frontiers | Research Topics

Application of fishes as biological models in genetic studies

Edited by
Tony Silveira, Sandra Isabel Moreno Abril, Mariana Härter Remião and Caroline Gomes Lucas

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-83251-366-8
DOI 10.3389/978-2-83251-366-8

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: frontiersin.org/about/contact

Application of fishes as biological models in genetic studies

Topic editors

Tony Silveira — Federal University of Rio Grande, Brazil
Sandra Isabel Moreno Abril — University of Vigo, Spain
Mariana Härter Remião — Federal University of Pelotas, Brazil
Caroline Gomes Lucas — University of Missouri, United States

Citation

Silveira, T., Abril, S. I. M., Remião, M. H., Lucas, C. G., eds. (2023). *Application of fishes as biological models in genetic studies*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83251-366-8

Table of contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 05 January 2023
DOI 10.3389/fgene.2022.1092160

OPEN ACCESS

EDITED BY
Jitendra Kumar Sundaray,
Central Institute of Freshwater Aquaculture, India

REVIEWED BY
Irfan Ahmad Bhat,
University of Iceland, Iceland

*CORRESPONDENCE
Tony Silveira,
silveira.tlr@gmail.com
Mariana Härter Remião,
mariana.remiao@ufpel.edu.br

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

RECEIVED 07 November 2022
ACCEPTED 28 November 2022
PUBLISHED 05 January 2023

CITATION
Silveira T, Moreno Abril SI, Lucas CG and Remião MH (2023), Editorial: Application of fishes as biological models in genetic studies.
Front. Genet. 13:1092160.
doi: 10.3389/fgene.2022.1092160

Editorial: Application of fishes as biological models in genetic studies

Tony Silveira[1]*, Sandra Isabel Moreno Abril[2,3], Caroline Gomes Lucas[4] and Mariana Härter Remião[5]*

[1]Biological Sciences Institute, Federal University of Rio Grande, Rio Grande, Brazil, [2]Marine Research Centre, University of Vigo, Vigo, Spain, [3]Department of Ecology and Animal Biology, University of Vigo, Vigo, Spain, [4]Division of Animal Sciences, Animal Science Research Center, National Swine Resource and Research Center, University of Missouri, Columbia, MO, United States, [5]Structural Genomics Laboratory, Technological Developmental Center, Federal University of Pelotas, Pelotas, Brazil

KEYWORDS

evolution, biotechnology, genomics, aquaculture, chromosomal diversity, microsatellites, microRNA, transcriptome

Editorial on the Research Topic
Application of fishes as biological models in genetic studies

Introduction

Although Gregor Mendel is considered the father of genetics, while alive, he never took credit for his principles on heredity. Mendel's treatises have been part of the collections of the largest European libraries since the 19th century but were apparently forgotten by science until their rediscovery in 1900 (Watson et al., 2013). Mendel's revolutionary ideas could have possibly given greater strength to the formulation of Charles Darwin's ideas regarding common descent and gradual evolution through natural selection, as presented in "The Origin of Species". Nonetheless, Darwin envisioned the possibility of genetic heredity by describing "invisible characters" emerging in atavistic situations and naming his hypothetical particles of heredity as "gemmules" (Darwin, 1868). It is remarkable that the "invisible characters" and "gemmules" referred to by Darwin are what we now know as genes—a term coined in 1909 by Wilhelm Johannsen.

During the 1930s and 40's, the findings of great proponents of genetics and evolution such as Mendel, Darwin, Wallace, Fisher, Haldane, Wright, Dobzhansky, Mayr, and others were brought together to form the neo-Darwinian synthesis (Mayr, 2004). In addition, during the 40's, a molecular revolution began in genetics, which, driven by sequencing technology, gave rise to the late 70's genomics era. It took approximately 100 years to formulate the theoretical foundations of genetics in order to understand how information is transmitted from generation to generation. Now, less than 45 years after

the beginning of the genomic era, it is possible to identify complete genomes in less than a week, as evidenced by the COVID-19 pandemic (Virological.org, 2020 [1]).

The comparative use of *in vivo* models such as mice, rats, zebrafish, and fruit flies have been fundamental for the advancement of medicine and biological research. Among animals, fishes fall between the most relevant groups in genetic studies, due to their diversity and plasticity. Although fish studies played a role in applying and corroborating Mendel's findings in the first few decades of the 19th century, these studies contributed little to the development of classical genetics. However, fish studies have been of great importance for the development of modern molecular genetics and other "omics" sciences.

Due to this increase in genomic knowledge, several fish species have become important biological models around the world. For example, zebrafish have been used as avatars of human patients in personalized medicine (Usai et al., 2020), tilapia islets have been studied in the treatment of diabetes (Wright et al., 2014), and tilapia skin has been studied for the healing of burn lesions (Costa et al., 2019; Dias et al., 2020). Moreover, fish can be used for the study of genes and genomes, epigenetics, and genetic expression, as presented in the articles of this Research Topic. The increase in knowledge and the importance of fishes in scientific research would certainly not be possible without studies like these. Together, they provide a glimpse of the *Application of Fishes as Biological Models in Genetic Studies* around the globe.

The broad diversity of the studies submitted in this Research Topic is organized into sub-topics that are listed below, but in fact, many of the articles fall under more than one sub-topic. This highlights the nowadays multifaceted character of the application of fishes in genetic studies.

Fishes as genetic models

The *Austrolebias* genus, a group of annual killifish, has been proposed by García et al. as an excellent model for evolutionary genomic processes due to its large genome size—probably associated with transposable elements—as well as its high chromosome instability, the occurrence of natural hybridization between sister species, and burst speciation events.

Another suggestion of fish as an animal model is *Astyanax altiparanae*, as reviewed by Yasui et al. This species, popularly known as yellowtrail tetra, can be used as an alternative for studies directed to neotropical fishes. *A. altiparanae* has been considered as the most advanced fish within the Neotropical region regarding fish biotechnology. It has already been used to describe artificial fertilization, germ cell transplantation, chromosome set manipulation, and other technologies with applications in aquaculture and conservation of genetic resources.

Beck et al. suggest that the threespine stickleback (*Gasterosteus aculeatus*) can be a model for human diseases related to mitochondrial DNA. In this study, the first complete coding region analysis of the two mitotypes of *G. aculeatus* is provided, from which, mitogenomic divergence can be extended to other mammal models including humans. Dohi and Matsui reinforce, in a brief research report, that small fish can be used for prospective studies on human aging, since genes involved in human pathways and diseases are shared with various fish species such as zebrafish, medaka, and the turquoise killifish (*Nothobranchius furzeri*).

Fishes as models in evolutionary studies

Genetic and molecular studies are also performed on fish species to clarify evolutionary aspects of the species. Fernandes et al. have applied conventional cytogenetic procedures and fluorescence *in situ* hybridization of 18S rDNA, 5S rDNA, H3, and H2B-H2A histone sequences in *Acanthurus* species. The authors describe the relations and evolutionary differentiation among *A. coeruleus*, *A. chirurgus*, and *A. bahianus* populations inhabiting coastal regions of the Southwest Atlantic, South Atlantic oceanic islands, Greater Caribbean, and Indo-Pacific Ocean.

In a similar manner, Silva et al. A) have performed a comparative study on two *Peckoltia* species and found that both species present 2n = 52 karyotypic formulas, although their chromosomal bands and the repetitive chromosomal sites between them were different. The same research group published another study on this topic, but with a comparison of two *Ancistrus* species from the Amazon region. Silva et al. B) found extensive chromosomal diversity between both analyzed species, presenting different karyotypes with distinct patterns of organization.

Through mitochondrial DNA sequencing and nuclear DNA microsatellites, Astorga et al. present the genetic diversity and divergence between landlocked and migratory populations of *Galaxias maculatus*, distributed in southern Chile. Additionally, the genetic diversity was significantly higher in migratory populations than in landlocked populations, presenting a higher differentiation among lakes than estuaries.

Using cytogenetic studies on *Gymnotus carapo*, Machado et al. hypothesize that it is not a single widespread species, but a group of cryptic species. Chromosome painting showed more complex rearrangements, a high number of repetitive DNA sites, and extensive karyotype reorganization in comparison with previous studies that used classical cytogenetics.

1 https://virological.org/t/first-cases-of-coronavirus-disease-covid-19-in-brazil-south-america-2-genomes-3rd-march-2020/409.

Fishes as models in biotechnology

In-depth knowledge of the fish genome can provide new alternatives to overcome old barriers. Merlano et al. present the production of recombinant cystatin-B (rCYST-B) from red piranha *Pigcentrus nattereri* to control bacterial growth in aquaculture. This protein exhibited bacteriostatic action, inhibiting *Escherichia coli* and *Bacillus subtilis* growth, suggesting its potential biotechnological use.

Molecular biology for aquaculture improvement

For Nile tilapia, two studies in this Research Topic have brought new clarifications in the production and aquaculture of this species, using a molecular approach. Martins et al. evaluate the expressions of genes responsible for appetite regulation, metabolic and physiological changes, and osmoregulation in tilapia that were exposed to different salinity concentrations. Gene modulation generated by a salinity increase may have contributed to a decrease in weight gain and growth rate, as well as an increase in oxidative damage to blood cells.

Chen et al., also regarding Nile tilapia, demonstrate that individuals with higher growth rates have significantly higher total protein, total triglyceride, total cholesterol, and high- and low-density lipoproteins, but significantly lower glucose levels when compared with individuals in the lower growth rate group. Transcriptomics has also been evaluated, showing that more than 1,000 genes were differentially expressed between the higher and lower growth rate groups.

Fishes as models in genomics

Graciano et al. have presented the first *de novo* genomic assembly for *Salminus brasiliensis*, one of the most important species for angling and consumption in southern South America. This study also presented the coding genome annotations of 12,962 putative genes from assembled genomic fragments over 10 kb as well as a genome-wide panel for predicted microsatellites.

The discovery of molecular markers is very important in fish studies, as they can be employed for the detection of adulteration and authenticity (Kotsanopoulos et al., 2021), toxicological research (Santos et al., 2018; Kar et al., 2021), conservation (Tan et al., 2019; Bourret et al., 2020), and phylogenetic studies (Saad, 2019; Prabhu et al., 2020). In this Research Topic, Goes et al. have characterized three new satellitomes from three species: *Psalidodon fasciatus*, *P. bockmanni*, and *Astyanax lacustris*. Finally, Pagano et al. have shown that the microRNA miR-429 can be differentially expressed in *Odontesthes humensis*, a possible biomarker for osmotic stress.

Conclusion

In summary, this Research Topic presents new advances in molecular studies and the proposal of fishes of different species as biological models in order to understand the evolution and distribution of species, to produce new molecules for biotechnological applications, and to optimize aquaculture and conservation of species. Fishes are gaining more importance in scientific research. Therefore, we hope this topic could instigate the readers to look for new questions and find new answers, walking in the Mendel and Darwin's footsteps and going beyond guided by knowledge that the fishes brought us after them.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and have approved it for publication.

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.

Publisher's note

References

Bourret, V., Albert, V., April, J., Côté, G., and Morissette, O. (2020). Past, present and future contributions of evolutionary biology to wildlife forensics, management and conservation. *Evol. Appl.* 13, 1420–1434. doi:10.1111/eva.12977

Costa, B. A., Lima Júnior, E. M., de Moraes Filho, M. O., Fechine, F. V., de Moraes, M. E. A., Silva Júnior, F. R., et al. (2019). Use of Tilapia skin as a xenograft for pediatric burn treatment: A case report. *J. Burn Care Res.* 40, 714–717. doi:10.1093/jbcr/irz085

Darwin, C. R. (1868). *The variation of animals and plants under domestication*. London: John Murray.

Dias, M. T. P. M., Bilhar, A. P. M., Rios, L. C., Costa, B. A., Lima Júnior, E. M., Alves, A. P. N. N., et al. (2020). Neovaginoplasty using nile Tilapia fish skin as a new biologic graft in patients with mayer-rokitansky-küster-hauser syndrome. *J. Minim. Invasive Gynecol.* 27, 966–972. doi:10.1016/j.jmig.2019.09.779

Kar, S., Sangem, P., Anusha, N., and Senthilkumaran, B. (2021). Endocrine disruptors in teleosts: Evaluating environmental risks and biomarkers. *Aquac. Fish.* 6, 1–26. doi:10.1016/j.aaf.2020.07.013

Kotsanopoulos, K. V., Exadactylos, A., Gkafas, G. A., Martsikalis, P. V., Parlapani, F. F., Boziaris, I. S., et al. (2021). The use of molecular markers in the verification of fish and seafood authenticity and the detection of adulteration. *Compr. Rev. Food Sci. Food Saf.* 20, 1584–1654. doi:10.1111/1541-4337.12719

Mayr, E. (2004). *What makes biology unique?: Considerations on the autonomy of a scientific discipline*. Cambridge: Cambridge University Press.

Prabhu, V. R., Singha, H. S., Kumar, R. G., Gopalakrishnan, A., and Nagarajan, M. (2020). Characterization of the complete mitochondrial genome of *Barilius malabaricus* and its phylogenetic implications. *Genomics* 112, 2154–2163. doi:10.1016/j.ygeno.2019.12.009

Saad, Y. M. (2019). Analysis of 16S mitochondrial ribosomal DNA sequence variations and phylogenetic relations among some serranidae fishes. *S. Afr. J. Anim. Sci.* 49, 80. doi:10.4314/sajas.v49i1.10

Santos, D., Vieira, R., Luzio, A., and Félix, L. (2018). "Zebrafish early life stages for toxicological screening: Insights from molecular and biochemical markers," in *Advances in molecular toxicology* (Netherlands: Elsevier), 151–179. doi:10.1016/B978-0-444-64199-1.00007-5

Tan, M. P., Wong, L. L., Razali, S. A., Afiqah-Aleng, N., Mohd Nor, S. A., Sung, Y. Y., et al. (2019). Applications of next-generation sequencing technologies and computational tools in molecular evolution and aquatic animals conservation studies: A short review. *Evol. Bioinform. Online* 15, 1176934319892284. doi:10.1177/1176934319892284

Usai, A., Di Franco, G., Colucci, P., Pollina, L. E., Vasile, E., Funel, N., et al. (2020). A model of a zebrafish avatar for Co-clinical trials. *Cancers* 12, 677. doi:10.3390/cancers12030677

Watson, J. D., Baker, T. A., Bell, S. P., Gann, A., Levine, M., and Losick, R. (2013). *Molecular biology of the gene*. 7th ed. San Francisco: CSHL Press & Benjamin Cummings.

Wright, J. R., Yang, H., Hyrtsenko, O., Xu, B., Yu, W., and Pohajdak, B. (2014). A review of piscine islet xenotransplantation using wild-type Tilapia donors and the production of transgenic Tilapia expressing a "humanized" Tilapia insulin. *Xenotransplantation* 21, 485–495. doi:10.1111/xen.12115

frontiers in Genetics

ORIGINAL RESEARCH
published: 28 October 2021
doi: 10.3389/fgene.2021.779464

Comparative Cytogenetics Analysis Among *Peckoltia* Species (Siluriformes, Loricariidae): Insights on Karyotype Evolution and Biogeography in the Amazon Region

Kevin Santos da Silva[1], *Augusto Cesar Paes de Souza*[2], *Ananda Marques Pety*[1], *Renata Coelho Rodrigues Noronha*[1], *Marcelo Ricardo Vicari*[3], *Julio Cesar Pieczarka*[1†] *and Cleusa Yoshiko Nagamachi*[1*†]

[1]*Laboratório de Citogenética, Centro de Estudos Avançados da Biodiversidade, Instituto de Ciências Biológicas, Universidade Federal Do Pará, Belém, Brazil,* [2]*Laboratório de Estudos da Ictiofauna da Amazônia, Instituto Federal de Educação Ciência e Tecnologia Do Pará, Abaetetuba, Brazil,* [3]*Laboratório de Biologia Cromossômica, Estrutura e Função, Departamento de Biologia Estrutural, Molecular e Genética, Universidade Estadual de Ponta Grossa, Ponta Grossa, Brazil*

Peckoltia is widely distributed genus in the Amazon and Orinoco basins and the Guiana Shield, containing 18 valid species, and distinct morphotypes still needing description in the scientific literature due to its great taxonomic complexity. This study performed a comparative chromosomal analysis of two undescribed *Peckoltia* species (*Peckoltia* sp. 3 Jarumã and *Peckoltia* sp. 4 Caripetuba) from the Brazilian Amazon using conventional chromosome bands methods and *in situ* localization of the repetitive DNA (*5S* and *18S rRNA* and *U1 snRNA* genes and telomeric sequences). Both species presented 2n = 52 but differed in their karyotype formula, probably due to inversions or translocations. The nucleolus organizer regions (NORs) showed distal location on a probably homeologous submetacentric pair in both species, besides an extra signal in a subtelocentric chromosome in *Peckoltia* sp. 4 Caripetuba. Heterochromatin occurred in large blocks, with different distributions in the species. The mapping of the 18S and 5S rDNA, and U1 snDNA showed differences in locations and number of sites. No interstitial telomeric sites were detected using the (TTAGGG)n probes. Despite 2n conservationism in *Peckoltia* species, the results showed variation in karyotype formulas, chromosomal bands, and locations of repetitive sites, demonstrating great chromosomal diversity. A proposal for *Peckoltia* karyotype evolution was inferred in this study based on the diversity of location and number of chromosomal markers analyzed. A comparative analysis with other *Peckoltia* karyotypes described in the literature, their biogeography patterns, and molecular phylogeny led to the hypothesis that the derived karyotype was raised in the left bank of the Amazon River.

Keywords: neotropical fish, snRNA, rDNA, biodiversity, amazon

OPEN ACCESS

Edited by:
Tony Silveira,
Federal University of Rio Grande, Brazil

Reviewed by:
Vladimir Pavan Margarido,
Universidade Estadual do Oeste do Paraná, Brazil
Analía Del Valle Garnero,
Federal University of Pampa, Brazil
Lucia Caetano,
State University of Londrina, Brazil

****Correspondence:***
Cleusa Yoshiko Nagamachi
cleusanagamachi@gmail.com
cleusa@ufpa.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: *18 September 2021*
Accepted: *18 October 2021*
Published: *28 October 2021*

Citation:
Santos da Silva K, de Souza ACP, Pety AM, Noronha RCR, Vicari MR, Pieczarka JC and Nagamachi CY (2021) Comparative Cytogenetics Analysis Among Peckoltia *Species (Siluriformes, Loricariidae): Insights on Karyotype Evolution and Biogeography in the Amazon Region. Front. Genet. 12:779464. doi: 10.3389/fgene.2021.779464*

INTRODUCTION

The Loricariidae is one of the most specious family of catfish within the order Siluriformes, containing 1,016 valid species (Fricke et al., 2021). They are endemic to the Neotropical region, distributed throughout South America and part of Central America, and occur in a great diversity of habitats (Armbruster, 2004; Armbruster, 2008; Armbruster and Lujan, 2016). Analyzes based on morphological and molecular data support the recognition of six subfamilies: Lithogeninae, Delturinae, Hypoptopomatinae, Neoplecostominae, Loricariinae, and Hypostominae grouped in the tribes: Corymbophanini, Rhinelepini, Hypostomini, Pterygoplichthyini and Ancistrini (Armbruster, 2004; Lujan et al., 2015).

Peckoltia Miranda Ribeiro, 1912 (*sensu* Lujan et al., 2015), comprises 18 valid species, in addition to distinct morphotypes that still lack description in the scientific literature. They are widely distributed in the Amazon and Orinoco basins and the Guiana Shield (Armbruster et al., 2015; Lujan et al., 2015; Armbruster and Lujan, 2016). According to phylogenetic analyzes proposed for Loricariidae, *Peckoltia* genus receive strong support as monophyletic lineage (Lujan et al., 2015; Lujan et al., 2017; Roxo et al., 2019); however, a complex taxonomic identification procedure at a specific level related to a wide geographic distribution and morphological similarity is observed for these species (Armbruster et al., 2015; Armbruster and Lujan, 2016). For example, representatives of *Peckoltia vittate* (the type species of the genus) collected in the Amazon region (Xingu, Madeira and Orinoco rivers) presented polyphyletic lineages in molecular phylogeny by Lujan et al. (2015). Armbruster and Lujan (2016), analyzed morphological and molecular characters of samples collected in the Orinoco basin associated with *P. vittate* by Lujan et al. (2015) and described a new species, *Peckoltia wernekei*. These recent analyses agree that the diversity of *Peckoltia* species can be underestimated for the Amazon region.

Cytogenetic markers are important tools to analyze fish species possessing complex taxonomy (Bertollo et al., 2000; Centofante et al., 2003; Mariotto and Miyazawa, 2006) or to understand evolutionary features in groups with highly rearranged karyotypes (Nagamachi et al., 2010; Deon et al., 2020). Cytogenetic data are available for eight lineages of the genus *Peckoltia,* including valid species and unidentified morphotypes, collected at different points in the Amazon region. Despite all *Peckoltia* species share 2n = 52 chromosomes, variations in chromosome morphology, the number and position of NORs, distribution of the constitutive heterochromatin (CH) regions, and the presence of B chromosomes are observed among species of this genus (**Table 1**). Therefore, in *Peckoltia* species, many unique karyotypic features are observed that can be useful in recognizing distinct taxonomic units.

Repetitive DNAs are found in most eukaryotic genomes, representing important markers for molecular diversity analysis at the chromosomal level; they are organized in blocks (e.g., satellites and multigene families) or dispersed (e.g., transposons and retrotransposons). The contribution of the repetitive DNA for fish genome evolution has been evidenced (Vicari et al., 2010; Schemberger et al., 2019). The eukaryotic ribosomal DNA (rDNA) represents two multigene families with an organization *in tandem*: *45S ribosomal RNA* (*18S + 5.8S + 28S* genes) and *5S ribosomal RNA* (Long and Dawin, 1980). These genes are widely used in chromosomal studies in several organisms, including *Peckoltia* species (Pety et al., 2018), showing great molecular chromosomal diversity involving these sequences.

Small nuclear *RNA* genes (snDNA) represent another multigene family involved in the splicing and maturation process of messenger RNA encoded by the *U1*, *U2*, *U4*, *U5* and *U6 snRNA* genes (Busch et al., 1982). The snDNA sequences have been used as chromosomal markers for detailed comparative chromosome analysis in several groups of organisms, including fish, reptiles and arthropods (Cabral-de-Melo et al., 2012; Almeida et al., 2017; Cavalcante et al., 2020; Dulz et al., 2020).

In this study, we describe the karyotypes of two undescribed *Peckoltia* species, first time sampled in the Tocantins River basin

TABLE 1 | Chromosomal diversity available in the literature and obtained in the present study for *Peckoltia* genus.

Species	Classic cytogenetics			Molecular cytogenetics				Loc.	Ref.
	2n	KF	Nor	18S rDNA	5S rDNA	U1 snDNA	Tel.		
Peckoltia sp. 1	52+1B	44m/sm+6st+2a+1B	multiple	-	-	-	-	A	1
Peckoltia sp. 2	52	32m/sm+18st+2a	multiple	-	-	-	-	A	1
P. vittata	52	36m/sm+14st+2a	simple	-	-	-	-	B	1
P. vittata	52	34m/sm+18st	simple	simple	multiple	-	-	B	2
P. sabaji	52	38m/sm+14st	multiple	multiple	multiple	-	-	B	2
P. oligospila	52	38m/sm+14st	multiple	multiple	simple	-	-	C	2
P. cavatica	52	38m/sm+14st	simple	multiple	multiple	-	-	C	2
P. multipinis	52	28m/sm+24st	simple	Simple	simple	-	-	B	2
Peckoltia sp. 3 Jarumã	52	46m/sm+6st	simple	simple	multiple	simple	distal	D	3
Peckoltia sp. 4 Caripetuba	52	40m/sm+12st	multiple	multiple	simple	multiple	distal	D	3

Locality: A—Monte Dourado, Pará state, Brazil; B—Altamira, Pará state, Brazil; C—Ourém, Pará state, Brazil; D—Abaetetuba, Pará state, Brazil. References: 1—de Souza et al., 2009; 2—Pety et al., 2018; 3—Present Study. Abbreviations: Diploid number (2n); Karyotype Formula (KF); Nucleolus Organizing Regions (NOR); Ribosomal RNA gene (rDNA); Small nuclear RNA gene (snDNA); Telomere (tel.); Supernumerary ou B chromosome (B); long arm (q); short arm (p); metacentric (m), submetacentric (sm), subtelocentric (st), acrocentric (a); Locality (Loc.); References (Ref.); (–) Data no available.

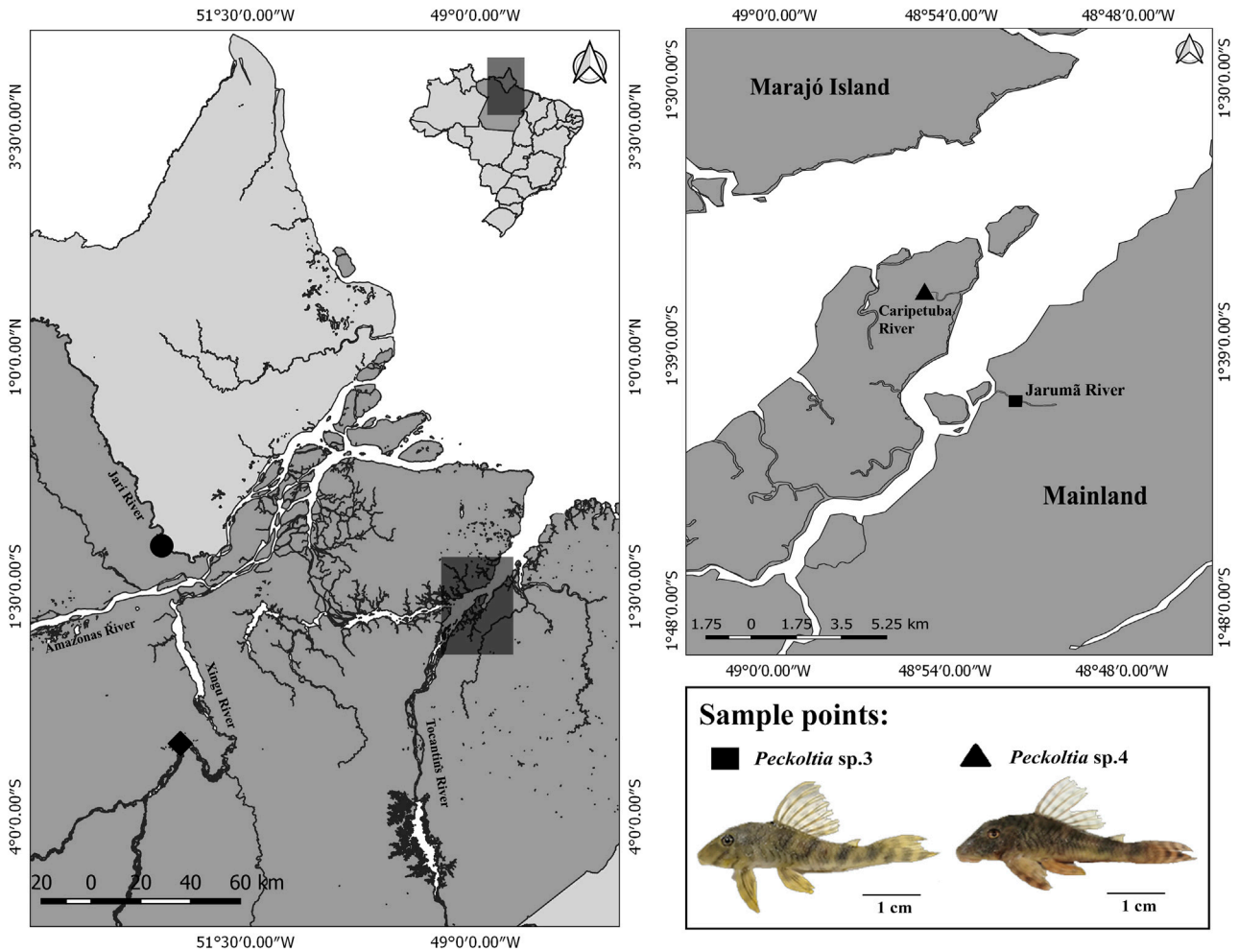


FIGURE 1 | Geographical location of sampling sites for *Peckoltia* sp. 3 Jarumã (square) and *Peckoltia* sp. 4 Caripetuba (triangle) in the present study. The coordinates of the sampling sites for *Peckoltia vittata* (diamond), *Peckoltia* sp. 1, and *Peckoltia* sp. 2 (circle) reported by de Souza et al. (2009) are also shown. The map was made using QUANTUM-GIS (Q-GIS) v. 3.4.5. The database was obtained from Instituto Brasileiro de Geografia e Estatística—IBGE. An *Peckoltia* specimen is shown below. Scale bar: 1 cm. Photo by KSS.

in Brazil, and compare them with cytogenetic data available in the literature. From this, we discuss the possible mechanisms of karyotypic diversification, biogeography and their evolutionary implications for this genus.

MATERIALS AND METHODS

Samples

Samples of the two morphologically different but still undescribed species of *Peckoltia* named *Peckoltia* sp. 3 Jarumã and *Peckoltia* sp. 4 Caripetuba, after the rivers they were collected in different hydrographic points in the Tocantins River basins of northern Brazil were analyzed (**Figure 1**). The taxonomic identification of the sample was checked using the identification key proposed by Armbruster and Lujan (2016), Armbruster (2008). The results show that the specimens do not fit into any of the species already **Figure 1**described. The collection points, number of individuals, sex, and voucher of deposits in the zoological collection are shown in **Table 2**. The samples were obtained under a permanent field permit obtained by JCP (number 13248 from "Instituto Chico Mendes de Conservação da Biodiversidade"). The Cytogenetics Laboratory from UFPA has permit number 19/2003 from the Ministry of Environment for sample transport and permit 52/2003 to use the samples for research. The Ethics Committee (Comitê de Ética Animal da Universidade Federal do Pará) approved this research (Permit 68/2015). The specimens have been deposited in the ichthyological collection of the Museu Paraense Emílio Goeldii (MPEG) (Belém, Brazil).

Chromosomal Analysis

Mitotic chromosomes were obtained from kidney cells after *in vivo* colchicine treatment as described (Bertollo et al., 1978). The animals were anesthetized with eugenol and subsequently sacrificed for the removal of kidney cells. Metaphases were analyzed by conventional Giemsa, C-banding (Sumner, 1972)

TABLE 2 | Samples and collection sites for *Peckoltia* species analyzed in this study.

Species	Sex		River	City/State	Voucher	Coordinates	
Peckoltia sp. 3	2♂	4♀	Jarumã River	Abaetetuba-PA	MPEG 38949	01°42′41.9″S	48°51′45.9″W
Peckoltia sp. 4	1♂	1♀	Caripetuba River	Abaetetuba-PA	MPEG 38950	01°37′23.49″S	48°55′33″W

MPEG—Museu Paraense Emilio Goeldi.

and $AgNO_3$ staining (Howell and Black, 1980). Fluorescence *in situ* hybridization (FISH) was undertaken as described (Martins and Galetti, 1999) using a general telomere probe for vertebrates, 18S rDNA, 5S rDNA, and U1 snDNA probes.

Probes Labeling and *in situ* Localization

DNA extraction was performed using PureLink Genomic DNA Mini Kit (Invitrogen) following the manufacturer's instructions. The probes were obtained from a PCR using genomic DNA of *Peckoltia* sp. 3 Jarumã and *Peckoltia* sp. 4 Caripetuba with primers previously described for 18S rDNA (Hatanaka and Galetti Jr, 2004), for 5S rDNA (Suarez et al., 2017) and U1 snDNA (Cabral-de Melo et al., 2012). These probes were labeled by nick-translation with biotin or digoxigenin. Telomeric probes were obtained from PCR using the set of primers F-5′$(TTAGGG)_5$-3′ and R-5′$(CCCTAA)_5$-3′ followed by labeling with Digoxigenin-11-dUTP (Roche Applied Science®) (Ijdo et al., 1991). Fluorescence *in situ* hybridization (FISH) was performed as described by Martins and Galetti (1999) using the following stringency conditions: 2.5 ng/μL of each probe, 50% formamide, 2 x SSC, 10% dextran sulfate, and hybridization at 42°C for 16 h. Fluorescent signals were detected using Streptavidin Alexa Fluor 488 (Molecular Probes, Carlsbad, CA, United States) and anti-digoxigenin rhodamine Fab fragments (Roche Applied Science, Penzberg, Germany). Chromosomes were counterstained with 0.2 μg/ml of 4′6-diamidino-2-phenylindole (DAPI) in Vectashield mounting medium (Vector, Burlingame, CA, United States).

Microscopic Analysis and Image Capture

At least 30 metaphases Giemsa-stained per individual were analyzed for confirming the diploid number, karyotypic structure, and chromosomal markers. Cytogenetic images of Giemsa-stained chromosomes were obtained using an Olympus BX41 microscope (bright field/phase) with a digital camera CCD 1300QDS and analyzed using GenASIs software version 7.2.7.34276 from ASI (Applied Spectral Imaging). FISH images were obtained using a Nikon H550S microscope and analyzed using the Nis-Elements software. Images were adjusted using Adobe Photoshop CS6 software. Chromosomal morphology was classified according to literature (Levan et al., 1964), with adaptations.

RESULTS

Peckoltia sp. 3 Jarumã and *Peckoltia* sp. 4 Caripetuba species presented 2n = 52 chromosomes and karyotype formulas (KF) 46m/sm + 6st, and 40m/sm + 12st, respectively. Heteromorphic sex chromosomes were not identified in the karyotypes described here (**Figures 2A,C**).

In *Peckoltia* sp. 3 Jarumã, heterochromatin occurred in the interstitial region of the short arm (p) of pairs 1p, 5p and 11p; in large blocks in the long arm (q) in pairs 9q with size heteromorphism and 12q; and in the pericentromeric and interstitial region of the 24q pair (**Figure 2B**). In *Peckoltia* sp. 4 Caripetuba, heterochromatin occurred in the interstitial region of pairs 1p, 14p, 19p and 20p; in the centromeric region of par 4; in large blocks in pairs 11q, 12q and 13q and in the pericentromeric and distal region of pair 21q, in which there is a marked heteromorphism in size between the homologues (**Figure 2D**). The Ag-NORs were located distal at pair 9q adjacent to the heterochromatin block in *Peckoltia* sp. 3 Jarumã (**Figure 2** in box), and at pair 11q, coincident with heterochromatin and in one of the homologues in pair 21q in *Peckoltia* sp. 4 Caripetuba (**Figure 2** in box).

Telomeric sequences occurred in the distal region of all chromosomes in both species, with no evidence of ITS vestiges (**Figures 3A,C**). The 18S rDNA, 5S rDNA, and U1 snDNA probes hybridized at the distal position of the chromosomes in the karyotypes of both species. In *Peckoltia* sp. 3 Jarumã, both 18S and 5S rDNA are colocalized in the 9q pair, in addition to a 5S rDNA site in one of the homologues in the 24q pair, and U1 snDNA is located in the distal region of the pair 13q (**Figure 3B**). In *Peckoltia* sp. 4 Caripetuba, 18S rDNA is located in pair 11q, in an additional site in one homologue of pair 21q; 5S rDNA is located in pair 1p; and U1 snDNA is located in the distal position of the pair 17q, with an additionnal site in one homologue of pair 21q colocalized with heterochromatin and 18S rDNA site (**Figure 3D**).

DISCUSSION

Karyotype Diversity in *Peckoltia*

Cytogenetic information for the species of *Peckoltia* is described in **Table 1** and, despite the occurrence of 2n = 52 chromosomes, all species of *Peckoltia* show different KF. These differences can be explained by inversions or translocations, which represent important mechanisms of karyotypic diversification in Loricariidae (Kavalco et al., 2005; Mariotto et al., 2011). Alternatively, centromeric repositioning has also been proposed to cause variations in chromosome morphology with no change in diploid number (Montefalcone et al., 1999; Rocchi et al., 2012). However, the occurrence of this mechanism in Loricariidae karyotypes still needs deep investigation.

The Loricariidae has extensive chromosomal diversity, with variation in diploid number from 34 to 96 chromosomes, with a

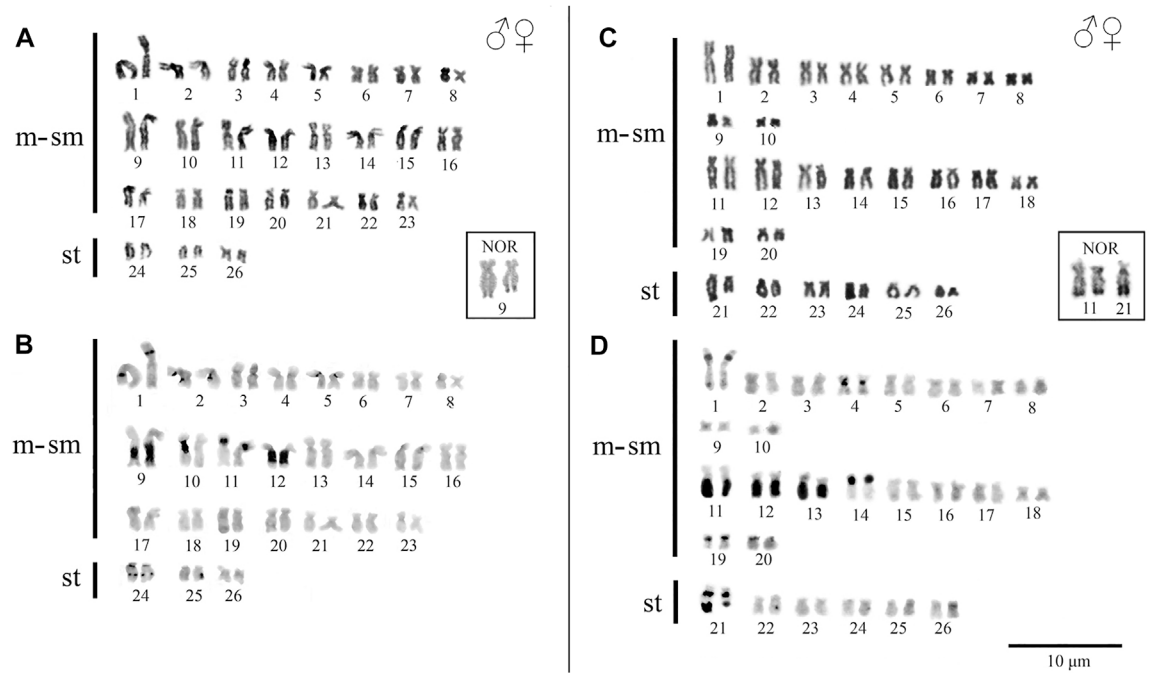


FIGURE 2 | Karyotypes of *Peckoltia* species. In **(A)** conventional staining and **(B)** C-banding of *Peckoltia* sp. 3 Jarumã; in **(C)** conventional staining and **(D)** C-banding of *Peckoltia* sp. 4 Caripetuba. In box the Nucleolus Organizing Regions (NORs).

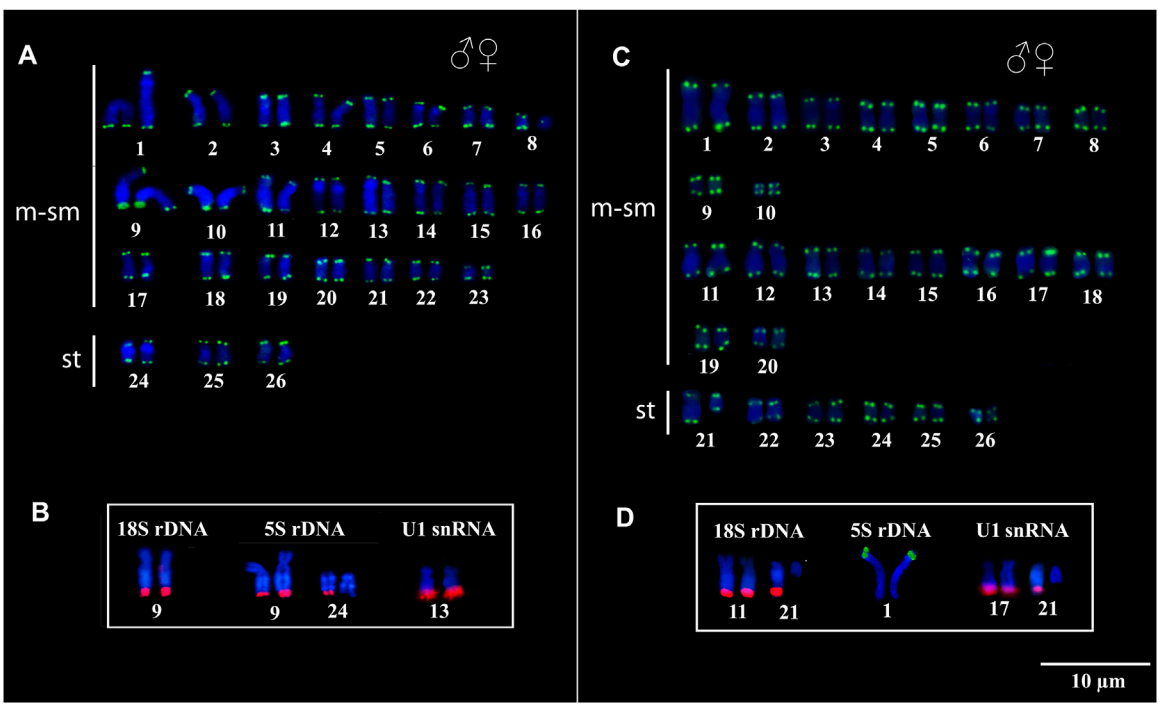


FIGURE 3 | Fluorescent *in situ* Hybridization indicating the physical location of the Telomeric probes, 18S rDNA, 5S rDNA and U1 snDNA probes in *Peckoltia* species. In *Peckoltia* sp. 3 Jarumã: **(A)** Telomeric sequence (green) and **(B)** 18S rDNA (red), 5S rDNA (red) and U1 snDNA (red). In *Peckoltia* sp. 4 Caripetuba: (**C**) Telomeric sequence (green) and **(D)** 18S rDNA (red), 5S rDNA (green) and U1 snDNA (red).

putative ancestral karyotype showing 2n = 54 chromosomes (Artoni and Bertollo, 2001). The reduction to 2n = 52 was probably due to a Robertsonian fusion in an ancient common ancestor in Ancistrini representatives with no ITS manutention (Bueno et al., 2018).

In Loricariidae, it is expected the presence of heterochromatic regions distributed in blocks on few chromosomes (Ziemniczak et al., 2012). Interestingly, in *Peckoltia*, extensive heterochromatic blocks are observed, some occupying a large part of the long arms of submetacentric/subtelocentric chromosomes (de Souza et al., 2009, present study). The presence of large heterochromatic blocks on morphologically similar chromosomes may suggest a shared character in *Peckoltia* karyotypes, as proposed previously (de Souza et al., 2009). However, it is known that heterochromatic regions are characterized by great diversity in highly repetitive DNA content (Dimitri et al., 2009) and may not reflect chromosomal homologies in *Peckoltia* species, as visualized among the species analyzed here by FISH with repetitive sequences. Noteworthy, the distribution of heterochromatin, and different repetitive sequences, observed in pairs 24 in *Peckoltia* sp. 3 Jarumã and 21 in *Peckoltia* sp. 4 Caripetuba, makes these chromosomes good cytotaxonomic markers, and both represent unique characteristics in each of these species.

Other plesiomorphic conditions for Loricariidae include the 18S and 5S rDNA sequences in a single pair of meta/submetacentric chromosomes (Ziemniczak et al., 2012). However, in the Ancistrini tribe, both the synteny and non-synteny between the 18S and 5S sequences are commonly observed (Mariotto et al., 2011; Ribeiro et al., 2015; Favarato et al., 2016; Prizon et al., 2016; Pety et al., 2018), showing the huge chromosome sites variation in the karyotypes of this group (Pansonato-Alves et al., 2013; Bueno et al., 2014; Pety et al., 2018). The mapping of 18S and the 5S sequences in the karyotypes here described is in agreement with that observed in other species of *Peckoltia*; in which extensive dispersion of these genes is observed (Pety et al., 2018) (**Table 1**). This dispersion of rDNA in the genomes of Loricariidae can be explained either by the association of these genes with other repetitive sequences, including transposable elements or by the evolutionary breakpoint regions close to rDNA sites promoting chromosome rearrangements (Barros et al., 2017; Glugoski et al., 2018, 2020; Deon et al., 2020). Furthermore, the heterochromatic condition involving clusters of rDNA suggests that other repetitive DNA classes, around 45S and 5S rDNA sequences, may promote their chromosomal dispersion in the *Peckoltia* species analyzed here, as shown for other species of Loricariidae (Glugoski et al., 2018; Deon et al., 2020).

In fish, the *snRNA* genes have shown great diversity of the pattern of chromosomal localization (Úbeda-Manzanaro et al., 2010; Cabral-de-Melo et al., 2012; Scacchetti et al., 2015; Yano et al., 2017). In this work, the U1 snDNA sequence was mapped for the first time in Loricariidae species, showing location in a pair of submetacentric chromosomes in both species, in addition to an extra site in one of the homologues of pair 21 of *Peckoltia* sp. 4 Caripetuba (**Figure 3D**). The snDNA sequences are considerably more stable at the chromosomal level when compared to rDNA (Cabral-de-Melo et al., 2012). However, we observed that among the karyotypes of *Peckoltia* analyzed here, the U1 snDNA probes show variation in the number of chromosomal sites similar to that observed for the rDNA (**Table 1**). These data indicate that these gene families can be equally dynamic in the genomes of species of *Peckoltia*. Several chromosomal sites of rDNA and snDNA sequences are observed in different groups of fish, such as species of the Loricariidae, Cichlidae and Anostomidae families; the emergence of new chromosomal sites is related to the association of these sequences with active mobile elements in these organisms (Kapitonov and Jurka, 2006; Cabral-de-Melo et al., 2012; Dulz et al., 2020). Future analyzes of rDNA and snDNA nucleotide sequences will be essential to verify the possible involvement of transposable elements in the movement of these sequences in the genomes of *Peckoltia* species.

Biogeography Hypothesis in *Peckoltia*

The putative ancestral karyotype for Loricariidae has 2n = 54, a single NOR and gene synteny for 5S and 18S rDNA sequences (Ziemniczak et al., 2012). The tribes belonging to the Hypostominae subfamily share a common ancestor (Armbruster, 2004; Lujan et al., 2015) that possibly had a 2n = 52 chromosomes (Bueno et al., 2018). Thus, variations in the 2n, multiple NOR and synteny break between 5S and 18S would represent derived characteristics that can be apomorphic or homoplasic. Analyses involving 18S and 5S rDNA and U1 snDNA show the importance of these sequences as markers of karyotype diversification in the *Peckoltia* genus.

Phylogenetic analyzes support the monophyly of the *Peckoltia* genus (Lujan et al., 2015, 2017; Roxo et al., 2019). A phylogeny for the *Peckoltia* clade proposed by Lujan et al. (2017), based on the concatenated sequences of two mitochondrial genes (16S, Cyt b) and three nuclear genes (RAG1, RAG2, MyH6), has two well-defined branches, which are sister groups and two branches with non-defined relationships. One of the defined branches presents *P. vittata* (single NOR, synteny of the 5S with 18S) (Pety et al., 2017), *P. compta*, *P. braueri* and *P. lineola* (karyotypes not described); and the other branch with *P. sabaji* (multiple NORs, non-synteny of the 5S with 18S) (Pety et al., 2018), *P. furcata* and *P. relictum* (karyotypes not described). Noteworthy, most of the specimens from the branch with *P. vittata* are on the right bank of the lower Amazon River, and those from the branch with *P. sabaji* are on the left bank. The karyotype of *Peckoltia* sp. 4 Caripetuba would be more similar to that of *Peckoltia* sp. 1 and 2 previously described (de Souza et al., 2009), collected on the left bank of the Amazon River and with karyotypic characteristics derived from the ancestral karyotype proposed for Loricariidae. Thus, it is possible that these derived features, such as multiple rDNA and NORs sites, have arisen in this region (**Figure 4**).

In the present work, the karyotype of *Peckoltia* sp. 4 Caripetuba has several derived characteristics, being found on the right bank of the Amazon River. This fact can be explained if we consider the paleogeography of the region. The continental portion of Abaetetuba (Rio Jarumã) comprises the Barreiras Formation (Miocene). In turn, the Post-Barreiras Formation (Plio-Pleistocene) filled the paleovale of the Tocantins River, diverting this river (which originally crossed Marajó Island to its northern portion) and thus splitting Marajó Island from the

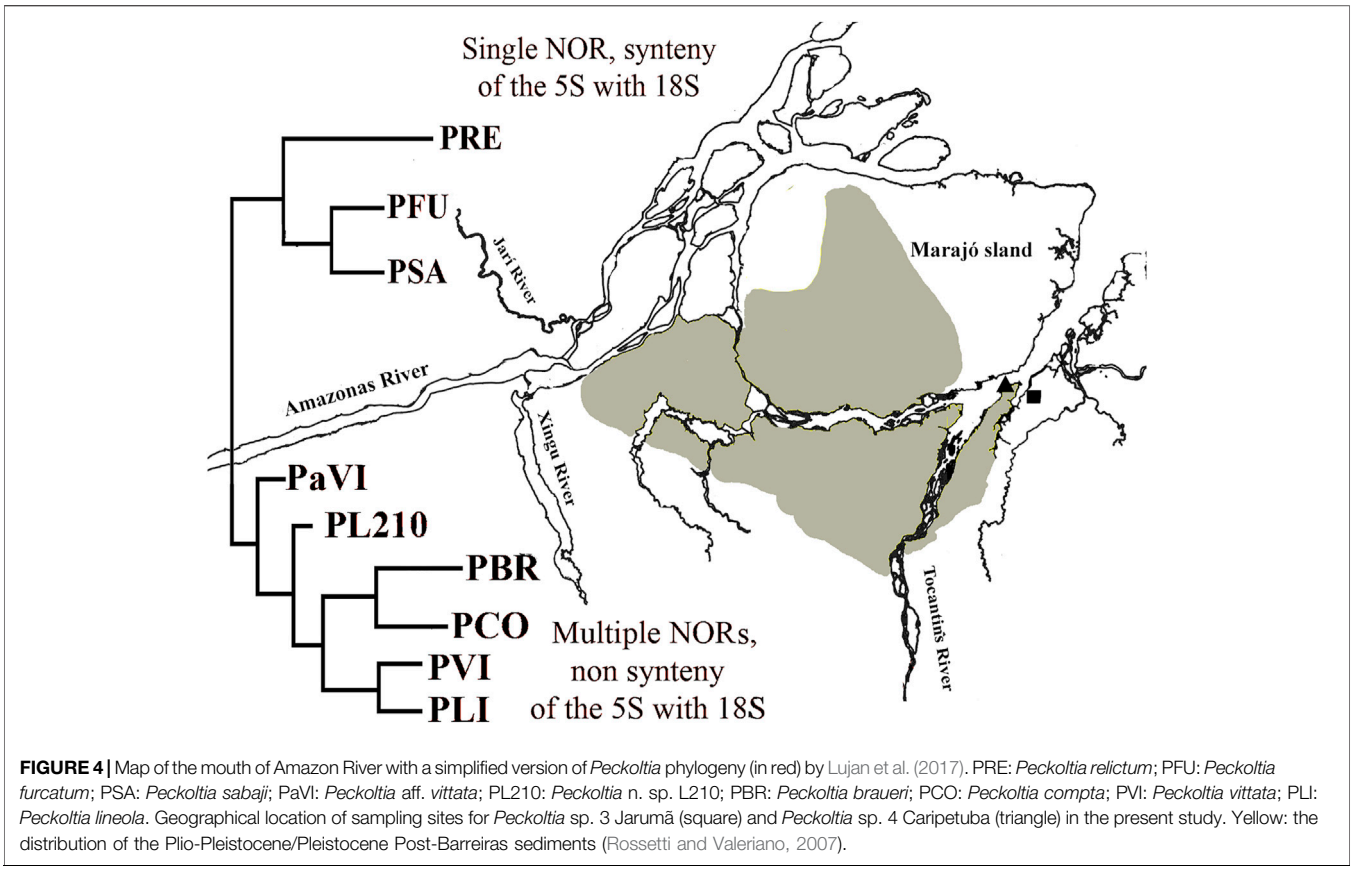


FIGURE 4 | Map of the mouth of Amazon River with a simplified version of *Peckoltia* phylogeny (in red) by Lujan et al. (2017). PRE: *Peckoltia relictum*; PFU: *Peckoltia furcatum*; PSA: *Peckoltia sabaji*; PaVI: *Peckoltia* aff. *vittata*; PL210: *Peckoltia* n. sp. L210; PBR: *Peckoltia braueri*; PCO: *Peckoltia compta*; PVI: *Peckoltia vittata*; PLI: *Peckoltia lineola*. Geographical location of sampling sites for *Peckoltia* sp. 3 Jarumã (square) and *Peckoltia* sp. 4 Caripetuba (triangle) in the present study. Yellow: the distribution of the Plio-Pleistocene/Pleistocene Post-Barreiras sediments (Rossetti and Valeriano, 2007).

mainland, originating Rio Pará and the islands in front of the city of Abaetetuba-PA (**Figure 4**), where the Caripetuba River is found (Rossetti and Valeriano, 2007). Therefore, the rivers where the karyotype of *Peckoltia* sp. 4 Caripetuba is found are in islands considerably newer than the one where *Peckoltia* sp. 3 Jarumã is found. These islands are covered by the Post-Barreiras Formation, being connected to the western portion of Marajó Island (Tatumi et al., 2007), not related with the continental region where the Jarumã River is located. The Marajó Island, in turn, is a communication corridor, connecting biodiversity on the left bank of the Amazon River with biodiversity on the right bank (Fernandes et al., 1995). This fact may explain why the most recent karyotype (*Peckoltia* sp. 4 Caripetuba) is located in this region. It will be important to test by molecular markers whether the *Peckoltia*-associated morphotypes analyzed in this study and by de Souza et al. (2009), which have derived karyotypes, belong to the *P. sabaji* branch.

This study describes the karyotypes of two undescribed species of *Peckoltia* and compares them with available chromosomal data for the genus. The maintenance of the same 2n in the species of *Peckoltia* may suggest that this genus has conserved karyotypes; however, the variations observed in the KF, NOR, heterochromatin and 18S and 5S rDNA sequences between the karyotypes of the species in this study compared to those previously described suggest great interspecific diversity in the genus. Furthermore, the differential localization of the U1 snDNA sequence among the karyotypes described here corroborates the involvement of repetitive sequences in the diversification of the genomes of these species. Therefore, due to the considerable cytogenetic diversity, species-specific characters were observed, showing great potential for identifying distinct taxonomic lineages in *Peckoltia*, in addition to demonstrating that the karyotypic variation in this genus is much greater than conventional staining suggests. Future analyzes considering the geographic distribution of *Peckoltia* species with primitive versus derived karyotypic characteristics compared to molecular phylogenies may provide relevant information about its evolutionary history.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by Comitê de Ética Animal da Universidade Federal do Pará.

AUTHOR CONTRIBUTIONS

KS: Conceptualization; Data Curation; Formal analysis; Investigation; Methodology; Visualization; Writing original draft; Writing review and editing. AS: Investigation; Methodology; Resources; Visualization; Writing review and editing. AP: Investigation; Methodology; Visualization; Writing review and editing. RN: Investigation; Visualization; Writing review and editing. MV: Methodology; Resources; Visualization; Writing review and editing. JP: Data Curation; Formal analysis; Funding acquisition; Resources; Visualization; Writing review and editing. CN: Data Curation; Formal analysis; Funding acquisition; Project administration; Resources; Supervision; Visualization; Writing review and editing.

FUNDING

The Pro-Reitoria de Pesquisa e Pos-graduação da Universidade Federal do Para supported by paying the open access publication fees. The authors thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support on project coordinated by CYN (Edital Pró-Amazônia Proc 047/2012); the FAPESPA for financial support (Edital Vale–Proc 2010/110447) and Banco Nacional de Desenvolvimento Econômico e Social–BNDES (2.318.697.0001) on a project coordinated by JCP. CYN (305880/2017-9), JCP (305876/2017-1) and MRV (305142/2019-4) are grateful to CNPq for Productivity Grants. This study is part of the Master dissertation of KS in Genetic and Molecular Biology who is recipient of CAPES Master Scholarship.

ACKNOWLEDGMENTS

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. Sample collections was authorized by Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) and Secretaria de Estado de Meio Ambiente do Pará (SEMA-PA) under permit 020/2005 (Registration: 207419).

REFERENCES

Almeida, B. R. R., Milhomem-Paixão, S. S. R., Noronha, R. C. R., Nagamachi, C. Y., Costa, M. J. R. D., Pardal, P. P. O., et al. (2017). Karyotype Diversity and Chromosomal Organization of Repetitive DNA in *Tityus Obscurus* (Scorpiones, Buthidae). *BMC Genet.* 18 (1), 35–11. doi:10.1186/s12863-017-0494-6

Armbruster, J. W., and Lujan, N. K. (2016). A New Species of *Peckoltia* from the Upper Orinoco (Siluriformes, Loricariidae). *ZooKeys* 569, 105–121. doi:10.3897/zookeys.569.6630

Armbruster, J. W., Werneke, D. C., and Tan, M. (2015). Three New Species of Saddled Loricariid Catfishes, and a Review of Hemiancistrus, Peckoltia, and Allied Genera (Siluriformes). *Zookeys* 480, 97–123. doi:10.3897/zookeys.480.6540

Armbruster, J. W. (2004). Phylogenetic Relationships of the Suckermouth Armoured Catfishes (Loricariidae) with Emphasis on the Hypostominae and the Ancistrinaefishes Loricariidae with Emphasis on the Hypostominae and the Ancistrinae. *Zoolog. J. Linn. Soc.* 141 (1), 1–80. doi:10.1111/j.1096-3642.2004.00109.x

Armbruster, J. W. (2008). The Genus *Peckoltia* with the Description of Two New Species and a Reanalysis of the Phylogeny of the Genera of the Hypostominae (Siluriformes: Loricariidae). *Zootaxa* 1822 (1), 1–76. doi:10.11646/zootaxa.1822.1.1

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

Barros, A. V., Wolski, M. A. V., Nogaroto, V., Almeida, M. C., Moreira-Filho, O., and Vicari, M. R. (2017). Fragile Sites, Dysfunctional Telomere and Chromosome Fusions: what Is 5S rDNA Role. *Gene* 608, 20–27. doi:10.1016/j.gene.2017.01.013

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 (7), 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 Erythrinidae). *Brazil. J. Genet.* 1, 103–120.

Bueno, V., Venere, P. C., Thums Konerat, J., Zawadzki, C. H., Vicari, M. R., and Margarido, V. P. (2014). Physical Mapping of the 5S and 18S rDNA in Ten Species of Hypostomus Lacépède 1803 (Siluriformes: Loricariidae): Evolutionary Tendencies in the Genus. *ScientificWorldJournal* 2014, 943825. doi:10.1155/2014/943825

Bueno, V., Konerat, J. T., Zawadzki, C. H., Venere, P. C., Blanco, D. R., and Margarido, V. P. (2018). Divergent Chromosome Evolution in Hypostominae Tribes (Siluriformes: Loricariidae): Correlation of Chromosomal Data with Morphological and Molecular Phylogenies. *Zebrafish* 15 (5), 492–503. doi:10.1089/zeb.2018.1612

Busch, H., Reddy, R., Rothblum, L., and Choi, Y. C. (1982). SnRNAs, SnRNPs, and RNA Processing. *Annu. Rev. Biochem.* 51 (1), 617–654. doi:10.1146/annurev.bi.51.070182.003153

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 (2), 279–292. doi:10.1007/s10577-011-9271-y

Cavalcante, M. G., Nagamachi, C. Y., Pieczarka, J. C., and Noronha, R. C. R. (2020). Evolutionary Insights in Amazonian Turtles (Testudines, Podocnemididae): Co-location of 5S rDNA and U2 snRNA and Wide Distribution of Tc1/Mariner. *Biol. open* 9 (4). bio049817, doi:10.1242/bio.049817

Centofante, L., Bertollo, L. A. C., Justi, A. J., and Moreira-Filho, O. (2003). Correlation of Chromosomal and Morphologic Characters in Two Astyanax Species. *Ichthyol. Exploration Freshwaters* 14 (4), 361–368.

De Souza, A., Nagamachi, C., Milhomem, S., Feldberg, E., and Pieczarka, J. (2009). Cytogenetic Analysis in Catfish Species of the Genus Peckoltia Miranda Ribeiro, 1912 (Teleostei: Siluriformes: Loricariidae). *Comp. Cytogenet.* 3, 103–109. doi:10.3897/compcytogen.v3i2.17

Deon, G. A., Glugoski, L., Vicari, M. R., Nogaroto, V., Sassi, F. d. M. C., Cioffi, M. d. B., et al. (2020). Highly Rearranged Karyotypes and Multiple Sex Chromosome Systems in Armored Catfishes from the Genus *Harttia* (Teleostei, Siluriformes). *Genes* 11, 1366. doi:10.3390/genes11111366

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 (4), 419–435. doi:10.1007/s00412-009-0211-y

Dulz, T. A., Azambuja, M., Nascimento, V. D., Lorscheider, C. A., Noleto, R. B., Moreira-Filho, O., et al. (2020). Karyotypic Diversification in Two *Megaleporinus* Species (Characiformes, Anostomidae) Inferred from *In Situ* Localization of Repetitive DNA Sequences. *Zebrafish* 17 (5), 333–341. doi:10.1089/zeb.2020.1918

Favarato, R. M., da Silva, M., de 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 (2), 103–111. doi:10.1089/zeb.2015.1140

Fernandes, M. E. B., Cardoso Da Silva, J. M., and De S. E Silva Junior, J., Jr. (1995). The Monkeys of the Islands of the Amazon Estuary, Brazil: a Biogeographic analysisA Biogeographic Analysis. *Mammalia* 59 (2), 213–221. doi:10.1515/mamm.1995.59.2.213

Fricke, R., Eschmeyer, W. N., and Van der Laan, R. (2021). Eschmeyer's Catalog of Fishes. Genera, Species, References. Available at: http://researcharchive.

calacademy.org/research/ichthyology/catalog/fishcatmain.asp (Electronic version accessed in March 25, 2021).

Glugoski, L., Deon, G. A., Schott, S., Vicari, M. R., Nogaroto, V., and Moreira-Filho, O. (2020). Comparative Cytogenetic Analyses in *Ancistrus* Species (Siluriformes: Loricariidae). *Neotropical Ichthyology* 18, e200013. doi:10.1590/1982-0224-2020-0013

Glugoski, L., Giuliano-Caetano, L., Moreira-Filho, O., Vicari, M. R., and Nogaroto, V. (2018). Co-located hAT Transposable Element and 5S rDNA in an Interstitial Telomeric Sequence Suggest the Formation of Robertsonian Fusion in Armored Catfish. *Gene* 650, 49–54. doi:10.1016/j.gene.2018.01.099

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 (3), 239–244. doi:10.1007/s10709-004-2039-y

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 (8), 1014–1015. doi:10.1007/BF01953855

Ijdo, J. W., Wells, R. A., Baldini, A., and Reeders, S. T. (1991). Improved Telomere Detection Using a Telomere Repeat Probe (TTAGGG)ngenerated by PCR. *Nucl. Acids Res.* 19 (17), 4780. doi:10.1093/nar/19.17.4780

Kapitonov, V. V., and Jurka, J. (2006). Self-synthesizing DNA Transposons in Eukaryotes. *Proc. Natl. Acad. Sci.* 103 (12), 4540–4545. doi:10.1073/pnas.0600833103

Kavalco, K. F., Pazza, R., Bertollo, L. A. C., and Moreira-Filho, O. (2005). Karyotypic Diversity and Evolution of *Loricariidae* (Pisces, Siluriformes). *Heredity* 94 (2), 180–186. doi:10.1038/sj.hdy.6800595

Levan, A., Fredga, K., and Sandberg, A. A. (1964). Nomenclature for Centromeric Position on Chromosomes. *Hereditas* 52 (2), 201–220. doi:10.1111/j.1601-5223.1964.tb01953.x

Long, E. O., and Dawid, I. B. (1980). Repeated Genes in Eukaryotes. *Annu. Rev. Biochem.* 49 (1), 727–764. doi:10.1146/annurev.bi.49.070180.003455

Lujan, N. K., Armbruster, J. W., Lovejoy, N. R., and López-Fernández, 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

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

Mariotto, S., Centofante, L., Vicari, M., Artoni, R., 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 (4), 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. Chromosomic Characterization of Four Populations and Occurence of Sexual Chromosomes of Type XX/XY, in the Pantanal basin of Mato Grosso, Brazil. *Caryologia* 59 (4), 299–304. doi:10.1080/00087114.2006.10797929

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

Montefalcone, G., Tempesta, S., Rocchi, M., and Archidiacono, N. (1999). Centromere Repositioning. *Genome Res.* 9 (12), 1184–1188. doi:10.1101/gr.9.12.1184

Nagamachi, C. Y., Pieczarka, J. C., Milhomem, S. S., O'Brien, P. C., de Souza, A. C., 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, 28. doi:10.1186/1471-2156-11-28

Pansonato-Alves, J. C., Serrano, É. A., Utsunomia, R., Scacchetti, P. C., Oliveira, C., and Foresti, F. (2013). Mapping Five Repetitive DNA Classes in Sympatric Species of *Hypostomus* (Teleostei: Siluriformes: Loricariidae): Analysis of Chromosomal Variability. *Rev. Fish. Biol. Fish.* 23 (4), 477–489. doi:10.1007/s11160-013-9303-0

Pety, A. M., Cardoso, A. L., Nagamachi, C. Y., Pieczarka, J. C., de Sousa, L. M., and Noronha, R. C. R. (2018). *In Situ* localization of Ribosomal Sites in *Peckoltia* and *Ancistomus* (Loricariidae: Hypostominae) from the Amazon Basin. *Zebrafish* 15 (3), 263–269. doi:10.1089/zeb.2017.1523

Prizon, A. C., Borin-Carvalho, L. A., Bruschi, D. P., Otávio Ribeiro, M., Magrinelli Barbosa, L., de Brito 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. *Ccg* 10 (4), 625–636. doi:10.3897/CompCytogen.v10i4.8532

Ribeiro, M. O., Noleto, R. B., Lorscheider, C. A., Porto, F. E., Prizon, A. C., Zawadzki, C. H., et al. (2015). Cytogenetic Description of Ancistrus Abilhoai (Siluriformes: Loricariidae) from Iguaçu River basin, Southern Brazil. *Genet. Mol. Res.* 14 (2), 4051–4057. doi:10.4238/2015.april.27.20

Rocchi, M., Archidiacono, N., Schempp, W., Capozzi, O., and Stanyon, R. (2012). Centromere Repositioning in Mammals. *Heredity* 108 (1), 59–67. doi:10.1038/hdy.2011.101

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 (2), 253–265. doi:10.1016/j.catena.2006.08.009

Roxo, F. F., Ochoa, L. E., Sabaj, M. H., Lujan, N. K., Covain, R., Silva, G. S. C., et al. (2019). Phylogenomic Reappraisal of the Neotropical Catfish Family Loricariidae (Teleostei: Siluriformes) Using Ultraconserved Elements. *Mol. Phylogenet. Evol.* 135, 148–165. doi:10.1016/j.ympev.2019.02.017

Scacchetti, P. C., Utsunomia, R., Pansonato-Alves, J. C., da Costa Silva, G. J., Vicari, M. R., Artoni, R. F., et al. (2015). Repetitive DNA Sequences and Evolution of ZZ/ZW Sex Chromosomes in Characidium (Teleostei: Characiformes). *PLoSOne* 10 (9), e0137231. doi:10.1371/journal.pone.0137231

Schemberger, M. O., Nascimento, V. D., Coan, R., Ramos, É., Nogaroto, V., Ziemniczak, K., et al. (2019). DNA Transposon Invasion and Microsatellite Accumulation Guide W Chromosome Differentiation in a Neotropical Fish Genome. *Chromosoma* 128 (4), 547–560. doi:10.1007/s00412-019-00721-9

Suárez, P., Pinto Barroso, I. C. G., Silva, D. 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 (3), 272–279. doi:10.1089/zeb.2016.1405

Sumner, A. T. (1972). A Simple Technique for Demonstrating Centromeric Heterochromatin. *Exp. Cel Res.* 75, 304–306. doi:10.1016/0014-4827(72)90558-7

Tatumi, S. H., da Silva, L. P., Pires, E. L., Rossetti, D. F., Góes, A. M., and Munita, C. S. (2007). Datação de Sedimentos Pós-Barreiras no Norte Do Brasil: implicações paleogeográficas. *Revista Brasileira de Geociências* 38 (3), 514–524. doi:10.25249/0375-7536.2008383514524

Úbeda-Manzanaro, M., Merlo, M. A., Palazón, J. L., Cross, I., Sarasquete, C., and Rebordinos, L. (2010). Chromosomal Mapping of the Major and Minor Ribosomal Genes, (GATA)n and U2 snRNA Gene by Double-Colour FISH in Species of the Batrachoididae Family. *Genetica* 138 (7), 787–794. doi:10.1007/s10709-010-9460-1

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 (5), 1094–1116. doi:10.1111/j.1095-8649.2010.02564.x

Yano, C. F., Bertollo, L. A. C., Rebordinos, L., Merlo, M. A., Liehr, T., Portela-Bens, S., et al. (2017). Evolutionary Dynamics of rDNAs and U2 Small Nuclear DNAs in Triportheus (Characiformes, Triportheidae): High Variability and Particular Syntenic Organization. *Zebrafish* 14 (2), 146–154. doi:10.1089/zeb.2016.1351

Ziemniczak, K., Barros, A. V., Rosa, K. O., Nogaroto, V., Almeida, M. C., Cestari, M. M., et al. (2012). Comparative Cytogenetics of Loricariidae (Actinopterygii: Siluriformes): Emphasis in Neoplecostominae and Hypoptopomatinae. *Ital. J. Zoolog.* 79 (4), 492–501. doi:10.1080/11250003.2012.676677

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Genetics

ORIGINAL RESEARCH
published: 29 October 2021
doi: 10.3389/fgene.2021.760244

Evolutionary Tracks of Chromosomal Diversification in Surgeonfishes (Acanthuridae: *Acanthurus*) Along the World's Biogeographic Domains

*Maria Aparecida Fernandes[1], Marcelo de Bello Cioffi[2], Luiz Antônio Carlos Bertollo[2], Gideão Wagner Werneck Félix da Costa[1], Clóvis Coutinho da Motta-Neto[1], Amanda Tôrres Borges[1], Rodrigo Xavier Soares[1], Allyson Santos de Souza[1], Krit Pinthong[3], Weerayuth Supiwong[4], Alongklod Tanomtong[5] and Wagner Franco Molina[1]**

[1]Department of Cell Biology and Genetics, Biosciences Center, Federal University of Rio Grande do Norte, Natal, Brazil, [2]Department of Genetics and Evolution, Federal University of São Carlos, São Carlos, Brazil, [3]Department of Fundamental Science, Faculty of Science and Technology, Surindra Rajabhat University, Muang, Thailand, [4]Applied Science Program, Faculty of Interdisciplinary Studies, Khon Kaen University, Nong Khai Campus, Nong Khai, Thailand, [5]Program of Biology, Faculty of Science, Khon Kaen University, Khon Kaen, Thailand

OPEN ACCESS

Edited by:
Mariana Härter Remião, Federal University of Pelotas, Brazil

Reviewed by:
Leandro Argolo, Vale Technological Institute (ITV), Brazil
Alessio Iannucci, University of Florence, Italy

***Correspondence:**
Wagner Franco Molina
molinawf@yahoo.com.br

Received: *17 August 2021*
Accepted: *18 October 2021*
Published: *29 October 2021*

Citation:
Fernandes MA, Cioffi MdB, Bertollo LAC, Costa GWWFd, Motta-Neto CCd, Borges AT, Soares RX, Souza ASd, Pinthong K, Supiwong W, Tanomtong A and Molina WF (2021) Evolutionary Tracks of Chromosomal Diversification in Surgeonfishes (Acanthuridae: Acanthurus*) Along the World's Biogeographic Domains. Front. Genet. 12:760244. doi: 10.3389/fgene.2021.760244*

Fishes of the genus *Acanthurus* (Acanthuridae) are strongly related to reef environments, in a broad biogeographic context worldwide. Although their biological aspects are well known, cytogenetic information related to this genus remains incipient. In this study, *Acanthurus* species from populations inhabiting coastal regions of the Southwest Atlantic (SWA), South Atlantic oceanic islands (Fernando de Noronha Archipelago and Trindade Island), Greater Caribbean (GC), and Indo-Pacific Ocean (the center of the origin of the group) were analyzed to investigate their evolutionary differentiation. For this purpose, we employed conventional cytogenetic procedures and fluorescence *in situ* hybridization of 18S rDNA, 5S rDNA, and H3 and H2B-H2A histone sequences. The Atlantic species (*A. coeruleus, A. chirurgus*, and *A. bahianus*) did not show variations among them, despite their vast continental and insular distribution. In contrast, *A. coeruleus* from SWA and GC diverged from each other in the number of 18S rDNA sites, a condition likely associated with the barrier created by the outflows of the Amazonas/Orinoco rivers. The geminate species *A. tractus* had a cytogenetic profile similar to that of *A. bahianus*. However, the chromosomal macrostructures and the distribution of rDNA and hisDNA sequences revealed moderate to higher rates of diversification when *Acanthurus* species from recently colonized areas (Atlantic Ocean) were compared to *A. triostegus*, a representative species from the Indian Ocean. Our cytogenetic data covered all *Acanthurus* species from the Western Atlantic, tracked phylogenetic diversification throughout the dispersive process of the genus, and highlighted the probable diversifying role of ocean barriers in this process.

