and Almeida-Toledo, L. F. (2016). Chromosomal diversification higher than molecular variation in *Astyanax* aff. *fasciatus* (Teleostei, Characidae). *Zebrafish* 13, 345–353. doi: 10.1089/zeb.2016.1272

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA 7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Mantovani, M., Abel, L. D. S., and Moreira-Filho, O. (2005). Conserved 5S and variable 45S rDNA chromosomal localisation revealed by FISH in *Astyanax scabripinnis* (Pisces, Characidae). *Genetica* 123, 211–216. doi: 10.1007/s10709-004-2281-3

Medrado, A. S., Affonso, P. R. A. M., Carneiro, P. L. S., Vicari, M. R., Artoni, R. F., and Costa, M. A. (2015). Allopatric divergence in *Astyanax* aff. *fasciatus Cuvier,* 1819 (Characidae, *Incertae sedis*) inferred from DNA mapping and chromosomes. *Zool. Anz.* 257, 119–129. doi: 10.4238/2011.November.22.5

Medrado, A. S., Figueiredo, A. V. A., Waldschmidt, A. M., Affonso, P. R. A. M., and Carneiro, P. L. S. (2008). Cytogenetic and morphological diversity in populations of *Astyanax fasciatus* (Teleostei, Characidae) from Brazilian northeastern river basins. *Genet. Mol. Biol.* 31, 208–214. doi: 10.1590/S1415-47572008000200007

Mello, R., Maniglia, T. C., Prioli, S. M. A. P., and Prioli, A. J. (2015). Genetic and biogeographical relationships among species of *Astyanax* (Teleostei, Characidae) in Brazilian river basins. *Genet. Mol. Res.* 14, 15356–15364. doi: 10.4238/2015.November.30.13

Melo, F. A. G. (2001). Revisão taxonômica das espécies do gênero *Astyanax* Baird e Girard, 1854 (Teleostei: Characiformes: Characidae) da região da Serra dos Órgãos. *Arq. Mus. Nac.* 59, 1–46.

Mendes, M. M., Rosa, R., Giuliano-Caetano, L., and Dias, A. L. (2011). Karyotype diversity of four species of the *incertae sedis* group (Characidae) from different hydrographic basins: analysis of AgNORs, CMA3 and 18S rDNA. *Genet. Mol. Res.* 10, 3596–3608. doi: 10.4238/2011.November.22.5

Mestriner, C. A., Galetti, P. M. Jr., Valentini, S. R., Ruiz, I. R. G., Abel, L. D. S., Moreira-Filho, O., et al. (2000). Structural and functional evidence that a B chromosome in the characid fish *Astyanax scabripinnis* is an isochromosome. *Heredity* 85, 1–9. doi: 10.1046/j.1365-2540.2000.00702.x

Mirande, J. M. (2009). Weighted parsimony phylogeny of the family Characidae (Teleostei: Characiformes). *Cladistics* 25, 574–613. doi: 10.1111/j.1096-0031.2009.00262.x

Mirande, J. M. (2010). Phylogeny of the family Characidae (Teleostei: Characiformes): from characters to taxonomy. *Neotrop. Ichthyol.* 8, 385–568. doi: 10.1590/S1679-62252010000300001

Mirande, J. M., Azpelicueta, M. M., and Aguilera, G. (2015). Redescription of *Astyanax correntinus* (Holmberg, 1891) (Teleostei: Characiformes: Characidae), more than one hundred years after original description. *Zool. Abh.* 55, 9–15.

Moreira-Filho, O., and Bertollo, L. A. C. (1991). *Astyanax scabripinnis* (Pisces, Characidae): a species complex. *Braz. J. Genet.* 14, 331–357.

Morelli, S., Bertollo, L. A. C., Foresti, F., Moreira-Filho, O., and Toledo Filho, S. A. (1983). Cytogenetic considerations on the genus *Astyanax* (Pisces, Characidae). I. Karyotypic variability. *Caryologia* 36, 235–244. doi: 10.1080/00087114.1983.10797664

Navarro, A., and Barton, N. H. (2003). Accumulation postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution* 57, 447–459. doi: 10.1111/j.0014-3820.2003.tb01537.x

Oliveira, O., Avelino, G. S., Abe, K. T., Mariguela, T. C., Benine, R. C., Ortí, G., et al. (2011). Phylogenetic relationships within the speciose family Characidae (Teleostei: Ostariophysi: Characiformes) based on multilocus analysis and extensive ingroup sampling. *BMC Evol. Biol.* 11:275. doi: 10.1186/1471-2148-11-275

Ornelas-García, C. P., Domínguez-Domínguez, O., and Doadrio, I. (2008). Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. *BMC Evol. Biol.* 8:340. doi: 10.1186/1471-2148-8-340

Paiz, L. M., Baumgärtner, L., Graça, W. J., and Margarido, V. P. (2015). Basic cytogenetics and physical mapping of ribosomal genes in four *Astyanax* species (Characiformes, Characidae) collected in Middle Paraná river, Iguassu National Park: considerations on taxonomy and systematics of the genus. *Comp. Cytogenet.* 9, 51–65. doi: 10.3897/CompCytogen.v9i1.9002

Pansonato-Alves, J. C., Hilsdorf, A. W. S., Utsunomia, R., Silva, D. M. Z. A., Oliveira, C., and Foresti, F. (2013). Chromosomal mapping of repetitive DNA and Cytochrome c Oxidase I sequence analysis reveal differentiation among sympatric samples of *Astyanax fasciatus* (Characiformes, Characidae). *Cytogenet. Genome Res.* 141, 133–142. doi: 10.1159/000354885

Pazza, R., and Kavalco, K. F. (2007). Chromosomal evolution in the Neotropical characin *Astyanax* (Teleostei, Characidae). *Nucleus* 50, 519–543.

Pazza, R., Kavalco, K. F., and Bertollo, L. A. C. (2006). Chromosome polymorphism in *Astyanax fasciatus* (Teleostei, Characidae). 1. Karyotype analysis, Ag-NORs and mapping of the 18S and 5S ribosomal genes in sympatric karyotypes and their possible hybrid forms. *Cytogenet. Genome Res.* 112, 313–319. doi: 10.1159/000089886

Pazza, R., Kavalco, K. F., and Bertollo, L. A. C. (2008). Chromosome polymorphism in *Astyanax fasciatus* (Teleostei, Characidae). 2 – Chromosomal location of a satellite DNA. *Cytogenet. Genome Res.* 122, 61–66. doi: 10.1159/000151317

Pereira, L. H. G., Hanner, R., Foresti, F., and Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genet.* 14:20. doi: 10.1186/1471-2156-14-20

Peres, W. A. M., Bertollo, L. A. C., Buckup, P. A., Blanco, D. R., Kantek, D. L. Z., and Moreira-Filho, O. (2012). Invasion, dispersion and hybridization of fish associated to river transposition: karyotypic evidence in "*bimaculatus* group" (Characiformes: Characidae). *Rev. Fish Biol. Fish.* 22, 519–526. doi: 10.1007/s11160-011-9246-2

Peres, W. A. M., Bertollo, L. A. C., and Moreira-Filho, O. (2008). Physical mapping of the 18S and 5S ribosomal genes in nine Characidae species (Teleostei, Characiformes). *Genet. Mol. Biol.* 31, 222–226. doi: 10.1590/S1415-47572008000200009

Piscor, D., Alves, A. L., and Parisi-Maltempi, P. P. (2015). Chromosomal microstructure diversity in three *Astyanax* (Characiformes, Characidae) species: comparative analysis of ghe chromosomal locations of the 18S and 5S rDNAs. *Zebrafish* 12, 81–90. doi: 10.1089/zeb.2014.1036

Piscor, D., and Parisi-Maltempi, P. P. (2016). Microsatellite organization in the B chromosome and A chromosome complement in *Astyanax* (Characiformes, Characidae) species. *Cytogenet. Genome Res.* 148, 45–51. doi: 10.1159/000444728

Rebordinos, L., Cross, I., and Merlo, A. (2013). High evolutionary dynamism in 5S rDNA of fish: state of the art. *Cytogenet. Genome Res.* 141, 103–113. doi: 10.1159/000354871

Rieseberg, L. H. (2001). Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* 16, 351–358. doi: 10.1016/S0169-5347(01)02187-5

Rossini, B. C., Oliveira, C. A. M., Melo, F. A. G., Bertaco, V. A., Astarloa, R. N. D., Rosso, J. J., et al. (2016). Highlighting *Astyanax* species diversity through DNA Barcoding. *PLoS One* 11:e0167203. doi: 10.1371/journal.pone.0167203

Santos, L. P., Castro, J. P., Francisco, C. M., Vicari, M. R., Almeida, M. C., Goll, L. G., et al. (2013). Cytogenetic analysis in the neotropical fish *Astyanax goyacensis* Eigenmann, 1908 (Characidae, *incertae sedis*): karyotype description and occurrence of B microchromosomes. *Mol. Cytogenet.* 6:48. doi: 10.1186/1755-8166-6-48

Santos, L. P., Francisco, C. M., Castro, J. P., Barbosa, P., Morelli, S., Oliveira, L. A., et al. (2016). Karyoevolutionary inferences in enigmatic taxon *Astyanax* (Teleostei: Characidae). *Zebrafish* 13, 578–583. doi: 10.1089/zeb.2016.1291

Sassi, F. M. C., Resende, S. V., Pazza, R., and Kavalco, K. F. (2018). "Natural Hybridization between chromosomal discrepant species and the role of hybrid speciation in the Genus *Astyanax*," in *An Essential Guide to Cytogenetics*, eds N. Norris and C. Miller (Hauppauge, NY: Nova Science Publisher).

Silva, D. M. Z. A., Pansonato-Alves, J. C., Utsonomia, R., Daniel, S. N., Hashimoto, D. T., Oliveira, C., et al. (2013). Chromosomal organization of repetitive DNA sequences in *Astyanax bockmanni* (Teleostei, Characiformes): dispersive location, association and co-localization in the genome. *Genetica* 141, 329–336. doi: 10.1007/s10709-013-9732-7

Silva, D. M. Z. A., Utsonomia, R., Pansonato-Alves, J. C., Oliveira, C., and Foresti, F. (2015). Chromosomal mapping of repetitive DNA sequences in five species of *Astyanax* (Characiformes, Characidae) reveals independent location of U1 and U2 snRNA sites and association of U1 snRNA and 5S rDNA. *Cytogenet. Genome Res.* 146, 144–152. doi: 10.1159/000438813

Silva, P. C. (2017). *Sistemática Integrativa - Diversidade e Relações de Deuterodon e Gêneros Afins*. Doctoral thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre, 188.

Sivasundar, A., Bermingham, E., and Ortí, G. (2001). Population structure and biogeography of migratory freshwater fishes (*Prochilodus*: Characiformes) in major South American Rivers. *Mol. Ecol.* 10, 407–417. doi: 10.1046/j.1365-294X.2001.01194.x

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Travenzoli, N. M., Silva, P. C., Santos, U., Zanuncio, J. C., Oliveira, C., and Dergam, J. A. (2015). Cytogenetic and molecular data demonstrate that the Bryconinae (Ostariophysi, Bryconidae) species from Southeastern Brazil form a phylogenetic and Phylogeographic unit. *PLoS One* 10:e0137843. doi: 10.1371/journal.pone.0137843

Vicari, M. R., Noleto, R. B., Artoni, R. F., Moreira-Filho, O., and Bertollo, L. A. C. (2008). Comparative cytogenetics among species of the *Astyanax scabripinnis* complex. Evolutionary and biogeographical inferences. *Genet. Mol. Biol.* 31, 173–179. doi: 10.3897/CompCytogen

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 27 April 2018
doi: 10.3389/fgene.2018.00138

Molecular Identification of Shark Meat From Local Markets in Southern Brazil Based on DNA Barcoding: Evidence for Mislabeling and Trade of Endangered Species

Fernanda Almerón-Souza[1†], Christian Sperb[2†], Carolina L. Castilho[2], Pedro I. C. C. Figueiredo[1], Leonardo T. Gonçalves[1], Rodrigo Machado[2], Larissa R. Oliveira[3], Victor H. Valiati[2†] and Nelson J. R. Fagundes[1*†]*

[1] Laboratório de Genética Médica e Evolução, Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, [2] Laboratório de Biologia Molecular, Centro de Ciências da Saúde, Universidade do Vale do Rio dos Sinos, São Leopoldo, Brazil, [3] Laboratório de Ecologia de Mamíferos, Centro de Ciências da Saúde, Universidade do Vale do Rio dos Sinos, São Leopoldo, Brazil

Elasmobranchs, the group of cartilaginous fishes that include sharks and rays, are especially vulnerable to overfishing due to low fecundity and late sexual maturation. A significant number of elasmobranch species are currently overexploited or threatened by fisheries activities. Additionally, several recent reports have indicated that there has been a reduction in regional elasmobranch population sizes. Brazil is an important player in elasmobranch fisheries and one of the largest importers of shark meat. However, carcasses entering the shark meat market have usually had their fins and head removed, which poses a challenge to reliable species identification based on the morphology of captured individuals. This is further complicated by the fact that the internal Brazilian market trades several different elasmobranch species under a common popular name: "cação." The use of such imprecise nomenclature, even among governmental agencies, is problematic for both controlling the negative effects of shark consumption and informing the consumer about the origins of the product. In this study, we used DNA barcoding (mtDNA, COI gene) to identify, at the species level, "cação" samples available in local markets from Southern Brazil. We collected 63 samples traded as "cação," which we found to correspond to 20 different species. These included two teleost species: *Xiphias gladius* ($n = 1$) and *Genidens barbus* ($n = 6$), and 18 species from seven elasmobranch orders (Carcharhiniformes, $n = 42$; Squaliformes, $n = 3$; Squatiniformes, $n = 2$; Rhinopristiformes, $n = 4$; Myliobatiformes, $n = 3$; Rajiformes, $n = 1$; and Torpediniformes, $n = 1$). The most common species in our sample were *Prionace glauca* ($n = 15$) and *Sphyrna lewini* ($n = 14$), while all other species were represented by four samples or less. Considering IUCN criteria, 47% of the elasmobranch species found are threatened at the global level,

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

Reviewed by:
Henrik R. Nilsson, University of Gothenburg, Sweden
Reinaldo A. De Brito, Universidade Federal de São Carlos, Brazil

***Correspondence:**
Victor H. Valiati
valiati@unisinos.br
Nelson J. R. Fagundes
nelson.fagundes@ufrgs.br

[†] These authors have contributed equally to this work.

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

Received: *01 November 2017*
Accepted: *03 April 2018*
Published: *27 April 2018*

Citation:
Almerón-Souza F, Sperb C, Castilho CL, Figueiredo PICC, Gonçalves LT, Machado R, Oliveira LR, Valiati VH and Fagundes NJR (2018) Molecular Identification of Shark Meat From Local Markets in Southern Brazil Based on DNA Barcoding: Evidence for Mislabeling and Trade of Endangered Species. Front. Genet. 9:138. doi: 10.3389/fgene.2018.00138

while 53% are threatened and 47% are critically endangered in Brazil. These results underline that labeling the meat of any shark species as "cação" is problematic for monitoring catch allocations from the fishing industry and discourages consumer engagement in conservationist practices through informed decision-making.

Keywords: cação, elasmobranch, cytochrome oxidase-1, shark fisheries, wildlife DNA forensics

INTRODUCTION

Elasmobranch (subclass Elasmobranchii) is a group of cartilaginous fishes that include sharks (superorder Selachii) and rays (superorder Batoidea). Even though elasmobranchs comprise less than 1% of the world fisheries catch (Food and Agriculture Organization of United Nations, 2014, 2016), these species have biological characteristics that make them particularly vulnerable to overfishing, such as a low fecundity and late sexual maturation (Bornatowski et al., 2014b). Indeed, several recent reports have indicated that there has been a reduction of elasmobranch populations, resulting in demographic collapse at a regional scale (Baum et al., 2003; Barausse et al., 2014). The overfishing of sharks is especially problematic because these top predators play a key role in marine ecosystems, and, therefore, their population dynamics may affect all local marine diversity (van der Elst, 1979; Heithaus et al., 2008; Gallagher et al., 2012; Pauly et al., 2013; Worm et al., 2013; Bornatowski et al., 2014a). In 1999, FAO (Food and Agriculture Organization) launched an international plan for the conservation and management of sharks and rays, recognizing the high vulnerability of these organisms (Vannuccini, 1999). However, despite this initiative, a significant number of elasmobranch species has remained overexploited or threatened by fisheries activities (Camhi et al., 2009; Cosandey-Godin and Morgan, 2011), which is illustrated by the 42% global increase in the shark meat trade from 2000 to 2011 (Food and Agriculture Organization of United Nations, 2015).

While shark fins are considered to be one of the most valuable products in the ocean (Gallagher and Hammerschlag, 2011), shark meat often attains only 20–60% of the price of tuna and mackerel meat (Bonfil, 1994). As a result, captured individuals usually have their fins removed for the shark fin market, the head is discarded, and the remaining central body part ("cigar") is then sold for the shark meat market with no special care (Kotas et al., 2008; Ward-Paige et al., 2012). From a taxonomic point of view, the removal of the head and fins represents a challenge to reliable species identification based on morphological features, allowing shark carcasses to be traded fraudulently (Holmes et al., 2009).

Brazil is among the six countries that have the highest capture rate for elasmobranchs (Lack and Sant, 2006), even though a thorough assessment of the impact of industrial fishing is made difficult by inaccurate records (Barreto et al., 2017). Southern Brazil is a region of high elasmobranch diversity (Lucifora et al., 2011), and has a large extractive marine fishing industry, with approximately 160 thousand metric tons of fish caught annually (MPA. Boletim estatístico da pesca e aquicultura, 2011). The two southernmost states, Santa Catarina (SC) and Rio Grande do Sul (RS), are responsible for 98% of the catches (MPA. Boletim estatístico da pesca e aquicultura, 2011). In addition, Brazil is a major player in the meat trade market, acting as the world's largest importer of shark meat in 2011 (Food and Agriculture Organization of United Nations, 2015). Internally, the Brazilian shark meat market trades several different elasmobranch species under the popular name "cação" (or other related popular terms such as "caçonete" and "anjo"), which is used to label several species (Figure S1). For example, Neto (2013) found 21 different species traded under the common name "cação", including hammerhead sharks (*Sphyrna* spp.), the blue shark (*Prionace glauca*), the tiger shark (*Galeocerdo cuvier*), the bull shark (*Carcharhinus leucas*), the Galapagos shark (*C. galapagensis*), and the blacktip shark (*C. limbatus*). Consumers value "cação" meat for its low cost and for being a "thornless fish" (Bornatowski et al., 2007). However, most consumers are not aware that "cação" is a synonym for sharks (or rays), and others believe that "cação" represents "a specific race of sharks" or even "a race of small sharks" (Bornatowski et al., 2015). Supermarkets, fisheries, and restaurants often omit any other information when selling "cação" meat. Indeed, the use of this term is so widespread that even Brazilian regulatory agencies categorize all elasmobranch species as "cação" without any species-specific information (MPA. Boletim estatístico da pesca e aquicultura, 2011).

The imprecise nomenclature of elasmobranchs makes it difficult to mitigate the negative effects of human shark consumption, as it becomes more difficult to inform the consumer if the product comes from a threatened species or from an illegal species trade. Since shark carcasses are sliced before being sold, it is virtually impossible to obtain accurate species diagnosis based on morphological traits for marketed elasmobranchs (Bornatowski et al., 2015). Therefore, there is an increasing need for fast, reliable, and cheap testing for determining the taxonomic identity of commercialized fishes (Rasmussen and Morrissey, 2008). A precise identification of marketed species also assures that the correct information is presented to the consumer, motivating him or her to take part in honest and regulated trade (Moretti et al., 2003; Martinez et al., 2005).

DNA barcoding uses a small fragment from a DNA sequence located within a standardized region of the genome to allow precise species identification (Hebert et al., 2003). In animals, the standard DNA barcode comes from a stretch of 650 base pairs (bp) from the 5′ end of the mitochondrial gene Cytochrome Oxidase Subunit I (COI or Cox 1) (Meyer and Paulay, 2005; Hajibabaei et al., 2007). This technique has been widely used in a range of studies of species identification (e.g., Meyer and

Paulay, 2005; Lowenstein et al., 2010; Carvalho et al., 2011; Rodrigues-Filho et al., 2012; Galimberti et al., 2013). Whilst DNA barcoding is a valuable tool for species identification, especially when the entire organism cannot be accessed for morphology, there are important limitations concerning its accuracy, which depend on the reference database available and on the degree of genetic difference among species (see Frézal and Leblois, 2008 for a review on the pros and cons of DNA barcoding). The aim of this study is to use DNA barcodes to identify, at the species level, samples of "cação" (or similarly labeled) meat available in local markets in Southern Brazil. Finally, we discuss the implications of these findings in the context of elasmobranch conservation in Brazil.

MATERIALS AND METHODS

Sample Collection

We studied samples sold under general names such as "cação," "caçonete," and "filé anjo," which usually refer to elasmobranch species. Between 2008-2013 and in 2016 we acquired filet samples from local fish markets and supermarkets in different cities from the RS and SC states in Southern Brazil (**Figure 1**,**Table 1**). We also included in the analysis samples from *Sphyrna lewini* ($n = 4$), *Pseudobatos horkelii* ($n = 2$), *Rhizoprionodon lalandii* ($n = 1$), *Narcine brasiliensis* ($n = 1$), *Zapteryx brevirostris* ($n = 2$), and *Gymnura altavela* ($n = 1$), collected from fishing vessels and morphologically identified according to Figueiredo (1977), to serve as controls for the DNA barcode identification. These samples are identified as E__ in **Table 1**. All samples were stored in 95% ethanol at −20°C.

Laboratory Procedures

DNA extraction started from a small portion (∼100 mg) of the tissue. For most samples we used the Wizard® Genomic DNA Purification Kit (Promega) modified to include an initial digestion step with 200 μg proteinase k (Aljanabi and Martinez, 1997). For the remaining samples, we used a protocol based on the CTAB method (Doyle, 1987). We used the COI primers FishF2 (5′ TCG ACT AAT CAT AAA GAT ATC GGC AC 3′) and FishR2 (5′ ACT TCA GGG TGA CCG AAG AAT CAG AA 3′) (Ward et al., 2005). Amplification reactions were prepared with 0.4 μM of each dNTP, 1.5 mM $MgCl_2$, 0.5 μM of each primer, 1 U Taq Polymerase, and ∼40 ng of genomic DNA. Cycling conditions included an initial denaturing step of 94°C for 5′, followed by 10 cycles of 94°C for 1′, 55°C (−0.5°C/cycle) for 1′, and 72°C for 1′30″, and 30 additional cycles of 94°C for 1′, 50°C for 1′, and 72°C for 1′30″, with a final extension step of 72°C for 5′. The amplification products were visualized on a 1% agarose gel stained with GelRed™ (Biotium). PCR products were purified enzymatically using 0.33U SAP (Shrimp Alkaline Phosphatase) and 3.33U ExoI (Exonuclease I). PCR products were sequenced by the Sanger method in Macrogen Inc. (Seoul, South Korea) and Ludwig Biotec (Porto Alegre, Brazil). DNA sequencing was performed on both strands using the primers mentioned above.

Data Analysis

The consensus sequence for each sample was assembled and trimmed in Geneious 9.1 (www.geneious.com). The reliability

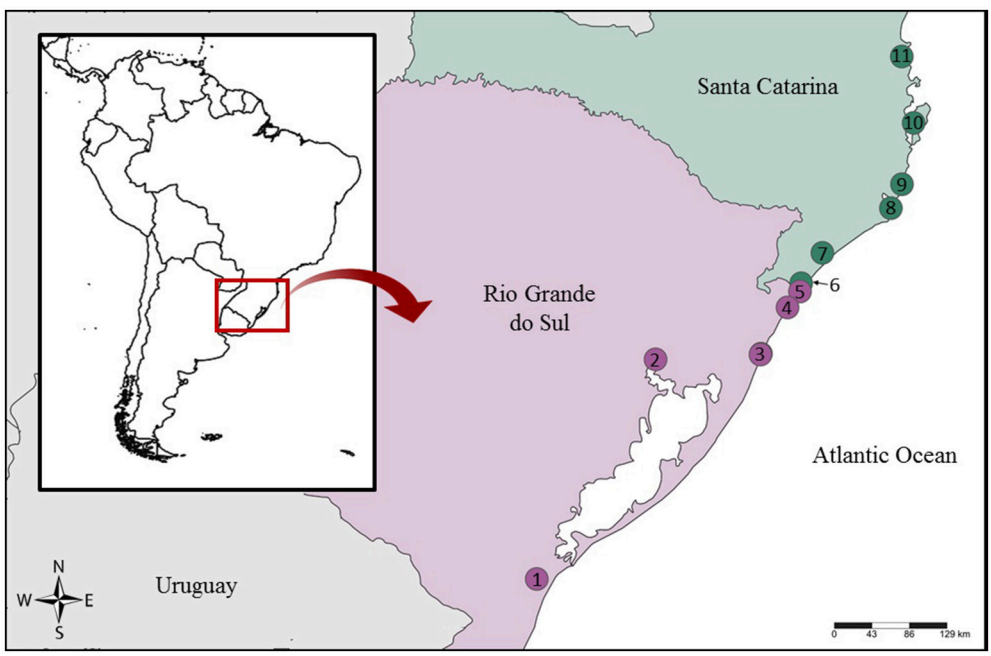


FIGURE 1 | Sampling locations in Southern Brazil. 1, Rio Grande; 2, Porto Alegre; 3, Tramandaí + Imbé; 4, Arroio do Sal; 5, Torres; 6, Passo de Torres; 7, Araranguá; 8, Laguna; 9, Imbituba; 10, Florianópolis; 11, Itajaí.

TABLE 1 | Sample information, species identification, average genetic distance, and results from the BLAST search.

Sample					Candidate species	Avg. distance[a]	% Coverage[b]	% Identity[b]
ID	Accession no.	Seq. Size	Location	Type				
IIL04	MG703512	650b	Itajaí, SC	fresh	*Carcharhinus brachyurus*	0.001	98	99
IIL05	MG703513	650b	Itajaí, SC	fresh	*Carcharhinus brachyurus*	0.001	97	99
IIL14	MG703514	650b	Itajaí, SC	fresh	*Carcharhinus brachyurus*	0.001	98	99
IIL04-2	MG703515	642b	Itajaí, SC	fresh	*Carcharhinus falciformis*	0.003	98	99
IIL27	MG703516	650b	Itajaí, SC	fresh	*Carcharhinus falciformis*	0.003	99	99
FA08	MG703517	650b	Porto Alegre, RS	fresh	*Galeorhinus galeus*	0.001	95	99
MP60	MG703518	650b	Porto Alegre, RS	fresh	*Genidens barbus*	NC	96	100
E14*	MG703519	615b	Arroio do Sal, RS	fresh	*Gymnura altavela*	0.020	99	99
IIL37	MG703520	650b	Laguna, SC	fresh	*Gymnura altavela*	0.021	98	99
IIL36	MG703521	650b	Laguna, SC	fresh	*Myliobatis goodei*	0.013	96	99
E13*	MG703522	650b	Torres, RS	fresh	*Narcine brasiliensis*	0.003	99	99
IIL15	MG703523	650b	Itajaí, SC	fresh	*Prionace glauca*	0.001	99	99
IIL30	MG703524	650b	Imbituba, SC	fresh	*Prionace glauca*	0.000	99	99
IIL31	MG703525	613b	Imbituba, SC	fresh	*Prionace glauca*	0.000	100	100
IIL34	MG703526	650b	Imbituba, SC	fresh	*Prionace glauca*	0.000	100	99
IIL35	MG703527	621b	Imbituba, SC	fresh	*Prionace glauca*	0.000	100	99
O22	MG703528	607b	Florianópolis, SC	fresh	*Prionace glauca*	0.000	97	100
FA02	MG703529	523b	Tramandaí, RS	fresh	*Prionace glauca*	0.002	100	99
FA03	MG703530	611b	Tramandaí, RS	fresh	*Prionace glauca*	0.000	96	100
FA23	MG703531	641b	Porto Alegre, RS	frozen	*Prionace glauca*	0.000	99	100
FA24	MG703532	588b	Porto Alegre, RS	frozen	*Prionace glauca*	0.002	100	99
FA25	MG703533	612b	Porto Alegre, RS	frozen	*Prionace glauca*	0.003	100	99
FA26	MG703534	634b	Porto Alegre, RS	frozen	*Prionace glauca*	0.000	100	100
FA27	MG703535	526b	Porto Alegre, RS	frozen	*Prionace glauca*	0.000	100	99
FA29	MG703536	556b	Porto Alegre, RS	frozen	*Prionace glauca*	0.000	100	100
FA31	MG703537	527b	Porto Alegre, RS	frozen	*Prionace glauca*	0.000	100	99
IIL26	MG703538	650b	Itajaí, SC	fresh	Rajiformes sp. BOLD AABB	0.000	96	100
E34*	MG703539	650b	Torres, RS	fresh	*Pseudobatos horkelii*	0.003	97	100
E36*	MG703540	610b	Torres, RS	fresh	*Pseudobatos horkelii*	0.002	98	100
E26*	MG703541	650b	Araranguá, SC	fresh	*Rhizoprionodon lalandii*	0.001	94	100
IIL13	MG703542	521b	Itajaí, SC	fresh	*Rhizoprionodon lalandii*	0.001	99	99
FA05	MG703543	590b	Imbé, RS	fresh	*Rhizoprionodon lalandii*	0.001	91	99
FA17	MG703544	512b	Porto Alegre, RS	fresh	*Rhizoprionodon lalandii*	0.001	100	99
O24	MG703545	650b	Florianópolis, SC	fresh	*Rhizoprionodon porosus*	0.001	95	100
E07*	MG703546	519b	Tramandaí, RS	fresh	*Sphyrna lewini*	0.024	100	100
E08*	MG703547	650b	Tramandaí, RS	fresh	*Sphyrna lewini*	0.021	98	99
E15*	MG703548	619b	Arroio do Sal, RS	fresh	*Sphyrna lewini*	0.022	97	100
E44*	MG703549	650b	Tramandaí, RS	fresh	*Sphyrna lewini*	0.021	97	99
MG04	MG703550	650b	Rio Grande, RS	fresh	*Sphyrna lewini*	0.034	97	99
MP55	MG703551	534b	Porto Alegre, RS	fresh	*Sphyrna lewini*	0.024	100	99
MP57	MG703552	542b	Porto Alegre, RS	fresh	*Sphyrna lewini*	0.023	97	100
MP58	MG703553	621b	Porto Alegre, RS	fresh	*Sphyrna lewini*	0.022	97	100
O06	MG703554	608b	Passo de Torres, SC	fresh	*Sphyrna lewini*	0.022	97	99
O07	MG703555	628b	Passo de Torres, SC	fresh	*Sphyrna lewini*	0.021	97	100
O08	MG703556	612b	Passo de Torres, SC	fresh	*Sphyrna lewini*	0.022	98	100
O09	MG703557	534b	Passo de Torres, SC	fresh	*Sphyrna lewini*	0.024	100	99
O27	MG703558	534b	Florianópolis, SC	fresh	*Sphyrna lewini*	0.024	100	100
O29	MG703559	650b	Florianópolis, SC	fresh	*Sphyrna lewini*	0.021	99	100
O28	MG703560	463b	Florianópolis, SC	fresh	*Sphyrna zygaena*	0.000	100	99

(Continued)

TABLE 1 | Continued

Sample					Candidate species	Avg. distance[a]	% Coverage[b]	% Identity[b]
ID	Accession no.	Seq. Size	Location	Type				
FA21	MG703561	642b	Porto Alegre, RS	fresh	*Shpyrna zygaena*	0.001	100	100
MP15	MG703562	630b	Porto Alegre, RS	fresh	*Squalus cubensis*	0.000	98	100
MP18	MG703563	603b	Porto Alegre, RS	fresh	*Squalus mitsukurii*	0.001	98	100
MP16	MG703564	641b	Porto Alegre, RS	fresh	*Squalus mitsukurii*	0.001	96	100
FA16	MG703565	593b	Porto Alegre, RS	fresh	*Squatina guggenhein*	0.001	96	100
MG08	MG703566	204b	Rio Grande, RS	fresh	*Squatina occulta*	0.000	95	100
IIL01	MG703567	650b	Itajaí, SC	frozen	*Xiphias gladius*	NC	100	99
IIL03	MG703568	650b	Itajaí, SC	frozen	*Xiphias gladius*	NC	100	99
IIL16	MG703569	650b	Itajaí, SC	frozen	*Xiphias gladius*	NC	100	99
IIL18	MG703570	589b	Itajaí, SC	frozen	*Xiphias gladius*	NC	100	99
IIL19	MG703571	458b	Itajaí, SC	frozen	*Xiphias gladius*	NC	100	100
IIL25	MG703572	622b	Itajaí, SC	frozen	*Xiphias gladius*	NC	100	100
E50*	MG703573	650b	Passo de Torres, SC	fresh	*Zapteryx brevirostris*	0.030	97	99
E54*	MG703574	232b	Passo de Torres, SC	fresh	*Zapteryx brevirostris*	0.000	100	100

[a] *Average genetic distance against all sequences from the same species in the final dataset.*
[b] *%Coverage and %Identity values considering the top-BLAST hit for the candidate species.*
[*] *All samples identified as E__ were obtained directly from fishing vessels, and were not purchased.*
NC, not computed.

of each consensus sequence was assessed by a thorough visual inspection of the chromatograms used in the assemblies to check for sequencing errors and artifacts. Low quality regions in the chromatograms, identified as a stretch of five or more contiguous bases having high background noise and uneven spacing, were trimmed and removed before sequence assembly. Because the assembly algorithm gives more weight to better quality reads, cases of sequence heterogeneity between strands are resolved in favor of the best quality read or, if both reads had similar quality for that position, marking it as an ambiguous base (N, R, Y, etc.). The consensus sequence was then used as a query for comparison with the NCBI database (http://www.ncbi.nlm.nih.gov/) using the Basic Local Alignment Search Tool—Nucleotide (BLASTn). In all cases, BLAST matched COI sequences from elasmobranchs (or, in some cases, from teleosts) with good coverage and identity (see Results), suggesting that we generated authentic COI sequences from our samples. We recorded the species representing the top BLAST hit for each query. Following this, we built a dataset of 2,877 COI sequences deposited in the GenBank including all species of all genera represented in the list of top BLAST hits. For example, if the top BLAST hit for a given sample was *S. lewini*, we included all sequences from all *Sphyrna* species (including eventual "*Sphyrna* sp." entries) in the dataset. We then picked at random 2–8 sequences for each species, which were aligned with the consensus sequences from the samples generated in this study using MAFFT 7.0 (Katoh and Standley, 2013), leading to a final dataset of 323 COI sequences for 147 species (including undescribed or unknown species). As a final quality control step, we checked the dataset for nonsense mutations and alignment gaps, as both could indicate the presence of nuclear mitochondrial translocations (Numts) (Triant and DeWoody, 2007). The final alignment file can be downloaded as Supplementary Material (File S1). The best substitution model (HKY+G+I) for this final dataset was estimated in jModelTest 2 based on the corrected Akaike Information Criterion (AICc) (Darriba et al., 2012). Pairwise genetic distances was estimated in PAUP* 4.0 (Swofford, 2002) based on most likely substitution model and its associated parameters [Lset base = (0.3624 0.2434 0.0914) nst = 2 tratio = 6.1561 rates = gamma shape = 0.8490 ncat = 4 pinvar = 0.4860]. For this final dataset, we inferred the maximum likelihood (ML) tree in RAxML 8.2.10 (Stamatakis, 2014). Node credibility was assessed based on 1,000 bootstrap replicates.