Keywords: marine fish, comparative cytogenetic, hisDNA, oceanic barrier, multigenic family

INTRODUCTION

Acanthuridae (surgeonfishes, tangs, and unicornfishes) represents a charismatic group of primarily large-bodied herbivorous fish species, which play an important ecological role in benthic communities and the resilience of coral environments in all tropical and subtropical seas (Randall, 2001; Green and Bellwood, 2009; Russ et al., 2018). The common name "surgeonfish" refers to the peculiar scalpel-like modified scales on both sides of the caudal peduncle that is used in inter- and intraspecific aggressive interactions (Randall, 2001). Due to their attractive colors and shapes, surgeonfishes dominate aquarium trade in several areas (Sadovy and Vincent, 2002; Papavlasopoulou et al., 2014).

The family comprises six genera and 85 species (Fricke et al., 2021). *Acanthurus* is the most diverse clade with 40 species, of which 90% are endemic to the Indo-Pacific, the origin and dispersion center of the group (Randall, 2001). Despite the growing set of genetic data on Acanthuridae, spanning the nuclear (Bernardi et al., 2018) and the mitochondrial (Ludt et al., 2020) genomes, the cytogenetic features of Acanthuridae are still largely incipient (<10% of species) (Arai and Inoue, 1976; Affonso et al., 2014; Fernandes et al., 2015).

Four species of *Acanthurus*: *A. coeruleus* (blue tang surgeonfish), *A. bahianus* (barber surgeonfish), *A. chirurgus* (doctorfish), and *A. tractus* (ocean surgeonfish), are found in the Western Atlantic, with an extensive distribution from the Caribbean to southern Brazil, including the island regions of Fernando de Noronha and Trindade and Martim Vaz (Rocha et al., 2002). In addition, the wide distribution of *Acanthurus* around the world offers a suitable model for investigating chromosomal specificities fixed among populations in the marine environment. Several species of *Acanthurus* are found in the Pacific and Indo-Pacific oceans, the center of origin of the genus, among which *A. triostegus* (convict surgeonfish) is one most representative species. *Acanthurus triostegus* has a remarkable ability to extend its larval stage by slowing metamorphosis (McCormick, 1999), with populations between the Indian and Pacific oceans exhibiting genetic structure (Grulois et al., 2020). Currently, the karyotype of this species is only known in representatives of the Pacific regions (Arai and Inoue, 1976; Ojima and Yamamoto, 1990); comparisons with samples from the Indian Ocean will be useful for estimating evolutionary divergences across the Indo-Pacific region. Concerning the Western Atlantic region, the freshwater outflows from the Amazonas (Brazil) and Orinoco (Venezuela) rivers delimit the Brazilian and the Greater Caribbean (GC) biogeographic provinces (Floeter et al., 2008). Although this barrier interferes with the genetic structure of some *Acanthurus* species (Rocha et al., 2002; Rocha 2003), it is not yet known whether the barrier contributes to karyotype differentiation between divided populations.

Due to the high evolutionary dynamic of multigene families, such as the ribosomal DNA genes (Gornung, 2013), with a vital role in protein synthesis, and of histones (H1, H2B-H2A, H3, and H4), acting in the structural organization of chromatin and regulation of gene expression in eukaryotes (Chioda et al., 2002), they have been largely used in evolutionary and population approaches (e.g., Motta-Neto et al., 2012; Amorim et al., 2017). The preferential participation of concerted evolution in the ribosomal DNA (Gonzalez and Sylvester, 2001) or by birth-and-death evolution in histone genes (Rooney and Ward, 2005) promotes differential arrays that can be followed in the genome of the species and populations (Lima-Filho et al., 2012; Costa et al., 2016).

Chromosomal variations result in the amplification of adaptive aspects of species (Kess et al., 2020). Additionally, micro or macrostructural cytogenetic patterns can be correlated with the species distribution, and action of biogeographic barriers (Motta-Neto et al., 2019), and thus associated to the diversification of Acanthuridae in the marine environment. Although mapping of repetitive DNA sequences has produced effective cytogenetic markers for detecting cryptic evolutionary diversification among Atlantic reef fishes (Costa et al., 2014; Amorim et al., 2016; Nirchio et al., 2017), little is known about the chromosomal organization in Acanthuridae. In this study, the chromosomal distribution of rDNA and histone genes among five species of *Acanthurus* from the Southwest Atlantic (SWA), Greater Caribbean (GC), and Indian Ocean (IO) were utilized to track their karyotype evolution and assess population stratifications. The results reveal a panel of increasing karyotype diversification associated with the historical biogeography of *Acanthurus* and present evidence of intraspecific variability between Atlantic areas.

MATERIALS AND METHODS

Sampling and Chromosome Preparation

Individuals of *A. coeruleus*, *A. bahianus*, *A. tractus*, and A. *chirurgus* were collected from different regions of the Western Atlantic, Brazilian Northeast coast (Rio Grande do Norte State), insular Atlantic regions (Fernando de Noronha archipelago and Trindade Island), and Florida Keys, an island archipelago in southern Florida (United States), belonging to GC province. *Acanthurus triostegus* individuals were obtained from the Andaman Sea in the Indian Ocean (**Table 1**).

All field and laboratory protocols used in this study, including specimen sampling, were approved by the Ethics Committee on the Use of Animals at the Federal University of Rio Grande do Norte (Proc.#044-15). Sample collections were authorized by the Chico Mendes Institute for Biodiversity Conservation (ICMBio), System of Authorization, and Information on Biodiversity (SISBIO-Licenses No. 19135-1, 131360-1 and 27,027-2).

The specimens were subjected to *in vivo* mitotic stimulation for 24 h using an attenuated antigen complex (Molina et al., 2010). The chromosome preparations were performed in short-term culture according to Gold et al. (1990). Silver-stained nucleolus organizer regions (Ag-NORs) and heterochromatic regions were visualized following the protocols described by Howell and Black (1980) and Sumner (1972), respectively.

TABLE 1 | Data of the species and populations of *Acanthurus* (Acanthuridae) used in cytogenetic analyses.

Species	Oceanic regions					N
	Southwest Atlantic (SWA)	Oceanic Atlantic islands		Greater Caribbean (GC)	Indian ocean	
	Brazilian coast (Rio Grande do Norte State)	Fernando de Noronha Archipelago	Trindade Island	Florida keys (Florida State, United States)	Andaman Sea	
A. coeruleus	31	01	—	08	—	40
A. bahianus	15	—	05	—	—	20
A. tractus	—	—	—	02	—	02
A. chirurgus	11	04	—	01	—	16
A. triostegus	—	—	—	—	03	03

Barcoding of Cryptic Species

Because of the cryptic identification of *A. bahianus* and the newly resurrected *A. tractus* species (Bernal and Rocha, 2011) and their sympatric occurrence in GC (Castellanos-Gell et al., 2012), genetic analyses using sequences of the cytochrome oxidase I (COI) were performed to confirm their taxonomic status. For this purpose, fragments of fins were removed, preserved in 95% ethanol, and stored at 4 °C. The total DNA of each specimen was extracted (Sambrook et al., 1989) and amplified by polymerase chain reaction (PCR) using primers for the COI gene. The PCR reactions consisted of 1 μL of total DNA, 0.5 U Taq polymerase, 0.4 μL of 50 mM $MgCl_2$, 1 μL of 10 x buffer, 0.5 μL 10 mM dNTP, 0.3 μL of each primer (10 μM) (L1987 and H2609) (Palumbi, 1991), and ultrapure water until the final volume of 25 μL. The amplification reactions were performed with an initial denaturation cycle at 95°C for 5 min; followed by 30 cycles at 94°C for 30 s, 49°C for 30 s, 72°C for 55 s, and a final extension at 72°C for 5 min. The PCR products were purified using the enzyme ExoSAP-IT (Applied Biosystems, Waltham, Massachusetts, EUA) and sequenced by ACTGene Ltd. The COI gene sequences of the individuals were compared to GenBank (www.ncbi.nlm.nih.gov/genbank/) and BOLD (www.boldsystems.org) databases employing BOLD identification tools and Blastn Search Tool, respectively, with those of *A. tractus*, confirming its taxonomic status.

Probes Preparation

The 5S rDNA (200 bp) and 18S rDNA (1,400 bp) probes were obtained by PCR from the nuclear DNA of *A. coeruleus* individuals from Northeast Brazilian coast using primers A 5′-TAC GCC CGA TCT CGT CCG ATC-3′ and B 5′ -CAG GCT GGT ATG GCC GTA AGC-3' (Pendás et al., 1994) and NS1 5′-GTA GTC ATA TGC TTG TCT C-3′ and NS8 5′-TCC GCA GGT TCA CCT ACG GA-3′ (White et al., 1990), respectively. The 5S rDNA and 18S rDNA probes were labeled with biotin-14-dATP and digoxigenin-11-dUTP, respectively, using nick translation according to the manufacturer's recommendations (Roche, Mannheim, Germany). Meanwhile, primers H2BAD 5′-CCC -CCC GAG ATG TGA TGG TAG A-3 ′ and H2BAR 5′-AGT ACA GCC TGG ATG TTT GGT AA-3′ were used to amplify the H2B-H2A sequences and primers H3D 5′-ATG GCT CGT ACC AAG CAG ACV GC-3′ and H3R 5′-ATA TCC TTR GGC ATR ATR GTG AC-3′ to amplify H3 sequences. Both primer sets were designed using the gene sequences of *Mytilus edulis* (Albig et al., 2003), and the genes were amplified according to Giribert and Distel (2003). Biotin-14-dATP and digoxigenin-11-dUTP were used to label H2B-H2A hisDNA and H3 hisDNA, respectively, using nick translation according to the manufacturer's recommendations (Roche, Mannheim, Germany).

Fluorescence *in situ* Hybridization (FISH)

FISH experiments were performed according to the protocols described by Pinkel et al. (1986). Mitotic chromosomes were treated with RNAse (20 μg/ml in 2 x SSC) at 37°C for 1 h and then with pepsin (0.005% in 10 mM HCl) at 37°C for 10 min, fixed with 1% formaldehyde for 10 min, and dehydrated using a series of alcohol solutions (70,85, and 100%) for 5 min. The chromosomal preparations were incubated in 70% formamide/2 × SSC at 72°C for 5 min. The hybridization solution ((50% formamide, 2 x SSC, and 10% dextran sulfate) and the denatured probe (5 ng/μL), with a final volume of 30 μL, were deposited on the slides, and hybridization was performed for 16 h at 37°C. Post-hybridization washes were performed using 15% formamide/0.2 × SSC at 42°C for 20 min, followed by washes using 0.1 × SSC at 60°C for 15 min and 0.5%/4 × SSC Tween 20 for 5 min at 25°C. Hybridization signals were detected using rhodamine-conjugated anti-digoxigenin for 18S rDNA and H3 hisDNA probes and FITC-conjugated streptavidin (Vector, Burlingame, CA, United States) for 5S rDNA and H2B-H2A hisDNA. Chromosomes were counterstained with Vectashield antifade with DAPI (4′,6-diamidine-2-phenylindole dihydrochloride) (1.5 μg/ml; Vector Laboratories Burlingame, CA, United States).

Microscopy and Image Analyses

Chromosomal images were obtained using an Olympus BX51 epifluorescence photomicroscope (Olympus, Tokyo, Japan) coupled to an Olympus DP73 digital capture system using the cellSens software (Olympus, Tokyo, Japan). Chromosomes were classified as metacentric (m), submetacentric (sm), subtelocentric (st), and acrocentric 1) according to their arms ratio (Levan et al., 1964). The fundamental number (FN) was established considering the occurrence of two arms on the meta-, submeta-, and subtelocentric chromosomes, and only one on the acrocentric

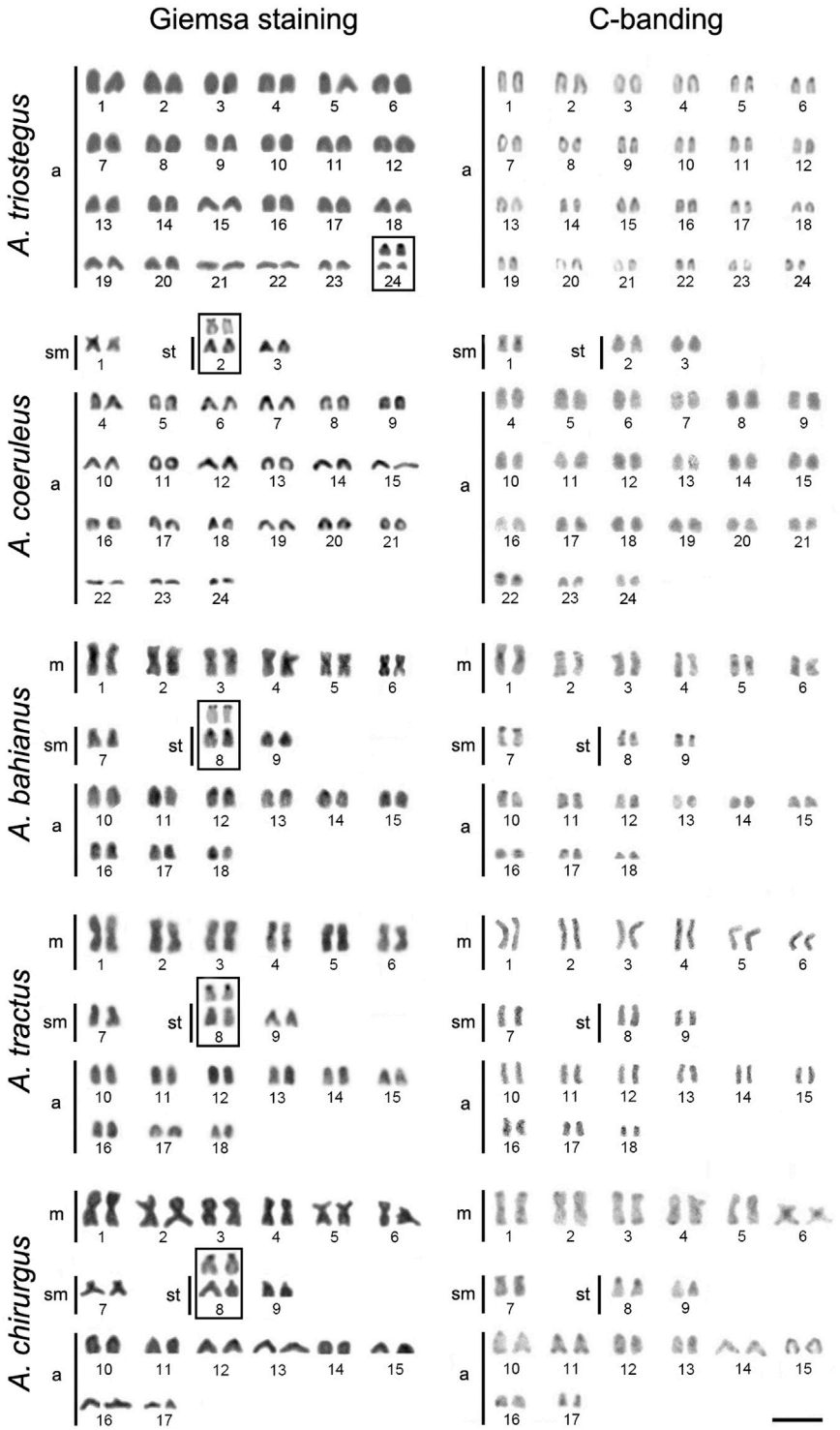


FIGURE 1 | Karyotypes of *Acanthurus triostegus*, *Acanthurus coeruleus*, *Acanthurus tractus*, *Acanthurus bahianus*, and *Acanthurus chirurgus* after Giemsa-staining and C-banding. Boxes highlight the chromosome pair with Ag-NORs (silver stained NORs). m–metacentric, sm–submetacentric, and a–acrocentric chromosomes. Scale bar = 5 μm.

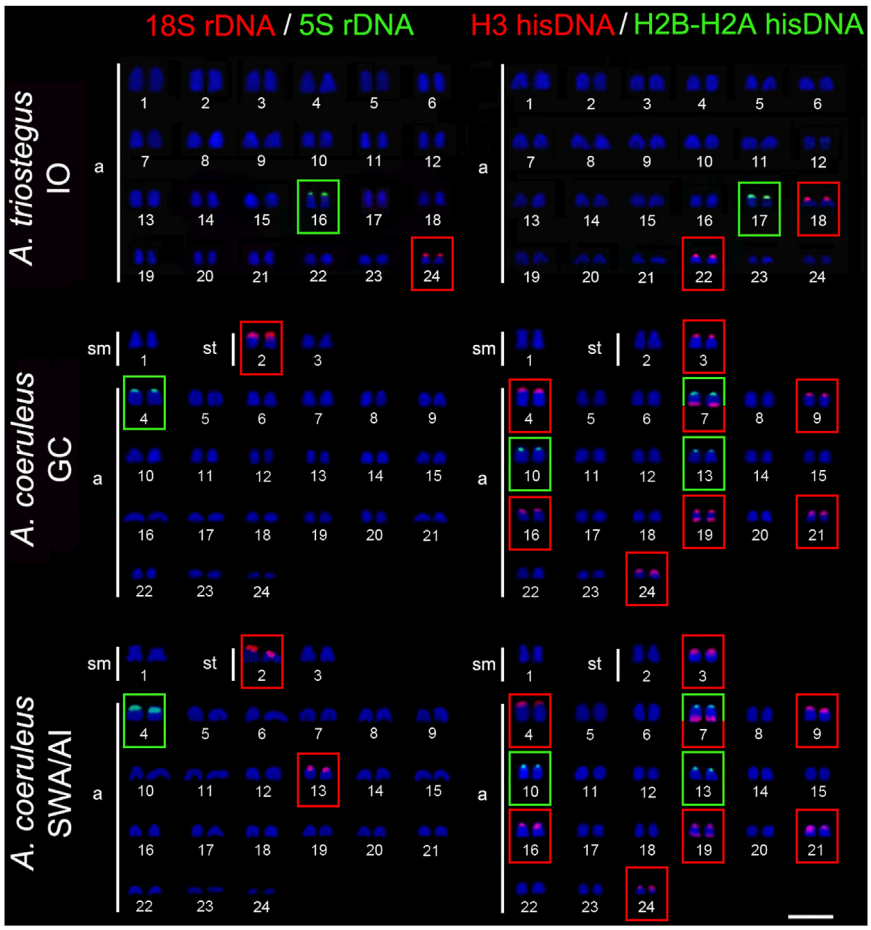


FIGURE 2 | Fluorescence *in situ* hybridization signals indicating the distribution of the 18S rDNA (red signals), 5S rDNA (green signals), H3 hisDNA (red signals) and H2B-H2A hisDNA (green signals) sites in chromosomes of *Acanthurus triostegus* and *Acanthurus coeruleus*. Scale bar = 5 μm. IO, Indian Ocean; SWA, Southwest Atlantic; AI, South Atlantic Islands; GC, Greater Caribbean.

chromosomes. Karyotypes were organized according to the decreasing order of size of the chromosomes within each of their respective morphological classes. Ideograms representing chromosomes and repetitive DNA class arrays were prepared using the Photoshop CS6 software.

RESULTS

Acanthurus triostegus from the Andaman Sea, analyzed for the first time has 2n = 48 acrocentric chromosomes (FN = 48), with Ag-NOR sites located in the short arm of pair 24. Meanwhile, *A. coeruleus* (2n = 48; 2 s + 4th + 42a; FN = 54), *A. tractus* (2n = 36; 12 m + 2 s + 4th + 18a; FN = 54), *A. bahianus* (2n = 36; 12 m + 2 s + 4th + 18a; FN = 54), and *A. chirurgus* (2n = 34; 12 m + 2 s + 4th + 16a; FN = 52) presented karyotype patterns similar to those previously reported (Affonso et al., 2014; Fernandes et al., 2015) (**Figure 1**). Different populations of *A. coeruleus*, *A. chirurgus*, and *A. bahianus* (**Table 1**) displayed similar karyotype structures when compared to each other.

The 18S rDNA sites occur exclusively in the short arm of pair 24 in *A. triostegus*, meanwhile, the sites are located in the short arms of the largest subtelocentric pair (pair 8), without inter-population variability, in *A. tractus*, *A. bahianus*, and *A. chirurgus*. In contrast, the number of 18S rDNA sites in *A. coeruleus* differs between individuals from the Brazilian coast (two loci in the short arms of pairs 2 and 13) –and from the Florida Keys (only one locus in the short arms of pair 2) (**Figures 2**, **3**).

The 5S rDNA sites are located in the short arms of pair 16 in *A. triostegus*. In the Atlantic species, *A. coeruleus*, *A. tractus*, *A. bahianus*, and *A. chirurgus*, the sites occur in the short arms of the largest acrocentric pair: pairs 4, 10, and 10 in individuals from Florida Keys (GC), Brazilian coast (SWA), and Atlantic oceanic islands, respectively (**Figures 2**, **3**).

In all species, the H2B-H2A hisDNA sites are located exclusively in the short arms of the chromosomes but with numerous divergences among them: only one locus in *A. triostegus* (pair 17), only one locus also in *A. chirurgus* (pair 11), two loci in *A. tractus* and *A. bahianus* (pairs 14 and 15), and

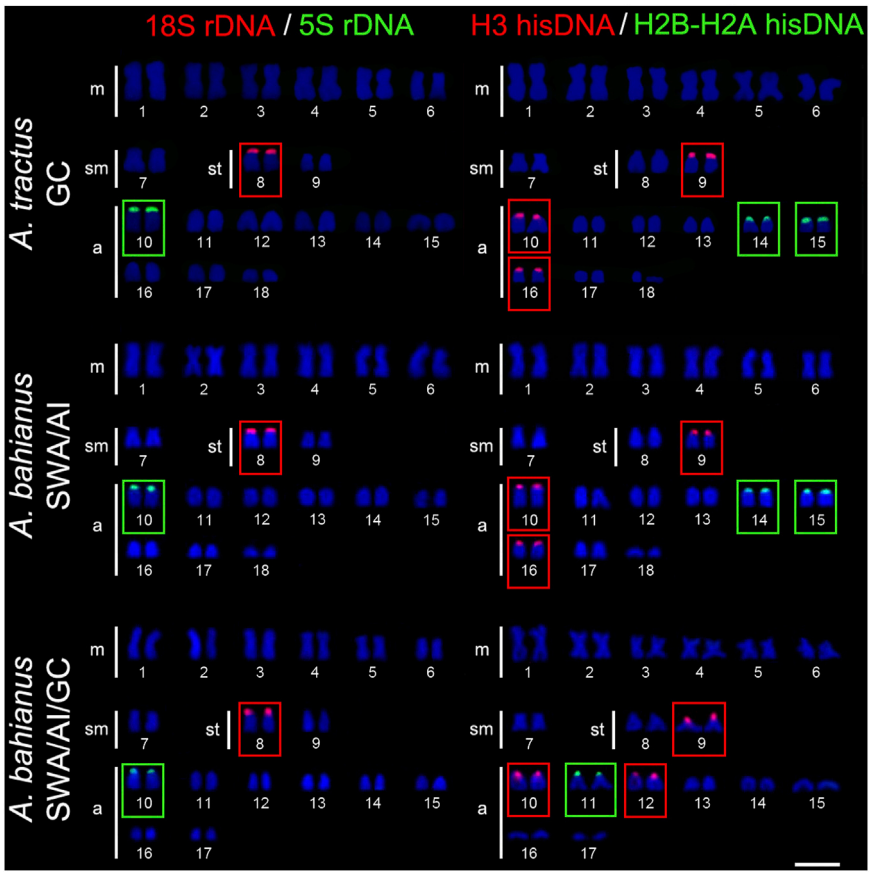


FIGURE 3 | Fluorescence *in situ* hybridization signals showing the distribution of the 18S rDNA (red signals), 5S rDNA (green signals), H3 hisDNA (red signals) and H2B-H2A hisDNA (green signals) sites in chromosomes of *Acanthurus tractus*, *Acanthurus bahianus*, and *Acanthurus chirurgus*. Scale bar = 5 μm. SWA, Southwest Atlantic; AI, South Atlantic Islands; GC, Greater Caribbean.

three loci in *A. coeruleus* (pairs 7, 10 and 13), highlighting a syntenic array with the H3 hisDNA in pair 7 (**Figures 2**, **3**). In turn, the H3 hisDNA showed notable variations in the number of loci and distribution among species. In *A. triostegus*, two loci were located in the short arms of pairs 18 and 22 (**Figure 2**). In *A. coeruleus* nine loci, the largest number among species, are distributed in the short arms of pairs 3, 4 (co-located with a 5S rDNA site), 16, 21, and 24, and in the terminal regions of the long arms of pairs 7 and 19, in this latter with a bitelomeric arrangement (**Figure 2**). Meanwhile, *A. tractus* and *A. bahianus* presented three loci: in the short arms of pairs 9, 10 (co-localized with 5S rDNA site), and in the short arms of pair 16. *Acanthurus chirurgus* also showed three loci: in the short arms of pairs 9, 10 (co-localized with 5S rDNA site) and in pair 12 (**Figure 3**).

DISCUSSION

The historical biogeography and current geographic context of *Acanthurus* species offer useful conditions for estimating the putative effects promoted by barriers and large oceanic spaces. In this context, our present data demonstrated cytogenetic variations and large-scale karyoevolutionary changes among populations and species, highlighting varied levels of divergence.

Karyotype Diversification in Western Atlantic Species of *Acanthurus*

Geographic barriers and ecological limitations can promote genetic structuring and endemism in coastal areas and oceanic islands of the Western Atlantic (Pinheiro et al., 2018). Karyotype divergences have been identified among several reef fish species divided by the Amazonas/Orinoco river plume (Nirchio et al., 2008; Rocha and Molina, 2008; Motta-Neto et al., 2019), probably promoted by divergent evolutionary forces under gene flow limitation.

Acanthurus species currently inhabiting the Western Atlantic are derived from two lineages that reached this oceanic region by different routes of colonization. Inferences based on mitochondrial and nuclear genetic sequences and fossil information suggest that one of them, reaching the region through the Tethys seaway, gave rise to *A. bahianus*, *A. tractus*, and *A. chirurgus*, at around 17.1 Mya (14.5–24.8 95% highest posterior density, HPD). The other one, coming through

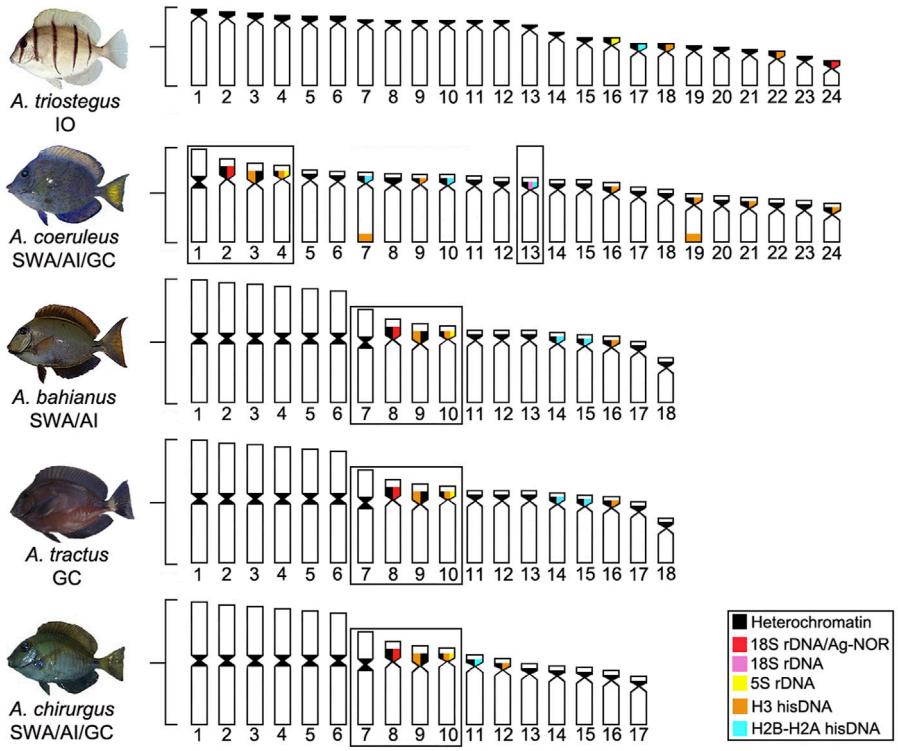


FIGURE 4 | Ideogram showing the organization of different classes of repetitive DNAs (rDNA and hisDNA) in five species and populations of *Acanthurus*. The boxes highlight a set of two-armed chromosomes shared by all Atlantic species. The dotted box highlights a chromosome pair bearing an extra 18S rDNA site in Southwest Atlantic individuals. GC, Greater Caribbean; SWA, Southwest Atlantic; IO, Indian Ocean.

the Isthmus of Panama, colonized more recently the Western Atlantic, at around 13.1 Mya (7.6–18.8 HPD) (Siqueira et al., 2019). *Acanthurus coeruleus* (2n = 48), whose Atlantic colonization is derived from the Pacific lineage, has the largest number of acrocentric chromosomes, sharing the greatest karyotype similarity with *A. triostegus* (2n = 48a), a basal *Acanthurus* species from the Indo-Pacific Ocean (Sorenson et al., 2013). In general, conspicuous series of sequential rearrangements mainly derived from pericentric inversions, centric and *in tandem* fusions, promoted the karyotype diversification in Atlantic *Acanthurus* (Affonso et al., 2014).

Given the current known biogeographic history, the evolutionary origin (homologous or homoplasic) of the set of three pairs of two-armed chromosomes in all the Atlantic *Acanthurus* species (**Figure 4**) deserves further investigation, given that reconstructed phylogenetic relationships based in molecular evidences indicate that *A. coeruleus* is phylogenetically distant from *A. bahianus*, *A. tractus*, and *A. chirurgus* (Sorenson et al., 2013). However, the common origin of this set of chromosomes by pericentric inversions in *A. tractus*, *A. bahianus*, and *A. chirurgus*, as well as of the identical set of six large metacentric pairs by Robertsonian fusions and with a similar distribution of repetitive sequences (Affonso et al., 2014; Fernandes et al., 2015; present work), indicate a synapomorphic condition fixed among these Western Atlantic species.

The evolutionary split between *A. tractus* and *A. bahianus* was preliminarily estimated at around one Mya. (Castellanos-Gell et al., 2012). Far from the origin of the Amazonas-Orinoco freshwater outflow barrier (around 10 Mya; Lovejoy et al., 1998), this recent event was not enough to fixate differential chromosomal characteristics between these two species. As these geminate species share similar karyotype patterns and have secondary contacts due to punctuated invasions of *A. bahianus* from the north to the south Atlantic regions in more recent times (Castellanos-Gell et al., 2012), they are very suitable targets for further evolutionary investigations. In fact, the sympatry between these geminate species is a scenario propitious to introgressive hybridization, a recurrent condition in *Acanthurus* species (Marie et al., 2007; DiBattista et al., 2016) and with evolutionarily consequences to Caribbean populations. On the other hand, the karyotype of *A. chirurgus* is significantly differentiated by its smaller diploid number (2n = 34) than that of *A. bahianus* and *A. tractus* (2n = 36). This autapomorphic condition is likely derived from *in tandem* fusion between a small acrocentric and the larger submetacentric chromosome pairs (Affonso et al., 2014).

Successive chromosomal differentiation in number and structure has occurred throughout the evolution of the Atlantic *Acanthurus* karyotypes; however, microstructural changes involving the repetitive fraction of the genome show diverse patterns. Repetitive sequences are important components

of genomic differentiation and evolutionary processes (López-Flores and Garrido-Ramos, 2012; Biscotti et al., 2015). Eventually, these redundant sequences are also efficient as population (Lima-Filho et al., 2012) or cytotaxonomic markers (Amorim et al., 2016), even in groups with marked chromosomal conservatism (Motta-Neto et al., 2011a; Calado et al., 2013).

Indeed, while the distribution of the rDNA sequences presents a more regular distribution pattern (Fernandes et al., 2015), our present data show that histone sequences have a more dynamic diversification among species. The histone multigene family is known to play a fundamental role in the structural organization of chromatin in eukaryotes, as well as in the regulation of gene expression (Chioda et al., 2002), showing a considerable level of organization on the chromosomes of various fish groups (Hashimoto et al., 2011; Costa et al., 2016; Borges et al., 2019). Within *Acanthurus*, the Atlantic species (*A. tractus*, *A. bahianus*, *A. chirurgus*, and mainly *A. coeruleus*) show a more variable chromosomal distribution for the H2B/H2A and H3 histones than the Indo-Pacific lineage (*A. triostegus*). In fact, among the Atlantic species, such histone sites are clustered in multiple chromosome pairs, including bitelomeric arrays in some chromosome pairs of *A. coeruleus*.

Except for a few H3 sites occupying telomeric positions in *A. coeruleus*, hisDNA sites are mainly present in centromeric heterochromatic regions. Similar patterns in other fish groups reinforce their functional co-localization with complex sets of repetitive DNAs (Hashimoto et al., 2011; Lima-Filho et al., 2012), including transposable elements (Roehrdanz et al., 2010; Costa et al., 2014, 2016). Based on provisional chromosome arrangements, our FISH signals also indicate the overlapping of H2B/H2A and H3 histones sequences with each other and with rDNA sequences, both 18S and 5S rDNA, in some heterochromatic regions. The association of distinct DNA classes highlights multiple evolutionary processes that model the repetitive DNA fraction in *Acanthurus* species. Additionally, H3 hisDNA showed more abundant sites than H2B-H2A. Despite its uncertain cause, this frequency has also been observed in other marine species, such as *Rachycentron canadum* (Rachycentridae) (Costa et al., 2014), *Ocyurus chrysurus* (Lutjanidae) (Costa et al., 2016), and *Centropomus* species (Centropomidae) (Borges et al., 2019). The number of H2B-H2A and H3 sites declined with respect to the evolutionary divergence of the Atlantic species: *A. coeruleus* with three and eight sites; *A. tractus* and *A. bahianus* with two and three sites; and *A. chirurgus* with one and three sites, respectively. Meanwhile, *A. triostegus* presents the lowest frequency: one and 2 sites. The distribution of hisDNA in these species indicates an evolution by a stochastic birth-and-death process (Novozhilov et al., 2006), with the reduction of sites along with their diversification steps.

Karyoevolution of the Genus *Acanthurus* in the Biogeographic Context

The cytogenetic profiles of *Acanthurus* species reveal evolutionary steps of chromosomal organization at the macro and microstructural levels, which are supported by historical biogeographic events and phylogenetic relationships. The occurrence of exclusive 2n = 48 acrocentric chromosomes in *A. triostegus*, a basal species with divergence contemporary to the genus *Acanthurus* (Sorenson et al., 2013), to the ancient genus *Prionurus* (*P. microlepidotus*; Arai and Inoue, 1976), and the paraphyletic *Ctenochaetus* (Ojima and Yamamoto, 1990), consolidates this karyotype as baseline for Acanthuridae in a most parsimonious manner. In addition, *A. triostegus* shares other symplesiomorphic traits with several Percomorpha groups, such as unique Ag-NORs/18S rDNA sites and heterochromatin restricted to centromeric regions (Galetti et al., 2000).

In addition to biological factors interfering with the gene flow of marine fish populations (McCormick, 1999; Otwoma et al., 2018; Grulois et al., 2020), stochastic physical events (tectonic processes, glaciations, and opening or closure of oceanic barriers) have modeled the complex biogeography of Acanthuridae (Siqueira et al., 2019) and its karyotype diversification. Cytogenetic comparisons in Atlantic *Acanthurus* highlight conspicuous synapomorphies regarding the karyotype structure and the organization of repetitive DNA classes, concerning the most basal pattern represented by the *A. triostegus* lineage.

The dispersive potential of *A. triostegus* (Fisher and Hogan, 2007; Liggins et al., 2016; Otwoma et al., 2018) favor gene flow in the marine environment (Rocha, 2003). These characteristics are possible explanations for the conservative basal karyotype of this species and the karyotype homogeneity between the Indian and Pacific populations. In contrast, *A. coeruleus*, *A. chirurgus*, and *A. tractus*/*A. bahianus*, with recent divergence in the Atlantic, and under different evolutionary and ecological changes, such as allopatry, population fragmentation, and niche displacement (Siqueira et al., 2019), show a more evident chromosomal diversification. Indeed, the divergence between the *A. coeruleus* lineage and the clade composed of *A. tractus*/*A. bahianus* and *A. chirurgus* (19 Mya), and those between the two last clades (10 Mya) (**Figure 5**), coincides with global changes in ocean dynamics (tectonic events) that affected the levels of richness and endemism of reef fish (Floeter et al., 2008). In the Atlantic Ocean, the isolation of coastal habitats during glaciations, which promoted recurrent variations at sea level, has influenced the patterns of genetic structuring (Souza et al., 2015) and, in more recent times, it may have contributed to a higher rate of chromosome evolution in reef fish.

In general, acanthurids have a very low population structure, even at great distances, such as the 3,540 km between the Brazilian region and Santa Helena Island (Rocha et al., 2002). It is believed that the high potential for dispersion (Lessios and Robertson, 2006) and colonization events of *Acanthurus* (Otwoma et al., 2018) promote genetic homogeneity and limited evolutionary dynamism among species karyotypes. In this context, the karyotype characteristics of Atlantic island populations are only now being clarified. Cytogenetic comparisons between *A. chirurgus* from the Fernando de Noronha Archipelago (FNA) and *A. bahianus* from Trindade Island, and among *A. coeruleus* from FNA, the Brazilian coast, and GC indicated that, despite the great

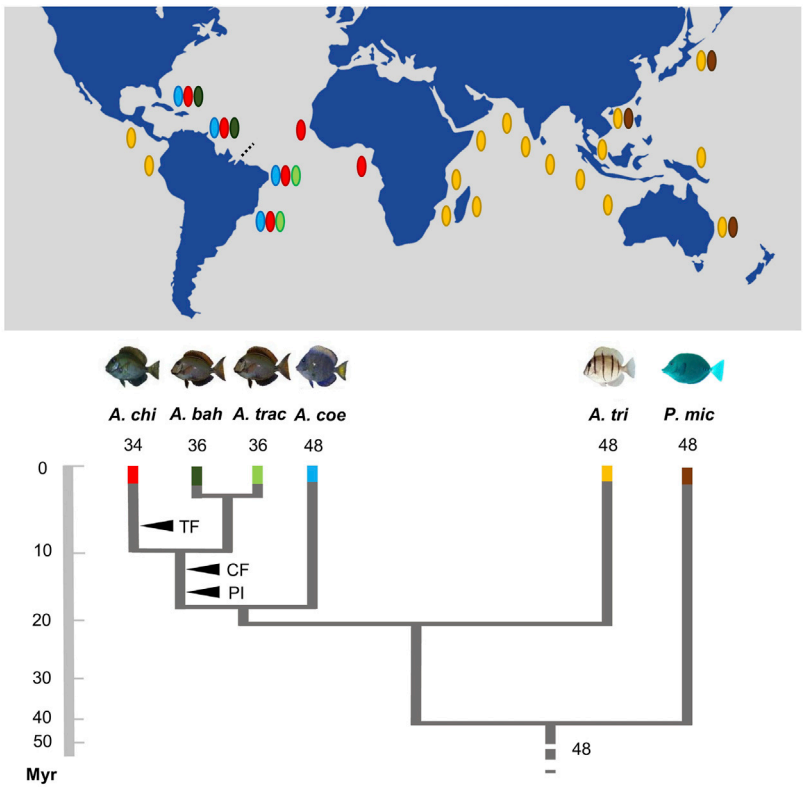


FIGURE 5 | Karyotype diversification, phylogenetic relationships, and schematic geographic distribution of the *Acanthurus* species analyzed. The black arrowheads indicate shared and exclusive chromosomal rearrangements in the Atlantic species (CF: six centric fusions events; PI: three pericentric inversions events; TF: in tandem fusion) regarding their phylogenetic diversification (adapted from Bernal and Rocha, 2011; Sorenson et al., 2013). The dotted line highlights the Amazonas River plume. The names of the species are identified by abbreviations - *Acanthurus chirurgus (A. chi); Acanthurus bahianus (A. bah); Acanthurus tractus (A. tra); Acanthurus coeruleus (A. coe); Acanthurus triostegus (A. tri); Prionurus microlepidotus (P. mic).*

oceanic distances and geographic isolation, there was no noticeable variation in their karyotypes.

In contrast, a certain level of variation occurs between specimens of *A. coeruleus* from GC and the SWA, in which there are two additional 18S rDNA sites in the latter. These additional sites correspond to non-active NORs (negative Ag-NORs) and a cytogeographic characteristic of this population. Significantly, this rDNA marker complements other phylogeographic pieces of evidence indicating population stratification between these areas (Rocha et al., 2002), as a consequence of the outflow barrier from the Amazon and Orinoco rivers. On the other hand, the geminate species, *A. bahianus* and *A. tractus*, resulting from their high sensibility to ecological effects of the Amazonas-Orinoco outows (Rocha et al., 2002), share a common cytogenetic pattern. Similar to some other reef fishes (Getlekha et al., 2018), these commonalities are derived from the recent evolutionary diversification of these groups, which impairs the fixation of chromosomal rearrangements. The sympatric occurrence of *A. bahianus* and *A. tractus* in some areas of the GC (Castellanos-Gell et al., 2012) opens a particular condition for investigating reproductive isolation in the absence of conspicuous karyotype diversification.

It is noteworthy that karyotype variability and diversification between populations and species, such as Lutjanidae (Nirchio et al., 2008; Rocha and Molina, 2008; Costa et al., 2016), Haemulidae (Motta-Neto et al., 2011b), and Grammatidae (Molina et al., 2012), have also been observed in Atlantic regions. These data increase the evidence of the primary interference of the Amazonas/Orinoco River plume in the karyotype differentiation of Atlantic reef fishes. Overall, our data highlight the use of integrating chromosomal patterns, including microstructural characters, phylogenetics, and historical biogeography, in elucidating the karyotype evolution of marine fishes, connectivity among biogeographic provinces and on the estimation of chromosomal divergences among marine populations.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by The animal study was reviewed and approved by the Committee of Ethics for the use of Animals of the Federal University of Rio Grande do Norte (Proc. #044/2015).

AUTHOR CONTRIBUTIONS

MF and WM contributed conception and design of the study; MF, GC, KP, WS and AT performed cytogenetic analyses. ASS performed sequence analyses. C-N, AB, GC, RS performed specimen's collection. MF and WM wrote the first draft of the manuscript; MC, LB, AT and WM, wrote sections of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING

This work received national funds through CNPq Brazilian National Council for Scientific and Technological Development through projects Proc. 442664/2015-0 and Proc. 442626/2019-3. This research was partly supported by INCT "Marine Sciences" (565054/2010-4).

ACKNOWLEDGMENTS

The authors thank ICMBio SISBIO for permits to collect specimens (licenses 19135-1, 27027-2 and 131360-1). We are also grateful to Jose Garcia Jr for taxonomic identification of specimens.

REFERENCES

Affonso, P. R. A. M., Fernandes, M. A., Almeida, J. S., and Molina, W. F. (2014). Sequential Steps of Chromosomal Differentiation in Atlantic Surgeonfishes: Evolutionary Inferences. *Scientific World J.* 2014, 1–7. doi:10.1155/2014/825703

Albig, W., Warthorst, U., Drabent, B., Prats, E., Cornudella, L., and Doenecke, D. (2003). *Mytilus edulis* Core Histone Genes Are Organized in Two Clusters Devoid of Linker Histone Genes. *J. Mol. Evol.* 56, 597–606. doi:10.1007/S00239-002-2428-8

Amorim, K. D. J., Cioffi, M. B., Bertollo, L. A. C., and Soares, R. X. (2017). Interregional Cytogenetic Comparisons in *Halichoeres* and *Thalassoma* Wrasses (Labridae) of Coastal and Insular Regions of the Southwestern Atlantic. *Genet. Mol. Res.* 16 (2). doi:10.4238/gmr16029650

Amorim, K. D. J., Cioffi, M. B., Soares, R. X., and Costa, G. W. W. (2016). Co-located 18S/5S rDNA Arrays: an Ancient and Unusual Chromosomal Trait in Julidini Species (Labridae, Perciformes). *Comp. Cytogenet.* 10, 555–570. doi:10.3897/CompCytogen.v10i4.10227

Arai, R., and Inoue, M. (1976). Chromosomes of Seven Species of Pomacentridae and Two Species of Acanthuridae from Japan. *Bull. Natl. Mus. Nat. Sci. A2*, 73–78.

Bernal, M. A., and Rocha, L. (2011). *Acanthurus Tractus*, Poey, 1860, a Valid Western Atlantic Species of Surgeonfish (Teleostei, Acanthuridae), Distinct from *Acanthurus Bahianus* Castelnau, 1855. *Zootaxa* 2905, 63–68. doi:10.11646/zootaxa.2905.1.5

Bernardi, G., Nelson, P., Paddack, M., Rulmal, J., Jr., and Crane, N. (2018). Genomic Islands of Divergence in the Yellow Tang and the Brushtail Tang Surgeonfishes. *Ecol. Evol.* 8, 8676–8685. doi:10.1002/ece3.4417

Biscotti, M. A., Olmo, E., and Heslop-Harrison, J. S. (2015). Repetitive DNA in Eukaryotic Genomes. *Chromosome Res.* 23, 415–420. doi:10.1007/s10577-015-9499-z

Borges, A. T., Cioffi, M. B., Bertollo, L. A. C., and Soares, R. X. (2019). Paracentric Inversions Differentiate the Conservative Karyotypes in Two *Centropomus* Species (Teleostei: Centropomidae). *Cytogenet. Genome Res.* 157, 239–248. doi:10.1159/000499748

Calado, L. L., Bertollo, L. A. C., Costa, G. W. W. F., and Molina, W. F. (2013). Cytogenetic Studies of Atlantic Mojarras (Perciformes: Gerreidae): Chromosomal Mapping of 5S and 18S Ribosomal Genes Using Double FISH. *Aquacult. Res.* 44, 829–835. doi:10.1111/j.1365-2109.2012.03089.x

Castellanos-Gell, J., Barcia, A. R., Monteagudo, P. C., and Metcalfe, C. (2012). The Surgeonfish, *Acanthurus Bahianus*, Has Crossed the Amazon–Orinoco Outflow Barrier. *Mar. Biol.* 159, 1561–1565. doi:10.1007/s00227-012-1942-5

Chioda, M., Eskeland, R., and Thompson, E. M. (2002). Histone Gene Complement, Variant Expression, and mRNA Processing in a Urochordate *Oikopleura dioica* that Undergoes Extensive Polyploidization. *Mol. Biol. Evol.* 19, 2247–2260. doi:10.1093/oxfordjournals.molbev.a004048

Costa, G. W. W. F., Cioffi, M. B., Bertollo, L. A. C., and Molina, W. F. (2016). The Evolutionary Dynamics of Ribosomal Genes, Histone H3, and Transposable Rex Elements in the Genome of atlantic Snappers. *J. Hered.* 107, 173–180. doi:10.1093/jhered/esv136

Costa, G. W. W. F., Cioffi, M. B., Bertollo, L. A. C., and Molina, W. F. (2014). Unusual Dispersion of Histone Repeats on the Whole Chromosomal Complement and Their Colocalization with Ribosomal Genes in *Rachycentron canadum* (Rachycentridae, Perciformes). *Cytogenet. Genome Res.* 144, 1–6. doi:10.1159/000366301

DiBattista, J. D., Whitney, J., Craig, M. T., Hobbs, J. A., Rocha, L. A., Feldheim, K. A., et al. (2016). Surgeons and Suture Zones: Hybridization Among Four Surgeonfish Species in the Indo-Pacific with Variable Evolutionary Outcomes. *Mol. Phylogenet. Evol.* 101, 203–215. doi:10.1016/j.ympev.2016.04.036

Fernandes, M. A., Affonso, P. R. A. M., Costa, G. W. W. F., and Cioffi, M. B. (2015). Atlantic Surgeonfishes bear Only Minor Microstructural Changes in Highly Derived Karyotypes. *Zool. Anz.* 254, 62–66. doi:10.1016/j.jcz.2014.11.003

Fisher, R., and Hogan, J. D. (2007). Morphological Predictors of Swimming Speed: a Case Study of Pre-settlement Juvenile Coral Reef Fishes. *J. Exp. Biol.* 210, 2436–2443. doi:10.1242/jeb.004275

Floeter, S. R., Rocha, L. A., Robertson, D. R., Joyeux, J. C., and Smith-Vaniz, W. F. (2008). Atlantic Reef Fish Biogeography and Evolution. *J. Biogeogr.* 35, 22–47. doi:10.1111/j.1365-2699.2007.01790.x

Fricke, R., Eschmeyer, W. N., and Fong, J. D. (2021). Eschmeyer's Catalog of Fishes: Genera/Species by Family/Subfamily. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp). Electronic version (accessed March 16, 2021).

Galetti, P. M., Jr., Aguilar, C. T., and Molina, W. F. (2000). An Overview of marine Fish Cytogenetics. *Hydrobiologia* 420, 55–62. doi:10.1023/A:1003977418900

Getlekha, N., Cioffi, M. B., Maneechot, N., Bertollo, L. A. C., Supiwong, W., Tonomtong, A., et al. (2018). Contrasting Evolutionary Paths Among Indo-Pacific *Pomacentrus* Species Promoted by Extensive Pericentric Inversions and Genome Organization of Repetitive Sequences. *Zebrafish* 15, 45–54. doi:10.1089/zeb.2017.1484

Giribert, G., and Distel, D. (2003). "Bivalve Phylogeny and Molecular Data," in *" in Molecular Systematics and Phylogeography of Mollusc*s. Editors C. Lydeard and D. R. Lindberg (Washington: Smithsonian Books), 45–90.

Gold, J. R., Li, Y. 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

Gonzalez, I. L., and Sylvester, J. E. (2001). Human rDNA: Evolutionary Patterns within the Genes and Tandem Arrays Derived from Multiple Chromosomes. *Genomics* 736, 255–263. doi:10.1006/geno.2001.6540

Green, A. L., and Bellwood, D. R. (2009). *Monitoring Functional Groups of Herbivorous Reef Fishes as Indicators of Coral Reef Resilience: A Practical Guide for Coral Reef Managers in the Asia Pacific Region. Manual. Gland.* Switzerland: The International Union for the Conservation of Nature and Natural Resources.

Grulois, D., Hogan, R. I., Paygambar, S., Planes, S., and Fauvelot, C. (2020). New Microsatellite DNA Markers to Resolve Population Structure of the Convict Surgeonfish, *Acanthurus triostegus*, and Cross-Species Amplifications on Thirteen Other Acanthuridae. *Mol. Biol. Rep.* 47, 8243–8250. doi:10.1007/s11033-020-05773-0

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

Howell, W. M., and Black, D. A. (1980). Controlled Silver-Staining of Nucleolus Organizer Regions with a Protective Colloidal a 1-step Method. *Experientia* 36, 1014–1015.

Kess, T., Bentzen, P., Lehnert, S. J., Sylvester, E. V. A., Lien, S., Kent, M. P., et al. (2020). Modular Chromosome Rearrangements Reveal Parallel and Nonparallel Adaptation in a marine Fish. *Ecol. Evol.* 10, 638–653. doi:10.1002/ece3.5828

Lessios, H. A., and Robertson, D. R. (2006). Crossing the Impassable: Genetic Connections in 20 Reef Fishes across the Eastern Pacific Barrier. *Proc. R. Soc. Lond. B Biol. Sci.* 7, 2201–2208. doi:10.1098/rspb.2006.3543

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

Liggins, L., Treml, E. A., Possingham, H. P., and Riginos, C. (2016). Seascape Features, rather Than Dispersal Traits, Predict Spatial Genetic Patterns in Co-distributed Reef Fishes. *J. Biogeogr.* 43, 256–267. doi:10.1111/jbi.12647

Lima-Filho, P. A., Cioffi, M. B., Bertollo, L. A. C., and Molina, W. F. (2012). Chromosomal and Morphological Divergences in Atlantic Populations of the Frillfin Goby *Bathygobius Soporator* (Gobiidae, Perciformes). *J. Exp. Mar. Biol. Ecol.* 434-435, 63–70. doi:10.1016/J.Jembe.2012.08.004

López-Flores, I., and Garrido-Ramos, M. A. (2012). "The Repetitive DNA Content of Eukaryotic Genomes," in *Genome Dynamics*. Editor M. A. Garrido-Ramos (Basel: Karger), 1–28.

Lovejoy, N. R., Bermingham, E., and Martin, A. P. (1998). Marine Incursion into South America. *Nature* 396, 421–422. doi:10.1038/24757

Ludt, W. B., Rocha, L. A., and Chakrabarty, P. (2020). The First Complete Mitochondrial Genomes of Sawtail Surgeonfishes (Acanthuridae: *Prionurus*). *Mitochondrial DNA B* 5, 212–213. doi:10.1080/23802359.2019.1699465

Marie, A. D., van Herwerden, L., and Choat, J. H. (2007). Hybridization of Reef Fishes at the Indo-Pacific Biogeographic Barrier: a Case Study. *Coral Reefs* 26, 841–850. doi:10.1007/s00338-007-0273-3

McCormick, M. I. (1999). Delayed Metamorphosis of a Tropical Reef Fish (*Acanthurus triostegus*): a Field experiment. *Mar. Ecol. Prog. Ser.* 76, 25–38.

Molina, W. F., Alves, D. E. O., Araujo, W. C., Martinez, P. A., Silva, M. F. M., and Costa, G. W. W. F. (2010). Performance of Human Immunostimulating Agents in the Improvement of Fish Cytogenetic Preparations. *Genet. Mol. Res.* 9, 1807–1814. doi:10.4238/Vol9-3gmr840

Molina, W. F., Costa, G. W. W. F., Cioffi, M. B., and Bertollo, L. A. C. (2012). Chromosomal Differentiation and Speciation in Sister-Species of Grammatidae (Perciformes) from the Western Atlantic. *Helgol. Mar. Res.* 66, 363–370. doi:10.1007/S10152-011-0276-X

Motta-Neto, C. C., Cioffi, M. B., Bertollo, L. A. C., and Molina, W. F. (2011b). Extensive Chromosomal Homologies and Evidence of Karyotypic Stasis in Atlantic Grunts of the Genus *Haemulon* (Perciformes). *J. Exp. Mar. Biol. Ecol.* 401, 75–79. doi:10.1016/J.Jembe.2011.02.044

Motta-Neto, C. C., Cioffi, M. B., Bertollo, L. A. C., and Molina, W. F. (2011a). Molecular Cytogenetic Analysis of Haemulidae Fish Evolutionary Conservation. *J. Exp. Mar. Biol. Ecol.* 407, 97–100. doi:10.1016/J.Jembe.2011.07.014

Motta-Neto, C. C., Cioffi, M. B., Costa, G. W. W. F., and Amorim, K. D. J. (2019). Overview on Karyotype Stasis in Atlantic Grunts (Eupercaria, Haemulidae) and the Evolutionary Extensions for Other marine Fish Groups. *Front. Mar. Sci.*, 61–12. doi:10.3389/fmars.2019.00628

Motta-Neto, C. C., Lima-Filho, P. A., Araújo, W. C., Bertollo, L. A. C., and Molina, W. F. (2012). Differentiated Evolutionary Pathways in Haemulidae (Perciformes): Karyotype Stasis versus Morphological Differentiation. *Rev. Fish. Biol. Fisher.* 22, 457–465. doi:10.1007/S11160-011-9236-4

Nirchio, M., Oliveira, C., Siccha Ramirez, Z. R., 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

Nirchio, M., Rondon, R., Oliveira, C., and Ferreira, I. A. (2008). Cytogenetic Studies in Three Species of *Lutjanus* (Perciformes: Lutjanidae: Lutjaninae) from the Isla Margarita, Venezuela. *Neotrop. Ichthyol.* 6, 101–108. doi:10.1590/S1679-62252008000100012

Ojima, Y., and Yamamoto, K. (1990). Cellular DNA Contents of Fishes Determined by Flow Cytometry. *La Kromosomo II* 57, 1871–1888.

Otwoma, L. M., Diemel, V., Reuter, H., Kochzius, M., and Meyer, A. (2018). Genetic Population Structure of the Convict Surgeonfish *Acanthurus triostegus*: a Phylogeographic Reassessment across its Range. *J. Fish. Biol.* 93, 597–608. doi:10.1111/jfb.13686

Palumbi, S. R. (1991). "Nucleic Acids II: the Polymerase Chain Reaction," in *" in Molecular Systematics*. Editors D. M. Hillis and C. Moritz (Sunderland: Mable Sinauer and Associates Inc.), 205–247.

Papavlasopoulou, I., Vardakas, L., Perdikaris, C., Kommatas, D., and Paschos, I. (2014). Ornamental Fish in Pet Stores in Greece: a Threat to Biodiversity?. *Mediterr. Mar. Sci.* 15, 126–134. doi:10.12681/mms.484

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. *Cell Genet.* 67, 31–36. doi:10.1159/000133792

Pinheiro, H. T., Rocha, L. A., and Macieira, R. M. (2018). South-western Atlantic Reef Fishes: Zoogeographical Patterns and Ecological Drivers Reveal a Secondary Biodiversity centre in the Atlantic Ocean. *Divers. Distrib* 24, 951–965. doi:10.1111/ddi.12729

Pinkel, D., Straume, T., and Gray, J. W. (1986). Cytogenetic Analysis Using Quantitative, High-Sensitivity, Fluorescence Hybridization. *P. Natl. Acad. Sci. U.S.A.* 83, 2934–2938. doi:10.1073/Pnas.83.9.2934

Randall, J. E. (2001). ""Acanthuridae. Surgeonfishes (Tangs, Unicornfishes)," in FAO Species Identification Guide for Fishery Purposes," in *The Living marine Resources of the Western Central Pacific*. Editors K. E. Carpenter and V. Niem (Rome: Bony fishes part 4 Labridae to LatimeriidaeFAO), Vol. 6, 3653–3683.

Rocha, E. C., and Molina, W. F. (2008). Cytogenetic Analysis in Western Atlantic Snappers (Perciformes, Lutjanidae). *Genet. Mol. Biol.* 31/2, 461–467. doi:10.1590/S1415-47572008000300011

Rocha, L. A., Bass, A. L., Robertson, D. R., and Bowen, B. W. (2002). Adult Habitat Preferences, Larval Dispersal, and the Comparative Phylogeography of Three Atlantic Surgeonfishes (Teleostei: Acanthuridae). *Mol. Ecol.* 11, 243–252. doi:10.1046/j.0962-1083.2001.01431.x

Rocha, L. A. (2003). Patterns of Distribution and Processes of Speciation in Brazilian Reef Fishes. *J. Biogeogr.* 30, 1161–1171. doi:10.1046/j.1365-2699.2003.00900.x

Roehrdanz, R., Heilmann, L., Senechal, P., Sears, S., and Evenson, P. (2010). Histone and Ribosomal RNA Repetitive Gene Clusters of the Boll Weevil Are Linked in a Tandem Array. *Insect Mol. Biol.* 19, 463–471. doi:10.1111/j.1365-2583.2010.01006.x

Rooney, A. P., and Ward, T. J. (2005). Evolution of a Large Ribosomal RNA Multigene Family in Filamentous Fungi: Birth and Death of a Concerted Evolution Paradigm. *PNAS* 102, 5084–5089. doi:10.1073/pnas.0409689102

Russ, G. R., Payne, C. S., Bergseth, B. J., Rizzari, J. R., Abesamis, R. A., and Alcala, A. C. (2018). Decadal-scale Response of Detritivorous Surgeonfishes (Family Acanthuridae) to No-Take marine reserve protection and Changes in Benthic Habitat. *J. Fish. Biol.* 93, 887–900. doi:10.1111/jfb.13809

Sadovy, Y. J., and Vincent, A. C. J. (2002). in *"Ecological Issues and the Trades in Live Reef Fishes," in Coral Reef Fishes*. Editor P.F. Sale (San Diego, USA: Academic Press), 391–420.

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

Siqueira, A. C., Bellwood, D. R., and Cowman, P. F. (2019). Historical Biogeography of Herbivorous Coral Reef Fishes: The Formation of an Atlantic Fauna. *J. Biogeogr.* 46, 1611–1624. doi:10.1111/jbi.13631

Sorenson, L., Santini, F., Carnevale, G., and Alfaro, M. E. (2013). A Multilocus Timetree of Surgeonfishes (Acanthuridae, Percomorpha), with Revised Family Taxonomy. *Mol. Phylogenet. Evol.* 68, 150–160. doi:10.1016/j.ympev.2013.03.014

Souza, A. S., Dias, E. A., Jr., Galetti, P. M., Jr., Garcia-Machado, E., Pichorim, M., and Molina, W. F. (2015). Wide-range Genetic Connectivity of Coney,

Cephalopholis Fulva (Epinephelidae), through Oceanic Islands and continental Brazilian Coast. *Acad. Bras. Ciênc.* 87, 121–136. doi:10.1590/0001-3765201520130411

Sumner, A. T. (1972). A Simple Technique for Demonstrating Centromeric Heterochromatin. *Exp. Cel Res.* 75, 304–306.

White, T. J., Bruns, T., Lee, S., and Taylor, J. (1990). "Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics," in *" in PCR Protocols: A Guide to Methods and Applications*. Editors M. A. Innis, D. H. Gelfand, J. J. Shinsky, and T. J. White (New York, NY: Academic Press), 315–322.

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Genetics

ORIGINAL RESEARCH
published: 24 March 2022
doi: 10.3389/fgene.2022.838462

Chromosomal Diversification in *Ancistrus* Species (Siluriformes: Loricariidae) Inferred From Repetitive Sequence Analysis

Kevin Santos da Silva[1], *Larissa Glugoski*[2,3], *Marcelo Ricardo Vicari*[3], *Augusto César Paes de Souza*[4], *Renata Coelho Rodrigues Noronha*[1], *Julio Cesar Pieczarka*[1] *and Cleusa Yoshiko Nagamachi*[1]*

[1]*Laboratório de Citogenética, Centro de Estudos Avançados da Biodiversidade, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Brazil,* [2]*Laboratório de Citogenética de Peixes, Universidade Federal de São Carlos, São Carlos, Brazil,* [3]*Laboratório de Biologia Cromossômica: Estrutura e Função, Departamento de Biologia Estrutural, Molecular e Genética, Universidade Estadual de Ponta Grossa, Ponta Grossa, Brazil,* [4]*Laboratório de Estudo da Ictiofauna Amazônica, Instituto Federal de Educação, Ciência e Tecnologia do Pará, Abaetetuba, Brazil*

OPEN ACCESS

Edited by:
Caroline Gomes Lucas, University of Missouri, United States

Reviewed by:
Vito Antonio Mastrochirico Filho, Universidade Estadual Paulista, Brazil
Camila U. Braz, University of Wisconsin-Madison, United States

***Correspondence:**
Cleusa Yoshiko Nagamachi
cleusa@ufpa.br
cleusanagamachi@gmail.com

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

Received: *17 December 2021*
Accepted: *25 February 2022*
Published: *24 March 2022*

Citation:
Santos da Silva K, Glugoski L, Vicari MR, de Souza ACP, Noronha RCR, Pieczarka JC and Nagamachi CY (2022) Chromosomal Diversification in Ancistrus Species (Siluriformes: Loricariidae) Inferred From Repetitive Sequence Analysis. Front. Genet. 13:838462. doi: 10.3389/fgene.2022.838462

The *Ancistrus* genus has extensive chromosomal diversity among species, including heteromorphic sex chromosomes occurrence. However, studies have been shown that chromosomal diversity may still be underestimated. Repetitive sequences represent a large part of eukaryotic genomes, associated with mechanisms of karyotypic diversification, including sex chromosomes evolution. This study analyzed the karyotype diversification of two *Ancistrus* species (*Ancistrus* sp. 1 and *Ancistrus* sp. 2) from the Amazon region by classical and molecular chromosomal markers. Conventional chromosome bands and fluorescence *in situ* hybridization using probes 18S and 5S rDNA, besides (CA)n, (CG)n, (GA)n, (CAC)n, (CAG)n, (CAT)n, (GAA)n, (GAC)n, (TAA)n, and (TTAGGG)n in tandem repeats were determined on the karyotypes. *Ancistrus* sp. 1 and *Ancistrus* sp. 2 presented karyotypes with 2n = 38 (20 m + 14sm+4st, XX/XY) and 2n = 34 (20 m + 14sm, without heteromorphic sex chromosomes), respectively. Robertsonian rearrangements can explain the diploid number difference. C-bands occurred in pericentromeric regions in some chromosomes, and a single 18S rDNA locus occurred in both species. The 5S rDNA showed variation in the number of loci between species karyotypes, suggesting the occurrence of unstable sites and rearrangements associated with these sequences in *Ancistrus.* The microsatellite mapping evidenced distinct patterns of organization between the two analyzed species, occurring mainly in the sex chromosomes in *Ancistrus* sp. 1, and in the centromeric and pericentromeric regions of chromosomes m/sm in *Ancistrus* sp. 2. These data shows the extensive chromosomal diversity of repetitive sequences in *Ancistrus*, which were involved in Robertsonian rearrangements and sex chromosomes differentiation.

Keywords: amazon fish, comparative cytogenetics, microsatellites, repetitive DNA, rDNA, sex chromosomes

INTRODUCTION

Loricariidae is one of the most specious neotropical fish families of the Siluriformes order, having 1,023 species (Fricke et al., 2021) distributed throughout Central and South America, from southeastern Costa Rica to northeastern Argentina (Armbruster, 2004; Armbruster, 2008; Armbruster and Lujan, 2016). Loricariidae is a monophyletic group organized into six subfamilies: Lithogeninae, Delturinae, Hypoptopomatinae, Neoplecostominae, Loricariinae, and Hypostominae (Roxo et al., 2019). The subfamily Hypostominae is the most speciose and widely distributed, organized into nine clades and tribes (Lujan et al., 2015).

The genus *Ancistrus* Kner, 1854 (Hypostominae, Ancistrini) occurs from Panama to Argentina, presenting 65 valid species, in addition to distinct lineages not formally identified in the scientific literature due to its taxonomic complexity (Ferraris, 2007; Armbruster, 2008; Prizon et al., 2018; Borba et al., 2019; Fricke et al., 2021). From a chromosomal point of view, *Ancistrus* represents one of the most diverse lineages of Loricariidae, emphasizing their extensive variation in the diploid number (2n = 34 to 54, **Supplementary Table S1**). The vast majority of *Ancistrus* species have karyotypes with $2n \leq 52$, probably due to the result of Robertsonian (Rb) fusions (see Glugoski et al., 2020), whose occurrence has been evidenced through the *in situ* localization of repetitive sequences (Barros et al., 2017). The presence of sex chromosomes is another striking feature of chromosomal diversity in *Ancistrus*, with species showing simple systems (XX/XY, XX/X0 and ZZ/ZW; Mariotto et al., 2004; Alves et al., 2006; Mariotto and Miyazawa, 2006; de Oliveira et al., 2007), multiple (XX/XY_1Y_2 and $Z_1Z_1Z_2Z_2/Z_1Z_2W_1W_2$; de Oliveira et al., 2008) or absence of differentiated sex chromosomes (**Supplementary Table S1**).

Sex chromosomes have emerged independently in different fish lineages, evolving through alternative mechanisms and showing various degrees of heteromorphism, even among closely related species (Charlesworth et al., 2005; Cioffi et al., 2010; Henning et al., 2011; Cioffi et al., 2012; Cioffi et al., 2013). Sex chromosomes in *Ancistrus* have been evidenced by size heteromorphism and accumulation of heterochromatic regions (Mariotto et al., 2004; Alves et al., 2006; Mariotto and Miyazawa, 2006; de Oliveira et al., 2007; de Oliveira et al., 2008; de Oliveira et al., 2009). Recently, *in situ* localization of repetitive sequences have provided insights into the differentiation of these chromosomes in several groups of fish, including *Ancistrus* (Cioffi and Bertollo, 2010; Schemberger et al., 2014; Cioffi et al., 2014; Favarato et al., 2017; Prizon et al., 2018; Schemberger et al., 2019).

Repetitive sequences represent the largest portion of eukaryotic genomes and may be organized in tandem repeats (e.g., microsatellites and multigene families, like ribosomal DNAs) or dispersed (e.g., transposons and retrotransposons). Ribosomal DNAs (rDNA) are represented by two gene families: 45S ribosomal RNA genes (18S, 5.8S and 28S genes), and 5S ribosomal RNA genes (Long and Dawid, 1980). The mapping of these sequences has shown intense variation in the location and number of chromosomal sites in Loricariidae. The participation of repetitive sequences, including rDNA and sequences microsatellites in chromosomal rearrangements has been evidenced, showing the importance of these markers in comparative analyzes (Pansonato-Alves et al., 2013; Barros et al., 2017; Primo et al., 2017; Bueno et al., 2018; Pety et al., 2018; Glugoski et al., 2020; Santos da Silva et al., 2021). Simple short tandem repeats of generally 1-6 nucleotides, known as,microsatellites, constitute another important tandem repeats group (Martins, 2007). The microsatellites are abundant in eukaryotic genomes, commonly a heterochromatin component, but they could also be found in euchromatic regions (Martins, 2007; Cioffi et al., 2010; Santos da Silva et al., 2021). They are helpful cytogenetic markers to demonstrate minor chromosomal variations into related species groups, including birds, amphibians, and fish (Cioffi et al., 2010; de Oliveira et al., 2017; Da Silva et al., 2021). Furthermore, the participation of these sequences in breakpoints regions for chromosomal rearrangements and evolution of sex chromosomes has been shown in many species, including *Ancistrus* (Farré et al., 2011; Cioffi et al., 2017; Favarato et al., 2017; Prizon et al., 2017).