RESULTS

In total, 63 samples were collected, amplified, sequenced, and compared to GenBank sequences (**Table 1**). High quality sequences ranged between 204 and 650 bases. There was no sequence heterogeneity between strands involving high quality bases from two or more reads. Overall, our analysis suggests the presence of 20 different species among the samples. Seven samples were identified as belonging to two Actinopterigii (ray-finned fishes) species: *Xiphias gladius* (Perciformes, swordfish; $n = 6$), and *Genidens barbus* (Siluriformes, white sea catfish; $n = 1$). The remaining samples may represent 18 elasmobranch species from three shark orders (Carcharhiniformes, Squaliformes, and Squatiniformes; $n = 42$, 3, 2, respectively) and four ray orders (Rhinopristiformes, Myliobatiformes, Rajiformes, and Torpediniformes; $n = 4$, 3, 1, 1, respectively). Three ray species (*P. horkelii*, *Z. brevirostris*, and *N. brasiliensis*) were only found in samples from fishing vessels (i.e., they were not purchased in the market).

Based on COI sequences, all but one elasmobranch samples were identified at the species level, representing 17 formally described species. One sample was associated with an undescribed or unsequenced species (which occurs in GenBank as "Rajiformes sp. BOLD: AABB1882"). The most common species found among market samples were *Prionace glauca* (blue shark, $n = 15$) and *S. lewini* (scalloped hammerhead shark, $n = 14$). All other species were far less common, including *R. lalandii* (Brazilian sharpnose shark, $n = 4$), *Carcharhinus brachyurus* (copper shark, $n = 3$), *Carcharhinus falciformis* (silky shark, $n = 2$), *Sphyrna zygaena* (smooth hammerhead shark, $n = 2$), *Squalus mitsukurii* (shortspine spurdog, $n = 2$), *Galeorhinus galeus* (school shark, $n = 1$), *Rhizoprionodon porosus* (Caribbean sharpnose shark, $n = 1$), *Squalus cubensis* (Cuban dogfish, $n = 1$), *Squatina occulta* (hidden angel shark, $n = 1$), and *Squatina guggenheim* (spiny angel shark, $n = 1$). All ray species identified in the study occurred once or twice among the samples: *G. altavela* (spiny butterfly ray, $n = 2$), *P. horkelii* (Brazilian guitarfish, $n = 2$), *Z. brevirostris* (shortnose guitarfish, $n = 2$), *Myliobatis goodei* (southern eagle ray, $n = 1$), and *N. brasiliensis* (Brazilian electric ray, $n = 1$).

The average genetic distance between each sample and representatives of its most likely candidate species (determined by its clustering in the ML tree) was always lower than 3.50%, and usually lower than 1% (**Table 1**). The ML tree showed cohesive clusters of conspecific sequences (**Figure 2**). The few exceptions, which had bootstrap support values lower than 90, included *S. mitsukurii*, *C. brachyurus*, *S. guggenheim*, and *S. occulta* (**Figure 2**). In all cases, however, the estimated genetic distance between our samples and reference sequences were used to indicate the most likely candidate species (shown in **Table 1**). An interesting case is sample MP16, whose top-hit in BLAST was *Squalus montalbani*, but clustered with *S. mitsukurii* in the ML tree (**Figure 2**). However, both MP16 and MP18 showed a much smaller distance from *S. mitsukurii* (0.0009) than to any other closely related species (0.0027 vs. *S. cf. megalops*; 0.0043 vs. *S. montalbani*; 0.0058 vs. *S. chloroculus*; and 0.0088 vs. *S. cf. mitsukurii*). Similarly, IIL04, IIL05, and IIL14 were much closer to *C. brachyurus* (0.0014) than to *C. brevipinna* (0.0150), MG08 was closer to *S. occulta* (0.0000) than to *S. guggenheim* (0.0071), while FA16 was closer to *S. guggenheim* (0.0008) than to *S. occulta* (0.0064). The complete distance matrix can be downloaded as Supplementary Material (File S2).

DISCUSSION

We found 18 Elasmobranchii and two Actinopterigii species among the samples acquired in Southern Brazilian fish markets as "cação," "caçonete," or "filé anjo." This represents 17% of all elasmobranch species registered for Southern Brazil and 13% of the species described for Brazil (Bornatowski et al., 2009). Other studies, based on other molecular markers, that aimed at species identification of shark filets from Northern Brazil have also shown the great number of species being trade without any taxonomic control (Rodrigues-Filho et al., 2009; Palmeira et al., 2013). Unfortunately, our DNA data does not allow us to conclude that these samples represent individuals captured in Southern Brazil. For example, most individuals included in the final dataset did not have location information. Additionally, even if this was available, it is unclear whether COI would have enough resolution to allow unambiguous recognition of regional stocks for these species (Antoniou and Magoulas, 2014). However, the fact that the vast majority of samples collected in this study were purchased fresh is a strong indication that these specimens may have been captured off Southern Brazil or in nearby areas.

The use of the COI DNA barcode allowed us to identify all samples at the specific level even though some cases deserve further discussion. The best match for IIL26 was an undescribed or unsequenced species, Rajiformes sp. BOLD:AAB1882 (Coverage = 96%, Identity = 100%). The sample MG08 resulted in a short DNA sequence, whose top-result in BLAST was against *S. occulta* (Coverage = 95%, Identity = 100%), but showed an inconclusive clustering with any *Squatina* species in the ML tree (**Figure 2**). Nevertheless, as occurred for other samples (FA16, IIL04, IIL05, IIL14, MP16, MP18), comparing the genetic distance among alternative candidate species allowed the identification of the most likely candidate for each sample. In the case of the samples associated to *Squatina*, species identification was corroborated by the fact that both *S. occulta* and *S. guggenheim* occur off Southern Brazil (Vaz and Vaz and De Carvalho, 2013) and that both samples were acquired as fresh filets, likely indicating a local catch. With a single exception, the species associated with the top-BLAST result also resulted in the lowest average genetic distance. The exception was MG16, whose top-BLAST result was *Squalus montalbani* (Coverage = 100%, Identity = 99%), but whose lowest average genetic distance was against *S. mitsukurii*, which also represented the second and third top-BLAST results (Coverage = 96%, Identity = 100%). The low genetic distance among *Squalus* species and the lack of a clear structure in the ML tree (**Figure 2**) may indicate that DNA barcoding for this genus may be more complicated than for other genera, and may require other genetic markers. From the taxonomic point of view, it is difficult to discriminate among *Squalus* species (Haddad and Gadig, 2005), which may be due to a shallow diversification time that is reflected in the low genetic distances among several species. The inherently difficult taxonomy of the genus may favor misnomers in reference databases. In this regard, *S. cubensis* presents a likely example of database confusion. There are two COI sequences for this species in GenBank. However, while the entry FJ519595 is close to *S. mitsukurii* (∼0.2% genetic distance) the other, FN431670, is distantly related to it (∼7.2% genetic distance) and associated with sample MP15 (**Figure 2**). These issues reinforce the importance of database curation and maintenance, with rigorous taxonomic criteria for the deposition of reference sequences (Ekrem et al., 2007; Teletchea, 2009; Dudgeon et al., 2012). It also highlights that in some cases it may be important to analyze additional genetic markers for a more accurate species identification (Mendonça et al., 2009; Moftah et al., 2011; Pérez-Jiménez et al., 2013).

The most abundant shark species in our samples were *Prionace glauca* and *S. lewini* (23.8 and 22.2%, respectively).

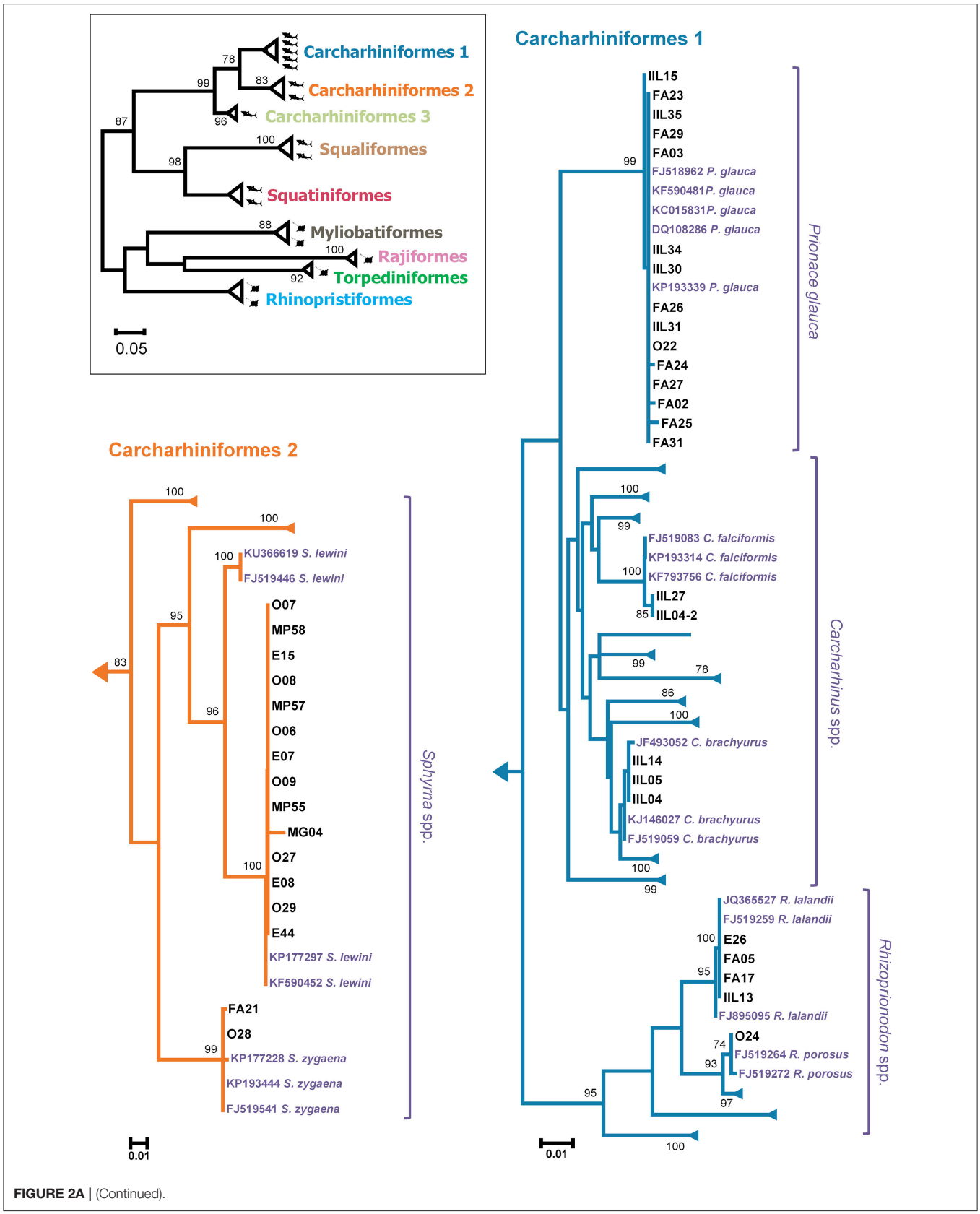


FIGURE 2A | (Continued).

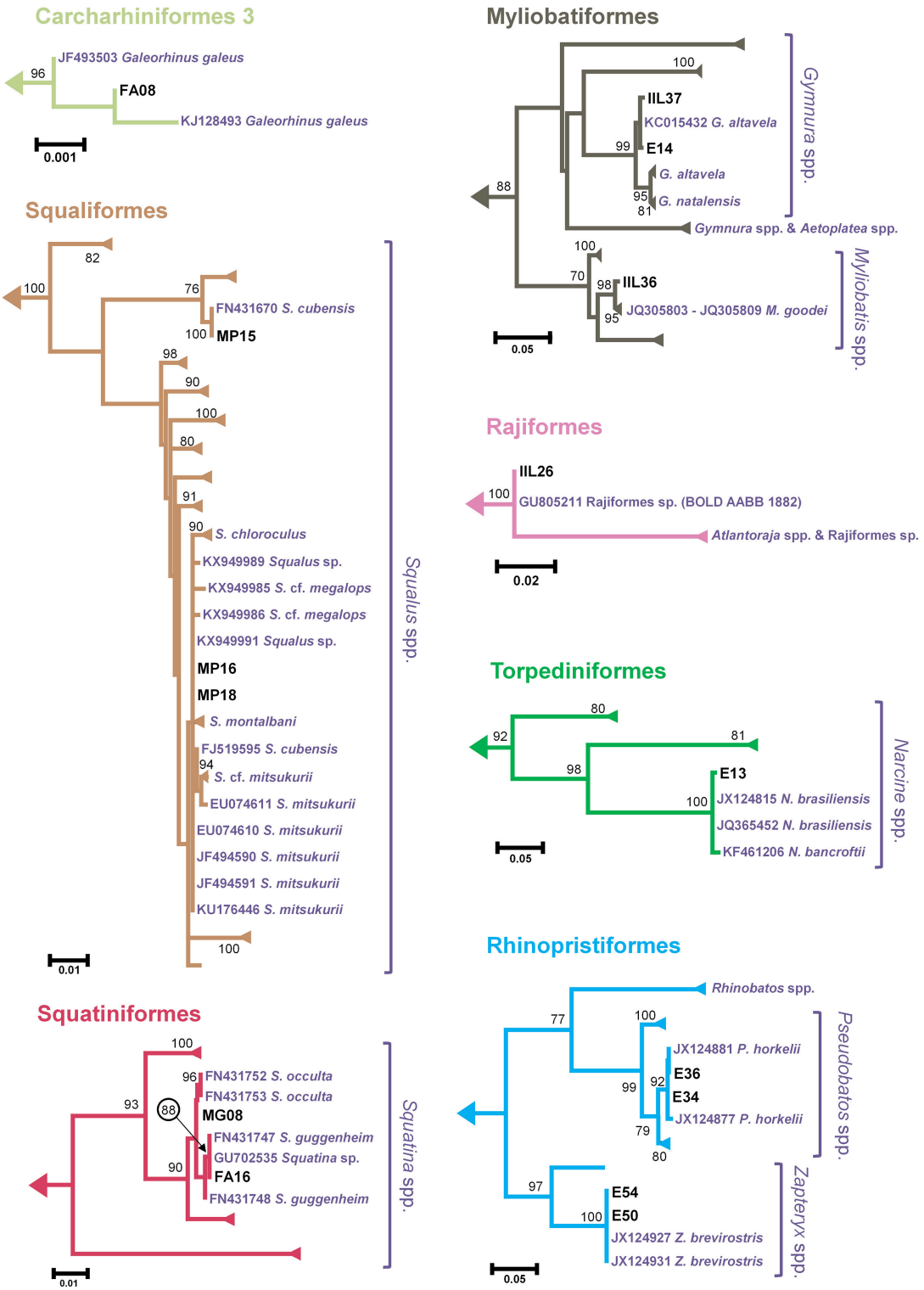


FIGURE 2B | ML tree based on HKY+G+I distance. The miniature on the upper left side shows major groups, displayed in more detail in individual panels. The number of shark and ray symbols represent the number of different species identified in the study for each group. Please note that this is an unrooted tree. Most entries were collapsed and the names were omitted for clarity. Samples from the present study are labeled according to **Table 1**. The most likely candidate species, together with other closely related species are shown in red. The numbers above the branches represent bootstrap percentage based on 1,000 replicates. Bootstrap values <70 were omitted. Please note the different scale among panels. The full ML tree is available as Supplementary Material (File S3).

TABLE 2 | Conservation status (global, national, and regional) of the species found in this study.

Species	Common name[a]	IUCN[b]	ICMBio[c]	RS[d]	SC[e]
Carcharhinus brachyurus	cooper shark	NT 2003	DD*	–	–
Carcharhinus falciformis	silky shark	NT 2016	NT*	–	–
Galeorhinus galeus	school shark	VU 2006	CR	CR	–
Gymnura altavela	butterfly ray	VU 2007	CR	EN	–
Myliobatis goodei	southern eagle ray	DD 2009	CR	CR	–
Narcine brasiliensis	Brazilian electric ray	DD 2007	DD*	–	–
Prionace glauca	blue shark	NT 2009	NT*	VU	–
Pseudobatos horkelli	Brazilian guitarfish	CR 2016	CR	CR	CR
Rhizoprionodon lalandii	Brazilian sharpnose shark	DD 2004	NT*	–	–
Rhizoprionodon porosus	Caribbean sharpnose shark	LC 2006	DD*	–	–
Sphyrna lewini	scalloped hammerhead shark	EN 2007	CR	CR	EN
Sphyrna zygaena	smooth hammerhead	VU 2005	CR	CR	EN
Squalus cubensis	Cuban dogfish	DD 2006	–	–	–
Squalus mitsukurii	shortspine spurdog	DD 2007	DD*	–	–
Squatina guggenheim	spiny angel shark	EN 2007	CR	CR	EN
Squatina occulta	smoothback angel shark	EN 2007	CR	CR	–
Zapteryx brevirostris	shortnose guitarfish	VU 2006	VU	CR	–

**Species included in the National List of Species Threatened of Extinction (available at Portaria MMA n° 445 of 2014).*
[a] IUCN Red List of Threatened Species (IUCN, 2017).
[b] Global conservation status according to IUCN (2017) criteria, followed by the year of assessment: (CR) critically endangered; (EN) endangered; (VU) vulnerable; (NT) near threatened; (DD) data deficient; (NE) not evaluated.
[c] National conservation status according to the Brazilian Red Book of Threatened Faunal Species (Instituto Chico Mendes de Preservação da Biodiversidade, 2016).
[d] Regional conservation status according to the List of Threatened Fauna of the Rio Grande do Sul State (Fundação Zoobotânica e Secretaria do Ambiente Desenvolvimento Sustentável, Decreto n° 51.797).
[e] Regional conservation status according to the List of Threatened Fauna of the Santa Catarina State (Fundação de Meio Ambiente – FATMA).

P. glauca is distributed globally and its capture volume has been estimated at approximately 20 million individuals per year (Mendonça et al., 2012). Despite its endangered status (**Table 2**), *P. glauca* is the most fished shark in the world, representing 56% of the total catch of pelagic sharks, especially by industrial fisheries in which the target species are tuna or swordfish (Rose, 1996; Dulvy et al., 2003; Camhi et al., 2009). After the increasing market demand for shark fins and the high prices paid for them, these animals began to be targeted for the removal of these parts, with the carcasses being sold worldwide (Domingues, 2011). Indeed, we identified *P. glauca* in all samples acquired as frozen filets, which may reflect that these individuals were captured in other parts of the world, such as Taiwan and subsequently imported to Brazil (Figure S1). However, we also found *P. glauca* among fresh samples, which more likely indicates local capture. On the other hand, *S. lewini* was the most abundant species among fresh samples, which may indicate a higher local impact on this species. Several authors have raised concerns of predatory fishing for this species off Brazil due to the high commercial value of its fins (Amorim et al., 2011). This results in fishing pressures occurring over all phases and life cycles of these animals, including neonates (Mader et al., 2007) both on the continental shelf and in oceanic waters (Kotas, 2004; Kotas et al., 2005; Vooren and Klippel, 2005).

Regarding their conservation status, IUCN estimates that 47% of the elasmobranch species found in this study are considered threatened at the global level, 53% are threatened at the national level, and 47% are critically endangered at the national level (**Table 2**). It is difficult, however, to present a more regional picture, given that the red list for both Rio Grande do Sul and Santa Catarina states include only 59 and 23.5% of the species identified in this study, even though there are records for most of these species off these Brazilian states (Gadig, 2001). The conservation status for *R. lalandii, S. mitsukurii, S. cubensis, M. goodei,* and *N. brasiliensis* is unknown due to data deficiency (DD). In the worst-case scenario, ∼50% of the species identified in this study would be threatened to some extent.

Our sampling was restricted to the south of Brazil due to a limited budget, but it would be important to perform similar studies in other Brazilian regions to provide a better picture of the shark fishing and trade in the country. It should be noted, however, that the Southern coast of Brazil is a hotspot for shark diversity, with high species richness, high endemism, and functional richness (Lucifora et al., 2011). Another future direction would be investigating how much of the shark meat market involves individuals fished locally.

Finally, an important issue in the conservation of these species is how local human populations will engage in more sustainable consumption practices. In this sense, labeling the meat of any shark species as "cação" may impose major barriers to conservation measures for this group, allowing the inadvertent consumption of protected species (Jacquet and Pauly, 2008). Indeed, Bornatowski et al. (2015), who interviewed fish meat consumers in Southern Brazil, reported that 61% of respondents

claimed that they have never tried shark meat, even though they ate "cação." In addition, 69% of respondents said they did not know that at least 25% of all elasmobranchs are threatened. Given these answers, it is evident that a significant portion of the population buying these products is not aware of the impact of their consumption habits, or of the current conservation status of elasmobranch species. Another issue for consumers is mislabeling of shark products, a common outcome of DNA barcode assessments of seafood products (Barbuto et al., 2010; Filonzi et al., 2010). This is illustrated by the presence of the two teleost species detected among our sample (**Table 1**). Therefore, it becomes essential and an ethical responsibility for the industry to label their products correctly and allow informed decision-making by the consumers. We suggest that all meat being sold as "cação" should be accompanied by the species common name, followed by its scientific name, and, whenever possible, the species threat categories according to the IUCN Red List. While fishing legislation may also have a positive impact on natural populations by suspending the capture and marketing of endangered elasmobranchs, environmental education measures focusing on the fishing community and on consumers will be fundamental for the effective protection of these species.

AUTHOR CONTRIBUTIONS

FA-S, CS, LO, VV, and NF designed the study; FA-S, CS, CC, PF, and RM executed experimental procedures; FA-S, CS, PF, LG, LO, VV, and NF performed data analysis and interpretation; FA-S, CS, VV, and NF wrote the paper.

FUNDING

This research was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação Grupo Boticário de Proteção à Natureza (Projeto 0921_20112), Programa de Pós-Graduação em Biologia Animal (PPGBAN-UFRGS), and Programa de Pós-Graduação em Genética e Biologia Molecular (PPGBM-UFRGS) and Programa de Pós-Graduação em Biologia (UNISINOS).

ACKNOWLEDGMENTS

The authors would like to thank Walter de Nisa e Castro Neto (Pró-Squalus), Maria João Veloso da Costa Ramos Pereira (UFRGS), Clarissa Fleck (McGill University), Andrea Thomaz (University of Michigan), Daniel Lyons (University of Michigan), Bonnie Armour (University of Adelaide), and two reviewers for contributions in earlier versions of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00138/full#supplementary-material

REFERENCES

Aljanabi, S. M., and Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* 25, 4692–4693. doi: 10.1093/nar/25.22.4692

Amorim, A. F., Della-Fina, N., and Piva-Silva, N. (2011). Hammerheads sharks, *Sphyrna lewini* and *S. zygaena* caught by longliners off Southern Brasil 2007-2008. *Collect. Vol. Sci. Pap. ICCAT* 66, 2121–2133.

Antoniou, A., and Magoulas, A. (2014). "Application of mitochondrial DNA in stock identification," in *Stock Identification Methods: Applications in Fishery Science*, eds S. X. Cadrin, L. A. Kerr, and S. Mariani (Waltham, MA: Academic Press). 257–295.

Barausse, A., Correale, V., Curkovic, A., Finotto, L., Riginella, E., Visentin, E., et al. (2014). The role of fisheries and the environment in driving the decline of elasmobranchs in the northern Adriatic Sea. *ICES J. Mar. Sci.* 71, 1593–1603. doi: 10.1093/icesjms/fst222

Barbuto, M., Galimberti, A., Ferri, E., Labra, M., Malandra, R., Galli, P., et al. (2010). DNA barcoding reveals fraudulent substitutions in shark seafood products: the Italian case of "palombo" (*Mustelus* spp.). *Food Res. Int.* 43, 376–381. doi: 10.1016/j.foodres.2009.10.009

Barreto, R. R., Bornatowski, H., Motta, F. S., Santander-Neto, J., Vianna, G. M. S., and Lessa, R. (2017). Rethinking use and trade of pelagic sharks from Brazil. *Mar. Policy* 85, 114–122. doi: 10.1016/j.marpol.2017.08.016

Baum, J. K., Myers, R. A., Kehler, D. G., Worm, B., Harley, S. J., and Doherty, P. A. (2003). Collapse and conservation of shark populations in the Northwest Atlantic. *Science* 299, 389–392. doi: 10.1126/science.1079777

Bonfil, R. (1994). *Overview of World Elasmobranch Fisheries*. Rome: FAO.

Bornatowski, H., Abilhoa, V., and Charvet-Almeida, P. (2009). Elasmobranchs of the Paraná Coast, southern Brazil, south-western Atlantic. *Mar. Biodivers. Rec.* 2:e158. doi: 10.1017/S1755267209990868

Bornatowski, H., Braga, R. R., Kalinowski, C., and Vitule, J. R. S. (2015). "Buying a pig in a poke": the problem of elasmobranch meat consumption in Southern Brazil. *Ethnobiol. Lett.* 6, 196–202. doi: 10.14237/ebl.6.1.2015.451

Bornatowski, H., Braga, R. R., and Vitule, J. R. S. (2014a). Threats to sharks in a developing country: the need for effective simple conservation measures. *Nat. Conserv.* 12, 11–18. doi: 10.4322/natcon.2014.003

Bornatowski, H., Costa, L., Robert, M. C., and Pina, J. V. (2007). Hábitos alimentares de tubarões-martelo jovens, *Sphyrna zygaena* (Carcharhiniformes: Sphyrnidae), no litoral sul do Brasil. *Biota Neotrop.* 7, 213–216. doi: 10.1590/S1676-06032007000100025

Bornatowski, H., Navia, A. F., Braga, R. R., Abilhoa, V., and Corrêa, M. F. M. (2014b). Ecological importance of sharks and rays in a structural foodweb analysis in southern Brazil. *ICES J. Mar. Sci.* 71, 1586–1592. doi: 10.1093/icesjms/fsu025

Camhi, M. D., Valenti, S. V., Fordham, S. V., Fowler, S. L., and Gibson, C. (2009). *The Conservation Status of Pelagic Sharks and Rays: Report of the IUCN Shark Specialist Group Pelagic Shark Red List Workshop*. Newbury: IUCN Species Survival Commission Shark Specialist Group, 78.

Carvalho, D. C., Neto, D. A., Brasil, B. S., and Oliveira, D. A. (2011). DNA barcoding unveils a high rate of mislabeling in a commercial freshwater catfish from Brazil. *Mitochondrial DNA* 22, 97–105. doi: 10.3109/19401736.2011.588219

Cosandey-Godin, A., and Morgan, A. (2011). *Fisheries Bycatch of Sharks: Options for Mitigation*. Washington, DC: Ocean Science Division; Pew Environment Group.

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772. doi: 10.1038/nmeth.2109

Domingues, R. R. (2011). *Identificação Molecular, Biologia e Pesca de Tubarões do Gênero Carcharhinus (Chondrichthyes – Carcharhiniformes): Uma Contribuição*

Para a Gestão de Pesca do Estado de São Paulo, Brasil. Master's thesis, Instituto de Pesca, São Paulo.

Doyle, J. J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19, 11–15.

Dudgeon, C. L., Blower, D. C., Broderick, D., Giles, J. L., Holmes, B. J., Kashiwagi, T., et al. (2012). A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *J. Fish Biol*. 80, 1789–1843. doi: 10.1111/j.1095-8649.2012.03265.x

Dulvy, N. K., Sadovy, Y., and Reynolds, J. D. (2003). Extinction vulnerability in marine populations. *Fish Fish*. 4, 25–64. doi: 10.1046/j.1467-2979.2003.00105.x

Ekrem, T., Willassen, E., and Stur, E. (2007). A comprehensive DNA sequence library is essential for identification with DNA barcodes. *Mol. Phylogenet. Evol*. 43, 530–542. doi: 10.1016/j.ympev.2006.11.021

Food and Agriculture Organization of United Nations (FAO) (2014). *The State of World Fisheries and Aquaculture* 2014. Rome: FAO.

Food and Agriculture Organization of United Nations (FAO) (2015). *State of the Global Market for Shark Products*. Rome: FAO.

Food and Agriculture Organization of United Nations (FAO) (2016). *The State of World Fisheries and Aquaculture 2016. Contributing to Food Security and Nutrition for All*. Rome: FAO.

Figueiredo, J. L. (1977). *Manual de Peixes Marinhos do Sudeste do Brasil–I. Introdução Cações, Raias e Quimeras*. São Paulo: Museu de Zoologia; Universidade de São Paulo.

Filonzi, L., Chiesa, S., Vaghi, M., and Marzano, F. N. (2010). Molecular barcoding reveals mislabelling of commercial fish products in Italy. *Food Res. Int*. 43, 1383–1388. doi: 10.1016/j.foodres.2010.04.016

Frézal, L., and Leblois, R. (2008). Four years of DNA barcoding: current advances and prospects. *Infect. Genet. Evol.* 8, 727–736. doi: 10.1016/j.meegid.2008.05.005

Gadig, O. B. F. (2001). *Tubarões da Costa Brasileira*. Ph.D. thesis, Instituto de Biociências, Rio Claro.

Galimberti, A., De Mattia, F., Losa, A., Bruni, I., Federici, S., Casiraghi, M., et al. (2013). DNA barcoding as a new tool for food traceability. *Food Res. Int*. 50, 55–63. doi: 10.1016/j.foodres.2012.09.036

Gallagher, A. J., and Hammerschlag, N. (2011). Global shark currency: the distribution, frequency, and economic value of shark ecotourism. *Curr. Issues Tour*. 14, 797–812. doi: 10.1080/13683500.2011.585227

Gallagher, A. J., Kyne, P. M., and Hammerschlag, N. (2012). Ecological risk assessment and its application to elasmobranch conservation and management. *J. Fish Biol*. 80, 1727–1748. doi: 10.1111/j.1095-8649.2012.03235.x

Haddad, V. Jr., and Gadig, O. B. F. (2005). The spiny dogfish ('cação-bagre'): description of an envenoming in a fisherman, with taxonomic and toxinologic comments on the Squalus gender. *Toxicon* 46, 108–110. doi: 10.1016/j.toxicon.2005.03.002

Hajibabaei, M., Singer, G. A., Hebert, P. D., and Hickey, D. A. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends Genet*. 23, 167–172. doi: 10.1016/j.tig.2007.02.001

Hebert, P. D., Cywinska, A., and Ball, S. L. (2003). Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B Biol. Sci*. 270, 313–321. doi: 10.1098/rspb.2002.2218

Heithaus, M. R., Frid, A., Wirsing, A. J., and Worm, B. (2008). Predicting ecological consequences of marine top predator declines. *Trends Ecol. Evol.* 23, 202–210. doi: 10.1016/j.tree.2008.01.003

Holmes, B. H., Steinke, D., and Ward, R. D. (2009). Identification of shark and ray fins using DNA barcoding. *Fish. Res.* 95, 280–288. doi: 10.1016/j.fishres.2008.09.036

Instituto Chico Mendes de Preservação da Biodiversidade (2016). *Brazil Red Book of Threatened Species of Fauna*. Available online at: http://www.icmbio.gov.br/portal/images/stories/comunicacao/publicacoes/publicacoes-diversas/dcom_sumario_executivo_livro_vermelho_da_fauna_brasileira_ameacada_de_extincao_2016.pdf (Accessed April 11, 2018).