In the present work, we studied the karyotypes of two not formally described *Ancistrus* species (*Ancistrus* sp. 1 and *Ancistrus* sp. 2) from the Amazon region, using different groups of repetitive sequences, in order to understand their mechanisms of chromosomal diversification.

MATERIALS AND METHODS

Samples

Samples of two species of *Ancistrus* (*Ancistrus* sp. 1 and *Ancistrus* sp. 2) were analysed in this study. These species are morphologically different but are not yet described in the scientific literature. The samples were collected in distinct locations of the Tocantins-Araguaia River basin, in the Brazilian Amazon (**Figure 1**). Details about the collection points, number of individuals, and sex are presented in **Table 1**. The collection permit (number 13248) was issued by the Chico Mendes Institute for Biodiversity Conservation, Brazil. The Cytogenetics Laboratory of the Federal University of Pará had licenses for transport (number 19/2003) and the use of animals for this research (52/2003) as granted by the Ministry of the Environment. This study was approved by the Animal Ethics Committee of the Federal University of Pará (permission 68/2015). The specimens analyzed in this study were deposited in the Ichthyology Collection of the Center for Advanced Studies in Biodiversity (CEABIO/UFPA), Belém, Pará, Brazil.

Chromosomal Analysis

Mitotic chromosomes were obtained from anterior kidney cells after *in vivo* treatment with colchicine solution (0.025%), as described by Bertollo et al. (1978). The animals were anesthetized with a eugenol solution (185 mg/L) (Vidal et al., 2008) and then sacrificed. Chromosomes were analyzed by

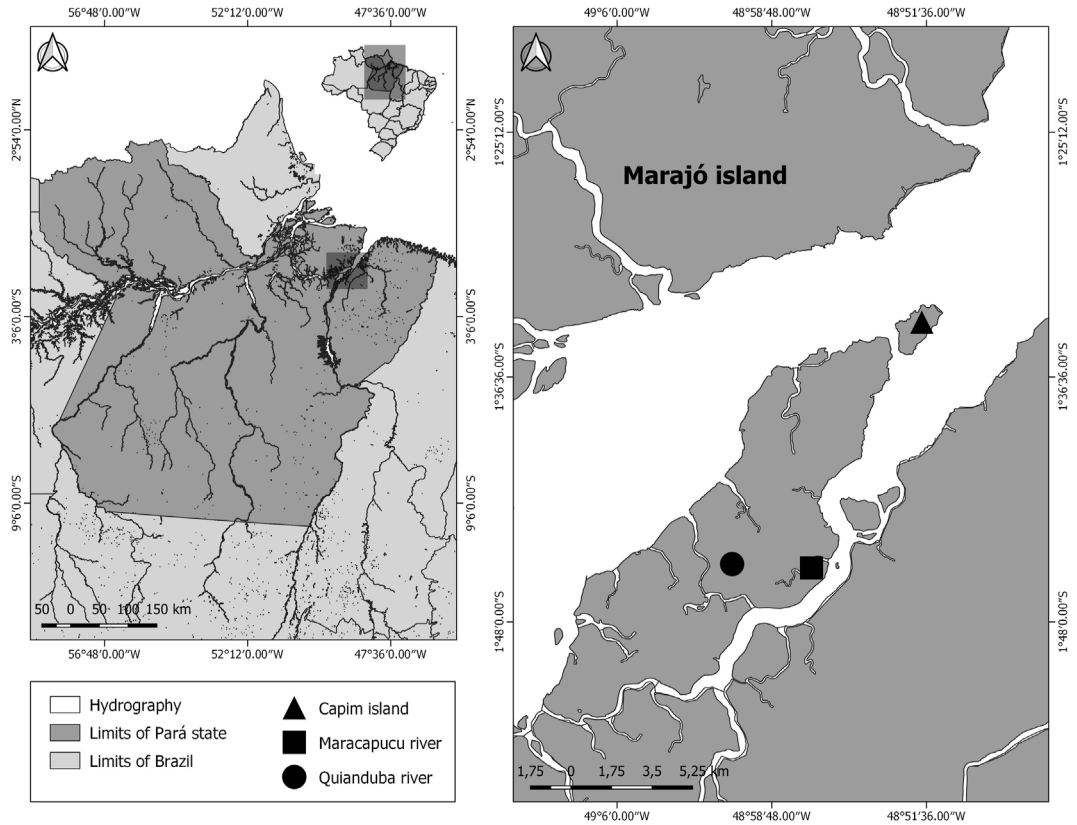


FIGURE 1 | The geographical location of the collection points of specimens of Ancistrus species analyzed in this study. The map was built using Q-GIS version 3.4.5. The database was obtained from Instituto Brasileiro de Geografia e Estatística (IBGE).

TABLE 1 | Sampling and collection sites of *Ancistrus* specimens analyzed in this study.

Species	Sex		River	Locality	Voucher	Coordinates	
Ancistrus sp. 1	5♂	2♀	A	Abaetetuba/PA	P4029	S01°45′18,2″	W49°00′38,8″
Ancistrus sp. 2	8♂	1♀	B	Abaetetuba/PA	P4263	S01°45′29,2″	W48°56′57″
	–♂	1♀	C	Abaetetuba/PA	P4251	S01°34′02,8″	W48°51′49,1″

Rivers: A—Quianduba River; B—Maracapucú River; C—Ilha do Capim; (–)—no samples

classical staining (conventional staining with 5% Giemsa), C-banding (Sumner, 1972), and also molecular methods (Fluorescence *in situ* hybridization, FISH).

Probes Labeling and Fluorescence *in situ* Hybridization

Genomic DNA was extracted using the PureLink Genomic DNA Kit (Promega), following the manufacturer's instructions. We used two rDNA sequences for *in situ* localization experiments: an 18S rDNA probe (1,400 bp segment) isolated from *Ancistrus* sp. 1 genomic DNA according described in (Gross et al., 2010), and a 5S rDNA (GenBank accession no. MT018470) probe obtained from *Ancistrus aguaboensis* (Glugoski et al., 2020). rDNA probes were labeled by nick-translation with biotin or digoxigenin. Telomeric probes were PCR labeled with digoxigeninin-11-dUTP (Roche Applid Science®) using primers F-5′(TTAGGG) 5-3′ and R-5′(CCCTAA)5-3′ without using template DNA (Ijdo et al., 1991). All PCR products were checked on 1% agarose gel electrophoresis. The nine microsatellite probes ((CA)n, (CAC)n, (CAG)n, (CG)n, (GA)n, (CAT)n, (GAA)n, (GAC)n and (TAA)n) were purchased already with direct labeling by Cy3 during synthesis. FISH was performed following the protocol proposed by Martins and Galetti (1999), with modifications, under the following stringency conditions: 2.5 ng/μL of each probe, 50% formamide, 2 x SSC, 10% dextran sulfate, and hybridization at 42°C for 16 h. Fluorescence signals were detected using Streptavidin Alexa Fluor 488 (Molecular

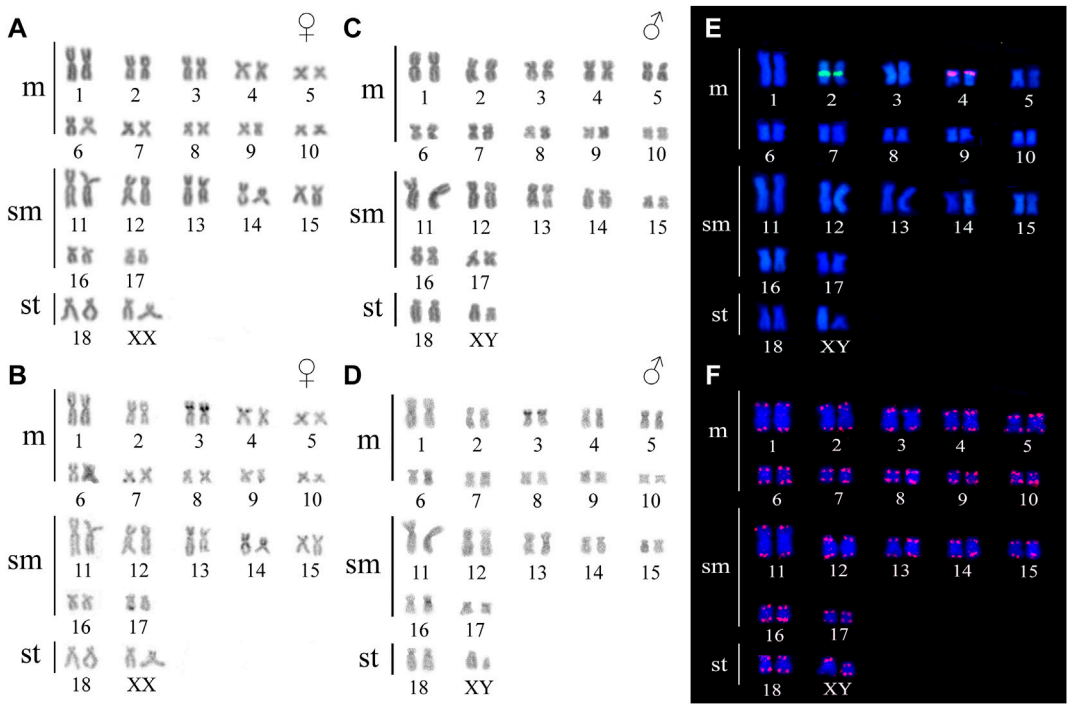


FIGURE 2 | Karyotype of Ancistrus sp. 1. In **(A,B)** male karyotype stained with Giemsa and C-banding, respectively; in **(C,D)** female karyotype stained with Giemsa and C-banding, respectively; in **(E)** *in situ* localization of 18S rDNA (green) and 5S rDNA (red) sequences; in **(F)** *in situ* localization of telomeric sequences. The probes were labeled with FITC (green) and Cy3 (red), the chromosomes were counterstained with DAPI (blue).

Probes, Carlsbad, CA, United States) and anti-digoxigenin rhodamine Fab fragments (Roche Applied Science, Penzberg, Germany). Chromosomes were counterstained with 0.2 µg/ml 4′6-diamidino-2-phenylindole (DAPI) in Vectashield mounting medium (Vector, Burlingame, CA, United States).

Image Capture and Analysis

Thirty metaphases per individual were analyzed to determine the diploid number, karyotype formula, and FISH experiments. Images of metaphases after Giemsa staining were obtained using an Olympus BX41 microscope (bright field) coupled to a CCD 1300QDS digital camera and analyzed using GenASIs ASI (Applied Spectral Imaging) software. FISH images were obtained using a Nikon H550S microscope and analyzed using Nis-Elements software. All images were adjusted using Adobe Photoshop CS6 software. The chromosome pairs were classified as metacentric (m), submetacentric (sm) and subtelocentric (st) following the criteria proposed by Levan et al. (1964). The count of the number of chromosome arms (Fundamental Number - FN) considered chromosomes m, sm and st as bi-armed.

RESULTS

Classical Cytogenetics

Ancistrus sp. 1 demonstrated diploid chromosome number (2n), fundamental number (FN) and karyotype formula (KF) as follow: 2n = 38, FN = 72, KF = 20m + 14sm+2st. A heteromorphic chromosome pair was identified in males while the female karyotypes were homomorphic, characterizing a XX/XY sex chromosome system. The small subtelocentric Y chromosome and a medium X subtelocentric chromosome were recorded (**Figures 2A,C**). Constitutive heterochromatin (CH) occurred in a few regions in the karyotype, and was not evidenced in the sex chromosomes (**Figures 2B,D**).

Ancistrus sp. 2 demonstrated diploid chromosome number (2n), fundamental number (FN) and karyotype formula (KF) as follow: 2n = 34, FN = 68, KF = 20m + 14sm. Additionally no morphologically differentiated of sex chromosomes was found (**Figure 3A**). Some heterochromatic blocks are mainly distributed in the centromeric and pericentromeric region of chromosomes 1, 2, 3, 4, 5, 6, 7, 8, 13, and 16, besides conspicuous blocks in the short arm (p) distal region of the chromosome 3, coincident with the 18S rDNA sites (**Figures 3B,C**).

Molecular Cytogenetics

In *Ancistrus* sp. 1, the 18S rDNA site was located in the proximal region of the long arm (q) of pair 2, while the 5S rDNA occurred in the interstitial region of 4p (**Figure 2E**). In *Ancistrus* sp. 2, the 18S rDNA was located in the distal region of 3p (**Figure 3C**). The 5S rDNA demonstrated multiple sites located in the proximal region of the 4p and 6p, and pericentromeric in the 7p and 16q (**Figure 3D**). Interstitial telomeric sites (ITS) were not observed in any analyzed karyotypes (**Figure 2F**; **Figure 3E**).

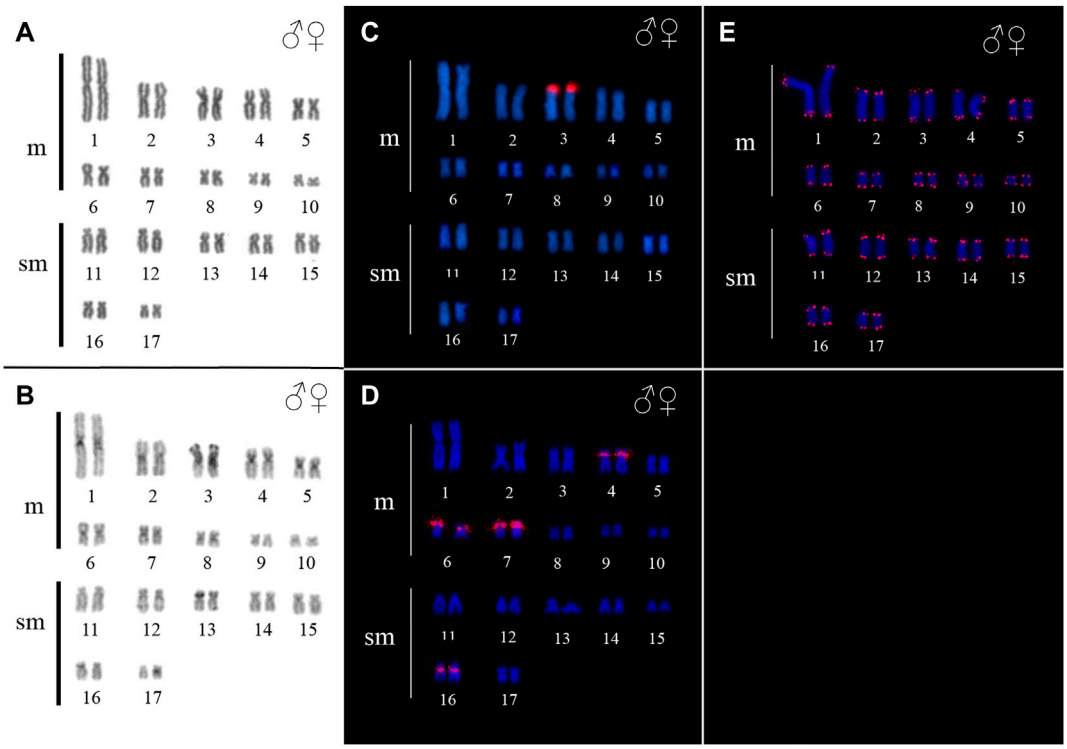


FIGURE 3 | Karyotype of Ancistrus sp. 2. In **(A,B)** karyotype stained with Giemsa and C-banding, respectively; in **(C–E)** *in situ* location of 18S rDNA, 5S rDNA and telomeric sequences, respectively. The probes were labeled with Cy3 (red), the chromosomes were counterstained with DAPI (blue).

TABLE 2 | Comparative analysis of the presence/absence and location of microsatellite sequences between the karyotypes of the *Ancistrus* species analyzed in this study.

Microsatellites	Ancistrus sp. 1				Ancistrus sp. 2	
	Autosomes		Sex chromosomes		Autosomes	
	♂	♀	Chr. X	Chr. Y	♂	♀
(CA)	–	–	–	–	+	+
(CG)	+	+	–	+	+	+
(GA)	–	–	–	–	+	+
(CAC)	–	–	–	–	+	+
(CAG)	–	–	–	–	–	–
(CAT)	–	–	+	–	+	+
(GAA)	–	–	+	–	+	+
(GAC)	–	–	+	–	+	+
(TAA)	–	–	+	–	+	+

(+)—presence of hybridization signal; (–)—no hybridization signal.

The FISH results of microsatellite sequences in the karyotypes of the two *Ancistrus* species were summarized in **Table 2**. In *Ancistrus* sp. 1, the microsatellites (CAT)n, (GAA)n, (GAC)n and (TAA)n were located exclusively in the pericentromeric region of the X chromosome; (CG)n showed signs in the pericentromeric region of the Y chromosome and in the centromeric region of the pair 13 in individuals of both sexes (**Figures 4A–E**). Microsatellites (CA)n, (GA)n, (CAC)n and (CAG)n did not demonstrate any hybridization signal in the karyotype of these species (**Table 2**). In *Ancistrus* sp. 2, the same sets of microsatellites were observed in other pairs of chromosomes, mainly in centromeric and pericentromeric regions associated with heterochromatic regions (**Figures 5A–H**, **Figure 6**). Some microsatellites, including (CAC)n, (CAT)n, (CG)n, (GAA)n, (GAC)n, and (TAA)n was associated with heterochromatic regions and coincident or adjacent to 5S rDNA sites in the pairs 4, 6 and 7 (**Figure 6**). The microsatellite (CAG)n did not show any hybridization signal in both karyotypes (**Table 2**).

DISCUSSION

Chromosomal Diversity in the Genus *Ancistrus*

The putative ancestral karyotype for Loricariidae presents 2n = 54, single nucleolus organizer regions (NOR) and few heterochromatic regions (Artoni and Bertollo, 2001; Ziemniczak et al., 2012). Nevertheless, considering the presence of 2n = 52 chromosomes in Pterygoplichthyini, the sister group for Ancistrini, Bueno et al. (2018) suggested that the putatively ancestral condition for Ancistrini is a diploid number of 52 chromosomes. Previously studies in *Ancistrus* have shown extensive chromosomal diversity with high variation in 2n values, the occurrence of multiple sites, as well as, 18S/5S rDNA synteny rupture (**Supplementary Table S1**). In this genus, species with

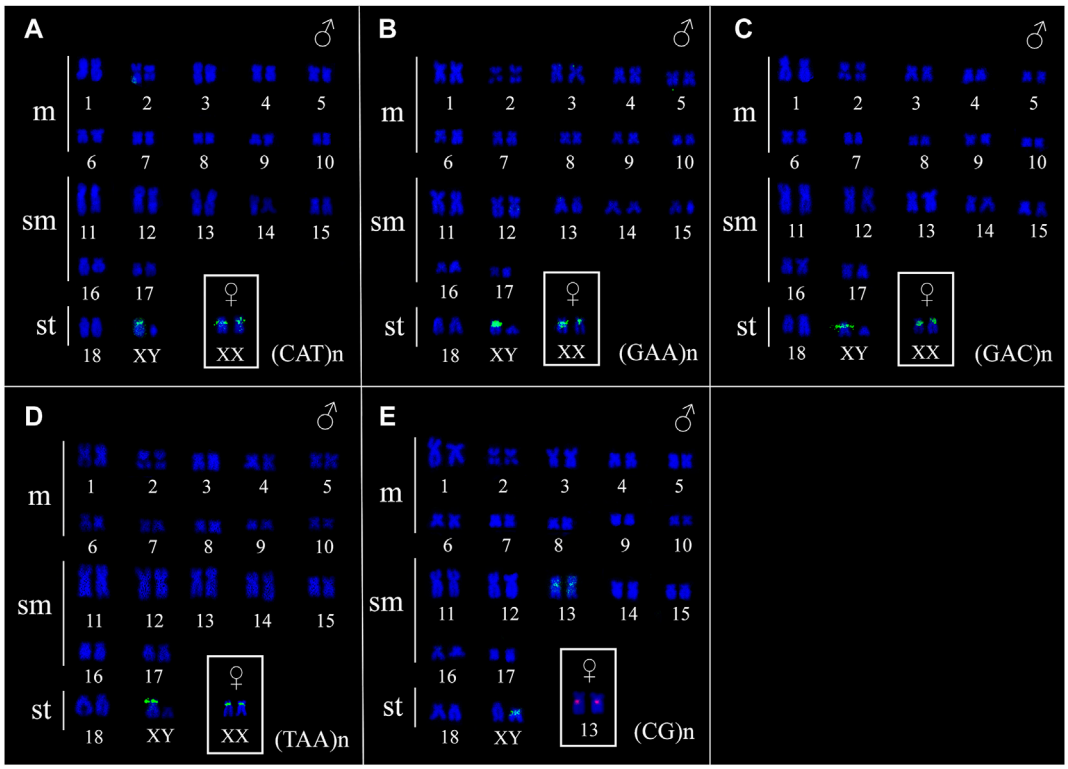


FIGURE 4 | Fluorescence *in situ* hybridization indicating the physical location of microsatellite sequences in the karyotype of Ancistrus sp. 1: **(A)** (CAT)n, **(B)** (GAA)n, **(C)** (GAC)n, **(D)** (TAA)n, **(E)** (CG)n. Karyotypes of males are presented; female sex chromosomes are in the boxes. The probes were labeled with FITC (green) and with Cy3 (red), chromosomes were counterstained with DAPI (blue).

lower 2n karyotypes have many chromosomes m/sm compared to those with higher 2n, which have more st/a chromosomes (Bueno et al., 2018; Glugoski et al., 2020). Thus, these findings suggests that chromosomal evolutions in this genus follows a tendency of 2n reduction due to the occurrence of Rb fusions (Barros et al., 2017).

Telomeric sequence mapping has shown the occurrence of ITS as a result of fusion events in Loricariidae (Barros et al., 2017; Primo et al., 2017). The species analyzed in this present study did not show the presence of ITS, which is in agreement with previous studies performed with other *Ancistrus* species (Prizon et al., 2018; Glugoski et al., 2020). The absence of ITS can be explained by the loss of these sequences during the fusion process (Slijepcevic, 1998).

Ribosomal genes have shown great diversity of location and number of sites among species and populations of different groups of fish (Gornung, 2013; Rebordinos et al., 2013). Mapping these sequences has revealed widely diverse chromosomal organizations in *Ancistrus* (Mariotto et al., 2011; Barros et al., 2017; Prizon et al., 2017; Prizon et al., 2018; Bueno et al., 2018; Glugoski et al., 2020). In Loricariidae, synteny between 18S/5S rDNA is considered a plesiomorphic character, with synteny break representing a derived state commonly observed in this group of fish (Bueno et al., 2018). This diversity of location and number of rDNA sites suggests the recurrent participation of these sequences in chromosomal reorganization events in Loricariidae (Pansonato-Alves et al., 2013; Prizon et al., 2016; Prizon et al., 2017; Barros et al., 2017; Prizon et al., 2018; Bueno et al., 2018; Santos da Silva et al., 2021).

Single 45S rDNA site represents a primitive character in Loricariidae (Artoni and Bertollo, 2001; Bueno et al., 2018). In the species described here and most of the analyzed *Ancistrus* species, these sequences are located in only one pair of chromosomes (see **Supplementary Table S1**), suggesting the maintenance of the primitive condition. Variations in the chromosomal position of these sequences were observed among the species in the present study, probably due to pericentric inversions.

Our results showed that the 5S rDNA was more dynamic than the 18S rDNA, varying in the number of sites among the analyzed karyotypes. This observation was consistent with previous observations performed in other *Ancistrus* species (Barros et al., 2017; Prizon et al., 2017; Glugoski et al., 2020; **Supplementary Table S1**). Studies have been showing that these sequences are involved in double-stranded DNA breaks and chromosomal rearrangements in *Ancistrus* (Mariotto et al., 2011; Favarato et al., 2016; Barros et al., 2017; Glugoski et al., 2020). Barros et al. (2017) demonstrated that the occurrence of multiple 5S rDNA sites is related to the emergence of pseudogenes in *Ancistrus* sp. (2n = 50). In addition, the involvement of 5S rRNA pseudogenes in Rb fusion events has

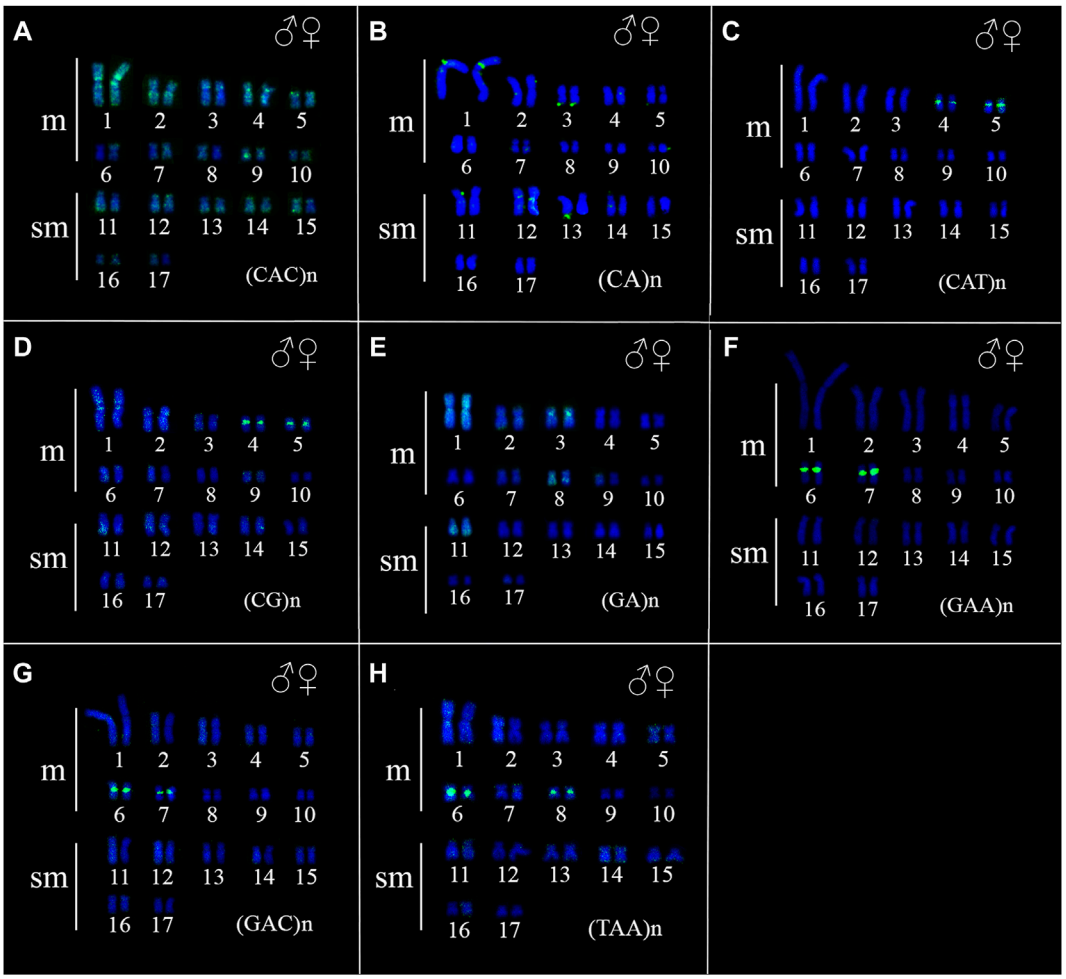


FIGURE 5 | Fluorescence *in situ* hybridization indicating the physical location of microsatellite sequences in the karyotype of Ancistrus sp. 2: **(A)** (CAC)n, **(B)** (CA)n, **(C)** (CAT)n, **(D)** (CG)n, **(E)** (GA)n, **(F)** (GAA)n, **(G)** (GAC)n, **(H)** (TAA)n. The probes were labeled with FITC (green), the chromosomes were counterstained with DAPI (blue).

been proposed in distinct genera of Loricariidae (Barros et al., 2017; Glugoski et al., 2018; Deon et al., 2020). Our results in *Ancistrus* sp. 2 showed multiple 5S rDNA sites located in the pericentromeric and proximal regions of some chromosome pairs. This data agrees with the hypothesis that 5S rDNA, or sequences derived from this gene family, may be involved in fusion events in *Ancistrus* sp. 2, as proposed previously for other species of this genus. Furthermore, these results supported the hypothesis that these sequences may represent evolutionary breakpoints regions (EBRs), which can be reused in chromosomal rearrangements in *Ancistrus* (Barros et al., 2017).

Comparative mapping of microsatellites revealed divergent patterns of organization between karyotypes of the *Ancistrus* species, occurring in euchromatic and heterochromatic regions in autosome and sex chromosomes. In *Ancistrus* sp. 2, the microsatellites mainly colonized heterochromatic blocks in centromeric and pericentromeric regions. Centromeric regions are characterized by the abundance of in tandem repeats, which are essential for maintaining the stability of this chromosomal region (Shang et al., 2010). However, the presence of different types of in tandem repeats, including microsatellites, at breakpoints for chromosomal rearrangements has been demonstrated previously (Kejnovsky et al., 2009; Cioffi and Bertollo, 2010; Ferré et al., 2011). The location of different microsatellites in the centromeric and pericentromeric region of metacentric chromosomes may indicate their association with EBRs (Ferré et al., 2011), suggesting the occurrence of Rb fusions during the evolution of the *Ancistrus* sp. 2. Furthermore, the association between microsatellites and rDNA sites has been observed in several organisms (Santos da Silva et al., 2021), corroborating to the chromosomal instability proposal to rDNA sites in *Ancistrus*. Therefore, the analyses carried out in this study suggest the participation of repetitive sequences in different mechanisms of chromosomal diversification in this group of neotropical fish.

Sex Chromosomes in *Ancistrus*

In general, sex chromosomes occur in only a small portion of neotropical fish species, having independent evolutionary origins and evolving from different mechanisms (Charlesworth et al., 2005;

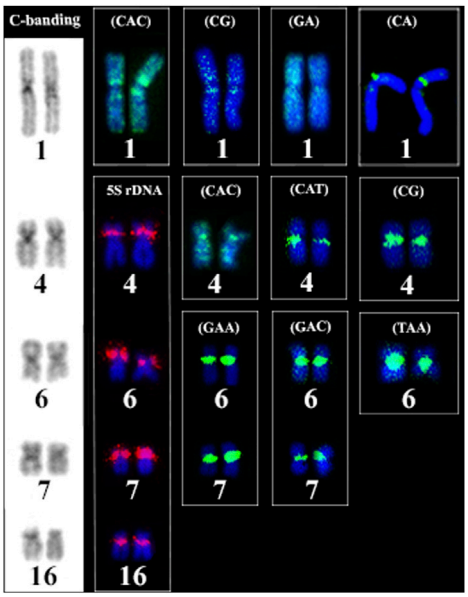


FIGURE 6 | Comparative analysis of the distribution of heterochromatic regions, 5S rDNA sites and microsatellite sequences and in the karyotype of Ancistrus sp. 2. Boxes indicate the repetitive sequences in heterochromatic regions of chromosome pairs 1, 4, 6, 7 and 16 of *Ancistrus* sp. 2. Chromosomes stained with Giemsa; hybridization signals with FITC (green) and Cy3 (red), chromosomes were counterstained with DAPI (blue).

Henning et al., 2008; Henning et al., 2011; Schamberger et al., 2019). In *Ancistrus*, extensive sex chromosome diversity is described, with different levels of morphological differentiation and DNA content (Mariotto et al., 2004; Mariotto and Miyazawa, 2006; de Oliveira et al., 2007; de Oliveira et al., 2008; de Oliveira et al., 2009). *Ancistrus* sex chromosomes have been analyzed mainly according to their size, heteromorphisms, and distribution of heterochromatic regions (Mariotto et al., 2004; Mariotto and Miyazawa, 2006; de Oliveira et al., 2007; de Oliveira et al., 2008; de Oliveira et al., 2009). Theoretically, heterochromatinization has been considered an essential step in proto sex chromosome differentiation due to the differential accumulation of repetitive sequences and its effects in decreasing the recombination rate (Cioffi et al., 2012). Partially or fully heterochromatic sex chromosomes could be considered a characteristic of well-differentiated systems in fish, as noted in *Eigenmannia*, *Tripohorteus*, *Characidium*, and Parodontidae (Henning et al., 2011; Cioffi et al., 2014; Schemberger et al., 2014; Ziemniczak et al., 2014; Pucci et al., 2016). However, the absence of heterochromatic regions is a frequent state in the sex chromosomes in *Ancistrus*, including XX/XY or ZZ/ZW systems (de Oliveira et al., 2007; de Oliveira et al., 2009; present study), suggesting that the sex chromosomes in *Ancistrus* evolved independently and, therefore, are at different stages of differentiation regarding the accumulation of repetitive sequences and heterochromatinization.

In situ localization of repetitive sequences represent an important approach for studying sex chromosome diversity and evolution in fish (Cioffi et al., 2010; Schemberger et al., 2014; Schemberger et al., 2019). In *Ancistrus*, this approach has been applied in some species (Favarato et al., 2017; Prizon et al., 2018), indicating the participation of different repetitive sequences in the sex chromosome differentiation (Prizon et al., 2018). Here, the microsatellite comparative *in situ* localization mapping highlights differences between heteromorphic X and Y sex chromosomes in *Ancistrus* sp. 1. Usually, in XX/XY sex chromosome systems, the heteromorphic sex chromosome differentiates by accumulating repetitive sequences and heterochromatinization (Charlesworth et al., 2005). Furthermore, the Y chromosome generally follows through a degeneration pathway due to the absence of recombination, leading to its reduced size (Charlesworth et al., 2005). In this study, a higher concentration of microsatellite sequences were found in the X when compared to the Y chromosome; however, the presence of Y-speceific microsatellite sequences was also observed. These results can be explained, in part, by partial recombination restriction between the homologs of the proto sex pair promoted by the differential accumulation of repetitive sequences (Kejnovsky et al., 2009).

Sex chromosomes are commonly rich in distinct families of transposable elements (TE) and in tandem repeats (Charlesworth et al., 2005). The invasion of sex chromosomes by TE occurs at different stages of the differentiation of these chromosomes (Charlesworth et al., 2005; Schemberger et al., 2019). These sequences are inactivated or degenerated leading to heterochromatinization of parts of the sex chromosomes (Charlesworth et al., 2005). In *Ancistrus* sp. 1, C-banding demonstrated the absence of heterochromatin in the sex chromosomes, suggesting a recent stage of differentiation when compared to sex chromosomes from other fish groups (Cioffi et al., 2010; Henning et al., 2011; Cioffi et al., 2014; Schemberger et al., 2014; Schemberger et al., 2019). On the other hand, heteromorphic size suggests the occurrence of Y chromosome degeneration in *Ancistrus* sp. 1, as demonstrated by the different repetitive DNA content between X and Y-chromosomes. Future analyses integrating *in situ* location data from different repetitive units classes and epigenetic analyses will be important to test the condition of repetitive DNA segments that have not undergone heterochromatinization in *Ancistrus* sex chromosomes.

CONCLUSION

Our study provided additional evidence on the evolutionary pathways to 2n reduction in *Ancistrus* species, highlighting specific chromosomal features that have emerged throughout their life. The obtained data also suggest the participation of repetitive sequences acting in *Ancistrus* sp. 1 and *Ancistrus* sp. 2 diversification, as those sequences can be involved in the Robertsonian rearrangements and sex chromosomes differentiation.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by The Animal Ethics Committee of the Federal University of Pará.

AUTHOR CONTRIBUTIONS

KS: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Visualization; Writing of the original manuscript; Redaction; Revision; Edition. LG: Investigation; Methodology; Visualization; Redaction; Revision; Edition. MV: Methodology; Resources; Visualization; Redaction; Revision; Edition. AdS: Investigation; Methodology; Resources; Visualization; Redaction; Revision; Edition. RN: Investigation; Methodology; Resources; Visualization; Redaction; Revision; Edition. JP: Data curation; Formal analysis; Acquisition of financing; Resources; Visualization; Writing, Proofreading and Editing. CN: Data curation; Formal analysis; Acquisition of financing; Project administration; Resources; Supervision; Visualization; Redaction; Revision; Edition.

FUNDING

The authors thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support on project coordinated by CN (Edital Pró-Amazônia Proc 047/2012); the FAPESPA for financial support (Edital Vale—Proc 2010/110447) and Banco Nacional de Desenvolvimento Econômico e Social—BNDES (2.318.697.0001) on a project coordinated by JP. CN (305880/2017-9), JP (305876/2017-1) and MV (305142/2019-4) are grateful to CNPq for Productivity Grants; the Pro-Reitoria de Pesquisa e Pos-Graduacao, Universidade Federal do Para (PROPESP—UFPA) for the payment of the article processing fee. This study is part of the Doctoral Thesis of KS in Genetic and Molecular Biology who is recipient of CAPES Doctor Scholarship.

ACKNOWLEDGMENTS

Sample collections was authorized by Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) and Secretaria de Estado de Meio Ambiente do Pará (SEMA-PA). 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. To David Santos da Silva for making the maps and helping with the fieldwork.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.838462/full#supplementary-material

REFERENCES

Alves, A. L., Oliveira, C., Nirchio, M., Granado, Á., 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), 1–9. doi:10.1007/s10709-005-0715-1

Armbruster, J. W., and Lujan, N. K. (2016). A New Species of *Peckoltia* from the Upper Orinoco (Siluriformes, Loricariidae). *ZooKeys* 569, 105–121. doi:10.3897/zookeys.569.6630

Armbruster, J. W. (2004). Phylogenetic Relationships of the Suckermouth Armoured Catfishes (Loricariidae) with Emphasis on the Hypostominae and the Ancistrinaeshes Loricariidae with Emphasis on the Hypostominae and the Ancistrinae. *Zoolog. J. Linn. Soc.* 141 (1), 1–80. doi:10.1111/j.1096-3642.2004.00109.x

Armbruster, J. W. (2008). The Genus *Peckoltia* with the Description of Two New Species and a Reanalysis of the Phylogeny of the Genera of the Hypostominae (Siluriformes: Loricariidae). *Zootaxa* 1822 (1), 1–76. doi:10.11646/zootaxa.1822.1.1

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

Barros, A. V., Wolski, M. A. V., Nogaroto, V., Almeida, M. C., Moreira-Filho, O., and Vicari, M. R. (2017). Fragile Sites, Dysfunctional Telomere and Chromosome Fusions: what Is 5S rDNA Role? *Gene* 608, 20–27. doi:10.1016/j.gene.2017.01.013

Bertollo, L. A. C., Takahashi, C. S., and Moreira-Filho, O. (1978). Cytotaxonomic Considerations on *Hoplias Lacerdae* (Pisces Erythrinidae). *Braz. J. Genet.* 1, 103–120.

de Borba, R. S., Mariotto, S., Centofante, L., Henrique Zawadzki, C., and Pasquali Parise-Maltempi, P. (2019). Molecular Discrimination of *Ancistrus* Lineages (Siluriformes: Loricariidae) Using Barcode DNA Tool. *Mitochondrial DNA A* 30 (4), 602–608. doi:10.1080/24701394.2019.1597071

Bueno, V., Konerat, J. T., Zawadzki, C. H., Venere, P. C., Blanco, D. R., and Margarido, V. P. (2018). Divergent Chromosome Evolution in Hypostominae Tribes (Siluriformes: Loricariidae): Correlation of Chromosomal Data with Morphological and Molecular Phylogenies. *Zebrafish* 15 (5), 492–503. doi:10.1089/zeb.2018.1612

Charlesworth, D., Charlesworth, B., and Marais, G. (2005). Steps in the Evolution of Heteromorphic Sex Chromosomes. *Heredity* 95 (2), 118–128. doi:10.1038/sj.hdy.6800697

Cioffi, M. B., and Bertollo, L. A. C. (2010). Initial Steps in XY Chromosome Differentiation in *Hoplias malabaricus* and the Origin of an X1X2Y Sex Chromosome System in This Fish Group. *Heredity* 105 (6), 554–561. doi:10.1038/hdy.2010.18

Cioffi, M. B., Martins, C., Vicari, M. R., Rebordinos, L., and Bertollo, L. A. C. (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 (3), 176–185. doi:10.1159/000309726

Cioffi, M. B., Moreira-Filho, O., Almeida-Toledo, L. F., and Bertollo, L. A. (2012). The Contrasting Role of Heterochromatin in the Differentiation of Sex Chromosomes: an Overview from Neotropical Fishes. *J. Fish. Biol.* 80 (6), 2125–2139. doi:10.1111/j.1095-8649.2012.03272.x

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 (2-3), 186–194. doi:10.1159/000354039

Cioffi, M., Fernanda Yano, C., Bertollo, L. A. C., Molina, W., and Liehr, T. (2014). Genomic Organization of Repetitive DNAs and its Implications for Male Karyotype and the Neo-Y Chromosome Differentiation in *Erythrinus Erythrinus* (Characiformes, Erythrinidae). *Ccg* 8 (2), 139–151. doi:10.3897/CompCytogen.v8i2.7597

Cioffi, M. d. B., Yano, C. F., Sember, A., and Bertollo, L. A. C. (2017). Chromosomal Evolution in Lower Vertebrates: Sex Chromosomes in Neotropical Fishes. *Genes* 8 (10), 258. doi:10.3390/genes8100258

Da Silva, D. S., da Silva Filho, H. F., Cioffi, M. B., de Oliveira, E. H. C., and Gomes, A. J. B. (2021). Comparative Cytogenetics in Four Leptodactylus Species (Amphibia, Anura, Leptodactylidae): Evidence of Inner Chromosomal Diversification in Highly Conserved Karyotypes. *Cytogenet. Genome Res.* 161 (1-2), 52–62. doi:10.1159/000515831

De Oliveira, R. R., Feldberg, E., Anjos, M. B. d., 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., Dos Anjos, M. B., and Zuanon, J. (2008). Occurrence of Multiple Sexual Chromosomes (XX/XY_1Y_2 and $Z_1Z_1Z_2Z_2/Z_1Z_2W_1W_2$) in Catfishes of the Genus Ancistrus (Siluriformes: Loricariidae) from the Amazon basin. *Genetica* 134 (2), 243–249. doi:10.1007/s10709-007-9231-9

De Oliveira, R. R., Feldberg, E., Dos Anjos, M. B., and Zuanon, J. (2009). Mechanisms of Chromosomal Evolution and its Possible Relation to Natural History Characteristics inAncistruscatfishes (Siluriformes: Loricariidae). *J. Fish Biol.* 75 (9), 2209–2225. doi:10.1111/j.1095-8649.2009.02450.x

De Oliveira, T. D., Kretschmer, R., Bertocchi, N. A., Degrandi, T. M., de Oliveira, E. H. C., Cioffi, M. d. B., et al. (2017). Genomic Organization of Repetitive DNA in Woodpeckers (Aves, Piciformes): Implications for Karyotype and ZW Sex Chromosome Differentiation. *PLoS ONE* 12 (1), e0169987. doi:10.1371/journal.pone.0169987

Deon, G. A., Glugoski, L., Vicari, M. R., Nogaroto, V., Sassi, F. d. M. C., Cioffi, M. d. B., et al. (2020). Highly Rearranged Karyotypes and Multiple Sex Chromosome Systems in Armored Catfishes from the Genus *Harttia* (Teleostei, Siluriformes). *Genes* 11 (11), 1366. doi:10.3390/genes11111366

Farré, M., Bosch, M., López-Giráldez, F., Ponsà, M., and Ruiz-Herrera, A. (2011). Assessing the Role of Tandem Repeats in Shaping the Genomic Architecture of Great Apes. *PLoS ONE* 6 (11), e27239. doi:10.1371/journal.pone.0027239

Favarato, R. M., da Silva, M., de 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 (2), 103–111. doi:10.1089/zeb.2015.1140

Favarato, R. M., Ribeiro, L. B., Feldberg, E., and Matoso, D. A. (2017). Chromosomal Mapping of Transposable Elements of the rex Family in the Bristlenose Catfish, *Ancistrus* (Siluriformes, Loricariidae), from the Amazonian Region. *Jhered* 108 (3), esw084. doi:10.1093/jhered/esw084

Ferraris, C. J., Jr. (2007). Checklist of Catfishes, Recent and Fossil (Osteichthyes: Siluriformes), and Catalogue of Siluriform Primary Types. *Zootaxa* 1418, 1–628. doi:10.11646/zootaxa.1418.1.1

Fricke, R., Eschmeyer, W. N., and Van der Laan, R. (2021). ESCHMEYER'S CATALOG of FISHES: Genera, Species, References. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (Accessed October 27, 2021).

Glugoski, L., Giuliano-Caetano, L., Moreira-Filho, O., Vicari, M. R., and Nogaroto, V. (2018). Co-located hAT Transposable Element and 5S rDNA in an Interstitial Telomeric Sequence Suggest the Formation of Robertsonian Fusion in Armored Catfish. *Gene* 650, 49–54. doi:10.1016/j.gene.2018.01.099

Glugoski, L., Deon, G., Schott, S., Vicari, M. R., Nogaroto, V., and Moreira-Filho, O. (2020). Comparative Cytogenetic Analyses in *Ancistrus* Species (Siluriformes: Loricariidae). *Neotrop. Ichthyol.* 18 (2), 1–16. doi:10.1590/1982-0224-2020-0013

Gornung, E. (2013). Twenty Years of Physical Mapping of Major Ribosomal RNA Genes across the Teleosts: a Review of Research. *Cytogenet. Genome Res.* 141 (2-3), 90–102. doi:10.1159/000354832

Gross, M. C., Schneider, C. H., Valente, G. T., Martins, C., and Feldberg, E. (2010). Variability of 18S rDNA Locus amongSymphysodonfishes: Chromosomal Rearrangements. *J. Fish Biol.* 76 (5), 1117–1127. doi:10.1111/j.1095-8649.2010.02550.x

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

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 (2), 391–400. doi:10.1038/hdy.2010.82

Ijdo, J. W., Wells, R. A., Baldini, A., and Reeders, S. T. (1991). Improved Telomere Detection Using a Telomere Repeat Probe (TTAGGG)ngenerated by PCR. *Nucl. Acids Res.* 19 (17), 4780. doi:10.1093/nar/19.17.4780

Kejnovsky, E., Hobza, R., Cermak, T., Kubat, Z., and Vyskot, B. (2009). The Role of Repetitive DNA in Structure and Evolution of Sex Chromosomes in Plants. *Heredity* 102 (6), 533–541. doi:10.1038/hdy.2009.17

Levan, A., Fredga, K., and Sandberg, A. A. (1964). Nomenclature for Centromeric Position on Chromosomes. *Hereditas* 52 (2), 201–220. doi:10.1111/j.1601-5223.1964.tb01953.x

Long, E. O., and Dawid, I. B. (1980). Repeated Genes in Eukaryotes. *Annu. Rev. Biochem.* 49 (1), 727–764. doi:10.1146/annurev.bi.49.070180.003455

Lujan, N. K., Armbruster, J. W., Lovejoy, N. R., and López-Fernández, 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., and Miyazawa, C. S. (2006). Ancistrus Cf. Dubius (Siluriformes, Ancistrinae), a Complex of Species. 1. Chromosomic Characterization of Four Populations and Occurence of Sexual Chromosomes of Type XX/XY, in the Pantanal basin of Mato Grosso, Brazil. *Caryologia* 59 (4), 299–304. doi:10.1080/00087114.2006.10797929

Sandra, M., Artoni, R. F., and Miyazawa, C. S. (2004). Occurrence of Sexual Chromosome, of the Type ZZ/ZW, in Ancistrus Cf. Dubius (Loricariidae, Ancistrinae) of the Paraguay River Basin, Mato Grosso, Brazil*Ancistrus* Cf. *Dubius* (Loricariidae, Ancistrinae) of the Paraguay River Basin, Mato Grosso, Brazil. *Caryologia* 57 (4), 327–331. doi:10.1080/00087114.2004.10589413

Mariotto, S., Centofante, L., Vicari, M., Artoni, R., 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. *Ccg* 5 (4), 289–300. doi:10.3897/compcytogen.v5i4.1757

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

Martins, C. (2007). "Chromosomes and Repetitive DNAs: A Contribution to the Knowledge of the Fish Genome," in *Fish Cytogenetics*. Editors E. Pisano, C. Ozouf-Costaz, F. Foresti, and B. G. Kapoor (Enfield: Science Publishers), 421–453. doi:10.1201/b10746-16

Pansonato-Alves, J. C., Serrano, É. A., Utsunomia, R., Scacchetti, P. C., Oliveira, C., and Foresti, F. (2013). Mapping Five Repetitive DNA Classes in Sympatric Species of *Hypostomus* (Teleostei: Siluriformes: Loricariidae): Analysis of Chromosomal Variability. *Rev. Fish. Biol. Fish.* 23 (4), 477–489. doi:10.1007/s11160-013-9303-0

Pety, A. M., Cardoso, A. L., Nagamachi, C. Y., Pieczarka, J. C., de Sousa, L. M., and Noronha, R. C. R. (2018). *In Situ* localization of Ribosomal Sites in *Peckoltia* and *Ancistomus* (Loricariidae: Hypostominae) from the Amazon Basin. *Zebrafish* 15 (3), 263–269. doi:10.1089/zeb.2017.1523

Primo, C. C., Glugoski, L., Almeida, M. C., Zawadzki, C. H., Moreira-Filho, O., Vicari, M. R., et al. (2017). Mechanisms of Chromosomal Diversification in Species of Rineloricaria (Actinopterygii: Siluriformes: Loricariidae). *Zebrafish* 14 (2), 161–168. doi:10.1089/zeb.2016.1386

Prizon, A. C., Borin-Carvalho, L. A., Bruschi, D. P., Otávio Ribeiro, M., Magrinelli Barbosa, L., de Brito 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. *Ccg* 10 (4), 625–636. doi:10.3897/CompCytogen.v10i4.8532

Prizon, A. C., Bruschi, D. P., Borin-Carvalho, L. A., Cius, A., Barbosa, L. M., Ruiz, H. B., et al. (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

Prizon, A. C., Bruschi, D. P., Gazolla, C. B., Borin-Carvalho, L. A., and Portela-Castro, A. L. D. B. (2018). Chromosome Spreading of the Retrotransposable

Rex-3 Element and Microsatellite Repeats in Karyotypes of the *Ancistrus* Populations. *Zebrafish* 15 (5), 504–514. doi:10.1089/zeb.2018.1620

Pucci, M. B., Barbosa, P., Nogaroto, V., Almeida, M. C., Artoni, R. F., Scacchetti, P. C., et al. (2016). Chromosomal Spreading of Microsatellites and (TTAGGG)n Sequences in the *Characidium Zebra* and *C. Gomesi* Genomes (Characiformes: Crenuchidae). *Cytogenet. Genome Res.* 149, 182–190. doi:10.1159/000447959

Rebordinos, L., Cross, I., and Merlo, A. (2013). High Evolutionary Dynamism in 5S rDNA of Fish: State of the Art. *Cytogenet. Genome Res.* 141 (2-3), 103–113. doi:10.1159/000354871

Roxo, F. F., Ochoa, L. E., Sabaj, M. H., Lujan, N. K., Covain, R., Silva, G. S. C., et al. (2019). Phylogenomic Reappraisal of the Neotropical Catfish Family Loricariidae (Teleostei: Siluriformes) Using Ultraconserved Elements. *Mol. Phylogenet. Evol.* 135, 148–165. doi:10.1016/j.ympev.2019.02.017

Santos da Silva, K., de Souza, A. C. P., Pety, A. M., Noronha, R. C. R., Vicari, M. R., Pieczarka, J. C., et al. (2021). Comparative Cytogenetics Analysis Among *Peckoltia* Species (Siluriformes, Loricariidae): Insights on Karyotype Evolution and Biogeography in the Amazon Region. *Front. Genet.* 12, 1–9. doi:10.3389/fgene.2021.779464

Schemberger, M. O., Oliveira, J. I. N., Nogaroto, V., Almeida, M. C., Artoni, R. F., Cestari, M. M., et al. (2014). Construction and Characterization of a Repetitive DNA Library in Parodontidae (Actinopterygii: Characiformes): A Genomic and Evolutionary Approach to the Degeneration of the W Sex Chromosome. *Zebrafish* 11 (6), 518–527. doi:10.1089/zeb.2014.1013

Schemberger, M. O., Nascimento, V. D., Coan, R., Ramos, É., Nogaroto, V., Ziemniczak, K., et al. (2019). DNA Transposon Invasion and Microsatellite Accumulation Guide W Chromosome Differentiation in a Neotropical Fish Genome. *Chromosoma* 128 (4), 547–560. doi:10.1007/s00412-019-00721-9

Shang, W.-H., Hori, T., Toyoda, A., Kato, J., Popendorf, K., Sakakibara, Y., et al. (2010). Chickens Possess Centromeres with Both Extended Tandem Repeats and Short Non-tandem-repetitive Sequences. *Genome Res.* 20 (9), 1219–1228. doi:10.1101/gr.106245.110

Slijepcevic, P. (1998). Telomeres and Mechanisms of Robertsonian Fusion. *Chromosoma* 107, 136–140. doi:10.1007/s004120050289

Sumner, A. T. (1972). A Simple Technique for Demonstrating Centromeric Heterochromatin. *Exp. Cel Res.* 75, 304–306. doi:10.1016/0014-4827(72)90558-7

Vidal, L. V. O., Albinati, R. C. B., Albinati, A. C. L., Lira, A. D. D., Almeida, T. R. D., and Santos, G. B. (2008). Eugenol como anestésico para a tilápia-Do-nilo. *Pesq. Agropec. Bras.* 43, 1069–1074. doi:10.1590/s0100-204x2008000800017

Ziemniczak, K., Barros, A. V., Rosa, K. O., Nogaroto, V., Almeida, M. C., Cestari, M. M., et al. (2012). Comparative Cytogenetics of Loricariidae (Actinopterygii: Siluriformes): Emphasis in Neoplecostominae and Hypoptopomatinae. *Ital. J. Zoolog.* 79 (4), 492–501. doi:10.1080/11250003.2012.676677

Ziemniczak, K., Traldi, J. B., Nogaroto, V., Almeida, M. C., Artoni, R. F., Moreira-Filho, O., et al. (2014). *In Situ* Localization of (GATA)n and (TTAGGG)n Repeated DNAs and W Sex Chromosome Differentiation in Parodontidae (Actinopterygii: Characiformes). *Cytogenet. Genome Res.* 144, 325–332. doi:10.1159/000370297

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Genetics

ORIGINAL RESEARCH
published: 24 March 2022
doi: 10.3389/fgene.2022.832495

Chromosome Painting in *Gymnotus carapo* "Catalão" (Gymnotiformes, Teleostei): Dynamics of Chromosomal Rearrangements in Cryptic Species

Milla de Andrade Machado[1], *Maelin da Silva*[2], *Eliana Feldberg*[3], *Patricia Caroline Mary O'Brien*[4], *Malcolm Andrew Ferguson-Smith*[4], *Julio Cesar Pieczarka*[1] *and Cleusa Yoshiko Nagamachi*[1]*

[1]*Laboratório de Citogenética, Centro de Estudos Avançados da Biodiversidade, Instituto de Ciências Biológicas, Universidade Federal Do Pará (UFPA), Belém, Brazil,* [2]*Departamento de Biologia Estrutural, Molecular e Genética, Universidade Estadual de Ponta Grossa, Ponta Grossa, Brazil,* [3]*Laboratório de Genética Animal, Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil,* [4]*Department of Veterinary Medicine, Cambridge Resource Centre for Comparative Genomics, University of Cambridge, Cambridge, United Kingdom*

OPEN ACCESS

Edited by:
Tony Silveira,
Federal University of Rio Grande, Brazil

Reviewed by:
Jack Craig,
Temple University, United States
Denisa Šimoníková,
Academy of Sciences of the Czech Republic, Czechia

***Correspondence:**
Cleusa Yoshiko Nagamachi
cleusa@ufpa.br
cleusanagamachi@gmail.com

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

Received: *09 December 2021*
Accepted: *10 February 2022*
Published: *24 March 2022*

Citation:
Machado MA, da Silva M, Feldberg E, O'Brien PCM, Ferguson-Smith MA, Pieczarka JC and Nagamachi CY (2022) Chromosome Painting in Gymnotus carapo *"Catalão" (Gymnotiformes, Teleostei): Dynamics of Chromosomal Rearrangements in Cryptic Species. Front. Genet. 13:832495. doi: 10.3389/fgene.2022.832495*

The genus *Gymnotus* is a large monophyletic group of freshwater weakly-electric fishes, with wide distribution in Central and South America. It has 46 valid species divided into six subgenera (*Gymnotus*, *Tijax*, *Tigre*, *Lamontianus*, *Tigrinus* and *Pantherus*) with large chromosome plasticity and diploid numbers (2n) ranging from 34 to 54. Within this rich diversity, there is controversy about whether *Gymnotus* (*Gymnotus*) *carapo* species is a single widespread species or a complex of cryptic species. Cytogenetic studies show different diploid numbers for *G. carapo* species, ranging from 40 to 54 chromosomes with varied karyotypes found even between populations sharing the same 2n. Whole chromosome painting has been used in studies on fish species and recently has been used for tracking the chromosomal evolution of *Gymnotus* and assisting in its cytotaxonomy. Comparative genomic mapping using chromosome painting has shown more complex rearrangements in *Gymnotus carapo* than shown in previous studies by classical cytogenetics. These studies demonstrate that multiple chromosome pairs are involved in its chromosomal reorganization, suggesting the presence of a complex of cryptic species due to a post zygotic barrier. In the present study, metaphase chromosomes of *G. carapo occidentalis* "catalão" (GCC, 2n = 40, 30m/sm+10st/a) from the Catalão Lake, Amazonas, Brazil, were hybridized with whole chromosome probes derived from the chromosomes of *G. carapo* (GCA, 2n = 42, 30m/sm+12st/a). The results reveal chromosome rearrangements and a high number of repetitive DNA sites. Of the 12 pairs of *G. carapo* chromosomes that could be individually identified (GCA 1–3, 6, 7, 9, 14, 16 and 18–21), 8 pairs (GCA 1, 2, 6, 7, 9, 14, 20, 21) had homeology conserved in GCC. Of the GCA pairs that are grouped (GCA [4, 8], [5, 17], [10, 11] and [12, 13, 15]), most kept the number of signals in GCC (GCA [5, 17], [10, 11] and [12, 13, 15]). The remaining chromosomes are rearranged in the GCC karyotype. Analysis of both populations of the *G. carapo* cytotypes shows extensive karyotype reorganization. Along with previous studies, this suggests that the different cytotypes analyzed here

may represent different species and supports the hypothesis that *G. carapo* is not a single widespread species, but a group of cryptic species.

Keywords: amazon, banded knife-fish, whole chromosome painting, FISH, complex of species

INTRODUCTION

The genus *Gymnotus* (Gymnotiformes, Teleostei) is a large monophyletic group of weakly-electric freshwater fish. It is the most speciose genus of the order, with 46 validated species (Craig et al., 2019; Kim et al., 2020; Fricke et al., 2021), widely distributed in the Neotropical region (Central and South America). The highest diversity is found in the Amazon-Orinoco-Guiana basins (Mago-Leccia, 1994; Albert, 2001; Albert and Crampton, 2005).

Gymnotus is divided into six subgenera: *Gymnotus*, *Tijax*, *Tigre*, *Lamontianus*, *Tigrinus* and *Pantherus* (Craig et al., 2019) with substantial chromosome plasticity. The diploid number (2n) varies 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). The growing number of studies based not only on the karyotypic formula but also on different kinds of repetitive DNA sequences such as rDNAs, satellites, microsatellites and transposable elements (Milhomem et al., 2007; Claro, 2008; Milhomem et al., 2008; Scacchetti et al., 2011; Milhomem et al., 2012a; Milhomem et al., 2012b; da Silva et al., 2014; Utsunomia et al., 2014; Almeida et al., 2015; da Silva, 2015; da Silva et al., 2016; Machado et al., 2017; Utsunomia et al., 2018), have shown many different species-specific karyotypes and even population variants.

This large variation in 2n, however, is mostly found in the *G. carapo* subgenus, the previous *Gymnotus carapo* clade (Craig et al., 2019). In all the other subgenera, the diploid number varies from 48 to 54 chromosomes, many sharing the 2n = 52 that is found in the basal subgenus, the species *G. (pantherus) pantherinus* (Scacchetti et al., 2011; da Silva et al., 2011; da Silva et al., 2014; Utsunomia et al., 2014; Almeida et al., 2015; da Silva et al., 2016; Machado et al., 2017; da Silva et al., 2019), and also found in *Electrophorus*, the sister genus of *Gymnotus* (Fonteles et al., 2008; Cardoso et al., 2015).

In the subgenus *Gymnotus*, the *G. carapo* species has huge karyotype diversity among populations, with 2n ranging from 40 to 54, and many karyotypic formulas within the 2n described (**Table 1**). The species *G. carapo*, based on morphology and distribution data, was divided into subspecies *Gymnotus c. australis, Gymnotus c. caatingaensis, Gymnotus c. carapo, Gymnotus c. madeirensis, Gymnotus c. occidentalis, Gymnotus c. orientalis,* and *Gymnotus c. septentrionalis* (Craig et al., 2017).

TABLE 1 | Cytogenetic studies of *Gymnotus arapaima, G. capanema and G. carapo*.

Species	2n (KF)	Authors	Localidades
G. arapaima	44 (24m/sm+20st/a)	Milhomem et al. (2013)	1. Reserva de Desenvolvimento Sustentável Mamirauá, AM
G. capanema	34 (20m/sm+14st/a)	Milhomem et al. (2012b), Milhomem et al. (2012a))	2. Capanema, PA
G. carapo	54 (54m/sm)	Foresti et al. (1984)	3. Miracatu, SP; 4. Botucatu, SP
	52 (50m/sm+2st/a)		5. Brotas, SP
	48 (34m/sm+14st/a)		6. Humaitá, AM
	42 (32m/sm+10st/a)		7. Belem, PA
	54 (52m/sm+2st/a)	Fernandes-Matioli et al. (1998)	8. Rio Mogi-Guaçu, SP
	54	Claro (2008)	9. Santa Albertina, SP;
			10. Cardoso, SP;
			11. Terra Roxa, SP;
			12. Mariápolis, SP
			13. Corumbataí, SP;
			4. Botucatu, SP;
			14. Angatuba, SP;
			15. Indaiatuba, SP
			16. São Lorenço, SP;
			17. Bertioga, SP;
			18. Piquete, SP;
			19. Cruzeiro, SP.
	42 (30m/sm+12st/a)	Milhomem et al. (2007)	20. Santa Cruz do Arari, PA;
	42 (30m/sm+12st/a)	Milhomem et al. (2008)	21. Ponta de Pedras, PA;
			22. São Miguel do Guamá, PA;
			2. Capanema, PA;
			23. Benfica, PA
	40 (28m/sm+12st/a)	Milhomem et al. (2008)	24. Almeirim, PA
G. cf. *carapo*	54 (50m/sm+4st/a)	Scacchetti et al. (2011)	25. Passo do Lontra, MS
G. carapo "Catalão"	40 (30m/sm+10st)	da Silva et al. (2014)	26. Lago Catalão, AM
G. carapo "Maranhão"	42 (30m/sm+12st/a)	da Silva et al. (2019)	27. Rio Munin, MA

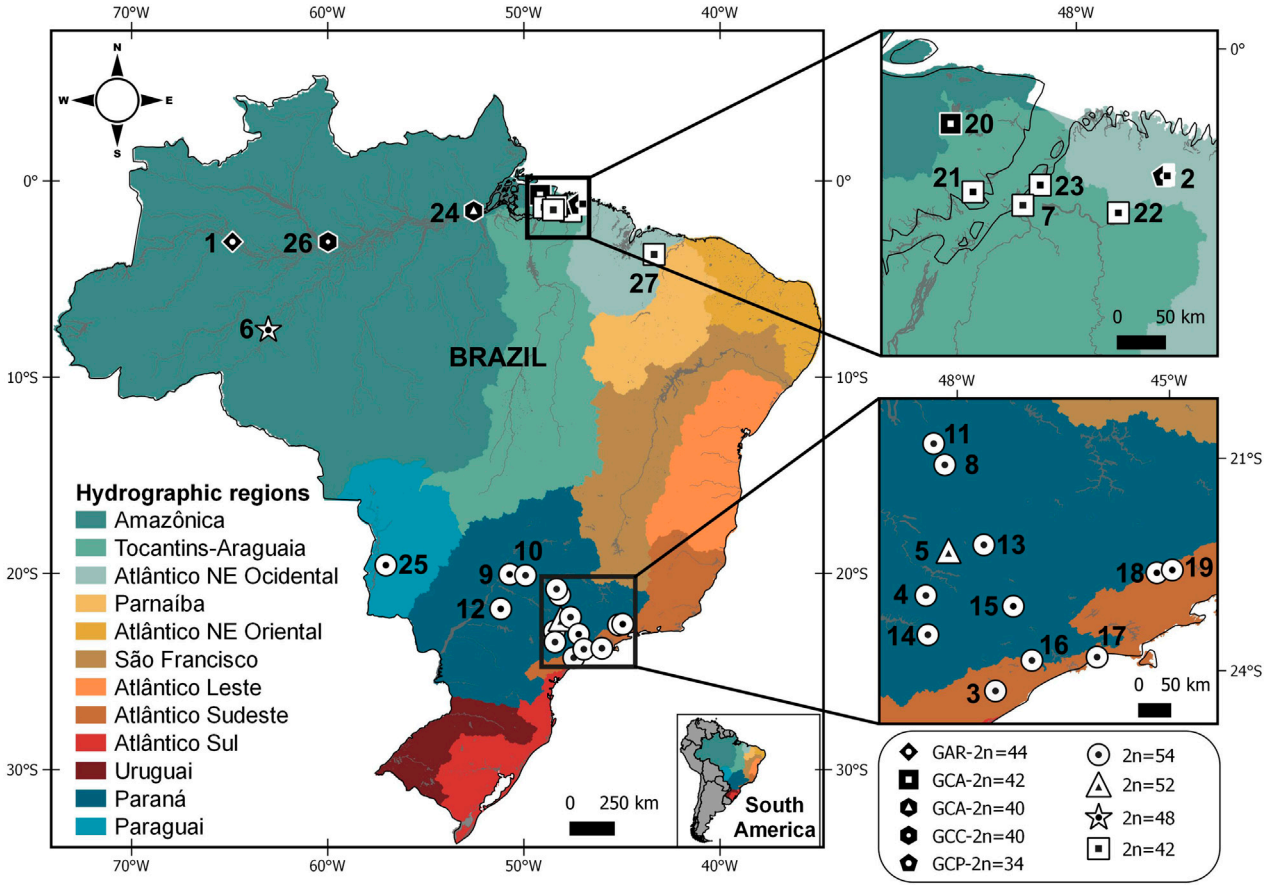


FIGURE 1 | Map showing the Brazilian hydrographic regions and sample points for *Gymnotus arapaima*, *G. capanema*, and *G. carapo*, with karyotypic information (2n = diploid number) from the present study and from the literature. Each symbol indicates the karyotype obtained from the locality; more than one symbol indicates that distinct karyotypes were collected at the same locality. Symbols in white denote specimens analyzed by classic cytogenetics; symbols in black denote specimens analyzed by chromosome painting with *G.* carapo (2n = 42) whole-chromosome probes (**Table 1**; Nagamachi et al., 2010; Nagamachi et al., 2013; Machado et al., 2018, and present study). The numbers refer to localities detailed on **Table 1**.

While most Neotropical fish species of South America have restricted geographic distributions, these species are distributed widely (Albert and Reis, 2011; Lehmberg, 2015; Craig et al., 2017).

Chromosome painting is as an important tool in comparative cytogenetics studies of fish species. The results have helped to solve various issues in this field and have giving insights into several evolutionary issues (Barby et al., 2019). It has helped to bring insights into the origin of B chromosomes (Vicari et al., 2011; Scudeler et al., 2015; Utsunomia et al., 2016), the origin and evolution of sex chromosomes (Carvalho et al., 2017; Cioffi et al., 2017; de Oliveira et al., 2017; de Moraes et al., 2017; Yano et al., 2017; de Freitas et al., 2018) and helped in the understanding of chromosomal evolution and relationships between closely related species (Ráb et al., 2008; Nagamachi et al., 2010; Nagamachi et al., 2013; Milhomem et al., 2013; Machado et al., 2018; Cioffi et al., 2019). Nagamachi et al. (2010) produced whole chromosome probes (WCP) from *G. c. orientalis* (GCA42, 2n = 42, 30m/sm+12st/a) by flow sorting, and hybridized these probes to the cytotypes of *G. c. orientalis* with 2n = 40 (GCA40, 34m/sm+6st/a). Two other studies with WCP from GCA42 mapped the karyotypes of *G. capanema* (GCP 2n = 34, Nagamachi et al., 2013) and *G. arapaima* (GAR 2n = 44, Machado et al., 2018). Those studies showed a higher level of chromosomal rearrangement than previously thought between these species.

In this study we used GCA42 WCP (Nagamachi et al., 2010) for mapping the karyotype of *G. c. occidentalis* "Catalão" (GCC 2n = 40), a distinctive population which has been proposed as a new species (da Silva et al., 2014). The results were compared with those obtained from GCA40 (Nagamachi et al., 2010), GCP34 (Nagamachi et al., 2013) and GAR44 (Machado et al., 2018).

MATERIAL AND METHODS

Sampling

Samples of *G. carapo* "Catalão" (GCC) were collected in Amazonas, Brazil (**Figure 1**). The Cytogenetics Laboratory from Centro de Estudos Avançados da Biodiversidade (UFPA) has permit number 19/2003 from the Ministry of Environment

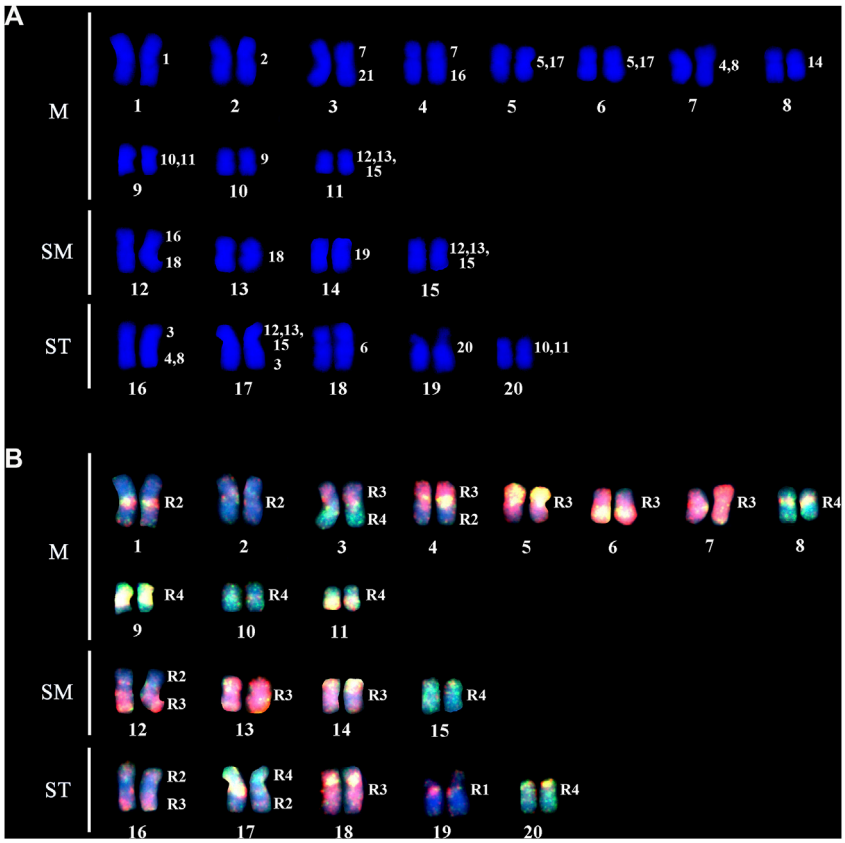


FIGURE 2 | (A) A DAPI stained karyotype of GCC; the numbers on the right represent the *G. carapo* (GCA42) equivalent chromosomes. **(B)** Dual color fish with the probes of R3 (pairs 4–8 and 17–19; red) and R4 (pairs 9–15 and 21; green). Chromosome segments hybridizing with 2 colors indicate repetitive DNA sequences. The chromosomes or segments in blue (DAPI) represent the NOR-bearing chromosome of GCA42 (pair 20) and the chromosomes corresponding to R2 of GCA42 (pairs 1–3 and 16).

for sample transport and permit 52/2003 for using the samples for research. The Ethics Committee from Para Federal University (Comitê de Ética Animal da Universidade Federal do Pará) approved this research (Permit 68/2015). Sample collections were authorized by Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) and Secretaria de Estado de Meio Ambiente do Pará (SEMA-PA) under permit 020/2005 (Registration: 207419).