IUCN (2017). *The IUCN Red List of Threatened Species*. Available online at: http://www.iucnredlist.org/ (Accessed October 01, 2016).

Jacquet, J. L., and Pauly, D. (2008). Trade secrets: renaming and mislabeling of seafood. *Mar. Policy* 32, 309–318. doi: 10.1016/j.marpol.2007.06.007

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kotas, J. E. (2004). *Dinâmica de Populações e pesca do Tubarão-martelo Sphyrna lewini (Griffith & Smith, 1834), Capturado no mar Territorial e zona Econômica Exclusiva do Sudeste-sul do Brasil.* Doctoral thesis, Universidade de São Paulo, São Paulo.

Kotas, J. E., Petrere, M. Jr., Azevedo, V. D., and Santos, S. D. (2005). *A Pesca de Emalhe e de Espinhel-de-Superfície na Região Sudeste-Sul do Brasil*. Instituto Oceanográfico-USP.

Kotas, J. E., Petrere Junior, M., Fiedler, F., Mastrochirico, V., and Sales, G. (2008). A pesca de emalhe-de-superficie de Santa Catarina direcionada à captura dos tubarões-martelo, *Sphyrna lewini* (Griffith & Smith 1834) e *Sphyrna zygaena* (Linnaeus 1758). *Atlântica* 30, 113–128. doi: 10.5088/atlântica.v30i2.1511

Lack, M., and Sant, G. (2006). *World Shark Catch, Production and Trade 1990-2003*. Department of the Environment and Heritage.

Lowenstein, J. H., Burger, J., Jeitner, C. W., Amato, G., Kolokotronis, S. O., and Gochfeld, M. (2010). DNA barcodes reveal species-specific mercury levels in tuna sushi that pose a health risk to consumers. *Biol. Lett.* 6, 692–695. doi: 10.1098/rsbl.2010.0156

Lucifora, L. O., García, V. B., and Worm, B. (2011). Global diversity hotspots and conservation priorities for sharks. *PLoS ONE* 6:e19356. doi: 10.1371/journal.pone.0019356

Mader, A., Sander, M., Casa, G. E. Jr., Altenhofen, R. J., and dos Anjos, C. S. (2007). Evidências de sobrepesca do tubarão martelo (*Sphyrna* spp.) no Rio Grande do Sul, Brasil. *Biodivers. Pampeana* 5.

Martinez, I., James, D., and Loréal, H. (2005). *Application of Modern Analytical Techniques to Ensure Seafood Safety and Authenticity*. FAO Fisheries Technical Paper. Food and Agriculture Organization of United Nations, Rome.

Mendonça, F. F., Hashimoto, D. T., Porto-Foresti, F., Oliveira, C., Gadig, O. B. F., and Foresti, F. (2009). Identification of the shark species *Rhizoprionodon lalandii* and *R. porosus* (Elasmobranchii, Carcharhinidae) by multiplex PCR and PCR-RFLP techniques. *Mol. Ecol. Resour*. 9, 771–773. doi: 10.1111/j.1755-0998.2009.02524.x

Mendonça, F. F., Ussami, L. H., Hashimoto, D. T., Pereira, L. H., Porto-Foresti, F., Oliveira, C., et al. (2012). Identification and characterization of polymorphic microsatellite loci in the blue shark *Prionace glauca*, and cross-amplification in other shark species. *J. Fish Biol*. 80, 2643–2646. doi: 10.1111/j.1095-8649.2012.03291.x

Meyer, C. P., and Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol*. 3:e422. doi: 10.1371/journal.pbio.0030422

Moftah, M., Aziz, S. H., Elramah, S., and Favereaux, A. (2011). Classification of sharks in the Egyptian Mediterranean waters using morphological and DNA barcoding approaches. *PLoS ONE* 6:e27001. doi: 10.1371/journal.pone.0027001

Moretti, V. M., Turchini, G. M., Bellagamba, F., and Caprino, F. (2003). Traceability issues in fishery and aquaculture products. *Vet. Res. Commun*. 27, 497–505. doi: 10.1023/B:VERC.0000014207.01900.5c

MPA. Boletim estatístico da pesca e aquicultura (2011). *Relatório do Ministério da Pesca e Aquicultura,* 98. Available online at: http://www.icmbio.gov.br/cepsul/images/stories/biblioteca/download/estatistica/est_2011_bol__bra.pdf

Neto, D. A. P. (2013). *Detecção de Adulteração de Espécies em Pescado e Derivados por Meio da Técnica de DNA Barcoding*, Master's thesis, Universidade Federal de Minas Gerais, Belo Horizonte.

Palmeira, C. A. M., Rodrigues-Filho, L. F., Sales, J. P. L., Vallinoto, M., Schneider, H., and Sampaio, I. (2013). Commercialization of a critically endangered species (largetooth sawfish, Pristis perotteti) in fish markets of northern Brazil: authenticity by DNA analysis. *Food Control* 34, 249–252. doi: 10.1016/j.foodcont.2013.04.017

Pauly, D., Hilborn, R., and Branch, T. A. (2013). Fisheries: does catch reflect abundance? *Nature* 494, 303–306. doi: 10.1038/494303a

Pérez-Jiménez, J. C., Rocha-Olivares, A., and Sosa-Nishizaki, O. (2013). Morphological and molecular differentiation of smooth-hound sharks (Genus *Mustelus*, Family Triakidae) from the Gulf of California. *J. Appl. Ichthyol.* 29, 268–270. doi: 10.1111/jai.12042

Portaria (2014). *MMA N° 443, de 17 de Dezembro de 2014*. Available online at: http://www.icmbio.gov.,br/cepsul/images/stories/legislacao/Portaria/2014/p_m ma_443_2014_lista_esp%C3%A9cies_amea%C3%A7adas_extin%C3%A7%C3 %A3o.pdf (Accessed September 03, 16)

Rasmussen, R. S., and Morrissey, M. T. (2008). DNA-based methods for the identification of commercial fish and seafood species. *Compr. Rev. Food Sci. Food Saf.* 7, 280–295. doi: 10.1111/j.1541-4337.2008.00046.x

Rodrigues-Filho, L. F., Pinhal, D., Sodré, D., and Vallinoto, M. (2012). "Shark DNA forensics: Applications and impacts on genetic diversity," in *Analysis of Genetic Variation in Animals*, ed M. Caliskan (Rijeka: InTech), 269–286.

Rodrigues-Filho, L. F., Rocha, T. C., Rêgo, P. S., Schneider, H., Sampaio, I., and Vallinoto, M. (2009). Identification and phylogenetic inferences on stocks of sharks affected by the fishing industry off the Northern coast of Brazil. *Genet. Mol. Biol.* 32, 405–413. doi: 10.1590/S1415-47572009005000039

Rose, D. A. (1996). *An Overview of World Trade in Sharks and Other Cartilaginous Fishes.* Traffic International.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Swofford, D. L. (2002). *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sunderland, MA: Sinauer Associates.

Teletchea, F. (2009). Molecular identification methods of fish species: reassessment and possible applications. *Rev. Fish Biol. Fish*. 19:265. doi: 10.1007/s11160-009-9107-4

Triant, D. A., and DeWoody, J. A. (2007). The occurrence, detection, and avoidance of mitochondrial DNA translocations in mammalian systematics and phylogeography. *J. Mammal.* 88, 908–920. doi: 10.1644/06-MAMM-A-204R1.1

van der Elst, R. P. (1979). A proliferation of small sharks in the shore-based Natal sport fishery. *Environ. Biol. Fish*. 4, 349–362. doi: 10.1007/BF00005524

Vannuccini, S. (1999). *Shark utilization, Marketing and Trade*. No. 389. Food & Agriculture Org. Rome: FAO.

Vaz, D. F., and De Carvalho, M. R. (2013). Morphological and taxonomic revision of species of *Squatina* from the Southwestern Atlantic Ocean (Chondrichthyes: Squatiniformes: Squatinidae). *Zootaxa* 3695, 1–81. doi: 10.11646/zootaxa.3695.1.1

Vooren, C. M., and Klippel, S. (2005). *Ações Para a Conservação de Tubarões e Raias no Sul do Brasil*. Porto Alegre: Igaré.

Ward-Paige, C. A., Keith, D. M., Worm, B., and Lotze, H. K. (2012). Recovery potential and conservation options for elasmobranchs. *J. Fish Biol.* 80, 1844–1869. doi: 10.1111/j.1095-8649.2012.03246.x

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., and Hebert, P. D. (2005). DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 360, 1847–1857. doi: 10.1098/rstb.2005.1716

Worm, B., Davis, B., Kettemer, L., Ward-Paige, C. A., Chapman, D., Heithaus, M. R., et al. (2013). Global catches, exploitation rates, and rebuilding options for sharks. *Mar. Policy* 40, 194–204. doi: 10.1016/j.marpol.2012.12.034

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 24 July 2018
doi: 10.3389/fgene.2018.00271

Hidden Diversity Hampers Conservation Efforts in a Highly Impacted Neotropical River System

Naiara G. Sales[1], Stefano Mariani[1], Gilberto N. Salvador[2], Tiago C. Pessali[3] and Daniel C. Carvalho[4]*

[1] Ecosystems and Environment Research Centre, School of Environment & Life Sciences, University of Salford, Salford, United Kingdom, [2] Laboratório de Ecologia e Conservação, Universidade Federal do Pará, Belém, Brazil, [3] Museu de Ciências Naturais PUC Minas, Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, Brazil, [4] Programa de Pós-graduação em Biologia de Vertebrados, Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, Brazil

Neotropical Rivers host a highly diverse ichthyofauna, but taxonomic uncertainty prevents appropriate conservation measures. The Doce River Basin (DRB), lying within two Brazilian threatened hotspots (Atlantic Forest and Brazilian Savanna) in south-east Brazil, faced the worst ever environmental accident reported for South American catchments, due to a dam collapse that spread toxic mining tailings along the course of its main river. Its ichthyofauna was known to comprise 71 native freshwater fish species, of which 13 endemic. Here, we build a DNA barcode library for the DRB ichthyofauna, using samples obtained before the 2015 mining disaster, in order to provide a more robust biodiversity record for this basin, as a baseline for future management actions. Throughout the whole DRB, we obtained a total of 306 barcodes, assigned to 69 putative species (with a mean of 4.54 barcodes per species), belonging to 45 genera, 18 families, and 5 orders. Average genetic distances within species, genus, and families were 2.59, 11.4, and 20.5%, respectively. The 69 species identified represent over 76% of the known DRB ichthyofauna, comprising 43 native (five endemic, of which three threatened by extinction), 13 already known introduced species, and 13 unknown species (such as *Characidium* sp., *Neoplecostomus* sp., and specimens identified only at the sub-family level Neoplecostominae, according to morphological identification provided by the museum collections). Over one fifth of all analyzed species (N = 16) had a mean intraspecific genetic divergence higher than 2%. An integrative approach, combining NND (nearest neighbor distance), BIN (barcode index number), ABGD (automatic barcode gap discovery), and bPTP (Bayesian Poisson Tree Processes model) analyses, suggested the occurrence of potential cryptic species, species complex, or historical errors in morphological identification. The evidence presented calls for a more robust, DNA-assisted cataloging of biodiversity-rich ecosystems, in order to enable effective monitoring and informed actions to preserve and restore these delicate habitats.

Keywords: barcode, biodiversity, cryptic diversity, Doce River, ichthyofauna, molecular identification

Abbreviations: BIN, barcode index number; BOLD, Barcode of Life Data System; DRB, Doce River Basin; NND, nearest neighbor distance.

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

Reviewed by:
Fernanda Simões De Almeida, Universidade Estadual de Londrina, Brazil
Carolina Machado, Universidade Federal de São Carlos, Brazil

****Correspondence:***
Naiara G. Sales
n.g.sales@edu.salford.ac.uk;
naiarasl@hotmail.com

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

Received: *01 November 2017*
Accepted: *03 July 2018*
Published: *24 July 2018*

Citation:
Sales NG, Mariani S, Salvador GN, Pessali TC and Carvalho DC (2018) Hidden Diversity Hampers Conservation Efforts in a Highly Impacted Neotropical River System. Front. Genet. 9:271. doi: 10.3389/fgene.2018.00271

INTRODUCTION

Neotropical rivers host an extremely diverse ichthyofauna, but anthropogenic impact associated with the occurrence of many still undescribed or unknown species may hamper conservation effort (Reis et al., 2016; Ely et al., 2017). Due to increasing, rapid anthropogenic environmental impacts (e.g., pollution, siltation, mining, damming), biodiversity in Neotropical rivers may be lost before scientists can fully describe and comprehend it (Agostinho et al., 2005).

Effective biodiversity conservation relies on unequivocal and precise species identification, especially in the case of ecosystems that underwent degradation and require restoration. However, high biodiversity regions, such as the neotropics, and the increasingly reduced budget for basic taxonomical research, have led to the so-called "taxonomic impediment" or "poor taxonomy", in which the shortage of funding and trained taxonomists, and the gaps in taxonomic knowledge, have delayed advances in assessment and description of biodiversity or even contributed to overestimate or underestimate species richness due to species misidentification or taxonomic confusions (Taylor, 1983; Ely et al., 2017).

The DNA barcoding initiative offers a powerful and cost-effective tool to assist with the detection of cryptic species and flag potentially problematic taxa, with the standard universal COI marker having proven particularly successful in invertebrates (Hebert et al., 2004a), birds (Hebert et al., 2004b), and fish (Ward et al., 2005; Hubert et al., 2008; Valdez-Moreno et al., 2009; Carvalho et al., 2011; Rosso et al., 2012). For effective DNA barcode performance, intraspecific variability must be lower than variability among congeneric species, the so-called 'Barcode Gap' (Meyer and Paulay, 2005). While the barcode gap tends to be around <1–2% sequence variability within species in most fish, there are exceptions (Hurst and Jiggins, 2005), especially in the case of recently diverged species (Vinas and Tudela, 2009; Shum et al., 2017). Moreover, the unambiguous identification of species from early larval stage to adulthood can aid a variety of conservation management actions. Accurate molecular identification may contribute to improving management and sustainability of long term fisheries (Metcalf et al., 2007), tracking invasive species (Corin et al., 2007; Carvalho et al., 2009), offer insights into community ecology (Pfenninger et al., 2007) and genetic certification of species used in restocking programs (Metcalf et al., 2007), as well as improving fundamental knowledge on cryptic and putatively new species (Pereira et al., 2011). Furthermore, molecular identification of eggs and larvae can provide data regarding spawning and recruitment areas, supporting a definition of priority areas for conservation (Becker et al., 2015; Frantine-Silva et al., 2015).

DNA barcode libraries have been developed for several Neotropical river systems as a biodiversity identification tool, and have contributed to reveal the existence of putatively cryptic/new fish species (Carvalho et al., 2011; Pereira et al., 2011; Gomes et al., 2015; Pugedo et al., 2016; Nascimento et al., 2016). However, the biodiversity complexity remains unknown in many already impacted catchments in Brazil. One emblematic case is that of the Doce River Basin (DRB), which faced the worst environmental accident reported for any South American catchment, in the form of the largest tailings dam burst in modern history; as a result, a toxic mud (i.e., extreme high concentration of iron) spread along its main river course, affecting wild communities, as well as the local human populations (Fernandes et al., 2016; Neves et al., 2016). As the local riverine human communities rely on fisheries for their livelihood (e.g., source of income and subsistence, resource for ecotourism), understanding the impacts of this disaster on the ichthyofauna is crucial for effective management actions (Ecoplan-Lume, 2010; GFT, 2015; Neves et al., 2016). Moreover, the recovery of fish populations in DRB, after the ecological disaster, relies on the recolonization of the main course of this river and on the diversity, size, and conservation status of the remnant fish populations in the tributaries (Olds et al., 2012).

The DRB runs through two Brazilian biodiversity hotspots (Atlantic forest and Brazilian Savanna) located in south-east Brazil (Myers et al., 2000). The river is 853 km long and the catchment covers a total drainage area of 83.400 km^2, between the states of Minas Gerais (86%) and Espírito Santo (14%), an area inhabited by three million people. DRB harbors a rich ichthyofauna, including several undescribed species, with the number of presently recognized native species summing up to 71 (Vieira, 2009). The Santo Antônio River, the second largest tributary of the Doce, was selected as a conservation priority area, since it hosts a great number of species considered endemic and threatened by extinction (Vieira et al., 2000; Vieira and Alves, 2001; Rosa and Lima, 2005). Historically, DRB is affected by human impacts by many ways. Native forest cover only 27% of DRB area (ANA, 2016), and the remained area is used to cattle, forestry, agriculture, and mining (Vieira, 2009), resulting in high rate of siltation (da Silva et al., 2011). Habitat fragmentation lead by hydroelectric construction is also affecting DRB, where there are 40 hydroelectric built along main channel of Doce River and its principal tributaries (ANEEL, 2010). However, without accurate biodiversity knowledge, species conservation may be hindered in this river system, and it had already been suggested that the environmental disaster involving the mining collapse could have led to the depletion/extinction of many still unknown endemic species (Fernandes et al., 2016). Here, we develop a DNA barcode library for the DRB ichthyofauna, using data obtained prior to the dam burst environmental disaster, contributing to an improved biodiversity baseline record for this recently impacted ecosystem.

MATERIALS AND METHODS

Sampling

We obtained fish tissue samples from 306 specimens collected between 2011 and 2015 along the main river channel and tributaries (**Figure 1**), identified and deposited by taxonomists in four Brazilian ichthyological collections: PUC Minas Natural History Museum (MCNIP), Museu de Biologia Professor Mello Leitão (MBML), Museu de Zoologia da Universidade Estadual de Campinas (ZUEC), and Núcleo de Pesquisas em Limnologia,

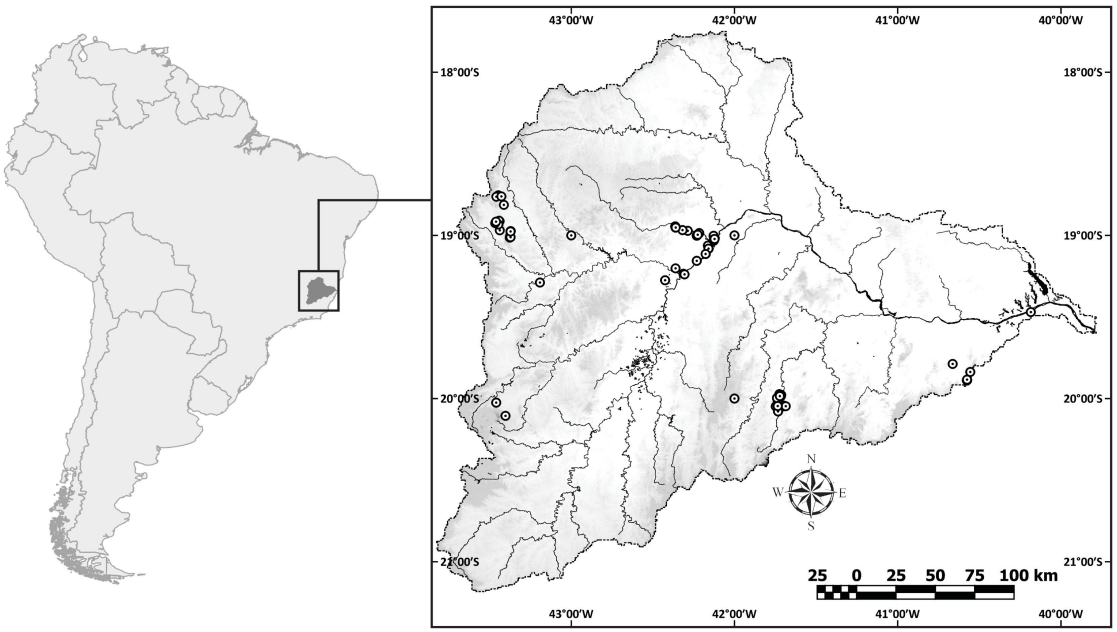


FIGURE 1 | Map of Doce River Basin, including sample sites distribution.

Ictiologia e Aquicultura (NUPELIA). All analyzed specimens were photographed, geo-referenced, and identified to the lowest taxonomic level from identification keys or previously published works (Vari, 1992; Albert et al., 1999; Castro and Vari, 2004; Zanata and Camelier, 2009).

Ethics Statement

All fish analyzed in this study were collected in accordance with Brazilian legislation (Collection license 6421-1, number 5498740) or obtained from Ichthyological collections. Fish were collected, and euthanized; samples of fins were clipped from each individual and stored in absolute ethanol for subsequent molecular analysis. Specimens were fixed in 10% formaldehyde and then stored in 70% ethanol.

DNA Extraction, Amplification, and Sequencing

Genetic analyses were conducted, whenever possible, on a minimum of five specimens from different sample sites per species. DNA extraction followed the salting out protocol (adapted from Aljanabi and Martinez (1997)). The cytochrome c oxidase I (COI) gene (650 bp) was amplified by polymerase chain reaction (PCR) using the primers FishF1/FishR1 described by Ward et al. (2005) and the Cocktail COI-3/C_FishF1t1-C_FishR1t1 described by Ivanova et al. (2007), and following the PCR protocol described in Gomes et al. (2015). The PCR products were visualized on 1% agarose gel, alongside negative controls and a size ladder, and positive amplifications were selected for DNA sequencing. DNA sequencing was conducted in both directions in an automated DNA analyzer ABI 3500 (Life Technologies).

Data Analysis

Barcode sequences were edited using DNA Baser® v.3.5.4 (DNA Sequence Assembler v4 (2013), Heracle BioSoft[1]) and SeqScape v.2.1.1 (Applied Biosystems, Foster City, CA, United States) (Díaz et al., 2016) softwares. DNA alignment was conducted using the CLUSTAL W alignment tool (Thompson et al., 1997). The neighbor-joining (NJ) trees (Saitou and Nei, 1987) and genetic distances estimations, using the K2P (Kimura-2-parameter) nucleotide evolution model (Kimura, 1980) were generated using MEGA 7 software (Kumar et al., 2016).

Intra- and inter-specific genetic distances, nearest neighbor distance (NND), and the barcode gap were calculated in the on-line Barcode of Life Data System (BOLD) Workbench[2] (Ratnasingham and Hebert, 2007). The NND was used to estimate the minimum genetic distance between pairs of species. Different approaches were used to delimitate the Molecular Operational Taxonomic Units (MOTUs), two clustering algorithms [barcode index number (BIN) and Automatic Barcode Gap Discovery (ABGD)] and one phylogenetic-coalescent method Bayesian Poisson Tree Processes model (bPTP). The BIN (Ratnasingham and Hebert, 2013) was estimated automatically in BOLD Workbench and allowed comparing DNA barcodes obtained here with other river basins that have a comprehensive DNA Barcode library, such as the São Francisco, the Mucuri, the Jequitinhonha, the Paraná, and the Paranaíba River Basins (Carvalho et al., 2011; Pereira et al., 2011; Gomes et al., 2015; Díaz et al., 2016; Pugedo et al., 2016). Using this approach, it is possible to identify endemic lineages and shared ichthyofauna. ABGD analyses (Puillandre et al., 2012)

[1] www.DnaBaser.com

[2] http://www.boldsystems.org

were performed using the web interface ([3]web version 'May 31 2017') with a relative gap width value of $X = 1.0$ and two available distance metrics [JC69 (Jukes and Cantor, 1969) and K2P (Kimura, 1980)], while the other parameter values employed default settings. The bPTP was conducted using both ML (maximum likelihood) and Bayesian approaches (Zhang et al., 2013). The PTP file input consisted in a nexus tree generated in MrBayes (Ronquist and Huelsenbeck, 2003) using six random parsimony trees, with the GTRGAMMA substitution model (obtained by MEGA 7 under BIC criteria), without rooting and applying the parameters of 20 million MCMC generations and a burn-in of 10%. Analysis was conducted applying default values through the bPTP server (500,000 generations, thinning = 100, burn-in = 10%).

All data, including fish photos, GPS coordinates of each sample site, vouchers numbers, detailed taxonomic identifications, and the corresponding sequence data and trace files were submitted to the BOLD ([2]see Ratnasingham and Hebert, 2007) within the project file 'DNA Barcoding of DRB'.

Species Delimitation and Hidden Biodiversity

Species delimitation based on integrative approaches that combine a diverse range of statistical methods has been extensively used to identify hidden biodiversity (i.e., Padial et al., 2010; Costa-Silva et al., 2015; Gomes et al., 2015; Rossini et al., 2016; Ramirez et al., 2017). Here, species with >2% of intraspecific genetic divergences, still undescribed or unknown and identified only at genus or family level were investigated individually to detect the occurrence of new molecular operational taxonomic units (MOTUs) according to the congruence among BIN, ABGD, bPTP outputs.

Undescribed species or those only identified at genus or family level were checked using the BIN and NND analyses in order to verify their occurrence in clusters composed by other nominal species, and their genetic divergence from the nearest neighbor (including species from DRB and/or distinct Brazilian basins). Were considered as new MOTUs when intraspecific genetic divergence was higher than 2% for described species and distinguished clusters identified by BIN, ABGD, and bPTP outputs.

RESULTS

Morphological identification on the 306 specimens yielded 69 species (see Supplementary Table S1) in of which 43 are native species (five endemic, three threatened by extinction and one endemic and threatened), 13 non-native species and 13 unknown species to the DRB (see Supplementary Table S2), representing over 76% of its known freshwater ichthyofauna (Vieira, 2009). We then obtained 306 partial sequences of the COI gene, consisting of 665 bp on average, and no insertions, deletions, or stop codons were detected, indicating that there was no case of NUMTS (Nuclear mitochondrial DNA sequences) (Song et al., 2008; Hazkani-Covo et al., 2010).

[3]http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html

TABLE 1 | Distance summary reports for sequence divergence between species, genus, and family level, including minimum, mean, and maximum genetic distances (K2P).

	Minimum distance (%)	Mean distance (%)	Maximum distance (%)
Within species	0	2.59	21.82
Within genera	0	11.4	24.2
Within families	0	20.5	30.99

A mean of 4.54 individuals per species were sequenced, comprising 45 genera, 18 families, and 5 orders [Characiformes (41.9%), Siluriformes (40.6%), Perciformes (9.4%), Gymnotiformes (4.7%), and Cyprinodontiformes (3.4%)]. Species represented by one or two specimens ($N = 19$) were not included in the estimation of intraspecific divergences (*Callichthys callichthys*, *Cichla kelberi*, *Clarias gariepinus*, *Hoplosternum littorale*, *Hyphessobrycon bifasciatus*, *H. eques*, *Hypostomus* sp., *Lophiosilurus alexandri*, *Metynnis maculatus*, *Parotocinclus maculicauda*, *Pimelodus maculatus*, *Poecilia vivipara*, *Prochilodus vimboides*, *Pygocentrus nattereri*, *Salminus brasiliensis*, *Steindachneridion doceanum*, *Trichomycterus* aff. *auroguttatus*, *T.* cf. *brasiliensis*, and *T. longibarbatus*). The NJ tree identified species-specific clades for 80.9% of all species. The mean genetic distances found within species, genera, and families were: 2.59, 11.4, and 20.5% (**Table 1**), respectively.

Over 65% of the analyzed species showed genetic distances lower than 1% and for 70% of the species the divergence value was below 2% (**Figure 2A**). When considering intra-generic distance, 19% of the species had a divergence higher than 20% (**Figure 2B**), suggesting the possibility of taxonomic errors or cryptic species.

Intra- and Inter-Specific Divergence

Intraspecific distance varied from 0 to 21.82%. Particularly high genetic distances (>10%) were recovered among specimens of *Astyanax fasciatus* (20.69%), *Astyanax scabripinnis* (21.82%), *Astyanax* sp. (20.5%), *Characidium* sp. (10.17%), *Crenicichla lacustris* (21.36%), *Harttia* sp. (12.2%), *Poecilia reticulata* (14.34%), and *Trichomycterus* aff. *alternatus* (18.49%), flagging possible new MOTUs (i.e., hidden diversity) or problems related with taxonomic morphological identification.

The NJ tree encompassing all species showed the occurrence of monophyletic clades and absence of shared haplotypes for 44 of the 69 analyzed species. The interspecific genetic distance showed that 63.2% of the analyzed species had a K2P divergence higher than 2% to their closest neighbor, with the exception of: *Astyanax* spp., *Deuterodon pedri*, *H. eques*, *Characidium* sp. and *Characidium* gr. *timbuiense*, *Gymnotus* spp., *Oligosarcus argenteus* and *O. acutirostris*, *P. reticulata* and *P. vivipara,* and *T.* aff. *alternatus* and *T. longibarbatus* (Supplementary Figure S1).

Incongruences between morphological and barcode identifications (BIN, ABGD, bPTP) (i.e., one BIN/ABGD/bPTP cluster containing more than one morphological species, morphological species represented by more than one

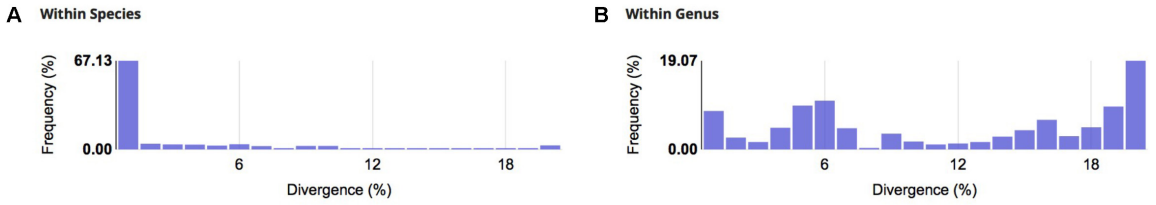


FIGURE 2 | Genetic divergences found for all sequences analyzed at species **(A)** and genus **(B)** levels.

BIN/ABGD/bPTP cluster, and/or >2% of intraspecific genetic distance and <1% of interspecific divergence) were observed within species of the genus *Astyanax*, *Characidium*, *Crenicichla*, *Deuterodon*, *Gymnotus*, *Harttia*, *Hoplias*, *Hyphessobrycon*, *Hypostomus*, *Knodus*, *Neoplecostomus*, *Oligosarcus*, *Pareiorhaphis*, *Poecilia*, *Prochilodus*, *Rhamdia*, and *Trichomycterus* (Supplementary Table S2).

Identification of Molecular Operational Taxonomic Units (MOTUs)

The BIN analysis identified 81 clusters, including 48 taxonomically concordant, 17 discordant, and 16 singletons. The ABGD analysis detected 54–133 MOTUs when varying the prior maximal distance from $P = 0.001$ to $P = 0.1000$ (applying both the K2P and JC69 nucleotide evolution methods). The partition that recovered 81 groups (intraspecific distance $P = 0.0077$) was chosen due to its consistency with our BIN analysis. The bPTP analyses (Bayesian and ML approaches) resulted in the same number of clusters obtained by BIN, except for *Harttia* sp. (three BIN and ABGD clusters and one bPTP) and *Prochilodus costatus* (two BIN, and one ABGD and bPTP clusters). ABGD species delineation was in agreement with all the BIN clusters with the following exceptions, which contain more than one BIN for each morpho-species: *A. scabripinnis* (BIN: AAC5910, ABGD: 36 and 81), *Knodus moenkhausii* (BIN: AAM1485, ABGD: 46 and 49), *P. costatus* (BIN: ADC2568 and ADC2571, ABGD: 10), *Trichomycterus* sp./*T.* aff. *alternatus*/*T.* aff. *auroguttatus*/*T. longibarbatus* (BIN: ACJ1164 and ACJ1161, ABGD: 64), *Trichomycterus* sp./*T.* cf. *brasiliensis* (BIN: ACK5393 and ACT6325, ABGD: 65) (Supplementary Table S2).

Identification of Hidden Biodiversity

Sequences from fifteen undescribed species or identified only at genus or family level were compared to other species available in BOLD database through NND and BIN analyses (**Table 2**). Within undescribed or unknown species, we recovered new MOTUs from the following genera: *Astyanax*, *Characidium, Gymnotus, Harttia, Hisonotus, Neoplecostomus, Pareiorhaphis, Phalloceros*, and *Trichomycterus*. The other six species were not considered new MOTUs (*Brycon* sp., *Hasemania* sp., *Hypostomus* sp., *Imparfinis* sp., Neoplecostominae, and *Pimelodella* sp.) since they were included in BINs composed by another nominal species and showed interspecific divergence <2% with the nearest neighbor.

Among species with deep intraspecific divergence (>2%) we recovered additionally at least three putative cryptic species due to the congruence among BIN, ABGD, bPTP, and genetic distance methods for *C. lacustris*, *Hoplias malabaricus*, and *Rhamdia* cf. *quelen* (**Table 3**). *A. fasciatus* and *A. scabripinnis* despite showing a congruence of BIN and ABGD analyses were included in clusters comprising another species of the genus. *K. moenkhausii* had a maximum intraspecific divergence of 3.07% and two distinct ABGD numbers, however, only one clade and one BIN was recovered for this species. *Astyanax lacustris*, *A. taeniatus*, *P. reticulata*, *P. costatus*, *T.* aff. *alternatus*, and *T.* aff. *immaculatus* despite showing a high intraspecific genetic distance were included in BINs comprised by another nominal species and thus, were not considered as putative cryptic species.

DISCUSSION

DNA Barcoding Effectiveness

We analyzed 306 fish specimens obtained before the dam burst in 2015 and provided genetic data for the ichthyofauna of the DRB, highlighting the occurrence of cryptic and previously unrecognized biodiversity. Therefore, we significantly extend the knowledge on this river system, whose previous surveys mostly focused on the middle course of the river and in lakes located inside the Doce State Park and its surroundings (Sunaga and Verani, 1987; Vieira, 1994; Vono and Barbosa, 2001; Latini and Petrere, 2004). This baseline offers a more robust platform for any future attempt to restore biodiversity and ecosystem functions to a level comparable to pre-disaster conditions.

Using DNA barcoding, we observed an intraspecific genetic distance considerably higher than previously reported for freshwater fish species from other Brazilian basins. On the other hand, intrageneric divergences were found to be similar to previous studies (Carvalho et al., 2011; Pereira et al., 2011; Pugedo et al., 2016). These results suggest a higher occurrence of hidden biodiversity in DRB when compared to other studied Brazilian basins (**Table 4**).

Hidden Biodiversity

DNA barcoding has already been used to reveal hidden biodiversity, such as cryptic species and new candidate fish species in the São Francisco (Carvalho et al., 2011), Mucuri (one species – Gomes et al., 2015), and Jequitinhonha (15 species – Pugedo et al., 2016) River catchments. In DRB, from 69 morphologically identified species, the barcode analyses

TABLE 2 | List of undescribed species, including the nearest neighbor, BIN, and genetic similarity (%).

Species	NND (nearest neighbor species)	BIN	BIN classification	Maximum similarity (%)
Astyanax sp.	*Astyanax fasciatus*	AAC5910	Discordant	99.2
	Deterodon pedri	ACJ9650	Discordant	99
	Astyanax intermedius	ACT0040	Singleton	93.7
	Astyanax fasciatus, A. bockmanni	AAY4812	Discordant	99.2
Brycon sp.	*Brycon ferox*	ACH8616	Concordant	100
Characidium sp.	*Characidium* sp.	ACS9348	Concordant	100
	Characidium cf. *timbuiense*	ACJ1226	Discordant	100
	Characidium cf. *timbuiense*	ACI3743	Discordant	100
Gymnotus sp.	*Gymnotus carapo*	AAB6216	Discordant	100
	Gymnotus sylvius	AAB6212	Concordant	100
	Gymnotus sp.	ACT0768	Concordant	100
Harttia sp.	*Harttia sp.*	ACJ1000	Singleton	100
	Harttia sp.	ACI6845	Concordant	100
	Harttia sp.	ACO6155	Singleton	100
Hasemania sp.	*Hasemania hanseni*	AAO6055	Concordant	100
Hisonotus sp.	*Hisotonus* sp.	ACW1732	Concordant	100
Hypostomus sp.	*Hypostomus auroguttatus*	AAB9690	Discordant	100
	Hypostomus heraldoi			98.52
	Hypostomus luetkeni			99.51
	Hypostomus strigaticeps			99.01
Imparfinis sp.	*Imparfinis minutus*	AAC2103	Concordant	99.28
	Imparfinis mirini			98.98
Neoplecostomus sp.	*Neoplecostomus* sp.	AAX6581	Concordant	100
	Neoplecostomus sp.	ACT2675	Concordant	100
Neoplecostominae	*Pareiohaphis* cf. *bahianus*	ACC0721	Concordant	98.3
Pareiorhaphis sp.	*Pareiohaphis scutula*	AAX0824	Discordant	99.8
	Pareiorhaphis sp.	ACI5663	Concordant	100
Phalloceros sp.	*Phalloceros* sp.	AAB7265	Concordant	100
Pimelodella sp.	*Pimelodella lateristriga*	AAC5327	Concordant	99.85
Trichomycterus sp.	*Trichomycterus* aff. *immaculatus/T.* cf. *pradensis*	ACI3868	Discordant	99.26
	Trichomycterus aff. *auroguttatus*	ACJ1164	Discordant	100
	Trichomycterus sp.	ACJ9705	Singleton	98
	Trichomycterus cf. *brasiliensis*	ACK5393	Singleton	98.57
	Trichomycterus cf. *brasiliensis*	ACT6325	Discordant	99.8

TABLE 3 | List of described species with high intraspecific divergence (>2%), showing the maximum and mean intraspecific genetic distance, clades and number of BIN, ABGD, and bPTP clusters.

Species	Maximum genetic distance (%)	Mean genetic distance (%)	Clades	BIN	ABGD	bPTP
Astyanax fasciatus	20.69	10.09	3	3	3	3
Astyanax lacustris	3.35	1.67	2	2	2	2
Astyanax scabripinnis	21.82	9.12	2	2	3	2
Astyanax taeniatus	3.96	1.48	2	2	2	2
Characidium sp./*Characidium* cf. *timbuiense**	10.17/9.9	5.51/5.98	4	4	4	4
*Crenicichla lacustris**	21.36	10.76	2	2	2	2
*Hoplias malabaricus**	6.7	3.27	2	2	2	2
Knodus moenkhausii	3.07	1.21	1	1	2	1
Poecilia reticulata	14.34	9.48	2	2	2	2
Prochilodus costatus	2.6	1.32	2	1	1	2
Rhamdia cf. *quelen**	3.48	1.25	2	2	2	2
Trichomycterus aff. *alternatus*	18.49	10.8	2	2	2	2
Trichomycterus aff. *immaculatus*	5.84	2.23	2	2	2	2

**Occurrence of cryptic species.*

TABLE 4 | Comparison among DNA barcoding studies conducted in Brazilian basins, including the number of sequences and species analyzed, and intraspecific and intrageneric distances (minimum and maximum. The mean is inside the parentheses).

Reference	Basin	Number of sequences	Number of species	Intraspecific distance (%)	Intrageneric distance (%)
Pugedo et al., 2016	Jequitinhonha	260	52	0–11.43 (0.44)	1.09–21.55 (12.16)
Nascimento et al., 2016	Itapecuru	440	64	0–8.9 (0.80)	2.65–7.70 (5.13)
Benzaquem et al., 2015*	Amazon	110	14	0–9.8 (2.8)	2.2–22.5 (19.0)
Gomes et al., 2015	Mucuri	141	37	0–3.24 (0.74)	4.29–18.44 (9.5)
Pereira et al., 2013	Upper Paraná	1244	254	0–8.5 (1.3)	0–24.9 (6.8)
Carvalho et al., 2011	São Francisco	431	101	0–10.54 (0.5)	0–22.88 (10.61)
Pereira et al., 2011	Paraíba do Sul	295	58	0–3.48 (0.13)	0.93–22.89 (10.36)
Present study	Doce	306	68	0–21.82 (2.59)	0–24.2 (11.4)

**Only Nannostomus spp.*

recovered 12 putative cryptic species within *Astyanax* sp., *Characidium* sp., *C.* gr. *timbuiense*, *C. lacustris*, *Gymnotus* sp., *Harttia* sp. (two putative cryptic species), *H. malabaricus*, *Neoplecostomus* sp., *R.* cf. *quelen*, *Trichomycterus* sp. (two putative cryptic species). The high intraspecific genetic distance estimation found for the DRB fish was related to the occurrence of cases of well-known species complexes – e.g., *Astyanax* spp. (maximum intraspecific distance reaching 21.82% in *A. scabripinnis*), *Gymnotus* sp. (6.32%), *H. malabaricus* (6.7%), *R.* cf. *quelen* (3.48%) and also due to the deep intraspecific barcode divergence found to putative overlooked cryptic MOTUs – e.g., *C. lacustris* (21.36%).

DNA barcoding allows for the identification of cryptic variation among morphologically similar species, indicating the occurrence of more than one species and reinforcing the need of an integrative approach combining molecular and morphological characters (Nascimento et al., 2016). By combining distinct species delimitation methods, we were able to identify new MOTUs from nine undescribed species *(Astyanax* sp.*, Characidium* sp.*, Gymnotus* sp.*, Harttia* sp.*, Hisonotus* sp.*, Neoplecostomus* sp.*, Pareiorhaphis* sp.*, Phalloceros* sp., and *Trichomycterus* sp.). Other species showed a high similarity with already described species from another river basins (e.g., specimens of *Brycon* sp. were assigned as *B. ferox* from Mucuri River basin) and were not considered as possible new MOTUs (**Table 2**) as shown by the BIN analysis.

Among the undescribed species, we were able to highlight new MOTUs within five morpho-species due to their high intraspecific genetic divergence and based on BIN, ABGD, and NND analyses. For instance, *Harttia* sp. showed mean divergence of 4.67% and three clades, which were congruent within the BIN and ABGD clustering methods, suggesting the occurrence of three new MOTUs in this genus. Specimens of *Hisonotus* sp. were included in the same BIN/ABGD/bPTP cluster and had an exclusive BIN containing only specimens from DRB suggesting a new MOTU exclusive to this catchment. *Neoplecostomus doceensis* is the only loricariid from this genus described for DRB, however, we found two possible cryptic MOTUs within this taxon, as the DNA barcodes from *Neoplecostomus* sp. did not cluster with barcodes available for this species and had two additional distinct BIN and ABGD clusters. Furthermore, exclusive BIN/ABGD clusters were recovered for *Pareiorhaphis* sp. and *Phalloceros* sp. suggesting at least one new MOTU for each genus endemic to the DRB.

Notwithstanding the high intraspecific genetic distance and species delimitation methods detecting more than one MOTUs, we did not consider new MOTUs for species showing a high similarity with another nominal species (e.g., species comprised in the same BIN cluster as another nominal species). Species with high intraspecific divergence were recovered within *Astyanax* spp. (*A. fasciatus*, *A. lacustris*, *A. scabripinnis*, and *A. taeniatus*). Despite showing a deep intraspecific divergence, and congruence of BIN/ABGD clusters, these species were not considered as comprising new MOTUs due to its high genetic similarity with another nominal species (e.g., *Astyanax parahybae*, *A. vermilion*, *Hyphessobrycon* spp., *Deuterodon* sp.) observed within the BIN and NND analysis, and also, because this highly diverse group is a well-known complex of species in need of more systematic studies (Garutti, 1995; Froese and Paulay, 2010; Eschmeyer, 2015).

High intraspecific divergence was also found for *T.* aff. *alternatus* and *T.* aff. *immaculatus*. These species, despite showing a high intraspecific distance (18.49 and 5.84%, respectively), were included in BINs comprised by another nominal species (e.g., *T. longibarbatus*) indicating it may be a case of morphological misidentification and not the occurrence of new MOTUs. This genus has an extensive geographical range and its morphological identification is complex due to the lack of consistent synapomorphies (Barbosa and Costa, 2003). Therefore, further studies combining an integrative approach focusing in these species are required in order to investigate the occurrence of putative cryptic species.

Prochilodus costatus showed a high intraspecific divergence (2.6%) and occurrence of two clusters (NJ and BIN analyses). However, this non-native species was not considered as a putative cryptic species since it was included in BINs comprising another non-native species (e.g., *Prochilodus argenteus*, *P. hartii*). As suggested in previous studies, the incongruence between morphological and molecular identification of *P. costatus* may indicate the occurrence of *Prochilodus* hybrids and not due to new MOTUs (Gomes et al., 2015; Sales et al., 2018).

Poecilia reticulata is a species introduced worldwide, occuring in more than 69 countries outside of its native range (Deacon et al., 2011). A high intraspecific divergence (14.34%) was found

for this species in the DRB. However, two specimens of *P. reticulata* were assigned to a BIN comprising specimens of *P. vivipara* (BIN AAC0279) and the high intraspecific divergence was due to the incongruence between morphological and molecular identification and not due to the occurrence of new MOTUs. Hybridization process between congeneric species of *Poecilia (P. velifera* or *P. petenensis* and *P. mexicana* or *P. orri)* and between different populations of *P. reticulata* have already been reported (Kittell et al., 2005; Lampert and Schartl, 2008; Sievers et al., 2012) and the incongruence detected in this study might be a case of hybridization between *P. reticulata* and *P. vivipara* or misidentification during the deposit in the museum collection and not due to the occurrence of cryptic species.

Hidden biodiversity was found within the genera *Characidium*, *Crenicichla*, *Gymnotus*, *Hoplias*, and *Rhamdia* due to high intraspecific genetic divergence and congruence among clustering methods BIN, ABGD, and bPTP (**Table 3**). Species of the genera *Rhamdia*, *Characidium*, *Pareiorhaphis*, *Gymnotus* were also flagged as cryptic and/or candidate species in other Brazilian basins (Carvalho et al., 2011; Gomes et al., 2015; Pugedo et al., 2016).

For instance, within the genera *Characidium* spp. we detected a mean intraspecific divergence of 5.82% and the occurrence of four clades, of which: two mixed clades comprising specimens identified as *C.* gr. *timbuiense* ($n = 3$ and $n = 4$) and *Characidium* sp. ($n = 1$), one clade exclusive to *C.* gr. *timbuiense* ($n = 1$) and one clade exclusive to *Characidium* sp. ($n = 4$). *C. lacustris* showed intraspecific divergence of 10.76% and presence of two different clades and BIN/ABGD/bPTP clusters (one for samples collected in Manhuaçu River and one for samples collected below the Baguari Dam). The electric knifefishes *Gymnotus* spp. had an intraspecific divergence above 2% and occurrence of three different clades corroborated by 3 BIN, ABGD, and bPTP clusters. All *Gymnotus* specimens were initially morphologically identified as *Gymnotus* sp. and *Gymnotus* cf. *carapo*. However, similarly to the findings obtained for this genus in Mucuri River Basin, these clusters may represent two different known species (*G. carapo* and the overlooked species *Gymnotus sylvius*) and a new MOTU yet to be analyzed and properly described (*Gymnotus* sp.). Two congruent BIN, ABGD, and bPTP clusters were identified for both *H. malabaricus* and *Rhamdia* cf. *quelen* (mean intraspecific divergence of 6.7 and 3.48%, respectively) suggesting the occurrence of cryptic species for each of these taxa. The divergence found in *H. malabaricus* may be due to allopatric speciation resulting from geographical barriers enhanced by its sedentary habitat, since one cluster comprised exclusively specimens from Jose Pedro River and the other was exclusive for specimens from Corrente Grande River. High genetic diversity was already reported for this species in other studied systems (Paraná and Tibagi Rivers) suggesting distinct evolutionary lineages, population structuring or occurrence of cryptic species (Dergam et al., 1998; Blanco et al., 2011; Oliveira et al., 2015).

The increase of available barcodes in BOLD database, including adjacent basins, may contribute to expose endemic cryptic species and reduce the risk of synonymies (Gomes et al., 2015). However, Pugedo et al. (2016) highlighted the concern of using solely DNA barcodes in defining species (e.g., using NND, BIN, ABGD, and bPTP analyses) due to the fact that Neotropical DNA barcode libraries are not yet complete. Furthermore, specimens included in BINs composed by different nominal species should be re-evaluated by a taxonomist to verify the data and check for potential misidentifications (Díaz et al., 2016).

Thus, a thorough analysis should be done for each flagged species to verify the correspondence of new MOTUs with putative new candidate species based on accurate morphological taxonomy analysis and to evaluate the divergence causes and the correlation of speciation process to natural or anthropogenic causes (e.g., presence of dams).

Importance of DNA Barcoding Library for the Doce River Ichthyofauna

This newly developed DNA barcode reference library for the DRB fish detected the occurrence of new MOTUs and suggested the existence of hidden biodiversity. This baseline information will provide a platform for several applications and management efforts, such as ichthyoplankton identification for the detection of fish recruitment areas, unambiguous choice of species to be used in restocking programs, and environmental DNA research. This data may contribute as a baseline for restoration programs in this catchment, by pointing out new MOTUs and suggesting the occurrence of overlooked and cryptic species among the DRB ichthyofauna, highlighting the complexity of Neotropical biodiversity.

The evidence presented here calls for a more robust, DNA-assisted cataloging of biodiversity-rich ecosystems, in order to enable effective monitoring and informed actions to preserve and restore delicate habitats, such as the DRB. Furthermore, studies should verify the extent to which fish biodiversity has been affected by the Doce dam collapse disaster, and what hotspots of diversity within the catchment can be identified as potential sources of replenishment. At the same time, the approaches used here, and additional high through-put methodologies (e.g., metabarcoding of water and sediment samples) should be increasingly employed to monitor biodiversity at a pace that can cater for the management needs of these increasingly impacted biodiverse habitats.

AUTHOR CONTRIBUTIONS

NS performed the molecular genetic analyses and drafted the manuscript. GS and TP collected the samples, conducted the morphological analyses and contributed to the correction of the text. DC designed and coordinates the study. DC and SM conceived the study, participated in its elaboration and helped to draft the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was funded by the Brazilian Barcode of Life Network BrBol/CNPq (564953/2010-5), Coordenação de

Aperfeiçoamento de Pessoal do Ensino Superior (CAPES-PRÓ EQUIPAMENTOS Grant Number 783380/2013), FIP PUC Minas Grant Number 010/2014, FAPEMIG, and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Grant Number 204620/2014-7).

ACKNOWLEDGMENTS

We thank Sergio Alexandre dos Santos, PUC Minas Natural History Museum (MCNIP), Museu de Biologia Professor Mello Leitão (MBML), Museu de Zoologia da Universidade Estadual de Campinas (ZUEC), and Núcleo de Pesquisas em Limnologia, Ictiologia e Aquicultura (NUPELIA) for the samples provided and fish identification. NS is grateful to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for her scholarship (204620/2014-7) and GS is supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00271/full#supplementary-material

REFERENCES

Agostinho, A. A., Thomaz, S. M., and Gomes, L. C. (2005). Conservação da biodiversidade em águas continentais do Brasil. *Megadiversidade* 1, 70–78.

Albert, J. S., Fernandes-Matioli, F. M. C., and Almeida-Toledo, L. F. (1999). New species of gymnotus (Gymnotiformes, Teleostei) from south eastern Brazil: toward the deconstruction of *Gymnotus carapo*. *Copeia* 1999, 410–442. doi: 10.2307/1447486

Aljanabi, S. M., and Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* 25, 4692–4693. doi: 10.1093/nar/25.22.4692

ANA (2016). *Encarte Especial Sobre a Bacia do Rio DOCE: Rompimento Da Barragem Em Mariana*. Technical Report. Brasília: Agência Nacional de Águas.

ANEEL (2010). *Sistema de Informações Georreferenciadas do Setor Elétrico – SIGEL*. Available at: http://sigel.aneel.gov.br [accessed March 2017].

Barbosa, M. A., and Costa, W. J. E. M. (2003). *Trichomycterus potschi* (Siluriformes: Loricarioidei): a new trichomycterid catfish from coastal streams of southeastern Brazil. *Ichthyol. Explor. Fres.* 14, 281–287.

Becker, R. A., Sales, N. G., Santos, G. M., and Carvalho, D. C. (2015). DNA barcoding and morphological of neotropical ichthyoplankton from the upper Paraná and São Francisco. *J. Fish Biol.* 87, 159–168. doi: 10.1111/jfb.12707

Benzaquem, D. C., Oliveira, C., Batista, J. S., Zuanon, J., and Porto, J. I. R. (2015). DNA barcoding in pencilfishes (Lebiasinidae: *Nannostomus*) reveals cryptic diversity across the Brazilian Amazon. *PLoS One* 10:e0112217. doi: 10.1371/journal.pone.0112217

Blanco, D. R., Lui, R. L., Vicari, M. R., Bertollo, L. A., and Moreira-Filho, O. (2011). Comparative cytogenetics of giant trahiras *Hoplias aimara* and *H. intermedius* (Characiformes, Erythrinidae): chromosomal characteristics of minor and major ribosomal DNA and cross-species repetitive centromeric sequences mapping differ among morphologically identical karyotypes. *Cytogenet. Genome Res.* 132, 71–78. doi: 10.1159/000320923

Carvalho, D. C., Oliveira, D. A. A., Pompeu, P. S., Leal, C. G., Oliveira, C., and Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco river Basin. *Mitochondrial DNA* 22, 80–86. doi: 10.3109/19401736.2011.588214

Carvalho, D. C., Oliveira, D. A. A., Santos, J. E., Teske, P., Beheregaray, L. B., Schneider, H., et al. (2009). Genetic characterization of native and introduced populations of the neotropical cichlid genus Cichla in Brazil. *Genet. Mol. Biol.* 32, 601–607. doi: 10.1590/S1415-47572009005000060

Castro, R. M., and Vari, R. P. (2004). *Detritivores of the South American Fish Family Prochilodontidae (Teleostei: Ostariophysi:Characiformes): A Phylogenetic and Revisionary Study*. Washington, WA: Smithsonian Contributions to Zoology, 189.

Corin, S. E., Lester, P. J., Abbott, K. L., and Ritchie, P. A. (2007). Inferring historical introduction pathways with mitochondrial DNA: the case of introduced Argentine ants (*Linepithema humile*) into New Zealand. *Divers. Distrib.* 13, 510–518. doi: 10.1111/j.1472-4642.2007.00355.x

Costa-Silva, G. J., Rodriguez, M. S., Roxo, F. F., Foresti, F., and Oliveira, C. (2015). Using different methods to access the difficult task of delimiting species in a complex Neotropical hyperdiverse group. *PLoS One* 10:e0135075. doi: 10.1371/journal. pone.0135075

da Silva, M. A., Silva, M. L. N., Curi, N., Avanzi, J. C., and Leite, F. P. (2011). Sistemas de manejo em plantios florestais de eucalipto e perdas de solo e água na região do vale do Rio Doce. MG. *Cien. Florest.* 21, 765–776. doi: 10.5902/198050984520

Deacon, A. E., Ramnarine, I. W., and Magurran, A. E. (2011). How reproductive ecology contributes to the spread of a globally invasive fish. *PLoS One* 6:e24416. doi: 10.1371/journal.pone.0024416

Dergam, J. A., Suzuki, H. I., Shibatta, O. A., Duboc, L. F., Júlio-Jr, H. F., Giuliano-Caetano, L., et al. (1998). Molecular biogeography of the neotropical fish hoplias malabaricus (Erythrinidae: Characiformes) in the Iguaçu, Tibagi, and Paraná rivers. *Genet. Mol. Biol.* 21:4. doi: 10.1590/S1415-47571998000400015

Díaz, J., Villanova, G. V., Brancolini, F., del Pazo, F., Posner, V. M., Grimberg, A., et al. (2016). First DNA barcode reference library for the identification of South American freshwater fish from the lower Paraná river. *PLoS One* 11:e0157419. doi: 10.1371/journal.pone.0157419

Ecoplan-Lume (2010). *Plano Integrado de Recursos Hídricos da Bacia Hidrográfica do Rio Doce*. Belo Horizonte: IGAM.

Ely, C. V., Bordignon, S. A. L., Trevisan, R., and Boldrini, I. I. (2017). Implications of poor taxonomy in conservation. *J. Nat. Conserv.* 36, 10–13. doi: 10.1016/j.jnc.2017.01.003

Eschmeyer, W. N. (2015). *Catalog of Fishes: World Wide Web Electronic Publication*. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp [accessed September 2017].

Fernandes, G. W., Goulart, F. F., Ranieri, B. D., Coelho, M. S., Dales, K., Boesche, N., et al. (2016). Deep into the mud: ecological and socio-economic impacts of the dam breach in Mariana, Brazil. *Nat. Conserv.* 14, 35–45. doi: 10.1016/j.ncon.2016.10.003

Frantine-Silva, W., Sofia, S. H., Orsi, M. L., and Almeida, F. S. (2015). DNA barcoding of freshwater ichthyoplankton in the neotropics as a tool for ecological monitoring. *Mol. Ecol. Resour.* 1, 1226–1237. doi: 10.1111/1755-0998.12385

Froese, R., and Paulay, D. (2010). *FishBase: World Wide Web Electronic Publication*. Available at: http://fishbase.org [accessed September 2017].

Garutti, V. (1995). *Revisão Taxonômica dos Astyanax (Pisces, Characidae), Com Mancha Umeral Ovalada e Mancha No Pedúnculo Caudal, Estendendo-se a Extremidade dos Raios Caudais Medianos, das Bacias do Paraná, São Francisco e Amazônica*. Ph.D. thesis, Universidade Estadual Paulista, São Paulo.

GFT (2015). *Avaliação dos Efeitos e Desdobramentos do Rompimento da Barragem de Fundão em Mariana-MG*. Available at: http://www.agenciaminas.mg.gov.br/ckeditor_assets/attachments/770/relatorio_final_ft_03_02_2016_15h5min.pdf [accessed September 2017].

Gomes, L. C., Pessali, T. C., Sales, N. G., Pompeu, P. S., and Carvalho, D. C. (2015). Integrative taxonomy detects cryptic and overlooked fish species in a neotropical river basin. *Genetica* 143, 581–588. doi: 10.1007/s10709-015-9856-z

Hazkani-Covo, E., Zeller, R. M., and Martin, W. (2010). Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet.* 6:e1000834. doi: 10.1371/journal.pgen.1000834

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. G., and Hallwachs, W. (2004a). Ten species in one: DNA barcoding reveals cryptic species in the

neotropical skipper butterfly *Astrapes fulgerator*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 14812–14817. doi: 10.1073/pnas.0406166101

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S., and Francis, C. M. (2004b). Identification of birds through DNA barcodes. *PLoS Biol.* 2:e312. doi: 10.1371/journal.pbio.0020312

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., Taylor, E., Burridge, M., et al. (2008). Identifying Canadian freshwater fishes through DNA barcodes. *PLoS One* 3:e2490. doi: 10.1371/journal.pone.0002490

Hurst, G. D. D., and Jiggins, F. M. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc. R. Soc. B* 272, 1525–1534. doi: 10.1098/rspb.2005.3056

Ivanova, V., Zemlak, T. S., Hanner, R. H., and Hebert, P. D. N. (2007). Universal primer cocktails for fish DNA barcoding. *Mol. Ecol. Resour.* 7, 544–548. doi: 10.1111/j.1471-8286.2007.01748.x

Jukes, T. H., and Cantor, C. R. (1969). "*Evolution of Protein Molecules*," in Mammalian protein metabolism, ed. H. N. Munro (New York, NY: Academic Press).

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–120. doi: 10.1007/BF01731581

Kittell, M. M., Harvey, M. N., Contreras-Balderas, S., and Ptacek, M. B. (2005). Wild-caught hybrids between sailfin and shortfin mollies (Poeciliidae, *Poecilia*): morphological and molecular verification. *Hidrobiológica* 15, 131–137.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lampert, K., and Schartl, M. (2008). The origin and evolution of a unisexual hybrid: *Poecilia formosa*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 2901–2909. doi: 10.1098/rstb.2008.0040

Latini, A. O., and Petrere, M. Jr. (2004). Reduction of a native fish fauna by alien species: an example from Brazilian freshwater tropical lakes. *Fish. Manag. Ecol.* 11, 71–79. doi: 10.1046/j.1365-2400.2003.00372.x

Metcalf, J. L., Pritchard, V. L., Silvestri, S. M., Jenkins, J. B., Wood, J. S., Cowley, D. E., et al. (2007). Across the great divide: genetic forensics reveals misidentification of endangered cutthroat trout populations. *Mol. Ecol.* 16, 4445–4454. doi: 10.1111/j.1365-294X.2007.03472.x

Meyer, C. P., and Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol.* 3:e422. doi: 10.1371/journal.pbio.0030422.t002

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Fonseca, G. A. B., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858. doi: 10.1038/35002501

Nascimento, M. H. S., Almeida, M. S., Veira, M. N. S., Limeira-Filho, D., Lima, R. C., Barros, M. C., et al. (2016). DNA barcoding reveals high levels of genetic diversity in the fishes of the Itapecuru Basin in Maranhão. Brazil. *Genet. Mol. Res.* 15, 1–11. doi: 10.4238/gmr.15038476

Neves, A. C. O., Nunes, F. P., Carvalho, F. A., and Fernandes, G. W. (2016). Neglect of ecosystems services by mining, and the worst environmental disaster in Brazil. *Nat. Conserv.* 14, 24–27. doi: 10.1016/j.ncon.2016.03.002

Olds, A. D., Pitt, K. A., Maxwell, P. S., and Connolly, R. M. (2012). Synergistic effects of reserves and connectivity on ecological resilience. *J. Appl. Ecol.* 49, 1195–1203. doi: 10.1111/jpe.12002

Oliveira, E. A., Bertollo, L. A. C., Yano, C. F., Liehr, T., and Cioffi, M. B. (2015). Comparative cytogenetics in the genus *Hoplias* (Characiformes, Erythrinidae) highlights contrasting karyotype evolution among congeneric species. *Mol. Cytogenet.* 8:56. doi: 10.1186/s13039-015-0161-4

Padial, J. M., Miralles, A., De la Riva, I., and Vences, M. (2010). The integrative future of taxonomy. *Front. Zool.* 7:16. doi: 10.1186/1742-9994-7-16

Pereira, L. H., Pazian, M. F., Hanner, R., Foresti, F., and Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the neotropical freshwater fish *Piabina argentea* (Characiformes: Characidae) from the upper Parana Basin of Brazil. *Mitochondrial DNA* 22, 87–96. doi: 10.3109/19401736.2011.588213

Pereira, L. H. G., Hanner, R., Foresti, F., and Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse neotropical freshwater fish fauna? *BMC Genet.* 14:20. doi: 10.1186/1471-2156-14-20

Pfenninger, M., Nowak, C., Kley, C., Steinke, D., and Streit, B. (2007). Utility of DNA taxonomy and barcoding for the inference of larval community structure in morphologically cryptic *Chironomus* (Diptera) species. *Mol. Ecol.* 16, 1957–1968. doi: 10.1111/j.1365-294X.2006.03136.x

Pugedo, M. L., Andrade-Neto, F. R., Pessali, T. C., Birindelli, J. L. O., and Carvalho, D. C. (2016). Integrative taxonomy supports new candidate fish species in a poorly studied neotropical region: the Jequitinhonha river Basin. *Genetica* 144, 341–349. doi: 10.1007/s10709-016-9903-4

Puillandre, N., Lambert, A., Brouillet, S., and Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Mol. Ecol.* 21, 1864–1877. doi: 10.1111/j.1365-294X.2011.05239.x

Ramirez, J. L., Birindelli, J. L., Carvalho, D. C., Affonso, P. R. A. M., Venere, P. C., Ortega, H., et al. (2017). Revealing hidden diversity of the underestimated neotropical ichthyofauna: DNA barcoding in the recently described genus *Megaleporinus* (Characiformes: Anostomidae). *Front. Genet.* 8:149. doi: 10.3389/fgene.2017.00149

Ratnasingham, S., and Hebert, P. D. N. (2007). BOLD: the barcode of life data system. *Mol. Ecol. Notes* 3, 355–364. doi: 10.1111/j.1471-8286.2007.01678.x

Ratnasingham, S., and Hebert, P. D. N. (2013). A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS One* 8:e66213. doi: 10.1371/journal.pone.0066213

Reis, R. E., Albert, J. S., Di Dario, F., Mincarone, M. M., Petry, P., and Rocha, L. A. (2016). Fish biodiversity and conservation in South America. *J. Fish Biol.* 89, 12–47. doi: 10.1111/jfb.13016

Ronquist, F., and Huelsenbeck, J. P. (2003). MRBAYES 3: bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574. doi: 10.1093/bioinformatics/btg180

Rosa, R. S., and Lima, F. C. T. (2005). *Livro da Fauna Brasileira Ameaçada de Extinção*. Belo Horizonte: Biodiversitas.

Rossini, B. C., Oliveira, C. A. M., Melo, F. A. G. D., Bertaco, V. A., Astarloa, J. M. D., Rosso, J. J., et al. (2016). Highlighting *Astyanax* species diversity through DNA barcoding. *PLoS One* 11:12. doi: 10.1371/journal.pone.0167203

Rosso, J., Mabragaña, E., Castro, M. G., and de Astarloa, J. M. (2012). DNA barcoding neotropical fishes: recent advances from the pampa plain. Argentina. *Mol. Ecol. Resour.* 12, 999–1011. doi: 10.1111/1755-0998.12010

Saitou, N., and Nei, M. (1987). The neighbor-joining method—a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Sales, N. G., Pessali, T. C., Andrade-Neto, F. R., and Carvalho, D. C. (2018). Introgression from non-native species unveils a hidden threat to the migratory neotropical fish *Prochilodus hartii*. *Biol. Invasions* 20, 555–566. doi: 10.1007/s10530-017-1556-4

Shum, P., Moore, L., Pampoulie, C., Di Muri, C., Vandamme, S., and Mariani, S. (2017). Harnessing mtDNA variation to resolve ambiguity in 'Redfish' sold in Europe. *PeerJ* 2, 1–18. doi: 10.7717/peerj.3746

Sievers, C., Willing, E. M., Hoffmann, M., Dreyer, C., Ramnarine, I., and Magurran, A. (2012). Reasons for the invasive success of a Guppy (*Poecilia reticulata*) population in Trinidad. *PLoS One* 7:e38404. doi: 10.1371/journal.pone.0038404

Song, H., Buhay, J. E., Whiting, M. F., and Crandall, K. A. (2008). Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13486–13491. doi: 10.1073/pnas.0803076105

Sunaga, T., and Verani, J. R. (1987). "Second report of comparative study on fish community of the Rio Doce Valley lakes," in *Limnological Studies in Rio Doce Valley Lakes and Pantanal Wetland*, eds Y. Saijo and J. G. Tundisi (Brazil: Nagoya University), 129–135.