Map

A distribution map was made using QGIS v.3.10.7. The shapefiles containing country limits were obtained from DIVA-GIS (Hijmans et al., 2004). We used the hydrographic regions limits provided by Braga et al. (2008) and we created the shapefiles on QGIS v.3.10.7. The localities numbered are shown on **Table 1**.

Whole Chromosome Painting

Whole Chromosome Probes (WCP) from *G. carapo* (2n = 42; 30m/sm+12st/a; Nagamachi et al., 2010) were hybridized onto metaphases of *G. carapo* "Catalão". Chromosome painting techniques followed Yang et al. (1995) with the modifications proposed by Nagamachi et al. (2010). Chromosomes were classified morphologically according to Levan et al. (1964) with modifications. The karyotype was organized following da Silva et al. (2014).

RESULTS

Chromosome Painting in *Gymnotus carapo Occidentalis* "Catalão"

Gymnotus carapo occidentalis "Catalão" (GCC) has 2n = 40 with 30m/sm+10st/a chromosomes (**Figure 2A**) without differentiated sex chromosomes in male and female specimens.

The regions of homology with GCA42 are indicated on the karyotype of GCC arranged from DAPI-stained chromosomes (**Figure 2A**).

Dual color FISH with the probes of GCA42 from R3 (pairs 4–8 and 17–19; red) and R4 (pairs 9–15 and 21; green) define the chromosome groups in GCC40 corresponding to the four groups in **Figure 2B**. Chromosome segments hybridizing with 2 colors

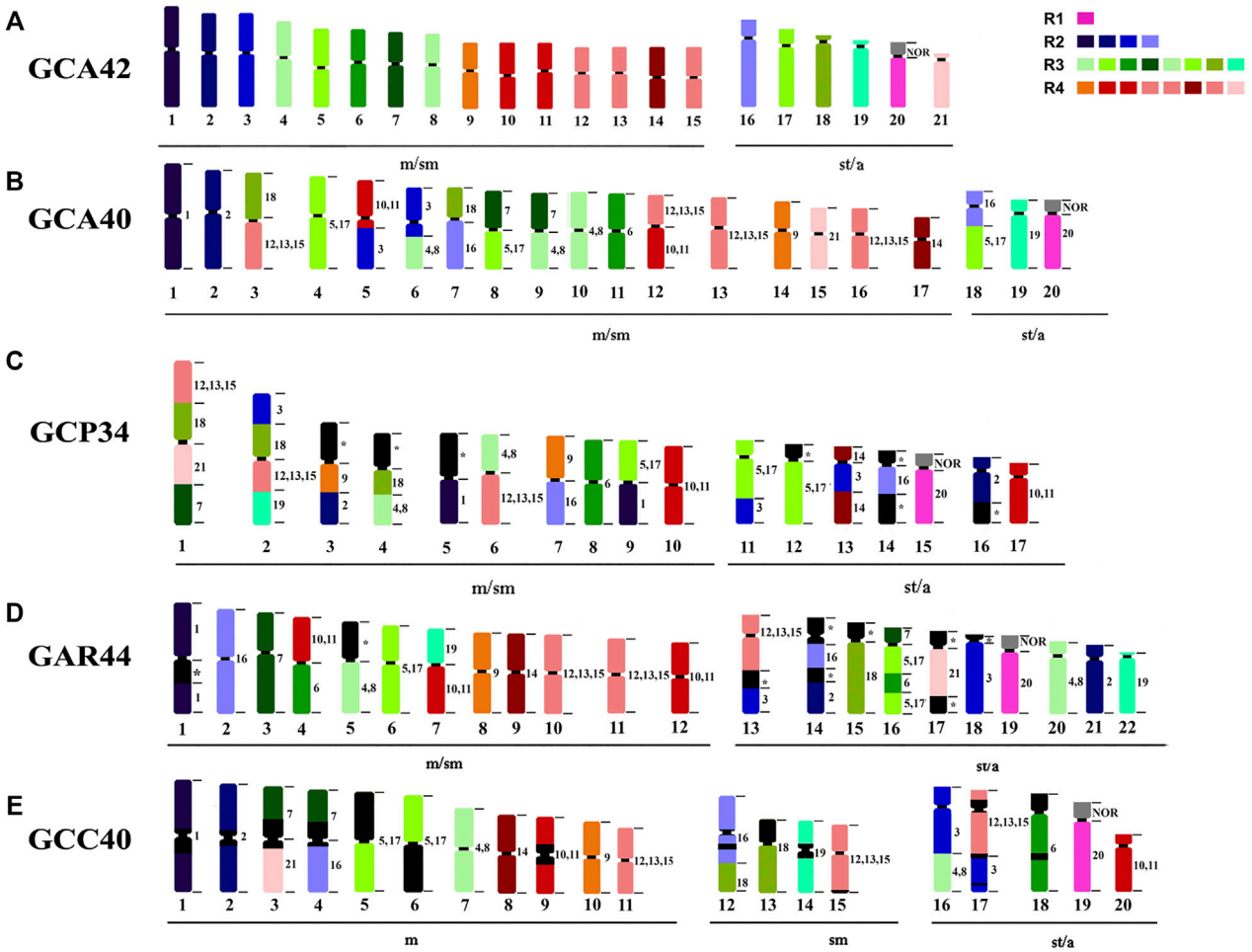


FIGURE 3 | Ideogram showing chromosome rearrangements obtained by WCP. R1, R2, R3 and R4 representing the colors of the four major regions of the *G. carapo* 2n = 42 probes. **(A)** *G. carapo* 2n = 42 (GCA42) utilized in the production of the probes, each different color represents a chromosome or chromosome group; **(B)** *G. carapo* 2n = 40 (GCA40) (Nagamachi et al., 2010); **(C)** *G. capanema* 2n = 34 (GCP34) (Nagamachi et al., 2013); **(D)** *G. arapaima* 2n = 44 (GAR44) (Machado et al., 2018); **(E)** *G. carapo* "catalão" 2n = 40 (GCC40) (present study).

indicate repetitive DNA sequences. The chromosomes or segments in blue (DAPI) represent the GCA42 NOR-bearing chromosome (pair 20) and the chromosomes corresponding to R2 (pairs 1–3 and 16).

From the 12 chromosome pairs of GCA42 that can be individually differentiated (pairs 1–3, 6, 7, 9, 14, 16 and 18–21), 8 pairs (1, 2, 6, 9, 14, 19, 20, 21) conserve homeology within GCC40 (pairs 1, 2, 3, 8, 10, 14, 18, 19). GCA42 pair 20 hybridizes one whole chromosome in GCC40, pair 19. Four chromosome pairs of GCA42 (3, 7, 16, and 18) show 2 signals on GCC40 (**Figure 2**).

The GCA42 probes that represent two chromosome pairs (4,8), (10,11) and (5,17) all reveal 2 signals, and the one that represents three pairs (12, 13, 15) reveals 3 signals on the GCC40 chromosomes.

The following chromosome associations of GCA42 are present in GCC40 pairs: 3 (7/C/21); 4 (7/C/16); 12 (16/C/16/18); 16 (3/C/3/[4, 8]); 17 ([12, 13, 15]/C/[12, 13, 15]/3), where C = centromere.

Comparative Chromosome Painting Analysis Among Different *Gymnotus* Karyotypes

From the 12 chromosome pairs in GCA42 that can be identified individually (1–3, 6, 7, 9, 14, 16, 18–21), GCC shows conserved synteny in 8 pairs (1, 2, 6, 9, 14, 19, 20, 21); GCC shares the same 8 pairs with GCA40, and also the same eight pairs shared between GCA40 and GCA42, grouping them together and closer to each other than the other analyzed karyotypes. From the pairs that are grouped in GCA42 [(4, 8), (10,11), (5,17), (12, 13, 15)], all keep the same number of signals in GCC40, while GCA40 has an extra signal for (5,17) and for (12, 14, 15) (**Figure 3**). The 16/18 syntenic associations are shared between GCC40 (16/C/16/18) and GCA40 (18/C/16) indicating pericentric inversions (**Figures 3**, **4**).

Compared to GAR, GCC shares five of the individually identified chromosomes (1, 9, 14, 20, 21), as well as the same number of signals in the groups (4, 8) and (12, 13, 15). There is

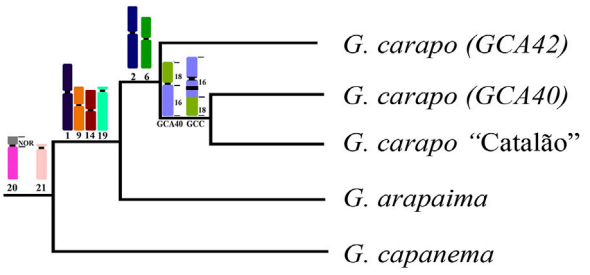


FIGURE 4 | Representative phylogeny based on da Silva et al. (2019), with the syntenic blocks shared by the nodes. Chromosome numbers refer to the *G. carapo* 2n = 42 chromosomes (see **Figure 3**).

TABLE 2 | Syntenic blocks shared among analyzed species with WCP. GCA42—*Gymnotus carapo* 2n = 42; GCA40—*G. carapo* 2n = 40; GCC40—*G. carapo* "Catalão" 2n = 40; GAR44 - *G. arapaima* 2n = 44; GCP34—*G. capanema* 2n = 34.

Species	Syntenic blocks
GCA42 x GCA40	GCA42 1, 2, 6, 9, 14, 19, 20, 21
GCC40 X GCA42	GCA42 1, 2, 6, 9, 14, 19, 20, 21
GCC40 X GCA40	GCA42 1, 2, 6, 9, 14, 19, 20, 21
GCC40 X GCA40 X GCA42	GCA42 1, 2, 6, 9, 14, 19, 20, 21
GCA42 x GAR44	GCA42 1, 9, 14, 18, 20, 21
GCA40 x GAR44	GCA42 1, 9, 14, 20, 21
GCC40 X GAR44	GCA42 1, 9, 14, 20, 21
GCC40 X GCA42 X GCA40 x GAR44	GCA42 1, 9, 14, 20, 21
GCA42 x GCP34	GCA42 6, 19, 20, 21
GCA40 x GCP34	GCA42 6, 19, 20, 21
GCC40 x GCP34	GCA42 6, 19, 20, 21
GAR44 x GCP34	GCA42 20, 21
ALL	GCA42 20, 21

also a similar chromosome rearrangement in GCC 17 and GAR 13 that is not shared with the other *G. carapo* cytotypes. However, as it is not possible to differentiate between GCA (2n = 42) 12,13, and 15, we cannot infer that the involved chromosome is the same or is different (**Figure 3**; **Table 2**).

Compared to GCP, GCC shares three individual pairs (GCA 1,20,21) and the same number of signals as GCA (4,8), (10,11) and (12, 13, 15). All species share homeology to GCA 1,20,21 (**Figure 3**; **Table 2**).

The syntenic block of GCA42 6 is conserved in four of the five analyzed karyotypes by painting, except for GAR (**Figure 3**; **Table 2**), in which it is divided into two signals in pairs 4 and 16, while the syntenic block 18 of GCA 42 is shared with GAR, but not with GCC or GCA 40 (**Figure 3**; **Table 2**).

DISCUSSION

Gymnotus carapo occidentalis "Catalão" has 2n = 40 (GCC, 30m/sm+10st/a), the same as *G. carapo orientalis* (GCA) 2n = 40 (28m/sm+12st/a), but with a different karyotype. It is hypothesized that the basal diploid number for Gymnotidae is 2n = 52 (da Silva et al., 2019), as the sister species *Electrophorus electricus* and *G. pantherinus* (sister species to all *Gymnotus*, Craig et al., 2019) both have 2n = 52.

This variation in karyotype can be consistently observed along the hydrographic regions. The 2n = 42 is found in the *G. c. orientalis* locations in the "Tocantins-Araguaia" region, while the 2n = 40 is observed in the *G. c. orientalis* (GCA40) located in the "Amazonica" region. While sharing the same 2n GCA40, GCC was sampled in a lake in an area close Negro river, in *G. c. occidentalis* occurrence area and also has significant karyotype differences. The 2n = 48 is found only in *G. c. madeirensis* (**Table 1**; **Figure 1**). The higher 2n = 52 and 54 is found only in *G. c. australis*, distributed along the Paraguai, "Paraná" and "Atlântico Sudeste" hydrographic regions.

When compared to all cytotypes of *G. carapo* in the literature, there is a tendency to a reduction in diploid number in the "Amazonica", "Tocantins-Araguaia" and "Atlântico Nordeste Ocidental" hydrographic regions, a trait also shared with *G. arapaima* (**Figure 1**). Whereas the cytotypes in the "Paraguai", "Paraná" and "Atlântico Sudeste" hydrographic regions have a higher chromosome number, with all locations having 2n = 54 except in one population that has 2n = 52 (**Table 1**; **Figure 1**) similar to more basal 2n in the genus. This suggests that the reduction in diploid number in the amazon region happened after colonization of the area.

Whole chromosome probes of GCA42 have been used in previous studies comparing two cytotypes of *G. carapo* (GCA42 and GCA40), *G. capanema* (GCP34) and *G. arapaima* (GAR44). The results demonstrate highly rearranged karyotypes, more than found by classical cytogenetics alone (Nagamachi et al., 2010; Nagamachi et al., 2013; Machado et al., 2018). The same results are observed in *G. carapo* "Catalão" (present study), confirming that the chromosomal evolution in this group is quite complex.

The karyotypes of the three *G. carapo* cytotypes analyzed by chromosome painting (GCA42, GCA40 and GCC40) are more similar to each other than to those of GAR44 or GCP34 and share a uniform amount of synteny. However, they have multiple species-specific rearrangements, which probably constitute a post zygotic barrier that would result in an inviable or infertile hybrid (**Figure 4**). We observe the same pattern when compared to *G. arapaima*, explained by the fact that they are sister species with recent divergence (Brochu, 2011) in relation to *G. capanema* (Craig et al., 2019).

This large number of chromosomal rearrangements in *Gymnotus carapo*, demonstrated by chromosome painting, indicates that the different cytotypes constitute a complex of cryptic species as already suggested (Milhomem et al., 2008; Nagamachi et al., 2010). The chromosomal speciation must have played a key role in this process that, if associated with small-inbred demes, could have facilitated the fixation of chromosomal rearrangements (King, 1993).

Currently these cryptic species (with the same morphology but different cytotypes) are in allopatry, which corroborates the pattern found in many Neotropical freshwater fish groups. This could be due to the dynamics of the river networks, including the fragmentation and merging of adjacent rivers, that led to increased species richness closer to the core region of the Amazon basin (Albert and Reis, 2011; Tagliacollo et al.,

2016; Albert et al., 2018a; Albert et al., 2018b; Albert et al., 2020).

The taxonomy of *G. carapo* is quite complex and has been discussed in many studies. Some authors report it as a single generalized species (Craig et al., 2017), other studies show a paraphyletic group within a monophyletic complex of related species (Brochu, 2011; da Silva et al., 2019) and still others suggest it to be a complex of cryptic species (Milhomem et al., 2008; Nagamachi et al., 2010). Within this context, some phylogenetic studies with molecular data show polytomy of the *G. carapo* complex, consisting of *G. carapo*, *G. arapaima*, and *G. ucamara* (Lovejoy et al., 2010; Brochu, 2011) and other studies show species of the subgenus *Gymnotus* nested within *G. carapo* lineages (Lehmberg et al., 2018; Craig et al., 2019; da Silva et al., 2019).

In conclusion, the results presented here support that these populations with different cytotypes of *G. carapo* analysed (*G. carapo occidentalis "Catalão"*, *G. c. orientalis* GCA42 and GCA 40, along with the geographic-specific 2n = 48 and 2n = 54) may be a cryptic species complex. Analyses by chromosome painting of more cytotypes of *G. carapo* as well as other species of this genus coupled with molecular studies of those samples could help elucidate the chromosomal evolution and pattern of speciation in the group and help identify same-species populations from endemic species that have recently diverged.

DATA AVAILABILITY STATEMENT

All data presented in this study are found in the article.

ETHICS STATEMENT

The animal study was reviewed and approved by The Ethics Committee from Para Federal University (Comitê de Ética Animal da Universidade Federal do Pará) approved this research (Permit 68/2015).

AUTHOR CONTRIBUTIONS

MM: Conceptualization; Data Curation; Formal analysis; Investigation; Methodology; Visualization; Writing original draft; Writing review and editing. MS: Investigation; Methodology; Visualization; Writing review and editing. EF: Investigation; Methodology; Funding acquisition; Visualization; Writing review and editing. PM: Investigation; Methodology; Visualization; Writing review and editing. MF-S: Investigation; Methodology; Resources; Visualization; Writing review and editing. JP: Data Curation; Formal analysis; Funding acquisition; Resources; Visualization; Writing review and editing. CN: Data Curation; Formal analysis; Funding acquisition; Project administration; Resources; Supervision; Visualization; Writing review and editing.

FUNDING

The authors thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support on a project coordinated by CN (Edital Pró-Amazônia Proc 047/2012); the FAPESPA for financial support (Edital Vale–Proc 2010/110447) and Banco Nacional de Desenvolvimento Econômico e Social–BNDES (2.318.697.0001) on a project coordinated by JP; the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support on project coordinated by EF (Edital Universal Proc 473812/2011-8). CYN (305880/2017-9), JP (305876/2017-1) and EF (301886/2019-9), are grateful to CNPq for Productivity Grants and MS (160155/2018-5) for a scholarship from CNPq. This study is part of the Doctoral Thesis in Genetic and Molecular Biology of MM who is recipient of a CNPq Doctor Scholarship.

ACKNOWLEDGMENTS

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.

REFERENCES

Albert, J. S., Craig, J. M., Tagliacollo, V. A., and Petry, P. (2018a). "Upland and lowland Fishes: a Test of the River Capture Hypothesis," in *Mountains, Climate and Biodiversity*. Editors C. Hoorn, A. Perrigo, and A. Antonelli (Chichester: Wiley Blackwell), 273–294.

Albert, J. S., and Crampton, W. G. R. (2005). "Diversity and Phylogeny of Neotropical Electric Fishes (Gymnotiformes)," in *Electroreception* (New York: Springer), 360–409. doi:10.1017/s1477200005001726

Albert, J. S. (2001). *Species Diversity and Phylogenetic Systematics of American Knifefishes (Gymnotiformes, Teleostei)*, 190. Ann Arbor, MI: Miscellaneous Publications. University of Michigan, 1–129.

Albert, J. S., Tagliacollo, V. A., and Dagosta, F. (2020). Diversification of Neotropical Freshwater Fishes. *Annu. Rev. Ecol. Evol. Syst.* 51 (1), 27–53. doi:10.1146/annurev-ecolsys-011620-031032

Albert, J. S., Val, P., and Hoorn, C. (2018b). The Changing Course of the Amazon River in the Neogene: center Stage for Neotropical Diversification. *Neotrop. Ichthyol.* 16, e180033. doi:10.1590/1982-0224-20180033

Almeida, J. S., Migues, V. H., Diniz, D., and Affonso, P. R. A. M. (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

Barby, F. F., Bertollo, L. A. C., de Oliveira, E. A., Yano, C. F., Hatanaka, T., Ráb, P., et al. (2019). Emerging Patterns of Genome Organization in Notopteridae Species (Teleostei, Osteoglossiformes) as Revealed by Zoo-FISH and Comparative Genomic Hybridization (CGH). *Sci. Rep.* 9 (1), 1112. doi:10.1038/s41598-019-38617-4

Braga, B. P. F., Flecha, R., Pena, D. S., and Kelman, J. (2008). Pacto federativo e gestão de águas. *Estud. Av.* 22 (63), 17–42. doi:10.1590/S0103-40142008000200003

Brochu, K. (2011). Molecular Phylogenetics of the Neotropical Electric Knifefish Genus Gymnotus (Gymnotidae, Teleostei): Biogeography and Signal Evolution of the Trans-andean Species. [Master dissertation]. Toronto (ON): University of Toronto.

Cardoso, A. L., Ready, J. S., Pieczarka, J. C., Milhomem, S. S. R., Figueiredo-Ready, W. M. B. d., Silva, F. H. R., et al. (2015). Chromosomal Variability between

Populations of Electrophorus Electricus Gill, 1864 (Pisces: Gymnotiformes: Gymnotidae). *Zebrafish* 12 (6), 440–447. doi:10.1089/zeb.2014.1059

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

Cioffi, M. B., Ráb, P., Ezaz, T., Bertollo, L. A. C., Lavoué, S., Oliveira, E. A., et al. (2019). Deciphering the Evolutionary History of Arowana Fishes (Teleostei, Osteoglossiformes, Osteoglossidae): Insight from Comparative Cytogenomics. *Ijms* 20, 4296. doi:10.3390/ijms20174296

Cioffi, M. d. B., Yano, C. F., Sember, A., and Bertollo, L. A. C. (2017). Chromosomal Evolution in Lower Vertebrates: Sex Chromosomes in Neotropical Fishes. *Genes* 8, 258. doi:10.3390/genes8100258

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. M., 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

Craig, J. M., Kim, L. Y., Tagliacollo, V. A., and Albert, J. S. (2019). Phylogenetic Revision of Gymnotidae (Teleostei: Gymnotiformes), with Descriptions of Six Subgenera. *Plos One* 14 (11), e0224599. doi:10.1371/journal.pone.0224599

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. (2019). Karyotypic Diversity and Evolutionary Trends in Neotropical Electric Fish of the Genus *Gymnotus* (Gymnotiformes: Gymnotidae). *Zebrafish* 16 (3), 308–320. doi:10.1089/zeb.2018.1716

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

de Freitas, N. L., Al-Rikabi, A. B. H., Bertollo, L. A. C., Ezaz, T., Yano, C. F., de Oliveira, E. A., et al. (2018). Early Stages of XY Sex Chromosomes Differentiation in the Fish *Hoplias malabaricus* (Characiformes, Erythrinidae) Revealed by DNA Repeats Accumulation. *Cg* 19 (3), 216–226. doi:10.2174/1389202918666170711160528

de 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 GenusPyrrhulina(Characiformes, Lebiasinidae). *Zebrafish* 14 (6), 536–546. doi:10.1089/zeb.2017.1465

de 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* 127 (1), 115–128. doi:10.1007/s00412-017-0648-3

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 ofGymnotus(Pisces: Gymnotiformes) from Southeastern Brazilian Basins. *Caryologia* 51, 221–234. doi:10.1080/00087114.1998.10797414

Fonteles, S. B. A., Lopes, C. E., Akama, A., Fernandes, F. M. C., Porto-Foresti, F., Senhorini, J. A., et al. (2008). Cytogenetic Characterization of the Strongly Electric Amazonian Eel, Electrophorus Electricus (Teleostei, Gymnotiformes), from the Brazilian Rivers Amazon and Araguaia. *Genet. Mol. Biol.* 31, 227–230. doi:10.1590/S1415-47572008000200010

Foresti, F., de Almeida Toledo, L. F., and de Almeida Toledo, S. (1984). Chromosome Studies inGymnotus CarapoandGymnotus Sp.(Pisces, Gymnotidae). *Caryologia* 37, 141–146. doi:10.1080/00087114.1984.10797693

Hijmans, R. J., Guarino, L., Bussink, C., Mathur, P., Cruz, M., Barrentes, I., et al. DIVA-GIS. Vsn 5.0. A Geographic Information System for the Analysis of Species Distribution Data. 2004. Available at: http://www.diva-gis.org.

J. S. Albert and R. Reis (Editors) (2011). *Historical Biogeography of Neotropical Freshwater Fishes* (Berkeley, USA: Univ of California Press). doi:10.1525/california/9780520268685.001.0001

Kim, L. Y., Crampton, W. G. R., and Albert, J. S. (2020). Two New Species of *Gymnotus* (Gymnotiformes: Gymnotidae) from Brazil and Historical Biogeography of the Subgenus Lamontianus. *Copeia* 108 (3), 468–484. doi:10.1643/ci-19-205

King, M. (1993). *Species Evolution - the Role of Chromosome Change*. Cambridge, England: Cambridge University Press, 335.

Lehmberg, E. (2015). *Biogeography of the Weakly Electric Knifefish* Gymnotus carapo *(Teleostei: Gymnotidae) in the Guianas*. Toronto: University of Toronto. Available at: https://tspace.library.utoronto.ca/bitstream/1807/69686/3/Lehmberg_Emma_S_201506_MSc_thesis.pdf (Accessed January 19, 2022).

Lehmberg, E. S., Elbassiouny, A. A., Bloom, D. D., López-Fernández, H., Crampton, W. G. R., and Lovejoy, N. R. (2018). Fish Biogeography in the "Lost World" of the Guiana Shield: Phylogeography of the Weakly Electric Knifefish *Gymnotus carapo* (Teleostei: Gymnotidae). *J. Biogeogr.* 45, 815–825. doi:10.1111/jbi.13177

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. L., 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., Pieczarka, J. C., Silva, F. H. R., O'Brien, P. C. M. O., Ferguson-Smith, M. A., and Nagamachi, C. Y. (2018). Extensive Karyotype Reorganization in the Fish *Gymnotus arapaima* (Gymnotiformes, Gymnotidae) Highlighted by Zoo-FISH Analysis. *Front. Genet.* 9 (8), 1–19. doi:10.3389/fgene.2018.00008

Machado, M. d. 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

Mago-Leccia, F. (1994). *Electric Fishes of the continental Waters of America: Classification and Catalogue of the Electric Fishes of the Order Gymnotiformes (Teleostei: Ostariophysi), with Descriptions of New Genera and Species*. Caracas, Venezuela: Fundacion para Desarrollo de las Ciencias Fiscias, Matematicas y Naturales.

Margarido, V. P., Bellafronte, E., and Moreira-Filho, O. (2007). Cytogenetic Analysis of Three sympatricGymnotus(Gymnotiformes, Gymnotidae) Species Verifies Invasive Species in the Upper Paraná River basin, Brazil. *J. Fish. Biol.* 70, 155–164. doi:10.1111/j.1095-8649.2007.01365.x

Milhomem, S. S., Pieczarka, J. C., Crampton, W. G., Silva, D. S., De Souza, A. C., 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., Crampton, W. G. R., Pieczarka, J. C., dos Santos Silva, D., 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. Fish.* 22, 485–497. doi:10.1007/s11160-011-9239-1

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., Pieczarka, J. C., Crampton, W. G. R., De Souza, A. C. P., Carvalho Jr, J. R., Jr., 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.1397.710.11646/zootaxa.1397.1.7

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

Nagamachi, C. Y., Pieczarka, J. C., Milhomem, S. S., O'Brien, P. C., de Souza, A. C., 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

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 betweenGymnotus capanemaandGymnotus carapo(Gymnotidae, Gymnotiformes). *Cytogenet. Genome Res.* 141, 163–168. doi:10.1159/000354988

Ráb, P., Rábová, M., Pereira, C. S., Collares-Pereira, M. J., and Pelikánová, Š. (2008). Chromosome Studies of European Cyprinid Fishes: Interspecific Homology of Leuciscine Cytotaxonomic Marker-The Largest Subtelocentric Chromosome Pair as Revealed by Cross-Species Painting. *Chromosome Res.* 16, 863–873. doi:10.1007/s10577-008-1245-3

R. Fricke, W. N. Eschmeyer, and R. van der Laan (Editors) (2021). *Eschmeyer's Catalog of Fishes: Genera, Species, References*. (San Francisco, United States: California Academy of Sciences). Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp.

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. *Ccg* 5, 223–235. doi:10.3897/compcytogen.v5i3.1375

Scudeler, P. E. S., Diniz, D., Wasko, A. P., Oliveira, C., and Foresti, F. (2015). Whole Chromosome Painting of B Chromosomes of the Red-Eye Tetra *Moenkhausia sanctaefilomenae* (Teleostei, Characidae). *Ccg* 9 (4), 661–669. doi:10.3897/CompCytogen.v9i4.5460

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

Utsunomia, R., Melo, S., Scacchetti, P. C., Oliveira, C., Machado, M. d. A., Pieczarka, J. C., et al. (2018). Particular Chromosomal Distribution of Microsatellites in Five Species of the Genus *Gymnotus* (Teleostei, Gymnotiformes). *Zebrafish* 15 (4), 398–403. doi:10.1089/zeb.2018.1570

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

Utsunomia, R., Silva, D. M. Z. d. 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 (3), e0150573. doi:10.1371/journal.pone.0150573

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

Yano, C. F., Bertollo, L. A. C., 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 (3), 276–283. doi:10.1038/hdy.2016.83

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Genetics

BRIEF RESEARCH REPORT
published: 28 March 2022
doi: 10.3389/fgene.2022.855718

Genomic Resources for *Salminus brasiliensis*

Raissa Cristina Dias Graciano[1], *Rafael Sachetto Oliveira*[2], *Isllas Miguel Santos*[3] *and Gabriel M. Yazbeck*[1,3]*

[1]*Laboratório de Recursos Genéticos, Programa de Pós Graduação Em Biotecnologia, Universidade Federal de São João Del Rei, São João Del Rei, Brazil,* [2]*Departamento de Ciência da Computação, Universidade Federal de São João Del Rei, São João Del Rei, Brazil,* [3]*Laboratório de Recursos Genéticos, Departamento de Zootecnia, Universidade Federal de São João Del Rei, São João Del Rei, Brazil*

The Neotropical region bears the most diverse freshwater fish fauna on the planet and is the stage for dramatic conservation struggles. Initiatives aiming for conservation of a single emblematic fish, a flagship species, to which different onlookers relate on a cultural/personal level, holds promise towards engagement and conservation actions benefiting whole biological communities and ecosystems. Here, we present the first comprehensive genomic resources for *Salminus brasiliensis*, a potential flagship Neotropical species. This fish faces pressing conservation issues, as well as taxonomic uncertainty, being a main species relevant to angling and commercial fisheries. We make available 178 million Illumina paired-end reads, 90 bases long, comprising 16 Gb (≈15X coverage) of filtered data, obtained from a primary genomic library of 500-bp fragments. We present the first *de novo* genomic assembly for *S. brasiliensis*, with ~1 Gb (N_{50} = 10,889), as well as the coding genome annotation of 12,962 putative genes from assembled genomic fragments over 10 kb, most of which could be identified from the Ostariophysi GenBank database. We also provide a genome-wide panel for more than 80,000 predicted microsatellite loci for low-cost, fast and abundant DNA marker development for this species. A total of 47, among 52 candidates, empirically assayed microsatellites were confirmed as polymorphic in this fish. All genomic data produced for *S. brasiliensis* is hereby made publicly accessible. With the disclosure of these results, we intend to foster general biology studies and to provide tools to be applied immediately in conservation and aquaculture in this candidate flagship Neotropical species.

Keywords: bioinformatics, sequencing-by-synthesis, population genetics resources, hatchery, environmental management, characiformes

Edited by:
Tony Silveira,
Federal University of Rio Grande, Brazil

Reviewed by:
Eric M. Hallerman,
Virginia Tech, United States
Carolina Machado,
Federal University of São Carlos, Brazil

***Correspondence:**
Gabriel M. Yazbeck
dna@ufsj.edu.br

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

Received: *15 January 2022*
Accepted: *18 February 2022*
Published: *28 March 2022*

Citation:
Graciano RCD, Oliveira RS, Santos IM and Yazbeck GM (2022) Genomic Resources for Salminus brasiliensis. Front. Genet. 13:855718. doi: 10.3389/fgene.2022.855718

INTRODUCTION

The Neotropical region bears the most diverse freshwater fish fauna on the planet, apparently as the result of rapid speciation process (Melo et al., 2021). Freshwater systems and fish biodiversity face ever-increasing impacts from human activity, a global issue that often does not muster as much societal awareness (Su et al., 2021) as other pressing conservation challenges such as accelerated rates of deforestation and normative deregulation (Gonçalves et al., 2020; Gonçalves-Souza et al., 2021). Associated with its particularly rich fish fauna, South America harbors some of the largest river basins on Earth and it is the stage for dramatic conservation struggles involving conflicting stakeholder concerns. The degree of cumulative anthropogenic degradation of its aquatic

FIGURE 1 | *Salminus brasiliensis* in its natural habitat, in Bonito, MS, Brazil (Author: André Seale).

systems is worrisome, as epitomized by two catastrophic mine-waste spills caused by obsolete tailing dam failures in Brazil in 2015 (Fernandes et al., 2016) and 2019 (Vergilio et al., 2020), leading to severe effects upon freshwater biodiversity (e.g., Gomes et al., 2019), in addition to social disruption and loss of human life (Polignano and Lemos, 2020).

Conservation initiatives aiming a single emblematic species, to which large numbers of diverse onlookers can relate on a personal or social level, holds promise towards engagement and successful actions for the benefit of whole biological communities and ecosystems (e.g., Leader-Williams et al., 1990; Dietz et al., 1994). For different taxonomic groups and geographic locations, some more well-known, culturally or economically important species can act as flagship species (e.g., Ochieng et al. 2021; Preston et al., 2021). Still, among those characterized as the 20 most charismatic species, none is a fin fish (Albert et al., 2021), and a recent assessment in Brazil noted only one Amazonian teleost fish, *Araipamas gigas*, as a flagship species, among 62 elected taxa (Wosnick et al., 2021). In this regard, the singular Neotropical migratory fish fauna, also known as *piracema* fish (Carolsfeld et al., 2004), are under-represented, especially given their prominent social-environmental importance, fragile conservation status and general public awareness.

One particular migratory species arguably fits the flagship bill better than most: *Salminus brasiliensis* Cuvier 1816 (Characiformes, Bryconidae), also known as *dourado*, *dorado*, *pirayú*, *saipé*, *pirajuba*, *dama* and *picudo* among other common names throughout South America (Froese and Pauley, 2000). It is an emblematic, large (females may reach as much as 26 kg), long-distance swimmer, top-predator with formidable aesthetic appeal (**Figure 1**). It also enjoys wide recognition from the general public and has established importance in angling, artisanal and commercial fisheries (Feitosa et al., 2004; Gagne et al., 2017). It commonly occurs in the La Plata River Basin, a vast hydrographic system which spans the Paraná, Paraguay and Uruguay Rivers and also in the Patos Lagoon in southern Brazil, being native to a total of five countries, including Argentina, Bolivia, Paraguay and Uruguay. It also has been targeted for aquaculture due to its favorable growth rate, size and gastronomic quality (Crescêncio et al., 2005; Zaniboni-Filho and Schulz, 2003; Zaniboni-Filho et al. 2017), despite being a voracious piscivorous species, which could pose challenges to commercial cultivation. *Salminus brasiliensis* is spawned mainly for stock supplementation purposes, along efforts to mitigate the environmental impact of hydroelectric dams upon migratory fish. It has suffered a sharp reduction of natural stocks also because of removal of the riparian zone along rivers, land and water pollution, introduction of invasive species and overfishing. The species has not yet been included, however, in the Red Book of Threatened Species (Rosa and Lima, 2008), mainly given its relatively high abundance in the Pantanal region of the Paraguay Basin. Yet, it has been considered vulnerable in the Paraná Basin (Marques et al., 2002; Abilhoa, 2004) and highly vulnerable in the La Plata River (Zayas and Cordiviola, 2007). It also has been classified as virtually extinct, with population size below a viable threshold, in some major Paraná Basin locations, such as the Grande, Tietê and Paranapanema Rivers (Rosa and Lima, 2008).

Rosso et al. (2012) observed the differentiation among *S. brasiliensis* stocks from Argentina's Pampa Plain and other parts of South America. According to Machado et al. (2017), the species displays distinct haplogroups, with the Upper Paraná containing a distinct Evolutionary Significant Unit from the remaining La Plata River Basin. Thus, there are taxonomic uncertainties associated with this species and further work is needed to inform its conservation and management plans.

Genetic and genomic resources can be applied to resolve conservation and taxonomic issues. So far, eight polymorphic microsatellites were described for this species by Rueda et al. (2011) and another 47 by our own group (*in* Cao et al., 2016) from next-generation sequencing (NGS) data first fully disclosed herein. Brandão-Dias et al. (2014) provided the first cytoplasmic genomic resources for *S. brasiliensis* with publication of its assembled and annotated mitogenome (17,721 bp long). Here, we present the first comprehensive genomic resources for *Salminus brasiliensis*, contributing to basic molecular genetics knowledge, and to resolution of conservation and taxonomic challenges.

MATERIALS AND METHODS

DNA Extraction and Sequencing

The genomic DNA sample from *S. brasiliensis* was obtained from a single female specimen from the Itutinga Hatchery at Grande River, in the Upper Paraná Basin, located in Minas Gerais–Brazil (21°17′040″S, 44°37′023″W). Following euthanasia, the fish was held on in ice during transport to the laboratory, where a muscle tissue sample was removed and placed in liquid nitrogen until DNA extraction. The individual was kept as a voucher (LARGE130547[1]). All

[1]https://www.ufsj.edu.br/large.

procedures were authorized by UFSJ's Ethics Committee on Animal Research (CEUA-UFSJ); specimen collection was conducted under SISBIO license number 37222 and genetic patrimony access was granted through license CGEN - A9D0E51. No experimentation on live animals was conducted in this research. DNA extraction, library preparation and NGS steps were conducted as presented in Yazbeck et al. (2018). In summary, a single primary library of random genomic fragments (~500 bp) was sequenced using an Illumina HiSeq 2000, generating 90-bp paired-end short reads. Raw data was filtered for quality and removal of duplicates and stored as two corresponding FASTQ files.

Bioinformatics

We verified FASTQ quality values with FastQC (ver. 0.11.09–Andrews, 2010) and we used KmerGenie (Chikhi and Medvedev, 2014) to assess *k* values from 41 to 61, in order to search which *k*-mer would result in a more diverse dataset. Then, different preliminary *de novo* genomic assemblies were attempted from the generated short reads, assaying best *k*-mer values with Minia (Chikhi and Rizk, 2012). The chosen *k*-mer was used for a *de novo* assembly using SOAPdenovo 2 (Luo et al., 2012). Assembly qualities were inferred by search for single-copy conserved core eukaryotic ortholog genes with BUSCO (Benchmarking universal Single-Copy Orthologs) (Simão et al., 2015), using the Zebrafish, *Danio rerio*, database (BuscoDB). The procedure to map perfect-repeat microsatellite loci (potential candidates for new molecular markers) in the chosen assembly was also conducted according to the methodology detailed in Yazbeck et al. (2018). In summary, an unknown genomic assembly performed with these short reads was used for potential microsatellite loci mining by a service provider. We used the data produced by this step, namely predicted PCR products, as queries to be searched in the new chosen assembly, using BLAST (Boratyn et al., 2013). The filtered short reads were mapped onto the chosen assembly using the SOAP aligner, and the resulting file was converted into BAM format with the aid of SAMtools (Li et al., 2009).

Based on an estimate of the average eukaryotic gene sequence length of 12,000 bp (Cooper and Hausman, 2016), due to the fragmented nature of the assemblies and computational power restrictions, we proceeded to perform functional annotation of contigs and scaffolds larger than 10,000 bases only, using the MAKER pipeline (Cantarel et al., 2008) using the Characiformes database[2] (peptide sequences and expressed sequence tags) from NCBI as a training set. Annotated sequences were then characterized by BLAST, from the superorder Ostariophysi data available from NCBI[3] and functionally classified with the aid of PANTHER (Mi et al., 2010). The mitogenome was obtained according to the steps presented in Yazbeck et al. (2014).

[2]https://www.ncbi.nlm.nih.gov/protein/?term=txid7991%5bOrganism:exp%5d.

[3]https://www.ncbi.nlm.nih.gov/nuccore/?term=txid32519[Organism:exp].

TABLE 1 | Results from the search of a set of 303 core eukaryotic orthologous genes in alternative assemblies performed with different *k*-mer values, according to BUSCO.

Gene	*k* = 41	*k* = 47	*k* = 55
Complete	98	194	58
Complete and single-copy	96	191	57
Complete and duplicated	2	3	1
Fragmented	148	80	113
Missing	57	29	132

RESULTS

The library sequencing resulted in a pair of twin FASTQ files consisting of 90-bp paired-end reads amounting to around 16 Gb of filtered data, with 97.4% exhibiting PHRED *Q* quality values of 20 or more (1% or less of wrong base calls). The average quality value per read was *Q* = 38 and CG content was 41.05%. This data is now available at NCBI's Sequence Read Archive (SRA) (Supplementary Material S1-under accession number PRJNA792751).

Even though *k*-mer analysis indicated *k* = 41 as a best candidate, three different assembly attempts were assayed, with *k* = 41, 47 and 55, using MINIA and SOAPdenovo 2. **Table 1** shows the number of complete core eukaryotic genes retrieved from each assembly, with the *k* = 41 assembly performing less well than *k* = 47 (98 and 194 complete genes, respectively). Both attempts performed with SOAPdenovo 2 (*k* = 47 and 55) resulted in 0.92 and 0.96 Gb assemblies (**Table 2**), with the former presenting a superior N_{50} value (10,889). Thus, the *k* = 47 assembly was selected for the annotation process and used as the basis for microsatellite mapping. It consisted of 560,090 contigs or scaffolds varying from 100 to 134135 bases (average 1,643) (Supplementary Material S2). Only 28 scaffolds were larger than 100 kb, 24,106 were larger than 10 kb, and 115,292 larger than 1 kb. This assembly showed a CG content of 41.25%.

An undisclosed assembly with these short reads by a service provider, followed by microsatellite search, revealed 86,832 potential microsatellite loci. These loci were searched with BLAST using the predicted PCR products and mapped in our new chosen *k* = 47 assembly, and the results are available in Supplementary Table S1. Around 76% of predicted loci could be found with a 100% match (66,206 microsatellites), another ~18% (14,527) were found with similarity of 90% or above and ~0.9% 742) with similarity within the 77–89% range. Around 6% of microsatellite loci determined from the unknown assembly (5,357) could not be mapped back on the assembly we utilized, including 13 loci empirically validated (out of 47) as polymorphic markers (Cao et al., 2016).

The MAKER prediction pipeline resulted in the annotation of 12,962 genes from around 485 Mb of analysed sequences from *S. brasiliensis*, with gene, exon and intron average sizes of 6,125, 194 and 760 bp, respectively. These data were consolidated as a general feature format file (GFF3), with detailed description of coding sequences (CDS), hereby made available through FigShare (Supplementary Material S3). The genes found then were searched through BLAST against the Ostariophysi databank from NCBI, allowing the characterization of 10,211 genes (Supplementary Table S2), which were functionally classified according to Gene Ontology

TABLE 2 | Description of the *S. brasiliensis* genomic assemblies performed with SOAPdenovo 2.

k-mer size	Genome size (bp-including N*)	Genome size (not including N*)	N_{50}	Number of scaffolds
k = 55	955,747,686	913,685,574	10,567	1,026,499
k = 47	920,486,504	874,381,160	10,889	560,090

*N = *unidentified base calls.*

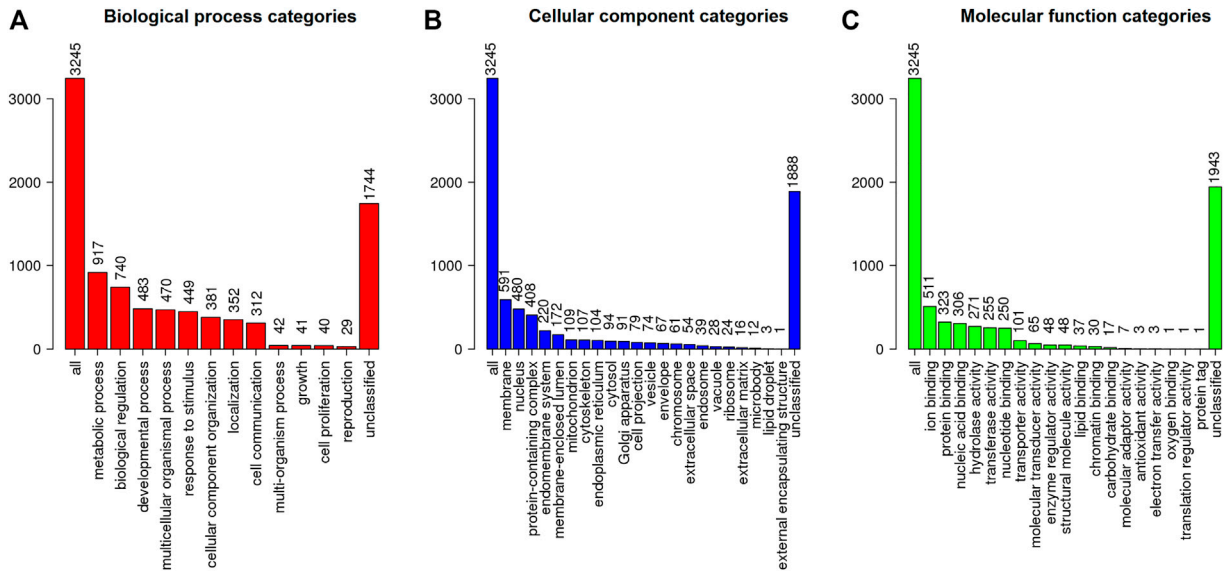


FIGURE 2 | Functional classification (according to Gene Ontology) regarding the **(A)** biological process, **(B)** cellular component, and **(C)** molecular function for predicted genes from *S. brasiliensis*.

regarding their potential biological role, resulting in the elucidation of function of 3,245 *S. brasiliensis* genes (**Figure 2**).

The sequenced individual's mitogenome is deposited in GenBank under accession KY825190 and amounts to 17,998 bp.

DISCUSSION

Genomics can produce powerful tools for fish conservation (Bernos et al., 2020), and we here provide the first comprehensive genomic data on *S. brasiliensis*. This study generated a massive set of short DNA reads with associated base-call quality scores, which is now publicly available at NCBI's Sequence Read Archive, allowing its application by other research groups. This data set, even though of low genomic coverage, is considered of high quality (average of one error per each 6,250 bp). It allowed the exploration of different *de novo* genomic assemblies, which showed that the adoption of different *k*-mer parameters resolve different genomic regions, as judged by complete core eukaryotic genes analysis. We make available one assembly (*k* = 47), while very fragmented, as shown by its N_{50} and short contigs/scaffolds, it still allowed the annotation of over 12,000 protein-encoding sequences, whereas only 13 mitochondrial genes were previously available (Brandão-Dias et al., 2014). Among arbitrary examples of fully annotated genes from this assembly is the *Leucine rich repeat containing 10* gene (LRRC10) involved in heart-tissue regeneration in *Astyanax mexicanus* (Stockdale et al., 2018) and a gonadotrophin-releasing hormone (GnRH) gene, which could support biotechnological development for induced hatchery spawning of this species, as it is involved in gonadal maturation. This functional annotation is made available as a GFF3 file. The number of annotated genes found (almost 13,000) leads us to believe that the presented assembly resolves approximately 40–50% of this species' genome, judging from comparison with completely characterized genomes from other fishes (*e.g.* 30,741 coding sequences from *Danio rerio* - Howe et al., 2013).

We also provided a panel of tens of thousands of candidate microsatellite loci, which can serve as a starting point for fast and inexpensive molecular marker development in this species, allowing development of marker loci for stock delimitation, linkage mapping, QTL characterization and association studies. Previously, Cao et al. (2016) validated 47 polymorphic microsatellite loci from this panel, attesting to its utility for efficient molecular marker development. Some microsatellite loci determined from a genomic assembly produced with the short reads presented here, which however could not be traced back to the *k* = 47 assembly, were among these empirically validated markers. We thus vouch for the retention of the non-mapped loci, since it has been shown to produce practical

results. There is inherent difficulty in resolving repetitive regions with short reads, which explains the varying results from assembly to assembly. Furthermore, among the more than 5,000 loci not found, it can be shown with other alignment tools (*e.g.* Swipe–Rognes, 2011), that many microsatellite loci are indeed present in different repetition resolutions, searching and matching primer pairs (results not shown), since BLAST has limited power working with low-complexity sequences, such as those present in microsatellite motifs.

The mitogenome of the sequenced specimen showed a larger size (277 bp) than the one previously published (Brandão-Dias et al., 2014), with extra DNA at the D-loop region, allowing rapid identification of putative variable sites to be explored in population genetics and other studies.

Together, the results presented here have the potential to be applied in conservation initiatives and taxonomic investigation of this important large fish, a candidate Neotropical migratory flagship species, aiding resolution of ongoing scientific and environmental issues faced by this organism. It will certainly aid future telomere-to-telomere genome characterization of *S. brasiliensis*.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/sra/, PRJNA792751, SRR17407720; https://www.ncbi.nlm.nih.gov/sra/, SRR17486808; https://figshare.com/, 10.6084/m9.figshare.18131495; https://figshare.com/, 10.6084/m9.figshare.11796468; https://figshare.com/, 10.6084/m9.figshare.17754188; https://figshare.com/, 10.6084/m9.figshare.19169756.

ETHICS STATEMENT

The animal study was reviewed and approved by the Comissão de Ética no Uso de Animais, Universidade Federal de São João Del Rei.

AUTHOR CONTRIBUTIONS

RG processed and analyzed NGS data and results, planned and conducted bioinformatic analysis, conducted quality control of data and results and wrote the manuscript. RO mobilized computational infrastructure, developed pipeline scripts, planned and conducted bioinformatic analysis. IS supported and conducted bioinformatic analysis. GY supervised the project, planned the NGS approach, secured funding, directed DNA extraction, conducted bioinformatic analysis, analyzed results and co-wrote the manuscript. All authors have approved the final version of the manuscript.

FUNDING

This work was developed with resources from The Brazilian National Council for Scientific and Technological Development (CNPq-308124/2020-0), *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES - Code 001), *Companhia Energética de Minas Gerais* and The Brazilian Electricity Regulatory Agency (CEMIG-ANEEL - GT345) and Minas Gerais State Agency for Research and Development (FAPEMIG - APQ-04569-10).

ACKNOWLEDGMENTS

We thank UFSJ (Pró Reitoria de Pesquisa e Pós Graduação, Programa de Pós Graduação em Biotecnologia and Programa de Pós Graduação em Ciência da Computação), CNPq, FAPEMIG and CAPES. We also thank José Mauro Ribeiro for conducting mtDNA assembly and Fausto Moreira da Silva Carmo for helping in collecting and conducting DNA extraction from the sequenced specimen. We also acknowledge André and Lucia Seale for their courtesy in providing a specimen photograph.

REFERENCES

Abilhoa, V. E. L. F. (2004). "Duboc. Peixes," in *Livro vermelho da fauna ameaçada no Estado do Paraná*. Editors S. B. Mikich and R. S. Bérnils (Curitiba: Instituto Ambiental do Paraná), 581–677.

Albert, J. S., Destouni, G., Duke-Sylvester, S. M., Magurran, A. E., Oberdorff, T., Reis, R. E., et al. (2021). Scientists' Warning to Humanity on the Freshwater Biodiversity Crisis. *Ambio* 50, 85–94. doi:10.1007/s13280-020-01318-8

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc (Accessed January 29, 2022).

Bernos, T. A., Jeffries, K. M., and Mandrak, N. E. (2020). Linking Genomics and Fish Conservation Decision Making: a Review. *Rev. Fish. Biol. Fish.* 30, 587–604. doi:10.1007/s11160-020-09618-8

Boratyn, G. M., Camacho, C., Cooper, P. S., Coulouris, G., Fong, A., Ma, N., et al. (2013). BLAST: a More Efficient Report with Usability Improvements. *Nucleic Acids Res.* 41, W29–W33. doi:10.1093/nar/gkt282

Brandão-Dias, P. F. P., Carmo, A. O. d., Martins, A. P. V., Pimenta, R. J. G., Alves, C. B. M., and Kalapothakis, E. (2014). Complete Mitochondrial Genome of *Salminus brasiliensis* (Characiformes, Characidae). Mitochondrial DNA, 1–2. doi:10.3109/19401736.2014.958676

Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., et al. (2008). MAKER: An Easy-To-Use Annotation Pipeline Designed for Emerging Model Organism Genomes. *Genome Res.* 18, 188–196. doi:10.1101/gr.6743907

Cao, Y.-L., Caputo, L. I., Cheng, H., da Silva Carmo, F. M., de Carvalho, L. C., Yazbeck, G. M., et al. (2016). Microsatellite Records for Volume 8, Issue 3. *Conservation Genet. Resour.* 8, 359–370. doi:10.1007/s12686-016-0581-4

Carolsfeld, J., Harvey, B., Ross, C., and Anton, B. (2004). *Migratory Fishes of South America: Biology, Fisheries and Conservation Status*. Ottawa: IDRC and the World Bank.

Chikhi, R., and Medvedev, P. (2014). Informed and Automated K-Mer Size Selection for Genome Assembly. *Bioinformatics* 30, 31–37. doi:10.1093/bioinformatics/btt310

Chikhi, R., and Rizk, G. (2012). "Space-efficient and exact de Bruijn graph representation based on a Bloom filter," in *Algorithms in Bioinformatics Lecture Notes in Computer Science*. Editors B. Raphael and J. Tang (Berlin, Heidelberg: Springer), 236–248. doi:10.1007/978-3-642-33122-0_19

Cooper, G. M., and Hausman, R. E. (2016). *A Célula - 3ed: Uma Abordagem Molecular*. Porto Alegre: Artmed.

Crescêncio, R., Ituassú, D. R., Roubach, R., Pereira Filho, M., Cavero, B. A. S., and Gandra, A. L. (2005). Influência Do período de alimentação no consumo e

ganho de peso Do pirarucu. *Pesq. Agropec. Bras.* 40, 1217–1222. doi:10.1590/S0100-204X2005001200009

Dietz, J. M., Dietz, L. A., and Nagagata, E. Y. (1994). "The Effective Use of Flagship Species for Conservation of Biodiversity: the Example of Lion Tamarins in Brazil," in *Creative Conservation*. Editors P. J. S. Olney, G. M. Mace, and A. T. C. Feistner (Dordrecht: Springer Netherlands), 32–49. doi:10.1007/978-94-011-0721-1_2

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. *Natureza & Conservação* 14, 35–45. doi:10.1016/j.ncon.2016.10.003

Froese, R., and Pauley, D. (2000). *FishBase 2000: Concepts, Design and Data Sources*. Los Baños, Laguna, Philippines: ICLARM.

Gagne, T. O., Ovitz, K. L., Griffin, L. P., Brownscombe, J. W., Cooke, S. J., and Danylchuk, A. J. (2017). Evaluating the Consequences of Catch-And-Release Recreational Angling on golden Dorado (Salminus Brasiliensis) in Salta, Argentina. *Fish. Res.* 186, 625–633. doi:10.1016/j.fishres.2016.07.012

Gomes, L. C., Chippari-Gomes, A. R., Miranda, T. O., Pereira, T. M., Merçon, J., Davel, V. C., et al. (2019). Genotoxicity Effects on Geophagus Brasiliensis Fish Exposed to Doce River Water after the Environmental Disaster in the City of Mariana, MG, Brazil. *Braz. J. Biol.* 79, 659–664. doi:10.1590/1519-6984.188086

Gomes, L. C., Feitosa, L. A., Fernandes, R., Da Costa, R. S., and Agostinho, A. A. (2004). Parâmetros populacionais e simulação Do rendimento por recruta de Salminus brasiliensis (Cuvier, 1816) Do alto rio Paraná. *Acta Sci. Biol. Sci.* 26, 317–323. doi:10.4025/actascibiolsci.v26i3.1593

Gonçalves-Souza, D., Vilela, B., Phalan, B., and Dobrovolski, R. (2021). The Role of Protected Areas in Maintaining Natural Vegetation in Brazil. *Sci. Adv.* 7, eabh2932. doi:10.1126/sciadv.abh2932

Gonçalves, P. R., Di Dario, F., Petry, A. C., Martins, R. L., da Fonseca, R. N., Henry, M. D., et al. (2020). Brazil Undermines parks by Relocating Staff. *Science* 368, 1199. doi:10.1126/science.abc8297

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

Leader-Williams, N., Harrison, J., and Green, M. J. B. (1990). Designing Protected Areas to Conserve Natural Resources. *Sci. Prog.* 74, 189–204.

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

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., et al. (2012). SOAPdenovo2: an Empirically Improved Memory-Efficient Short-Read De Novo Assembler. *GigaSci* 1, 18. doi:10.1186/2047-217X-1-1810.1186/2047-217X-1-18

Machado, C. D. B., Ishizuka, T. K., Freitas, P. D. D., Valiati, V. H., and Galetti, P. M. (2017). DNA Barcoding Reveals Taxonomic Uncertainty in Salminus (Characiformes). *Syst. Biodiversity* 15, 372–382. doi:10.1080/14772000.2016.1254390

Marques, A. A. B., Fontana, C. S., Vélez, E., Bencke, G. A., Schneider, M., and dos Reis, R. E. (2002). Lista das Espécies da Fauna Ameaçadas de Extinção no Rio Grande Do Sul. Porto Alegre: FZB/MCT–PUCRS/PANGEA, 52p. (Publicações Avulsas FZB, no11). Available at: http://www.tecniflora.com.br/fauna_ameacada_RS.pdf (Accessed January 29, 2022).

Melo, B. F., Sidlauskas, B. L., Near, T. J., Roxo, F. F., Ghezelayagh, A., Ochoa, L. E., et al. (2021). Accelerated Diversification Explains the Exceptional Species Richness of Tropical Characoid Fishes. *Syst. Biol.* 71, 78–92. doi:10.1093/sysbio/syab040

Mi, H., Dong, Q., Muruganujan, A., Gaudet, P., Lewis, S., and Thomas, P. D. (2010). PANTHER Version 7: Improved Phylogenetic Trees, Orthologs and Collaboration with the Gene Ontology Consortium. *Nucleic Acids Res.* 38, D204–D210. doi:10.1093/nar/gkp1019

Polignano, M. V., and Lemos, R. S. (2020). Rompimento da barragem da Vale em Brumadinho: impactos socioambientais na Bacia Do Rio Paraopeba. *Cienc. Cult.* 72, 37–43. doi:10.21800/2317-66602020000200011

Preston, S. D., Liao, J. D., Toombs, T. P., Romero-Canyas, R., Speiser, J., and Seifert, C. M. (2021). A Case Study of a Conservation Flagship Species: the Monarch Butterfly. *Biodivers Conserv* 30, 2057–2077. doi:10.1007/s10531-021-02183-x

Rognes, T. (2011). Faster Smith-Waterman Database Searches with Inter-sequence SIMD Parallelisation. *BMC Bioinformatics* 12, 221. doi:10.1186/1471-2105-12-221

Rosa, R. S., and Lima, F. C. (2008). Os peixes brasileiros ameaçados de extinção. *Livro vermelho da fauna brasileira ameaçada de extinção* 2, 9–275.

Rosso, J. J., Mabragaña, E., González Castro, M., and Díaz de Astarloa, J. M. (2012). DNA Barcoding N Eotropical Fishes: Recent Advances from the P Ampa P Lain, A Rgentina. *Mol. Ecol. Resour.* 12, 999–1011. doi:10.1111/1755-0998.12010

Rueda, E. C., Amavet, P., Brancolini, F., Sommer, J., and Ortí, G. (2011). Isolation and Characterization of Eight Polymorphic Microsatellite Markers for the Migratory Characiform Fish, Salminus Brasiliensis. *J. Fish Biol.* 79, 1370–1375. doi:10.1111/j.1095-8649.2011.03109.x

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs. *Bioinformatics* 31, 3210–3212. doi:10.1093/bioinformatics/btv351

Stockdale, W. T., Lemieux, M. E., Killen, A. C., Zhao, J., Hu, Z., Riepsaame, J., et al. (2018). Heart Regeneration in the Mexican Cavefish. *Cell Rep.* 25, 1997–2007. e7. doi:10.1016/j.celrep.2018.10.072

Su, G., Logez, M., Xu, J., Tao, S., Villéger, S., and Brosse, S. (2021). Human Impacts on Global Freshwater Fish Biodiversity. *Science* 371, 835–838. doi:10.1126/science.abd3369

Tobias Ochieng, N., Elizabeth, K. N., and Nigel, L.-W. (2021). Measuring the Conservation Attitudes of Local Communities towards the African Elephant *Loxodonta africana*, a Flagship Species in the Mara Ecosystem. *PLOS ONE* 16, e0253234. doi:10.1371/journal.pone.0253234

Vergilio, C. d. S., Lacerda, D., Oliveira, B. C. V. d., Sartori, E., Campos, G. M., Pereira, A. L. d. S., et al. (2020). Metal Concentrations and Biological Effects from One of the Largest Mining Disasters in the World (Brumadinho, Minas Gerais, Brazil). *Sci. Rep.* 10, 5936. doi:10.1038/s41598-020-62700-w

Wosnick, N., Leite, R. D., Giareta, E. P., Nunes, A. R. O. P., Nunes, J. L. S., Charvet, P., et al. (2021). Evaluating Conservation Status and Governmental Efforts towards Regional Flagship Species in Brazil. *J. Environ. Manage.* 292, 112732. doi:10.1016/j.jenvman.2021.112732

Yazbeck, G. M., Oliveira, R. S., Ribeiro, J. M., Carmo, F. M. S., and Carvalho, M. B. (2016). First complete mitochondrial genome for any anostomid fish: *Leporinus piavussu*, a recently described piracema species. *Mitochondrial DNA Part A* 27, 2293–2294. doi:10.3109/19401736.2014.987239

Yazbeck, G. M., Oliveira, R. S., Ribeiro, J. M., Graciano, R. D., Santos, R. P., Carmo, F. M. S., et al. (2018). A Broad Genomic Panel of Microsatellite Loci from *Brycon orbignyanus* (Characiformes: Bryconidae) an Endangered Migratory Neotropical Fish. *Sci. Rep.* 8, 8511. doi:10.1038/s41598-018-26623-x

Zaniboni-Filho, E., Ribolli, J., Hermes-Silva, S., and Nuñer, A. P. O. (2017). Wide Reproductive Period of a Long-Distance Migratory Fish in a Subtropical River, Brazil. *Neotrop. Ichthyol.* 15, 1. doi:10.1590/1982-0224-20160135

Zaniboni-Filho, E., and Schulz, U. H. (2003). "Migratory Fishes of the Uruguay River," in *Migratory Fishes of the South America: Biology, Social Importance and Conservation Status* (Victoria: World Fisheries Trust), 157–194.

Zayas, M. A., and Cordiviola, E. (2007). The conservation state of Characidaefish (Pisces: Characiformes) in anareaofthe Plata basin, Argentina estado de conservación de pecesCharacidae (Pisces: Characiformes) enunárea de lacuenca del plata, argentina. *Gayana* 71 (2), 178–186. doi:10.4067/s0717-65382007000200006

Conflict of Interest: GY filled for a patent request involving certain combinations of specific microsatellites, from the set previously described in Cao et al. (2016), at INPI (*Instituto Nacional da Propriedade Industrial*), Brazil (BR1020180128558), to be used as PCR multiplexes in commercial genotyping operations. The remaining 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 19 May 2022
doi: 10.3389/fgene.2022.854362

Contrasting Patterns of Genetic Diversity and Divergence Between Landlocked and Migratory Populations of Fish *Galaxias maculatus*, Evaluated Through Mitochondrial DNA Sequencing and Nuclear DNA Microsatellites

Marcela P. Astorga[1], *Andrea Valenzuela*[1], *Nicolás I. Segovia*[2,3], *Elie Poulin*[3,4], *Luis Vargas-Chacoff*[3,5,6] *and Claudio A. González-Wevar*[3,5,6]*

[1]*Instituto de Acuicultura, Universidad Austral de Chile, Puerto Montt, Chile,* [2]*Departamento de Biología Marina, Facultad de Ciencias del Mar, Universidad Católica Del Norte, Coquimbo, Chile,* [3]*Instituto Milenio Biodiversidad de Ecosistemas Antárticos y Sub-Antárticos (BASE), Santiago, Chile,* [4]*Departamento de Ciencias Ecológicas, Facultad de Ciencias, Universidad de Chile, Santiago, Chile,* [5]*Instituto de Ciencias Marinas y Limnológicas (ICML), Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile,* [6]*Centro FONDAP de Investigaciones en Dinámica de Ecosistemas Marinos de Altas Latitudes (IDEAL), Universidad Austral de Chile, Valdivia, Chile*

Galaxias species are interesting biogeographic models due to their distribution and different types of life cycles, with migratory and landlocked populations. To obtain a better understanding of the genetic consequences of the Quaternary glacial cycles in *Galaxias maculatus*, in this work we compared landlocked and migratory populations collected in areas that were differentially affected by ice advances and retreats. We included nine populations of *G. maculatus*, four collected from lakes (landlocked) and five from their associated estuaries/rivers (migratory) in three estuary-lake systems across southern Chile. Genetic analyses were performed using the mitochondrial control region and nine microsatellite loci. Genetic diversity measured with both markers was significantly higher in migratory than in landlocked populations across the study area. The levels of genetic differentiation showed higher differentiation among lakes than estuaries. Genetic diversity was higher in migratory populations located in areas that were less impacted by ice during Quaternary glacial processes. These results may be the consequence of recent recolonization of small freshwater bodies following the Last Glacial Maximum (LGM). Finally, the greatest differentiation was observed in populations that were exposed to continental ice advances and retreats during the LGM. Thus, in the present work we corroborate a pattern of differentiation between lakes and estuaries, using mtDNA sequences and microsatellite nuclear markers. This pattern may be due to a combination of biological factors, i.e., resident non-migratory behaviour or landlocking and natal homing-in, as well as geological factors, i.e., Expansion-Contraction Quaternary glacial biogeographic processes.

Keywords: mtDNA, Chile, estuary, lake, last glacial maximum

OPEN ACCESS

Edited by:
Sandra Isabel Moreno Abril, University of Vigo, Spain

Reviewed by:
Fernando Alda, University of Tennessee at Chattanooga, United States
Marta Barluenga, Spanish National Research Council (CSIC), Spain

***Correspondence:**
Claudio A. González-Wevar
claudio.gonzalez@uach.cl

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

Received: *13 January 2022*
Accepted: *08 April 2022*
Published: *19 May 2022*

Citation:
Astorga MP, Valenzuela A, Segovia NI, Poulin E, Vargas-Chacoff L and González-Wevar CA (2022) Contrasting Patterns of Genetic Diversity and Divergence Between Landlocked and Migratory Populations of Fish Galaxias maculatus, Evaluated Through Mitochondrial DNA Sequencing and Nuclear DNA Microsatellites. Front. Genet. 13:854362. doi: 10.3389/fgene.2022.854362

1 INTRODUCTION

Galaxias maculatus is one of the freshwater fish species with the broadest distribution in cold and temperate regions of the Southern Hemisphere, including populations in southern Australia, Tasmania, New Zealand, South America, and the Falkland/Malvinas Islands (McDowall, 1970; Dyer, 2000; Habit et al., 2006). This species presents amphidromous behaviour, i.e., it reproduces in lakes and rivers and the larval stages migrate to the sea, where the individual develops before returning to complete its growth in freshwater ecosystems (McDowall, 1997). In addition, the presence of resident or landlocked populations has been observed in lakes and rivers, as well as migratory populations that maintain amphidromous behaviour (Delgado et al., 2019). This species is therefore considered an interesting model for historical biogeographic studies due to its particular distribution (Waters and Burridge, 1999; Waters et al., 2000; Zattara et al., 2005; Burridge et al., 2012; Waters et al., 2020), its response to glacial events (Zemlak et al., 2010; Zemlak et al., 2011; Carrea et al., 2012, 2013; González-Wevar et al., 2015a; González-Wevar et al., 2015b; Victoriano et al., 2020), and its migratory and non-migratory behaviour in the same habitat (Delgado et al., 2019; Delgado and Ruzzante, 2020). This variability makes *G. maculatus* a suitable model for evaluating genetic indicators associated with these unique characteristics, as this species has been considered to represent an intermediate evolutionary step between migratory behaviours of marine and freshwater organisms (Corush, 2019).

The glacial history of southern South America is relatively well understood. Continental ice sheet advances and retreats during the Quaternary occurred across the Pacific margin of Patagonia and generated major shifts in sea level, climate and landscape (Hulton et al., 2002; Rabassa et al., 2005, 2011; Hein et al., 2010). During the Last Glacial Maximum the Patagonian Ice Sheet expanded to an area stretching from 35°S to 56°S, and covered most of the Pacific fjords and channels of Patagonia (Clapperton, 1994; McCulloch et al., 2000). Accordingly, the Quaternary geomorphology of Patagonia varied significantly during this period and major glacial changes resulted in periodic extinctions of fauna associated with these near-shore ecosystems, allowing the colonization of vacant niches and creating opportunities for geographical isolation and speciation (Valdovinos et al., 2003; Zemlak et al., 2010; González-Wevar et al., 2011; Fraser et al., 2012; González-Wevar et al., 2012; González-Wevar et al., 2016; González-Wevar et al., 2018; Fernández-Iriarte et al., 2020). The Expansion-Contraction (E-C) model of Quaternary biogeography (Provan and Bennett, 2008) proposes that species and populations contracted their distributions to glacial refugia during glacial maxima. During interglacial periods they expanded their distributions towards previously glaciated areas following the deglaciation process (Provan and Bennett, 2008; Marko et al., 2010). Accordingly, unglaciated areas are expected to harbour higher levels of genetic diversity than ice-impacted areas, or newly founded ones. In contrast, glaciated areas should exhibit evidence of recent postglacial demographic expansions including lower levels of genetic polymorphism with small numbers of haplotypes dominating disproportionally large areas because of recolonization processes (Marko, 2004; Maggs et al., 2008). Moreover, recolonized areas should exhibit low divergence among haplotypes and lower levels of genetic differentiation than refuges (Provan and Bennett, 2008; Marko, 2004; Marko et al., 2010; González-Wevar et al., 2012; González-Wevar et al., 2013). This model provides a relatively simple paradigm to test demographic hypotheses during the Quaternary, and allows us to understand the response of populations and species to major climate change by recognizing distribution range shifts, potential refuge areas and recolonization routes (Hewitt, 2000, Hewitt, 2004; Maggs et al., 2008; Marko et al., 2010; Fraser et al., 2012; González-Wevar et al., 2013). *G. maculatus* includes migratory and landlocked populations distributed along the Pacific margin of South America between 30°S and the southern tip of Chilean Patagonia at 56°S. Phylogeographic and population-based analyses have recorded a small influence of Quaternary glacial cycles on the demography of the species (Zattara et al., 2005; Zemlak et al., 2010; Zemlak et al., 2011; González-Wevar. et al., 2015a; González-Wevar et al., 2015b). In general, *G. maculatus* is characterized by its high levels of genetic diversity and strong phylogeographic structure across its distribution (Zemlak et al., 2010; González-Wevar et al., 2015a; Delgado et al., 2019). Across non-glaciated areas between 38°S and 41°S each estuarine or riverine population represents a different genetic unit characterized by divergent haplogroups (genetic pool derived from shared mtDNA haplotypes). The diversity of *G. maculatus* includes four main haplogroups, three of them in the ice-free zone. The fourth haplogroup is the one that dominates Patagonia (south of 42°S). In contrast, ice-impacted populations along the Pacific margin, between 43°S and 53°S, were all included in a single haplogroup, with lower levels of polymorphism and genetic structure (González-Wevar et al., 2015a). This phylogeographic pattern is consistent with the Expansion-Contraction (E-C) Quaternary biogeographic model (Provan and Bennett, 2008).