Taylor, R. W. (1983). "Descriptive taxonomy: past, present, and future," in *Australian Systematic Entomology: A Bicentenary Perspective*, eds E. Highley and R. W. Taylor (Victoria, VIC: CSIRO), 93–134.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The clustalx windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic. Acids Res.* 24, 4876–4882. doi: 10.1093/nar/25.24.4876

Valdez-Moreno, M., Ivanova, N. V., Elias-Gutierrez, M., Contreras-Balderas, S., and Hebert, P. D. N. (2009). Probing diversity in freshwater fishes from mexico and Guatemala with DNA barcodes. *J. Fish. Biol.* 74, 377–402. doi: 10.1111/j.1095-8649.2008.02077.x

Vari, R. (1992). Systematics of the neotropical characiform genus *Cyphocharax* fowler (Pisces:Ostariophysi). *Smithon. Contrib. Zool.* 529, 137. doi: 10.5479/si.00810282.529

Vieira, F. (1994). *Estrutura de Comunidade e Aspectos da Alimentação e Reprodução dos Peixes em Dois Lagos do Médio Rio Doce, MG*. Doctoral dissertation, Universidade Federal de Minas Gerais, Belo Horizonte.

Vieira, F. (2009). Distribuição, impactos ambientais e conservação da fauna de peixes da bacia do rio Doce. *MG BIOTA* 2, 5–22.

Vieira, F., and Alves, C. B. M. (2001). Threatened fishes of the world: *Henochilus wheatlandii* Garman 1980 (Characidae). *Env. Biol. Fish.* 62, 414. doi: 10.1023/A:1012233128362

Vieira, F., Alves, C. B. M., and Santos, G. B. (2000). Rediscovery and first record of *Henochilus wheatlandii* (Teleostei: Characiformes) a rare neotropical fish, in rio Doce basin of Southeastern Brazil. *Ichthyol. Explor. Fres.* 11, 201–206.

Vinas, J., and Tudela, S. (2009). A validated methodology for genetic identification of tuna species (*Genus Thunnus*). *PLoS One* 4:e7606. doi: 10.1371/journal.pone.0007606

Vono, V., and Barbosa, F. A. R. (2001). Habitats and littoral zone fish community structure of two natural lakes in Southeast Brazil. *Env. Biol. Fish.* 61, 371–379. doi: 10.1023/A:1011628102125

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., and Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. B* 360, 1847–1857. doi: 10.1098/rstb.2005.1716

Zanata, A. M., and Camelier, P. (2009). Astyanax vermilion and Astyanax burgerai: new characid fishes (Ostariophysi: Characiformes) from Northeastern Bahia. Brazil. *Neotrop. Ichthyol.* 7, 175–184. doi: 10.1590/S1679-62252009000200007

Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29, 2869–2876. doi: 10.1093/bioinformatics/btt499

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 14 August 2018
doi: 10.3389/fgene.2018.00295

Remarkable Geographic Structuring of Rheophilic Fishes of the Lower Araguaia River

Tomas Hrbek[1]*, *Natasha V. Meliciano*[1,2], *Jansen Zuanon*[3] and *Izeni P. Farias*[1]

[1] Laboratório de Evolução e Genética Animal, Departmento de Genética, Instituto de Ciências Biológicas, Universidade Federal do Amazonas, Manaus, Brazil, [2] Instituto de Saúde e Biotecnologia, Universidade Federal do Amazonas, Coari, Brazil, [3] Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil

Rapids and waterfalls, and their associated fauna and flora are in peril. With the construction of each new hydroelectric dam, more rapids and waterfalls are destroyed, leading to the disappearance of associated fauna and flora. Areas of rapids harbor distinct, highly endemic rheophilic fauna and flora adapted to an extreme environment. Rheophilic habitats also have disjunct distribution both within and across rivers. Rheophilic habitats thus represent islands of suitable habitat separated by stretches of unsuitable habitat. In this study, we investigated to what extent, if any, species of cichlid and anostomid fishes associated with rheophilic habitats were structured among the rapids of Araguaia River in the Brazilian Amazon. We tested both for population structuring as well as non-random distribution of lineages among rapids. Eight of the nine species had multiple lineages, five of these nine species were structured, and three of the eight species with multiple lineages showed non-random distribution of lineages among rapids. These results demonstrate that in addition to high levels of endemicism of rheophilic fishes, different rapids even within the same river are occupied by different lineages. Rheophilic species and communities occupying different rapids are, therefore, not interchangeable, and this realization must be taken into account when proposing mitigatory/compensatory measures in hydroelectric projects, and in conservation planning.

Keywords: Cichlidae, Anostomidae, Tocantins river basin, mitochondrial control region, micro endemism, diversity loss, rapids

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

Reviewed by:
Henrique Batalha-Filho, Universidade Federal da Bahia, Brazil
Rubens Pazza, Federal University of Viçosa, Campus Rio Paranaíba, Brazil

***Correspondence:**
Tomas Hrbek
hrbek@evoamazon.net

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

Received: *01 November 2017*
Accepted: *13 July 2018*
Published: *14 August 2018*

Citation:
Hrbek T, Meliciano NV, Zuanon J and Farias IP (2018) Remarkable Geographic Structuring of Rheophilic Fishes of the Lower Araguaia River. Front. Genet. 9:295. doi: 10.3389/fgene.2018.00295

INTRODUCTION

Brazil, as many other South American nations, has experienced a period of rapid economic growth accompanied by expanding energy needs. To meet its energetic needs, Brazil has invested heavily in hydroelectric power generation such that ≈ 80% of Brazil's electricity production is currently met by hydropower (International Energy Agency, 2013a). At the same time, it is estimated that up to 55% the hydroelectric potential still could be exploited (International Energy Agency, 2013b). While the hydroelectric potential of the Paraná and São Francisco river basins in southern, central and northeastern Brazil has largely been realized, the Amazon basin is the next and last hydroelectric frontier (Lees et al., 2016; Latrubesse et al., 2017). However, several large dams, such as the Tucuruí, Balbina, Santo Antônio, Jirau and the recently completed Belo Monte

have already been implemented in the Brazilian Amazon. An additional 200+ dams, such as the Tapajós hydroelectric complex, have been proposed by South American governments (Finer and Jenkins, 2012; Castello et al., 2013; Lees et al., 2016). The construction of the Tapajós hydroelectric complex was only not initiated mostly due to the economic crisis Brazil is experiencing since 2014 (Fearnside, 2015). If these plans were to be enacted, only three Amazon tributaries would remain unimpounded (Castello and Macedo, 2016). Although hydroelectric projects have been lauded as cheap and clean energy alternatives, dam construction and operation result in substantial environmental stresses including the destruction of rheophilic habitats by permanently submerging rapids and waterfalls (Clausen and York, 2008; Castello and Macedo, 2016; Pelicice et al., 2017).

Rheophilic habitats are a unique geomorphological feature of Amazonian River tributaries that descend the Brazilian and Guiana shields, cutting through the rocky surface to create a series of waterfalls and rapids, and a complex matrix of rocky habitats. Rapids are characterized as river sections of supercritical flow, where surface tension breaks at the water/air interface (Hawkins et al., 1993); the presence of rocky substratum is a fundamental component of this type of environment. These river sections are characterized by not only high velocity but also by highly heterogeneous water flow, oxygen–rich waters, and a very complex substrate matrix consisting of rocky slabs, caves, cracks, and crevices with lodged tree trunks harboring a highly diverse array of niches. It is in these high–energy sections of the rivers where hydroelectric projects are developed. Damming of the river at these sites to create a reservoir (most Amazonian hydroelectric projects) or diverting the river from these sites into a holding reservoir (as done in the Belo Monte hydroelectric project), changes permanently the hydrology of the river. Rheophilic habitats generally are permanently submerged underneath the reservoir, turning a high–energy, oxygen–rich lotic habitat into an hypoxic lentic habitat. In the case of the Belo Monte hydroelectric project, much of the rheophilic habitat in the Volta Grande has become permanently dewatered (emerged), and those sections that had not, no longer are subject to seasonal water level fluctuations necessary for the flowering and thus the reproduction of the unique Podostemaceae flora and the fish fauna that depend on this resource and which use seasonal water level fluctuations as reproductive clues (Lowe-McConnell, 1987). Under both scenarios, the specialized rheophilic fauna and flora suffers local or potentially even global extinctions (Winemiller et al., 2016).

From an ichthiofaunistic perspective, rapids had historically been poorly sampled and are poorly known environments (Böhlke et al., 1978; Menezes, 1996). On the other hand, it is known that rapids areas harbor their own specialized and commonly endemic ichthyofaunas, adapted to life in turbulent water environments (e.g., Hora, 1930; Roberts and Stewart, 1976; Kullander, 1988; Isbrücker and Nijssen, 1991; Jégu, 1992, 2004; Zuanon, 1999; Flausino Junior et al., 2016; Collins et al., 2018; Fitzgerald et al., 2018; Machado et al., 2018). Studies of natural history and ecology focusing on associations of fish species in rapids and waterfalls are rare (Balon, 1974; Balon and Stewart, 1983; Casatti and Castro, 1998; Fitzgerald et al., 2018). Mainly due to sampling difficulties, little is known about the way of life and structure of fish communities in these environments. Most of the information on the fish assemblages of rapids in the Amazon region is concentrated in unpublished technical reports, especially those related to studies of environmental impacts of proposed hydroelectric projects. In addition, existing information is usually restricted to occurrence records based on poorly resolved taxonomy, with data obtained without standardized sampling/experimental design among studies, making data interpretation and quantitative comparisons difficult. In Brazil, publications of rheophilic communities are restricted to a natural history study of rheophilic fish fauna of a stream in the headwaters of the São Francisco River (Casatti and Castro, 1998), a study (Fitzgerald et al., 2018) and an unpublished doctoral thesis (Zuanon, 1999) focusing on the Volta Grande region of the Xingu River, and a study of the fishes associated with podostomacean mats in the Aripuanã River (Flausino Junior et al., 2016). Despite the scarcity of studies, it is known that rheophilic environments harbor species–rich and highly diverse assemblages, comprised of many endemics.

Dearth of studies of rheophilic taxa is in part a consequence of the still poor general knowledge of distribution of freshwater biodiversity (Vörösmarty et al., 2010). In the Neotropics, this impediment is the result of severe undersampling of freshwater taxa in general (Lundberg et al., 2000; Lévêque et al., 2007; Alofs et al., 2014), leading to many taxa not being evaluated by the IUCN or national conservation agencies, or when evaluated being listed as data deficient (Collen et al., 2014). This, in turn, hampers prioritizing freshwater habitats for conservation (Abell, 2002; Darwall et al., 2011; Frederico et al., 2016, 2018).

Perhaps the most emblematic groups of fishes occupying rheophilic habitats in the Neotropics are the suckermouth armored catfishes (Loricariidae). The loricariids have acquired spectacular adaptations in their mouth, labial and dental morphologies that allow them to occupy rheophilic habitats and diversify within them (Lujan et al., 2012, 2017; Roxo et al., 2017; Collins et al., 2018). They probably are the most strictly associated group of fishes with rheophilic habitats (Reis et al., 2003). However, the loricariids are not the only group to have colonized and diversified within rheophilic habitats. Several clades of the plant eating pacus (Serrasalmidae) are also rheophilic specialists (e.g., Jégu, 2004; Andrade et al., 2016, 2017; Machado et al., 2018), as are some cichlid (Cichlidae) (e.g., Kullander, 1988) and anostomid (Anostomidae) (e.g., Santos and Jégu, 1987) species.

Among–river discontinuity of rheophilic habitats causes high degree of endemism, with few species occupying more than a single river basin (Reis et al., 2003), and those that occur in more than one river generally represent divergent, independently evolving lineages (Collins et al., 2018). By the same token, discontinuous rheophilic habitats within the same river system could lead to structuring of rheophilic fish species. Rheophilic habitats are islands of suitable habitat within a waterscape of unsuitable habitat, analogous to the rock islands of the great

African lakes (Salzburger et al., 2014). It is with this objective that we decided to study to what degree, if any, the rheophilic fish fauna occupying distinct and unconnected rheophilic habitat is genetically structured along the Araguaia River. The implication of population structuring of rheophilic fishes, or the occupation of different rapids by different lineages would make reophilic communities occupying different rapids even more distinct and non–substitutable when planning environmental mitigatory measures.

For this purpose we chose to sample and analyze the less well studied cichlid and anostomiid species rather than the obvious candidate loricariid species. If the rheophilic habitats really are analogous to the rocky islands and outcrops in the great African lakes, then we should observe significant amount of population structuring between rheophilic cichlids and anostomiids occupying different rapids along the river.

MATERIALS AND METHODS

Field Sampling

Over an eight day period, we sampled six stretches of rapids in the lower Araguaia River (**Table 1**; **Figure 1**). The sampled rapids were Cachoeira do Zé dos Gatos, São Miguel, Remanso dos Botos, Santa Isabel, São Bento, and Rebojo (**Figure 2**). Distance between the Cachoeira do Zé dos Gatos and Rebojo rapids is 150 km, while the São Bento and Rebojo rapids are separated by less than 1 km.

After a study of the prevailing conditions in the field, we decided to divide the team into a group of two people collecting fish with gilnets (24–60 mm mesh between opposing nodes) and castnets (10 mm mesh between opposing nodes) and one person catching fish with line and hook. This strategy aimed to optimize capture effort and to sample the available diversity of habitats occupied by anostomids and cichlids. Sampling time at each site was approximately 8 h with the the exception of the Cachoeira do Zé dos Gatos and Santa Isabel rapids where we spent approximately 16 h each.

We focused on sampling two groups: cichlids which generally are sedentary and some species are rheophilic specialists, and anostomids which also have rheophilic specialists but as a group may be considered vagil. Final choice of species included in this study was conditioned by our ability to obtain sufficient number of specimens from at least four of the six rapids.

Preliminary identification of specimens was carried out in the field. Final identification of the collected specimens was made based on comparisons with samples deposited in the INPA Fish Collection (Manaus, AM), consultation of original descriptions and use of published dichotomous keys (when available).

All field collections were authorized by IBAMA/SISBIO 11325-1, and access to genetic resources was authorized by permit No. 034/2005/IBAMA.

DNA Extraction, Molecular Markers and Data Pre-processing

DNA extraction was performed using standard phenol/chloroform protocol (Sambrook and Russell, 2001). For the molecular analyzes we chose to amplify the control region *CR* of mitochondrial DNA (mtDNA), since it is the fastest evolving section of the mitochondrial genome, and thus is most likely to register fine-scale population structuring (Meyer, 1993). We obtained approximately 650–750 base pairs for each individual. We analyzed the same region of mtDNA for all individuals and thus eliminated any confounding effects of using different mtDNA gene regions with potentially different rates of molecular evolution in different species.

A common forward primer was used for all species (ProF: 5′-AACYCCCRCCCCTAACYCCCAAAG-3′). For anostomids we used the reverse primer DLOstariR.1 (5′-gtaaaacgacggccagTCCT GGTTTHGGGGTTTRAC6AG-3′), while for cichlids we used the reverse primer DLPercoR.1 (5′-gtaaaacgacggccagTCCTG



FIGURE 1 | A close up of an island within the Santa Isabel rapid complex showing habitat structure and flowering Podostemaceae.

TABLE 1 | Location of sampled rapids in the Araguaia River.

Locality	State	Municipality	Date	Latitude	Longitude
Zé dos Gatos rapid	Pará	Piçarra	8–9/sept/09	−6.78927	−48.97752
Santa Isabel rapid	Pará	Santa Isabel	10–11/sept/09	−6.14894	−48.36074
São Miguel rapid	Tocantins	Xambioá	12/sept/09	−6.38009	−48.53014
Remanso dos Botos rapid	Tocantins	Xambioá	13/sept/09	−6.36741	−48.37892
São Bento rapid	Tocantins	Araguatins	14/sept/09	−5.46001	−48.34394
Rebojo rapid	Tocantins	Araguatins	15/sept/09	−5.39970	−48.39315

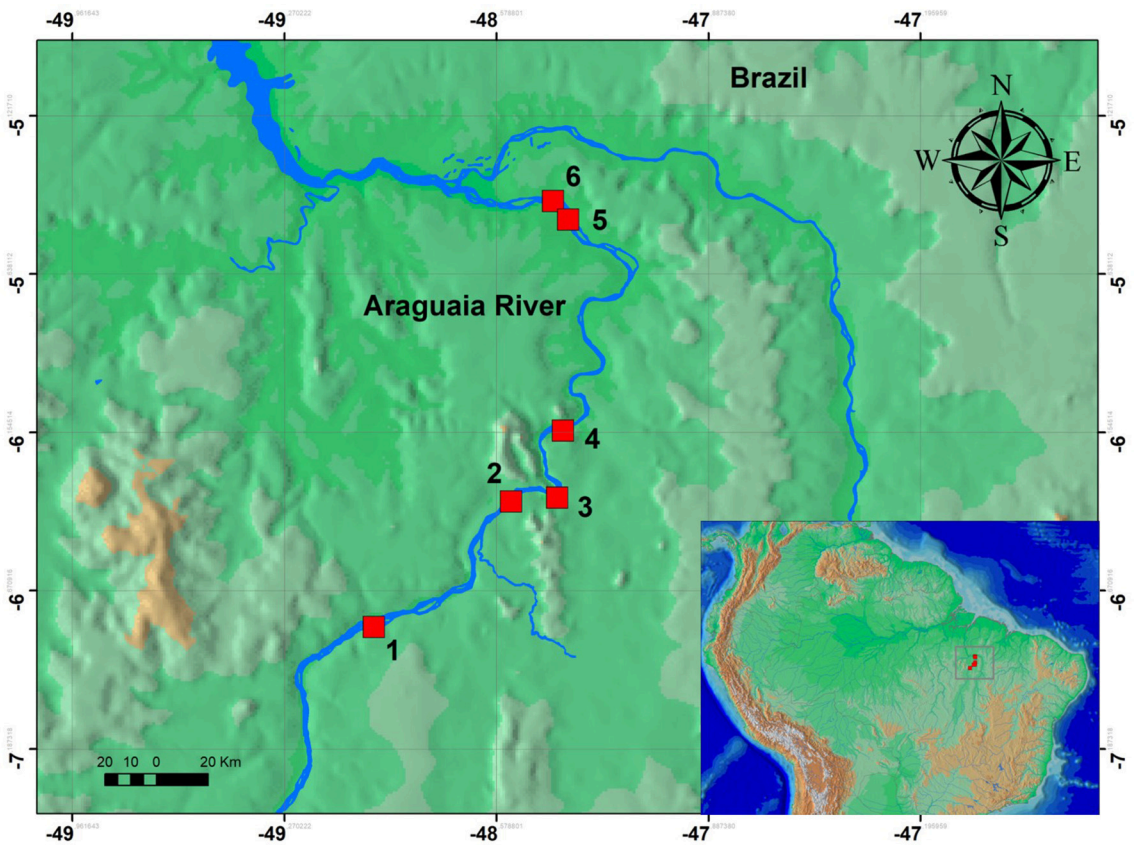


FIGURE 2 | Map of the six regions of rapids sampled in this study. Localities are: (1) Zé dos Gatos rapid; (2) São Miguel rapid; (3) Remanso dos Botos rapid; (4) Santa Isabel rapid; (5) São Bento rapid; (6) Rebojo rapid.

TTTCCGGGGGGTTTACAG-3′). Both reverse primers incorporated an M13(-21) tail on their 5′ end. The 15 μL PCR mix included 1.2 μL of 10 mM dNTPs (2.5 mM each DNTP), 1.5 μL 10× buffer (75 mM Tris HCL, 50 mM KCL, 20 mM $(NH_4)_2SO_4$), 1.2 μL 25 mM $MgCl_2$, 1.5 μL of primer cocktails (2 pmol each), 0.5 μL of Taq DNA polymerase, 1 μL of template DNA and 6.6 μL ddH_2O. PCR conditions were: 94°C (30 s); 35 cycles of 94°C (30 s), 50°C (35 s), 68°C (90 s); followed by 68°C (5 min).

PCR products were evaluated on a 1% percent agarose gel, and then purified using Exo-SAP (Exonuclease—Shrimp Alcaline Phosphatase) following the manufacturer's suggested protocol (Werle et al., 1994). The control region products were sequenced using the M13(-21) primer. Sequencing reactions were carried out according to the manufacturer's recommendation for the ABI BigDye Terminator cycle sequencing mix, using an annealing temperature of 50°C (Platt et al., 2007). Sequencing reactions were precipitated using standard EDTA/EtOH protocol, and was resolved on the ABI 3130xl (Life Technologies) automatic sequencer. Sequence products were edited, concatenated and aligned using the Clustal W algorithm (Thompson et al., 1994) followed by manual adjustments as implemented in the program GENIOUS v8.1.7 (Kearse et al., 2012). Sequences are deposited in GenBank under accession numbers MH514035–MH514287.

Analyses

Using the mtDNA control region we tested whether the populations of the species found at different rapids are differentiated from one another. We tested this hypothesis using the Analysis of Molecular Variance (AMOVA) (Excoffier et al., 1992) as implemented in the program ARLEQUIN v3.5 (Excoffier and Lischer, 2010). Significant differentiation was interpreted as lack of gene exchange between populations of individuals of the same species found in different rapids.

We also delimited lineages using the Bayesian approach implemented in the program BAPS v6.0 (Corander et al., 2008). Individual level mixture analysis was performed for different maximum number of lineages ($k = 1$–5), with 10 independent runs for each value of k. The k with the highest posterior probability was selected as representing the correct data partition. We subsequently tested for a non-random distribution of lineages across the different rapids from which these lineages were sampled using the Fisher's exact test as implemented in the R statistical language (R Development Core Team, 2011).

Haplotype network and geographic occurrence of haplotypes was estimated in the program HAPVIEW (Salzburger et al., 2011) using a phylogeny of haplotypes generated in RAXML v8.2.0 (Stamatakis, 2014).

Pairwise uncorrected p–distances were calculated between all individuals of a given species using GENIOUS v8.1.7 (Kearse et al., 2012).

Finally, we calculated the Tajima's D (Tajima, 1989) for each species.

All analyses were carried out separately for each species, assuming $\alpha = 0.05$.

RESULTS

The aligned DNA data matrix comprised nine taxa, 253 individuals by 640–778 bp. Number of individuals per species varied from 16 to 43 (**Table 2**). Eight of the nine taxa had more than one lineage, with upto five lineages being observed in *Crenicichla cametana*. In total we observed 175 unique haplotypes (**Figures 3**, **4**) grouped into 29 lineages (**Figure 5**). The existence of multiple lineages within rheophilic taxa and *Leporinus desmotes* is further reforced by significantly positive Tajima's D metrics. Significant population structuring was observed in five of the nine species analyzed, and non-randon association of lineages and rapids was observed in three of the eight species with multiple lineages. Results of tests of among-rapids population differentiation, Tajima's D and Fu's Fs, and non-random distribution of lineages among rapids is summarized in **Tables 3**, **4**, respectively.

Species Accounts

Retroculus lapidifer (Cichlidae) (Figure 6A)

This species is a sand-sifter and a rheophilic specialist. The collections of this species occurred in the localities of Cachoeira do Zé dos Gatos ($N = 2$), Santa Isabel ($N = 9$), São Miguel ($N = 2$) and Rebojo and São Bento ($N = 7$). Within this species we observe at least four genetic lineages with up to 2.24% p–distance sequence divergence. Analysis of Molecular Variance (AMOVA) indicated high levels of geographic structuring ($\phi_{ST} = 0.35119$; $p = 0.00040 \pm 0.00019$). Tajima's D was positive (0.53176). The hypothesis of random associations of genetic lineages with rapids was rejected (Fisher's exact test, $p = 0.00100$). These lineages were largely restricted to specific rapids.

Geophagus altifrons (Cichlidae) (Figure 6B)

This species is a habitat generalist but is commonly found in shallow sandy areas of stretches of rapids. The collections of this species occurred in the localities of Cachoeira do Zé dos Gatos ($N = 1$), Santa Isabel ($N = 12$), Remanso dos Botos ($N = 1$) and Rebojo and São Bento ($N = 29$). We observed only one genetic lineage within this species. AMOVA indicated geographic structuring ($\phi_{ST} = 0.11732$; $p = 0.00802 \pm 0.00086$). Tajima's D was negative and significant (-1.79152).

Cichla piquiti (Cichlidae) (Figure 6C)

This species is usually found in structured habitats, including rapids, where it lurks and waits for its prey. The collections of this species occurred in the localities of Cachoeira do Zé dos Gatos ($N = 1$), Santa Isabel ($N = 19$), São Miguel ($N = 1$), Remanso dos Botos ($N = 1$) and Rebojo and São Bento ($N = 6$). We observed three genetic lineages with up to 2.58% p–distance sequence divergence. AMOVA indicated high levels of geographic structuring ($\phi_{ST} = 0.29184$; $p = 0.01050 \pm 0.00096$). Tajima's D was positive and significant (2.35861). The hypothesis of random associations of genetic lineages with rapids was not rejected (Fisher's exact test, $p = 0.08546$).

Crenicichla cametana (Cichlidae) (Figure 6D)

This species is a relatively small ambush predator, and a specialist of rapids. The collections of this species occurred in the localities of São Miguel ($N = 9$), Remanso dos Botos ($N = 5$) and Rebojo and São Bento ($N = 8$). We observed five divergent genetic lineages with up to 1.57% p–distance sequence divergence. AMOVA indicated high levels of geographic structuring ($\phi_{ST} = 0.86498$; $p < 0.00001$). Tajima's D was positive and significant (1.57836). The hypothesis of random association of genetic lineages with rapids was rejected (Fisher's exact test, $p = 0.00050$). These lineages were largely restricted to specific rapids.

Hypomasticus cf. *pachycheilus* (Anostomidae) (Figure 7A)

This species is a rheophilic specialist, an algae scraper, found in fast flowing waters of rapids. The collections of this species

TABLE 2 | Number of individuals of each species analyzed from the rapids of the Araguaia River.

Species	Zé dos Gatos	Santa Isabel	São Miguel	Remanso dos Botos	São Bento & Rebojo	Total
Retroculus lapidifer	2	9	2	0	7	20
Geophagus altifrons	1	12	0	1	29	43
Cichla piquiti	1	19	1	1	6	28
Crenicichla cametana	0	0	9	5	8	22
Hypomasticus cf. *pachycheilus*	27	1	2	4	9	43
Leporinus maculatus	4	9	0	9	6	28
Leporinus affinis	0	3	0	2	29	34
Leporinus desmotes	2	2	2	4	6	16
Leporinus unitaeniatus	4	4	0	0	11	19

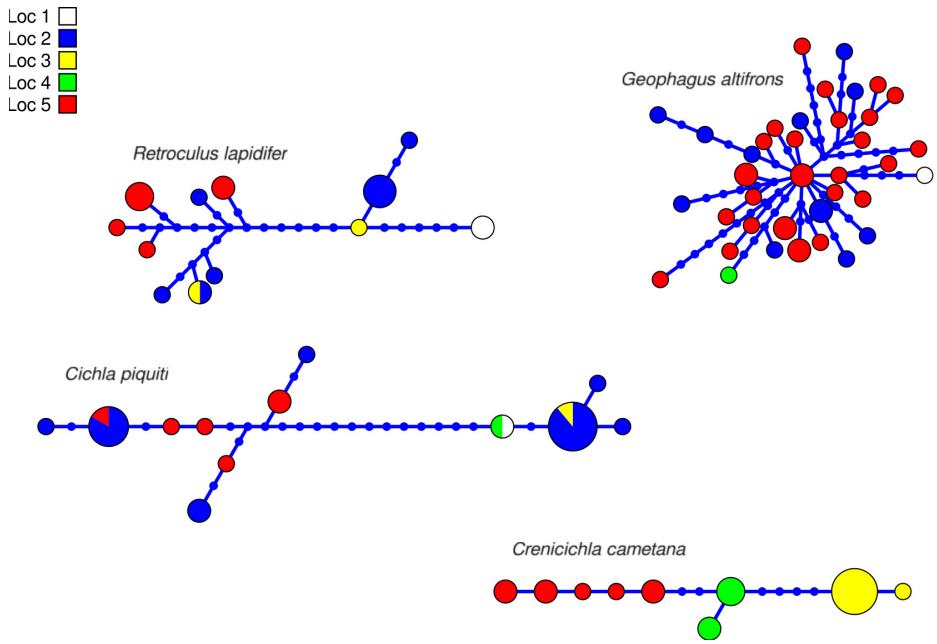


FIGURE 3 | Haplotype network of four species of *Cichlidae* indicating geographic occurance of the haplotypes. Networks were generated in HAPVIEW, and haplotype size is proportional among networks and to the number of individuals with that haplotype. Localities are: (1) Zé dos Gatos rapid; (2) São Miguel rapid; (3) Remanso dos Botos rapid; (4) Santa Isabel rapid; (5) São Bento rapid and Rebojo rapid.

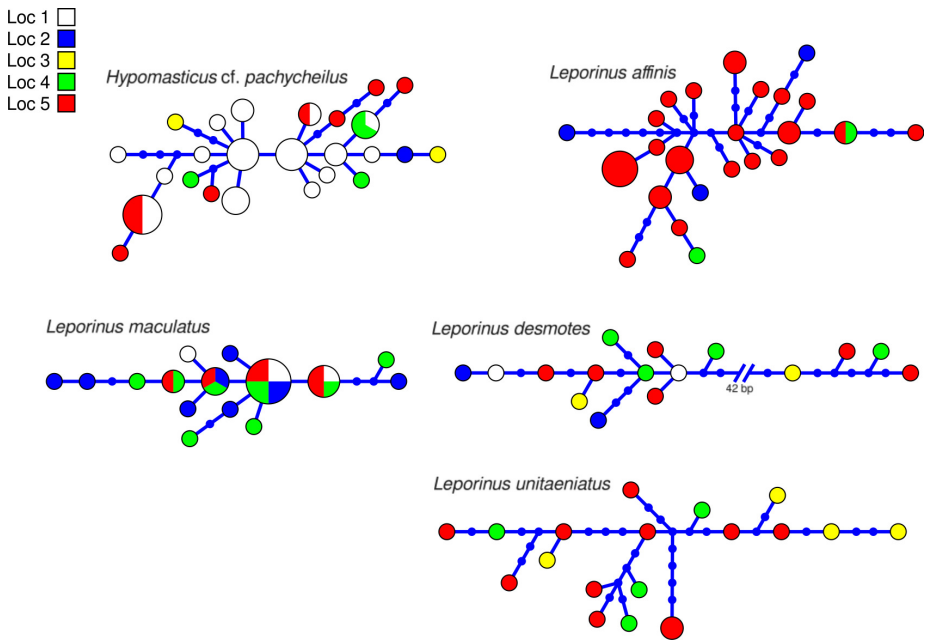


FIGURE 4 | Haplotype network of five species of *Anostomidae* indicating geographic occurance of the haplotypes. Networks were generated in HAPVIEW, and haplotype size is proportional among networks and to the number of individuals with that haplotype. Localities are: (1) Zé dos Gatos rapid; (2) São Miguel rapid; (3) Remanso dos Botos rapid; (4) Santa Isabel rapid; (5) São Bento rapid and Rebojo rapid.

occurred in the localities of Cachoeira do Zé dos Gatos ($N = 27$), Santa Isabel ($N = 1$), São Miguel ($N = 2$), Remanso dos Botos ($N = 4$) and Rebojo and São Bento ($N = 9$). We observed four genetic lineages with up to 1.04% p–distance sequence divergence. AMOVA indicated geographic structuring ($\phi_{ST} = 0.15674$; $p = 0.00406 \pm 0.00066$). Tajima's D was

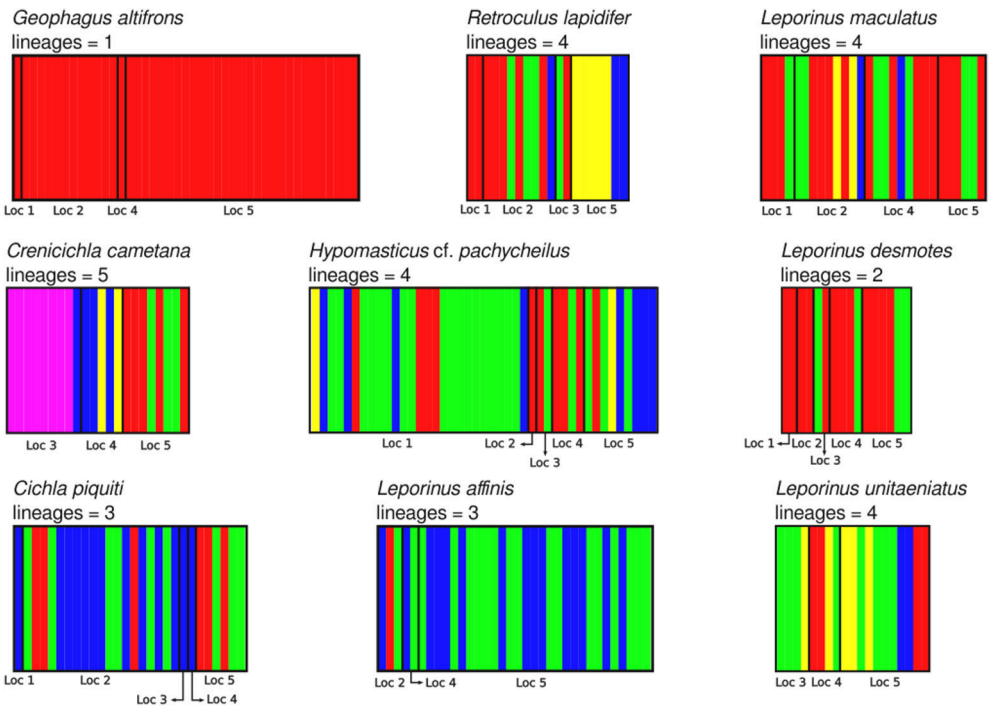


FIGURE 5 | Identification of lineages of all nine species analyzed in this study using BAPS v6.0. Localities are: (1) Zé dos Gatos rapid; (2) São Miguel rapid; (3) Remanso dos Botos rapid; (4) Santa Isabel rapid; (5) São Bento rapid and Rebojo rapid.

TABLE 3 | Analysis of population differentiation among rapids of the Araguaia River, and demography of each species.

Species	ϕ_{ST}	AMOVA	Tajima's D
Cichlidae			
Retroculus lapidifer	0.35119	$p = 0.00040 \pm 0.00019$	0.53176†
Geophagus altifrons	0.11732	$p = 0.00802 \pm 0.00086$	−1.79152*
Cichla piquiti	0.29184	$p = 0.01050 \pm 0.00096$	2.35861‡
Crenicichla cameta	0.86498	$p < 0.00001$	1.57836‡
Anostomidae			
Hypomasticus cf. *pachycheilus*	0.15674	$p = 0.00406 \pm 0.00066$	−0.80127
Leporinus maculatus	−0.04735	$p = 0.87851 \pm 0.00363$	−1.04746
Leporinus affinis	0.00075	$p = 0.43525 \pm 0.00483$	−1.09098
Leporinus desmotes	−0.19958	$p = 0.95950 \pm 0.00191$	0.96519†
Leporinus unitaeniatus	0.02545	$p = 0.29604 \pm 0.00478$	−0.19675

*,significantly negative; †,positive; ‡,significantly positive.

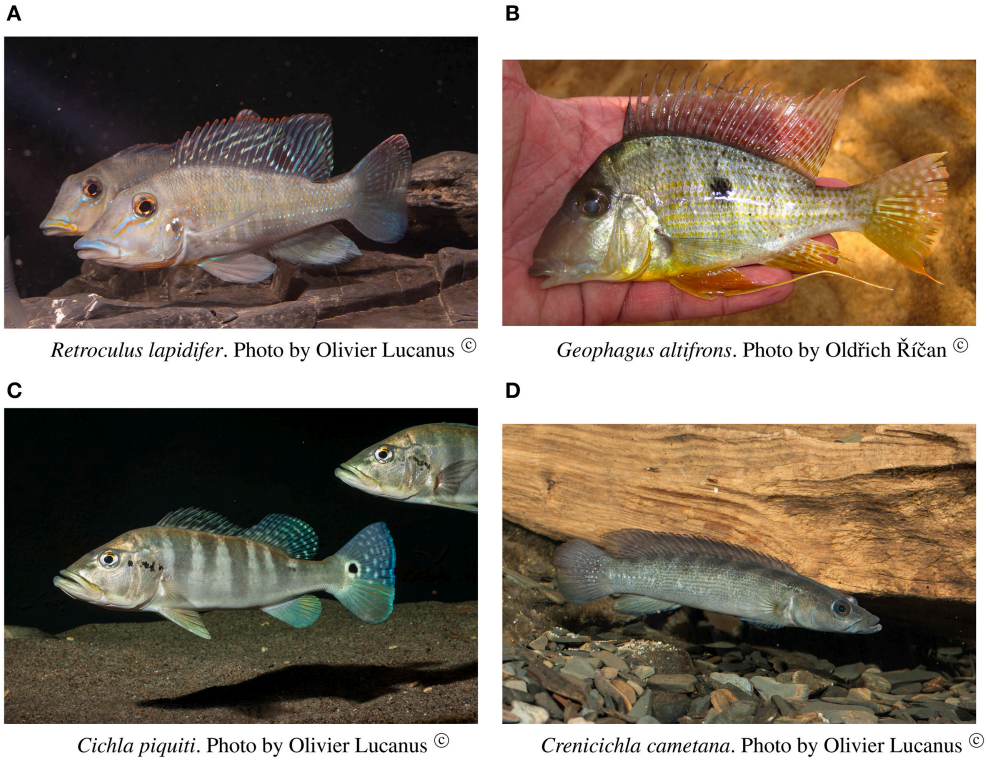
TABLE 4 | Analysis of non-random distribution of lineages among rapids of the Araguaia River.

Species	# of lineages	Fisher's exact test	lineage p–distance (%)
Cichlidae			
Retroculus lapidifer	4	$p = 0.00100$	2.24
Geophagus altifrons	1	NA	1.46
Cichla piquiti	3	$p = 0.08546$	2.58
Crenicichla cameta	5	$p = 0.00050$	1.57
Anostomidae			
Hypomasticus cf. *pachycheilus*	4	$p = 0.00700$	1.04
Leporinus maculatus	4	$p = 0.94250$	0.89
Leporinus affinis	3	$p = 0.16490$	1.50
Leporinus desmotes	2	$p = 1$	7.33
Leporinus unitaeniatus	4	$p = 0.71860$	1.54

negative (-0.80127). The hypothesis of random association of genetic lineages with rapids was rejected (Fisher's exact test, $p = 0.00700$).

Leporinus maculatus (Anostomidae) (Figure 7B)

This species is found in relatively slow and fast water regions of rapids; although it is not a rapids specialist, shows a strong association with rocky substrates. The collections of this species occurred in the localities of Cachoeira do Zé dos Gatos ($N = 4$), Santa Isabel ($N = 9$), Remanso dos Botos ($N = 9$) and Rebojo and São Bento ($N = 6$). We observed four genetic lineages with up to 0.89% p–distance sequence divergence. AMOVA indicated no genetic structuring in the species ($\phi_{ST} = -0.04735$; $p = 0.87851 \pm 0.00363$). Tajima's D was negative (−1.04746). The hypothesis of random association of lineages with the rapids was not rejected (Fisher's exact test, $p = 0.94250$).



Retroculus lapidifer. Photo by Olivier Lucanus ©

Geophagus altifrons. Photo by Oldřich Říčan ©

Cichla piquiti. Photo by Olivier Lucanus ©

Crenicichla cametana. Photo by Olivier Lucanus ©

FIGURE 6 | Cichlid species analyzed in this study.

Leporinus affinis (Anostomidae) (Figure 7C)

This species is a habitat generalist that is found in both slow and fast water regions of rapids, and also in a wide variety of habitats and water flows within the Araguaia-Tocantins river system. The collections of this species occurred in the localities of Santa Isabel ($N = 3$), Remanso dos Botos ($N = 2$) and Rebojo and São Bento ($N = 29$). We observed three divergent genetic lineages with up to 1.50% p–distance sequence divergence. AMOVA indicated absence of geographic structuring ($\phi_{ST} = 0.00075$; $p = 0.43525 \pm 0.00483$). Tajima's D was negative (−1.09098). The hypothesis of random association of lineages with the rapids was not rejected (Fisher's exact test, $p = 0.16490$).

Leporinus desmotes (Anostomidae) (Figure 7D)

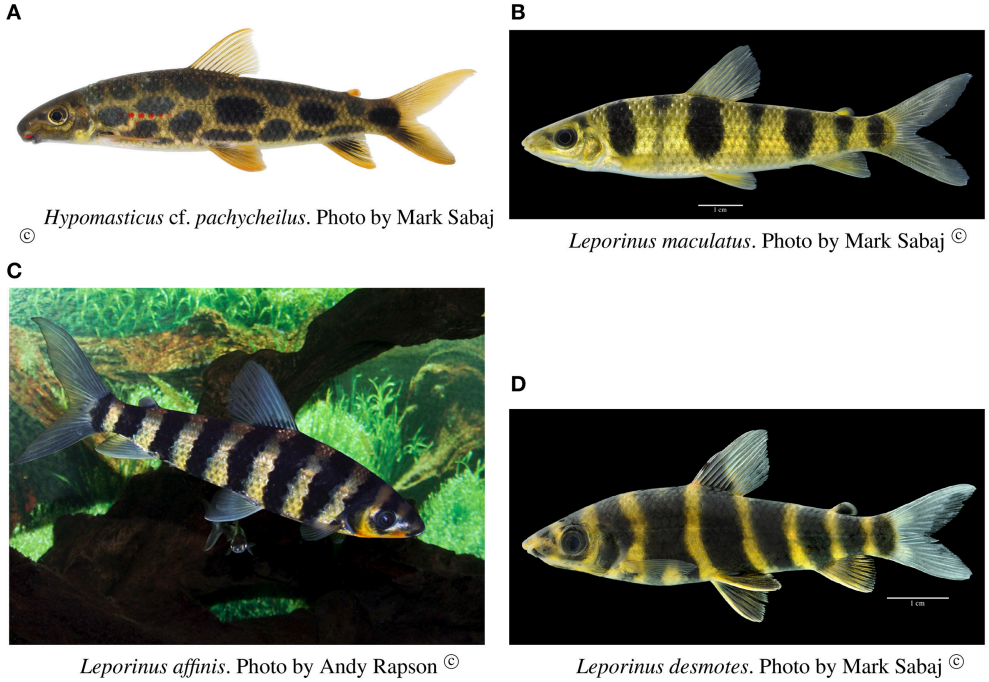
This species is found in both slow and fast water regions of rapids; although it is not a rapids specialist, it shows strong association with rocky substrates. The collections of this species occurred in the localities of Cachoeira do Zé dos Gatos ($N = 2$), Santa Isabel ($N = 2$), São Miguel ($N = 2$), Remanso dos Botos ($N = 4$) and Rebojo and São Bento ($N = 6$). We observed two divergent genetic lineages with upto 7.33% p–distance sequence divergence. AMOVA indicated lack of genetic structuring ($\phi_{ST} = -0.19958$; $p = 0.95950 \pm 0.00191$). Tajima's D was positive (0.96519). The hypothesis of random association of lineages with the rapids was not rejected (Fisher's exact test, $p = 1.00$).

Leporinus unitaeniatus (Anostomidae)

This species is found in both slow and fast water regions of rapids, and although it is not a rapids specialist, it is a rocky substrate dweller. The collections of this species occurred at the localities of Cachoeira do Zé dos Gatos ($N = 4$), Santa Isabel ($N = 4$), Remanso dos Botos ($N = 0$) and Rebojo and São Bento ($N = 11$). We observed two divergent genetic lineages with upto 1.54% p–distance sequence divergence. AMOVA indicated lack of genetic structuring ($\phi_{ST} = 0.02545$; $p = 0.29604 \pm 0.00478$). Tajima's D was negative (−0.19675). The hypothesis of random association of lineages with the rapids was not rejected (Fisher's exact test, $p = 0.71860$).

DISCUSSION

In this study we have shown an unprecedented level of fine-scale population structuring among fish species that occupy rapids stretches of a large Amazonian river. Five of the nine studied species showed either non–random distribution of genetic diversity among rapids, non–random distribution of lineages among rapids or both. Among these, three rheophilic specialists (the cichlids *Retroculus lapidifer* and *Crenicichla cametana* and the anostomid *Hypomasticus* cf. *pachycheilus*) show strong population level differentiation between rapids and non–random distribution of lineages among rapids. The remaining two species of cichlids (*Geophagus altifrons* and *Cichla piquiti*) are not rapids specialists but both are phylopatric and sedentary species which



Hypomasticus cf. *pachycheilus*. Photo by Mark Sabaj ©

Leporinus maculatus. Photo by Mark Sabaj ©

Leporinus affinis. Photo by Andy Rapson ©

Leporinus desmotes. Photo by Mark Sabaj ©

FIGURE 7 | Anostomid species analyzed in this study.

favors the occurrence of genetic structuring in habitat patches along the river. Geneflow among populations of these species occupying distinct rapids is low, in all five instances $Nm<2$ and in three instances $Nm<1$. In addition to the nine analyzed species, two additional species of rheophilic anostomids, *Leporellus vittatus* and *Leporinus bistriatus*, probably exhibit a similar pattern of genetic differentiation as *Hypomasticus* cf. *pachycheilus*. However, this could not be tested because of the low sample size resulting from the difficulty in collecting samples. Finally, *Leporinus desmotes* was comprised of two deeply divergent—7.33% p–distance—lineages that may represent two cryptic species. These results parallel those of the only other fine–scaled population study by Markert et al. (2010) of two rheophilic cichlid fish species collected at five rapids–associated sites of the lower Congo River separated by no more than 100 km. Therefore, significant levels of within river system geographic structuring is likely to be the norm for rheophilic fish fauna inhabiting rapids of not just the Araguaia, but all Amazonian rivers.

In summary, rheophilic and sedentary species such as cichlids showed high levels of population differentiation among rapids, and consequently low levels of geneflow, in some cases sufficiently low to prevent divergence of these populations by genetic drift. These species were also comprised of distinct lineages non–randomly distributed among rapids. When non-random distribution of lineages was observed, in many cases these lineages were allopatric, i.e., there was no overlap in the distribution of these lineages. This signifies that there are some lineages that only occur in specific rapids and/or groups of rapids, i.e., they are micro-endemic. However, none of these lineages are readily identifiable using traditional external morphological traits.

The occurrence of such remarkable population structuring of rheophilic fishes in the rapids along the Araguaia River points to a possible high level of extinction risk for a considerable portion of the Amazon fish fauna, at an unexpected spatial scale. The Amazon River, as other major rivers, is rapidly being developed and its hydroelectric potential is starting to be realized (Lees et al., 2016; Winemiller et al., 2016). However, more so than any other major river, the Amazon and its affluents have been largely unimpacted by large–scale anthropogenic actions and flow through predominantly pristine landscapes (Winemiller et al., 2016). Most of the basin's hydroelectric potential has also been largely untapped, but there is strong pressure to do so both in Brazil as well as other Amazonian countries (Finer and Jenkins, 2012; Fearnside, 2016). At the moment hydroelectric projects have concentrated in regions with very large energy generating potential—e.g., Tucurui or Belo Monte supplying electricity to energy intensive industries—or on those supplying local needs—e.g., Balbina supplying electricity to Manaus and Curua-una supplying electricity to Santarém, both in the Brazilian Amazon.

Dam construction is subject to environmental impact studies which generally consist of faunal and floral surveys. More complex biological processes acting beyond the seasonal scale are generally not addressed by time constraints of the licencing process. Part of the licensing process requires a plan for the mitigation of environmental disturbances caused by the construction and functioning of the hydoelectric complex, although it is questionable to what extent, if any, this goal can be accomplished (Fearnside, 2012). Rarely are molecular tools

employed, and when they are, they do not necessarily address relevant conservation–driven questions. If molecular tools are used, they are predominantly employed to test to what extent, if any, the area of rapids functions as barriers to geneflow to commercially important species with distributions upstream and downstream of the rapids. The presumed motivation is to determine if fish ladders should be constructed or not. The effectiveness of fish ladders in Amazonian context is questionable, however (Pelicice and Agostinho, 2008). In the only case in the Brazilian Amazon where fish ladders were constructed and their effectiveness evaluated—the Santo Antônio hydroelectric complex on the upper Madeira River—many of the migratory species, including the goliath catfishes (*Brachyplatystoma* spp.), that used to migrate upstream of the Teôtonio rapids for reproduction scarcely do so any more, while species previously isolated by the rapids have used the ladders to invade upstream (Cella-Ribeiro et al., 2017).

Molecular tools are not, however, used to evaluate the effects on rheophilic fauna and flora, and consequently to what extent, if any, can the destruction of this fauna and flora can be compensated with mitigatory measures. To our knowledge this study is the first to attempt to specifically address this question.

Implication for Conservation

Rapids–dwelling fishes have ecological specializations uniquely linked to the occupation of these habitats. One of these specializations is the use of podostemaceous plants as a place of shelter, growth, foraging and direct consumption by herbivorous species (Flausino Junior et al., 2016). As an example, strongly rheophilic pacus species (Characiformes: Serrasalmidae), such as those of the genera *Mylesinus* and *Tometes*, present cutting teeth used in leaf pruning of these plants (e.g., Jégu and Santos, 1988; Santos et al., 1997; Vitorino Júnior et al., 2016; Andrade et al., 2017). The strong dependence on podostemaceous mats as food makes these pacus particularly vulnerable to the environmental impacts resulting from dam construction. The reduction of densities and numbers of the Podostemaceae in rivers with artificially regulated flow seems to have caused a significant population reduction of the pacu *Mylesinus paraschomburgi*, a specialized consumer of these plants (Santos et al., 1997). Similar responses to environmental perturbations have been observed in other rheophilic fish species, such as the anostomids (Santos et al., 1997), loricariids (Zuanon, 1999) and cichlids (Kullander, 1988, 1991; Zuanon, 1999).

In addition to the Podostemaceae, many rapids–dwelling fish species use epiliton, a layer of organic debris, algae and small invertebrates that cover the rocky substratum in the rapids, as food. There is evidence that rheophilic fishes are strongly dependent on this type of food (Casatti and Castro, 1998; Zuanon, 1999), and environmental impacts that alter the quality and quantity of this resource negatively affect the local ichthyofauna, with loss of species and trophic relationships.

In addition to the close and direct ecological relationship between fishes and rapids, the spatial distribution of rheophilic habitats within a course of a river also contributes to the fragility of the system. In many rivers, rapids are environments distributed in a mosaic, isolated from each other by other habitat types unsuitable for rheophilic fauna and flora. Not only are rheophilic habitats characterized by high levels of endemism and taxonomic diversity, our study suggests that rheophilic habitats are also characterized by high levels of populational structuring and phylogenetic diversity. Thus while fish assemblages and communities may be similar among the different rapids of a same river, our study indicates that they are composed of different lineages, and in some cases even ecologically equivalent but morphologically cryptic species.

Areas of rapids of several Amazonian rivers have already been radically altered by the construction of hydroelectric dams. In all these cases, the rapids and much of the associated ichthyofauna disappeared (Agostinho et al., 2008), and based on our study, this fauna is not substituitable by fauna occuring outside the area of direct impact of the hydroelectric dam. The construction of reservoirs causes a permanent change of lotic to lentic habitats, ensuing in the local extinction of the specialized rheophilic fauna and flora whose habitat became submerged under the hydroelectric reservoir. The elevated levels of phylogenetic diversity, the non-random distribution of this diversity among the rapids, with lineages often being restricted to a specific rapid or groups of rapids, implies that the disappearance of the rheophilic ichthyofauna from areas of hydroelectric projects probably leads to the extiction of these lineages, even if other lineages survive elsewhere.

We May Ask: Why Does This Matters?

Although ecologically equivalent lineages may persist in other areas, we are loosing unique and singular evolutionary heritage. Once extinct, these lineages can never become "unextinct." Even more importantly, these lineages are an inevitable consequence of evolution, and thus are intricately linked to the underlying evolutionary processes that gave rise to them. By destroying these lineages, we, inevitable, impair or even destroy our potential to understand these processes and the origins of this amazing biodiversity.

An argument may also be made that we should not focus on lineages, but rather on species, since species are the fundamental units of biodiversity (Soulé and Wilcox, 1980). However, species are lineages (Simpson, 1961; de Queiroz, 2007), and need to be understood in this light. If we treat species as classes, or do not recognize that species may be morphologically cryptic, often relying on incomplete taxonomic knowledge, we limit ourselves in what we can understand of the evolutionary history and the evolutionary potential of the studied groups.

Ultimately we must recognize the lineage-like nature of species and focus our studies on them if we hope to understand the evolution and maintenance of biodiversity (Willis, 2017), and if we hope to minimize our contribution to the sixth mass biological extinction (Ceballos et al., 2015).

AUTHOR CONTRIBUTIONS

TH, IF, and JZ conceived the experiment and obtained funding. TH, NM, and JZ conducted fieldwork and collected specimens. JZ morphotyped and revised specimens. NM and TH collected molecular data. TH analyzed the results. TH wrote the

manuscript with substantial contributions from JZ. All authors contributed to and reviewed the manuscript.

FUNDING

Funding for this study was provided by the GESAI consortium, CNPq 483155/2010-1, and 482662/2013-1 to TH.

ACKNOWLEDGMENTS

We thank Marla Soares Carvalho and Kiara Formiga for helping to collect and process fish in the field, and Oliver Lucanius, Mark Sabaj, Andry Rapson, and Oldřich Říčan for representative photos of studied species.

REFERENCES

Abell, R. (2002). Conservation biology for the biodiversity crisis: a freshwater follow-up. *Conserv. Biol.* 16, 1435–1437. doi: 10.1046/j.1523-1739.2002.01532.x

Agostinho, Â. A., Pelicice, F. M., and Gomes, L. C. (2008). Dams and the fish fauna of the Neotropical region: impacts and management related to diversity and fisheries. *Braz. J. Biol.* 68(4 Suppl):1119–1132. doi: 10.1590/S1519-69842008000500019

Alofs, K. M., Liverpool, E. A., Taphorn, D. C., Bernard, C. R., and López-Fernández, H. (2014). Mind the (information) gap: the importance of exploration and discovery for assessing conservation priorities for freshwater fish. *Divers. Distribut.* 20, 107–113. doi: 10.1111/ddi.12127

Andrade, M. C., Machado, V. N., Jégu, M., Farias, I. P., and Giarrizzo, T. (2017). A new species of *Tometes* Valenciennes 1850 (Characiformes: Serrasalmidae) from Tocantins–Araguaia River Basin based on integrative analysis of molecular and morphological data. *PLoS ONE* 12:e0170053. doi: 10.1371/journal.pone.0170053

Andrade, M. C., Sousa, L. M., Ota, R. P., Jégu, M., and Giarrizzo, T. (2016). Redescription and geographical distribution of the endangered fish *Ossubtus xinguense* Jégu 1992 (Characiformes, Serrasalmidae) with comments on conservation of the rheophilic fauna of the Xingu River. *PLoS ONE* 11:e0161398. doi: 10.1371/journal.pone.0161398

Balon, E. K. (1974). Fishes from the edge of Victoria Falls, Africa: demise of a physical barrier for downstream invasions. *Copeia* 1974, 643–660. doi: 10.2307/1442678

Balon, E. K., and Stewart, D. J. (1983). Fish assemblages in a river with unusual gradient (Luongo, Africa - Zaire system), reflections on river zonation, and description of another new species. *Environ. Biol. Fish.* 9, 225–252. doi: 10.1007/BF00692373

Böhlke, J. E., Weitzman, S. H., and Menezes, N. A. (1978). Estado atual da sistemática dos peixes de água doce da América do Sul. *Acta Amazon.* 8, 657–677.

Casatti, L., and Castro, R. M. C. (1998). A fish community of the São Francisco River headwaters riffles, southeastern Brazil. *Ichthyol. Explor. Freshw.* 9, 229–242.

Castello, L., and Macedo, M. N. (2016). Large-scale degradation of Amazonian freshwater ecosystems. *Glob. Change Biol.* 22, 990–1007. doi: 10.1111/gcb.13173

Castello, L., McGrath, D. G., Hess, L. L., Coe, M. T., Lefebvre, P. A., Petry, P., et al. (2013). The vulnerability of Amazon freshwater ecosystems. *Conserv. Lett.* 6, 217–229. doi: 10.1111/conl.12008

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., and Palmer, T. M. (2015). Accelerated modern human-induced species losses: entering the sixth mass extinction. *Sci. Adv.* 1:e1400253. doi: 10.1126/sciadv.1400253

Cella-Ribeiro, A., da Costa Doria, C. R., Dutka-Gianelli, J., Alves, H., and Torrente-Vilara, G. (2017). Temporal fish community responses to two cascade run-of-river dams in the Madeira River, Amazon basin. *Ecohydrology* 10, 1–11. doi: 10.1002/eco.1889

Clausen, R., and York, R. (2008). Global biodiversity decline of marine and freshwater fish: a cross-national analysis of economic, demographic, and ecological influences. *Soc. Sci. Res.* 37, 1310–1320. doi: 10.1016/j.ssresearch.2007.10.002

Collen, B., Whitton, F., Dyer, E. E., Baillie, J. E. M., Cumberlidge, N., Darwall, W. R. T., et al. (2014). Global patterns of freshwater species diversity, threat and endemism. *Glob. Ecol. Biogeogr.* 23, 40–51. doi: 10.1111/geb.12096

Collins, R. A., Ribeiro, E. D., Sousa, L. M., and Hrbek, T. (2018). Hydroelectric dams on the Xingu and Tapajós rivers imperil a diverse and unique fauna of suckermouth catfishes. *Aquat. Conserv.: Mar. Freshw. Ecosyst.*

Corander, J., Marttinen, P., Sirén, J., and Tang, J. (2008). Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* 9:539. doi: 10.1186/1471-2105-9-539

Darwall, W. R. T., Holland, R. A., Smith, K. G., Allen, D., Brooks, E. G. E., Katarya, V., et al. (2011). Implications of bias in conservation research and investment for freshwater species. *Conserv. Lett.* 4, 474–482. doi: 10.1111/j.1755-263X.2011.00202.x

de Queiroz, K. (2007). Species concepts and species delimitation. *Systemat. Biol.* 56, 879–886. doi: 10.1080/10635150701701083

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Excoffier, L., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.

Fearnside, P. M. (2012). *Belo Monte Dam: A Spearhead for Brazil's dam Building Attack on Amazonia?, GWF Discussion Paper 1210*. Technical report, Global Water Forum, Canberra, ACT.

Fearnside, P. M. (2015). Amazon dams and waterways: Brazil's Tapajós Basin plans. *Ambio* 44, 426–439. doi: 10.1007/s13280-015-0642-z

Fearnside, P. M. (2016). Brazilian politics threaten environmental policies. *Science* 353, 746–748. doi: 10.1126/science.aag0254

Finer, M., and Jenkins, C. N. (2012). Proliferation of hydroelectric dams in the Andean Amazon and implications for Andes-Amazon connectivity. *PLoS ONE* 7:e35126. doi: 10.1371/journal.pone.0035126

Fitzgerald, D. B., Sabaj Perez, M. H., de Sousa, L. M., Gonçalves, A. P., Rapp Py-Daniel, L., Lujan, N. K., et al. (2018). Diversity and community structure of rapids-dwelling fishes of the Xingu River: implications for conservation amid large-scale hydroelectric development. *Biol. Conserv.* 222, 104–112. doi: 10.1016/j.biocon.2018.04.002

Flausino Junior, N., Machado, F. A., Zuanon, J., and Ferreira, E. J. G. (2016). The fish fauna of sessile hydrophytes stands (*Mourera* spp.: Podostemaceae) in the Dardanelos Waterfalls, Rio Aripuanã, Brazil. *Aqua Int. J. Ichthyol.* 22, 133–144.

Frederico, R. G., Olden, J. D., and Zuanon, J. (2016). Climate change sensitivity of threatened, and largely unprotected, Amazonian fishes. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 26, 91–102. doi: 10.1002/aqc.2658

Frederico, R. G., Zuanon, J., and De Marco, P. (2018). Amazon protected areas and its ability to protect stream-dwelling fish fauna. *Biol. Conserv.* 219, 12–19. doi: 10.1016/j.biocon.2017.12.032

Hawkins, C. P., Kershner, J. L., Bisson, P. A., Bryant, M. D., Decker, L. M., Gregory, S. V., et al. (1993). A hierarchical approach to classifying stream habitat features. *Fisheries* 18, 3–12. doi: 10.1577/1548-8446(1993)018<0003:AHATCS>2.0.CO;2

Hora, S. L. (1930). Ecology, bionomics and evolution of the torrential fauna, with special reference to the organs of attachment. *Philos. Trans. R. Soc. Lond. Ser. B* 218, 171–282. doi: 10.1098/rstb.1930.0005

International Energy Agency (2013a). *World Energy Outlook – 2013*. Paris: International Energy Agency.

International Energy Agency (2013b). *World Energy Resources – 2013 Survey*. Paris: International Energy Agency.

Isbrücker, I. J. H., and Nijssen, H. (1991). *Hypancistrus zebra*, a new genus and species of uniquely pigmented ancistrinae loricariid fish from the Rio Xingu, Brazil (Pisces: Siluriformes: Loricariidae). *Ichthyol. Explor. Freshw.* 1, 345–350.

Jégu, M. (1992). *Ossubtus xinguensis*, nouveaux genre et espèce du Rio Xingu, Amazonie, Brésil (Teleostei: Serrasalmidae). *Ichthyol. Explor. Freshw.* 3, 235–252.

Jégu, M., and Santos, G. M. d. (1988). Une nouvelle espèce du genre *Mylesinus* (Pisces, Serrasalmidae), *Mylesinus paucisquamatus*, décrite du bassin du rio Tocantins (Amazonie, Brésil). *Cybium* 12, 331–341.

Jégu, M. L. A. M. F. (2004). *Taxinomie des Serrasalminae Phytophages et Phylogenie des Serrasalminae (Teleostei: Characiformes: Characidae)*. Ph.D. thesis, Paris.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kullander, S. O. (1988). *Teleocichla*, a new genus of South American rheophilic cichlid fishes with six new species (Teleostei: Cichlidae). *Copeia* 1988, 196–230. doi: 10.2307/1445938

Kullander, S. O. (1991). *Crenicichla phaiospilus* and *C. percina*, two new species of pike cichlids (Teleostei: Cichlidae) from the rio Xingu, Brazil. *Ichthyol. Explor. Freshw.* 1, 351–360.

Latrubesse, E. M., Arima, E. Y., Dunne, T., Park, E., Baker, V. R., D'Horta, F. M., et al. (2017). Damming the rivers of the Amazon basin. *Nature* 546, 363–369. doi: 10.1038/nature22333

Lees, A. C., Peres, C. A., Fearnside, P. M., Schneider, M., and Zuanon, J. A. S. (2016). Hydropower and the future of Amazonian biodiversity. *Biodivers. Conserv.* 25, 451–466.

Lévêque, C., Oberdorff, T., Paugy, D., Stiassny, M. L. J., and Tedesco, P. A. (2007). Global diversity of fish (Pisces) in freshwater. *Hydrobiologia* 595, 545–567. doi: 10.1007/s10750-007-9034-0

Lowe-McConnell, R. H. (1987). *Ecological Studies of Tropical Fish Communities*. Cambridge: Cambridge University Press.

Lujan, N. K., Cramer, C. A., Covain, R., Fisch-Muller, S., and López-Fernández, H. (2017). Multilocus molecular phylogeny of the ornamental wood-eating catfishes (Siluriformes, Loricariidae, *Panaqolus* and *Panaque*) reveals undescribed diversity and parapatric clades. *Mol. Phylogenet. Evol.* 109, 321–336. doi: 10.1016/j.ympev.2016.12.040

Lujan, N. K., Winemiller, K. O., and Armbruster, J. W. (2012). Trophic diversity in the evolution and community assembly of loricariid catfishes. *BMC Evol. Biol.* 12:124. doi: 10.1186/1471-2148-12-124

Lundberg, J. G., Kottelat, M., Smith, G. R., Stiassny, M. L. J., and Gill, A. C. (2000). So many fishes, so little time: an overview of recent ichthyological discovery in continental waters. *Ann. Missouri Bot. Garden* 87, 26–62. doi: 10.2307/2666207

Machado, V. N., Collins, R. A., Ota, R. P., Andrade, M. C., Farias, I. P., and Hrbek, T. (2018). One thousand DNA barcodes of piranhas and pacus reveal geographic structure and unrecognised diversity in the Amazon. *Sci. Rep.* 8:8387. doi: 10.1038/s41598-018-26550-x

Markert, J. A., Schelly, R. C., and Stiassny, M. L. J. (2010). Genetic isolation and morphological divergence mediated by high-energy rapids in two cichlid genera from the lower Congo rapids. *BMC Evol. Biol.* 10:149. doi: 10.1186/1471-2148-10-149

Menezes, N. A. (1996). "Methods for assessing freshwater fish diversity," in *Biodiversity in Brazil - A First Approach*, eds C. E. de Bicudo and N. A. Menezes (Brasília: CNPq), 289–296.

Meyer, A. (1993). "Evolution of mitochondrial DNA in fishes," in *Molecular Biology Frontiers: Biochemistry and Molecular Biology of Fishes*. Vol. 2. eds P. W. Hochachka and T. P. Mommsen (Amsterdam: Elsevier), 1–38.

Pelicice, F. M., and Agostinho, A. A. (2008). Fish-passage facilities as ecological traps in large Neotropical rivers. *Conserv. Biol.* 22, 180–188. doi: 10.1111/j.1523-1739.2007.00849.x

Pelicice, F. M., Azevedo-Santos, V. M., Vitule, J. R. S., Orsi, M. L., Lima Junior, D. P., Magalhães, A. L. B., et al. (2017). Neotropical freshwater fishes imperilled by unsustainable policies. *Fish Fish.* 18, 1119–1133. doi: 10.1111/faf.12228

Platt, A. R., Woodhall, R. W., and George, A. L. Jr. (2007). Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol. *Biotechniques* 43, 58–62. doi: 10.2144/000112499

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Reis, R. E., Kullander, S. O., and Ferraris Jr, C. J. (2003). *Check List of the Freshwater Fishes of South and Central America*. Porto Alegre: EDIPUCRS.