Previous studies have shown diverse patterns of genetic differentiation in this species. Carrea et al. (2012) examined spatial patterns of genetic and phenotypic variability of *G. maculatus* from postglacial lakes in north-western Argentinian Patagonia and recognized three different genetic clusters using microsatellite markers. Moreover, genetic analyses of migratory and landlocked populations across a latitudinal gradient recorded higher levels of genetic structure in northern Argentinian Patagonia than in southern Patagonia (Carrea et al., 2013). Northern Patagonian populations of *G. maculatus* could have survived Quaternary glacial cycles *in situ*, while higher latitude populations were probably eradicated by extensive ice-sheet advances and retreats. Recently, Delgado et al. (2019), using neutral and putatively adaptive SNP loci, found that migratory populations in estuarine zones of the Chilean coast are highly differentiated from their landlocked counterparts. Moreover, migratory populations showed higher levels of gene flow and absence of site fidelity, while landlocked populations showed evidence of different colonization events with relatively low genetic diversity and varying levels of gene flow. Victoriano et al. (2020), using mtDNA sequences, observed that *G. maculatus* showed higher levels of genetic diversity and structure than the species *G. platei,* in the same river basin on both sides of the Andes. This result was explained by the different migratory and non-migratory behaviour of the two species.

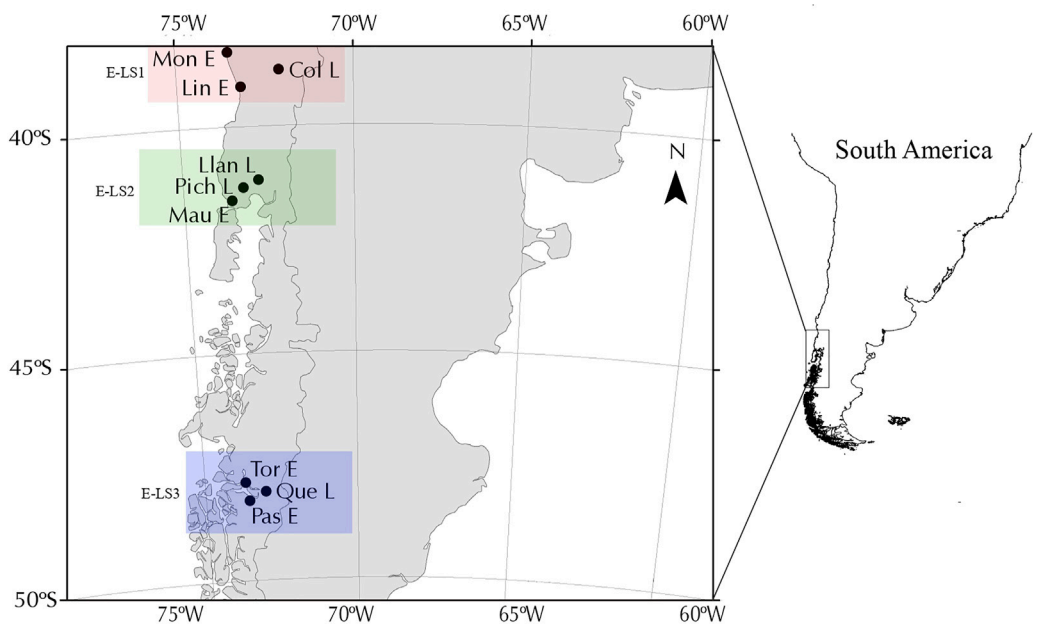


FIGURE 1 | Map showing *Galaxias maculatus* sampling locations in western Patagonia separated into the three systems analysed.

To gain a better understanding of the genetic consequences of the Quaternary glacial in *G. maculatus*, and corroborate the genetic diversity and structure observed in the species, we compared landlocked and migratory populations from western Patagonia in areas that were differentially affected by ice advances and retreats. The populations were evaluated using two types of molecular markers, mitochondrial DNA sequences and nuclear DNA microsatellites. Following the E-C Quaternary biogeographic model, non-glaciated areas are expected to exhibit higher levels of genetic diversity and structure than heavily ice-impacted ones, for both migratory and landlocked populations. Similarly, migratory estuarine populations are expected to harbour higher levels of genetic polymorphism associated with older demographic histories, while landlocked populations should exhibit lower levels of genetic diversity and recent population expansions probably derived from associated estuarine/riverine ecosystems.

2 MATERIALS AND METHODS

2.1 Sampling, DNA Extraction and Amplification

Our analyses included nine populations of *G. maculatus* collected from lakes (n = 4) and associated estuaries (n = 5) in different biogeographical regions across the species distribution in southern Chile, in three areas that were differentially affected by Quaternary glacial processes (**Figure 1**). Firstly, we included two estuary-lake systems (E-LS) located in non-glaciated areas between 38°S and 42°S (**Figure 1**), also known as the Intermediate Area following Camus (2001). Localities in the first system (E-LS1) were Moncul Estuary (Mon-E); Lingue Estuary (Lin-E) and Colico Lake (Col-L) (**Table 1**). Localities in the second system (E-LS2) were: Maullín Estuary (Mau-E); Llanquihue Lake (Lla-L) and Pichilaguna Lake (Pic-L). We also included a third estuary-lake system located in an area that was heavily impacted by the Patagonian Ice Sheet in Magellan Province; localities in this third system (E-LS3) were: Tortel Estuary (Tor-R); Pascua Estuary (Pas-R); and Quetru Lake (Que-L) (**Figure 1**). Based on previous results (Gonzalez-Wevar et al., 2015a; Gonzalez-Wevar et al., 2015b), estuary populations are assumed to represent migratory populations and those located in lakes, landlocked ones. The geo-references of the sites are shown in **Table 1** and their locations in **Figure 1**.

About 300 individuals were sampled during the year 2013, varying in a range of 30–50 individuals per locality, using whitebait nets and storage in ethanol (95%) for DNA preservation.

DNA was extracted from muscle tissue using 1) the salting-out method described by Aljanabi and Martinez (1997), which was applied to the samples analysed in the Genomics and Molecular Ecology Lab in Valdivia; or 2) the standard DNA extraction protocol with ethanol precipitation (Taggart et al., 1992), applied to the samples analysed in the Molecular Genetic Lab in Puerto Montt. Both methods showed the same quality of DNA collection and were used in both markers. A partial fragment of the mitochondrial Control Region (D-loop) was amplified using the specific primers GAL-F and GAL-R designed from the complete mitochondrial genome of the species (González-Wevar et al., 2015b). Nine microsatellite loci were amplified using specific primers developed by Carrea et al. (2009). PCR conditions for mtDNA and microsatellite amplifications were defined following González-Wevar et al., 2015a, Gonzalez-Wevar et al., 2015b and Carrea et al. (2009) respectively. Mitochondrial

TABLE 1 | *Galaxias maculatus* sampling sites in the estuaries and lakes of western Patagonia, separated into the three systems analysed, with sample size for molecular analysis.

System	Locality	Habitat	Code	Latitud	Longitud	N Mit	N Nucl
E-LS1	Moncul	Estuary	MON-E	38°42′S	73°24′W	24	25
	Lingue	Estuary	LIN-E	39°26′S	73°09′W	24	23
	Colico	Lake	COL-L	39°03′S	72°03′W	27	24
E-LS2	Maullín	Estuary	MAU-E	41°35′S	73°38′W	27	31
	Llanquihue	Lake	LLA-L	41°12′S	73°02′W	32	37
	Pichilaguna	Lake	PIC-L	41°16′S	73°03′W	29	19
E-LS3	Tortel	Estuary	TOR-E	47°45′S	73°32′W	27	27
	Pascua	Estuary	PAS-E	48°15′S	73°18′W	45	25
	Quetru	Lake	QUE-L	48°11′S	73°13′W	31	30

DNA amplification products were purified using QIAquick Gel Extraction Kit (QIAGEN) and sequenced in both directions with an Automatic Sequencer ABI3730 × 1 at Macrogen Inc. (Seoul, South Korea). The sizes of the amplified microsatellite alleles were determined using an automatic DNA sequencer (ABI Prism 377; Applied Biosystems).

2.2 Mitochondrial DNA Analysis

Mitochondrial sequences in *G. maculatus* were edited using GENEIOUS v.9.0.4 (http://www.geneious.com) and aligned with MUSCLE (Edgar, 2004). New D-loop sequences were deposited in GenBank under Accession Numbers: OM743508—OM743773. DNA saturation analysis was performed using DAMBE (Xia, 2013) to determine how saturation of transitions accumulates in relation to nucleotide divergence. Levels of genetic diversity of mtDNA marker were estimated through standard indices including: the number of haplotypes (k), the number of segregating sites (S), haplotype diversity (H), the average and number of pairwise differences ($\prod$) for each locality, each estuary-lake system, and for the whole D-loop data set using DnaSP (Librado and Rozas 2009). We performed neutrality statistical tests (Tajima's D and Fu's FS) in DnaSP for each locality, each E-LS and the whole data set to estimate whether D-loop sequences in the species deviate from expectations under mutation-drift equilibrium. Considering the high levels of genetic diversity recorded in *G. maculatus* (Zemlak et al., 2010; González-Wevar CA. et al., 2015, Gonzalez-Wevar et al., 2015b, we determined levels of genetic differentiation between the localities analysed using mean pairwise differences (N_{ST}) following Pons and Petit (1996) in ARLEQUIN v. 3.5 (Excoffier and Lischer, 2010). The statistical significance of these analyses was estimated through 20,000 permutations. We estimated phylogeographic structure using the nearest neighbour statistic (S_{nn}), which measures how often nearest neighbours in sequences (in sequence space) are from the same locality in geographic space (Hudson, 1990). The statistical significance of S_{nn} was determined through 10,000 permutations. We determined the spatial genetic structure in *G. maculatus* through the number and composition of groups that were most differentiated based on mtDNA sequence data using SAMOVA (Spatial Analysis of Molecular Variance) (Dupanloup et al., 2002). SAMOVA is a popular analysis that uses multiple spatial scales in statistical methods to characterize spatial genetic structure based on pairwise genetic differences. We reconstructed mtDNA genealogical relationships in *G. maculatus* using Maximum Parsimony Networks in HAPVIEW (Salzburger et al., 2011). We also estimated the patterns of demographic history in the species by comparing the distribution of pairwise differences between haplotypes (mismatch distribution) for each locality, each E-LS and for the whole data set, up to the expected distribution under the sudden expansion growth model of Rogers and Harpending (1992) in DnaSP. Moreover, we estimated the mitochondrial age of the last demographic expansion in landlocked populations (Colico, Pichilaguna, and Quetru) under the sudden population growth model using the formula $\tau = 2\,\mu t$, where τ = the date of growth/decline measured in units of mutational time, t = time in years, and μ = the mutation rate per sequence per year. For this, we used a specific phylogeographic mutation rate estimated for the D-loop in galaxiid fish estimated by Burridge et al. (2008).

Finally, we performed phylogenetic reconstructions including the complete D-loop data set in *G. maculatus*, together with a sequence of *G. platei* as the outgroup. Bayesian reconstructions of haplotype relationships were done in MrBAYES v.3.1.2 (Huelsenbeck and Ronquist, 2001) using the GTR + I + G substitution model as previously selected with JModeltest v.2.1.10 (Darriba et al., 2012). Bayesian posterior probabilities (BPP) were estimated using the Metropolis coupled Markov-chain Monte-Carlo algorithm (MCMC). For this, we ran four chains for 150×10^6 generations and trees were sampled every 1,000 generations. Stationarity of this analysis was inferred when the average standard deviation of split frequencies was lower than 0.01 as suggested by Huelsenbeck and Ronquist (2001). The first 5% of the parameter values were discarded as burn-in and posterior probabilities were calculated as the fraction of trees showing a particular node. Finally, posterior probability-density was summarized as a maximum clade credibility tree using TreeAnnotator v.1.6.1 (http://beast.bio.ed.ac.uk/TreeAnnotator) and edited using the FigTree v.1.4.3 programme (http://tree.bio.ed.ac. uk/software/figtree).

2.3 Microsatellite Analyses

Microsatellite fragment analysis was performed using Peak Scanner™ Software from Applied Biosystems and allele sizes were assigned to bins using FLEXIBIN (Amos et al., 2007). Finally, the evaluation of null alleles was reviewed and corrected using MICROCHECKER software v.2.2.3 (Van Oosterhout et al., 2004).

The statistical independence between loci was assessed using GENEPOP 4.7.5 (Rousset, 2008). Genotypic linkage disequilibrium between each pair of loci within populations was tested using Fisher's exact test with a Markov chain. The deviation from genotypic proportion expected under the Hardy-Weinberg equilibrium was tested with exact *p* values by the Markov chain method using GENALEX 6.501 (Smouse and Peakall, 2012). Allele richness (Rs), standardized for the number of data, was calculated with FSTAT 2.9.4 (Goudet, 2002); the observed heterozygosity (Ho) and the expected heterozygosity (He) were estimated with GENALEX 6.501 (Smouse and Peakall, 2012). Microsatellite population genetic differentiation was assessed through F_{ST} comparisons between pairs of localities based on the total of polymorphic loci (Weir and Cockerham, 1984) using the sum of squared size differences (F_{ST}-like) with *p*-value significance, implemented in ARLEQUIN 3.5 (Excoffier and Lischer, 2010).

We used different approaches to estimate the spatial distribution of genetic diversity using multilocus data. We established patterns of genetic differentiation among localities in *G. maculatus,* through Bayesian analysis implemented in STRUCTURE v.2.3.2 (Pritchard et al., 2000). To determine which K best fitted the data, we used ΔK methods described by Evanno et al. (2005); this method specifies that the most likely K is that found when this probability is plotted against successive values of K asymptotes, and does not increase significantly with increasing K. This analysis was estimated using the Structure Harvester web page (Earl and VonHoldt, 2012).

Finally, we performed a discriminant analysis of principal components (DAPC) using ADEGENET 3.1-1 (Jombart, 2008; Jombart and Ahmed, 2011) in the R platform (R Core Team, 2014). For this analysis, the number of clusters (K) was identified using the *find. clusters* function, which is based on the lowest values of the Bayesian information criterion (BIC). This type of analysis defines a model in which genetic variation is divided within and between groups, and which at the same time maximises the synthesis of the variables within groups and minimises it between groups using the information of the precedence of each sample as *a priori* information. Discriminant analysis (DA) achieves the best separation of the individuals within predefined groups and makes a probabilistic assignment of individuals to each group.

3 RESULTS

3.1 Mitochondrial DNA

The complete mitochondrial DNA set in *G. maculatus* included 266 individuals collected from nine populations, with an average of 28 individuals per locality in a range from 19 to 45 individuals per locality (**Table 2**). The set consisted of 928 nucleotide positions. Considering that the analyses included a non-coding region of the mtDNA, several insertions and deletions were detected but not considered for future analyses. Control region sequences in *G. maculatus* were A—T rich (57.6%) compared to the G—C content (42.4%). Genetic diversity in *G. maculatus* was high with 197 variable positions (21.22%), of which 147 (74.61%) were parsimoniously informative. Levels of genetic diversity were much higher in migratory populations (estuaries) than in most of the landlocked populations (lakes), except for Llanquihue Lake, which exhibited high levels of polymorphism (**Table 2**). Haplotype diversity (H) ranged between 1.0 (Moncul, Maullín, Tortel and Pascua estuaries) and 0.729 (Quetru Lake) (**Table 2**). Similarly, the average numbers of nucleotide differences (Π) were higher in migratory than in landlocked populations, ranging from 30.33 (Maullín Estuary) to 1.72 (Colico Lake).

Pairwise N_{ST} comparisons detected a high degree of genetic differentiation between migratory and landlocked populations of *G. maculatus* along a latitudinal gradient. All pairwise population comparisons with the exception of one (between Pascua and Tortel Estuaries), were statistically significant, showing the high degree of genetic differentiation reported in the species (**Table 3**). Across the sampling area, levels of mtDNA genetic differentiation were higher among lakes ($\bar{X}$ = 0.947) than among estuaries ($\bar{X}$ = 0.513), followed by the differentiation between lake and estuary populations ($\bar{X}$ = 0.669). Specifically, migratory populations of the southern system (Tortel and Pascua Estuaries E-LS3) showed no significant genetic differentiation (**Table 3**). The nearest neighbour statistic S_{nn} in *G. maculatus* (S_{nn} = 0.815) showed a high and significant phylogeographic signal in the whole D-loop data set ($p < 0.001$). SAMOVA supported the general pattern of mtDNA structure and identified eight groups with maximum differences accounting for 66.09% of the total genetic variance, while only 0.06% of the variance was due to within-group variation among localities (**Table 4**). SAMOVA recognized each locality, except for Pascua and Tortel Estuaries (E-LS3), as different populations.

Considering the levels of genetic diversity recorded in the species, haplotype network analyses were divided into the different estuary-lake systems (E-LS1—E-LS3) analysed here (**Figures 2A–C**). Maximum parsimony haplotype network in E-LS1 (**Figure 2A**) included a total of 59 different haplotypes, and Moncul Estuary showed the highest diversity with a total of 24 different haplotypes (H = 1.00). Lingue Estuary exhibited a total of 18 different haplotypes (H = 0.94), three of which were shared by more than one individual. Considering the expanded genealogies recorded at Moncul (Π = 20.50) and Lingue (Π = 22.77) estuaries, both localities exhibited multimodal distributions of pairwise differences between haplotypes. In contrast, the Colico Lake parsimony network included a total of 17 different haplotypes. A dominant haplotype was recorded in 12 individuals (44.4%), with several closely associated singletons (**Figure 2A**). Considering the short genealogy with a typical star-like topology recorded at Colico Lake, the distribution of pairwise differences between haplotypes presented a unimodal distribution (not shown). Maximum parsimony

TABLE 2 | Diversity indices and neutrality tests in *Galaxias maculatus* populations in the different E-LS across the species distribution in western Patagonia.

Locality	D-Loop mitochondrial DNA						Microsatellites nuclear DNA			
	K	H	S	Π	D	F_S	Rs	Ho	He	F
MON-E	24	1.00	85	20.50	−0.76	−10.09*	10.62 ± 3.46	0.45 ± 0.07	0.78 ± 0.07	0.39 ± 0.10
LIN-E	18	0.94	77	22.77	−0.44	−0.61	10.56 ± 4.36	0.42 ± 0.04	0.76 ± 0.08	0.43 ± 0.06
COL-L	17	0.84	18	1.72	−2.29*	−14.48*	5.86 ± 3.67	0.45 ± 0.11	0.57 ± 0.09	0.25 ± 0.14
E-LS1	**59**	**0.97**	**103**	**20.75**	**−0.57**	**−21.70***	**9.01 ± 4.34**	**0.44 ± 0.06**	**0.78 ± 0.06**	**0.42 ± 0.08**
MAU-E	27	1.00	109	30.33	−0.22	−9.18*	9.68 ± 3.39	0.34 ± 0.06	0.71 ± 0.08	0.47 ± 0.11
LLA-L	28	0.99	86	11.65	−1.91*	−13.31*	9.23 ± 3.79	0.59 ± 0.06	0.75 ± 0.04	0.20 ± 0.09
PIC-L	10	0.77	19	3.23	−1.15	−1.26	4.67 ± 2.84	0.57 ± 0.13	0.50 ± 0.09	−0.09 ± 0.18
E-LS2	**65**	**0.98**	**162**	**36.63**	**−0.14**	**−12.53***	**7.86 ± 3.97**	**0.48 ± 0.05**	**0.79 ± 0.05**	**0.34 ± 0.10**
TOR-E	26	1.00	71	15.88	−1.01	−11.95*	10.37 ± 3.18	0.48 ± 0.06	0.81 ± 0.03	0.39 ± 0.07
PAS-E	43	1.00	75	13.24	−1.28	−32.40*	10.04 ± 3.48	0.42 ± 0.05	0.80 ± 0.05	0.48 ± 0.06
QUE-L	12	0.73	22	3.11	−1.60	−2.02	6.68 ± 2.40	0.49 ± 0.09	0.62 ± 0.08	0.27 ± 0.06
E-LS3	**81**	**0.98**	**94**	**14.38**	**−1.18**	**−61.83***	**9.03 ± 3.39**	**0.46 ± 0.05**	**0.79 ± 0.04**	**0.42 ± 0.06**
Total	205	0.99	184	31.71	−0.45	−164.5*	11.70 ± 3.57	0.47 ± 0.03	0.70 ± 0.03	0.31 ± 0.04

Indices for D-loop from mitochondrial DNA where K, number of haplotypes; H, haplotype diversity; S*, number of polymorphic sites; Π, average number of pairwise differences; D, Tajima's D neutrality test and F_S, Fu's F_S neutrality test and indicators with standard error of the nine microsatellite loci; Rs, allele richness; Ho, observed heterozygosity; He, expected heterozygosity; F, inbreeding coefficient.*

Global genetic indices for each E-LS are marked in bold. Statistical significance for Neutrality Test (Fs) are indicated with an asterisk.

TABLE 3 | Matrix of genetic differentiation values between pairs of *Galaxias maculatus* locations, estimated by mean general pairwise values of differentiation (N_{ST}) of Mitochondrial DNA data (above diagonal) and estimated by F_{ST} obtained from the nine microsatellite loci (below diagonal).

	MON-E	LIN-E	COL-L	MAU-E	LLA-L	PIC-L	TOR-E	PAS-E	QUE-L
MON-E		**0.315**	**0.650**	**0.509**	**0.249**	**0.788**	**0.594**	**0.634**	**0.738**
LIN-E	0.014		**0.431**	**0.391**	**0.412**	**0.730**	**0.417**	**0.467**	**0.618**
COL-L	**0.216**	**0.446**		**0.683**	**0.750**	**0.949**	**0.807**	**0.805**	**0.944**
MAU-E	0.016	0.044	**0.146**		**0.583**	**0.711**	**0.289**	**0.355**	**0.519**
LLA-L	**0.100**	**0.124**	**0.264**	**0.139**		**0.848**	**0.664**	**0.686**	**0.796**
PIC-L	**0.270**	**0.333**	**0.460**	**0.296**	**0.135**		**0.810**	**0.812**	**0.936**
TOR-E	0.033	0.007	**0.430**	**0.077**	**0.135**	**0.386**		0.000	**0.463**
PAS-E	**0.057**	0.029	**0.350**	0.045	**0.186**	**0.288**	**0.101**		**0.473**
QUE-L	**0.086**	**0.051**	**0.507**	**0.156**	**0.123**	**0.399**	0.018	**0.161**	

Statistical significant pairwise comparisons are marked in bold.

TABLE 4 | Spatial Analysis of Molecular Variance (SAMOVA) showing the percentage of genetic variation explained among groups (Moncul E, Lingue E, Colico L, Maullín E, Llanquihue L, Pichilaguna L, Quetru L, Pascua/Tortel E), among populations within groups and within populations using mtDNA. Where F_{SC} represents differentiation within populations among groups and F_{CT} represents differentiation among groups (***$p < 0.001$, **$p < 0.01$).

Source of variation	d.f.	Sum of squares	Variance component	Percentage of variation
Among groups	7	3067.993	13.34842 Va	66.09
Among populations within groups	1	7.237	0.01188 Vb	0.06
Within populations	257	1756.879	6.83611 Vc	33.85
Total	265	4832.109	20.19641	

*Fixation index FSC: 0.00173*** FCT: 0.66093**.*

Statistical significant pairwise comparisons are marked in bold.

network in E-LS2 (**Figure 2B**) included a total of 65 haplotypes; Maullín Estuary showed the highest diversity with 27 different haplotypes (H = 1.00). Llanquihue Lake showed high levels of genetic polymorphism with 28 different haplotypes (H = 0.99), three of which were found in more than one individual. Considering the expanded genealogy recorded in Maullín Estuary (Π = 30.33) and Llanquihue Lake (Π = 11.65), both localities exhibited multimodal distributions of pairwise differences between haplotypes. Interestingly, two individuals collected in Llanquihue Lake are closely related to the diversity recorded in Maullín Estuary. In contrast, the parsimony network for Pichilaguna Lake included a total of only 10 haplotypes. Two dominant haplotypes were recorded in this locality, the first was found in 11 specimens (37.93%) and the second in 9 (31.39%). Another eight low-frequency haplotypes were closely associated with the dominant ones (**Figure 2B**). The pairwise differences in Pichilaguna Lake showed a bimodal

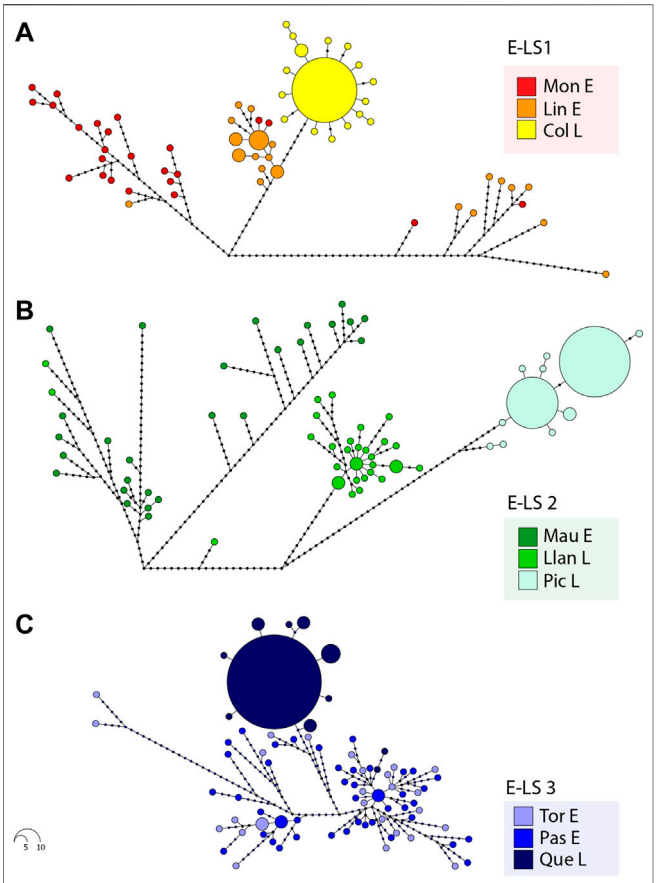


FIGURE 2 | Maximum parsimony haplotype networks in *Galaxias maculatus* based on mtDNA D-loop sequences from: **(A)** Estuary-Lake System 1: E-LS 1 (Moncul Estuary: Mon E, Lingue Estuary: Lin E and Colico Lake: Col L); **(B)** Estuary-Lake System 2: E-LS 2(Maullín Estuary: Mau E, Llanquihue Lake: Llan L and Pichilaguna Lake: Pic L); and **(C)** Estuary-Lake System 3: E-LS3 (Tortel Estuary: Tor E, Pascua Estuary: Pas E and Quetru Lake: Que L). Each haplotype is represented by a coloured circle showing the locality where it was collected. The size of the circles is proportional to the frequency in the sampling effort.

distribution. The maximum parsimony network in E-LS3 (**Figure 2C**) included a total of 81 different haplotypes. Tortel and Pascua Estuaries showed the highest diversity with 26 and 43 different haplotypes respectively. Considering the expanded genealogy recorded at Tortel (Π = 15.88) and Pascua (Π = 13.24) Estuaries, both localities exhibited multimodal distributions of pairwise differences between haplotypes. In contrast, the parsimony network in Quetru Lake included a total of 12 different haplotypes. A dominant haplotype was recorded in 16 individuals (51.61%) surrounded by several singletons and low-frequency haplotypes. Interestingly, two individuals from Quetru Lake were closely associated with the diversity recorded in the associated tributaries (Pascua and Tortel Estuaries) (**Figure 2C**). Considering this, the distribution of pairwise differences between haplotypes from Quetru Lake exhibited a bimodal distribution. The age of population expansions estimated for landlocked populations of *G. maculatus* were 13,470 years for Colico lake, 16,840 years for Pichilaguna lake and 12,120 years for Quetru lake.

Phylogenetic reconstructions recognized three (HI, HIII and HIV) of the four main haplogroups previously recorded in *G. maculatus* in southern South America (Zemlak et al., 2010; González-Wevar et al., 2015a, Gonzalez-Wevar et al., 2015b. Haplogroup I was more abundant across the Intermediate Area localities while HIII was present in Lingue and Maullín estuaries (**Supplementary Figure S1**). Haplogroup IV was the dominant in estuarine and lacustrine populations located within the Patagonian Ice Sheet (Tortel Estuary, Pascua Estuary, and Quetru Lake) in areas that were heavily impacted by ice advances and retreats during glacial periods (**Supplementary Figure S1**). Some specimens from the localities of the Intermediate Area (Moncul Estuary, Lingue Estuary, Maullín Estuary and Llanquihue Lake) fell within the diversity of HIV, confirming the asymmetrical gene flow pattern (from north to south) previously found in the species (González-Wevar et al., 2015a).

3.2 Microsatellites

Genetic data for the microsatellite loci analysed were obtained from a total of 241 individuals collected from nine populations, of which 131 were collected in estuaries (migratory populations) and 110 in lakes (landlocked populations). Deviation from the Hardy-Weinberg equilibrium was observed in different loci combined with different populations (**Supplementary Table S1**). Excess homozygosity was observed, mainly through the presence of null alleles in some loci. Allele richness was higher in migratory localities (Rs = 10.25) than in landlocked populations (Rs = 6.61), with significant differences ($p < 0.01$) between the groups and when estuaries and lakes were compared within each E-LS ($p < 0.01$). In general, estuarine migratory populations showed greater allele richness than landlocked ones, except for Llanquihue Lake (E-LS2), which showed similar diversity to those recorded in migratory estuarine localities. The highest levels of allele richness were observed in migratory populations from Moncul and Lingue Estuaries (**Table 2**), and the lowest in landlocked populations including Pichilaguna, Colico and Quetru Lakes. The inbreeding coefficient (F; **Table 2**) showed no significant differences across a latitudinal gradient. When the genetic diversity was compared among systems (E-LS), no major differences were observed using microsatellites.

Overall comparison between locations showed a value of F_{ST} = 0.149. *A posteriori* pairwise analysis between locations based on the F_{ST} estimator (**Table 3**) showed significant differences between localities; this is explained mainly by the landlocked populations which differ between all the locations, except for Quetru Lake which showed no significant differences from Tortel Estuary.

Comparing only migratory populations (estuaries) with each other, we observed a low although significant genetic differentiation with F_{ST} = 0.058 ($p < 0.001$). When landlocked populations (lakes) were compared with each other, a higher genetic differentiation was observed (F_{ST} = 0.228), with

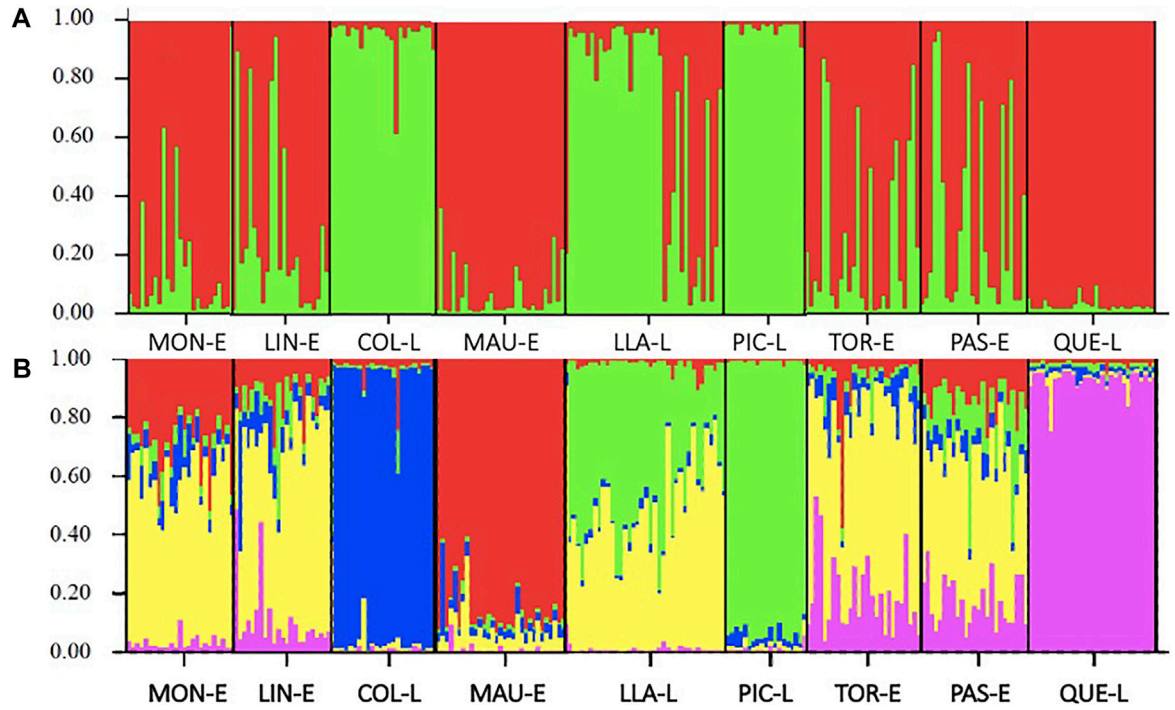


FIGURE 3 | Genetic differentiation between locations, assessed by Bayesian analysis: **(A)** Diagram of the estimated population structure of Galaxias maculatus for the nine locations sampled with K = 2, the colours represent the genetic groups; **(B)** Diagram of population structure with K = 5. The name of the localities is observed on the lower horizontal axis: MON-E: Moncul Estuary; LIN-E: Lingue Estuary; COL-L: Colico Lake; MAU-E: Maullín Estuary; LLA-L: Llanquihue Lake; PIC-L Pichilaguna Lake; TOR-E: Tortel Estuary; PAS-E: Pascua Estuary and QUE-L Quetru Lake.

significant differences ($p < 0.001$). When the estuaries were compared with the lakes within each system, the lowest genetic differentiation was observed between the localities of E-LS 3.

Bayesian analysis of genetic structuring in STRUCTURE showed two genetic groups differentiated from one another, represented by the colours distributed among the nine locations sampled (**Figure 3**) in which estuaries are mainly separated from lakes. The correlation proposed by Evanno et al. (2005) shows the point where ΔK reaches its maximum value, estimating that K = 2 (Ln P(K) = −9377.4) is the most probable K for the genetic groups identified in the sample (**Figure 3A**). Genetic differentiation was structured with two genetic clusters: one group with the individuals belonging to Colico, Llanquihue and Pichilaguna lakes and other cluster formed by all the rivers and Quetru Lake. However, when the results were observed with a higher value of K (K = 5) (**Figure 3B**), it was possible to identify a pattern of differentiation, consistent with all previous statistical results, in which high similarity among the estuaries except for Maullin Estuary and high differentiation between the lakes is observed, except for Lake Llanquihue.

The results of the discriminant analysis of principal components (DAPC) (**Figure 4**) showed a consistent pattern of genetic differentiation, similar to that observed in the Fst analysis. The highest differentiation was observed in Colico Lake and the lowest among estuarine localities. The results again group the individuals from Moncul, Lingue and Pascua Estuaries together as single groups, but close to Tortel Estuary. However, a clear separation is observed between estuary-lake systems. The individuals from Colico Lake clearly separate into a differentiated group, Llanquihue and Pichilaguna Lakes are seen to be segregated jointly, with these landlocked populations highly differentiated. The exception was Maullín Estuary, which is seen to form a more differentiated group.

4 DISCUSSION

4.1 Genetic Diversity

4.1.1 Mitochondrial DNA

As previously demonstrated in *Galaxias maculatus* (Zemlak et al., 2010; Zemlak et al., 2011; González-Wevar et al., 2015a; Gonzalez-Wevar et al., 2015b; Delgado et al., 2019), this species shows high levels of genetic diversity across its distribution in western South America. Among 266 individuals analysed with mtDNA, we recorded a total of 205 different haplotypes. On the one hand, levels of genetic diversity recorded in migratory populations of *G. maculatus* were higher across formerly unglaciated areas, while populations within the limits of the Patagonian Ice Sheet exhibited lower levels of polymorphism; accordingly, unglaciated areas represent more stable populations that were less affected by demographic processes associated with Quaternary glacial ice advances and retreats. On the other hand, the general diversity

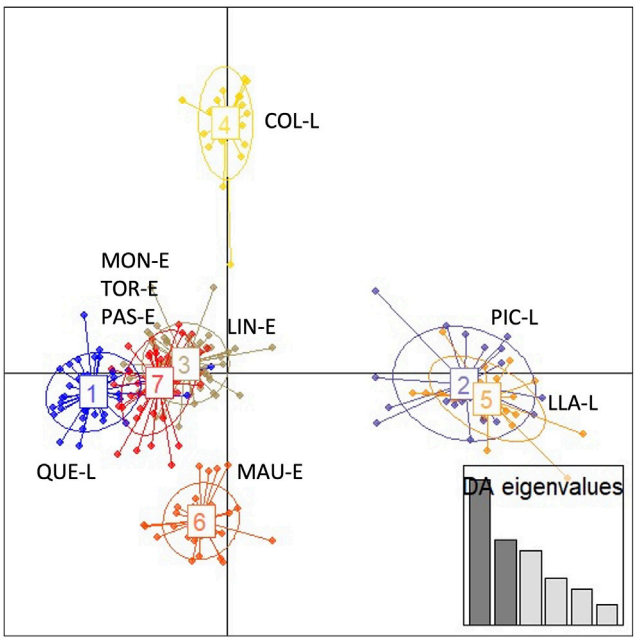


FIGURE 4 | Ordering of the *Galaxias maculatus* clusters identified by the discriminant analysis of principal components (DAPC) of the different locations sampled. 1) Quetru Lake (QUE-L); 2: Pichilaguna Lake (PIC-L); 3: Lingue Estuary (LIN-E); 4: Colico Lake (COL-L); 5: Llanquihue Lake (LLA-L); 6: Maullín Estuary (MAU-E); 7: Moncul Estuary (MON-E)—Tortel Estuary (TOR-E)- Pascua Estuary (PAS-E).

pattern recorded in migratory populations along a latitudinal gradient was not found in landlocked populations; levels of genetic diversity in landlocked populations of *G. maculatus* were lower and similar across our sampling effort, and probably a consequence of analogous demographic trajectories associated with recent bottlenecks and/or founder effects. Even though Colico and Pichilaguna lakes are located in unglaciated areas of western Patagonia, these lakes were covered by ice during the LGM (Moreno et al., 2018). Accordingly, the expansions of landlocked *G. maculatus* populations occurred after the Last Glacial Maximum, and hence are associated with post-glacial recolonization processes following the deglaciation process. As previously stated, the case of Llanquihue Lake is interesting and requires further study. Nevertheless, studies on *G. maculatus* have demonstrated a direct relationship between genetic diversity and the area of the lake analysed. Llanquihue is the second largest lake in South America; it is located at the limit of the expansion of glacial ice and probably served as a genetic reservoir for the species.

4.1.2 Microsatellites

Microsatellite analysis in *G. maculatus* showed similar diversity patterns to those recorded in populations of the species from lakes and estuaries across Argentinian Patagonia (Carrea et al., 2012; Carrea et al., 2013). Our mean diversity and allele richness estimations are higher than those recorded by Carrea et al. (2013) in Argentinian Patagonia (Rs = 6.45). This study showed that migratory localities exhibited higher levels of genetic diversity than landlocked ones. It is to be expected that those populations with migratory behaviour will present greater genetic diversity than closed or landlocked populations, due to higher gene flow in the former and greater probability of genetic drift in the latter. Delgado et al. (2019) observe that estuarine migratory populations exhibited higher genetic diversity than resident populations, which is explained by the (probably low) diversity of the colonizing populations. Therefore, these differences in diversity between types of samples can be explained by bottlenecks generated by population reduction during glacial maxima, similar to those recorded by Zemlak et al. (2010) and Vera-Escalona et al., 2020.

4.2 Genetic Structure

4.2.1 Mitochondrial DNA

As previously shown in this species (González-Wevar et al., 2015a), we also recorded contrasting patterns of genetic structure in migratory populations between the main biogeographic areas analysed. Each of the estuaries analysed north of the 42°S boundary, and across an area less than 250 km wide, represents a different genetic unit; the estuaries analysed in the southern system (E-LS3) of Patagonia exhibited lower levels of genetic differentiation. Again, major differences in terms of the genetic structure between unglaciated and glaciated areas are in basic agreement with the predictions of the Expansion-Contraction Quaternary biogeographic model in the south-eastern Pacific (Provan and Bennett, 2008; Pardo-Gandarillas et al., 2018; Fernández-Iriarte et al., 2020). González-Wevar CA. et al., 2015, González-Wevar et al., 2015b analysed populations from Chilean Patagonia and recorded low degrees of genetic structuring in estuary locations south of 42°S, belonging to the Magellan Biogeographical Province. These results coincide with observations in E-LS3 in this work, where the two estuaries showed no significant genetic differentiation. Thus, the genetic differences observed between the types of localities can be explained by the life cycles of the fish and their resident or migratory behaviour observed in the estuaries and rivers (Wallis et al., 2001; Delgado et al., 2019). Our results suggest a high probability of observing migratory behaviours in the estuary samples, and resident behaviours in lakes; the exception would be Maullín Estuary (MAU-E), where the slight genetic differentiation observed can be explained by the possible presence of a high proportion of individuals from resident populations.

Our results also show that the greatest differentiation is observed among all the landlocked populations. Quetru Lake, located in the area influenced by the LGM, showed the highest differentiation; it could have been affected by a process of extinction and recolonization from possible glacial refuges (Fraser et al., 2012).

4.2.2 Microsatellites

Comparative analyses using microsatellites in *G. maculatus* detected significant differences among the populations analysed. However, pairwise analyses showed that these differences are mainly explained by the high genetic differentiation recorded in the landlocked populations, particularly those from Pichilaguna and Quetru Lakes. When

landlocked and migratory populations were compared separately, we observed that the estuaries exhibit lower genetic structuring; in contrast, landlocked populations showed higher levels of genetic structure and fewer migrants. These observations were corroborated by the multivariate analyses (DAPC), where migratory populations from estuaries were grouped in a big central cloud; however, there are two clusters of higher similarity, one including Moncul, Lingue and Pascua and another one Tortel, all geographically distant from each another. These multivariate analyses showed that individuals sampled from landlocked populations could be separated from each other, and from migratory populations. Colico, Llanquihue and Pichilaguna Lakes showed the greatest similarity. Llanquihue and Pichilaguna Lakes are 3.5 km apart and are probably connected by streams. Llanquihue is the second-largest lake in South America, covering 860 km^2; accordingly, it may represent a source for the *G. maculatus* population of Pichilaguna lake, which has an area of approximately 1 km^2, although mtDNA data do not show shared genotypes. Zattara et al., 2005 found a positive correlation between genetic diversity, measured by isoenzymes, and lake area and perimeter; this agrees with the observations in our work, in which the lake with the largest area and perimeter (Llanquihue) has the highest allele richness, while the smallest lake (Pichilaguna) exhibited the lowest diversity of alleles. This is mainly due to the sizes reached by the populations in each lake; it is probable that populations of smaller effective size are more exposed to genetic drift or to the combined effects of bottlenecks, founder effects, endogamy, and restricted levels of gene flows, which would increase the inter-population genetic divergence in comparison with larger populations (Hartl and Clark, 1997). Just as was observed from the sequence data, Quetru Lake (E-LS3), located in the zone affected by the Last Glacial Maximum, shows greater differentiation and genetic structure than the lakes and estuaries located in the other two systems; this can be explained by the processes of extinction and recolonization in this area.

Finally, Bayesian Structure analysis showed a high differentiation among lakes and estuaries, separating *G. maculatus* populations into two main differentiated genetic groups. However, when evaluating this analysis with a greater number of differentiated groups (K = 5), it is possible to detect a differentiation pattern with high separation between lakes while the similarity between estuaries is maintained, which is consistent with our observations in the other analyses. The low differentiation detected among lakes by this Bayesian analysis may result from the high genetic diversity detected by microsatellite markers, due to their high mutational rate, which may generate allele similarity by homoplasies (Jarne and Lagoda, 1996; Estoup et al., 2002; Bhargava and Fuentes, 2010). The results obtained by the two types of markers are highly consistent; however, as observed in the structure result, a difference may be observed between the two molecular markers used. This could be due to the greater genetic diversity of microsatellites, which makes them very suitable for population analysis; however, in cases where the populations present high differentiation, it may be that the use of sequences as molecular markers can better detect the differentiation pattern than that shown by microsatellites. This would be due mainly to the possibility that the latter show high genetic similarity due to alleles that have already presented various mutations; in other words, for highly differentiated groups, the microsatellites could show homoplasies (Jarne and Lagoda, 1996; Estoup et al., 2002; Bhargava and Fuentes, 2010).

The different genetic statistical analyses show a consistent pattern of separation between the lakes analysed, and between lakes and estuaries. This pattern of genetic differentiation between lakes is similar to those recorded by Zattara et al., 2005, Delgado et al. (2019) and Rojo et al. (2020) using different types of molecular markers and analysing different watershed systems, and may be explained by the non-migratory behaviour of lacustrine populations. A similar pattern of genetic differentiation was recorded by Waters and Burridge (1999) when comparing mtDNA sequences from two Chilean lakes located in the same basin: Riñihue and Saval.

This high genetic differentiation, observed mainly in lakes and not in estuaries, may be because each lake is an "island", and the separation between them acts as a barrier to gene flow. These results show that lakes behave as landlocked zones with little capacity for admixture, as there is probably little or no fish migration between lakes and estuaries. It has also been proposed that the level of structuring found in some basins (connected lakes, rivers, and estuaries) may be due to differential effects of Quaternary climate cycles in the different locations studied, since in some cases the population adopts a resident behaviour or life history, as observed in the variation to optional diadromous behaviour in a basin in Argentinian Patagonia (Carrea et al., 2013). This contrasts with the findings of Hollin and Schilling (1981) in locations at high latitudes which were strongly affected by glaciation in the mid-Pleistocene, the last glacial period when glaciers covered vast areas of Patagonia. This factor may have generated a refuge which now presents high genetic differentiation, as observed in the most southerly basin in Argentinian Patagonia analysed by Zemlak et al. (2008) and Carrea et al. (2013). In the present work, we observed a similar pattern of differentiation increasing with the latitude of the possible refuge zones affected by glaciation (Zemlak et al., 2008), mainly in Estuary-Lake System 3 (Tortel and Pascua Estuary with Quetru Lake), which would have been more strongly affected by the last glacial period than E-LS1 and E-LS2. However, a pattern of genetic differentiation was observed in the genetic diversity of E-LS3, explained by an effect of the last glacial period. In the case of lakes, which present high differentiation between one another and as compared to estuaries, the differentiation observed may be due to the possible presence of a refuge in each. In post-glaciation periods, these refuges would reconstruct highly differentiated populations. Thus, in the present work, using mtDNA sequences and microsatellite nuclear markers, we corroborated a pattern of differentiation between lakes and estuaries, similar to that already detected on a small scale by Zemlak et al. (2008) using mitochondrial DNA sequencing, and Delgado et al. (2019) using SNP markers. This corroborated the

genetic separation of the lakes, which had already been evaluated through transplants by Delgado et al. (2020), where these differences come to be explained by evolutionary processes (natural selection or genetic drift); and by the presence of two types of behaviours in the same species, resident and migratory (Delgado et al., 2019). In addition, the effect of the Quaternary climate cycles on population structure and the loss of genetic diversity can be detected with both types of markers. Therefore, these patterns may be due to a combination of biological factors, i.e., resident non-migratory behaviour and/or landlocking and natal homing-in, as well as geological factors, i.e., E-C Quaternary biogeographic glaciation. It is therefore important to establish protection measures for this species, mainly in the lake populations due to the low diversity and high genetic differentiation observed in the lakes. Our findings need to be confirmed, to ensure the protection or conservation of these isolated populations.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: National Center for Biotechnology Information (NCBI) BioProject database under accession numbers OM743508—OM743773.

ETHICS STATEMENT

The animal study was reviewed and approved by the Comite de Etica-Bioetica Universidad Austral de Chile.

AUTHOR CONTRIBUTIONS

MPA wrote the manuscript, analyzed and interpreted data; AV carried out laboratory analysis and data analysis; NS carried out data analysis and prepared figures and tables and reviewed the manuscript; EP and LV-C contributed substantially to the drafting of the manuscript; CG-W conceived the study, wrote the manuscript, and analyzed and interpreted data.

FUNDING

This study was supported by the postdoctorate 3120075 fellowship. At the same time, this research was partially funded by the Projects P05002 ICM and PFB 023 (IEB, Universidad de Chile), to EP and CG-W and projects Fondecyt 1110798 and 1151375 to LV-C. Fondap IDEAL program (15150003), Fondecyt Regular project 1210787 and ANID—Millenium Science Initiative—BASE. ANID — Programa Iniciativa Científica Milenio — ICN2021_002.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.854362/full#supplementary-material

Supplementary Figure S1 | Bayesian maximum clade credibility tree of *Galaxias maculatus* haplotype relationships based on D-loop sequences. Bayesian posterior probabilities are shown above the nodes. Colored circles indicate sampling localities of the specimens analysed. Main haplogroups (HI, HIII, and HIV) recorded in *G. maculatus* are also indicated.

Supplementary Table S1 | Deviation from the Hardy-Weinberg equilibrium in different microsatellites loci in different localities.

REFERENCES

Aljanabi, S., 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

Amos, W., Hoffman, J. I., Frodsham, A., Zhang, L., Best, S., and Hill, A. V. (2007). Automated Binning of Microsatellite Alleles: Problems and Solutions. *Mol. Ecol. Notes* 7 (1), 10–14. doi:10.1111/j.1471-8286.2006.01560.x

Bhargava, A., and Fuentes, F. F. (2010). Mutational Dynamics of Microsatellites. *Mol. Biotechnol.* 44 (3), 250–266. doi:10.1007/s12033-009-9230-4

Burridge, C. P., Craw, D., Fletcher, D., and Waters, J. M. (2008). Geological Dates and Molecular Rates: Fish DNA Sheds Light on Time Dependency. *Mol. Biol. Evol.* 25 (4), 624–633. doi:10.1093/molbev/msm271

Burridge, C. P., McDowall, R. M., Craw, D., Wilson, M. V. H., and Waters, J. M. (2012). Marine Dispersal as a Pre-requisite for Gondwanan Vicariance Among Elements of the Galaxiid Fish Fauna. *J. Biogeogr.* 39 (2), 306–321. doi:10.1111/j.1365-2699.2011.02600.x

Camus, P. A. (2001). Biogeografía marina de Chile continental. *Rev. Chil. Hist. Nat.* 74 (3), 587–617. doi:10.4067/s0716-078x2001000300008

Carrea, C., Barriga, J. P., Cussac, V. E., and Ruzzante, D. E. (2012). Genetic and Phenotypic Differentiation Among *Galaxias Maculatus* Populations in a Patagonian Postglacial lake System. *Biol. J. Linn. Soc. Lond.* 107 (2), 368–382. doi:10.1111/j.1095-8312.2012.01939.x

Carrea, C., Cussac, V. E., and Ruzzante, D. E. (2013). Genetic and Phenotypic Variation among Galaxias Maculatus populations Reflects Contrasting Landscape Effects between Northern and Southern Patagonia. *Freshw. Biol.* 58 (1), 36–49. doi:10.1111/fwb.12036

Carrea, C., Paterson, I. G., Cussac, V. E., and Ruzzante, D. E. (2009). Ten Novel Microsatellite Loci Characterized for a Remarkably Widespread Fish: *Galaxias Maculatus* (Galaxiidae). *Mol. Ecol. Resour.* 9, 1503–1505. doi:10.1111/j.1755-0998.2009.02702.x

Clapperton, C. M. (1994). The Quaternary Glaciation of Chile: a Review. *Rev. Chil. Hist. Nat.* 67, 369–383.

Corush, J. B. (2019). Evolutionary Patterns of Diadromy in Fishes: More Than a Transitional State between marine and Freshwater. *BMC Evol. Biol.* 19 (1), 168–213. doi:10.1186/s12862-019-1492-2

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: More Models, New Heuristics and Parallel Computing. *Nat. Methods* 9 (8), 772. doi:10.1038/nmeth.2109

Delgado, M. L., and Ruzzante, D. E. (2020). Investigating Diadromy in Fishes and its Loss in an -Omics Era. *iScience* 23, 101837. doi:10.1016/j.isci.2020.101837

Delgado, M. L., Górski, K., Habit, E., and Ruzzante, D. E. (2019). The Effects of Diadromy and its Loss on Genomic Divergence: The Case of Amphidromous *Galaxias Maculatus* Populations. *Mol. Ecol.* 28 (24), 5217–5231. doi:10.1111/mec.15290

Delgado, M. L., Manosalva, A., Urbina, M. A., Habit, E., Link, O., and Ruzzante, D. E. (2020). Genomic Basis of the Loss of Diadromy in cursive: Insights from Reciprocal Transplant Experiments. *Mol. Ecol.* 29 (24), 4857–4870. doi:10.1111/mec.15686

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

Dyer, B. (2000). Systematic Review and Biogeography of the Freshwater Fishes of Chile. *Estud. Oceanol.* 19, 77–98.

Earl, D. A., and VonHoldt, B. M. (2012). STRUCTURE HARVESTER: a Website and Program for Visualizing STRUCTURE Output and Implementing the Evanno Method. *Conservation Genet. Resour.* 4 (2), 359–361. doi:10.1007/s12686-011-9548-7

Edgar, R. C. (2004). MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* 32 (5), 1792–1797. doi:10.1093/nar/gkh340

Estoup, A., Jarne, P., and Cornuet, J.-M. (2002). Homoplasy and Mutation Model at Microsatellite Loci and Their Consequences for Population Genetics Analysis. *Mol. Ecol.* 11 (9), 1591–1604. doi:10.1046/j.1365-294x.2002.01576.x

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 (8), 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 (3), 564–567. doi:10.1111/j.1755-0998.2010.02847.x

Fernández Iriarte, P. J., González-Wevar, C. A., Segovia, N. I., Rosenfeld, S., Hüne, M., Fainburg, L., et al. (2020). Quaternary Ice Sheets and Sea Level Regression Drove Divergence in a marine Gastropod along Eastern and Western Coasts of South America. *Sci. Rep.* 10 (1), 844. doi:10.1038/s41598-020-57543-4

Fraser, C. I., Nikula, R., Ruzzante, D. E., and Waters, J. M. (2012). Poleward Bound: Biological Impacts of Southern Hemisphere Glaciation. *Trends Ecol. Evol.* 27 (8), 462–471. doi:10.1016/j.tree.2012.04.011

González-Wevar, C. A., Rosenfeld, S., Segovia, N. I., Hüne, M., Gérard, K., Ojeda, J., et al. (2016). Genetics, Gene Flow, and Glaciation: The Case of the South American Limpet *Nacella Mytilina*. *PLoS ONE* 11, e0161963. doi:10.1371/journal.pone.0161963

González-Wevar, C. A., Salinas, P., Hüne, M., Segovia, N. I., Vargas-Chacoff, L., Astorga, M., et al. (2015a). Phylogeography in *Galaxias Maculatus* (Jenyns, 1848) along Two Biogeographical Provinces in the Chilean Coast. *PloS ONE* 10 (7), e0131289. doi:10.1371/journal.pone.0131289

González-Wevar, C. A., Hüne, M., Cañete, J. I., Mansilla, A., Nakano, T., and Poulin, E. (2012). Towards a Model of Postglacial Biogeography in Shallow marine Species along the Patagonian Province: Lessons from the Limpet *Nacella Magellanica* (Gmelin, 1791). *BMC Evol. Biol.* 12, 139. doi:10.1186/1471-2148-12-139

González-wevar, C. A., Nakano, T., Cañete, J. I., and Poulin, E. (2011). Concerted Genetic, Morphological and Ecological Diversification in *Nacella* Limpets in the Magellanic Province. *Mol. Ecol.* 20, 1936–1951. doi:10.1111/j.1365-294x.2011.05065.x

González-Wevar, C., Segovia, T., Morley, S., Chown, S., and Poulin, E. (2013). Extinction and Recolonization of Maritime Antarctica in the Limpet Nacella Concinna (Strebel, 1908) During the Last Glacial Cycle: Toward a Model of Quaternary Biogeography in Shallow Antarctic Invertebrates. *Mol. Ecol.* 22, 5221–5236. doi:10.1111/mec.12465

González-Wevar, C. A., Segovia, N. I., Rosenfeld, S., Ojeda, J., Hüne, M., Naretto, J., et al. (2018). Unexpected Absence of Island Endemics: Long-Distance Dispersal in Higher Latitude Sub-Antarctic Siphonaria (Gastropoda: Euthyneura) Species. *J. Biogeogr.* 45, 874–884. doi:10.1111/jbi.13174

González-Wevar, C., Salinas, P., Hüne, M., Segovia, N., Vargas-Chacoff, L., Oda, E., et al. (2015b). Contrasting Genetic Structure and Diversity of Galaxias maculatus (Jenyns, 1848) along the Chilean Coast: Stock Identification for Fishery Management. *Jhered* 106, 439–447. doi:10.1093/jhered/esv005

Goudet, J. (2002). FSTAT (Version 2.9. 3.2): A Program to Estimate and Test Gene Diversities and Fixation Indices. Available at: http://www. unil. ch/izea/software/fstat. html.

Habit, E., Dyer, B., and Vila, I. (2006). Estado de conocimiento de los peces dulceacuícolas de Chile. *Gayana* 70 (1), 100–113. doi:10.4067/s0717-65382006000100016

Hartl, D. L., and Clark, A. G. (1997). *Principles of Population Genetics*. Second edition. Sunderland: Sinauer Associates, 682.

Hein, A. S., Hulton, N. R. J., Dunai, T. J., Sugden, D. E., Kaplan, M. R., and Xu, S. (2010). The Chronology of the Last Glacial Maximum and Deglacial Events in central Argentine Patagonia. *Quat. Sci. Rev.* 29, 1212–1227. doi:10.1016/j.quascirev.2010.01.020

Hewitt, G. M. (2004). Genetic Consequences of Climatic Oscillations in the Quaternary. *Phil. Trans. R. Soc. Lond. B* 359, 183–195. doi:10.1098/rstb.2003.1388

Hewitt, G. (2000). The Genetic Legacy of the Quaternary Ice Ages. *Nature* 405, 907–913. doi:10.1038/35016000

Hollin, J. T., and Schilling, D. H. (1981). "Late Wisconsin-Wichselian Mountain Glaciers and Small Ice Caps," in *The Last Great Ice Sheets: 179-206*. Editors J H Denton and T J Hughes (New York: Wiley). Available at: https://oceanrep.geomar.de/id/eprint/37507.

Hudson, R. (1990). Gene Genealogies and the Coalescent Process. *Oxf. Surv. Evol. Biol.* 7, 1–44.

Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian Inference of Phylogenetic Trees. *Bioinformatics* 17 (8), 754–755. doi:10.1093/bioinformatics/17.8.754

Hulton, N., Purves, P., McCulloch, R., Sugden, D., and Bentley, M. (2002). The Last Glacial Maximum and Deglaciation in Southern South America. *Quat. Sci. Rev.* 21, 233–241. doi:10.1016/S0277-3791(01)00103-2

Jarne, P., and Lagoda, P. J. L. (1996). Microsatellites, from Molecules to Populations and Back. *Trends Ecology Evolution* 11 (10), 424–429. doi:10.1016/0169-5347(96)10049-5

Jombart, T. (2008). Adegenet: a R Package for the Multivariate Analysis of Genetic Markers. *Bioinformatics* 24 (11), 1403–1405. doi:10.1093/bioinformatics/btn129

Jombart, T., and Ahmed, I. (2011). Adegenet 1.3-1: New Tools for the Analysis of Genome-wide SNP Data. *Bioinformatics* 27 (21), 3070–3071. doi:10.1093/bioinformatics/btr521

Librado, P., and Rozas, J. (2009). DnaSP V5: a Software for Comprehensive Analysis of DNA Polymorphism Data. *Bioinformatics* 25 (11), 1451–1452. doi:10.1093/bioinformatics/btp187

Maggs, C. A., Castilho, R., Foltz, D., Henzler, C., Jolly, M. T., Olsen, J., et al. (2008). Evaluating Signatures of Glacial Refugia for North Atlantic Benthic marine Taxa. *Ecology* 89, S108–S122. doi:10.1890/08-0257.1

Marko, P. B. (2004). 'What's Larvae Got to Do with it?' Disparate Patterns of post-glacial Population Structure in Two Benthic marine Gastropods with Identical Dispersal Potential. *Mol. Ecol.* 13, 597–611. doi:10.1046/j.1365-294x.2004.02096.x

Marko, P. B., Hoffman, J. M., Emme, S. A., McGovern, T. M., Keever, C. C., and Nicole Cox, L. (2010). The 'Expansion-Contraction' Model of Pleistocene Biogeography: Rocky Shores Suffer a Sea Change? *Mol. Ecol.* 19, 146–169. doi:10.1111/j.1365-294x.2009.04417.x

McCulloch, R. D., Bentley, M. J., Purves, R. S., Hulton, N. R. J., and Clapperton, D. E. (2000). Climatic Inferences from Glacial and Palaeoecological Evidence at the Last Glacial Termination, Southern South America. *J. Quat. Sci.* 15 (4), 409–417. doi:10.1002/1099-1417(200005)15:4<409::aid-jqs539>3.0.co;2

McDowall, R. M. (1970). Comments on a New Taxonomy ofRetropinna(Galaxioidei: Retropinnidae). *New Zealand J. Mar. Freshw. Res.* 4, 312–324. doi:10.1080/00288330.1970.9515349

McDowall, R. M. (1997). The Evolution of Diadromy in Fishes (Revisited) and its Place in Phylogenetic Analysis. *Rev. Fish. Biol. Fish.* 7 (4), 443–462. doi:10.1023/a:1018404331601

Moreno, P. I., Videla, J., Valero-Garcés, B., Alloway, B. V., and Heusser, L. E. (2018). A Continuous Record of Vegetation, Fire-Regime and Climatic Changes in Northwestern Patagonia Spanning the Last 25,000 Years. *Quat. Sci. Rev.* 198, 15–36. doi:10.1016/j.quascirev.2018.08.013

Pardo-Gandarillas, M. C., Ibáñez, C. M., Torres, F. I., Sanhueza, V., Fabres, A., Escobar-Dodero, J., et al. (2018). Phylogeography and Species Distribution Modelling Reveal the Effects of the Pleistocene Ice Ages on an Intertidal Limpet from the South-eastern Pacific. *J. Biogeogr.* 45 (8), 1751–1767. doi:10.1111/jbi.13362

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research-Aan Update. *Bioinformatics* 28 (19), 2537–2539. doi:10.1093/bioinformatics/bts460

Pons, O., and Petit, R. J. (1996). Measwring and Testing Genetic Differentiation with Ordered versus Unordered Alleles. *Genetics* 144 (3), 1237–1245. doi:10.1093/genetics/144.3.1237

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155 (2), 945–959. doi:10.1093/genetics/155.2.945

Provan, J., and Bennett, K. (2008). Phylogeographic Insights into Cryptic Glacial Refugia. *Trends Ecol. Evol.* 23 (10), 564–571. doi:10.1016/j.tree.2008.06.010

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. Second edition. Vienna, Austria: R Foundation for Statistical Computing.

Rabassa, J., Coronato, A., and Martínez, O. (2011). Late Cenozoic glaciations in Patagonia and Tierra del Fuego: an updated review. *Biol. J. Linn. Soc.* 103, 316–335. doi:10.1111/j.1095-8312.2011.01681.x

Rabassa, J., Coronato, A. M., and Salemme, M. (2005). Chronology of the Late Cenozoic Patagonian Glaciations and Their Correlation with Biostratigraphic Units of the Pampean Region (Argentina). *J. South Am. Earth Sci.* 20, 81–103. doi:10.1016/j.jsames.2005.07.004

Rogers, A. R., and Harpending, H. (1992). Population Growth Makes Waves in the Distribution of Pairwise Genetic Differences. *Mol. Biol. Evol.* 9, 552–569. doi:10.1093/oxfordjournals.molbev.a040727

Rojo, J. H., Fernández, D. A., Figueroa, D. E., and Boy, C. C. (2020). Phenotypic and Genetic Differentiation between Diadromous and Landlocked puyenGalaxias Maculatus. *J. Fish. Biol.* 96 (4), 956–967. doi:10.1111/jfb.14285

Rousset, F. (2008). genepop'007: a Complete Re-implementation of the Genepop Software for Windows and Linux. *Mol. Ecol. Resour.* 8 (1), 103–106. doi:10.1111/j.1471-8286.2007.01931.x

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

Taggart, J. B., Hynes, R. A., Prodöuhl, P. A., and Ferguson, A. (1992). A Simplified Protocol for Routine Total DNA Isolation from Salmonid Fishes. *J. Fish. Biol.* 40 (6), 963–965. doi:10.1111/j.1095-8649.1992.tb02641.x

Valdovinos, C., Navarrete, S. A., and Marquet, P. A. (2003). Mollusk Species Diversity in the Southeastern Pacific: Why Are There More Species towards the Pole? *Ecography* 26, 139–144. doi:10.1034/j.1600-0587.2003.03349.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. Notes* 4 (3), 535–538. doi:10.1111/j.1471-8286.2004.00684.x

Vera-Escalona, I., Delgado, M. L., Habit, E., and Ruzzante, D. E. (2020). Historical and Contemporary Diversity of Galaxiids in South America: Biogeographic and Phylogenetic Perspectives. *Diversity* 12 (8), 304. doi:10.3390/d12080304

Victoriano, P. F., Muñoz-Ramírez, C. P., Canales-Aguirre, C. B., Jara, A., Vera-Escalona, I., Burgos-Careaga, T., Muñoz-Mendoza, C., and Habit, E. M. (2020). Contrasting Evolutionary Responses in Two Co-distributed Species of *Galaxias* (Pisces, Galaxiidae) in a River from the Glaciated Range in Southern Chile. *R. Soc. Open Sci.* 7 (7), 200632. doi:10.1098/rsos.200632

Wallis, G. P., Judge, K. F., Bland, J., Waters, J. M., and Berra, T. M. (2001). Genetic Diversity in New Zealand *Galaxias Vulgaris* Sensu Lato (Teleostei: Osmeriformes: Galaxiidae): a Test of a Biogeographic Hypothesis. *J. Biogeogr.* 28 (1), 59–67. doi:10.1046/j.1365-2699.2001.00535.x

Waters, J. M., Burridge, C. P., and Craw, D. (2020). River Capture and Freshwater Biological Evolution: A Review of Galaxiid Fish Vicariance. *Diversity* 12 (6), 216. doi:10.3390/d12060216

Waters, J. M., and Burridge, C. P. (1999). Extreme Intraspecific Mitochondrial DNA Sequence Divergence inGalaxias maculatus(Osteichthys: Galaxiidae), One of the World's Most Widespread Freshwater Fish. *Mol. Phylogenet. Evol.* 11, 1–12. doi:10.1006/mpev.1998.0554

Waters, J. M., Dijkstra, L. H., and Wallis, G. P. (2000). Biogeography of a Southern Hemisphere Freshwater Fish: How Important Is marine Dispersal? *Mol. Ecol.* 9, 1815–1821. doi:10.1046/j.1365-294x.2000.01082.x

Weir, B. S., and Cockerham, C. C. (1984). Estimating F -Statistics for the Analysis of Population Structure. *Evolution* 38 (6), 1358–1370. doi:10.1111/j.1558-5646.1984.tb05657.x

Xia, X. (2013). DAMBE5: a Comprehensive Software Package for Data Analysis in Molecular Biology and Evolution. *Mol. Biol. Evol.* 30 (7), 1720–1728. doi:10.1093/molbev/mst064

Zattara, E., and Premoli, A. (2005). Genetic Structuring in Andean Landlocked Populations of *Galaxias Maculatus*: Effects of Biogeographic History. *J. Biogeo.* 32, 5–14. doi:10.1111/j.1365-2699.2004.01164.x

Zemlak, T. S., Habit, E. M., Walde, S. J., Battini, M. A., Adams, E. D. M., 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 ofGalaxias Plateiin Patagonia. *Mol. Ecol.* 17 (23), 5049–5061. doi:10.1111/j.1365-294x.2008.03987.x

Zemlak, T. S., Habit, E. M., Walde, S. J., Carrea, C., and Ruzzante, D. E. (2010). Surviving Historical Patagonian Landscapes and Climate: Molecular Insights from *Galaxias Maculatus*. *BMC Evol. Biol.* 10 (1), 67. doi:10.1186/1471-2148-10-67

Zemlak, T. S., Walde, S. J., Habit, E. M., and Ruzzante, D. E. (2011). Climate-induced Changes to the Ancestral Population Size of Two Patagonian Galaxiids: the Influence of Glacial Cycling. *Mol. Ecol.* 20 (24), 5280–5294. doi:10.1111/j.1365-294x.2011.05352.x

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 03 June 2022
doi: 10.3389/fgene.2022.812971

Heterologous Production and Evaluation of the Biological Activity of Cystatin-B From the Red Piranha *Pygocentrus nattereri*

Juan Antonio Ramirez Merlano[1] *and Daniela Volcan Almeida*[2]*

[1]Aquaculture Institute, University of the Llanos, Meta, Colombia, [2]Laboratory of Molecular Biology, Institute of Biological Sciences, Federal University of Rio Grande, Rio Grande, Brazil

Cystatin proteins are known to form a superfamily of cysteine protease inhibitors, which play a key role in protein degradation and are related to different physiological processes, such as development and immunity. Currently, numerous immunoregulatory proteins, such as cystatins, are being used in the control and prevention of diseases in aquaculture. Thus, the objective of this study was to produce recombinant cystatin (rCYST-B) from the red piranha *Pygocentrus nattereri* and to evaluate its effect on bacterial growth. The gene that encodes cystatin-B was isolated from the spleen of *P. nattereri* and cloned in an expression system. The protein was produced via a heterologous system involving the yeast *Pichia pastoris* X-33. The inhibitory activity of purified cystatin-B was evaluated on papain using different concentrations (0–80.0 μg/μL) of rCYST-B. The bacteriostatic action of the protein was evaluated using the *Kirby-Bauer* method on the growth of *Escherichia coli* and *Bacillus subtilis*. rCYST-B showed 100% inhibition at a concentration of 60 μg/μL. Moreover, the bacteriostatic activity of *E. coli* and *B. subtilis* showed inhibition of 40.36 and 49.08% compared to the negative control (phosphate buffer), respectively. These results suggest that recombinant CYST-B has biotechnological potential for use in aquaculture.

Keywords: aquatic diseases, fish, recombinant protein, transgene, stefins

OPEN ACCESS

Edited by:
Sandra Isabel Moreno Abril, University of Vigo, Spain

Reviewed by:
Guilherme Campos Tavares, Federal University of Minas Gerais, Brazil
William Borges Domingues, Federal University of Pelotas, Brazil

***Correspondence:**
Daniela Volcan Almeida
danivolcan@pq.cnpq.br

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

Received: *11 November 2021*
Accepted: *21 April 2022*
Published: *03 June 2022*

Citation:
Ramirez Merlano JA and Almeida DV (2022) Heterologous Production and Evaluation of the Biological Activity of Cystatin-B From the Red Piranha Pygocentrus nattereri. Front. Genet. 13:812971. doi: 10.3389/fgene.2022.812971

HIGHLIGHTS

- Cystatin-B was successfully expressed in methylotrophic yeast, *Pichia pastoris* X-33.
- Functional analysis of purified rCystatin-B demonstrated papain inhibition activity.
- rCystatin-B efficiently inhibited the growth of *E. coli* and *B. subtilis*.
- Recombinant Cystatin-B exhibited potential for therapeutic use in fish.