Roberts, T. R., and Stewart, D. J. (1976). An ecological and systematic survey of fishes in the rapids of the lower Zaïre or Congo River. *Bull. Museum Comp. Zool. Harvard Univ.* 147, 239–317.

Roxo, F. F., Lujan, N. K., Tagliacollo, V. A., Waltz, B. T., Gabriel, S., Silva, C., et al. (2017). Shift from slow- to fast-water habitats accelerates lineage and phenotype evolution in a clade of Neotropical suckermouth catfishes (Loricariidae: hypoptopomatinae). *PLoS ONE* 12:e0178240. doi: 10.1371/journal.pone.0178240

Salzburger, W., Ewing, G. B., and Von Haeseler, A. (2011). The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Mol. Ecol.* 20, 1952–1963. doi: 10.1111/j.1365-294X.2011.05066.x

Salzburger, W., Van Bocxlaer, B., and Cohen, A. S. (2014). Ecology and evolution of the African Great Lakes and their faunas. *Ann. Rev. Ecol. Evol. Systemat.* 45, 519–545. doi: 10.1146/annurev-ecolsys-120213-091804

Sambrook, J., and Russell, D. (2001). *Molecular Cloning: A Laboratory Manual, 3rd Edn*. Cold Springs Harbor, NY: Cold Springs Harbor Laboratory Press.

Santos, G. M. D., and Jégu, M. (1987). Novas ocorrências de *Gnathodolus bidens*, *Synaptolaemus cingulatus* e descrição de duas espécies novas de *Sartor* (Characiformes, Anostomidae). *Amazoniana* 10, 181–196.

Santos, G. M. D., Pinto, S. S., and Jégu, M. (1997). Alimentação do pacu-cana, *Mylesinus paraschomburgkii* (Teleostei, Serrasalmidae) em rios da Amazônia brasileira. *Rev. Brasil. Biol.* 57, 311–315.

Simpson, G. G. (1961). *Principles of Animal Taxonomy*. New York, NY: Columbia University Press.

Soulé, M. E., and Wilcox, B. (1980). *Conservation Biology: An Evolutionary-Ecological Perspective*. Sunderland, MA: Sinauer Associates, Inc.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Vitorino Júnior, O. B., Agostinho, C. S., and Pelicice, F. M. (2016). Ecology of *Mylesinus paucisquamatus* Jégu & Santos, 1988, an endangered fish species from the rio Tocantins basin. *Neotrop. Ichthyol.* 14:e150124. doi: 10.1590/1982-0224-20150124

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., et al. (2010). Global threats to human water security and river biodiversity. *Nature* 467, 555–561. doi: 10.1038/nature09440

Werle, E., Schneider, C., Renner, M., Völker, M., and Fiehn, W. (1994). Convenient single-step, one tube purification of PCR products for direct sequencing. *Nucleic Acids Res.* 22, 4354–4355. doi: 10.1093/nar/22.20.4354

Willis, S. C. (2017). One species or four? Yes!... and, no. Or, arbitrary assignment of lineages to species obscures the diversification processes of Neotropical fishes. *PLoS ONE* 12:e0172349. doi: 10.1371/journal.pone.0172349

Winemiller, K. O., McIntyre, P. B., Castello, L., Fluet-Chouinard, E., Giarrizzo, T., Nam, S., et al. (2016). Balancing hydropower and biodiversity in the Amazon, Congo, and Mekong. *Science* 351, 128–129. doi: 10.1126/science.aac7082

Zuanon, J. A. S. (1999). *História Natural da Ictiofauna de Corredeiras do rio Xingú, na região de Altamira, Pará*. Ph.D. thesis, Campinas.

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 14 August 2018
doi: 10.3389/fgene.2018.00263

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

Maria da Conceição Freitas Santos[1], *Tomas Hrbek*[2] *and Izeni P. Farias*[2]*

[1] Departamento de Biologia, Universidade do Estado do Amazonas, Manaus, Brazil, [2] Laboratório de Evolução e Genética Animal, Departamento de Genética, Universidade Federal do Amazonas, Manaus, Brazil

OPEN ACCESS

Edited by:
Rodrigo A. Torres, Universidade Federal de Pernambuco, Brazil

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

***Correspondence:**
Izeni P. Farias
izeni@evoamazon.net

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

Received: *08 November 2017*
Accepted: *28 June 2018*
Published: *14 August 2018*

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

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

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

INTRODUCTION

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

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

Previous studies have found a high degree of genetic variability of populations of *C. macropomum* (Santos et al., 2007; Farias et al., 2010), indicating that overfishing has not yet affected genetic diversity of wild populations, nor were signals of population reduction detectable. The authors suggested that the absence of a genetic sign of population reduction was likely related to the large effective population size of the species. Moreover, *C. macropomum* has migratory behavior and moves through the rivers of the Amazon seasonally for the purposes of feeding and breeding (Araújo-Lima and Ruffino, 2004). The behavior of *C. macropomum* and the hydrological dynamic of the floodplain habitat it predominantly occupies may partially explain the panmixia reported for this species. This apparent lack of population structuring is found throughout the mainstream of the Amazon basin, with the exception of individuals found upstream of the series of rapids delimiting the Bolivian sub-basin from the Amazon basin (Farias et al., 2010). The authors also suggested that populations of *C. macropomum* from Bolivian sub-basin were largely demographically stable, while the Brazilian Amazon basin populations evidenced a historical population growth from the Pleistocene onward.

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

MATERIALS AND METHODS

Samples and Data Collection

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

Total genomic DNA was extracted using Proteinase K/Phenol-chloroform/isoamyl alcohol protocol and precipitated with 70% ethanol (Sambrook et al., 1989). Approximately 50 to 100 ng of genomic DNA was used as a template for PCR reactions. We amplified the mitochondrial DNA control region (mtDNA control region) and the ATPase subunits 6 and 8, using the primers Chara_LDloop and Chara_RDloop; CMF2 and CMR2 (control region) and ATP 8.2_L8331 and CO3.2_H9236 (ATPase genes) listed in **Supplementary Table S1**. The PCR reactions for the two regions were performed in a final volume of 15 μL containing 1.5 μl of the forward and reverse primer (2 mM), 1.5 μl of buffer (Tris-KCL 200 mM, pH 8.5), 1.5 μL of 25 mM MgCl, 1.5 μL of 25 mM dNTP, 0.3 μL of 5 U/μL Taq polymerase and 6.2 μL ddH_2O. PCR conditions (for control region and ATPase gene) were as follows: denaturation at 94°C for 60 s, primer annealing

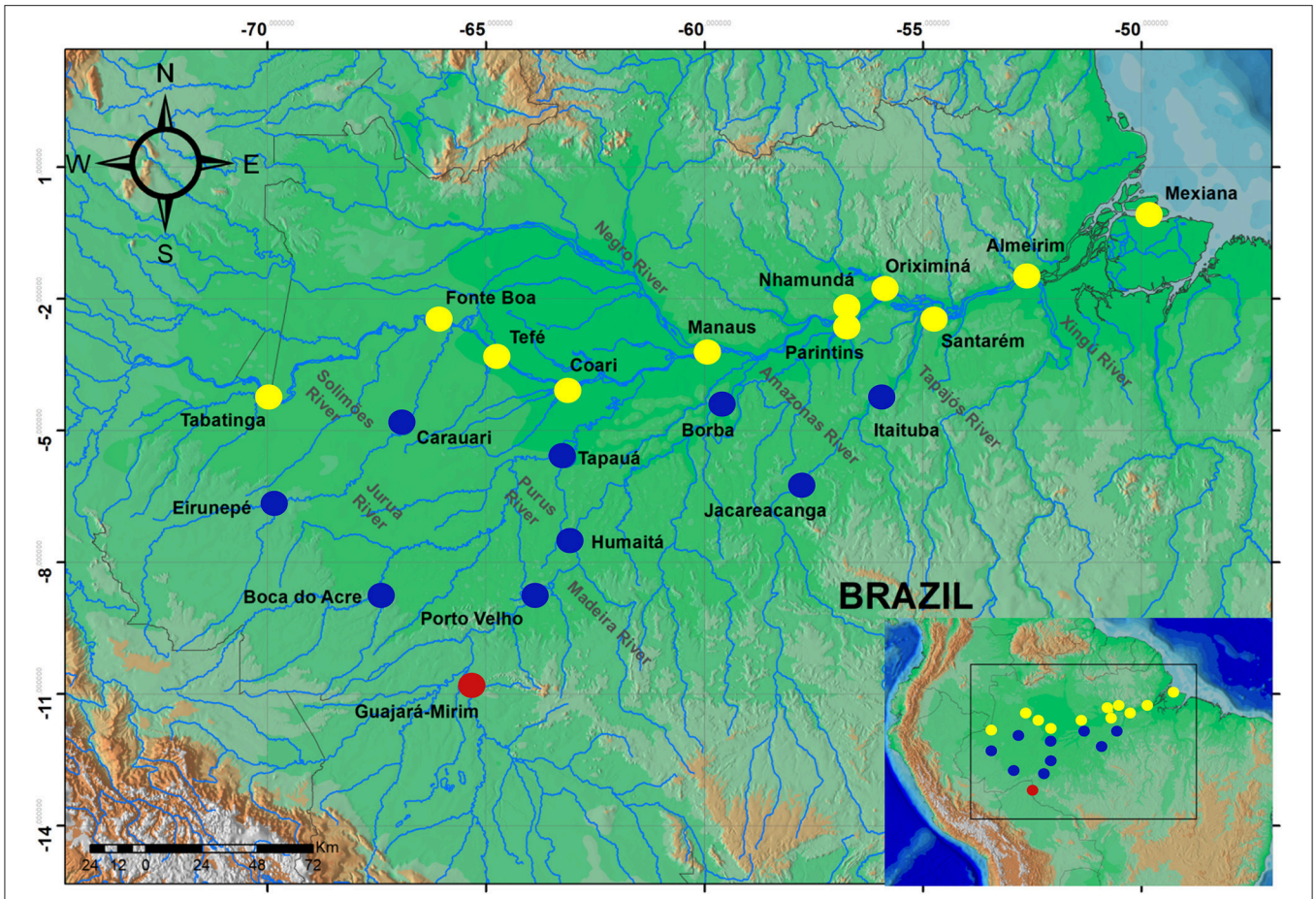


FIGURE 1 | Map of the Amazon basin showing sampled localities. Circles represent localities in the mainstream of Amazon River (yellow), tributaries of the Amazon River (blue), and the locality of Bolivian sub-basin (red).

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

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

Microsatellite genotypes were generated using a multiplex design (**Supplementary Table S2**) using 13 pairs of primers developed by Santos et al. (2009) for *C. macropomum*. The amplification conditions for each multiplex were: For three pairs of primers: 1.5 μl $MgCl_2$ (25 mM), 1.5 μl dNTPs (10 mM), 1.5 μl 10x buffer (100 mM Tris-HCl, 500 mM KCl), 1.0 μL of each forward primer containing one of the two M13 tails (2 μM), 1.5 μL of each reverse primer, 1.5 μL of fluorescence-labeled M13f (FAM) primer, 0.7 μL of primer fluorescence-labeled M13r (HEX), 0.8 μl Taq DNA Polymerase (5 U/μl) and 1 μl DNA (50–100 ng), with a final volume of 14.5 μl. For two pairs of primers: 3.0 μl ultra pure water, 1.5 μl $MgCl_2$ (25 mM), 1.5 μl dNTPs (10 mM), 1.5 μl 10x buffer (100 mM Tris-HCl, 500 mM KCl), 1.0 μL of each forward primer containing one of the M13 tails (2 μM), 1.5 μL of each reverse primer, 1.5 μL of the fluorescently labeled M13f primer (FAM), or 0.7 μL of fluorescently labeled M13r primer (HEX), 0.6 μl of Taq DNA Polymerase (5 U/μl), and 1 μl of the DNA (50–100 ng), with a final volume of 14 μl.

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

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

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

Mitochondrial DNA Analyses

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

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

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

The Tajima's *D* (Tajima, 1989) and the Fu's *Fs* (Fu, 1997) tests were used to examine whether populations are at a mutation-drift equilibrium assuming no selective differences among haplotypes. Both tests were performed using the program ARLEQUIN v3.5 (Excoffier and Lischer, 2010). Demographic history may also be inferred from frequency distribution of pairwise haplotype differences. In populations that are at a demographic equilibrium, the distribution of differences are generally multimodal, while populations that have undergone recent expansion or reduction typically have a unimodal distribution (Slatkin and Hudson, 1991). In order to distinguish population reduction and expansion, we used the results of two tests. The first test evaluates the distribution of the sum of the squares of differences (SSD) between the mismatch distribution observed for each locality and the expected distribution for a null expansion model, where significant values for SSD indicate deviations from the population expansion model (Schneider and Excoffier, 1999). The other test is based on the Harpending inequality index ($Hri = r$) (Harpending, 1994), which quantifies the variance of the mismatch distribution, assuming that the mismatch distribution is unimodal. These analyzes were performed in the programs DNASP v5.0 (Librado and Rozas, 2009) and ARLEQUIN v3.5 (Excoffier and Lischer, 2010); significance was tested via 10,000 permutations with a $P = 0.05$ cut-off.

Microsatellite DNA Analyses

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

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

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

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

Coalescent analyses implemented in the program IMa2 (Hey and Nielsen, 2007) were used to partition allele sharing between populations due to ongoing geneflow and ancestral haplotype sharing. We estimated the parameters t (splitting time), m (migration rates), and theta (θ) where $\theta = 2\ Ne\mu$. We sampled 20,000,000 Monte Carlo Chain Markov Chain Monte Carlo (MCMC) generations after discarding the first 1,000,000 generations as burn-in. Two independent runs were carried out with different starting points, in order to verify convergence. The two independent runs converged and thus were combined and the parameters θ, m, and t were estimated. Then, these were converted into demographic parameters: contemporary effective population size, number of migrants per generation, and time of divergence of the populations in generations. The analyses using IMa2 were performed with pairs of sampling sites located at the geographic extremes along the main channel of the Amazon River and between upstream and downstream localities of principal tributaries. We also estimated geneflow using the program MIGRATE 3.1.6 (Beerli and Felsenstein, 2001), where for diploid data $\theta = 4\ Ne\mu$ and $M = \mathrm{m}/\mu$ migration rate ratio and mutation rate. For the Bayesian analysis, we ran ten short chains, sampling each chain 10,000 times. Then we ran six long chains of 2,000,000 steps, sampling each chain 200,000 times and discarding the first 2,000 samples. The runs were replicated, and the convergence between the chains was evaluated using the Gelman-Rubin statistic implemented in the program. We estimated the historical migration rates (M) between the localities and the relative number of migrants per generation $Nm = \mathrm{M}\theta/2$. To convert the results into biological information, we assumed a 3-year generation time for *C. macropomum* (previously justified with mtDNA data), and a mutation rate $\mu = 5 \times 10^{-4}$ (mean rate of evolution of microsatellites, Di Rienzo et al., 1994).

To detect, quantify, and date the historical and contemporaneous demographic changes in *C. macropomum* populations we implemented the coalescent sampler implemented in the program MSVar v1.3 (Beaumont, 1999; Storz et al., 2002). We ran 11 independent parallel chains sampling every 1,000th proposal to collect 20,000 proposals in the MCMC chain in each parallel run. Priors for current and historical population size means and variances were equal, and variances encompassed three orders of magnitude. Prior for mean time of population size change was set at 1,000 with variance encompassing time range from 1,000,000 to 0 generations ago. The runs were evaluated for convergence and were pooled to provide an estimate of current and historical effective population size. Convergence was assessed using the Gelman–Rubin criterion (Gelman and Rubin, 1992) and the test of alternative hypotheses (population decline vs. stable population size) as suggested by Beaumont (1999) was tested using Bayes factors. Calculations and plots were performed in the R statistical programming language (R Development Core Team, 2011) using the packages CODA and ggplot2.

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

RESULTS

Genetic Diversity of *C. macropomum*

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

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

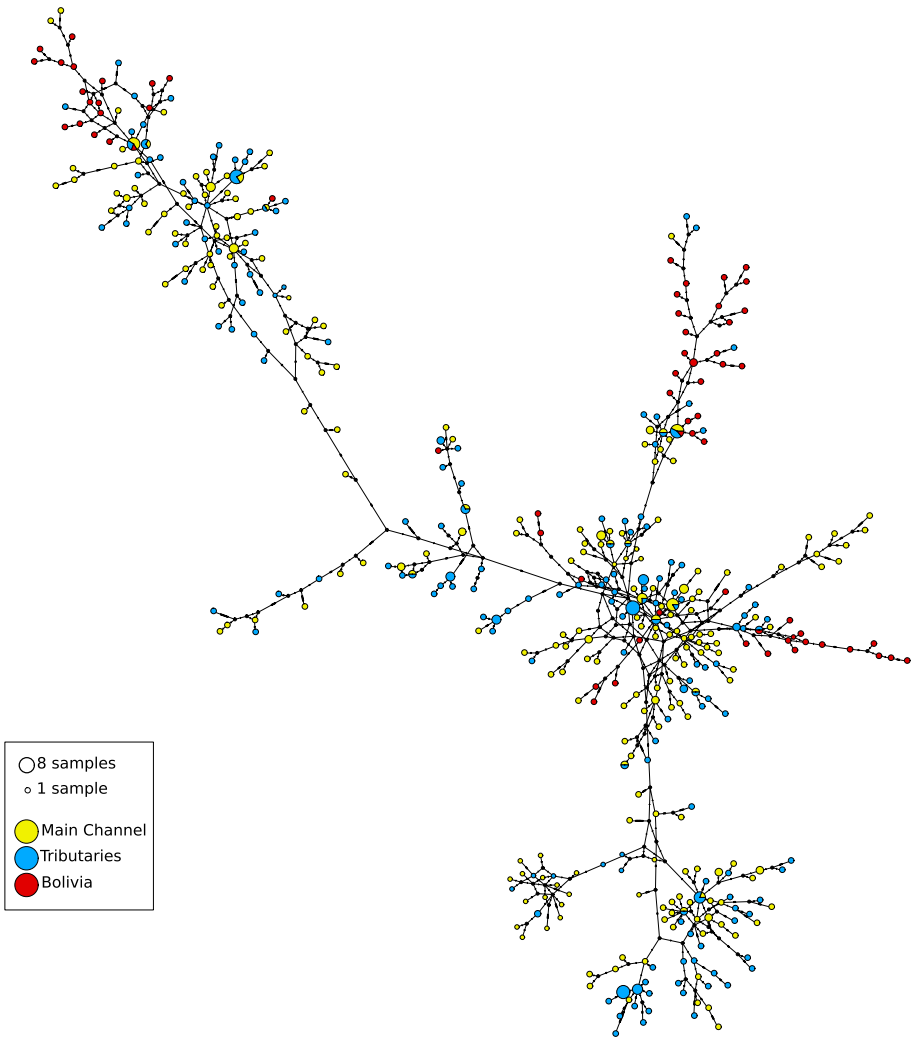


FIGURE 2 | Haplotype network of *Colossoma macropomum* haplotypes estimated using Network. Each line represents a single mutation. Circle size correspond to the number of observations, and missing haplotypes remain unfilled. Colors corresponds to observation of a haplotype in one of three main regions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

Population Structure

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

Global AMOVA of microsatellite data resulted in partitioning more than 98% variance within sites ($F_{ST} = 0.0111$, $P = 0.00124$). Based in this result, additional AMOVA tests were implemented assuming two main groups: Amazon basin vs. Bolivia basin and, within the Amazon basin, tributaries vs. main channel. The AMOVA results were significant for both analyses ($F_{ST} = 0.0192$, $P = 0.0478$; $F_{ST} = 0.0086$, $P = 0.9277$; $F_{ST} = 0.0026$, $P = 0.0004$), respectively.

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

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

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

				mtDNA										Microsatellites				
Basin	**Locality**	**River**	**Water type***	**N**	***H***	***S***	***h***	**π**	**N**	**A**	**Ana**	**AR**	**PAR**	**AGD**	**Ho-He**	**FIS**	**Ne (Cis)**	**p_SMM**	**Mvalue**
	Mexiana	Amazon	White	15	15	14	1.000 ± 0.024	0.009 ± 0.005	31	139	11.5	6.23	0.12	0.791 ± 0.407	0.779–0.791	**0.049**	4183.1 (140.7-inf)	**0.040**	0.71
	Almeirim	Amazon	White	21	20	14	0.995 ± 0.016	0.009 ± 0.005	36	150	12.5	6.31	0.13	0.781 ± 0.402	0.754–0.781	**0.070**	−234.4 (1094.3-inf)	**0.008**	0.80
	Santarém	Lower Tapajós	Clear	26	26	24	1.000 ± 0.010	0.012 ± 0.006	36	143	11.9	6.29	0.13	0.771 ± 0.397	0.680–0.771	**0.144**	−266.3 (845.1-inf)	0.243	0.78
	Itaituba	Tapajós	Clear	32	24	15	0.977 ± 0.014	0.011 ± 0.006	32	138	11.5	6.01	0.20	0.762 ± 0.393	0.734–0.762	**0.102**	476.9 (102.6-inf)	**0.011**	0.68
	Jacareacanga	Tapajós	Clear	29	19	6	0.958 ± 0.020	0.009 ± 0.005	30	134	11.1	6.04	0.13	0.757 ± 0.391	0.730–0.757	**0.078**	−2610.5 (169.4-inf)	0.277	0.79
	Oriximiá	Lower Trombetas	Clear	34	34	30	1.000 ± 0.007	0.012 ± 0.006	32	139	11.5	6.04	0.12	0.773 ± 0.399	0.715–0.773	**0.102**	275.7 (89.5-inf)	0.118	0.79
	Nhamundá	Lower Nhamundá	Black	22	21	16	0.995 ± 0.015	0.011 ± 0.005	22	127	10.5	6.25	0.19	0.766 ± 0.398	0.727–0.766	**0.078**	100.3 (45.1-inf)	**0.034**	0.74
	Parintins	Amazon	White	20	20	14	1.000 ± 0.015	0.012 ± 0.006	35	148	12.3	6.13	0.17	0.768 ± 0.396	0.719–0.768	**0.081**	96.2 (57.9-233)	**0.012**	0.76
Amazon	Borba	Madeira	White	19	19	13	1.000 ± 0.017	0.013 ± 0.007	23	131	10.9	6.13	0.21	0.776 ± 0.403	0.775–0.776	**0.023**	22.3 (16.4-32.3)	**0.037**	0.74
	Humaitá	Madeira	White	22	20	17	0.991 ± 0.016	0.011 ± 0.005	22	110	9.10	5.66	0.10	0.780 ± 0.405	0.746–0.780	**0.077**	11.6 (8.8-15.6)	**0.042**	0.75
	Porto Velho	Madeira	White	21	14	9	0.895 ± 0.061	0.011 ± 0.006	21	103	8.5	5.59	0.15	0.737 ± 0.385	0.674–0.737	**0.143**	39.7 (24.3-86.4)	0.249	0.64
	Manaus	lower Negro	Black	30	29	22	0.997 ± 0.009	0.012 ± 0.006	35	152	12.6	6.38	0.14	0.785 ± 0.404	0.738–0.785	**0.090**	−437.1 (333.5-inf)	**0.002**	0.77
	Tapauá	Purus	White	28	28	27	1.000 ± 0.009	0.014 ± 0.007	28	144	12.0	6.37	0.18	0.766 ± 0.396	0.755–0.766	**0.056**	−168.7 (73.5-inf)	**0.037**	0.73
	Boca do Acre	Purus	White	21	21	19	1.000 ± 0.014	0.013 ± 0.007	21	130	10.8	6.35	0.27	0.771 ± 0.401	0.742–0.771	**0.076**	240.3 (62.6-inf)	0.112	0.73
	Coari	Lake Coari	Black	20	20	15	1.000 ± 0.015	0.012 ± 0.006	29	141	11.7	6.20	0.16	0.762 ± 0.394	0.683–0.762	**0.123**	924.8 (119.3-inf)	**0.002**	0.78
	Tefé	Lake Tefé	Black	18	16	9	0.986 ± 0.022	0.012 ± 0.006	20	117	9.70	6.04	0.16	0.751 ± 0.392	0.733–0.751	**0.058**	652.6 (86.4-inf)	0.439	0.64
	Carauari	Juruá	White	25	25	25	1.000 ± 0.011	0.013 ± 0.007	25	141	11.6	6.53	0.18	0.790 ± 0.409	0.760–0.790	**0.067**	−448.2 (174.2-inf)	0.244	0.71
	Eirunepé	Juruá	White	16	15	13	0.991 ± 0.025	0.010 ± 0.005	20	113	9.40	6.00	0.08	0.797 ± 0.414	0.754–0.797	**0.066**	47.9 (30.1-100.3)	0.248	0.73
	Fonte Boa	Amazon	White	20	18	10	0.989 ± 0.019	0.011 ± 0.005	31	142	11.8	6.18	0.16	0.773 ± 0.399	0.736–0.773	**0.050**	−309.0 (250.3-inf)	**0.036**	0.73
	Tabatinga	Amazon	White	31	30	24	0.997 ± 0.008	0.013 ± 0.007	37	146	12.1	6.15	0.15	0.779 ± 0.401	0.734–0.779	**0.077**	40.0 (77.0-527.4)	0.290	0.76
Bolivian	Guajará-Mirim	Guaporé	White	69	68	64	0.999 ± 0.002	0.012 ± 0.006	38	123	10.2	5.32	0.09	0.714 ± 0.369	0.677–0.714	**0.087**	−329.8 (443.5-inf)	**0.038**	0.75

*N, number of individuals; H, number of haplotypes; S, singletons; h, gene diversity; π, nucleotide diversity; A, number of alleles; A_{NA}, average number of alleles per locus; AR, allelic richness; PAR, private allelic richness; AGD, average genetic diversity; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, endogamy index; Ne [CI], effective population size [95% confidence intervals], p_SMM= Stepwise Mutation Model from Bottleneck program, Mvalue= M-ratio using ARLEQUIN (Garza and Williamson index). Values in bold are significant at 0.05. *Classification of water types based in Sioli (1984) and Venticinque et al. (2016).*

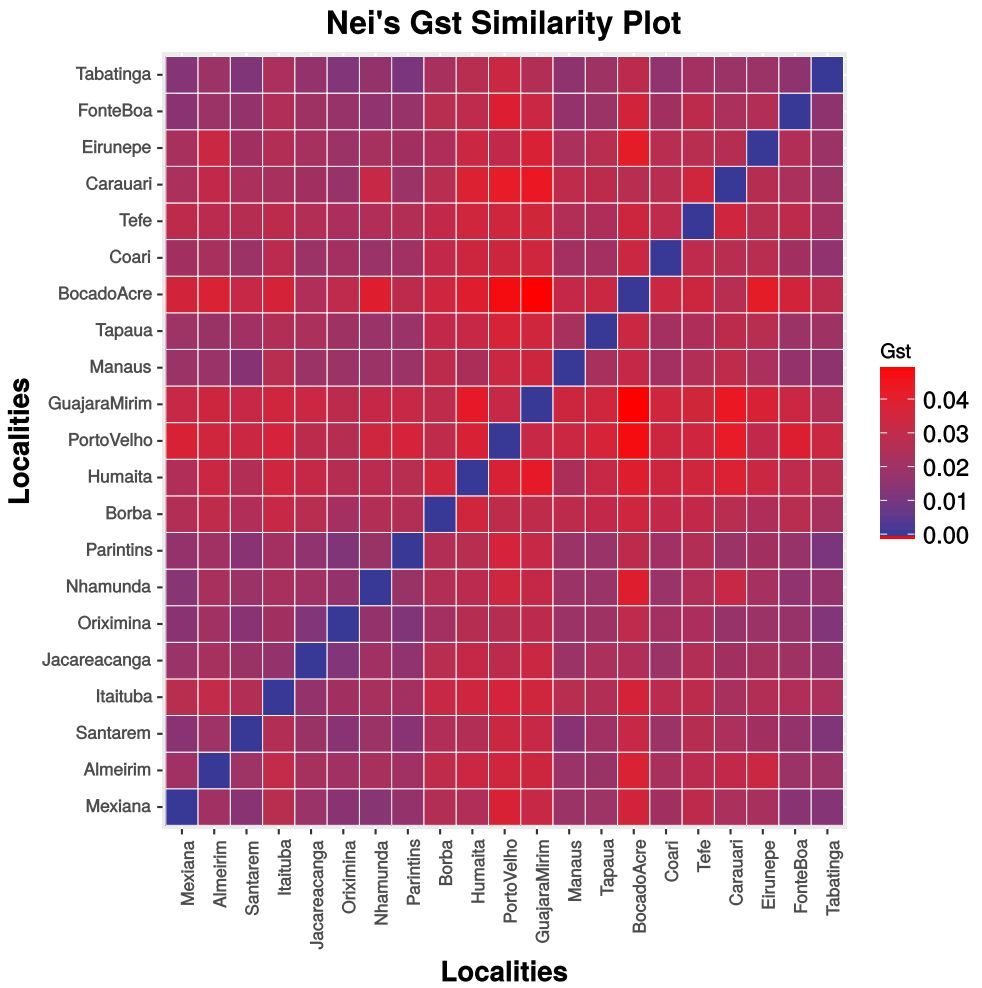


FIGURE 3 | Graphic representation of pairwise Hedrick's G_{ST} values.

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

Gene Flow

Results of the isolation-with-migration analyses using microsatellite data are in **Supplementary Table S6**. Thus, Mexiana and Tabatinga (on the Amazon River) were paired, and a group denominated the main channel was formed by randomly sampling 30 individuals from among the sampling sites of the main Amazon River channel, which was then analyzed with upper-most tributary localities: Jacareacanga (Tapajós River), Guajará-Mirim (Bolivian sub-basin), Boca do Acre (Purus River), and Eirunepé (Juruá River) (**Table 2**). The result indicated bidirectional gene flow between all localities. In all cases, the direction of migration from upstream areas of tributaries to the central Amazon basin predominated except in the case of the Jacareacanga locality.

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

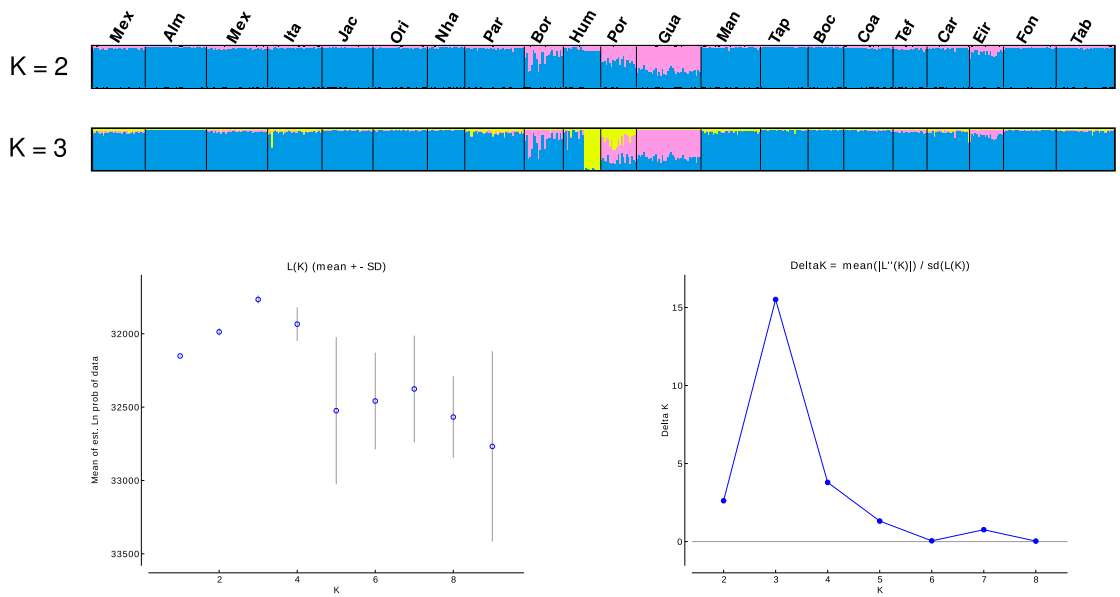


FIGURE 4 | Population structure based on STRUCTURE analysis ($K = 2$, and $K = 3$) of 13 microsatellites. Each vertical bar represents an individual. Estimates of the number of populations (K) based on the mean likelihood Ln (K) and the delta K statistic (Evanno et al., 2005). Locality codes are: Mexiana (Mex), Almeirim (Alm), Santarém (San), Itaituba (Ita), Jacareacanga (Jac), Oriximiná (Ori), Nhamundá (Nha), Parintins (Pin), Borba (Bor), Humaitá (Hum), Porto Velho (Pve), Guajará-Mirim (Gua), Manaus (Mao), Tapauá (Tap), Boca do Acre (Bda), Coari (Coa), Tefé (Tef), Carauari (Car), Eirunepé (Eir), Fonte Boa (Fbo), Tabatinga (Tab).

Population Demography

Using the genetic parameters from IMa2 analyses (**Table 2**), the coalescent effective population size did not differ substantially for almost all pair of localities examined. As a whole, effective population sizes were of thousands of individuals, with the exception of Guajará-Mirim.