INTRODUCTION

Aquaculture is a productive sector with accelerated growth, an activity that has become fundamental for the achievement of the Sustainable Development Goals (SDGs) proposed by the United Nations (FAO 2017). According to the Food and Agriculture Organization of the United Nations (FAO) global fisheries production included approximately 179 million tons of fish, of which aquaculture represented production of 82.1 million tons, corresponding to 46% of the total production in 2018 (FAO SOFIA 2020). However, this growth was accompanied by the emergence or re-emergence of

several infectious diseases (Pérez-Sánchez et al., 2018; FAO SOFIA 2020), such as acute hepatopancreatic necrosis (AHPND) (OIE 2013), enterocytozoon hepatopenaei (EHP) (Tang et al., 2017), emerging bacteria of the genus *Acinetobacter*, a potential pathogen in shrimp *Penaeus vannamei* (Huang et al., 2020) and some parasitic infections such as acanthocephalosis, reported for several groups of fish (Valladão et al., 2019), which have a detrimental influence on their health (Rosny et al., 2016). To overcome this problematic it is necessary to apply good management techniques and to develop new technologies. Biotechnology methods and genetic engineering enable the production of recombinant proteins, which can be used both in the prevention and treatment of diseases, and are considered promising for use in the aquaculture industry.

Cystatins are a large group of proteins that function as protease inhibitors involving cysteine residues in the proteolytic reaction (Berti and Storer 1995). Typically, they are small proteins that inactivate the protease substrate in a specific manner, forming reversible complexes (Turk and Bode 1991). Cysteines play essential roles in the physiology of all living organisms, from protozoa to mammals. In pathogenic microorganisms including bacteria, fungi, and parasites, cysteine protease can act as virulence factor, causing diseases in host organisms (Mottram et al., 2004; Rudenskaya and Pupov 2008). The cystatin superfamily is grouped into three families (Barrett 1986; Magister and Kos 2013), namely, family I (*Stefin*), family II (cystatins), and family III (kininogen). A fourth family has been reported of invertebrate origin, mainly of nematode parasites (Khaznadji et al., 2005; Li et al., 2010). Family I, composed of *Stefin* A and B, also known as cystatin A and B, are single-chain polypeptides of approximately 100 residues and molecular weight between 10 and 11 kDa without disulfide bonds or carbohydrate side chains. These are intracellular protease inhibitors present in the cytosol (Barrett 1986; Turk and Bode 1991; Abrahamson et al., 2003). Family 2 cystatins include cystatin C, D, E/M, F, G, S, SA, and SN with a molecular weight between 13–14 kDa and exhibit a signal peptide and two disulfide bridges (Cornwall and Hsia 2003). Family III (kininogens), are large glycoproteins (60–120 kDa) and complex in structure and are found in body fluids, especially in plasma (Barrett 1986; Prunk et al., 2016). All cystatins, regardless of family classification, contain several conserved regions, including an N-terminal glycine segment, a QXVXG sequence that constitutes part of the β-hairpin *loop* structure, and a proline-tryptophan-containing region that forms a second *hairpin loop* (Margis et al., 1998; Rzychon and Chmiel 2004). Cystatins play important defensive and regulatory roles during various cellular events. They modulate and stimulate TNF-α and IL-10 synthesis as a defense strategy in response to pathogen infection (Verdot et al., 1999).

In fish, cystatins exhibit protease inhibitory activity and appear to be involved in immune responses against infectious agents (Xiao et al., 2010; Premachandra et al., 2012; Ahn et al., 2013). To explore the functions of fish cystatin, some recombinant cystatins from Keta salmon (Yamashita and Konagaya 1991), trout (Li et al., 2000), and carp (Tzeng et al., 2001) were generated by heterologous expression in *Escherichia coli*. Although these recombinant proteins have already been developed for fish, the use of a species-specific protein or that obtained from a phylogenetically close species may decrease the side effects, besides potentiating the expected immunoregulatory effect.

Currently, genome sequencing allows the use of genetic information to produce specific proteins for application in aquaculture. However, few species of commercial interest have their genome completely sequenced. One such species is the red piranha *Pygocentrus nattereri*. The red piranha is a widely distributed carnivorous fish species that is observed in the rivers and lakes of South America (Behr and Signor 2008). In addition to being a popular aquarium fish, piranhas are often sold for human consumption in local markets in the Amazon basin (Duponchelle et al., 2007). Another important characteristic of this species is its similarity phylogenetic with other freshwater species such as tambaqui (*Colossoma macropomum* of the family Serrasalmidae). Associating the genetic knowledge of species of commercial interest to the development of biotechnologies allows the production of recombinant proteins of industrial interest, such as cystatin, for applications in aquaculture to maintain animal health. Thus, in this study, the main objective was the heterologous production and evaluation of the biological activity of the Cystatin-B derived from the red piranha *Pygocentrus nattereri*.

MATERIAL AND METHODS

Source of Animals and Sample Collection

Red piranha *Pygocentrus nattereri* juveniles, weighing approximately 200 g were captured at Janauacá Lake, municipality of Manaquiri - Amazonas, Brazil. The fish were euthanized (n = 13) with eugenol (200 mg/L), and the spleen tissue was harvested, preserved in RNAlater (Ambion, United States) and stored at −80°C.

RNA Isolation and cDNA Synthesis

Total RNA extraction from the red piranha spleen samples was performed using the total RNA extraction protocol using Trizol Reagent ® (Invitrogen, United States), according to the manufacturer's recommendations. The quality of the extracted RNA was evaluated by 1% (w/v) agarose gel electrophoresis and quantified using a spectrophotometer (BioDrop - Biodrop μLite). Total RNA was treated with the DNAse I (RQ1 RNase-Free DNase - Promega) following the protocol described by the manufacturer. Complementary DNA (cDNA) was generated using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, United States) from 1 μg of total RNA using oligo d (T) primer.

Isolation of *Cystatin*-B Gene

Isolation of the genes encoding *cystatin*-B was performed by polymerase chain reaction (PCR). Specific primers were previously designed with the Primer-BLAST tool from GenBank (www.ncbi.nlm.nih.gov) based on the sequences

XM_017717432 of *Pygocentrus nattereri* and used in this experiment. The primer sequences are listed as follows: Forward: 5′ CAGCAGGAGAGCAGAAGTTGA 3′ e Reverse 5′ TGTTAGTACGGTTTGTTAAGGGGA 3'. GoTaq® (Promega Corporation. Madison, United States) was used in the reaction, following the manufacturer's recommendations. PCR cycles were set as follows: initial cycle at 94°C for 2 min, 35 cycles at 94°C for 30 s, 60°C for 30 s, and 72°C for 30 s, followed by a final extension at 72°C for 5 min.

After verifying the correct fragment size using electrophoresis in 1% agarose gel, the samples were purified using the PureLink™ PCR Purification Kit® (Invitrogen, Brazil), following the manufacturer's protocol. The purified PCR products were sequenced using the Sanger method on the Applied Biosystems® Sanger Sequencing 3500 Genetic Analyzer. Automated sequencing was performed by the company ACTGene Molecular Analyses (UFRGS, Brazil). The sequencing result was first analyzed by the quality of the electropherograms obtained using the Chromas software and then aligned using the BLASTn tool (https://blast.ncbi.nlm.nih.gov).

Bioinformatics Analysis

After confirmation by sequencing, protein analyzes were performed based on the predicted protein sequences using the Genebank database. Was used the cystatin-B isoform X1 (XP_017572921.1) protein. Multiple alignment analysis of the amino acid sequences of the CYSTATIN-B protein (named CYST-B herein) was performed using Geneious Prime Software (version 2020.1.2). The signal peptide deduced from the amino acid sequence was predicted by XtalPred (https://xtalpred.godziklab.org) and SignalP 5.0 Software (http://www.cbs.dtu.dk/services/SignalP). The secondary structure of the amino acid sequence of the protein was constructed using PredictProtein (https://predictprotein.org) and Phyre2 system (http://www.sbg.bio.ic.ac.uk/phyre2). The 3D structural model of the protein was established using the Phyre2 system. The phylogenetic tree was constructed based on the amino acid sequence of Cystatin-B, using MEGA X (Kumar et al., 2018) and the *Neighbour-joining* method with 1000 *bootstrap* replicates.

Plasmid Construction

Specific primers were re-designed by inserting restriction enzyme sites of XhoI and XbaI (cystBXhoI- For = 5′AACTCGAGCAGC AGCAGGAGAGCAGAAGTTGA 3′ and cystBXbaI-Rev = 5′AATCTAGATGTTAGTTAGTACGGTTTGTTAAGGGGA 3′). The insertion of these sites allowed the cloning of fragments in the pPICZα expression vector (**Figure 1**). GoTaq® (Promega Corporation. Madison, United States) was used in the reaction, following the manufacturer's recommendations. PCR cycles were set as follows: initial cycle at 94°C for 2 min, 35 cycles at 94°C for 30 s, 60°C for 30 s, and 72°C for 30 s, plus a final extension at 72°C for 5 min. Verification of the size of the PCR products was performed using agarose gel electrophoresis (1%). After confirmation, the PCR products were purified using the PureLink PCR Purification Kit (Thermo Fisher Scientific, United States) and cleaved using the enzymes XhoI and XbaI (Promega Corporation, United States), following the protocol described by the manufacturer (**Figure 1B**). The cleaved fragments were further purified (as described above) for binding to the pPICZα expression vector. The fragment vector was obtained by PCR using the pPICZα plasmid as the template (**Figure 1C**). The primers pPIC-Forward (XbaI) ATCTAGAAC AAAAACTCATCTCAGAGAAGAGG and pPIC-Rever (XhoI) AGCTTCAGCCTCTCTTTTCTCGAGAG were used. PCR conditions were set as follows: an initial cycle of 94°C for 2 min, 35 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 4 min in addition to a final extension of 72°C for 10 min. The pPIC vector was also cleaved with the same restriction enzymes (XbaI and XhoI) to allow cohesive sequence formation and binding at complementary sites. The binding of the cleaved and purified insert to the vector was performed using the enzyme T7 ligase (Invitrogen, Brazil), following the manufacturer's recommendations (**Figure 1D**).

Transformation Into *Escherichia coli* With the Cloning Vector

The obtained genetic constructs were used for *Escherichia coli* TOP10 transformation. The calcium chloride method was used to generate chemically competent cells. Plasmid DNA was extracted using the plasmid DNA kit PureYield ™ Plasmid Miniprep System (Promega). After transformation, bacteria were grown overnight (37°C) in culture plates in LB (Luria-Bertani) solid medium and 50 μg/ml of the selection antibiotic Zeocin™ (Invitrogen, Brazil). *E. coli* colonies transformed with the cloned plasmid pPIC_cystatin -B were inoculated in LB liquid culture medium containing Zeocin (50 $\mu g/ml^{-1}$) with shaking at 225 rpm and a temperature of 37°C. Plasmid DNA extraction was performed using the PureYield™ Plasmid Miniprep System plasmid DNA kit (Invitrogen, Brazil) according to the manufacturer's recommendations. Plasmids extracted were cleaved with XhoI, XbaI, BamHI, and *Bg*LII and visualized using 1% agarose gel electrophoresis, as well as by PCR using different combinations of primers for specific regions both in the gene of interest and regions of the pPIC_cystatin B plasmid: 5′AOX- (3′ GACTGGTTCCAATTGACAAGC 5′) and 3′AOX -(5′ GCAAATGGCATTCTGACATCC 3′).

Cloning in the Yeast *Pichia pastoris*

Following verification, the plasmid was linearized using the restriction enzyme BamHI, the products were visualized by 1% agarose gel electrophoresis, quantified using a spectrophotometer (BioDrop, Isogen Life Science, Netherlands), and purified. Competent cells were prepared following the procedure in the Easy Select ™ Pichia Expression Kit manual (Thermo Fisher Scientific, United States). The transformation was performed in *P. pastoris* X33 yeast using the PichiaPink™ Expression System protocol (Thermo Fisher Scientific, United States) with some modifications (was used linearized plasmid DNA). A total of 80 μL of the cells with linearized plasmid DNA were mixed into electroporation cuvettes (0.2 cm) and incubated for 5 min on ice. They were subsequently electroporated (1500 V, 25 μF, 400 Ω). Immediately, 1 ml of ice-cold sorbitol was added to the cuvette,

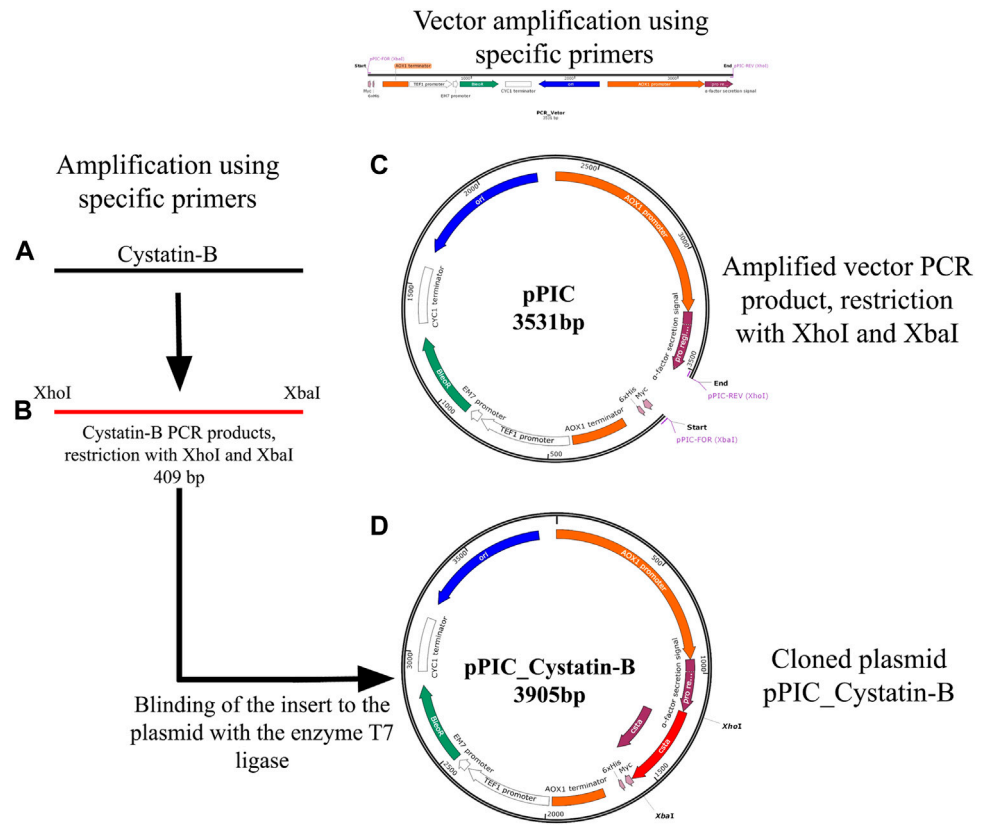


FIGURE 1 | Schematic representation of the genetic construct containing *cystatin-B* **(A)** Amplification by PCR of the *cystatin-B* gene of juveniles of *Pygocentrus nattereri*, using specific primers, designed based on the sequence XM_017717432. **(B)** Amplified PCR products and cleaved with XhoI and XbaI enzymes. **(C)**. Amplified vector (pPIC) products by PCR, cleaved with XhoI and XbaI. **(D)**. Cloned plasmid pPIC_Cystatin-B (Binding of the insert to the plasmid with the enzyme T7 ligase).

and the contents were transferred to a sterile 1.5 ml Eppendorf-type tube and incubated for 2 h at 30°C without shaking. A total of 100 µL of the transformed cells were plated onto Minimal Dextrose Medium (2% agar; 2% glucose; 4 × 10^{-5}% biotin; yeast nitrogen base YNB 1.34% and 125 µg/ml Zeocin). The plates were incubated at 30°C for a maximum of 4 days until the development of colonies and selection of clones for further induction analyses.

Expression and Purification of Recombinant Cystatin-B Protein

To increase the cell mass the selected clones and the negative control (without plasmid) of *P. pastoris* were inoculated in 200 ml of Buffered Glycerol Complex Medium (yeast extract 1%; peptone 2%; phosphate buffer pH 6.0 100 mM; YNB 1.34%; biotin 4 × 10^{-5}%; glycerol 1%) and incubated at 30°C with shaking at 250 rpm until the culture showed an OD_{600} between 2–6, in approximately 18 h. Upon reaching the required OD_{600}, the cultures were centrifuged at 3000 rpm for 10 min at 4°C and resuspended in 300 ml of Buffered Methanol Complex Medium (yeast extract 1%; peptone 2%; phosphate buffer pH 6.0 100 mM; YNB 1.34%; biotin 4 × 10^{-5}%; methane 1%) in 1 L Erlenmeyer flask. The expression of alcohol oxidase enzyme promoter 1 (AOX1) was induced by adding absolute methanol every 24 h to the cell cultures, maintaining a final concentration of 1% (v/v) for 120 h at 30°C with shaking at 250 rpm. Aliquots containing 50 ml of the cultures were collected, centrifuged at 3000 rpm for 10 min at 4°C, and preserved at −80°C for further analysis. The recombinant protein was purified using Ni Sepharose high-Performance nickel-based resin (GE Healthcare, United States) with polyhistidine tail (6xHis). The purified recombinant rCYST-B protein was analyzed via denaturing 15% polyacrylamide gel electrophoresis (SDS-PAGE). Subsequently, the gels were stained with Coomassie blue (0.25%) (**Figure 3A**), and silver nitrate (20%) (**Figure 3B**). The concentration of the purified recombinant protein was evaluated according to the method described by Lowry et al. (1951).

Protease Inhibition Assay

The inhibitory effect of the recombinant rCYST-B protein was analyzed using papain as a cysteine protease, performed as described in Xiao et al. (2010) modified for (Yu et al., 2019). The assay was performed using different amounts of rCYST-B proteins (0, 0.1; 0.5; 1.0; 10.0; 15.0; 20.0; 30.0; 60.0 and 80.0 µg), which were incubated with 10 µL of papain (0.1 µg/µL) (Sigma-Aldrich, United States)at 28°C for 30 min. The reaction was initiated by adding 200 µL of Azocasein (0.2% W/V) (Sigma-

Aldrich, United States), followed by incubation at 37°C for 2 h. The reaction was inactivated with 200 µL of 10% trichloroacetic acid (TCA). The contents were chilled on ice for 15 min. The precipitate was separated by centrifugation at 15,000 rpm for 5 min. The absorbance was evaluated by spectrophotometry at 450 nm. As a negative control, αGHT protein was used in place of rCYST-B. The relative activity of rCYST-B was determined by: 100x (1–($_{A440}$ of rCYST-B)/($_{A440}$ control).

Bacteriostatic Action Test

The assay for the identification of bacteriostatic action was performed using the *Kirby-Bauer* method (Bauer et al., 1966). Two types of bacteria (*Escherichia coli* or *Bacillus subtilis*) were plated on the surface of Petri dishes with LB medium and incubated at 37°C for 12–16 h. Disks were added with rCYST-B (60 and 80 µg/µL), Spectinomycin 50 ug/uL (positive control), and Potassium Phosphate Buffer 50 mM (negative control). The bacteriostatic action was evaluated qualitatively and quantitatively. The qualitative analysis was related to the presence of inhibition halos around the disk, and the quantitative analysis evaluated the halo diameter (mm), attributed to the bacteriostatic action of the protein. For halo diameter analysis, the plates were photographed (n = 4) and analyzed using ImageJ software.

Statistical Analysis

All experiments were performed in at least triplicates. All values are shown as mean ± standard deviation of the mean. Factorial type ANOVA was applied between the different treatments used to determine the bacteriostatic effect of rCYST-B; significant differences were determined with a $p < 0.05$. The analyses were performed using GraphPad Prism 8.0 software (GraphPad Software).

Ethical and Legal Aspects

All procedures adopted in this study were performed according to the protocols previously approved by the Ethics Committee on Animal Use (CEUA) of Universidade Nilton Lins with approval protocol 015/2017.

RESULTS

Characterization of CYST-B Structure

The PCR amplified gene products presented a fragment of 409 base pairs (bp) with 100% similarity to the partial sequence available in GeneBank (XM_017717432.2). The sequence of CYST-B showed 104 amino acid residues, with an estimated molecular weight of 11.755 kDa, a theoretical isoelectric point of 6.04, stability index of 33.98, without the presence of a signal peptide. The secondary structure analysis of CYST-B (**Figure 2A**), showed a structure composed of: *loop* 49.0%, helix 17.3%, and strand 33.7%. The 3D structure of CYST-B was predicted by Phyre2 (**Figure 2B**). The tertiary structure of CYST-B was colored using the rainbow command, blue at the N-terminal and red at the C-terminal. The rainbow command colors residues; 99 residues were modeled up to 100.0% confidence and a reliable model was obtained with the following characteristics: (Å): X: 44.068 Y:34.294 Z:30.445. Multiple homologous alignments of the amino acid sequences of CYST-B from *P. nattereri* (**Figure 2C**) revealed conserved regions typical of the cystatin superfamily, including two conserved N-terminal glycine residues (G10G11) as well as the glutamine-valine-glycine motif (QXVXG), crucial for the biological activity of the molecule. The motif represented by glutamine-leucine-valine-alanine-glycine (Q52LVAG56) exhibited homology with species such as *Colossoma macropomum* (GeneBank accession code XP_036414104.1) with 100% identity. In addition, a typical proline-tryptophan (PW) motif was identified at the C-terminus, where the tryptophan residue was replaced by a tyrosine residue (P80Y81). Based on the total amino acid sequence length of CYST-B, phylogenetic tree analysis was performed using MEGA X (**Figure 2D**). The result indicated a sequence similarity to that of tambaqui *Colossoma macropomum*, a phylogenetically close species.

Expression and Purification of the Recombinant rCYST-B Protein

SDS-PAGE showed that purified rCYST-B produced a band (**Figure 3B**, Lanes 3–8) with the expected size of approximately 11.8 kDa. These results confirm the expression of the recombinant rCYST-B protein in *P. pastoris* X33 cells (**Figure 3A**), demonstrating a suitable platform for the production of heterologous cystatin proteins.

Inhibitory Activity

The *in vitro* inhibitory activity of the recombinant rCYST-B protein on papain (a cysteine protease), revealed a concentration-dependent inhibition response of the inhibitor, with the increase starting at a concentration of 30 µg/µL and reaching 100% inhibition at the concentration of 60 µg/µL (**Figure 4**).

Analysis of Growth Inhibition and Bacteriostatic Action

The bacteriostatic activity of the recombinant rCYST-B protein on *E. coli* (Gram-negative bacteria) and *B. subtilis* (Gram-positive bacteria) demonstrated the presence of inhibition halos (zone of inhibition) at concentrations of 60 and 80 µg/µL of rCYST-B (**Figure 5A**). The diameter of the inhibition halo at the concentration of 60 µg/µL was 9.66 ± 0.61 mm for *E. Coli* and 10.50 ± 2.94 mm ($p > 0.05$) for *B. Subtilis* (**Figure 5B**). At the concentration of 80 µg/µL, the halo diameter was 8.64 ± 0.70 mm and 12.41 ± 5.53 mm for *E. coli* and *B. subtilis*, respectively (**Figure 5C**). These values represent a bacteriostatic action in comparison to the negative control (PBS), which did not show a halo of inhibition for any of the replicates. The positive control (spectinomycin antibiotic) showed a halo with a diameter of 23.93 ± 1.42 mm and 21.39 ± 1.83 mm at a concentration of 60 µg/µL for *E. coli* and *B. subtilis*, respectively. At the concentration of 80 µg/µL, the halo of inhibition in the

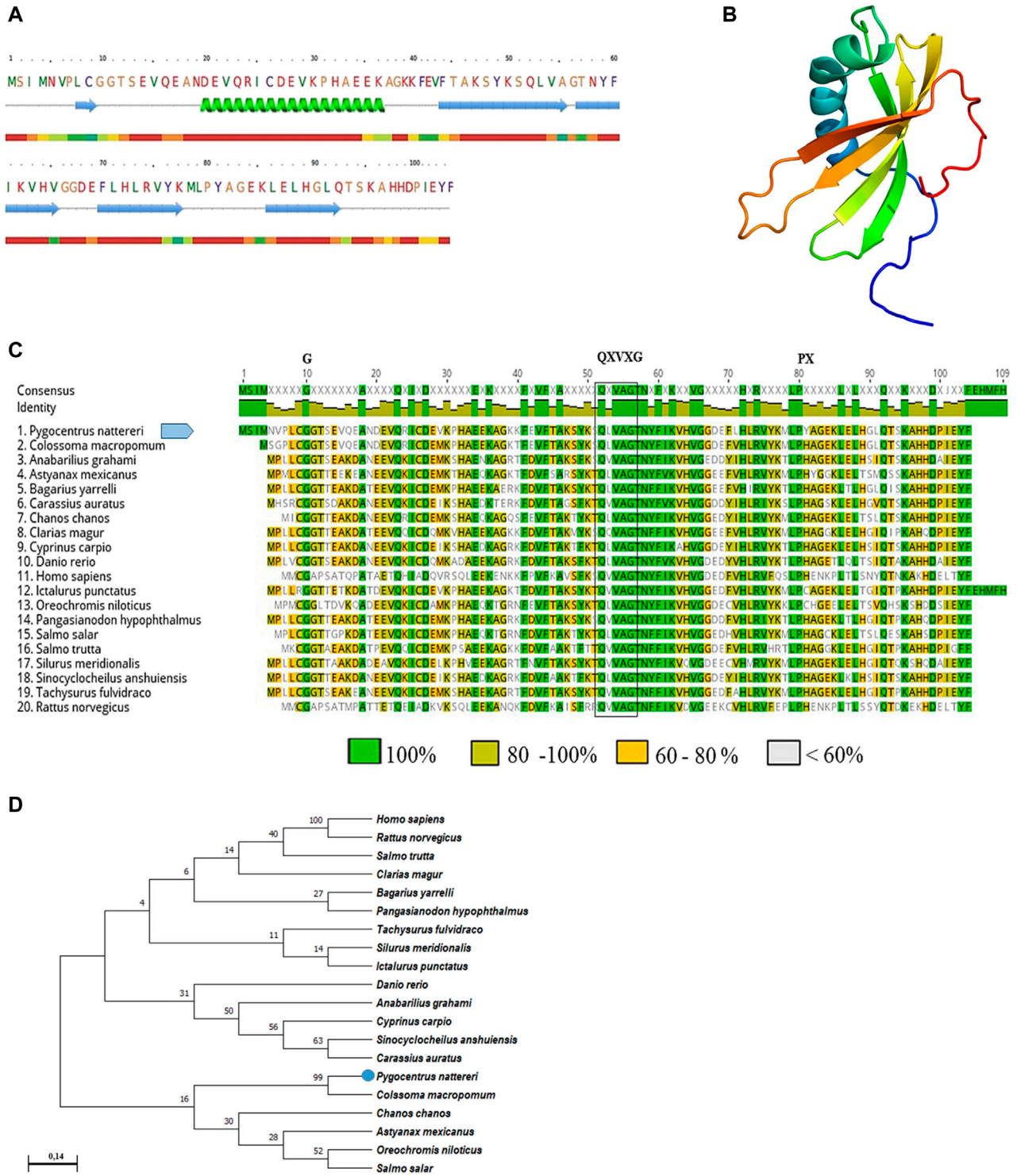


FIGURE 2 | Bioinformatic analysis of *Pygocentrus nattereri* cystatin-B **(A)**. Secondary structure of cystatin-B of *P. nattereri* generated by PredictProtein and Phyre2 system. fx1 Alpha helix, fx2Beta Strand. **(B)**. The 3D tertiary structure of Cystatin-B generated by the Phyre2 system. Model dimensions (Å): X:44,068 Y:34.294 Z: 30.445. **(C)**. Multiple alignments of the cystatin-B amino acid sequence of *Pygocentrus nattereri* and other sequences. *Colossoma macropomum,* XP_036414104.1; *Anabarilius grahami,* ROK23366.1*; Astyanax mexicanus,* XP_007249105.1; *Bagarius yarrelli,* TSK98422.1; *Carassius auratus*, XP_026093937.1; *Chanos*, XP_030625480.1; *Clarias magur*, KAF5903616.1; *Cyprinus carpio*, XP_018952861.1; *Danio rerio*, NP_001096599.1; *Homo sapiens*, NP_000091.1; *Ictalurus punctatus*, XP_017308724.1; *Oreochromis niloticus*, XP_003443657.1; *Pangasianodon hypopthalmus*, XP_026793986.1; *Salmo salar*, XP_014062104.1; *Salmo trutta*, XP_029625791.1; *Silurus meridionalis*, KAF7691695.1; *Sinocyclocheilus anshuiensis*, XP_016296734.1; *Tachysurus fulvidraco*, XP_027030842.1; *Rattus norvegicus*, NP_036970.1. Characteristic and conserved residues of glycine (**G**) and proline (P) are indicated by the letter "G" and "P", respectively. The characteristic and preserved residues of the QXVXG motif are marked with a box. **(D)**. Phylogenetic tree of cystatin-B and sequences of their counterparts constructed using the neighbor-joining method and 1000 bootstrap replicates. The numbers presented in the branches represent the bootstrap value (%).

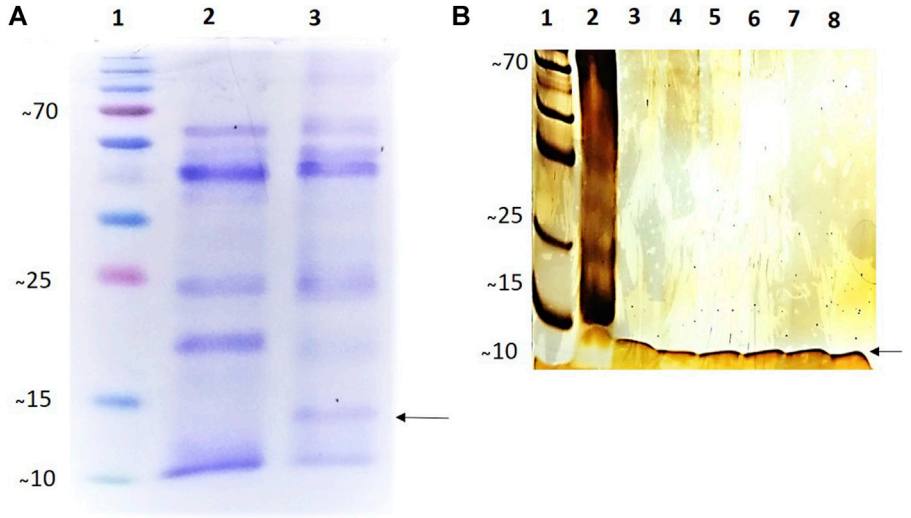


FIGURE 3 | Analysis of the expression and purification of rCystatin-B with SDS-PAGE (120 h of induction). **(A)**. Gel stained with 0.25% Coomassie blue; (Lane 1). PMM - Molecular Mass Standard - kDa (Invitrogen). (Lane 2). Lyophilized fraction (crude extract) of *Pichia pastoris* cell culture without cloned plasmid (negative control) subjected to the induction and expression process; (Lane 3). Transformed and lyophilized fraction of the culture of *Pichia pastoris* cells subjected to the induction and expression process (inclusion of cloned plasmid: pPIC_Cystatin-B). **(B)**. Gel stained with Silver Nitrate (20%). (Lane 1). PMM - Molecular Mass Standard - kDa (Invitrogen). (Lane 2). Transformed and lyophilized fraction of the culture of *Pichia pastoris* cells subjected to the induction and expression process (inclusion of cloned plasmid: pPIC_Cystatin-B). (Lanes 3–8). Purified rCystatin-B protein (indicated by the arrow). The recombinant protein was purified using Ni Sepharose High-Performance Nickel based resin.

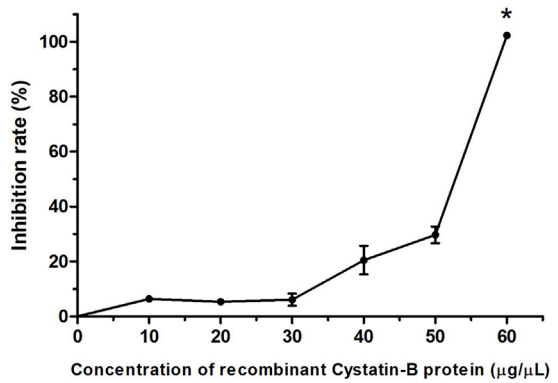


FIGURE 4 | Inhibition pattern of rCystatin-B against papain activity. The curve represents the change in residual papain activity along with the addition of rCystatin-B. αGTH (negative control), did not show any inhibitory effect (data not shown). Values are shown as mean ± standard deviation from the mean (n = 3). Asterisc (*) indicate statistical difference (p = 0.003).

positive control was 20.04 ± 2.55 mm and 24.53 ± 3.91 mm for *E. coli* and *B. subtilis*, respectively.

DISCUSSION

Developing new strategies for disease management and control is one of the current challenges in aquaculture. Biotechnology is considered an important tool for the treatment of emerging diseases associated with aquaculture production. In this study, the objective was to produce a protein with possible immunoregulatory potential for fish. For this, a 409 bp fragment related to the *cystatin* B gene was isolated from the spleen of the fish *Pygocentrus nattereri*. This is the first report of the isolation of this gene for this species.

Typically, cystatins in fish are composed of approximately 100 amino acids, varying somewhat according to the species and isoform. For example, in species such as *Scophthalmus maximus* (Xiao et al., 2010), the gene size of *cystatin* B was 300 bp length. Ahn et al. (2013) reported the stefin B gene with a size of 297 bp for *Paralichthys olivaceus*. Cystatin homologs or *stefin* A have been isolated from species such as *Ctenopharyngodon idella* with a nucleotide sequence length of 294 bp (Li et al., 2021). Another important feature of cystatins involves the three conserved domains that form the inhibitory sites of interaction with proteases: an N-terminal glycine, a glutamine-X-valine-X-glycine motif, and a C-terminal proline-tryptophan amino acid pair (Bode et al., 1988; Stubbs et al., 1990). In the present study, multiple alignments identified three conserved regions typical of the cystatin superfamily. In rCYST-B, the N-terminal domain containing the conserved glycine residue (G10G11) and the presence of similarity to its homologs from other species were found. The 2 G residues are known to constitute a wedge-shaped border, involved in the inhibition of protease activity, complementary to the active site of papain-like cysteine proteases (Bode et al., 1988; Abrahamson 1993). In addition, it is the domain involved in the interaction between the protease and the inhibitor, which allows it to interact directly with the S3, S2, and S1 substrate-binding regions of the protease (Hall et al., 1995; Bjork et al., 1995). The QXVXG motif, highly conserved in cystatins, was observed in rCYST-B from *P. Nattereri*. This motif

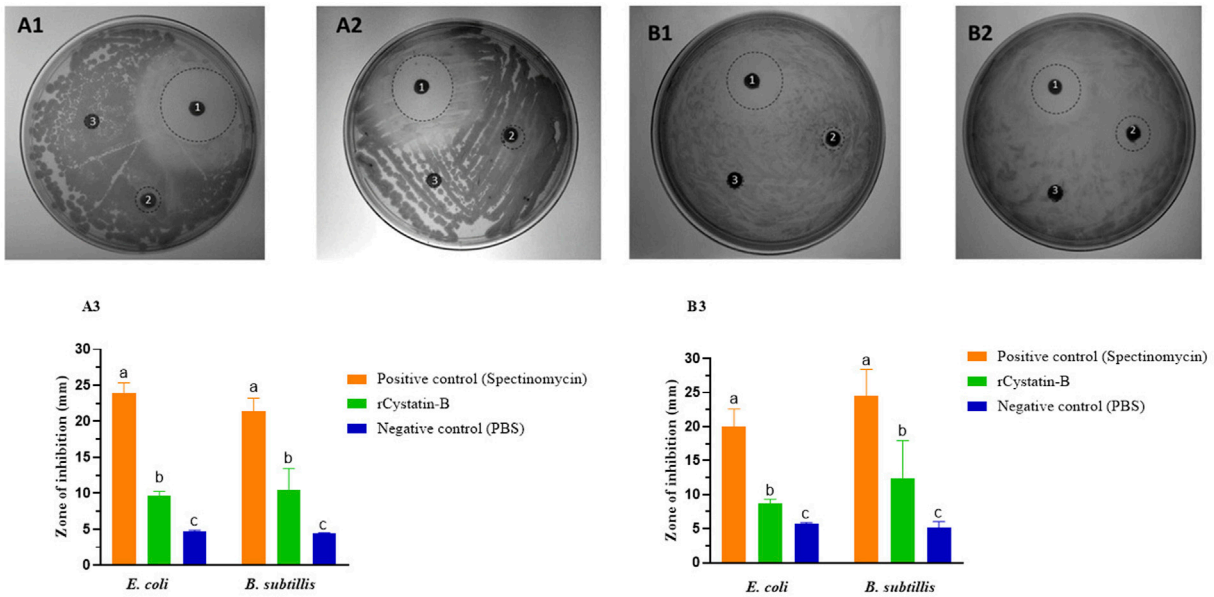


FIGURE 5 | Bacteriostatic action of rCystatin-B. (**A**). 60 μg/μL. (A1). *E. coli.* (A2). *B. subtilis*. (A3). Zone of inhibition (diameter mm). (**B**). 80 μg/μL. (B1). *E. coli*. (B2). *B. subtilis*. (B3). Inhibition diameter (mm). The dashed line circles correspond to the observed zone of inhibition. 1). Spectinomycin (positive control). 2) recombinant rCystatin-B protein. 3). PBS (negative control). Values are shown as mean ± standard deviation from the mean (n = 3). Different letters show statistical differences ($p < 0.05$).

binds to the active site of cysteine proteases, thereby interfering with their interaction with the substrate and inhibiting their activity (Bode et al., 1988; Turk and Bode 1991). The variation in the presence of amino acids proline - tyrosine in the C-terminal region of Cystatin B theoretically does not affect the inhibitory activity, both amino acids have hydrophilic characteristics and very similar isoelectric points. The inhibitory response was observed in the recombinant rCYST-B produced in the present study. Different types of C-terminal variations have been observed in human *Stefin* B, with replacements by the histidine (H) residue (Bjork et al., 1995) as well as in *Acipenser sinensis* (Bai et al., 2006) with the replacement by the Leucine (L) residue. Analysis of the phylogenetic tree performed in this study (**Figure 2D**) confirms the evolutionary relationship and proximity of Cystatin-B to its cystatin homologs as well as the functions among type 1 cystatin (*Stefins*) (Pemberton 2006).

The estimated molecular weight of rCYST-B was 11.8 kDa, similar results are observed for recombinant cystatin B protein obtained from *Ctenopharyngodon idella* with 11.48 KDa (Li et al., 2021), *Pseudosciaena crocea* with 11.4 KDa (Li et al., 2009), *Scophthalmus maximus* with 11.1 kDa (Xiao et al., 2010), *Paralichthys olivaceus* with 12.7 KDa (Ahn et al., 2013), reported values for *Stefins* family cystatins (Ochieng and Chaudhuri 2010).

After bioinformatics analyses, we evaluated the protease inhibition activity of the purified recombinant CYST-B protein. For this, we chose papain as cysteine protease and azocasein as substrate. The rCYST-B from *P. Nattereri* exhibited concentration-dependent inhibition of papain activity. These results confirmed that the purified recombinant protein was able to maintain its biological functions. The concentration at which 100% inhibition was observed was 60 μg/μL. However, a previous study reported a concentration of 0.5 μg/μL for *Stefins* B homologs from *Oplegnathus fasciatus* (Premachandra et al., 2012). Recently, Li et al. (2021), reported a dose-dependent inhibition pattern for the proteolytic activity of papain, cathepsin B, and cathepsin L, showing a decrease in the activity with the increasing amount of *Stefin* A isolated from *Ctenopharyngodon idella*. In the case of species such as *Crassostrea gigas*, the concentration of 50 μg/μL of cystatin A showed an inhibition rate of around 80% (Mao et al., 2018). In contrast, Premachandra et al. (2012), produced and purified a recombinant cystatin B from rock bream *Oplegnathus fasciatus* and demonstrated 82% inhibition of papain activity at a concentration of 0.5 μg/μL. Even though there is a wide variation in the concentration of cystatin showing 100% inhibition of protease activity, all results suggest that these recombinant proteins may show regulation of exogenous proteases from microorganisms and parasites that infect the host (Scott et al., 2007). These differences in the potential or response of the inhibitory activity of cystatin are supported by the wide spectrum of its structural variation and distribution in different types of cells and tissues (Lefebvre et al., 2008; Xiao et al., 2010); structure characterized by highly conserved regions that form a discrete wedge-shaped structure that blocks the active site of cysteine proteases (Gregory and Maizels, 2008). On the other hand, G11 introduces some flexibility in the N-terminal region of inhibitors, allowing the N-terminal motif to adopt an optimal conformation for enzyme-inhibitor interaction (Hall et al., 1993). The greater inhibitory potential exerted on cysteine proteases has allowed the use of the specificity of

Cysteine protease inhibitory activity substrates, as well as different evaluation protocols under *in vitro* conditions (temperature and enzymatic reaction times) (Bai et al., 2006; Xiao et al., 2010; Wickramasinghe et al., 2020).

In our study, it was evidenced that the protein heterologously produced by *P. pastoris* was able to inhibit the growth of *E. coli* and *B. subtilis*. Inhibition of *E. coli* and *B. subtilis* bacterial growth were also observed with *Paralichthys olivaceus* Cystatin-C (Yu et al., 2019) as well as on *Scophthalmus maximus* Cystatin-B (Xiao et al., 2010). Previous results confirm the antiviral and antibacterial role of cystatins (Wesierska et al., 2005). In general, proteins with antimicrobial capacity use their own cationic charge to attack and enter in a faster and more effective way through the bacteria membrane, increasing membrane permeability and flow of intracellular contents and subsequently, cell lysis (Izadpanah and Gallo 2005). We believe that this mechanism is used by rCYST-B, however, further research is needed to prove this affirmation. It should be noted that Cystatin-B is involved in immune responses to bacterial and fungal infections and anti-apoptotic processes in the brain (Takahashi et al., 1994; Di Giaimo et al., 2002). They are also expressed at high levels in follicular dendritic cells, which are involved in the immune and inflammatory reaction, present in the germinal centers of secondary lymphoid organs (Rinne et al., 1986). Furthermore, they can prevent apoptosis of B cells containing high-affinity receptors for the antigen present on the surface of follicular dendritic cells (van Eijk and Groot 1999).

Based on the confirmation of the antibacterial potential of rCYST-B, we believe that this protein can be used as an immunoregulatory protein for fish, and also as a food supplement in the fish feed. This hypothesis is based on the fact that recombinant fish cystatins have already been produced and showed potential in preventing fish diseases (Bjorklund et al., 1997). Recombinant cystatins obtained from Keta salmon (Yamashita and Konagaya 1991), trout (Li et al., 2000), and carp (Tzeng et al., 2001) were generated by heterologous expression in *Escherichia coli*. Even though these recombinant proteins have already been developed for fish, the use of a species-specific protein or that obtained from a phylogenetically close species may decrease the side effects besides potentiating the expected immunoregulatory effect.

The results of this study constitute the first reports on the production of a recombinant rCYST-B protein for a species native to South America. In conclusion, the systematic study of cloning, expression, and characterization of the biological activity of the recombinant cystatin B protein from *Pygocentrus nattereri* suggests its role as a possible immunoregulatory protein in biological defense against invaders and cellular protection against proteolysis mediated by cysteine proteases.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by Ethics Committee on Animal Use (CEUA) of Universidade Nilton Lins with approval protocol 015/2017.

AUTHOR CONTRIBUTIONS

JM Collected the data–Performed the analysis–Discussed the results and wrote to the final manuscript DA Conceived and designed the analysis–Performed the analysis–Discussed the results and contributed to the final manuscript. Supervised the findings of this work.

FUNDING

This work was supported by Brazilian CNPq (National Council for Scientific and Technological Development) (Proc. Number 432071/2018) and FAPEAM (Foundation for Research Support of the State of Amazonas). Special thanks to Professor Luis Fernando Marins for helping with the molecular analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.812971/full#supplementary-material

REFERENCES

Abrahamson, M., Alvarez-Fernandez, M., and Nathanson, C.-M. (2003). Cystatins. *Biochem. Soc. Symp.* 70, 179–199. doi:10.1042/bss0700179

Abrahamson, M. (1993). Cystatins - Protein Inhibitors of Papain-like Cysteine Proteinases Cienc Cult. *Sao Paulo)* 45, 299.

Ahn, S. J., Bak, H. J., Park, J. H., Kim, S. A., Kim, N. Y., Lee, J. Y., et al. (2013). Olive Flounder (*Paralichthys olivaceus*) Cystatin B: Cloning, Tissue Distribution, Expression and Inhibitory Profile of Piscine Cystatin B. *Comp. Biochem. Physiology Part B Biochem. Mol. Biol.* 165 (3), 211–218. doi:10.1016/j.cbpb.2013.04.007

Bai, J., Ma, D., Lao, H., Jian, Q., Ye, X., Luo, J., et al. (2006). Molecular Cloning, Sequencing, Expression of Chinese Sturgeon Cystatin in Yeast *Pichia pastoris* and its Proteinase Inhibitory Activity. *J. Biotechnol.* 125 (2), 231–241. doi:10.1016/j.jbiotec.2006.02.023

Barrett, A. J. (1986). The Cystatins: a Diverse Superfamily of Cysteine Peptidase Inhibitors. *Biomed. Biochim. Acta* 45 (11-12), 1363–1374.

Bauer, A. W., Kirby, W. M. M., Sherris, J. C., and Turck, M. (1966). Antibiotic Susceptibility Testing by a Standardized Single Disk Method. *Am. J. Clin. Pathology* 45, 493–496. doi:10.1093/ajcp/45.4_ts.493

Behr, E. R., and Signor, C. A. (2008). Distribuição e alimentação de duas espécies simpátricas de piranhas Serrasalmus maculatus e Pygocentrus nattereri (Characidae, Serrasalminae) Do rio Ibicuí, Rio Grande Do Sul,

Brasil. *Iheringia, Sér. Zool.* 98 (4), 501–507. doi:10.1590/S0073-47212008000400014

Berti, P. J., and Storer, A. C. (1995). Alignment/phylogeny of the Papain Superfamily of Cysteine Proteases. *J. Mol. Biol.* 246, 273–283. doi:10.1006/jmbi.1994.0083

Björk, I., Brieditis, I., and Abrahamson, M. (1995). Probing the Functional Role of the N-Terminal Region of Cystatins by Equilibrium and Kinetic Studies of the Binding of Gly-11 Variants of Recombinant Human Cystatin C to Target Proteinases. *Biochem. J.* 306, 513–518. doi:10.1042/bj3060513

Björklund, H. V., Johansson, T. R., and Rinne, A. (1997). Rhabdovirus-induced Apoptosis in a Fish Cell Line Is Inhibited by a Human Endogenous Acid Cysteine Proteinase Inhibitor. *J. Virol.* 71, 5658–5662. doi:10.1128/jvi.71.7.5658-5662.1997

Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., et al. (1988). The 2.0 A X-Ray Crystal Structure of Chicken Egg White Cystatin and its Possible Mode of Interaction with Cysteine Proteinases. *EMBO J.* 7, 2593–2599. doi:10.1002/j.1460-2075.1988.tb03109.x

Cornwall, G. A., and Hsia, N. (2003). A New Subgroup of the Family 2 Cystatins. *Mol. Cell. Endocrinol.* 200, 1–8. doi:10.1016/S0303-7207(02)00408-2

Di Giaimo, R., Riccio, M., Santi, S., Galeotti, C., Ambrosetti, D. C., and Melli, M. (2002). New Insights into the Molecular Basis of Progressive Myoclonus Epilepsy: a Multiprotein Complex with Cystatin B. *Hum. Mol. Genet.* 11, 2941–2950. doi:10.1093/hmg/11.23.2941

Duponchelle, F., Lino, F., Hubert, N., Panfili, J., Renno, J.-F., Baras, E., et al. (2007). Environment-related Life-History Trait Variations of the Red-Bellied Piranha *Pygocentrus Nattereri* in Two River Basins of the Bolivian Amazon. *J. Fish. Biol.* 71, 1113–1134. doi:10.1111/j.1095-8649.2007.01583.x

FAO (2017). Iniciativa de Crecimiento Azul. Asociación con los países para lograr los Objetivos de Desarrollo Sostenible. Available at: http://www.fao.org/publications/card/en/c/8801244b-8c61-469e-ab33-7774ae232285/.

FAO. SOFIA (2020). *State of Fisheries and Aquaculture in the World 2018*.

Gregory, W., and Maizels, R. (2008). Cystatins from Filarial Parasites: Evolution, Adaptation and Function in the Host-Parasite Relationship☆. *Int. J. Biochem. Cell Biol.* 40, 1389–1398. doi:10.1016/j.biocel.2007.11.012

Hall, A., Dalbøge, H., Grubb, A., and Abrahamson, M. (1993). Importance of the Evolutionarily Conserved glycine Residue in the N-Terminal Region of Human Cystatin C (Gly-11) for Cysteine Endopeptidase Inhibition. *Biochem. J.* 291, 123–129. doi:10.1042/bj2910123

Hall, A., Håkansson, K., Mason, R. W., Grubb, A., and Abrahamson, M. (1995). Structural Basis for the Biological Specificity of Cystatin C. *J. Biol. Chem.* 270, 5115–5121. doi:10.1074/jbc.270.10.5115

Huang, X., Gu, Y., Zhou, H., Xu, L., Cao, H., and Gai, C. (2020). *Acinetobacter Venetianus*, a Potential Pathogen of Red Leg Disease in Freshwater-Cultured Whiteleg Shrimp *Penaeus Vannamei*. *Aquac. Rep.* 18, 100543. doi:10.1016/j.aqrep.2020.100543

Izadpanah, A., and Gallo, R. L. (2005). Antimicrobial Peptides. *J. Am. Acad. Dermatology* 52, 381–390. doi:10.1016/j.jaad.2004.08.026

Josiah Ochieng, J., and Gautam Chaudhuri, G. (2010). Cystatin Superfamily. *J. Health Care Poor Underserved* 21, 51–70. doi:10.1353/hpu.0.0257

Khaznadji, E., Collins, P., Dalton, J. P., Bigot, Y., and Moiré, N. (2005). A New Multi-Domain Member of the Cystatin Superfamily Expressed by Fasciola Hepatica. *Int. J. Parasitol.* 35, 1115–1125. doi:10.1016/j.ijpara.2005.05.001

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi:10.1093/molbev/msy096

Lefebvre, C., Vandenbulcke, F., Bocquet, B., Tasiemski, A., Desmons, A., Verstraete, M., et al. (2008). Cathepsin L and Cystatin B Gene Expression Discriminates Immune Cœlomic Cells in the Leech Theromyzon Tessulatum. *Dev. Comp. Immunol.* 32, 795–807. doi:10.1016/j.dci.2007.11.007

Li, F., An, H., Seymour, T. A., and Barnes, D. W. (2000). Rainbow Trout (*Oncorhynchus mykiss*) Cystatin C: Expression in *Escherichia coli* and Properties of the Recombinant Protease Inhibitor. *Comp. Biochem. Physiology Part B Biochem. Mol. Biol.* 125 (4), 493–502. doi:10.1016/S0305-0491(00)00156-5

Li, F., Gai, X., Wang, L., Song, L., Zhang, H., Qiu, L., et al. (2010). Identification and Characterization of a Cystatin Gene from Chinese Mitten Crab *Eriocheir Sinensis*. *Fish Shellfish Immunol.* 29, 521–529. doi:10.1016/j.fsi.2010.05.015

Li, R., Li, S., Chen, Z., Jin, Y., Li, S., Li, S., et al. (2021). Grass Carp (*Ctenopharyngodon Idella*) Stefin A: Systematic Research on its Cloning, Expression, Characterization and Tissue Distribution. *Food Chem.* 335, 127564. doi:10.1016/j.foodchem.2020.127564

Li, S., Yang, Z., Ao, J., and Chen, X. (2009). Molecular and Functional Characterization of a Novel Stefin Analogue in Large Yellow Croaker (*Pseudosciaena Crocea*). *Dev. Comp. Immunol.* 33 (12), 1268–1277. doi:10.1016/j.dci.2009.07.008

Lowry, O., Rosebrough, N., Farr, A. L., and Randall, R. (1951). Protein Measurement with the Folin Phenol Reagent. *J. Biol. Chem.* 193, 265–275. doi:10.1016/s0021-9258(19)52451-6

Magister, Š., and Kos, J. (2013). Cystatins in Immune System. *J. Cancer* 4, 45–56. doi:10.7150/jca.5044

Mao, F., Lin, Y., He, Z., Li, J., Xiang, Z., Zhang, Y., et al. (2018). Dual Roles of Cystatin A in the Immune Defense of the Pacific Oyster, *Crassostrea gigas*. *Fish Shellfish Immunol.* 75, 190–197. doi:10.1016/j.fsi.2018.01.041

Margis, R., Reis, E. M., and Villeret, V. (1998). Structural and Phylogenetic Relationships Among Plant and Animal Cystatins. *Archives Biochem. Biophysics* 359, 24–30. doi:10.1006/abbi.1998.0875

Mottram, J. C., Coombs, G. H., and Alexander, J. (2004). Cysteine Peptidases as Virulence Factors of Leishmania. *Curr. Opin. Microbiol.* 7, 375–381. doi:10.1016/j.mib.2004.06.010

OIE (2013). Acute Hepatopancreatic Necrosis Disease: A Etiology, Epidemiology, Diagnosis, Prevention and Control References. Retrieved on 11 Nov 2019. Available at: https://www.oie.int/doc/ged/D14023.PDF.

Pemberton, P. A. (2006). PROTEINASE INHIBITORS | Cystatins. 511–517. doi:10.1016/B0-12-370879-6/00329-X

Pérez-Sánchez, T., Mora-Sánchez, B., and Balcázar, J. L. (2018). Biological Approaches for Disease Control in Aquaculture: Advantages, Limitations and Challenges. *Trends Microbiol.* 26 (11), 896–903. doi:10.1016/j.tim.2018.05.002

Premachandra, H. K. A., Whang, I., Lee, Y.-d., Lee, S., De Zoysa, M., Oh, M.-J., et al. (2012). Cystatin B Homolog from Rock Bream *Oplegnathus Fasciatus*: Genomic Characterization, Transcriptional Profiling and Protease-Inhibitory Activity of Recombinant Protein. *Comp. Biochem. Physiology Part B Biochem. Mol. Biol.* 163 (1), 138–146. doi:10.1016/j.cbpb.2012.05.012

Prunk, M., Perišić Nanut, M., Sabotič, J., and Kos, J. (2016). Cystatins, Cysteine Peptidase Inhibitors, as Regulators of Immune Cell Cytotoxicity. *Period Biol.* 118, 9. doi:10.18054/pb.v118i4.4504

Rinne, A., Dorn, A., Järvinen, M., Alavaikko, M., Jokinen, K., and Hopsu-Havu, V. K. (1986). Immunoelectron Microscopical Location of the Acid Cysteine Proteinase Inhibitor in the Lymphatic Tissue of the Tonsils. *Acta Histochem.* 79 (2), 137–145. doi:10.1016/S0065-1281(86)80072-1

Rosny, H. S., Hossain, M. M. M., Hasan-Uj-Jaman, M., Roy, H. S., Aurpa, I. A., Khondoker, S., et al. (2016). Dietary Supplementation of Garlic (*Allium Sativum*) to Prevent Acanthocephala Infection in Aquaculture. *Int. J. Fish. Aquat. Stud.* 4 (3), 188–192. doi:10.22271/fish.2016.v4.i3c.746

Rudenskaya, G. N., and Pupov, D. V. (2008). Cysteine Proteinases of Microorganisms and Viruses. *Biochem. Mosc.* 73, 1–13. doi:10.1134/s000629790801001x

Rzychon, M., and Chmiel, D. (2004). Stec-Niemczyk, JModes of Inhibition of Cysteine Proteases. *Acta Biochim. Pol.* 51, 861–873. doi:10.18388/abp.2004_3519

Scott, D. K., Lord, R., Muller, H. K., Malley, R. C., and Woods, G. M. (2007). Proteomics Identifies Enhanced Expression of Stefin A in Neonatal Murine Skin Compared with Adults: Functional Implications. *Br. J. Dermatol* 156, 1156–1162. doi:10.1111/j.1365-2133.2007.07875.x

Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., et al. (1990). The Refined 2.4 A X-Ray Crystal Structure of Recombinant Human Stefin B in Complex with the Cysteine Proteinase Papain: A Novel Type of Proteinase Inhibitor Interaction. *EMBO J.* 9 (6), 1939–1947. doi:10.1002/j.1460-2075.1990.tb08321.x

Takahashi, M., Tezuka, T., and Katunuma, N. (1994). Inhibition of Growth and Cysteine Proteinase Activity ofStaphylococcus aureusV8 by Phosphorylated Cystatin α in Skin Cornified Envelope. *FEBS Lett.* 355, 275–278. doi:10.1016/0014-5793(94)01196-6

Tang, K. F. J., Aranguren, L. F., Piamsomboon, P., Han, J. E., Maskaykina, I. Y., and Schmidt, M. M. (2017). Detection of the Microsporidian Enterocytozoon Hepatopenaei (EHP) and Taura Syndrome Virus in *Penaeus Vannamei*

Cultured in Venezuela. *Aquaculture* 480, 17–21. doi:10.1016/j.aquaculture.2017.07.043

Turk, V., and Bode, W. (1991). The Cystatins: Protein Inhibitors of Cysteine Proteinases. *FEBS Lett.* 285, 213–219. doi:10.1016/0014-5793(91)80804-C

Tzeng, S.-S., Chen, G.-H., Chung, Y.-C., and Jiang, S.-T. (2001). Expression of Soluble Form Carp (*Cyprinus carpio*) Ovarian Cystatin in *Escherichia coli* and its Purification. *J. Agric. Food Chem.* 49 (9), 4224–4230. doi:10.1021/jf0105135

Valladão, G. M. R., Gallani, S. U., Jerônimo, G. T., and Seixas, A. T. d. (2019). Challenges in the Control of Acanthocephalosis in Aquaculture: Special Emphasis onNeoechinorhynchus Buttnerae. *Rev. Aquacult*, 1–13. doi:10.1111/raq.12386

van Eijk, M., and de Groot, C. (1999). Germinal Center B Cell Apoptosis Requires Both Caspase and Cathepsin Activity. *J. Immunol.* 163, 2478–2482.

Verdot, L., Lalmanach, G., Vercruysse, V., Hoebeke, J., Gauthier, F., and Vray, B. (1999). Chicken Cystatin Stimulates Nitric Oxide Release from Interferon-γ-Activated Mouse Peritoneal Macrophages via Cytokine Synthesis. *Eur. J. Biochem.* 266, 1111–1117. doi:10.1046/j.1432-1327.1999.00964.x

Wesierska, E., Saleh, Y., Trziszka, T., Kopec, W., Siewinski, M., and Korzekwa, K. (2005). Antimicrobial Activity of Chicken Egg White Cystatin. *World J. Microbiol. Biotechnol.* 21 (1), 59–64. doi:10.1007/s11274-004-1932-y

Wickramasinghe, P. D. S. U., Kwon, H., Elvitigala, D. A. S., Wan, Q., and Lee, J. (2020). Identification and Characterization of Cystatin B from Black Rockfish, *Sebastes Schlegelii*, Indicating its Potent Immunological Importance. *Fish Shellfish Immunol.* 104, 497–505. doi:10.1016/j.fsi.2020.05.068

Xiao, P.-p., Hu, Y.-h., and Sun, L. (2010). *Scophthalmus maximus* Cystatin B Enhances Head Kidney Macrophage-Mediated Bacterial Killing. *Dev. Comp. Immunol.* 34 (12), 1237–1241. doi:10.1016/j.dci.2010.07.008

Yamashita, M., and Konagaya, S. (1991). A Comparison of Cystatin Activity in the Various Tissues of Chum Salmon *Oncorhynchus keta* between Feeding and Spawning Migrations. *Comp. Biochem. Physiology - Part A Physiology* 100 (3), 749–751. doi:10.1016/0300-9629(91)90402-X

Yu, H., Xu, X., Zhang, Q., and Wang, X. (2019). Molecular Characterization, Expression and Functional Analysis of Cystatin C in Japanese Flounder (*Paralichthys olivaceus*). *Fish. Shellfish Immunol.* 86, 695–701. doi:10.1016/j.fsi.2018.12.015

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

MINI REVIEW
published: 20 June 2022
doi: 10.3389/fgene.2022.903683

Living in Temporary Ponds Loading Giant Genomes: The Neotropical Annual Killifish Genus *Austrolebias* as New Outstanding Evolutionary Model

Graciela García, Verónica Gutiérrez† and Néstor Ríos*

Sección Genética Evolutiva, Facultad de Ciencias, UdelaR, Montevideo, Uruguay

The term Annual killifish describes a short-lived and amazing group of vertebrates inhabiting temporary ponds exposed to an extremely variable environment during its short lifespan in South America and Africa, leading to the death of the entire adult population during the dry season. *Austrolebias* is a specious genus of the family Rivulidae, with ~58 currently recognized species, extensively distributed in the temperate Neotropical region. Herein, we reviewed different aspects of the evolutionary biology with emphasis on the genome dynamic linked to the burst speciation process in this genus. *Austrolebias* constitutes an excellent model to study the genomic evolutionary processes underlying speciation events, since all the species of this genus analyzed so far share an unusually large genome size, with an average DNA content of 5.95 ± 0.45 picograms per diploid cell (mean C-value of about 2.98 pg). The drastic nuclear DNA–increasing would be associated with a considerable proportion of transposable elements (TEs) found in the *Austrolebias* genomes. The genomic proportion of the moderately repetitive DNA in the *A. charrua* genome represents approximately twice (45%) the amount of the repetitive components of the highly related sympatric and syntopic rivulinae taxon *Cynopoecilus melanotaenia* (25%), as well as from other rivulids and actinopterygian fish. These events could explain the great genome instability, the high genetic diversity, chromosome variability, as well as the morphological diversity in species of *Austrolebias*. Thus, species of this genus represent new model systems linking different evolutionary processes: drastic genome increase, massive TEs genomic representation, high chromosome instability, occurrence of natural hybridization between sister species, and burst speciation events.

Keywords: neotropical, killifish, *Austrolebias*, giant genomes, evolutionary model

OPEN ACCESS

Edited by:
Tony Silveira,
Federal University of Rio Grande, Brazil

Reviewed by:
Daiana Kaster Garcez,
Federal University of Santa Maria, Brazil
Luis Esteban Krause Lanés,
Instituto Pró-Pampa, Brazil

****Correspondence:***
Graciela García
ggarcia@fcien.edu.uy

†Present Address:
Verónica Gutiérrez,
Departamento de Biodiversidad y Genética, Instituto de Investigaciones Biológicas Clemente Estable, Ministerio de Educación y Cultura, Montevideo, Uruguay

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

Received: *24 March 2022*
Accepted: *05 May 2022*
Published: *20 June 2022*

Citation:
García G, Gutiérrez V and Ríos N (2022) Living in Temporary Ponds Loading Giant Genomes: The Neotropical Annual Killifish Genus Austrolebias as New Outstanding Evolutionary Model. Front. Genet. 13:903683. doi: 10.3389/fgene.2022.903683

INTRODUCTION

It has been proposed that model organisms are widely used in research as accessible and convenient systems to study a particular area or question in biology (Russell et al., 2017). In recent years they have accelerated the proliferation of experimental approaches, such as high-throughput sequencing, CRISPR gene editing, transgenesis, and other technologies which have enabled new insights, particularly when a trait of interest is most readily observed in a non-traditional model organism (Juntti, 2019).

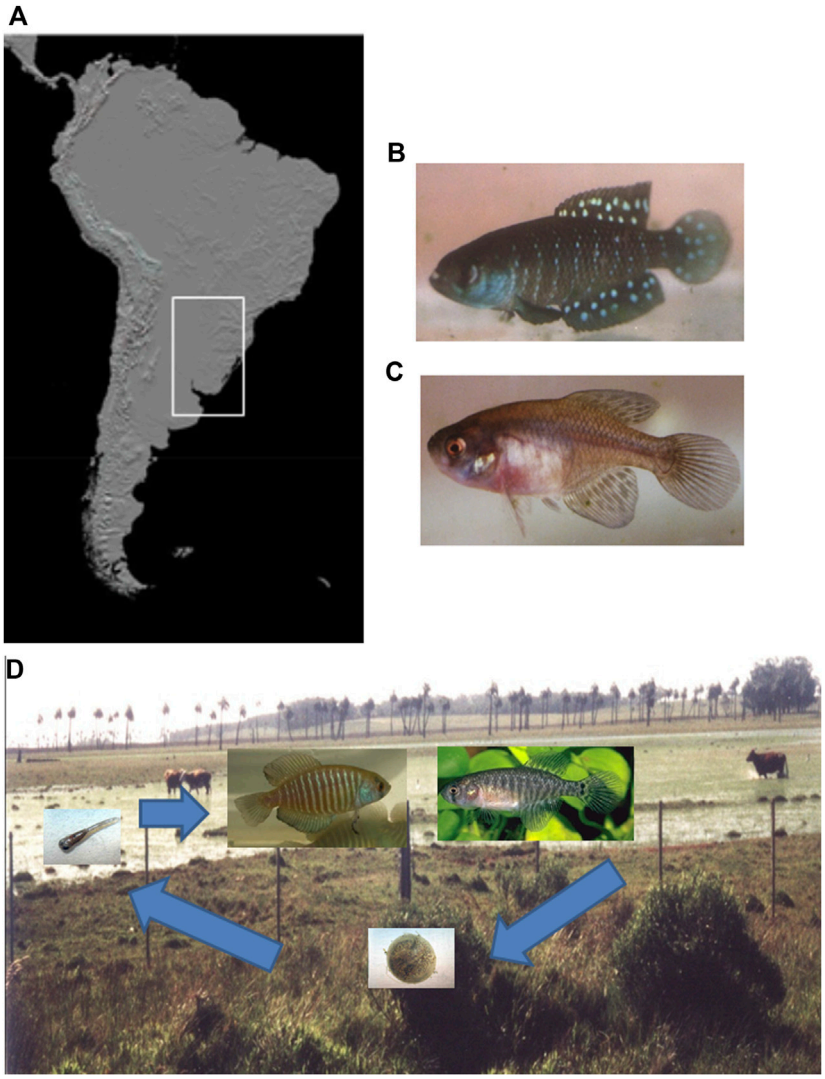


FIGURE 1 | Continental occurrence area, specimens, and life cycle of annual fish *Austrolebias*. **(A)** Distribution area of the *Austrolebias* genus in South America (white rectangle). **(B)** Adult male of the *A. affinis* species group. **(C)** Adult female of the *A. affinis* species group. **(D)** Life cycle of annual fish *Austrolebias* in temporary ponds from the R.16, Rocha Department, in Uruguay. *A. charrua* male (left) and female (right) are shown, the embryos remain in a prehatching stage of developmental arrest and hatch in the next rainy season when the ponds are flooded. The resulting juveniles reach sexual maturity in a few weeks.

In this sense, annual fish have been focused on as an important model organism as a consequence of an evolutionary adaption to its extreme habitat (Berois et al., 2014; Reichard and Polačik, 2019). This fish group inhabiting temporary ponds in South America and Africa is a unique, short-lived, vertebrate model presenting developmental, ecological, physiological, genetics, and evolutionary peculiar adaptations (García et al., 2014; Loureiro and de Sá, 2016; Berois et al., 2016, Reichard and Polačik, 2019). In particular, *Austrolebias* is a specious genus of the family Rivulidae, with ~58 currently recognized species (Froese and Pauly, 2021), extensively distributed in the temperate Neotropical region in two different South American basins (**Figure 1A**): La Plata-Paraná-Paraguay-Uruguay basin and Patos-Merin system (Costa, 2006; Loureiro et al., 2011; Volcan et al., 2014). Seasonal ponds inhabited by *Austrolebias* range from rounded to irregularly shaped shallow depressions in grasslands and seasonal flooded areas adjacent to rivers or large permanent wetlands (Loureiro et al., 2016). Seasonal patches strongly push individuals to accommodate spatially to optimize food supply, metabolic efficiency, mating opportunities, and predator avoidance, all in a few months (Wilbur, 1997). Therefore, they are freshwater teleosts exposed to an extremely variable environment during their short lifespan (Berois et al., 2016). During the rainy season the adults reproduce and generate desiccation-resistant embryos that remain buried in the muddy bottom of the dry ponds (**Figures 1B–D**). The embryos hatch in the next rainy season since the ponds are flooded, whereas the entire adult population dies during the dry season (Berois et al., 2016). The resulting juveniles reach sexual maturity in a few weeks, until a new reproductive cycle begins (Wourms, 1967;

Arezo et al., 2005). Therefore, the survival of the species becomes entirely dependent upon buried embryos and this unique life cycle is correlated to the seasonal environments (Berois et al., 2016).

One of the strengths of using a model organism is the ability to dispose of all possible information, including different sources of data about the analyzed system, improving the interpretation of results. Currently, studies on the evolutionary biology (García et al., 2002, 2009, 2014, 2019, 2020), systematics and conservation (Loureiro et al., 2018), ethology (Passos et al., 2013), ecology (Arim et al., 2011), reproductive strategies (Arezo et al., 2005, 2014), regulation of developmental pathways (Berois et al., 2016), aging (Gutiérrez, 2014), and neurogenesis (Fernández et al., 2020) on several species of *Austrolebias* are in progress. Herein, we present different findings about the evolutionary biology of annual fish with emphasis on the genome dynamic and the population genomics linked to the speciation process in the genus *Austrolebias*.

LIVING IN EXTREME ENVIRONMENTS DISPLAYING HIGH VARIABILITY AT DIFFERENT LEVELS

Ponds in the temperate region of South America are heterogeneous environments that cover a large range of areas. They have variable depths, contain organic matter, are rich in invertebrates, plant diversity, and biomass, and are usually isolated from other ponds (Laufer et al., 2009; Arim et al., 2011). The fragmentary nature of this kind of habitat can induce the occurrence of evolutionary mechanisms such as genetic drift (Gillespie, 2000), bottlenecks, founder effects, and endogamy (Carvalho and Hauser, 1999) in annual fish. Long-term isolation among population also produces local differentiation driving to high levels of genetic and morphological diversity and endemisms in this South American region (García et al., 2002; Loureiro, 2004; Loureiro et al., 2018). However, previous genetic analyses have proposed that the *Austrolebias* population behave as a metapopulation. During the rainy season, floods mix the populations across large distances yielding a high intrapopulation genetic and morphological variation (García, 2006; García et al., 2009). The wide range of gene flow values revealed by analyses carried out with mitochondrial and nuclear markers indicated that it is not homogeneous among ponds (García et al., 2019, 2020). In fact, the estimates of gene flow reinforced the hypothesis that the populations under study live in a region with a complex geological history and in an irregular rainy regime yielding to chaotic local dynamics, which facilitate the persistence of such a metapopulation (García, 2006). On the other hand, this area has been affected by sea level fluctuations at least since the Miocene, between 15 and 30 million of years ago (Sprechmann, 1978), while the last transgression being five-thousand years ago (García-Rodriguez and Witkowaki, 2003). In this complex geomorphological scenario in the wetlands from South America the proposed explosive speciation events for the genus *Austrolebias* would be favored (García et al., 2002, 2014). The evolutionary history together with their peculiar development strategy and short generation time (1 year) could also favor high morphologic and molecular variability among and within ponds (García, 2006). Species size ranges from small (max. 35 mm of standard length) to large (more than 200 mm of standard length) (Huber, 1996; Loureiro and de Sá, 2016). Most large species are considered to form a monophyletic group of top predators (García et al., 2002; Costa, 2006, 2010). Moreover, striking sexual dimorphism has been reported for several species of rivulids fish and most Neotropical annual killifish. In *Austrolebias*, males are usually larger than females (**Figures 1B–C**). A study on intra- and intersexual selection revealed that a larger body size favors *A. charrua* males (Passos et al., 2013).

KARYOTYPIC DIVERGENCE AMONG SPECIES OF THE GENUS *AUSTROLEBIAS*

In general, chromosome numbers vary little within and among teleost groups (Mank and Avise, 2006) and do not differ greatly from a widely proposed ancestral karyotype of 48 acrocentric chromosomes (Mank and Avise, 2006). This is also the case for some Neotropical rivulid species that include annual species (Elder et al., 1993; García et al., 1993, 1995, 2001). However, among annual killifish, the chromosome numbers range from 2N = 48 to 2N = 28 with arm numbers (NF) as high as 80 (García et al., 1993, 1995, 2001; García, 2016). In particular, chromosome variation at intra- and interspecific level was earlier described in species of the genus *Austrolebias* (García et al., 1993, 2001). Based on phylogenetic analysis using mitochondrial genes two major clades within *Austrolebias* could represent different repatterning pathways of the karyotypes (García et al., 2001; García, 2016). This analysis suggested that Clade I, corresponding to the *Austrolebias alexandri-affinis* species group, has differentiated basally during the late Miocene (García et al., 2014). This clade is characterized by a 2N = 48, different numbers of NF in each species, and the presence of extralarge acrocentric and subtelocentric chromosomes. Therefore, the occurrence of predominantly pericentric inversions, including perhaps heterochromatic loss/addition was proposed. On the other hand, Clade II is composed of all the remaining species groups clustered in four subclades showing the reduction of the diploid number from 2N = 48 to 2N = 34 and the maintenance of NF values nearly 48. Therefore, other rearrangement types could explain the chromosome evolution in these species groups (García et al., 1993, 2001).

GIANT GENOMES AND SPECIATION IN THE NEOTROPICAL GENUS *AUSTROLEBIAS*

Mank and Avise (2006) proposed that the genomes of ray-finned fishes (Actinopterygii) are well known for their evolutionary dynamism as reflected by drastic alterations in DNA content in general via partial or whole-genome duplication. Also, Mable et al. (2011) reported C-values for teleosts ranging between 0.35

and 4.9 pg, with an average of 1.2 pg. An amazing finding revealed that at least 16 species of *Austrolebias* show a C-value average of about 5.95 ± 0.45 pg/diploid cell (mean C-value of about 2.98 pg). In the same study, the genome size reported in the putative sister sympatric and syntopic taxon *C. melanotaenia* was of 2.72 ± 0.06 pg/diploid cell (García et al., 2014). Both C-values were corroborated in more recent findings through NGS of the total genome and *de novo* assembly and RNA-seq. This last analysis revealed haploid genomes size of 3.4 and 1.0 Gb for *A. charrua* and *C. melanotaenia* respectively (Valdivieso et al., 2017). The *C. melanotaenia* value fell within the range of most other cyprinodontoid and rivulid fish genomes, as in the early study in *Anablepsoides urophthalmus* (Hinegardner and Rosen, 1972), and more recently as in *Austrofundulus limnaeus* (Wagner et al., 2018), *Orestias ascotanensis* (Di Genova et al., 2022), and *Nematolebias whitei* (Thompson et al., 2022). Also, it is consistent with that of 3.11 pg/diploid cell reported in the African annual killifish *Nothobranchius furzeri* (Reichwald et al., 2009). Therefore, *Austrolebias* exhibit larger genomes compared to nearly all other reported diploid, i.e., non-(paleo) polyploid species of actinopterygian fishes (García et al., 2014). Previous analyses based on different mitochondrial genes supported the possible occurrence of burst cladogenetic processes in this genus. The sudden speciation in *Austrolebias* species, was preceded by events of divergence since a hypothetic ancestral rivulid genome which contained approximately 3.0 pg/diploid nucleus and a basal karyotype constituted by 48 small chromosomes of acrocentric type (García et al., 2001, 2002, 2014). These hypothesized events could be occurred since the Quaternary sharing drastic nuclear DNA increasing in all species analyzed, great genome instability and high levels of chromosomal divergence as previously mentioned (García et al., 2002; García, 2006; 2016). Moreover, they could explain the high morphological diversity described in species of the genus (Loureiro and de Sá, 1998; Loureiro et al., 2018).

MASSIVE ENRICHMENT OF TRANSPOSABLE ELEMENTS AND GENOME INSTABILITY AS AN EVOLUTIONARY DRIVING FORCE IN *AUSTROLEBIAS*

The extensive genetic variation by means of chromosome rearrangements involving both Robertsonian and non-Robertsonian changes at intra- and interspecific levels provides evidence for the genome instability occurring in *Austrolebias* (García et al., 1993, 2001). Remarkably, the frequency of these different types of chromosomal rearrangements between the two major clades occurred without statistically significant differences in nuclear DNA content. This finding reinforces the hypothesis that all *Austrolebias* species share similar nuclear DNA content since a hypothetic common ancestor (García, 2016).

Moreover, the drastic nuclear DNA–increasing events found in *Austrolebias* would be associated with considerable increase in the proportion of TEs. In fact, a comparative analysis of partial repetitive DNA content by means of NGS (New Generation Sequencing) revealed that the proportion of moderately repetitive DNA in *A. charrua* (45%) is approximately twice than that of the genus *C. melanotaenia* (25%) (García et al., 2015; Gutiérrez et al., 2016). Both species inhabit the same ponds and are distributed in the same South America temperate region. Most recently, total NGS genome analyses revealed a high number of repetitive elements present in *A. charrua* (68.1% of the total genome size), the most repetitive genome reported for a teleost (Valdivieso et al., 2017). Similar to the information reported by Chalopin et al. (2015) and Rhee et al. (2017), almost all classes of repetitive DNA are present in many teleosts groups and in the Rivulidae genomes. In the aforementioned comparative partial genome analysis implemented by García et al. (2015) in *A. charrua* and *C. melanotaenia,* retroelements make up most of the repetitive DNA. In particular, the *A. charrua* genome was predominantly enriched by LINE retroelements of the REX-Babar, Jockey, and L2 type (García et al., 2015; Valdivieso et al., 2017). Moreover, Rhee et al. (2017) reported that approximately one-fourth of the highly related mangrove killifish *Kryptolebias marmoratus* genome is composed of TEs, corroborating the previous finding in *C. melanotaenia*. Consistently, the most recent genome analysis of *O. ascotanensis* revealed that 21% (~142 Mb) correspond to repetitive sequences, including LINE, LTR, and SINES sequences (Di Genova et al., 2022).

These results are in contrast to those reported in other related rivulid species as in the non-annual genome of mangrove killifish in which DNA transposons (approximately 10–14%) are relatively common (Rhee et al., 2017). TEs insertions can be responsible for the disruption of genes or regulatory sequences, and can also cause chromosomal rearrangements, representing a threat to their host genome integrity (Hedges and Deininger, 2007). The importance of TEs in the structure and evolution of vertebrate genomes and their major impact on genome diversity between and within lineages, were revisited by comparing the mobilomes of 23 vertebrate genomes (Chalopin et al., 2015). It was hypothesized that TE activation could promote or intensify morphological and karyotypical changes, some of which may be potentially important for the process of microevolution, and allow species with plastic genomes to survive as new forms or even as species in times of rapid climatic change (Belyayev, 2014). As proposed by this author, the scenario of events preceding speciation by different models in small marginal populations could be as follows: under the influence of unusual ecological conditions, TEs become active; the mobilization of TEs produces genetic variations, epigenetic alterations, and high rates of karyotypic change (including changes in species-specific chromosomal pattern). All these evolutionary hypothetic scenarios would be taken place in the genus *Austrolebias*.

Whereas the annotation and characterization of the complex genome sequence of the *A. charrua* and that of *C. melanotaenia* are in progress (Valdivieso et al., 2017), the comparative analysis of these genomes and mobilomes could help to interpret their possible association with the extensive phenotypic plasticity detected at all levels in species of the genus *Austrolebias*,

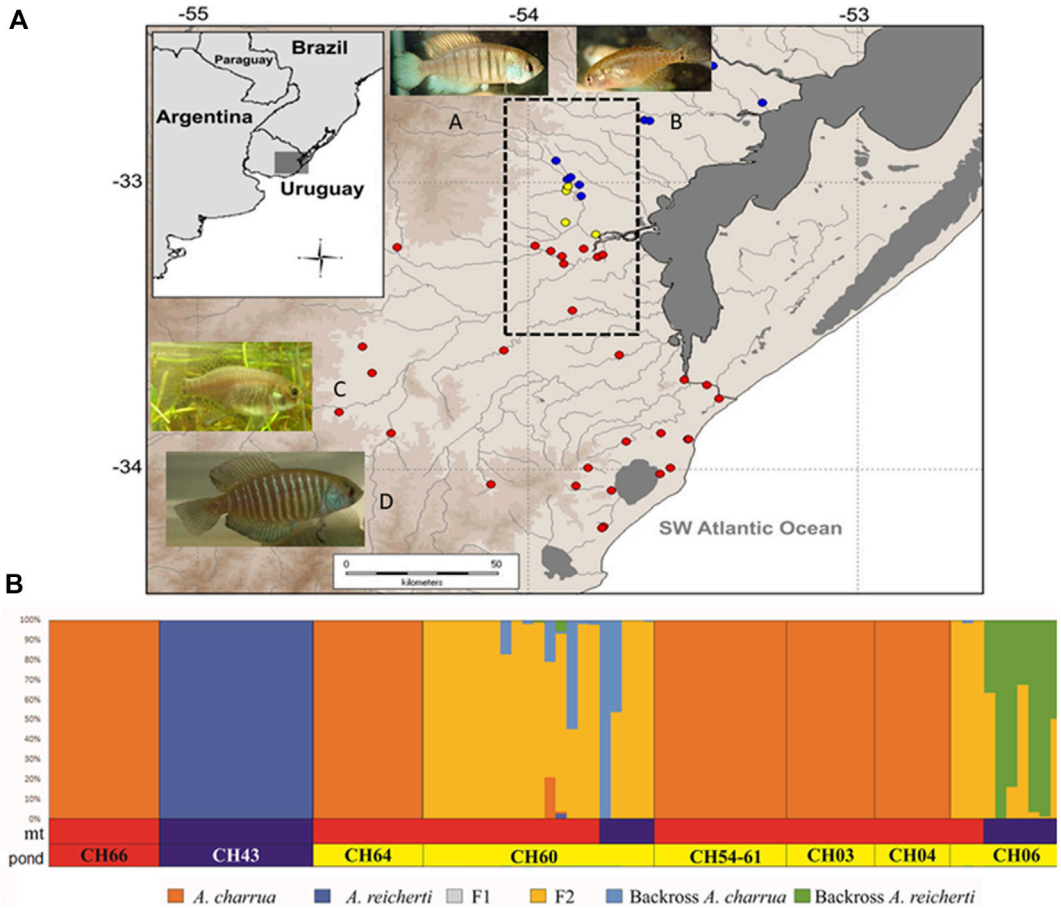


FIGURE 2 | Geographic distribution and genotype assignment of individuals sampled in the hybrid zone of the Patos-Merin coastal lagoons basin system, South America. **(A)** The contact area between both taxa is delimited for the dotted line rectangle. Ponds are represented by circles: *A. charrua* (red circles), *A. reicherti* (blue circles) and hybrid populations (yellow circles) In the top, male **(A)** and female **(B)** of the parental species *A. reicherti*; in the bottom, female **(C)** and male **(D)** of the parental species *A. charrua*. **(B)** Posterior probabilities of the ancestral genotype class estimated with NewHybrids, under Uniform prior assumption. Each individual is represented as a vertical bar divided into six segments. Each color indicates the posterior probability of an individual assignment to pure *A. reicherti* (blue), pure *A. charrua* (red), F1 (grey), F2 (yelow), and first generation backcross of a F1 hybrid with a pure *A. reicherti* (BC1R, green) or with a pure *A. charrua* (BC1C, sky). Each of the sampled ponds in the hybrid zone are labeled below the bar plots, and named as follows: CH66, CH43, CH64, CH60, CH54-61, CH03, CH04, and CH06. mt (mitochondrial) bars: represent the Cytb haplotype assignation of each individual to *A. charrua* (red) or *A. reicherti* (blue) species. Modified from García et al. (2019).

adapted to extreme environmental conditions in temporary ponds.

ADDITIONAL GENOMIC REVOLUTION AND PHENOTYPIC INNOVATIONS IN NATURAL HYBRID SWARMS OF A HYBRID ZONE BETWEEN TWO *AUSTROLEBIAS* SPECIES

Most recent population genetics analyses have detected other sources of morphological and genetic variability present in a contact area between two parapatric species *A. charrua* and *A. reicherti* (**Figure 2A**). This hybrid zone is the first described among Neotropical killifish and it was located in the Cebolattí River basin in Patos-Merin coastal lagoon system (García et al., 2019, 2020). The RNA-seq-based sequencing of the transcriptomes from pools of individuals of the two parental species and their putative natural hybrids allowed to identified a set of 111,725 SNP (single nucleotide polymorphism) markers, representing presumably fixed allelic differences among the two species (García et al., 2019). From these detected markers the first panel of 106 SNPs, in a single diagnostic multiplex assay to validate their capacity to reconstruct the patterns of the hybrid zone between both taxa, was performed. High-quality transcriptomes and a large set of gene-linked SNPs greatly facilitate functional and population genomics studies in the hybrid zone of these endangered species. Over the previously characterized bimodal hybrid zone (García et al., 2020) the new approach including more extensive sampling individuals and molecular markers, combined with morphological and biogeographic analyses, detected a population structure in which some groups among the hybrid swarms showed different level of introgression towards one or the other

parental species according to their geographic distribution. In fact, after 14 years of fieldwork and laboratory analysis, the present hybrid zone remains localized and spatially reduced to two patches: 1) individuals showed intergradation of the morphological and pigmentation patterns towards *A. charrua*; 2) individuals morphologically *A. reicherti*-like. New Hybrid analysis (**Figure 2B**) suggested that the combination of hybrid genotypes and introgression generated new genomic entities, which are different from both parental taxa. Abundant empirical evidence shows that hybridization frequently leads to transgressive phenotypes in plants, animals, and fungi (Rieseberg et al., 1999; Langdon et al., 2019). Natural introgression between sympatric or parapatric sister species could be considered an *in situ* conservation strategy (Becker et al., 2013) and Hamilton and Miller (2015) have suggested that genetic variation that persists within natural hybrids may have conservation value. The present results reinforce the importance of the natural hybridization and introgression in the analyzed contact zone between *A. charrua* and *A. reicherti* to preserve *in situ* biodiversity by increasing the observed phenotypic variation in each temporal pond where these endangered killifish populations inhabit. Finally, hybridization between species could cause genomic stress, which can lead to several genome reorganizations that seem to be driven by TEs (Fontdevila, 2005; Michalak, 2009; Romero-Soriano et al., 2017). In this sense, earlier, Barbara McClintock (1984) proposed the genomic shock hypothesis in which the hybridization between two species constitutes a source of stress that could disrupt the control mechanisms of TEs and cause their activation.

GENERAL CONSIDERATIONS AND OPEN QUESTIONS

Among different genera of annual and non-annual killifish from the New and Old World, giant genomes (~6 Gb) were only detected in *Austrolebias* (García et al., 2015; Valdivieso et al., 2017). Genome amplification has not occurred by polyploidization, since *Austrolebias* species are true diploids (García et al., 2014). Therefore, how can this formidable increase in genome size in the Neotropical genus *Autrolebias* be explained? Could it be related to the fact that *Austrolebias* is the unique rivulid genus that inhabit temperate Neotropical region? Could the extremely variable pond environments under different stress conditions have triggered unexpected TEs mobilization events? Are natural hybridization events between para/sympatric species linked to TEs mobilization in the genomes?

Previous findings in other taxa could contribute to our understanding of the possible underlying mechanisms to explain the existence of giant genomes and their dynamics in the genus *Austrolebias*. For example, the adaptation of *Drosophila* species to temperate climates was associated to widespread TEs (González et al., 2010). TEs invasion affected putatively genes which were also highly diverse in terms of their molecular and cellular function. Increasing the genome size as well as the timing of the development could have an impact also on its complexity (Gregory, 2005). Finally, an interesting and additional avenue to explore is the parallelism in the genomic size and the proportion of TEs (particularly, the abundance of LINE-type elements) between the human genome associated with different disorders and diseases and the *Austrolebias* one. For all the aforementioned issues, *Austrolebias* genus represents an outstanding species model in eco-evo-devo-research linking different evolutionary processes: drastic genome increase, massive presence of TEs, high chromosome instability, occurrence of natural hybridization between sister species, and the burst speciation process.

AUTHOR CONTRIBUTIONS

GG wrote the manuscript. All authors contributed to different works included in the article and approved the submitted version.

FUNDING

The research of GG, VG, and NR was supported by SNI (ANII, Agencia Nacional de Investigación e Innovación). Furthermore, NR was supported by DT (UdelaR, Universidad de la República Oriental del Uruguay), and GG was supported by DT_CSIC_ Project (2002–2019) from UdelaR, Uruguay.

ACKNOWLEDGMENTS

We are grateful to grants for molecular research to GG, furnished by Killi-Data Organization and Museum National d'Histoire Naturelle of France (2006–2007). Further, we are grateful to PEDECIBA (Programa de Desarrollo de las Ciencias Básicas) and CONICYT (Consejo Nacional de Innovación, Ciencia y Tecnología) Uruguay, for additional financial support. The authors are grateful to reviewers and to the editor for comments and suggestions improving the manuscript.

REFERENCES

Arezo, M. J., Papa, N., Gutiérrez, V., García, G., and Berois, N. (2014). Sex Determination in Annual fishes: Searching for the Master Sex-Determining Gene in Austrolebias *Charrua* (Cyprinodontiformes, Rivulidae). *Genet. Mol. Bio* 37, 364–374. doi:10.1590/S1415-47572014005000009

Arezo, M. J., Pereiro, L., and Berois, N. (2005). Early Development in the Annual Fish Cynolebias Viarius. *J. Fish. Bio.* 66, 1357–1370. doi:10.1111/j.0022-1112.2005.00688.x

Arim, M., Berazategui, M., Barreneche, J. M., Ziegler, L., Zarucki, M., and Abades, S. R. (2011). Determinants of Density-Body Size Scaling within Food Webs and Tools for Their Detection. *Adv. Ecol. Res.* 45, 1–43. doi:10.1016/B978-0-12-386475-8.00001-0

Becker, M., Gruenheit, N., Steel, M., Voelcke, C., Deusch, O., Heenan, P. B., et al. (2013). Hybridization May Facilitate *in situ* Survival of Endemic Species

through Periods of Climate Change. *Nat. Clim. Chang.*, 3. 1039–1043. Letters. doi:10.1038/nclimate2027

Belyayev, A. (2014). Bursts of Transposable Elements as an Evolutionary Driving Force. *J. Evol. Biol.* 27, 2573–2584. doi:10.1111/jeb.12513

Berois, N., Arezo, M. J., and de Sá, R. O. (2014). The Neotropical Genus Austrolebias: An Emerging Model of Annual Killifishes. *Annu. Rev. Cell Dev. Biol.* 3, 1–9. doi:10.1146/annurev.cellbio.3.1.1

Berois, N., Arezo, M. J., Papa, N. G., and Chalar, C. (2016). "Life Cycle, Reproduction, and Development in Annual Fishes: Cellular and Molecular Aspects," in *Annual Fishes: Life History Strategy, Diversity, and Evolution*. Editors N. Berois, G. García, and R. De Sá (Florida: CRC Press: Francis & Taylor Group), 33–45. ISBN 9780367575311.

Carvalho, G. R., and Hauser, L. (1999). Molecular Markers and the Species Concept: New Techniques Toresolve Old Disputes? *Rev. Fish. Biol. Fish.* 9, 379–382.

Chalopin, D., Naville, M., Plard, F., Galiana, D., and Volff, J. N. (2015). Comparative Analysis of Transposable Elements Highlights Mobilome Diversity and Evolution in Vertebrates. *Genome Biol. Evol.* 7, 567–580. doi:10.1093/gbe/evv005

Costa, W. J. E. M. (2010). Historical Biogeography of Cynolebiasine Annual Killifishes Inferred from Dispersal-Vicariance Analysis. *J. Biogeogr.* 37, 1995–2004. doi:10.1111/j.1365-2699.2010.02339.x

Costa, W. J. E. M. (2006). The South American Annual Killifish Genus Austrolebias(Teleostei: Cyprinodontiformes: Rivulidae): Phylogenetic Relationships, Descriptive Morphology and Taxonomic Revision. *Zootaxa* 1213, 1–162. doi:10.11646/zootaxa.1067.1.1

Di Genova, A., Nardocci, G., Maldonado-Agurto, R., Hodar, C., Valdivieso, C., Morales, P., et al. (2022). Genome Sequencing and Transcriptomic Analysis of the Andean Killifish *Orestias Ascotanensis* Reveals Adaptation to High-Altitude Aquatic Life. *Genomics* 114, 305–315. doi:10.1016/j.ygeno.2021.12.018

Elder, J. F., Bruce, J. T., Thomerson, J. E., and Taphorn, D. C. (1993). Karyotypes of Nine Venezuel Anannual Killifishes (Cyprinodontidae) with Comments on Karyotype Differentiation in Annual Killifishes. *Ichthyol. Explor. Freshwaters* 4, 261–268.

Fernández, A. S., Rosillo, J. C., Berrosteguieta, I., Herrera, M. L., and Torres-Pérez, M. (2020). The Ciliary Marginal Zone of the Retina of *Austrolebias Charrua* Fish is a Highly Proliferative Neurogenic Niche. *Microsc. Microanal.* 26, 171–172. doi:10.1017/S1431927620001075

Fontdevila, A. (2005). Hybrid Genome Evolution by Transposition. *Cytogenet. Genome Res.* 110, 49–55. doi:10.1159/000084937

Froese, R., and Pauly, D. (2021). FishBase. World Wide Web Electronic Publication. Available at: www.fishbase.org.version.(08/2021).

García, G., Gutiérrez, V., Ríos, N., Papa, N., Serra, S., and Loureiro, M. (2020). Characterization of a Hybrid Zone between Two Annual Killifish Genus *Austrolebias* from the Biosphere Reserve and Ramsar Sites in South America. *Hydrobiologia* 847, 421–442. doi:10.1007/s10750-019-04104-0

García, G., Alvarez-Valin, F., and Gómez, N. (2002). Mitochondrial Genes: Signals and Noise inPhylogenetic Reconstruction within Killifish Genus Cynolebias (Cyprinodontiformes, Rivulidae). *Biol. J. Linn. Soc.* 76, 49–59. doi:10.1111/j.1095-8312.2002.tb01713.x

García, G. (2016). "Chromosome Evolution, Divergent Meiotic Pathways, and Speciation in the Family Rivulidae," in *Annual Fishes: Life History Strategy, Diversity, and Evolution*. Editors N. Berois, G. García, and R. De Sá (Florida: CRC Press: Francis & Taylor Group), 275–287. ISBN 9780367575311.

García, G., Gutiérrez, V., Ríos, N., Turner, B., Santiñaque, F., López-Carro, B., et al. (2014). Burst Speciation Processes and Genomic Expansion in the Neotropical Annual Killifish Genus *Austrolebias* (Cyprinodontiformes, Rivulidae). *Genetica* 142, 87–98. doi:10.1007/s10709-014-9756-7

García, G., Lalanne, A. I., Aguirre, G., and Cappetta, M. (2001). Chromosome Evolution in Annual Killifish Genus *Cynolebias* and Mitochondrial Phylogenetic Analysis. *Chromosome Res.* 9, 93–100. doi:10.1023/a:1011664009509

García, G., Loureiro, M., Berois, N., Arezo, M. J., Casanova, G., and Olivera, A. (2009). Pattern of Differentiation in the Annual Killifish Genus *Austrolebias* (Cyprinodontiformes: Rivulidae) from a Biosphere Reserve Site in South America: a Multidisciplinary Approach. *Biol. J. Linn. Soc.* 98, 620–635. doi:10.1111/j.1095-8312.2009.01303.x

García, G. (2006). Multiple Simultaneous Speciation in Killifishes of the Cynolebias Adloffi Speciescomplex (Cyprinodontiformes, Rivulidae) from Phylogeography and Chromosome Data. *J. Zool. Syst. Evol. Res.* 44, 75–87. doi:10.1111/j.1439-0469.2005.00346.x

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

García, G., Ríos, N., Gutiérrez, V., Serra, S., and Loureiro, M. (2019). Transcriptome-Based SNP Discovery and Validation in the Hybrid Zone of the Neotropical Annual Fish Genus *Austrolebias*. *Genes (Basel).* 10, 789. doi:10.3390/genes10100789

García, G., Scvortzoff, E., and Hernández, A. (1995). Karyotypic Heterogeneity in South American Annual Killifishes of the Genus *Cynolebias* (Pisces, Cyprinodontiformes, Rivulidae). *Cytologia* 60, 103–110.

García, G., Scvortzoff, E., Maspoli, M. C., and Vaz-Ferreira, R. (1993). Analysis of Karyotypic Evolution in Natural Populations of *Cynolebias* (Pisces: Cyprinodontiformes, Rivulidae) Using Banding Techniques. *Cytologia* 58, 85–94.

García-Rodriguez, F., and Witkowaki, A. (2003). Inferring Sea Level Variation from Relative Percentages of Pseudopodosira Kosugii in Rocha Lagoon, SE Uruguay. *Diatom. Res.* 18, 49–59. doi:10.1080/0269249X.2003.9705572

Gillespie, J. H. (2000). Genetic Drift in an Infinite Population: the Pseudohitchhiking Model. *Genetics* 155, 909–919. doi:10.1093/genetics/155.2.909

González, J., Karasov, T. L., Messer, P. W., and Petrov, D. A. (2010). Genome-Wide Patterns of Adaptation to Temperate Environments Associated with Transposable Elements in *Drosophila*. *PLOS Genet.* 6, e1000905. doi:10.1371/journal.pgen.1000905

Gregory, T. R. (2005). Synergy between Sequence and Size in Largescale Genomics. *Nat. Rev. Genet.* 6, 699–708. doi:10.1038/nrg1674

Gutiérrez, V. (2014). *Patterns of Molecular Senescence in the Neotropical Annnual Killifish Genus Austrolebias*. UdelaR, Uruguay: Facultad de Ciencias. Umpublished Ph.D. Dissertation Thesis PEDECIBA.

Gutiérrez, V., Ríos, N., and García, G. (2016). "Genome of the Annual Fishes Austrolebias Dynamized byTransposable Elements," in *Annual Fishes: Life History Strategy, Diversity, and Evolution*. Editors N. Berois, G. García, and R. De Sá (Florida: CRC Press: Francis & Taylor Group), 245–251. ISBN 9780367575311.

Hamilton, J. A., and Miller, J. M. (2015). Adaptive Introgression as a Resource for Management and Geneticconservation in a Changing Climate. *Conserv. Biol.* 30, 33–41. doi:10.1111/cobi.12574

Hedges, D. J., and Deininger, P. L. (2007). Inviting Instability: Transposable Elements, Double-Strand Breaks, and the Maintenance of Genome Integrity. *Mutat. Research/Fundamental Mol. Mech. Mutagen.* 616, 46–59. doi:10.1016/j.mrfmmm.2006.11.021

Hinegardner, R., and Rosen, D. E. (1972). Cellular DNA Content and Evolution of Teleostean Fishes. *Am. Nat.* 106, 621–644.

Huber, J. H. (1996). *Updated Checklist of Taxonomic Names, Collecting Localities and Bibliographic References of Oviparous Cyprinodont Fishes, Bibliographic References of Oviparous Cyprinodont Fishes (Atherinomorpha, Pisces). Killi-Data 1996*. Paris, France: Société Française d'Ichtyologie, Muséum National d'Histoire Naturelle, 399.

Juntti, S. (2019). The Future of Gene-Guided Neuroscience Research in Non-Traditional ModelOrganisms. *Brain Behav. Evol.* 93, 108–121. doi:10.1159/000500072

Langdon, Q. K., Peris, D., Baker, E. P., Opulente, D. A., Nguyen, H. V., Bond, U., et al. (2019). Fermentation Innovation through Complex Hybridization of Wild and Domesticated Yeasts. *Nat. Ecol. Evol.* 3, 1576–1586. doi:10.1038/s41559-019-0998-8

Laufer, G., Arim, M., Loureiro, M., Piñeiro-Guerra, J. M., Clavijo-Baquet, S., and Fagúndez, C. (2009). Diet of Four Annual Killifishes: An Intra and Interspecific Comparison. *Neotrop. Ichthyol.* 7, 77–86. doi:10.1590/S1679-62252009000100010

Loureiro, M., Borthagaray, A., Hernández, D., Duarte, A., Pinelli, V., and Arim, M. (2016). "Austrolebias in Space Scaling from Ponds to Biogeographical Regions," in *Annual Fishes: Life History Strategy, Diversity, and Evolution*. Editors

N. Berois, G. García, and R. De Sá (Florida: CRC Press: Francis & Taylor Group), 111–132. ISBN 9780367575311.

Loureiro, M., de Sá, R. O., Berois, N., García, G., and De Sá, R. (2016). "Diversity of Aplocheiloidei," in *Annual Fishes: Life History Strategy, Diversity, and Evolution* (Florida: CRC Press: Francis & Taylor Group), 3–31. ISBN 9780367575311.

Loureiro, M., and de Sá, R. O. (1998). Osteological Analysis of the Killifish Genus Cynolebias (Cyprinodontiformes: Rivulidae). *J. Morphol.* 238, 109–262.

Loureiro, M., de Sá, R. O., Serra, S. W., Alonso, F., Krause Lanés, L. E., Volcan, M. V., et al. (2018). Review of the Family Rivulidae (Cyprinodontiformes, Aplocheiloidei) and a Molecular and Morphological Phylogeny of the Annual Fish Genus *Austrolebias* Costa 1998. *Neotrop. Ichthyol.* 16, e180007. doi:10.1590/1982-0224-20180007

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

Loureiro, M. (2004). *Sistemática y biogeografía de los peces anuales de la subtribu Cynolebiatina(Cyprinodontiformes: Rivulidae: Cynolebiatinae)*. Montevideo: Universidad de la República, 119. Unpublished Ph.D. Thesis, PEDECIBA.

Mabble, B. K., Alexandrou, M. A., and Taylor, M. I. (2011). Genome Duplication in Amphibians and Fish: An Extended Synthesis. *J. Zoology* 284, 151–182. doi:10.1111/j.1469-7998.2011.00829.x

Mank, J. E., and Avise, J. C. (2006). Cladogenetic Correlates of Genomic Expansions in the Recent Evolution of Actinopterygiian Fishes. *Proc. R. Soc. B* 273, 33–38. doi:10.1098/rspb.2005.3295

McClintock, B. (1984). The Significance of Responses of the Genome to Challenge. *Science* 226, 792–801. doi:10.1126/science.15739260

Michalak, P. (2009). Epigenetic, Transposon and Small RNA Determinants of Hybrid Dysfunctions. *Heredity* 102, 45–50. doi:10.1038/hdy.2008.48

Passos, C., Tassino, B., Loureiro, M., and Rosenthal, G. G. (2013). Intra- and Intersexual Selection on Male Body Size in the Annual Killifish *Austrolebias Charrua*. *Behav. Process* 96, 20–26. doi:10.1016/j.beproc.2013.01.008

Reichard, M., and Polacik, M. (2019). *Nothobranchius Furzeri*, an 'instant' Fish from an Ephemeral Habitat. *eLife* 8, e41548. doi:10.7554/eLife.41548

Reichwald, K., Lauber, C., Nanda, I., Kirschner, J., Hartmann, N., Schories, S., et al. (2009). High Tandem Repeat Content in the Genome of the Shortlived Annual FIsh Nothobranchius Furzeri: A New Vertebrate Model for Aging Research. *Genome Biol.* 10, R16. doi:10.1186/gb-2009-10-2-r16

Rhee, J.-S., Choi, B.-S., Kim, J., Kim, B.-M., Lee, Y.-M., Kim, I.-C., et al. (2017). Diversity, Distribution, and Significance of Transposable Elements in the Genome of the Only Selfing Hermaphroditic Vertebrate *Kryptolebias Marmoratus*. *Sci. Rep.* 7, 40121. doi:10.1038/srep40121

Rieseberg, L., Archer, M., and Wayne, R. (1999). Transgressive Segregation, Adaptation and Speciation. *Heredity* 83, 363–372 .doi:10.1038/sj.hdy.688617

Romero-Soriano, V., Modolo, L., Lopez-Maestre, H., Mugat, B., Pessia, E., Chambeyron, S., et al. (2017). Transposable Element Misregulation Is Linked to the Divergence between Parental piRNA Pathways in Drosophila Hybrids. *Genome Biol. Evol.* 9, 1450–1470. doi:10.1093/gbe/evx091

Russell, J. J., Theriot, J. A., Sood, P., Marshall, W. F., Landweber, L. F., Fritz-Laylin, L., et al. (2017). Non-model Model Organisms. *BMC Biol.* 15, 55. doi:10.1186/s12915-017-0391-5

Sprechman, P. (1978). The Paleoecology and Paleogeography and Stratigraphy of the Uruguayan Coastal área during the Neogene and Quaternary. *Zitteliana* 4, 3–72.

Thompson, A. W., Wojtas, H., Davoll, M., and Braasch, I. (2022). Genome of the Rio Pearlfish (Nematolebias Whitei), a Bi-annual Killifish Model for Eco-Evo-Devo in Extreme Environments. *G3, Genes,Genomes, Genet*. 12 (5), jkac045. doi:10.1093/g3journal/jkac045

Valdivieso, C., Pereiro, L., Di Genova, A., Gajardo, F., García, G., Arezo, M. J., et al. (2017). "The Genome Sequence of the Annual Killifish *Austrolebias Charrua* Reveals Mechanisms of Genomic Expansion via Retrolelements," in *IX Latin American Society for Developmental Biology Meeting* (Colombia: Medellín), 192.)

Volcan, M. V., Gonçalves, A. C., and Lanés, L. E. K. (2014). Austrolebias Quirogai (Actinopterygii: Cyprinodontiformes: Rivulidae) in Brazil: Occurence, Population Parameters, Habitat Characteristics, and Conservation Status. *Acta Ichtyol. Pisc.* 44, 37–44. doi:10.3750/AIP2014.44.1.05

Wagner, J. T., Singh, P. P., Romney, A. L., Riggs, C. L., Minx, P., Woll, S. C., et al. (2018). The Genome of *Austrofundulus Limnaeus* Offers Insights into Extreme Vertebrate Stress Tolerance and Embryonic Development. *BMC Genomics* 19, 155. doi:10.1186/s12864-018-4539-7

Wilbur, H. M. (1997). Experimental Ecology of Food Webs: Complex Systems in Temporary Ponds. *Ecology* 78, 2279–2302. doi:10.2307/2265892

Wourms, J. P. (1967). "Annual Fishes," in *Methods in Developmental Biology*. Editors F.H. Wilt and N. Wessels (New York: Thomas and Crowell Company), 123–137.

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 21 June 2022
doi: 10.3389/fgene.2022.884072

Revealing the Satellite DNA History in *Psalidodon* and *Astyanax* Characid Fish by Comparative Satellitomics

Caio Augusto Gomes Goes[1], *Rodrigo Zeni dos Santos*[1], *Weidy Rozendo Clemente Aguiar*[1], *Dálete Cássia Vieira Alves*[2], *Duílio Mazzoni Zerbinato de Andrade Silva*[3], *Fausto Foresti*[3], *Claudio Oliveira*[3], *Ricardo Utsunomia*[2] *and Fabio Porto-Foresti*[1]*

[1]*Laboratório de Genética de Peixes, Faculdade Estadual Paulista "Júlio de Mesquita Filho", Departamento de Ciências Biológicas, Faculdade de Ciências, Bauru, Brazil,* [2]*Instituto de Ciências Biológicas e da Saude, Universidade Federal Rural do Rio de Janeiro, Seropédica, Brazil,* [3]*Laboratório de Biologia e Genética de Peixes, Faculdade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências, Botucatu, Brazil*

OPEN ACCESS

Edited by:
Tony Silveira,
Federal University of Rio Grande, Brazil

Reviewed by:
Érica Ramos,
Serviço Social da Indústria (SESI), Brazil
Pablo Mora Ruiz,
University of Jaén, Spain

***Correspondence:**
Fabio Porto-Foresti
fp.foresti@unesp.br

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

Received: *25 February 2022*
Accepted: *31 May 2022*
Published: *21 June 2022*

Citation:
Goes CAG, dos Santos RZ, Aguiar WRC, Alves DCV, Silva DMZdA, Foresti F, Oliveira C, Utsunomia R and Porto-Foresti F (2022) Revealing the Satellite DNA History in Psalidodon and Astyanax Characid Fish by Comparative Satellitomics. Front. Genet. 13:884072. doi: 10.3389/fgene.2022.884072

Eukaryotic genomes are usually enriched in repetitive DNA sequences, which can be classified as dispersed or tandemly repeated elements. Satellite DNAs are noncoding monomeric sequences organized in a head-to-tail fashion that are generally located on the subtelomeric and/or pericentromeric heterochromatin. In general, a single species incorporates a diverse group of satellite DNA families, which collection is called satellitome. Here, we characterized three new satellitomes from distinct characid fish (*Psalidodon fasciatus, P. bockmanni,* and *Astyanax lacustris*) using a combination of genomic, cytogenetic, and bioinformatic protocols. We also compared our data with the available satellitome of *P. paranae.* We described 57 satellite DNA (satDNA) families of *P. fasciatus* (80 variants), 50 of *P. bockmanni* (77 variants), and 33 of *A. lacustris* (54 variants). Our analyses demonstrated that several sequences were shared among the analyzed species, while some were restricted to two or three species. In total, we isolated 104 distinctive satDNA families present in the four species, of which 10 were shared among all four. Chromosome mapping revealed that the clustered satDNA was mainly located in the subtelomeric and pericentromeric areas. Although all *Psalidodon* species demonstrated the same pattern of clusterization of satDNA, the number of clusters per genome was variable, indicating a high dynamism of these sequences. In addition, our results expand the knowledge of the As51 satellite DNA family, revealing that *P. bockmanni* and *P. paranae* exhibited an abundant variant of 39 bp, while *P. fasciatus* showed a variant of 43 bp. The majority of satDNAs in the satellitomes analyzed here presented a common library repetitive sequence in *Psalidodon* and *Astyanax*, with abundance variations in each species, as expected for closely related groups. In addition, we concluded that the most abundant satDNA in *Psalidodon* (As51) passed through a diversification process in this group, resulting in new variants exclusive of *Psalidodon*.

Keywords: satellitome, satDNA evolution, characiforms, fish, neotropical fishes

INTRODUCTION

A significant part of eukaryotic genomes is composed of repetitive DNA sequences (Charlesworth et al., 1994), represented by transposable elements and tandemly arrayed sequences, such as multigene families and satellite DNAs (satDNAs) (Charlesworth et al., 1994; Jurka et al., 2005). SatDNAs are noncoding sequences organized in tandem arrays of up to hundreds of thousands of nucleotides (López-Flores and Garrido-Ramos, 2012; Plohl et al., 2012) that are typically observed in subtelomeric and pericentromeric heterochromatin areas, although short arrays of satDNAs dispersed in euchromatin have been documented (Kuhn, 2015; Ruiz-Ruano et al., 2016; Garrido-Ramos, 2017; Pita et al., 2017; Rodrigues et al., 2019; Montiel et al., 2021). They are originated from duplication of a new sequence by replication slippage or rolling circle replication, with posterior dispersion throughout the genome and massive local amplification (Ruiz-Ruano et al., 2016; Vondrak et al., 2020).

The evolution of satDNAs is characterized by a process called concerted evolution (Elder and Turner, 1995; Lorite et al., 2017). In this sense, satDNAs go through a step of homogenization, following mechanisms such as unequal crossing over and gene conversion (Smith, 1976; Dover, 1986), with homogenized variants fixed in populations by reproduction. In addition, related species may share an ancestral set of different conserved satDNA families (Fry and Salser, 1977; Ruiz-Ruano et al., 2016), according to the library hypothesis. Different variants may be amplified or depleted in each species, generating different collections of detectable satDNAs. Therefore, a part of the library may appear as an abundant satDNA, while others remain at low amounts (Camacho et al., 2022). To name this whole collection of satDNA families observed in a single genome, Ruiz-Ruano et al. (2016) proposed the term satellitome.

The first satellitome of a Neotropical fish was described for *Psalidodon paranae*, with 45 satDNA families (Silva et al., 2017), and some were detected by Fluorescence *in situ* hybridization (FISH) in congeneric species, corroboration of the library hypothesis. This identification of common satDNAs in three species of *Psalidodon* (Silva et al., 2017) and the recent diversification and phylogenetic proximity between *P. paranae*, *P. fasciatus*, and *P. bockmanni*, incites interest in this group for the study of the evolution of satDNAs. *Psalidodon* was part of the *Astyanax*, until recently (Terán et al., 2020). Prior to the description of the *P. paranae* satellitome, the only known satDNA in this group was As51 (Mestriner et al., 2000), which is widely used as a cytogenetic marker, and present in several species of *Psalidodon*, usually in subtelomeric chromatin regions (Abel et al., 2006; Kantek et al., 2009). In the *P. paranae* satellitome, As51 represents the most abundant satDNA, corresponding to ApaSat01-51 (Silva et al., 2017).

The satellitome of *P. paranae* is unique to the genus *Psalidodon*, although some of its satDNAs are observed in other related species (Silva et al., 2017). Previous evolutionary studies involving the characterization of the entire satellitome of a species are scarce. The objective of this study was to conduct, for the first time, a comparison among the catalogs of satDNAs, involving the satellitome of *P. paranae*, and three new satellitomes: *P. fasciatus*, *P. bockmanni*, and *A. lacustris*, to expand the knowledge of *Psalidodon* satDNA families and to describe the *Astyanax* satellitome for the first time.

MATERIALS AND METHODS

Sampling and Cytogenetic Analyses

We analyzed individuals of *P. bockmanni*, *P. fasciatus*, and *A. lacustris* that were collected from natural populations of the Alambari (22°27′07.4"S; 4914′30.4″W) (*P. bockmanni* and *P. fasciatus*), Araras (22°27′49″S; 47°44′44″W) (*P. fasciatus*), and Batalha (22°23′40.8"S; 4906′34.7″W) (*A. lacustris*) rivers. Five individuals of each species were collected, with exception of *P. fasciatus*, in which we collected five individuals of each population analyzed (Alambari and Araras). The samples were collected, maintained, and analyzed following the protocols for the care and use of animals of the Brazilian Society for Laboratory Animal Science (SBCAL), and approved (protocol 1227) by Bioscience Institute/UNESP Ethics Committee on the Use of Animals (CEEAA/IBB/UNESP). The samples were collected with authorization from the relevant organizations (MMA/IBAMA/ICMBio/SISBIO—18884-1, registered with IBAMA No. 2567470). The individuals were deposited at the fish collection of Laboratório de Genética de Peixes, at UNESP, Bauru, São Paulo, Brazil, under the vouchers LGP12524—LGP12529 (*P. bockmanni*), LGP13006 and LGP14052—LGP14055 (*A. lacustris*), LGP12554—LGP12558 (*P. fasciatus*–Alambari), and LGP85536 (*P. fasciatus*–Araras).

The cytogenetic analysis to obtain mitotic chromosomes followed a protocol previously established (Foresti et al., 1981), using cells of anterior kidney tissue.

DNA Extraction and Whole Genome Sequencing

We extracted total genomic DNA from the livers of *P. bockmanni*, *P. fasciatus* (Alambari, 2n = 46), and *A. lacustris* using the Wizard Genomic Kit (Promega, Madison, United States), following the manufacturer's instructions. RNA was removed using RNase A (Invitrogen, Waltham, United States). Genomic DNA sequencing from an individual of each of *P. fasciatus*, *P. bockmanni*, and *A. lacustris* was performed using an Illumina MiSeq (paired-end 2 × 250 bp). In addition, the library of *P. paranae* was used in our analyses (Silva et al., 2017).

Satellitome Characterization and Additional Analyses

After quality and adapter trimming using Trimmomatic v0.33 (Bolger et al., 2014), we performed a high-throughput analysis of satDNAs within the genomes of *P. fasciatus*, *P. bockmanni*, and *A. lacustris*, using the satMiner bioinformatic protocol (Ruiz-Ruano et al., 2016). We performed a clustering of 2 × 500,000 reads, randomly selected using RepeatExplorer (Novák et al., 2013) with

the default options to select clusters with a structure typical of satDNA. A search for contigs with tandem repetitions was performed using the dotplot tool, integrated into Geneious 8.1 software (Biomatters). We used the Deconseq software (Schmieder and Edwards, 2011) to filter the assembled contigs of all clusters designated as *in tandem* by RepeatExplorer, and the remaining sequences were clustered using RepeatExplorer. These processes were repeated until no new satDNA sequences appeared.

Subseqently, we filtered the obtained sequences and removed other tandemly repeated elements, such as multigene families, and used the software rm_homology (https://github.com/fjruizruano/satminer/blob/master/rm_homology.py) to eliminate redundancies of isolated contigs, and to group the sequences into the same variant, different variants of the same family, or superfamilies (similarities greater than 95, 80, and 50%, respectively), as established by Ruiz-Ruano et al., 2016. The same method was used to compare the satellitomes of *P. paranae*, *P. fasciatus*, *P. bockmanni*, and *A. lacustris*. The abundance and divergence of each satDNA variant were estimated using RepeatMasker software (Smit et al., 2017), using randomly selected reads (2 × 5,000,000 reads). In RepeatMasker, the reads of species were separately mapped against concatenated monomers of satDNAs consensus sequences (spanning 200 nucleotides). SatDNA families were named according to a previous study (Ruiz-Ruano et al., 2016), with the species abbreviations Pfa, Pbo, and Ala, for *P. fasciatus*, *P. bockmanni*, and *A. lacustris*, respectively, in addition to the term "Sat" and a catalog number in order of decreasing abundance. The catalogs of satDNA families were deposited on the GenBank with accession numbers OM793143-OM793191 (*P. bockmanni*), OM793192-OM793247 (*P. fasciatus*), and OM793248-OM793279 (*A. lacustris*). In addition, we generated repeat landscapes to estimate the average divergence, considering the distances between sequences based on the Kimura-2-parameter model using the script calcDivergenceFromAlign.pl, of the RepeatMasker suite (**Supplementary Figures S1–S3**).

To better understand the satDNA families observed in all four satellitomes described, we generated variant profiles and coverage depths using the RepeatProfiler pipeline (Negm et al., 2021) to analyze the sequence variation in the genomes studied. In addition, we included a library of *A. mexicanus*, a species model of the *Astyanax* genus, obtained in NCBI, SRA database under the access number SRR6386652. We randomly selected 2 × 1,000,000 reads for all five species, and the target satDNA families were concatenated to a minimum of 200 bp. We used dimers when the satDNA were greater than 200 bp. Read mapping was performed with Bowtie2 (Langmead and Salzberg, 2012) with the preset values "-sensitive" and "no-mixed." We utilized 10 single-copy fish genes to be mapped for single-copy normalization of the read coverage, as described previously (dos Santos et al., 2021). The genes used here were *ppfia1* (XM_022685633.1), *foxl2* (XM_007232295.3), *prospero* (XM_017708821.1), *msh4* (XM_017711771.1), *zdhhc22* (XM_017711775.1), *coq6* (XM_017711829.1), *znf106* (XM_017711848.1), *lactamase* (XM_022682177.1), *gastrula zinc finger* (XM_022685636.1), and *tubulin kinase* (XM_017711762.1).

After verifying a satDNA family related to the cytogenetic marker As51 on the satellitome of *P. bockmanni* (PboSat03-39), characterized by a deletion of 12 bp, we investigated the presence of this variant in the genomes of *P. paranae*, *P. bockmanni*, *P. fasciatus*, *A. lacustris*, and *A. mexicanus*. In addition, a variant of 43 bp was observed in the satellitome of *P. fasciatus*, and we included this sequence in our analysis. We collected monomers of the three variants cited (51 bp, 39 bp, and 43 bp) from the genomes of the five species, using a random selection of 2 × 250,000 reads for each species. We then aligned the isolated reads against each satDNA variant, to only select full reads (Utsunomia et al., 2019). Subsequently, we discarded singletons using CD-HIT software (Li and Godzik, 2006). Due to the large number of monomers obtained, principally in *P. paranae*, we performed a random proportional selection of monomers for each species using Seqtk software (https://github.com/lh3/seqtk). A total of 2044 monomers were utilized, and information on the monomers obtained for each species and the quantity of monomers utilized in our analyses are shown in **Supplementary Table S1**. Finally, we constructed a minimum spanning tree (MST) of the pairwise differences and considered the relative abundance of the haplotypes using software PHYLOViZ 2.0 (Nascimento et al., 2017). The images were produced using the Inkscape software.

Fluorescence *in situ* Hybridization

FISH experiments were performed with eight satDNA families, which were common to the four analyzed species. We utilized primers described by Silva et al. (2017) and probes were labeled with digoxigenin-11dUTP in PCR reactions. FISH experiments were performed following the protocol established by Pinkel et al. (1986), with some modifications (Utsunomia et al., 2017). The metaphasic plate was treated with RNase A (50 μg/ml), for 50 min, with subsequent chromosomal DNA denaturation in 70% formamide/2 × SSC for 2 min, at 70°C. After hybridization, the slides were washed in 0.2 × SSC/15% formamide for 5 min at 42°C, with subsequent washes in 4 × SSC/0.5% Tween-20, at room temperature. Probe detection was performed with anti-digoxigenin-rhodamine (Roche, Basiléia, Switzerland) and the chromosomes were counterstained with DAPI (4′,6-diamino-2-phenylindole, Vector Laboratories, Burlingame, United States). The results were analyzed using an optical microscope (Olympus BX61). Images were captured using the DP Controller software (Olympus®, Hamburg, Germany).

RESULTS

Cytogenetic Analysis

Five individuals from each species were collected. Diploid numbers observed were consistent with the literature, with 2n = 50 for *P. bockmanni* and *A. lacustris*. Populations of *P. fasciatus* demonstrated differential diploid numbers, with 2n = 46 for those from the Alambari River and 2n = 48 for the Araras River. The karyotype formulas were 3m+5sm+6st+11a for *P. bockmanni*,

TABLE 1 | Main characteristics of *Psalidodon fasciatus* satellitome.

satDNA family	RUL	A + T (%)	V	Abundance	Divergence (%)
PfaSat01-51	51	58.8	4	0.09111751	6.16
PfaSat02-237	237	64.5	1	0.03233236	5.25
PfaSat03-97	97	54.6	3	0.02888047	13.62
PfaSat04-51	51	54.9	1	0.01037053	14.5
PfaSat05-71	71	54.9	2	0.01018242	18.04
PfaSat06-85	85	57.6	1	0.00988488	13.97
PfaSat07-31	31	64.9	2	0.00987813	19.69
PfaSat08-42	42	57.1	2	0.00919605	14.18
PfaSat09-177	177	67.2	1	0.00684888	12.49
PfaSat10-61	61	70.4	1	0.00675362	5.6
PfaSat11-21	21	76.1	2	0.00635465	14.7
PfaSat12-68	68	61.7	1	0.00468916	4.77
PfaSat13-91	91	50.5	3	0.00423541	6.39
PfaSat14-40	40	57.5	1	0.00411315	3.02
PfaSat15-187	187	67.3	1	0.00402267	9.25
PfaSat16-61	61	70.4	1	0.00396132	3.81
PfaSat17-59	59	66.1	1	0.00374524	6.13
PfaSat18-27	27	66.6	3	0.00358966	6.42
PfaSat19-22	22	54.5	1	0.00342889	13.93
PfaSat20-76	76	68.4	3	0.00342035	7.47
PfaSat21-109	109	60.5	1	0.00335123	4.62
PfaSat22-24	24	70.8	1	0.00329913	7.18
PfaSat23-236	236	64.4	1	0.00315699	5.99
PfaSat24-83	83	51.8	1	0.00304989	9.26
PfaSat25-46	46	78.2	1	0.00269382	4.78
PfaSat26-54	54	46.2	2	0.00259426	10.15
PfaSat27-197	197	63.4	1	0.00234954	4.03
PfaSat28-51	51	54.9	1	0.00225298	6.8
PfaSat29-190	190	62.1	2	0.00218743	8.99
PfaSat30-85	85	57.6	1	0.00218497	10.27
PfaSat31-42	42	50.0	2	0.00205879	8.6
PfaSat32-65	65	66.1	2	0.00180669	4.15
PfaSat33-286	286	67.8	1	0.0018058	6.02
PfaSat34-103	103	72.8	1	0.00172698	7.39
PfaSat35-142	142	73.9	1	0.00160352	10
PfaSat36-33	33	75.7	1	0.00148574	5.83
PfaSat37-166	166	70.4	1	0.00145701	7.53
PfaSat38-52	52	71.1	2	0.001455	14.65
PfaSat39-100	100	66.0	1	0.00143516	12.24
PfaSat40-143	143	76.2	1	0.00141934	6.73
PfaSat41-6	6	50.0	1	0.00141474	20.4
PfaSat42-51	51	64.7	1	0.00141112	11.93
PfaSat43-191	191	64.9	1	0.00133484	8.45
PfaSat44-141	141	64.5	1	0.0012583	4.02
PfaSat45-41	41	63.4	1	0.00122238	12.8
PfaSat46-54	54	50.0	1	0.00113283	6.39
PfaSat47-35	35	68.5	2	0.00109669	8.86
PfaSat48-27	27	74.0	2	0.00101737	10
PfaSat49-42	42	42.8	1	0.00096175	7.35
PfaSat50-29	29	51.7	1	0.00076152	6.71
PfaSat51-22	22	59.0	2	0.00073328	6.65
PfaSat52-21	21	57.1	1	0.00067703	5.45
PfaSat53-52	52	59.6	1	0.00067471	10.81
PfaSat54-56	56	33.9	1	0.00061054	4.88
PfaSat55-43	43	62.7	1	0.00057551	9.08
PfaSat56-55	55	67.2	1	0.00041117	4.47
PfaSat57-51	51	60.7	1	1.41E-05	20.46

TABLE 2 | Main characteristics of *Psalidodon bockmanni* satellitome.

satDNA family	RUL	A + T (%)	V	Abundance	Divergence (%)
PboSat01-51	51	56.8	4	0.01690405	3.46
PboSat02-235	235	64.2	1	0.00606882	1.96
PboSat03-39	39	48.7	2	0.00547919	1.91
PboSat04-235	235	62.5	1	0.00132786	13.92
PboSat05-84	84	55.9	1	0.00115793	13.99
PboSat06-23	23	52.1	1	0.00083767	11.66
PboSat07-31	31	64.5	1	0.00079599	19.14
PboSat08-188	188	67.5	2	0.00075067	15.74
PboSat09-35	35	62.8	5	0.00071892	7.09
PboSat10-40	40	57.5	2	0.00064097	2.62
PboSat11-27	27	62.9	1	0.00057536	6.64
PboSat12-190	190	61.0	1	0.00053385	9.01
PboSat13-106	106	60.3	2	0.00048688	3.76
PboSat14-61	61	72.1	1	0.00046441	5.84
PboSat15-87	87	72.4	1	0.00043171	5.58
PboSat16-63	63	69.8	1	0.00039598	7.36
PboSat17-69	69	60.8	1	0.00038312	4.45
PboSat18-52	52	46.1	3	0.00035021	8.76
PboSat19-22	22	50.0	1	0.00032650	10.62
PboSat20-107	107	40.1	1	0.00032312	13.37
PboSat21-82	82	56.0	1	0.00030316	8.35
PboSat22-22	22	40.9	2	0.00029609	14.61
PboSat23-54	54	46.2	2	0.00028669	6.36
PboSat24-50	50	66.0	2	0.00027735	14.36
PboSat25-42	42	52.3	2	0.00026625	6.50
PboSat26-21	21	57.1	1	0.00025531	6.30
PboSat27-51	51	54.9	1	0.00024676	6.54
PboSat28-62	62	58.0	1	0.00023092	12.37
PboSat29-142	142	77.4	1	0.00022172	5.90
PboSat30-55	55	67.2	1	0.00020175	3.57
PboSat31-657	657	53.5	1	0.00019655	2.74
PboSat32-193	193	56.4	1	0.00016677	7.80
PboSat33-42	42	42.8	2	0.00015628	12.87
PboSat34-56	56	32.1	4	0.00015562	4.74
PboSat35-584	584	58.3	1	0.00015463	1.98
PboSat36-419	419	49.8	1	0.00014438	4.05
PboSat37-188	188	65.9	1	0.00014396	9.31
PboSat38-91	91	51.6	3	0.00014333	7.47
PboSat39-59	59	64.4	1	0.00014129	2.57
PboSat40-78	78	56.4	3	0.00014103	5.13
PboSat41-204	204	49.5	1	0.00013263	9.14
PboSat42-112	112	63.3	1	0.00012000	4.51
PboSat43-52	52	67.3	1	0.00011740	12.35
PboSat44-6	6	50.0	1	0.00011095	17.28
PboSat45-220	220	57.7	1	0.00010466	8.41
PboSat46-90	90	67.7	1	0.00010049	3.87
PboSat47-52	52	69.2	2	9.18E-06	6.15
PboSat48-32	32	50.0	1	8.58E-06	4.15
PboSat49-42	42	57.1	1	8.56E-06	7.17
PboSat50-62	62	66.1	2	7.25E-06	6.69

3m+5sm+12st+4a for *P. fasciatus* (Araras), 4m+7sm+9st+3a for *P. fasciatus* (Alambari), and 3m+9sm+9st+4a for *A. lacustris*. The species *P. fasciatus* is part of a "species complex", with diploid numbers varying between 45 and 49, so our results are consistent with the literature (Kantek et al., 2009; Pasa et al., 2021). None of the analyzed individuals had B chromosomes.

Description of Two New Satellitomes of *Psalidodon* and the First Satellitome of *Astyanax*

After several iterations with the satMiner toolkit protocol (6 for *P. fasciatus*, seven for *P. bockmanni*, and three for *A. lacustris*), until

TABLE 3 | Main characteristics of *Astyanax lacustris* satellitome.

satDNA family	RUL	A + T (%)	V	Abundance	Divergence (%)
AlaSat01-91	91	54.9	8	0.0019966	5.1
AlaSat02-186	186	64.5	1	0.00186961	0.98
AlaSat03-3028	3028	64.0	1	0.00191953	5.63
AlaSat04-151	151	63.5	2	0.00163407	2.49
AlaSat05-364	364	50.2	1	0.00164291	5.98
AlaSat06-42	42	54.7	1	0.00095201	14.46
AlaSat07-189	189	64.5	4	0.00095719	9.32
AlaSat08-236	236	63.9	1	0.00098154	12.74
AlaSat09-31	31	64.5	1	0.00093329	19.82
AlaSat10-84	84	57.1	2	0.0004786	15.78
AlaSat11-6	6	50.0	1	0.00046441	17.28
AlaSat12-177	177	66.6	1	0.00035606	16.02
AlaSat13-24	24	70.8	1	0.00031664	8.12
AlaSat14-62	62	70.9	2	0.00031979	7.96
AlaSat15-69	69	59.4	1	0.00027686	4.55
AlaSat16-251	251	56.9	1	0.00030137	11.4
AlaSat17-58	58	67.2	3	0.00028017	8.85
AlaSat18-80	80	71.2	1	0.00029079	9.31
AlaSat19-106	106	60.3	3	0.00022585	5.99
AlaSat20-85	85	57.6	2	0.0002287	13.25
AlaSat21-899	899	56.5	1	0.0002351	4.64
AlaSat22-22	22	45.4	2	0.00019073	14.02
AlaSat23-1242	1242	52.8	1	0.00022293	0.86
AlaSat24-577	577	58.7	1	0.00021396	2.13
AlaSat25-52	52	67.3	2	0.00017473	13.82
AlaSat26-418	418	52.3	1	0.00016463	4.59
AlaSat27-35	35	68.5	2	0.00013577	7.57
AlaSat28-574	574	64.9	1	0.00014314	5.37
AlaSat29-185	185	40.0	1	0.00013656	6.23
AlaSat30-352	352	69.3	1	0.00013661	8.13
AlaSat31-552	552	52.8	1	0.00013736	0.89
AlaSat32-187	187	68.4	3	0.00011427	9.54
AlaSat33-22	22	63.6	1	1.00E-07	29.07

no satDNA was uncovered, we found 57 families of satDNAs for *P. fasciatus* (80 variants), 50 for *P. bockmanni* (77 variants) and 33 for *A. lacustris* (54 variants). The length distribution of satDNA families revealed the predominance of short satDNAs (<100 bp) for *P. fasciatus* (42) and *P. bockmanni* (35), corresponded to 73.9 and 70.0% the satellitomes, respectively. In contrast, long satDNAs predominated in *A. lacustris* (18), corresponding to 54.5% of the satellitome. The repeat unit length (RUL) ranged from 6 to 286 bp for *P. fasciatus* (median 82.08); 6 to 584 for *P. bockmanni* (median 111.3) and 6 to 3028 in *A. lacustris* (median 316.27). The A + T content varied between 33.9 and 78.2% for *P. fasciatus*, (median 62.1%); 32.1–72.4% in *P. bockmanni* (median 57.9%), and 40.0–71.2% in *A. lacustris* (median 60.4%), indicating a predominance of A + T-rich satDNAs. Complete information on the three new satellitomes is presented in **Tables 1**–**3**. The Shapiro-Wilks test demonstrated that only the A + T content showed a normal distribution (W = 0.973, $p = 0.235$ for *P. fasciatus*, W = 0.984, $p = 0.34$ for *P. bockmanni*, and W = 0.945, $p = 0.115$ for *A. lacustris*). Kendall's rank correlation test demonstrated that the only correlation between traces was a negative correlation between RUL and divergence in *A. lacustris* (tau = −0.465).

TABLE 4 | Superfamilies characterized in satellitomes of *Psalidodon fasciatus*, *Psalidodon bockmanni,* and *Astyanax lacustris*.

	SatDNA	SatDNA	SatDNA	Similarity
P. bockmanni				
SF1	PboSat01-51	PboSat03-39	-	68.27%
SF2	PboSat02-235	PboSat04-235	-	78.90%
SF3	PboSat06-23	PboSat22-22	-	65.21%
P. fasciatus				
SF1	PfaSat01-51	PfaSat55-43	PfaSat57-51	78.43%/69.23%
SF2	PfaSat02-237	PfaSat23-236	-	79.83%
SF3	PfaSat17-59	PfaSat42-51	-	60.65%
SF4	PfaSat22-24	PfaSat36-33	-	72.72%
A. lacustris				
SF1	AlaSat22-22	AlaSat33-22	-	72.72%

Comparisons between satDNA families for each species detected homology among sequences. Four superfamilies were detected in *P. fasciatus*, three in *P. bockmanni,* and only one in *A. lacustris* (**Table 4**). In most cases, the variation between sequences involved in a superfamily was caused by nucleotide substitutions. However, deletions of segments from 8 to 10 bp were observed in one of the sequences of SF1 in *P. bockmanni* and SF1, SF3, and SF4 in *P. fasciatus* (**Supplementary Material**).

Comparative Analysis Demonstrated the Conservation of Several Satellite DNA Families Between *Psalidodon* and *Astyanax*

We performed a comparative analysis between the three satellitomes described in this study and that of *P. paranae*, using the RepeatMasker software. Of a total of 104 satDNA families present in the four species, 10 were observed in all species analyzed. One of them were identified as the telomeric sequence (ApaSat07-06-tel) that was included in other fish satellitomes (Silva et al., 2017; Utsunomia et al., 2019; Stornioli et al., 2021), and other was identified as CharSat01-52 (ApaSat29-52), which was conserved of the several species in Characidae family (dos Santos et al., 2021). Eight other sequences were observed in all four species, with at least 50% similarity (**Table 5**). In addition, several sequences were detected in only two or three species analyzed, as shown in **Table 6**, along with their degrees of similarity.

RepeatProfiler Reveals Highly Conserved Satellite DNA Families Between Genera

We generated RepeatProfiler plots of the ten satDNAs shared between the four species. In addition, we included the genome of *Astyanax mexicanus* for this analysis. We represent profiles of ApaSat12-69, and ApaSat30-50 in **Figure 1**, and remaining are in **Supplementary Material**. Our results revealed a high degree of conservation of ApaSat11-22 and ApaSat12-69 for all five species analyzed, with a similar degree of abundance in all groups. As expected, in all cases, the profiles demonstrated greater similarity between species of the *Psalidodon* than those of the *Astyanax* (ApaSat03-91, ApaSat29-52, ApaSat30-50, and ApaSat40-189). In addition, the two satDNA families (ApaSat02-236 and

TABLE 5 | SatDNA families with at least 50% of similarity in three species of *Psalidodon* and one species of *Astyanax*.

P. paranae	*P. bockmanni*	*P. fasciatus*	*A. lacustris*
ApaSat02-236	Pbosat02-235/Pbosat04-235	Pfasat02-237/Pfasat23-236	Alasat08-236
Apasat03-91	Pbosat38-91	Pfasat13-91	Alasat01-91
Apasat04-233	Pbosat02-235/Pbosat04-235	Pfasat02-237/Pfasat23-236	Alasat08-236
Apasat07-6-tel	Pbosat44-6	Pfasat41-6	Alasat11-6
Apasat08-35	Pbosat09-35	Pfasat47-35	Alasat27-35
Apasat11-22	Pbosat19-22	Pfasat19-22	Alasat22-22/Alasat33-22
Apasat12-69	Pbosat17-69	Pfasat12-68	Alasat15-69
Apasat29-52	Pbosat43-52	Pfasat38-52	Alasat25-52
Apasat30-50	Pbosat24-50	Pfasat17-59/Pfasat42-51	Alasat17-58
Apasat40-189	Pbosat37-188	Pfasat43-191	Alasat07-189

TABLE 6 | SatDNAs similarity in two or three species analyzed. Similarity in superfamilies level (between 50 and 80%) are highlighted (*).

P. paranae	*P. bockmanni*	*P. fasciatus*	*A. lacustris*
Apasat01-51	Pbosat01-51	Pfasat01-51	-
	Pbosat03-39*	Pfasat55-43*	
		Pfasat57-51*	
Apasat05-23	Pbosat06-23	-	-
Apasat13-23	Pbosat22-22*		
Apasat06-86	-	Pfasat24-83	-
Apasat09-21	-	Pfasat11-21	-
Apasat10-179	-	-	Alasat02-186
Apasat15-51	Pbosat27-51	Pfasat28-51	-
Apasat16-54	Pbosat18-52	Pfasat26-54	-
Apasat17-365	-	-	Alasat05-364
Apasat18-58	Pbosat23-54	Pfasat46-54	-
Apasat19-77*	Pbosat16-63*	-	-
Apasat22-62	Pbosat28-62	-	-
Apasat24-78	Pbosat40-78	-	-
Apasat27-178*	-	-	Alasat29-185*
Apasat33-112	Pbosat42-112	-	-
Apasat36-21	Pbosat26-21	Pfasat52-21	-
Apasat38-107	Pbosat20-107	-	-
Apasat39-32	Pbosat48-32	-	-
Apasat42-90	Pbosat46-90	Pfasat32-65*	-
-	Pbosat05-84	Pfasat06-85	Alasat10-84
-	Pbosat07-31	Pfasat07-31	Alasat09-31
-	Pbosat08-188*	Pfasat15-187	Alasat32-187
-	Pbosat10-40	Pfasat14-40	-
-	Pbosat11-27	Pfasat18-27	-
-	Pbosat12-190	Pfasat29-190	-
	Pbosat13-106	Pfasat21-109	Alasat19-106
-	Pbosat14-61	Pfasat10-61	Alasat14-62
-	Pbosat15-87	Pfasat20-76	Alasat18-80
-	Pbosat21-82	Pfasat30-85	Alasat20-85
-	Pbosat25-42	Pfasat31-42	-
-	Pbosat29-142	Pfasat40-143	-
-	Pbosat30-55	Pfasat56-55	-
-	Pbosat31-657	-	Alasat31-552
-	Pbosat33-42	Pfasat49-42	-
-	Pbosat34-56	Pfasat54-56	-
-	Pbosat35-584	-	Alasat24-577
-	Pbosat36-419	-	Alasat26-418
-	-	Pfasat08-42	Alasat06-42
-	-	Pfasat09-177	Alasat12-177
-	-	Pfasat22-24	Alasat13-24
		Pfasat36-33*	
-	-	Pfasat27-197*	Alasat04-151*

ApaSat04-233) observed as centromeric sequences, each demonstrated a large deletion in the *Astyanax* species, as well as different abundances in the *Psalidodon* and different mutations fixed on the species of this group.

Diversification of Abundant SatDNA in *Psalidodon*: New Variants Observed in *Psalidodon*

The As51 satDNA family was present in the satellitomes of *P. paranae*, *P. fasciatus*, and *P. bockmanni*, corresponding to the most abundant satDNA in these three species. However, this sequence is part of the superfamilies of *P. fasciatus* and *P. bockmanni*. New variants in these two species were produced mainly by deletion of parts of the original sequence, resulting in variants of 39 bp (PboSat03-39) and 43 bp (Pfasat55-43) (**Figure 2**). We produced a minimum spanning tree (MST) of the As51 satDNA and its variants using monomers extracted from *P. paranae*, *P. fasciatus*, *P. bockmanni*, *A. lacustris*, and *A. mexicanus*, excluding the sequence variants found only once (singletons) (**Figure 2**). The MST of the 39 bp variant was restricted to *P. bockmanni* and *P. paranae*, although this sequence was missing from the *P. paranae* satellitome, with several haplotypes shared between these two species, including the most abundant. The variant of 43 bp was restricted to *P. fasciatus*, with low abundance. None of the As51 monomers were shared among more than two species, and most only between *P. paranae* and *P. bockmanni* or *P. paranae* and *P. fasciatus*, corroborating the phylogeny of the group. Monomers of As51 were isolated in the genomes of *A. lacustris* and *A. mexicanus*, despite the absence of these satDNAs in the satellitomes of these species and the absence of FISH signals on their chromosomes.

Cytogenetic Mapping of Conserved Satellite DNA Families in *Astyanax* and *Psalidodon*

We performed cytogenetic mapping of eight of the conserved satDNA families in the three species (**Figure 3**), except for the telomeric sequence (Apasat07-6-tel) and Apasat08-35, in

which PCR amplification failed. We utilized the metaphase plates of *A. lacustris*, *P. bockmanni*, and two cytotypes of *P. fasciatus*. None of the satDNA families analyzed here demonstrated clusters on the chromosomes of *A. lacustris*, so we considered them as non-clustered in this species (**Supplementary Figure S11**). Additionally, Apasat29-52 did not cluster with any individual in our analysis. Clustered satDNAs were mainly present in heterochromatic subtelomeric and centromeric areas. All species demonstrated the same pattern of clusterization of satDNAs, but the number of chromosomes with cluster signals varied. We highlight the following: 1—Apasat02-236 and Apasat04-233 were clustered in pericentromeric regions, with Apasat02-236 present in all chromosomes of *P. fasciatus* and *P. bockmanni*, and Apasat04-233 in approximately half of the chromosomes of *P. bockmanni* and absent only on in par 12 in *P. fasciatus;* 2—Apasat30-50 demonstrated conserved clustered positions on the short arms of a pair of metacentric chromosomes in *P. bockmanni* and *P. fasciatus*; and 3—all other satDNAs demonstrated clusters in subtelomeric regions, with the exception of Apasat11-22 that had clusters in the interstitial regions of a pair of acrocentric chromosomes in *P. bockmanni* and two pairs of subtelocentric chromosomes in *P. fasciatus*.

DISCUSSION

In this study, we performed, for the first time for Neotropical fishes, an evolutionary comparison of the complete satellitome in four species. We observed a high retention of satDNAs in *Psalidodon* and *Astyanax*, demonstrated by the low number of species-exclusive satDNAs (8 for *P. bockmanni*, 15 for *P. fasciatus*, six for *A. lacustris,* and 16 for *P. paranae*).

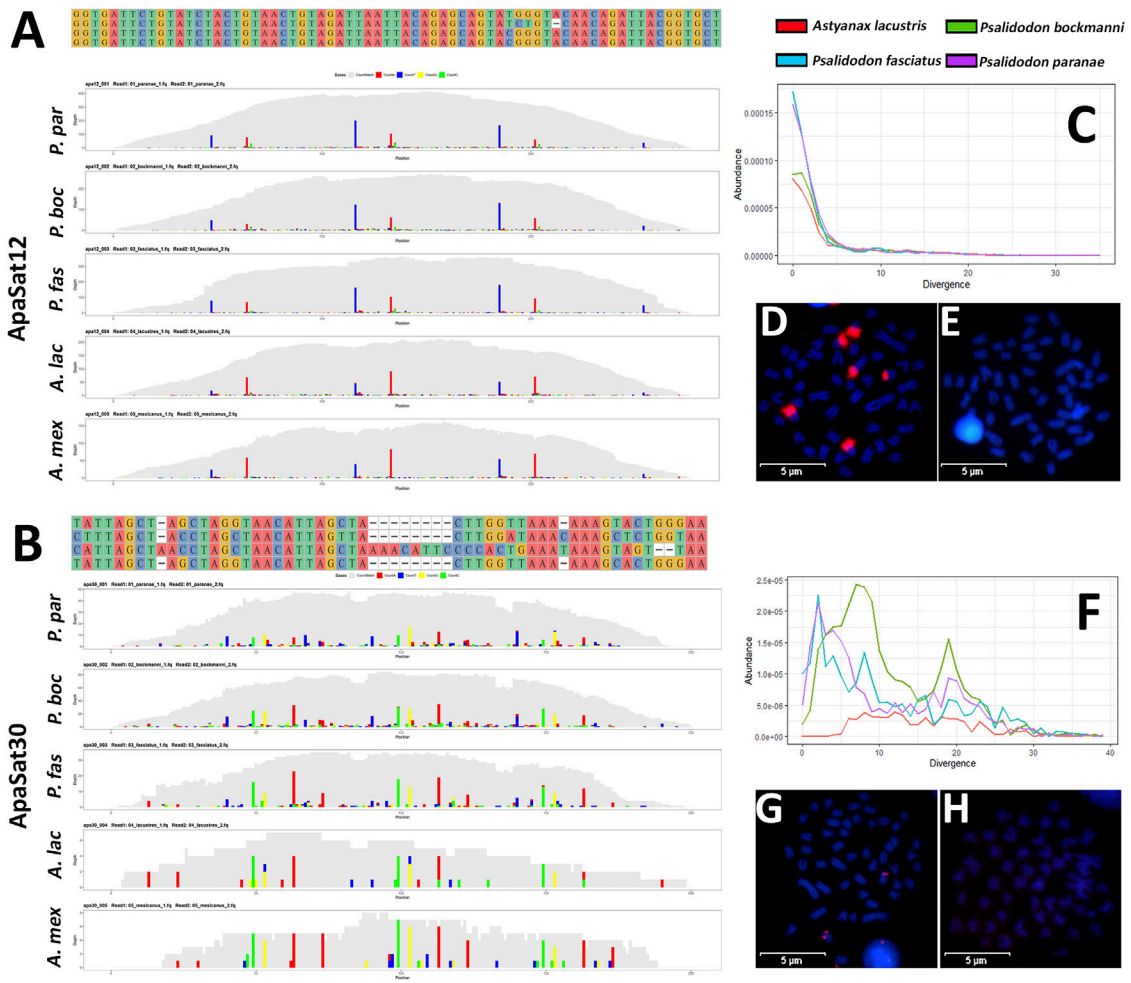


FIGURE 1 | Repeats landscapes of conserved satDNA families conserved between Psalidodon and Astyanax. Alignments and repeat profilers of ApaSat12-69 **(A)** and Apasat30-50 **(B)**. Repeat landscapes of Apasat12-69 **(C)**, Apasat30-50 **(F)** demonstrate the abundance and divergence of these satDNAs in the analyzed species. In addition, these satDNAs demonstrated FISH signals in all Psalidodon species **(D and G)**, but not in Astyanax **(E and H)**.

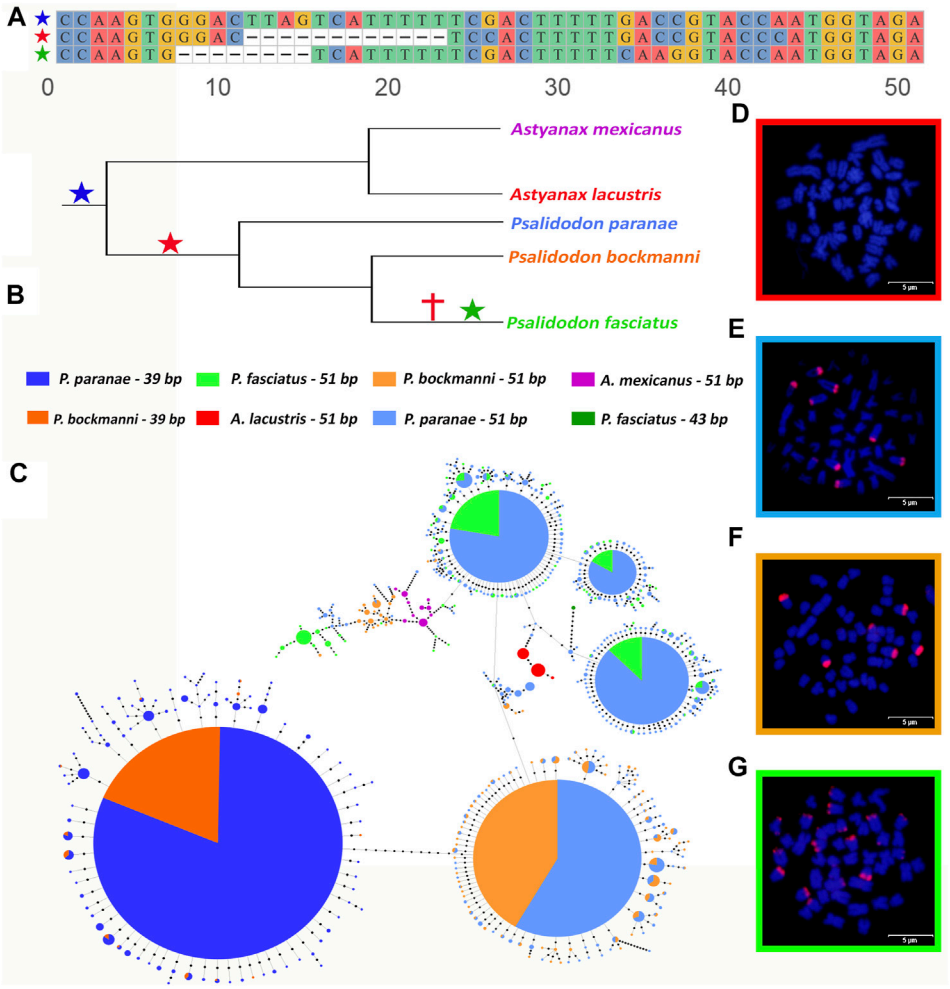


FIGURE 2 | Structure and conservation of As51 monomers between *Psalidodon* and *Astyanax*. Alignment of As51 variants, with 51 bp, 39 bp, and 43 bp, respectively **(A)**. The emergence and disappearance of each variant is demonstrated in the phylogeny of species (blue = 51 bp, red = 39 bp, green = 43 bp) **(B)**. Stars represents the surging of a variation of As51 and crosses represents the elimination of a variation of As51. Linear MST demonstrating the haplotypes of variants of As51 in *Astyanax* and *Psalidodon* species **(C)**. Highlight of the dominance of a haplotype of As51-39 bp shared between *Psalidodon bockmanni* and *Psalidodon paranae*, and the most relevant of four haplotypes of As51-51 bp shared between *P. bockmanni* and *P. paranae* (1) and *P. paranae* and *Psalidodon fasciatus* (3) **(C)**. As51 do not demonstrates FISH signals in *Astyanax lacustris* **(D)**, but this marker forms greats cluster on chromosomes of *P. bockmanni*, *P. paranae*, and *P. fasciatus* **(E–G)**. Fish images were captured in a magnification of ×1000.

According to the library hypothesis (Fry and Salser, 1977), a group of related species should share a common library of satDNAs, and satellitomes can demonstrate quantitative differences among species due to differential amplification. Therefore, in addition to the possibility that these species-specific satDNAs have appeared *de novo*, future studies could find monomers of these sequences in low abundance on the genome of the other correlated species. A high degree of satDNA families found in the four satellitomes analyzed were shared between the species (**Tables 5**, **6**), supporting the existence of a common library. However, divergences in the abundance of correlated satDNA families were found, as predicted by the library hypothesis (Fry and Salser, 1977). As an example, Alasat07-189 was correlated with Apasat40-189. Changes in satDNA abundance can occur by unequal crossing-over (Garrido-Ramos, 2017), replication slippage (Walsh, 1987; Stephan, 1989; Ruiz-Ruano et al., 2018), replication of extrachromosomal circles of tandem repeats by rolling-circle replication (Cohen et al., 2005; 2010), and transposition element actions (Jurka et al., 2005; Šatović and Plohl, 2013). Comparative analyses of satellitomes of correlated species also found a high degree of shared satDNA families, as in the grasshoppers *Locusta migratoria* and *Oedaleus decorus* (Camacho et al., 2022), In this case, association between satDNAs families and transposable elements were observed, as LmiSat02-176 and OdeSat17-176 associated with Helitron TEs (Camacho et al., 2022).

The A + T content was the only characteristic with a normal distribution in the satellitomes of *P. bockmanni*, *P. fasciatus*, and *A. lacustris*, similar to that found in the satellitomes of *P. paranae*

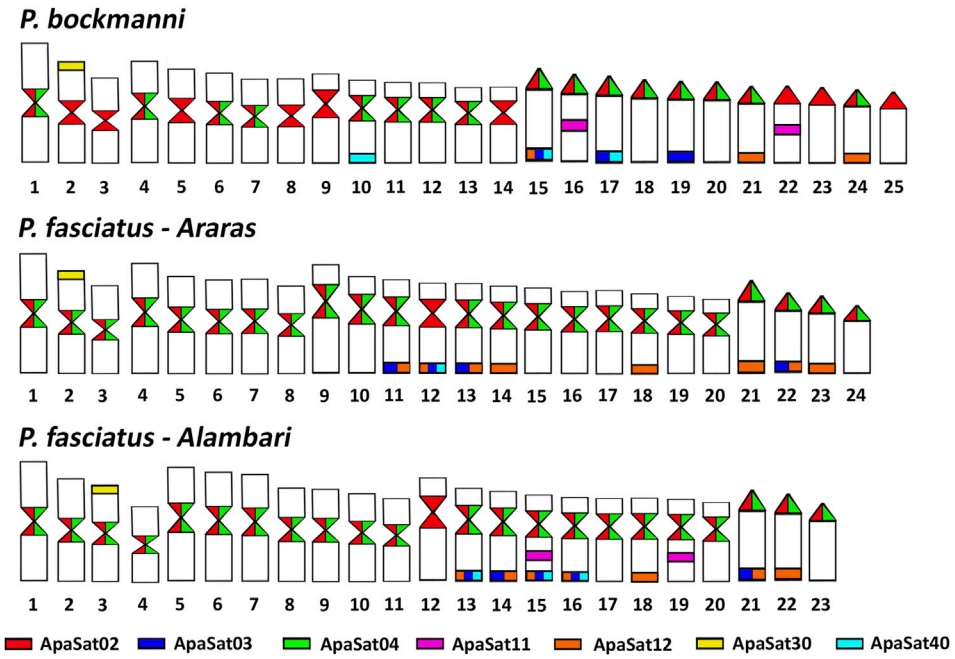


FIGURE 3 | Ideogram of clustered conserved satDNAs in *Psalidodon bockmanni* and two populations of *Psalidodon fasciatus*. No signal was observed in *Astyanax lacustris*.

(Silva et al., 2017) and *Megaleporinus macrocephalus* (Utsunomia et al., 2019). However, no correlations were observed between A + T content and RUL, as identified in *P. paranae*, or divergence and abundance, as identified in *M. macrocephalus*.

In addition to the Charsat01-52 and telomeric sequence, we found another eight satDNA families present on the four satellitomes. These satDNAs were maintained from 11.2 mya, when *A. lacustris* diverged from *Psalidodon* (Piscor et al., 2019). The maintenance of satDNA families across different species can occur through the biological function of determinate satDNA (Fry and Salser, 1977) or independence of natural selection (Stephan, 1986; Stephan, 1987; Walsh, 1987; Stephan, 1989; Harding et al., 1992). Despite the occurrence of the transcribed monomers of Charsat01-52 in *P. paranae* (dos Santos et al., 2021), we did not test the transcription of conserved satDNA in our satellitomes. However, the presence of Apasat02-236 and Apasat04-233 in the centromeres of all chromosomes in *Psalidodon* individuals suggests some structural function of these satDNAs. These sequences were related to Alasat08-236, and no FISH signal was observed in *A. lacustris*. Evidence in grasshoppers demonstrated that a satDNA family may be involved in centromeric function in one species, but not in other related species, suggesting that the species had replaced the centromeric satDNA during the evolution process (Camacho et al., 2022). It is common that the more abundant satDNAs are probably involved in centromeric function (Melters et al., 2013), as observed in *P. paranae* (ApaSat02-236 and ApaSat04-233) (Silva et al., 2017); however, in *Eumingus monticola*, the eighth satDNAs in order of decreasing abundance is located only in the pericentromeric regions (Camacho et al., 2022). In addition, examples of species with different satDNAs present in centromeres are common, such as those of chickens (Shang et al., 2010), plants (Iwata et al., 2013), and fishes (*Prochilodus lineatus;* Stornioli et al., 2021).

During the description of the satellitome of *P. paranae* (Silva et al., 2017), those authors obtained FISH signals of *P. paranae* satDNAs from *P. bockmanni* and *P. fasciatus*. Our analyses corroborated their results, with the addition of clustered signals of Apasat12-69 and Apasat40-189 in *P. bockmanni* and *P. fasciatus*. However, FISH signals were not observed in *A. lacustris*. According to the species tree (Silva et al., 2014) *P. paranae* and *P. bockmanni* are closely related species. We observed that approximately 50% of the satDNA families of *P. bockmanni* had some similarity with satDNAs of *P. paranae*, corroborating these results. In addition, the 39 bp variant of As51 (Apasat01-51) was present only in *P. paranae* and *P. bockmanni*.

The As51 satDNA was characterized by digestion of the *KpnI* restriction enzyme in *P. scabripinnis* (Mestriner et al., 2000), and is the most commonly used satDNA cytogenetic marker in this group, with FISH signals in *P. paranae* (Silva et al., 2014), *P. scabripinnis* (Mestriner et al., 2000), *P. fasciatus* (Kantek et al., 2009), and several other species. The description of the satellitome of *P. paranae* revealed that As51 was the most abundant satDNA in this species, corresponding to Apasat01-51. Our results demonstrated that it was the most also abundant in the satellitomes of *P. bockmanni* and *P. fasciatus*, despite its absence in *A. lacustris* satellitome. In addition, variants of this sequence have been described for the first time in *P. fasciatus* (Afasat55-43) and *P. bockmanni* (Abosat03-39, also present in *P. paranae*). Our data suggest the

presence of As51 in an ancestor of *Psalidodon* and *Astyanax*, due to the identification of As51-51 monomers in the genomes of *A. lacustris* and *A. mexicanus*, where it remains as a relic in these species, with the absence of other variants in *Astyanax*. Therefore, we suggest the emergence of As51-51 in an ancestor of *Psalidodon*, and *Astyanax* (11.2 mya), with subsequent amplification and diversification of this satDNA in *Psalidodon* resulting in variants of 39 bp (6.5 mya) and 43 bp (2 mya). The absence of a 39 bp variant in *P. fasciatus* may be derived from stochastic processes that have led to significant nucleotide divergence. Similar cases were observed in *Drosophila*, with 1.688 satDNA conserved in a subgroup of species, with the exception of *D. kikkawai* and *D. leontia* (de Lima et al., 2020).

Our results expand the knowledge of the conservation and evolution of satDNAs in *Psalidodon* and *Astyanax*, demonstrating a large degree of sharing of sequences between these genera. In addition, we describe the evolutionary history of As51 with expansion and diversification of this sequence in *Psalidodon*.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

ETHICS STATEMENT

The animal study was reviewed and approved by Society for Laboratory Animal ScienceCollege of Animal Experimentation (SBCAL), and approved (protocol 1227) by Bioscience Institute/ UNESP Ethics Committee on the Use of Animals (CEEAA/IBB/ UNESP).

AUTHOR CONTRIBUTIONS

CG, RU, and FP-F conceived the study and wrote the manuscript. CG, RdS, WA, and DA conduced the experiments. CG, DS, and RU analysed the data. All authors read and approval the final version.

FUNDING

This study was supported by the FAPERJ (Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro, grants 201.289/2021 and 211.475/2019), FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, grant 2018/ 03365-3), CAPES, and CNPq.

ACKNOWLEDGMENTS

We would like to thank the financial support of FAPESP, CAPES, CNPq, and FAPERJ, and Francisco Ruiz-Ruano and Juan Pedro Camacho to support to this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.884072/full#supplementary-material

REFERENCES

Abel, L. D. D. S., Mantovani, M., and Moreira-Filho, O. (2006). Chromosomal Distribution of the As51 Satellite DNA in Two Species Complexes of the Genus *Astyanax* (Pisces, Characidae). *Genet. Mol. Biol.* 29 (3), 448–452. doi:10.1590/s1415-47572006000300008

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30 (15), 2114–2120. doi:10.1093/bioinformatics/btu170

Camacho, J. P. M., Cabrero, J., López-León, M. D., Martín-Peciña, M., Perfectti, F., Garrido-Ramos, M. A., et al. (2022). Satellitome Comparison of Two Oedipodine Grasshoppers Highlights the Contingent Nature of Satellite DNA Evolution. *BMC Biol.* 20, 36. doi:10.1186/s12915-021-01216-9

Charlesworth, B., Sniegowski, P., and Stephan, W. (1994). The Evolutionary Dynamics of Repetitive DNA in Eukaryotes. *Nature* 371 (6494), 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 (1), 11. doi:10.1186/1759-8753-1-11

Cohen, S., Agmon, N., Yacobi, K., Mislovati, M., and Segal, D. (2005). Evidence for Rolling Circle Replication of Tandem Genes in *Drosophila*. *Nucleic Acids Res.* 33 (14), 4519–4526. doi:10.1093/nar/gki764

de Lima, L. G., Hanlon, S. L., and Gerton, J. L. (2020). Origins and Evolutionary Patterns of the 1.688 Satellite DNA Family in *Drosophila* Phylogeny. G3-Genes. *Genom. Genet.* 10 (11), 4129–4146. doi:10.1534/g3.120.401727

dos Santos, R. Z., Calegari, R. M., Silva, D. M. Z. A., Ruiz-Ruano, F. J., Melo, S., Oliveira, C., et al. (2021). A Long-Term Conserved Satellite DNA that Remains Unexpanded in Several Genomes of Characiformes Fish Is Actively Transcribed. *Genome Biol. Evol.* 13 (2), evab002. doi:10.1093/gbe/evab002

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

Elder, J. F., Jr, and Turner, B. J. (1995). Concerted Evolution of Repetitive DNA Sequences in Eukaryotes. *Q. Rev. Biol.* 70 (3), 297–320. doi:10.1086/419073

Foresti, F., Almeida Toledo, L. F., and Toledo-Filho, S. A. (1981). Polymorphic Nature of Nucleolus Organizer Regions in Fishes. *Cytogenet. Genome Res.* 31 (3), 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 (4), 1069–1084. doi:10.1016/0092-8674(77)90170-2

Garrido-Ramos, M. (2017). Satellite DNA: An Evolving Topic. *Genes* 8 (9), 230. doi:10.3390/genes8090230

Harding, R. M., Boyce, A. J., and Clegg, J. B. (1992). The Evolution of Tandemly Repetitive DNA: 1137 Recombination Rules. *Genetics* 132 (3), 847–859. doi:10.1093/genetics/132.3.847

Iwata, A., Tek, A. L., Richard, M. M., Abernathy, B., Fonsêca, A., Schmutz, J., et al. (2013). Identification and Characterization of Functional Centromeres of the Common Bean. *Plant J.* 76 (1), 47–60. doi:10.1111/tpj.12269

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 (1–4), 462–467. doi:10.1159/000084979

Kantek, D. L. Z., Vicari, M. R., Peres, W. A. M., Cestari, M. M., Artoni, R. F., Bertollo, L. A. C., et al. (2009). Chromosomal Location and Distribution of As51 Satellite DNA in Five Species of the Genus *Astyanax* (Teleostei, Characidae, Incertae Sedis). *J. Fish. Biol.* 75 (2), 408–421. doi:10.1111/j.1095-8649.2009.02333.x

Kuhn, G. C. S. (2015). Satellite DNA Transcripts Have Diverse Biological Roles in *Drosophila*. *Heredity* 115 (1), 1–2. doi:10.1038/hdy.2015.12

Langmead, B., and Salzberg, S. L. (2012). Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* 9 (4), 357–359. doi:10.1038/nmeth.1923

Li, W., and Godzik, A. (2006). Cd-hit: A Fast Program for Clustering and Comparing Large Sets of Protein or Nucleotide Sequences. *Bioinformatics* 22 (13), 1658–1659. doi:10.1093/bioinformatics/btl158

López-Flores, I., and Garrido-Ramos, M. A. (2012). "The Repetitive DNA Content of Eukaryotic Genomes," in *Repetitive DNA*. Editor M.A. Garrido-Ramos (Basel: Karger), 7, 1–28. doi:10.1159/000337118

Lorite, P., Muñoz-López, M., Carrillo, J. A., Sanllorente, O., Vela, J., Mora, P., et al. (2017). Concerted Evolution, a Slow Process for Ant Satellite DNA: Study of the Satellite DNA in the *Aphaenogaster* Genus (Hymenoptera, Formicidae). *Org. Divers. Evol.* 17 (3), 595–606. doi:10.1007/s13127-017-0333-7

Melters, D. P., Bradnam, K. R., Young, H. A., Telis, N., May, M. R., Ruby, J., et al. (2013). Comparative Analysis of Tandem Repeats from Hundreds of Species Reveals Unique Insights into Centromere Evolution. *Genome Biol.* 14 (1), R10. doi:10.1186/gb-2013-14-1-r10

Mestriner, C. A., Galetti, P. M., 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

Montiel, E. E., Panzera, F., Palomeque, T., Lorite, P., and Pita, S. (2021). Satellitome Analysis of *Rhodnius Prolixus*, One of the Main Chagas Disease Vector Species. *Int. J. Mol. Sci.* 22 (11), 6052. doi:10.3390/ijms22116052

Nascimento, M., Sousa, A., Ramirez, M., Francisco, A. P., Carriço, J. A., and Vaz, C. (2017). PHYLOViZ 2.0: Providing Scalable Data Integration and Visualization for Multiple Phylogenetic Inference Methods. *Bioinformatics* 33 (1), 128–129. doi:10.1093/bioinformatics/btw582

Negm, S., Greenberg, A., Larracuente, A. M., and Sproul, J. S. (2021). RepeatProfiler: A Pipeline for Visualization and Comparative Analysis of Repetitive DNA Profiles. *Mol. Ecol. Resour.* 21 (3), 969–981. doi:10.1111/1755-0998.13305

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 (6), 792–793. doi:10.1093/bioinformatics/btt054

Pasa, R., Menegídio, F. B., Rodrigues-Oliveira, I. H., da Silva, I. B., de Campos, M. L. C. B., Rocha-Reis, D. A., et al. (2021). Ten Complete Mitochondrial Genomes of *Gymnocharacini* (Stethaprioninae, Characiformes). Insights into Evolutionary Relationships and a Repetitive Element in the Control Region (*D-Loop*). *Front. Ecol. Evol.* 9, 456. doi:10.3389/fevo.2021.650783

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 (9), 2934–2938. doi:10.1073/pnas.83.9.2934

Piscor, D., Pozzobon, A. P. B., Fernandes, C. A., Centofante, L., and Parise-Maltempi, P. P. (2019). Molecular Clock as Insight to Estimate the Evolutionary History and Times of Divergence for 10 Nominal *Astyanax* Species (Characiformes, Characidae): An Evolutionary Approach in Species with 2n= 36, 46, 48, and 50 Chromosomes. *Zebrafish* 16 (1), 98–105. doi:10.1089/zeb.2018.1647

Pita, S., Panzera, F., Mora, P., Vela, J., Cuadrado, Á., Sánchez, A., et al. (2017). Comparative Repeatome Analysis on *Triatoma Infestans* Andean and Non-Andean Lineages, Main Vector of Chagas Disease. *PLoS One* 12 (7), e0181635. doi:10.1371/journal.pone.0181635

Plohl, M., Meštrović, N., and Mravinac, B. (2012). "Satellite DNA Evolution," in *Repetitive DNA*. Editor M. A. Garrido-Ramos (Basel: Karger Publishers), 7, 126–152. doi:10.1159/000337122

Rodrigues, P. H. D. M., dos Santos, R. Z., Silva, D. M. Z. D. A., Goes, C. A. G., Oliveira, C., Foresti, F., et al. (2019). Chromosomal and Genomic Dynamics of Satellite DNAs in Characidae (Characiformes, Teleostei) Species. *Zebrafish* 16 (4), 408–414. doi:10.1089/zeb.2019.1738

Ruiz-Ruano, F. J., Castillo-Martínez, J., Cabrero, J., Gómez, R., Camacho, J. P. M., and López-León, M. D. (2018). High-throughput Analysis of Satellite DNA in the Grasshopper *Pyrgomorpha Conica* Reveals Abundance of Homologous and Heterologous Higher-Order Repeats. *Chromosoma* 127 (3), 323–340. doi:10.1007/s00412-018-0666-9

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 (1), 28333. doi:10.1038/srep28333

Šatović, E., and Plohl, M. (2013). Tandem Repeat-Containing MITEs in the Clam *Donax Trunculus*. *Genome Biol. Evol.* 5 (12), 2549–2559. doi:10.1093/gbe/evt202

Schmieder, R., and Edwards, R. (2011). Quality Control and Preprocessing of Metagenomic Datasets. *Bioinformatics* 27 (6), 863–864. doi:10.1093/bioinformatics/btr026

Shang, W.-H., Hori, T., Toyoda, A., Kato, J., Popendorf, K., Sakakibara, Y., et al. (2010). Chickens Possess Centromeres with Both Extended Tandem Repeats and Short Non-Tandem-Repetitive Sequences. *Genome Res.* 20 (9), 1219–1228. doi:10.1101/gr.106245.110

Silva, D. M. Z. D. 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 (4), e94896. doi:10.1371/journal.pone.0094896

Silva, D. M. Z. D. A., Utsunomia, R., Ruiz-Ruano, F. J., Daniel, S. N., Porto-Foresti, F., Hashimoto, D. T., et al. (2017). High-Throughput Analysis Unveils a Highly Shared Satellite DNA Library Among Three Species of Fish Genus *Astyanax*. *Sci. Rep.* 7, 12726. Available at: https://www.nature.com/articles/s41598-017-12939-7. doi:10.1038/s41598-017-12939-7

Smit, A. F. A., Hubley, R., and Green, P. (2017). *1996–2010. RepeatMasker Open-3.0*.

Smith, G. P. (1976). Evolution of Repeated DNA Sequences by Unequal Crossover. *Science* 191 (4227), 528–535. doi:10.1126/science.1251186

Stephan, W. (1987). Quantitative Variation and Chromosomal Location of Satellite DNAs. *Genet. Res.* 50 (1), 41–52. doi:10.1017/S0016672300023326

Stephan, W. (1986). Recombination and the Evolution of Satellite DNA. *Genet. Res.* 47, 167–174. doi:10.1017/S0016672300023089

Stephan, W. (1989). Tandem-Repetitive Noncoding DNA: Forms and Forces. *Mol. Biol. Evol.* 6 (2), 198–212. doi:10.1093/oxfordjournals.molbev.a040542

Stornioli, J. H. F., Goes, C. A. G., Calegari, R. M., dos Santos, R. Z., Giglio, L. M., Foresti, F., et al. (2021). The B Chromosomes of *Prochilodus Lineatus* (Teleostei, Characiformes) Are Highly Enriched in Satellite DNAs. *Cells* 10 (6), 1527. doi:10.3390/cells10061527

Terán, G. E., Benitez, M. F., and Mirande, J. M. (2020). Opening the Trojan Horse: Phylogeny of *Astyanax*, Two New Genera and Resurrection of *Psalidodon* (Teleostei: Characidae). *Zool. J. Linn. Soc.* 190 (4), 1217–1234. doi:10.1093/zoolinnean/zlaa019

Utsunomia, R., Ruiz-Ruano, F. J., Silva, D. M. Z. A., Serrano, E. A., Rosa, I. F., Scudeler, P. E. S., et al. (2017). A Glimpse into the Satellite DNA Library in Characidae Fish (Teleostei, Characiformes). *Front. Genet.* 8, 103. doi:10.3389/fgene.2017.00103

Utsunomia, R., Silva, D. M. Z. A., Ruiz-Ruano, F. J., Goes, C. A. G., Melo, S., Ramos, L. P., et al. (2019). Satellitome Landscape Analysis of *Megaleporinus Macrocephalus*

(Teleostei, Anostomidae) Reveals Intense Accumulation of Satellite Sequences on the Heteromorphic Sex Chromosome. *Sci. Rep.* 9 (1), 1–10. doi:10.1038/s41598-019-42383-8

Vondrak, T., Ávila Robledillo, L., Novák, P., Koblížková, A., Neumann, P., and Macas, J. (2020). Characterization of Repeat Arrays in Ultra-Long Nanopore Reads Reveals Frequent Origin of Satellite DNA from Retrotransposon-Derived Tandem Repeats. *Plant J.* 101 (2), 484–500. doi:10.1111/tpj.14546

Walsh, J. B. (1987). Persistence of Tandem Arrays: Implications for Satellite and Simple-Sequence DNAs. *Genetics* 115 (3), 553–567. doi:10.1093/genetics/115.3.553

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 12 July 2022
doi: 10.3389/fgene.2022.879570

Comparing Transcriptomes Reveals Key Metabolic Mechanisms in Superior Growth Performance Nile Tilapia (*Oreochromis niloticus*)

*Binglin Chen†, Wei Xiao†, Zhiying Zou, Jinglin Zhu, Dayu Li, Jie Yu and Hong Yang**

Key Laboratory of Freshwater Fisheries and Germplasm Resources Utilization, Ministry of Agriculture and Rural Affairs, Freshwater Fisheries Research Center, Chinese Academy of Fishery Sciences, Wuxi, China

Metabolic capacity is intrinsic to growth performance. To investigate superior growth performance in Nile tilapia, three full-sib families were bred and compared at the biochemical and transcriptome levels to determine metabolic mechanisms involved in significant growth differences between individuals under the same culture environment and feeding regime. Biochemical analysis showed that individuals in the higher growth group had significantly higher total protein, total triglyceride, total cholesterol, and high- and low-density lipoproteins, but significantly lower glucose, as compared with individuals in the lower growth group. Comparative transcriptome analysis showed 536 differentially expressed genes (DEGs) were upregulated, and 622 DEGs were downregulated. These genes were significantly enriched in three key pathways: the tricarboxylic acid cycle (TCA cycle), fatty acid biosynthesis and metabolism, and cholesterol biosynthesis and metabolism. Conjoint analysis of these key pathways and the biochemical parameters suggests that Nile tilapia with superior growth performance have higher ability to consume energy substrates (e.g., glucose), as well as higher ability to biosynthesize fatty acids and cholesterol. Additionally, the fatty acids biosynthesized by the superior growth performance individuals were less active in the catabolic pathway overall, but were more active in the anabolic pathway, and might be used for triglyceride biosynthesis to store excess energy in the form of fat. Furthermore, the tilapia with superior growth performance had lower ability to convert cholesterol into bile acids, but higher ability to convert it into sterols. We discuss the molecular mechanisms of the three key metabolic pathways, map the pathways, and note key factors that may impact the growth of Nile tilapia. The results provide an important guide for the artificial selection and quality enhancement of superior growth performance in tilapia.

Keywords: tilapia, growth, energy metabolism, transcriptomes, substance metabolism

OPEN ACCESS

Edited by:
Tony Silveira,
Federal University of Rio Grande, Brazil

Reviewed by:
Agustin Barria,
University of Edinburgh,
United Kingdom
Bruna Félix Nornberg,
Federal University of Rio Grande, Brazil

****Correspondence:***
Hong Yang
yanghong@ffrc.cn

†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: *19 February 2022*
Accepted: *20 June 2022*
Published: *12 July 2022*

Citation:
Chen B, Xiao W, Zou Z, Zhu J, Li D, Yu J and Yang H (2022) Comparing Transcriptomes Reveals Key Metabolic Mechanisms in Superior Growth Performance Nile Tilapia (Oreochromis niloticus). Front. Genet. 13:879570. doi: 10.3389/fgene.2022.879570

INTRODUCTION

Nile tilapia (*Oreochromis niloticus*) is one of the most farmed fish worldwide, with 4.53 million tonnes produced in 2018 accounting for 8.3% of the world's total aquaculture production (The State of World Fisheries and Aquaculture 2020, 2020). Although there are established breeding models and management methods for this species, and it has long been selectively bred for growth

performance (Circa et al., 1995), large differences in growth still occur among individuals of the same genetic background and under the same rearing conditions (Chen et al., 2021), which effects their average breeding cycle and size, in turn making it difficult to control the costs of breeding and processing. Compared with higher-priced fish fillets, smaller tilapia was more often made into lower-priced fish meal, fish oil, or other byproducts, which has a great impact on the economic benefits of the tilapia industry. Simple phenotypes, such as body weight and activity, are the criteria still used by the majority of farmers to judge and establish quality breeding stocks, but, with the expansion of aquaculture and the increase in quality requirements, such criteria are no longer sufficient to meet farmers' requirements for selecting superior breeding stock (Neira et al., 2004; Rutten et al., 2005). Therefore, a better understanding of the molecular mechanisms of tilapia growth will provide more possibilities to develop this industry.

The growth of organisms is inextricably linked to metabolism as the ordered series of chemical reactions continuously needed to sustain life and enable the body to grow, reproduce, and maintain stability of its internal environment in response to environmental change (Enquist et al., 2003). Cell metabolism can be divided into pathways of substance metabolism and energy metabolism, which denote the ability of cells to exchange substances and transfer energy, and ultimately gives the organism its different phenotypic traits (DeBerardinis et al., 2008). Most studies have described key metabolic mechanisms that are vital to animal growth, but, owing to the complex metabolic network of an organism, elaboration of the relationship between metabolism and growth has largely focused on descriptions of signaling pathways (Yang et al., 2008; Mihaylova and Shaw, 2011; Wang M.-C. et al., 2022, Wang et al., 2022 M.; Yin et al., 2022). However, with the advancement of genomics and transcriptome analysis now provide feasibility for the exploration of key genes that determine key mechanisms in biological metabolism.

The development of a range of high-throughput sequencing techniques that includes transcriptome sequencing (RNA-seq) provides possibilities for relating an increasing number of phenotypic traits and molecular function. With the completion and updating of the tilapia reference genome, comparative RNA-seq-based analyses have revealed a variety of molecular mechanisms of tilapia, including: disease resistance (Wang et al., 2016), sexual differentiation (Tao et al., 2018), body color (Wang et al., 2018), and environmental adaptations (Liu et al., 2018), but the metabolic mechanisms of Nile tilapia with superior growth performance and the possible relationships to growth have been rarely reported.

Here, we report RNA-seq data obtained from full-sib families of Nile tilapia that showed significant growth differences under the same culture environment and feeding regime. We aimed to screen for differentially expressed genes (DEGs) between tilapia with higher average body weight (faster growth) and lower average body weight (slower growth), to identify key metabolic genes that may be associated with body weight gain, and ultimately to describe and explain the metabolic mechanisms that involve a range of key genes. The results provide insight into the genetic and molecular mechanisms potentially associated with effective weight gain in Nile tilapia with superior growth performance.

MATERIALS AND METHODS

Animal Breeding and Sampling

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the ethics committee of laboratory animal welfare and ethics of Freshwater Fisheries Research Center, Chinese Academy of Fishery Sciences (FFRC, CAFS), No. SYXK (SU) 2017-0007. Nile tilapia were obtained from three full-sib families (F1, F2, F3) from the tilapia genetic breeding base of the Freshwater Fisheries Research Centre of the Chinese Academy of Fishery Sciences. The tilapia used in the study was the GIFT population, which consisted of 60 families at the time of introduction. FFRC mixed all the families and randomly retained 5,000 offspring every 3 years but did not select for growth performance, in order to ensure that genetic diversity was not destroyed. The three tilapia pairs used to establish the full-sib families in this study were selected from the 4th generation (born in 2017) GIFT population. In May 2019, three pairs of 2-year-old Nile tilapia (♀:♂ = 1:1) were artificially inseminated and incubated to establish full-sib families, and all fertilized eggs were broken within one week and standardized for 2 weeks for fry breeding. Next, 500 juveniles of body weight (BW) 2.0 ± 0.5 g were randomly selected from each family and then released into a 35 m^2 pond for rearing; each family was raised in a separate pond with water temperature 28-32°C, about 13 h of natural daylight per day, pH of 7.0-8.0, and dissolved oxygen >5.0 mg/L. These juveniles were hand-fed a commercial feed of expanded (floating) pellets (30% crude protein, Nanjing ADM Animal Nutrition Co., Ltd., Nanjing, China), twice a day (at 8:00 and 18:00) until apparent satiation, for 3 months; satiation was determined by observing when feed remained after 60 min (uneaten food was then removed) according to Chen et al. (Chen et al., 2021).

To exclude the influence of sexual dimorphism in tilapia, the sex of the individuals in each family was identified and their growth traits were measured before tissue and blood sampling. Fifteen males with the highest growth performance (i.e., maximum weight) and 15 males with the lowest growth performance (minimum weight) were then selected in each family. Thus, a total of 90 extreme-growth individuals were obtained from the three families: 45 individuals with high growth performance were classified as the higher growth group (HG), and 45 individuals with low growth as the lower growth group (LG). These groups showed normal distribution ($p > 0.05$) and significant differences in BW, total length (TL), standard length (SL), head length (HL), body depth (BD), caudal peduncle length (CPL), caudal peduncle depth (CPD), and body width (BWD) ($p < 0.01$) (**Figure 1B**). In addition, we add weight gain (WG) [(Final body weight− Initial body weight)/Initial body weight], specific growth rate (SGR) {100×[ln(Final body weight)−ln(Initial body weight)]/No. of days}, average daily growth (ADG) [(Final body weight− Initial body weight)/No. of days], and condition factor (CF) ($100\times BW/BL^3$) to provide a more

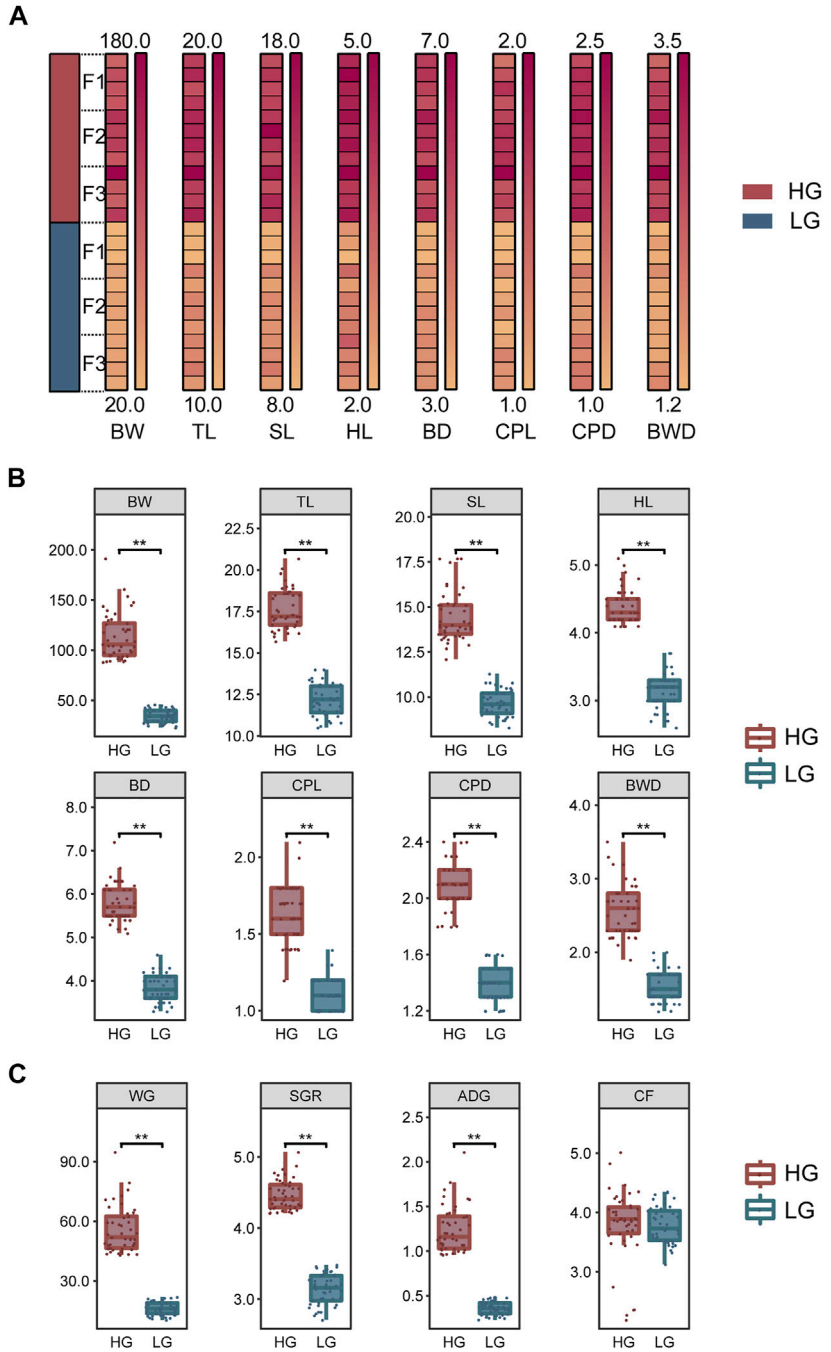


FIGURE 1 | Growth traits and analysis of significant differences among samples. **(A)** Differences in growth traits of the Nile tilapia used for transcriptome sequencing. **(B)** Significant differences in growth traits between the higher growth (HG) and lower growth (LG) groups of tilapia (n = 45, each group). **(C)** Differences in weight gain (WG), specific growth rate (SGR), average daily growth (ADG), and condition factor (CF) between the HG and LG groups (n = 45, each group). Levene's test was used to assess the equality of variances in the data, and the independent samples t-test showed a normal distribution ($p > 0.05$). $^{**}p < 0.01$.

comprehensive description of growth according to Du and Turchini (2022) (**Figure 1C**). All growth data were shown in **Supplementary Table S1**. Four individuals were randomly selected from the HG and LG in each family (i.e., a total of 12 HG individuals and 12 LG individuals from three families) for transcriptomic sequencing and subsequent biochemical parameters and qRT-PCR analysis; The mean BW (Mean ± SD) of the individuals used for sampling was 125.4 ± 25.2 g and 35.9 ± 7.6 g for HG and LG, respectively (**Figure 1A**). These 24 fish were anesthetized with Tricaine (MS-222) at a concentration of 13.5 g/m^3 in 25°C water, to exclude the effect of stress on the experiment. Liver tissue was collected and stored

at −80°C for RNA extraction and transcriptome analysis. Blood was collected (>500 μL) and centrifuged (12,000 g, 5 min, 4°C), and the supernatant was stored at −80°C for the biochemical analysis.

Biochemical Assays

The 24 samples (12 from HG and 12 from LG) used for transcriptome analysis were continued to be used for blood biochemical analysis to ensure the accuracy and consistency of the experiment. The activities of total protein (TP), total triglyceride (TG), total cholesterol (TC), glucose (Glu), high-density lipoprotein (HDL), and low-density lipoprotein (LDL) were measured using commercial kits manufactured by Mindray Biomedical Electronics Co., Ltd. (Shenzhen, China), the volumes used for parameter tested was consistent between each sample. All kits were passed the quality control of the manufacturer. The coefficient of variation between replicates and the relative deviation between batches are less than 3 and 5%, respectively. And these kits have been shown to be widely used for the determination of fish biochemical parameters (Xiao et al., 2022).

RNA Extraction, Library Preparation, and Sequencing

For transcriptome analysis, each group contained three samples (higher growth group: HG1, HG2, and HG3; lower growth group: LG1, LG2, and LG3), and each sample mixed liver tissue from four individuals showing the same growth performance and from the same family. Total RNA in liver tissue was isolated with TRIzol reagent (Invitrogen, Carlsbad, United States), following the manufacturer's instructions, checked using 1% agarose gel electrophoresis, and its concentration and purity determined with a NanoDrop™ Lite spectrophotometer (Thermo Fisher Scientific, Waltham, United States). The RNA-seq transcriptome library was prepared using a TruSeq RNA Sample Prep Kit (Illumina, San Diego, United States) with 1 μg of total RNA. The libraries were sequenced on an Illumina HiSeq 4000 platform (Illumina, San Diego, United States).

Identification and Functional Annotation of DEGs

Before data analysis, quality control of raw sequencing data using FASTP software (Chen et al., 2018). To obtain high-quality data, we removed adapter sequences, sequences of unknown nucleotides (>10%), and low-quality reads (Q-value ≤ 20). The clean reads were mapped to the reference genome (NCBI: GCF_001858045.2, *Oreochromis niloticus*) (Conte et al., 2017) using HISAT2.2.4 (Kim et al., 2015). To quantify the expression abundance and variation of each transcription region, we used RSEM software to calculate transcripts per million (TPM) (Li and Dewey, 2011). The differentially expressed genes (DEGs) between two groups were analyzed using DESeq2 software (Love et al., 2014). Gene expression with a false discovery rate (FDR) of <0.05 and fold change (FC) of >2 or <0.5 was considered to be a significant difference. To recognize the main biological functions of DEGs, we applied Gene Ontology (GO) enrichment analysis with GOseq software (Young et al., 2010); to identify the enrichment signaling pathways of DEGs, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was applied using KOBAS 2.0 software (Xie et al., 2011). FDR ≤0.05 was taken as the threshold for GO and KEGG enrichment. The RNA-Seq data has been submitted in Sequence Read Archive (SRA) database (No. PRJNA787719).

Correlation Network Analysis

Growth performance, DEGs with research potential (29 DEGs that were significantly enriched in three major metabolic pathways: TCA cycle, Fatty acid biosynthesis and metabolism, and cholesterol biosynthesis and metabolism), and biochemical parameters were included in correlation analysis. Significant correlations ($p < 0.01$) with Pearson's correlation coefficient more than 0.80 were used for network visualization in Gephi 0.9.2 (Lin et al., 2020).

Quantitative Real-Time PCR Analysis

The 24 samples (12 from HG and 12 from LG) used for transcriptome analysis were continued to be used for qRT-PCR analysis to ensure the accuracy and consistency of the experiment. We performed qRT-PCR on DEGs that are significantly enriched in three major metabolic pathways: TCA cycle, Fatty acid biosynthesis and metabolism, and cholesterol biosynthesis and metabolism, to verify their relative expression trends. First-strand cDNA for each sample was synthesized with equal amounts of 900 ng of total RNA (The RNA extraction method used for gene expression is the same as described in 2.3) using a PrimeScript™ RT Reagent Kit (TaKaRa, Dalian, China); mRNA expression levels of DEGs in the liver tissue of the different groups were measured using quantitative real-time PCR (qRT-PCR) in a 7900HT Fast Real-Time PCR System (Applied Biosystems, Waltham, United States). The qRT-PCR was performed using SYBR® Green qPCR Master Mix (Vazyme, Nanjing, China) in a 20 μL total reaction volume containing 10 μL of SYBR® Green qPCR Master Mix, 1 μL of cDNA template, 0.4 μL of each primer (10 μM), and 8.2 μL of sterile water. The qRT-PCR program began with an initial denaturation at 95°C for 30 s, followed by 40 cycles of 95°C for 10 s, 60°C for 30 s, and a melting curve at 95°C for 30 s, 60°C for 60 s, and 95°C for 15 s, three replicates were set for each sample. Both *actin beta* (*actb*) and *ubiquitin-conjugating enzyme* (*ubce*) genes were used as reference genes following the studies used in Nile tilapia by Deloffre et al. (2012), Yang et al. (2013), and Chen et al. (2020). The qRT-PCR specific primers were designed by Primer Premier 5 and are listed in **Supplementary Table S2**. A pretest of each primer pair is required to obtain a standard curve before the experiment. The standard curves of all genes were constructed using tilapia liver cDNA. The standard was

diluted in five gradients as 1×, 5×, 25×, 125×, and 625×, and three replicates were set for each gradient, the system and reaction conditions are the same as above. Multiple primer pairs for each target gene were designed, and primers with R^2 between 0.99 and 1.00 was selected for the transcriptome validation experiments. The relative mRNA levels of target genes were calculated using the $2^{-\triangle\triangle Ct}$ method.

Statistical Analysis

SPSS 24.0 software (Hansen, 2005) was used for Levene's Test for Equality of Variances and Independent Samples *t*-test (Sedgwick, 2010) for growth traits ($n = 45$), biochemical parameters ($n = 12$), and qRT-PCR ($n = 12$) comparison between HG and LG in this study, results are presented as box plots providing sample distribution to clearly demonstrate the distribution of results and differences were considered significant at $p < 0.01$. To describe the clustering among sequenced samples and the differences in DEGs, M-versus-A plot (MA plot) were using ggpubr package, Principal component analysis (PCA) were using ggplot 2 package (Ginestet, 2011) and clusters samples based on their expression in all annotated genes according to Ren et al. (Ren et al., 2022), the above packages are based on R software (Version 4.1.0).

RESULTS

Growth Comparison of HG and LG Groups

Growth comparison analysis using the 90 samples of extreme-growth individuals (45 from each the HG and LG groups) revealed significant differences in all growth traits ($p < 0.01$) (**Figure 1B**). In addition, significant differences were shown in WG, SGR, and ADG between the HG and LG fish ($p < 0.01$). However, there was no significant difference in CF ($p > 0.05$) (**Figure 1C**).

Comparison of Biochemical Parameters Under Growth Differences

The TP, TC, TG, HDL, and LDL contents were significantly higher, but Glu was significantly lower, in HG individuals compared with LG individuals ($p < 0.01$) (**Figure 2**).

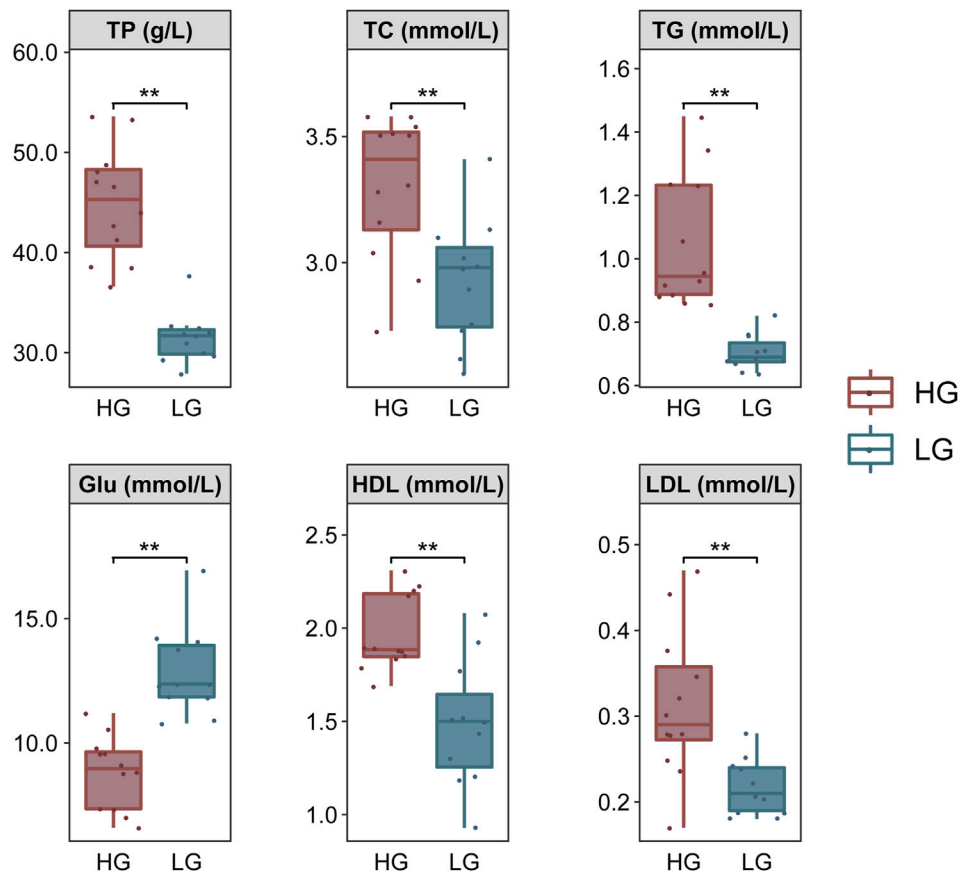


FIGURE 2 | Comparison of biochemical parameters of Nile tilapia, grouped according to extreme growth differences ($n = 12$). Levene's test was used to assess the equality of variances in the data, and the independent samples *t*-test showed a normal distribution ($p > 0.05$). $^{**}p < 0.01$.

TABLE 1 | Statistics of total reads in RNA-Seq of higher growth group (HG) and lower growth group (LG) in the evaluation of superior growth performance in Nile tilapia (*Oreochromis niloticus*).

Sample	Raw Reads	Clean Reads	Q20 (%)	Q30 (%)	GC Content (%)
HG1	45,402,242	44,991,790	98.17	94.53	49.37
HG2	47,032,260	46,608,926	98.24	94.69	49.68
HG3	51,795,966	51,295,092	98.17	94.51	49.58
LG1	47,661,686	47,131,572	97.57	93.10	49.43
LG2	49,847,744	49,355,034	97.96	94.01	49.55
LG3	46,985,888	46,504,688	98.11	94.36	49.77

Transcriptome Sequencing Quality Assessment and Identification of DEGs

After filtering of raw data, the RNA sequence generated 44,991,790 to 51,295,092 clean reads in the liver of HG and LG. The Q20 and Q30 values for each sample were greater than 97.57 and 93.10%, respectively, and the average GC content was 48.98%. The alignment statistics results showed that the ratio of the mapped reads was approximately 91.92% when compared with the reference genome of Nile tilapia (**Table 1**). A total of 34,893 genes were annotated and used for DEG analysis, and a total of 1,158 DEGs were identified, including 536 upregulated

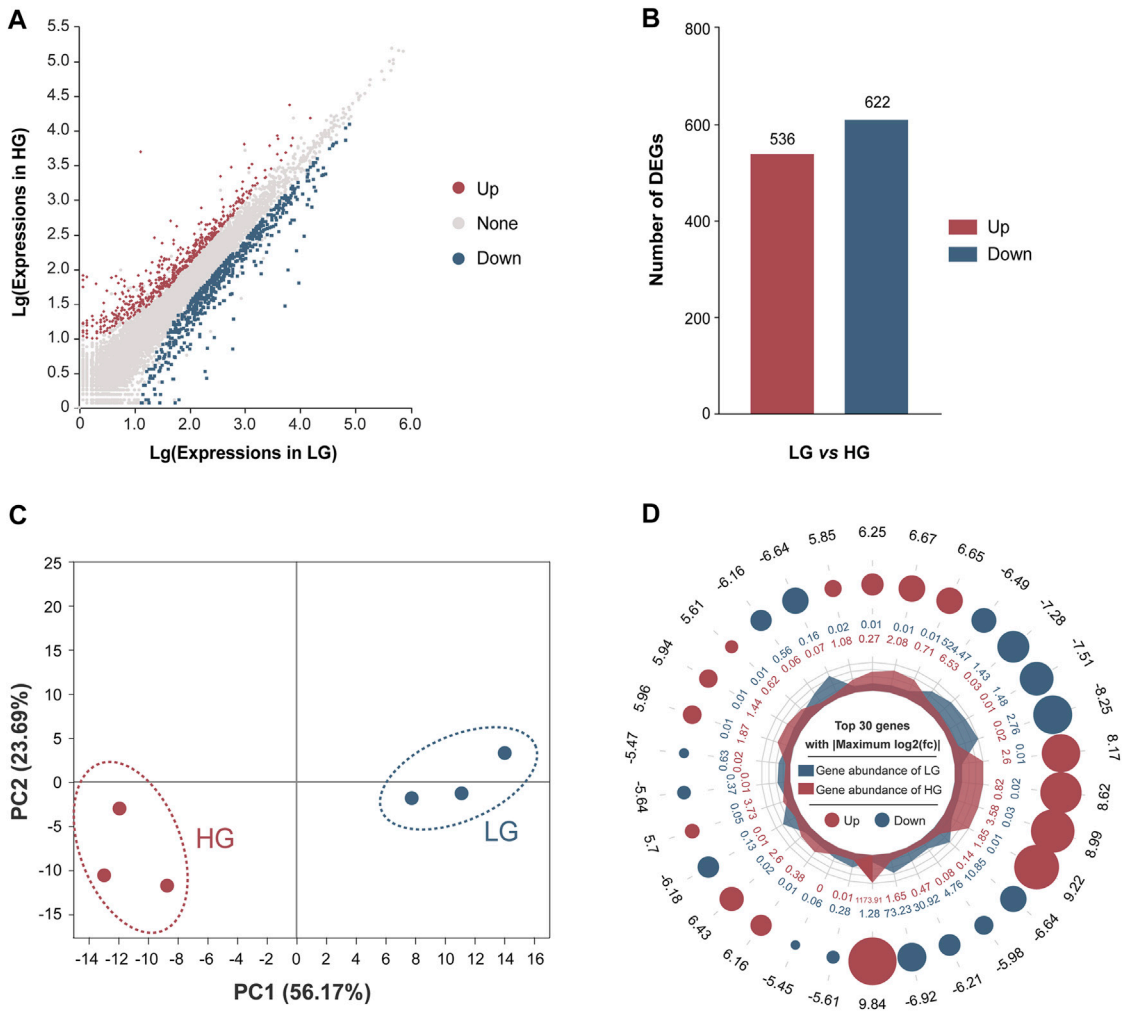


FIGURE 3 | Comparison of differentially expressed genes (DEGs) between the higher growth (HG) and lower growth (LG) groups of Nile tilapia, with the LG group as the control. **(A)** M-versus-A plot (MA plot) of DEGs of the HG versus the LG group. **(B)** Numbers of DEGs between the LG and HG groups, with LG as the control group and HG as the comparison group, with $\log_2$ fold change (FC) of ≥2 or ≤0.5 ($p < 0.01$). **(C)** Principal components analysis of the HG and LG groups based on expression of all annotated genes. **(D)** Ploidy radar plot of the top-30 significantly different genes between the HG and LG groups. The numbers in the outermost circle are the differential multiples of genes; red and blue dots represent upregulation and downregulation, respectively; a larger circle diameter represents a larger absolute value of the multiples; the inner red and blue numbers represent the gene abundance within the HG and LG groups, respectively.

genes and 622 downregulated genes (with LG as the control group and HG as the comparison group, FC ≥ 2 or ≤0.5; $p < 0.01$) (**Figures 3A,B**). Results of principal component analysis showed that the samples were clustered into two different groups according to high growth and low growth, which was consistent with the grouping of the samples (**Figure 3C**). Radar mapping showed that the multiplicity of the top-30 significantly upregulated or downregulated genes between the two groups ranged from 5.45 to 9.84 (**Figure 3D**). These results indicated that the data from these samples qualified for the DEG analysis.

Functional Analysis Using GO and KEGG Enrichment

To better understand the metabolic mechanisms in individuals that showed significant growth differences under the same culture environment and feeding regime, we performed GO enrichment analysis for the three main categories: molecular functions (MF), cellular components (CC), and biological processes (BP). The upregulated DEGs were mainly enriched in "organic substance metabolic process", "small molecule metabolic process" of BP, "membrane-bounded organelle", "intracellular membrane-bounded organelle", "nucleus" of CC, and "isomerase activity", and 'steroid hormone receptor activity" of MF. The downregulated DEGs were mainly enriched in "iron ion binding", "tetrapyrrole binding", "heme binding" of MF, "collagen trimer" of CC and "alcohol metabolic process", and "organic hydroxy compound metabolic process" of BP (**Figure 4A**). The KEGG annotation results showed that DEGs were concentrated in 32 pathways involving Cellular Processes, Environmental Information Processing, Genetic Information Processing, Metabolism, and Organismal Systems (**Figure 4B**).

To further elucidate the function of significant DEGs in signaling pathways, we annotated the DEGs significantly upregulated and downregulated in HG in the KEGG database to analyze the significantly enriched KEGG pathway. The results showed that three metabolism-related pathways were among the 10 pathways significantly enriched in upregulated DEGs (**Figures 5A,C**), namely Pyruvate metabolism, Citrate cycle (TCA cycle), Steroid biosynthesis, and Fatty acid biosynthesis. Cell cycle and DNA replication, two pathways associated with amplification of genetic information, were also defined as significantly enriched. Downregulated DEGs were significantly enriched in two pathways (**Figures 5B,C**), namely Primary bile acid biosynthesis and AMPK signaling (which is significantly associated with metabolic regulation). The above results suggest that the significantly enriched KEGG signaling pathway plays an important role in the metabolic system, and might be concentrated in three metabolic pathways: TCA cycle, Fatty acid biosynthesis and metabolism, and cholesterol biosynthesis and metabolism.

Key Factors of the TCA Cycle Signaling Pathway in Superior Growth Performance Tilapia

Four genes in the upregulated DEGs of HG individuals were annotated to the TCA cycle signaling pathway, namely: *phda*

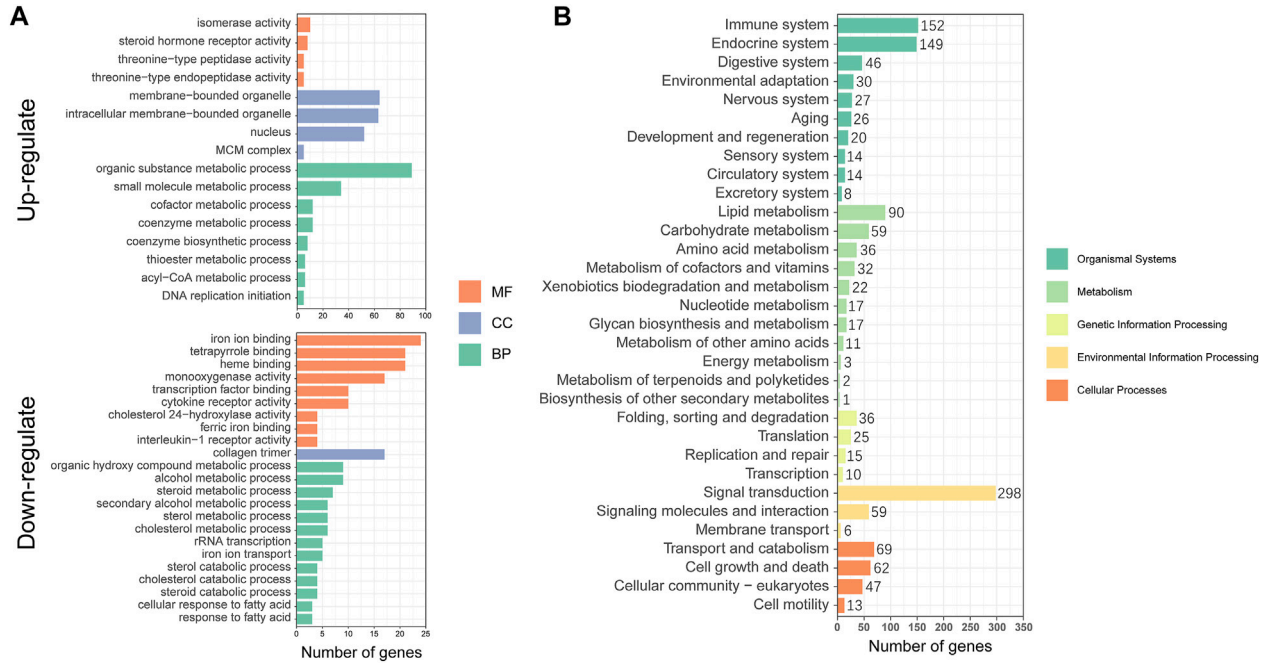


FIGURE 4 | Functional annotation and enrichment analysis of differentially expressed genes (DEGs): **(A)** Gene Ontology (GO) functional enrichment analysis; **(B)** Kyoto Encyclopedia of Genes and Genomes (KEGG) functional annotation analysis.

and *pdhb*, encoding the pyruvate dehydrogenase E1 component, which promotes acetyl-CoA biosynthesis upstream of the TCA cycle; *idh*, encoding isocitrate dehydrogenase; and *mdh* encoding malate dehydrogenase, which is involved in the TCA cycle. The reactions involved in all three of these enzymes promote the production of nicotinamide adenine dinucleotide (NADH), which accounts for 75% of the total number of NADH synthesis pathways in the TCA cycle (**Figure 6**). These results indicate that genes encoding key enzymes in pathways involved in energy metabolism (NADH-ATP) were expressed as upregulated in Nile tilapia with superior growth performance.

Key Factors of the Fatty Acid Biosynthesis and Metabolism Pathway in Superior Growth Performance Tilapia

There were 11 significant DEGs annotated to the Fatty acid biosynthesis and metabolism pathway in HG individuals, of which the upregulated DEGs were ATP citrate lyase (*acly*), acetyl-CoA carboxylase (*acc*), fatty acid synthase (*fasn*), long-chain acyl-coenzyme A synthetase (*acsls*), acyl-coenzyme A thioesterase (*acot1*), and sterol-regulatory element-binding proteins 1 (*srebp1*); the proteins encoded by these genes are mainly used to promote fatty acid biosynthesis and to maintain the stability of free fatty acids. Downregulated DEGs were carnitine palmitoyl transferase 1a (*cpt1a*), hormone-sensitive lipase (*lipe*), NAD^+-dependent protein deacetylase sirtuin 1 (*sirt1*), very long-chain acyl-CoA dehydrogenase (*acadvl*), and forkhead box O protein (*foxos*). The functions of the proteins encoded by these genes are different, as considered in the Discussion (**Figure 7**).

Key Factors of the Cholesterol Biosynthesis and Metabolism Pathway in Superior Growth Performance Tilapia

There were 10 significant DEGs annotated to the cholesterol biosynthesis and metabolism pathway in HG individuals, including the upregulated genes acetyl-CoA acetyltransferase 2 (*acat2*), isopentenyl-diphosphate delta isomerase 1 (*idi1*), lanosterol synthase (*erg7*), cholesterol delta-isomerase (*ebp*), and 7-dehydrocholesterol reductase (*dhcr7*), which encode proteins or enzymes that are primarily used to promote cholesterol biosynthesis. The proteins or enzymes encoded by the downregulated genes, including hydroxysteroid dehydrogenase type 12 (*hsd17b12*), steroid 5β-reductases (*akr1d1*), and UDP-glucuronosyltransferase (*ugt*) are mainly used to promote steroid hormone biosynthesis, and the enzymes encoded by sterol 12alpha-hydroxylase (*cyp8b1*) and

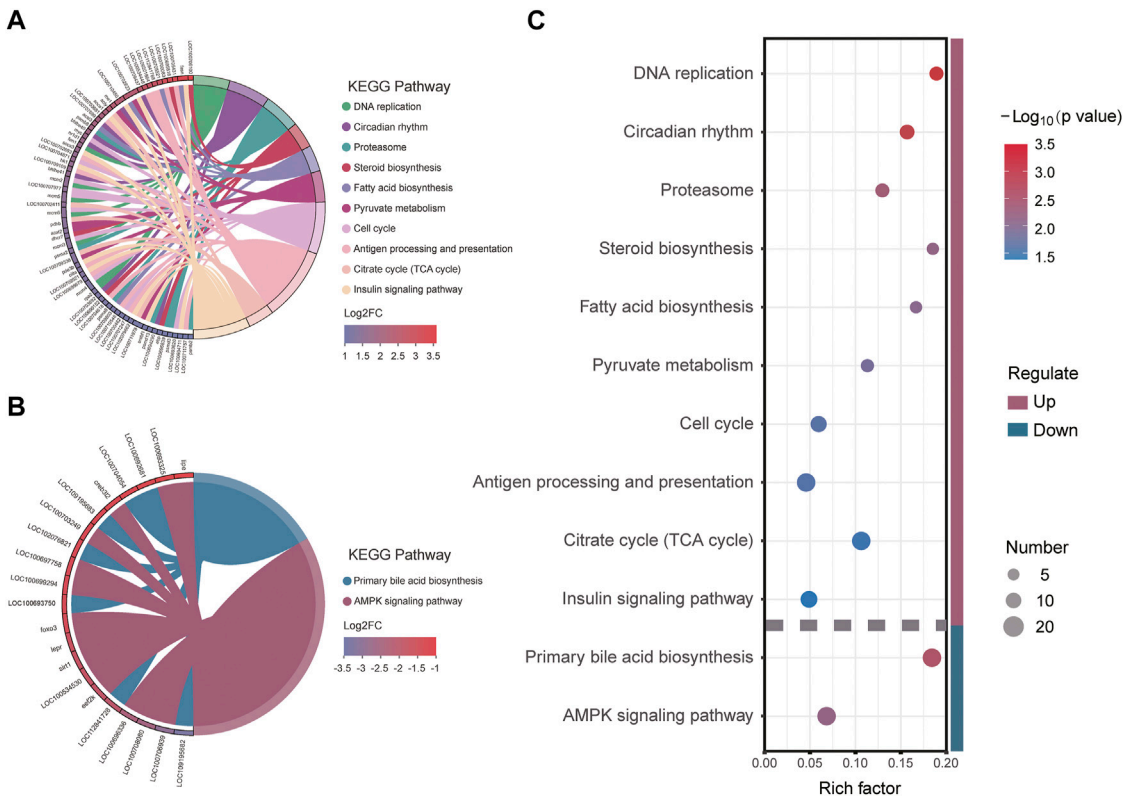


FIGURE 5 | Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment and its major enrichment genes: **(A)** upregulated differentially expressed genes (DEGs) significantly enriched; **(B)** downregulated DEGs significantly enriched; **(C)** significantly enriched KEGG pathways and the Rich factors.