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

Population expansion was also supported by Harpending's raggedness index, which was significantly small ($r = 0.0046$, $P = 1.0000$) considering all samples as well as when considering the two basins separately (Amazon basin: $r = 0.00060$, $P = 1.0000$; Bolivian basin: $r = 0.0018$, $P = 0.9990$) (**Table 4**). These indexes statistically support the inference of population expansion based on the observation of the distribution of mismatch distribution (Harpending, 1994) for all samples of *C. macropomum*. However, when mismatch distribution was investigated for the basins separately, unimodal distribution was found only within the Amazon basin, whereas multimodal distribution was found for the Bolivian basin (results not shown), which fits a pattern expected under stable population size, although this was not supported by Harpending's raggedness index. The sum of squared deviations (Schneider and Excoffier, 1999) was non-significant for the overall sample (SSD = 0.0020, $P = 0.6260$) as well as the inference performed for the two basins separately (Amazon basin: SSD = 0.0021, $P = 0.6430$; Bolivian basin: SSD = 0.0047, $P = 0.8420$). Thus, these values neither support nor reject the null hypothesis of a demographic population expansion for *C. macropomum* (**Table 4**). The Tajima's *D* was non-significant for all sampling localities. The same was found for Fu's *Fs*. When considering the basins separately, Fu's *Fs* was significantly negative for the Bolivian basin, suggesting a population expansion.

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

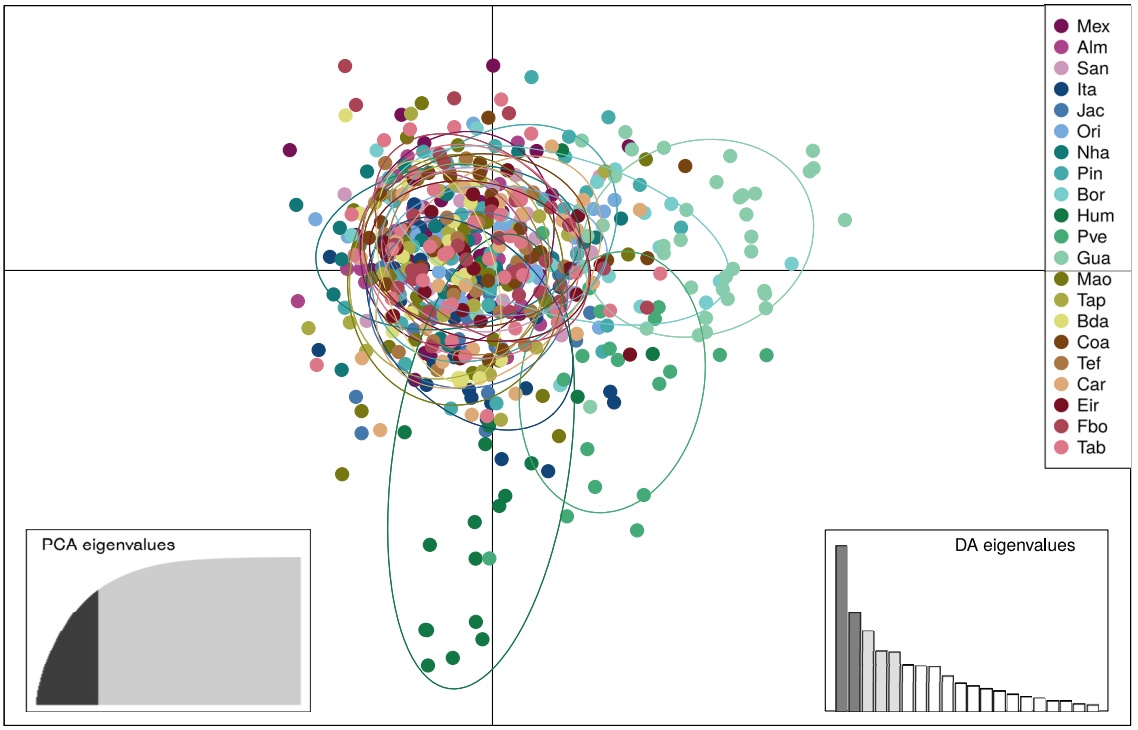


FIGURE 5 | Results of the Discriminant Analysis of Principal Components (DAPC) showing the scatterplot of the first two principal components based on 13 microsatellite loci of 604 individuals of *Colossoma macropomum* from 21 sampling locations. Discriminant function 2 on the x axis and discriminant function 1 on the y axis. In the DAPC graph circles represent different individuals, and colors different sampling localities. Locality codes are: Mexiana (Mex), Almeirim (Alm), Santarém (San), Itaituba (Ita), Jacareacanga (Jac), Oriximiná (Ori), Nhamundá (Nha), Parintins (Pin), Borba (Bor), Humaitá (Hum), Porto Velho (Pve), Guajará-Mirim (Gua), Manaus (Mao), Tapauá (Tap), Boca do Acre (Bda), Coari (Coa), Tefé (Tef), Carauari (Car), Eirunepé (Eir), Fonte Boa (Fbo), Tabatinga (Tab).

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

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

Effective population size Ne = 4 Neμ/4 μx3 (Ne_1 of the locality1, Ne_2 of the locality 2 and Ne_A of the ancestor of the two localities); current time in years T = (t x μ) × 3; number of migrants per generation Nm = Mθ/2. Values in parentheses = minimum and maximum of confidence limits. Three year generation time is assumed.

DISCUSSION

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

The use of more variable genetic markers, such as microsatellites, has confirmed some of our earlier findings. *Colossoma macropomum* is not panmictic throughout its distribution area. Considering the entire sample, AMOVA, SAMOVA, STRUCTURE, Φ_{ST} (DNAmt), and F_{ST}/G_{ST} (microsatellites) analyses, suggested genetic differentiation of the Bolivian sub-basin and Brazilian Amazon basin localities.

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

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

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

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

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

The genetic differentiation evidenced between the two basins (Brazilian and Bolivian) is associated with the upper Madeira River rapids, which serve as a natural barrier that restricts, but does not prevent, geneflow between the populations of *C. macropomum* of the two basins. The origin of the Bolivian sub-basin is related to the elevation of the Fitzcarrald arch at the beginning of the middle Pliocene (4 to 3 Ma), which gradually isolated the Bolivian basin, resulting in considerable changes in the drainage pattern, in which the main rivers north of Bolivia drain into the Amazon River through the Madeira River (Hoorn et al., 1995; Campbell et al., 2001), which is the largest tributary of the southern margin of the Solimões-Amazonas basin (Lundberg et al., 1998). The Bolivian sub-basin includes the main Beni and Mamoré rivers, as well as approximately 60% of the entire drainage area of the Madeira River (upper part of Madeira) and is separated from the Amazon basin by a set of 18 rapids and cataracts located between Guajará-Mirim and Porto Velho (Cella-Ribeiro et al., 2013). The largest of these cataracts, the Teotônio cataract, constituted the greatest barrier to navigation on this river, as well as to the movement of many species of fish (Goulding et al., 2003). The rapids of the upper Madeira River play an important role in the structuring of populations of other Amazonian aquatic species, such as the river turtle *Podocnemis expansa* (Pearse et al., 2006), river dolphins (Gravena et al., 2014, 2015), the black Amazonian flanelmouth characin *Prochilodus nigricans* (Machado et al., 2017), and the catfish *Brachyplatystoma rousseauxii* (Batista, 2010; Carvajal-Vallejos et al., 2014). The Teotônio cataract, as well as the Jirau cataract, the second largest of the Madeira River, have been submerged by hydroelectric reservoirs.

Population Genetic Structure in the Amazonas River

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

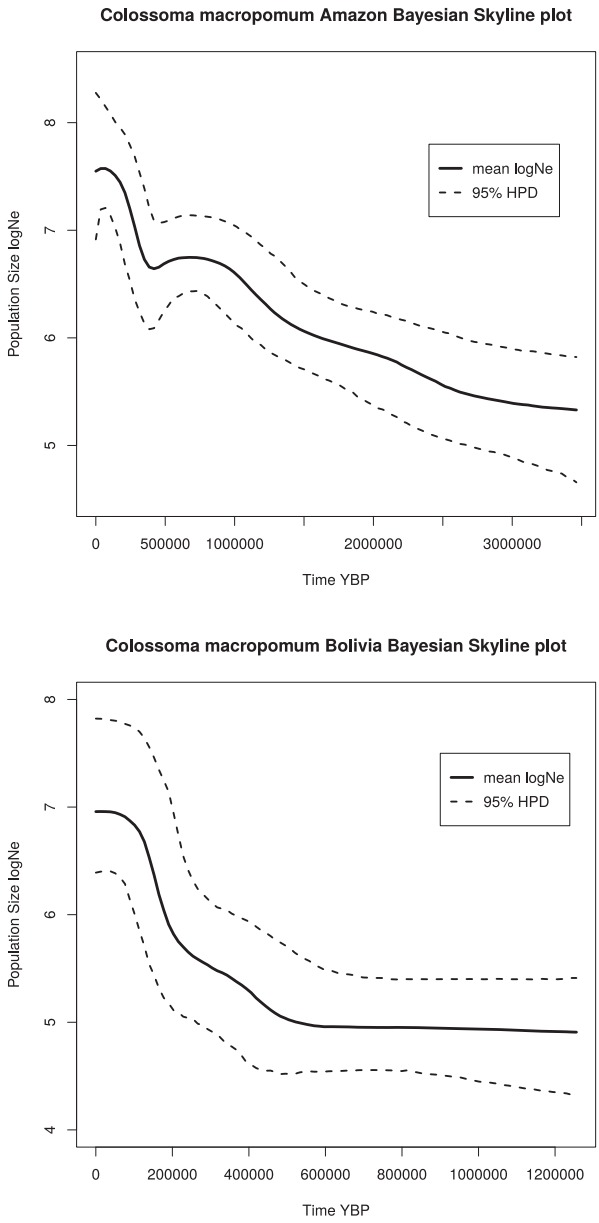


FIGURE 6 | Bayesian skyline plots for Amazon and Bolivian basins.

and tributary headwaters and black water sites generally is smaller than one effective migrant per generation which also confirms the migration patterns observed in IMa and MIGRATE analyses. Contrary to this pattern, STRUCTURE analysis shows population differentiation only for upper Madeira (localities below the rapids). STRUCTURE program uses a Bayesian approach to investigate the number of biological groups in the dataset. The discordance, between these analyses, could be due to the fact that G_{ST} analyses are based on variance in allelic frequencies between/among groups and Fisher's exact test uses contingency tables to test null hypothesis that the alleles are drawn from the same distribution in all populations. These two analyses are more sensitive to detecting smaller and finer levels of genetic differentiation, while STRUCTURE tests for shared system of mating among individuals of the same group. The algorithm of STRUCTURE will not necessarily detect weak structure (Evanno et al., 2005). The population structure observed in *C. macropomum* falls within the category of weak to moderate level of population differentiation (according to Wright, 1965), which could be limiting the sensitivity of the analysis to find more refined substructuring.

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

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

Genetic Diversity: Large or Small?

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

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

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

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

from 0.71 to 0.79 with an average of 0.78. For all microsatellite data, variability measures, such as expected heterozygosity and allelic diversity, were similar to those observed in other exploited migratory Amazonian fishes (Carvajal-Vallejos et al., 2014; Ochoa et al., 2015; Oliveira et al., 2018). A compilation of average H_E values for microsatelite loci obtained from the literature for exploited Amazonian migratory fish shows average H_E values ranging from 0.50 of *Brachyplatystoma platynemum* (Ochoa et al., 2015), to 0.87 of *Semaprochilodus insignis* (Passos et al., 2010). Between this minimum and maximum, one can observe $H_E = 0.61$ of *Brachyplatystoma rousseauxii* (Batista, 2010), $H_E = 0.75$ of *Brachyplatystoma vaillantii* (Rodrigues et al., 2009), and $H_E = 0.83$ of *Brycon amazonicus* (Oliveira et al., 2018). An average of these values (mean $H_E = 0.72$) is not in agreement with DeWoody and Avise (2000), who report an average $H_E = 0.54$ for freshwater fishes. The $H_E = 0.78$ of *C. macropomum* and most of the H_E reported for exploited Amazonian migratory fishes are more similar to the H_E of marine fishes reported by DeWoody and Avise (2000) with a mean $H_E = 0.77$. In the review of DeWoody and Avise (2000), the authors summarized microsatelite data from North American and European freshwater species, that in comparison to the Amazon basin are geographically very restricted. At 5.5 million km^2, the Amazon basin is by far the largest hydrographic basin on the planet, and the size of area occupied by many fish species is comparable to that for marine fishes. The Amazon is in a sense a "sea" that provides an expansive and effectively continuous environment for migratory freshwater fishes such as *C. macropomum* and the other cited species. Migratory fishes, whether freshwater or marine, are usually *r* strategists, that is, they have high dispersal capacity, produce a lot of offspring, and in general have a large effective population size, which is turn is reflected in high heterozigosity levels.

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

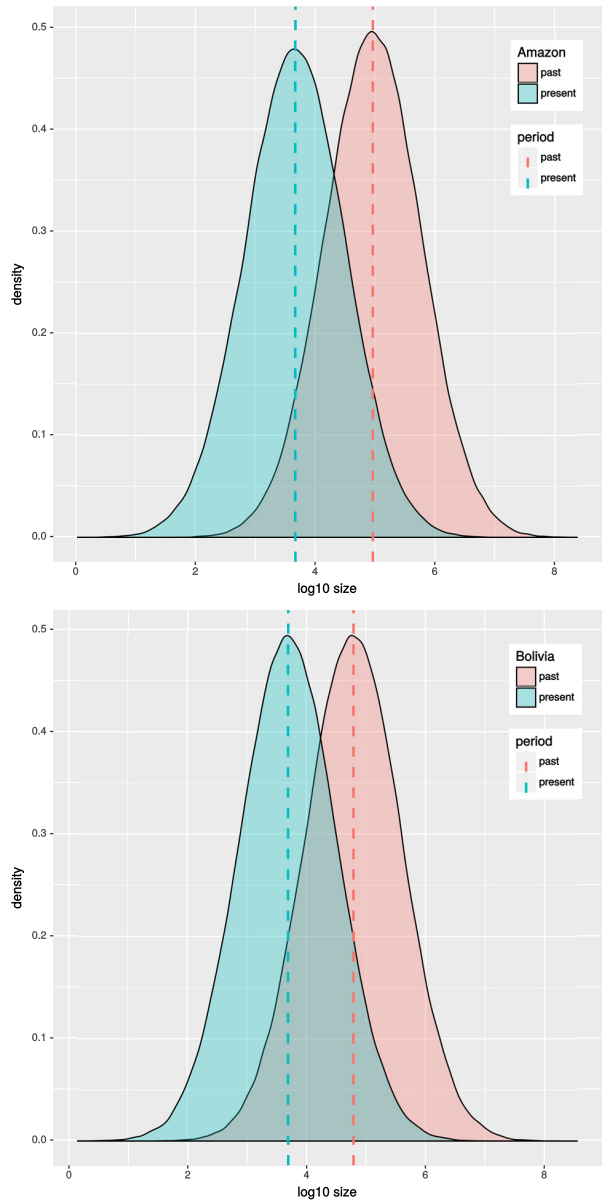


FIGURE 7 | MSVar results for Amazon and Bolivian basins.

Population Genetic Demography

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

Most studies conducted in the Amazon involving fish species such as *Prochilodus nigricans* (Machado et al., 2017), *Brachyplatystoma rousseauxii* (Batista and Alves-Gomes, 2006; Carvajal-Vallejos et al., 2014), and *Brachyplatystoma platynemum* (Ochoa et al., 2015) indicate recent population expansion, at least as indicated by Fu's *Fs* test. The difference in effective population size of *C. macropomum* prior estimated for the Bolivian and Amazon basin broadly corresponded to the relative proportion of potential habitat in the Bolivian basin. This basin accounts for approximately 20% of the total Amazon basin, suggesting that the Bolivian basin had a *C. macropomum* population approximately 20% smaller than the rest of the Amazon basin during the Holocene.

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

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

The demographic decrease was from approximately 100,000 effective individuals to approximately 5,000 individuals, an approximate 1.5 orders of magnitude decrease. Similar values for current effective population size were inferred using IMa2 (**Table 2**). By any measure, the effective population size is small, and much smaller than in the last several thousand years.

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

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

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

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

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

FUNDING

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

ACKNOWLEDGMENTS

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

SUPPLEMENTARY MATERIAL

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

REFERENCES

Araújo-Lima, C. A. R. M. (2002). "Piscicultura extensiva de tambaqui na floresta de várzea," in *Livro de Resultados dos Projetos de Pesquisa Dirigida (PPDS)-PPG7* (Brasília: MCT), 131–135.

Araújo-Lima, C. A. R. M., and Goulding, M. (1998). *Os Frutos do Tambaqui. Ecologia, Conservação e Cultivo na Amazônia*. Belem: Sociedade Civil Mamirauá-MCT-CNPq.

Araújo-Lima, C. A. R. M., and Ruffino, M. L. (2004). "Migratory fishes of the Brazilian Amazon," in *Migratory Fishes of South America. Biology, Fisheries, and Conservation Status*, eds J. Carolsfield, B. Harvey, C. Ross, and A. Baer (Victoria, BC: Co-published by World Fisheries Trust/World Bank/International Development Research Center), 233–302.

Barthem, R. B., and Fabré, N. N. (2003). "Biologia e diversidade dos recursos pesqueiros da Amazônia," in *A Pesca e os Recursos Pesqueiros na Amazônia Brasileira*, ed M. L. Ruffino (Manaus, AM: ProVarzea), 11–55.

Barthem, R. B., and Goulding, M. (2007). *An Unexpected Ecosystem: The Amazon as Revealed by Fisheries*. St. Louis, MO: Amazon Conservation Association and Missouri Botanical Garden Press.

Batista, J. S. (2010). *Caracterização Genética da Dourada-Brachyplatystoma Rousseauxii, Castelnau, 1855 (Siluriformes - Pimelodidae) na Amazônia por Meio de Marcadores Moleculares Mitocondriais e Microssatélites: Subsídios Para Conservação e Manejo*. Tese de Mestrado do Programa de Pós-graduação de Genética, Conservação e Biologia Evolutiva, Instituto Nacional de Pesquisas da Amazônia, Manaus, Brasil.

Batista, J. S., and Alves-Gomes, J. A. (2006). Phylogeography of *Brachyplatystoma rousseauxii* (Siluriformes - Pimelodidae) in the Amazon Basin offers preliminary evidence for the first case of "homing" for an Amazonian migratory catfish. *Gen. Mol. Res.* 5, 723–740.

Beaumont, M. A. (1999). Detecting population expansion and decline using microsatellites. *Genetics* 153, 2013–2029.

Beerli, P., and Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4563–4568. doi: 10.1073/pnas.081068098

Campbell, K. E. Jr., Heizler, M., Frailey, C. D., Romero Pittman, L., and Prothero, D. R. (2001). Upper Cenozoic chronostratigraphy of the southwestern Amazon Basin. *Geology* 29, 595–598. doi: 10.1130/0091-7613(2001)029<0595:UCCOTS>2.0.CO;2

Carvajal-Vallejos, F. M., Duponchelle, F., Desmarais, E., Cerqueira, F., Querouil, S., Nuñez, J., et al. (2014). Genetic structure in the Amazonian catfish *Brachyplatystoma rousseauxii*: influence of life history strategies. *Genetica* 142, 323–336. doi: 10.1007/s10709-014-9777-2

Cella-Ribeiro, A., Torrente-Vilara, G., Hungria, D. B., and Oliveira, M., de (2013). "As corredeiras do Rio Madeira," in *Peixes do Rio Madeira*, eds L. J. de Queiroz, G. Torrente-Vilara, W. M. Ohara, T. H. da S. Pires, J. Zuanon, and C. R. C. Doria (Santo Antonio Energia), 47–53.

Corander, J., and Tang, J. (2007). Bayesian analysis of population structure based on linked molecular information. *Math. Biosci.* 205, 19–31. doi: 10.1016/j.mbs.2006.09.015

DeWoody, J. A., and Avise, J. C. (2000). Microsatellite variation in marine, freshwater and anadromus fishes compared with other animals. *J. Fish Biol.* 56, 461–473. doi: 10.1006/jfbi.1999.1210

DeWoody, J. A., Schupp, J., Kenefic, L., Busch, J., Murfitt, L., and Keim, P. (2004). Universal method for producing ROX-labeled size standards suitable for automated genotyping. *BioTechniques* 37, 348–352. doi: 10.2144/04373BM02

Di Rienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M., and Freimer, N. B. (1994). Mutational processes of simple-sequence repeat

loci in human populations. *Proc. Natl. Acad. Sci. U.S.A.* 91, 3166–3170. doi: 10.1073/pnas.91.8.3166

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., and Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Mol. Ecol. Res.* 14, 209–214. doi: 10.1111/1755-0998.12157

Drummond, A. J., and Rambaut, A. (2007). BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214. doi: 10.1186/1471-2148-7-214

Dupanloup, I., Schneider, S., and Excoffier, L. (2002). A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* 11, 2571–2581. doi: 10.1046/j.1365-294X.2002.01650.x

Estoup, A., Tailliez, C., Cornuet, J.-M., and Solignac, M. (1995). Size homoplasy and mutational processes of interrupted microsatellite in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Mol. Biol. Evol.* 12, 1074–1084.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Excoffier, L., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.

Farias, I. P., Torrico, J. P., García-Dávila, C., da Santos, M. C. F., Hrbek, T., and Renno, J.-F. (2010). Are rapids a barrier for floodplain fishes of the Amazon basin? A demographic study of the keystone floodplain species *Colossoma macropomum* (Teleostei: Characiformes). *Mol. Phylogenet. Evol.* 56, 1129–1135. doi: 10.1016/j.ympev.2010.03.028

Frankham, R., Ballou, J. D., and Briscoe, D. A. (2002). *Introduction to Conservation Genetics.* Cambridge: Cambridge University Press.

Frederico, R. G., Farias, I. P., Araújo, M. L. G., Charvet-Almeida, P., and Alves-Gomes, J. A. (2012). Phylogeography and conservation genetics of the Amazonian freshwater stingray *Paratrygon aiereba* Müller & Henle, 1841 (Chondrichthyes: Potamotrygonidae). *Neotrop. Ichthyol.* 10, 71–80. doi: 10.1590/S1679-62252012000100007

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.

Garza, J. C., and Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. *Mol. Ecol.* 10, 305–318. doi: 10.1046/j.1365-294X.2001.01190.x

Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–472. doi: 10.1214/ss/1177011136

Goulding, M., Barthem, R. B., and Ferreira, E. J. G. (2003). *The Smithsonian Atlas of the Amazon*. Washington, DC: Smithsonian Institution Press.

Goulding, M., and Carvalho, M. L. (1982). Life history and management of the tambaqui (*Colossoma macropomum*, Characidae): an important amazonian food fish. *Rev. Bras. Zool.* 1, 107–133. doi: 10.1590/S0101-81751982000200001

Gravena, W., da Silva, V. M. F., da Silva, M. N. F., Farias, I. P., and Hrbek, T. (2015). Living between rapids: genetic structure and hybridization in botos (Cetacea: Iniidae: Inia spp.) of the Madeira River, Brazil. *Biol. J. Linn. Soc.* 114, 764–777. doi: 10.1111/bij.12463

Gravena, W., Farias, I. P., da Silva, M. N. F., da Silva, V. M. F., and Hrbek, T. (2014). Looking to the past and the future: were the Madeira River rapids a geographic barrier to the boto (Cetacea: Iniidae)? *Conserv. Genet.* 15, 619–629. doi: 10.1007/s10592-014-0565-4

Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98. Available online at: http://www.mbio.ncsu.edu/BioEdit/bioedit.html

Harpending, H. C. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum. Biol.* 66, 591–600.

Hasegawa, M., Kishino, H., and Yano, T. A. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160–174. doi: 10.1007/BF02101694

Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution* 59, 1633–1638. doi: 10.1111/j.0014-3820.2005.tb01814.x

Hey, J., and Nielsen, R. (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2785–2790. doi: 10.1073/pnas.0611164104

Hoorn, C., Guerrero, J., Sarmiento, G. A., and Lorente, M. A. (1995). Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. *Geology* 23, 237–240. doi: 10.1130/0091-7613(1995)023<0237:ATAACF>2.3.CO;2

Hoorn, C., Wesselingh, F. P., ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., et al. (2010). Change, landscape evolution, and biodiversity Amazonia through time: Andean uplift, climate. *Science* 330, 927–931. doi: 10.1126/science.1194585

Irion, G., and Kalliola, R. J. (2010). "Long-term landscape development processes in Amazonia," in *Amazonia: Landscape and Species Evolution: A Look into the Past*, eds C. Hoorn and F. P. Wesselingh (Chichester: Wiley-Blackwell), 185–197.

Isaac, V. J., Milstein, A., Ruffino, M. L. (1996). A pesca artesanal no baixo Amazonas: análise multivariada da captura por espécies. *Acta Amazon.* 26, 185–208. doi: 10.1590/1809-43921996263208

Isaac, V. J., and Ruffino, M. L. (1996). Population dynamics of tambaqui, *Colossoma macropomum* Cuvier, in the lower Amazon, Brazil. *Fish. Manag. Ecol.* 1996, 315–333. doi: 10.1046/j.1365-2400.1996.d01-154.x

Isaac, V. J., and Ruffino, M. L. (2000). "Biologia pesqueira do tambaqui, Colossoma macropomum, no Baixo Amazonas," in *Recursos Pesqueiros do Médio Amazonas: Biologia e Estatística Pesqueira* (Brasília: Edições IBAMA; Coleção meio ambiente; Série Estudos Pesca), 65–88.

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11:94. doi: 10.1186/1471-2156-11-94

Junk, W. J. (1997). *The Central Amazon System: Ecology of a Pulsing System*. Berlin: Springer Verlag.

Kalinowski, S. T. (2005). Hp-Rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol. Ecol. Notes* 5, 187–189. doi: 10.1111/j.1471-8286.2004.00845.x

Kitada, S., Kitakado, T., and Kishino, H. (2007). Empirical bayes inference of pairwise F_{ST} and its distribution in the genome. *Genetics* 177, 861–873. doi: 10.1534/genetics.107.077263

Kitada, S., Nakamichi, R., and Kishino, H. (2017). The empirical Bayes estimators of fine-scale population structure in high gene flow species. *Mol. Ecol. Resour.* 17, 1210–1222. doi: 10.1111/1755-0998.12663

Leberg, P. L. (2002). Estimating allelic richness: effects of sample size and bottlenecks. *Mol. Ecol.* 11, 2445–2449. doi: 10.1046/j.1365-294X.2002.01612.x

Ledru, M.-P., Braga, P. I. S., Soubiès, F., Fournier, M., Martin, L., Suguio, K., et al. (1996). The last 50,000 years in the Neotropics (Southern Brazil): evolution of vegetation and climate. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 123, 239–257. Available online at: http://www.sciencedirect.com/science/article/B6V6R-3VVT2P4-D/2/b91b6aa40adb8b4b114c150b06404376. doi: 10.1016/0031-0182(96)00105-8

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187

Lovejoy, N. R., Willis, S. C., and Albert, J. S. (2010). "Molecular signatures of Neogene biogeographical events in the Amazon fish fauna," in *Amazonia: Landscape and Species Evolution: A Look into the Past*, eds C. Hoorn and F. P. Wesselingh (Oxford: Wiley-Blackwell), 405–417.

Lundberg, J. G., Marshall, L. G., Guerrero, J., Horton, B., Malabarba, M. C. S. L., and Wesselingh, F. P. (1998). "The stage for Neotropical fish diversification: a history of tropical South American rivers," in *Phylogeny and Classification of Neotropical Fishes*, eds L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M. S. Lucena, and C. A. S. Lucena (Porto Alegre: EDIPUCRS), 13–48.

Machado, V. N., Willis, S. C., Hrbek, T., and Farias, I. P. (2017). Population genetic structure of the Amazonian black flannelmouth characin (Characiformes, Prochilodontidae: Prochilodus nigricans Spix and Agassiz, 1829): contemporary and historical gene flow of a migratory and abundant fishery species. *Environ. Biol. Fishes* 100, 1–16. doi: 10.1007/s10641-016-0547-0

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27, 209–220.

Ochoa, L. E., Pereira, L. H. G., Costa-Silva, G. J., Roxo, F. F., da Batista, J. S., Formiga, K., et al. (2015). Genetic structure and historical diversification

of catfish Brachyplatystoma platynemum (Siluriformes: Pimelodidae) in the Amazon basin with implications for its conservation. *Ecol. Evol.* 5, 2005–2020. doi: 10.1002/ece3.1486

Ohta, T., and Kimura, M. (1973). A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Gen. Res.* 22, 201–204. doi: 10.1017/S0016672300012994

Oliveira, R. C., Santos, M. C. F., Bernardino, G., Hrbek, T., and Farias, I. P. (2018). From river to farm: an evaluation of genetic diversity in wild and aquaculture stocks of Brycon amazonicus (Spix & Agassiz, 1829), Characidae, Bryconinae. *Hydrobiologia* 805, 75–88. doi: 10.1007/s10750-017-3278-0

Palstra, F. P., and Fraser, D. J. (2012). Effective/census population size ratio estimation: a compendium and appraisal. *Ecol. Evol.* 2, 2357–2365. doi: 10.1002/ece3.329

Passos, K. B., Leão, A. S. A., Oliveira, D. P., Farias, I. P., and Hrbek, T. (2010). Polymorphic microsatellite markers for the overexploited Amazonian fish, Semaprochilodus insignis (Jardine and Schomburgk 1841). *Conserv. Gen. Res.* 2, 231–234. doi: 10.1007/s12686-010-9245-y

Pearse, D. E., Arndt, A. D., Valenzuela, N., Miller, B. A., Cantarelli, V. H., and Sites, J. W. (2006). Estimating population structure under nonequilibrium conditions in a conservation context: continent-wide population genetics of the giant Amazon River turtle, *Podocnemis expansa* (Chelonia; Podocnemididae). *Mol. Ecol.* 15, 985–1006. doi: 10.1111/j.1365-294X.2006.02869.x

Piry, S., Luikart, G.,, and Cornuet, J. (1999). BOTTLENECK: a computer program for detecting recent reduction in the effective size using allele frequency data. *J. Hered.* 90, 502–503. doi: 10.1093/jhered/90.4.502

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Putman, A. I., and Carbone, I. (2014). Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecol. Evol.* 4, 4399–4428. doi: 10.1002/ece3.1305

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing Available online at: www.R-project.org

Reis, R. E., Albert, J. S., Di Dario, F., Mincarone, M. M., Petry, P., and Rocha, L. A. (2016). Fish biodiversity and conservation in South America. *J. Fish Biol.* 89, 12–47. doi: 10.1111/jfb.13016

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* 43, 223–225.

Rodrigues, F. C., Farias, I. P., Batista, J. S., and Alves-Gomes, J. (2009). Isolation and characterization of microsatellites loci for "piramutaba" (*Brachyplatystoma vaillantii*, Siluriformes: Pimelodidae), one of the commercially most important migratory catfishes in the Amazon Basin. *Conserv. Genet. Res.* 1:365. doi: 10.1007/s12686-009-9084-x

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Springs Harbor, NY: Cold Springs Harbor Laboratory Press.

Santos, M. C. F., Ruffino, M. L., and Farias, I. P. (2007). High levels of genetic variability and panmixia of the tambaqui *Colossoma macropomum* (Cuvier, 1818) in the main channel of the Amazon River. *J. Fish Biol.* 71A, 33–44. doi: 10.1111/j.1095-8649.2007.01514.x

Santos, M. D., Hrbek, T., and Farias, I. P. (2009). Microsatellite markers for the tambaqui (*Colossoma macropomum*, *Serrasalmidae*, *Characiformes*), an economically important keystone species of the Amazon River floodplain. *Mol. Ecol. Resour.* 9, 874–876. doi: 10.1111/j.1755-0998.2008.02331.x

Schneider, S., and Excoffier, L. (1999). Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 152, 1079–1089. Available online at: http://www.genetics.org/cgi/content/abstract/152/3/1079

Sioli, H. (1984). "The Amazon and its main affluents: hydrography, morphology of the river courses and river types," in *The Amazon, Limnology and Landscape Ecology of a Mighty Tropical River and its Basin,* ed H. Sioli (New York, NY: Springer Verlag), 127–165.

Slatkin, M., and Hudson, R. R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129, 555–562.

Storz, J. F., Beaumont, M. A., and Alberts, S. C. (2002). Genetic evidence for long-term population decline in a savannah-dwelling primate: inferences from a hierarchical bayesian model. *Mol. Biol. Evol.* 19, 1981–1990. doi: 10.1093/oxfordjournals.molbev.a004022

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.

Tamura, K., Stecher, G., Peterson, D. G., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4, 535–538. doi: 10.1111/j.1471-8286.2004.00684.x

Venticinque, E., Forsberg, B., Barthen, B. R., Petry, P., Hess, L., Mercado, A., et al. (2016). An explicit GIS-based river basin framework for aquatic ecosystem conservation in the Amazon. *Earth Syst. Sci. Data* 8, 651–661. doi: 10.5194/essd-2016-17

Wang, X., Lawrence Edwards, R., Auler, A. S., Cheng, H., Kong, K., Wang, Y., et al. (2017). Hydroclimate changes across the Amazon lowlands over the past 45,000 years. *Nature* 541, 204–207. doi: 10.1038/nature20787

Waples, R. S., and Do, C. (2008). Ldne: a program for estimating effective population size from data on linkage disequilibrium. *Mol. Ecol. Res.* 8, 753–756. doi: 10.1111/j.1755-0998.2007.02061.x

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19, 395–420.

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

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

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership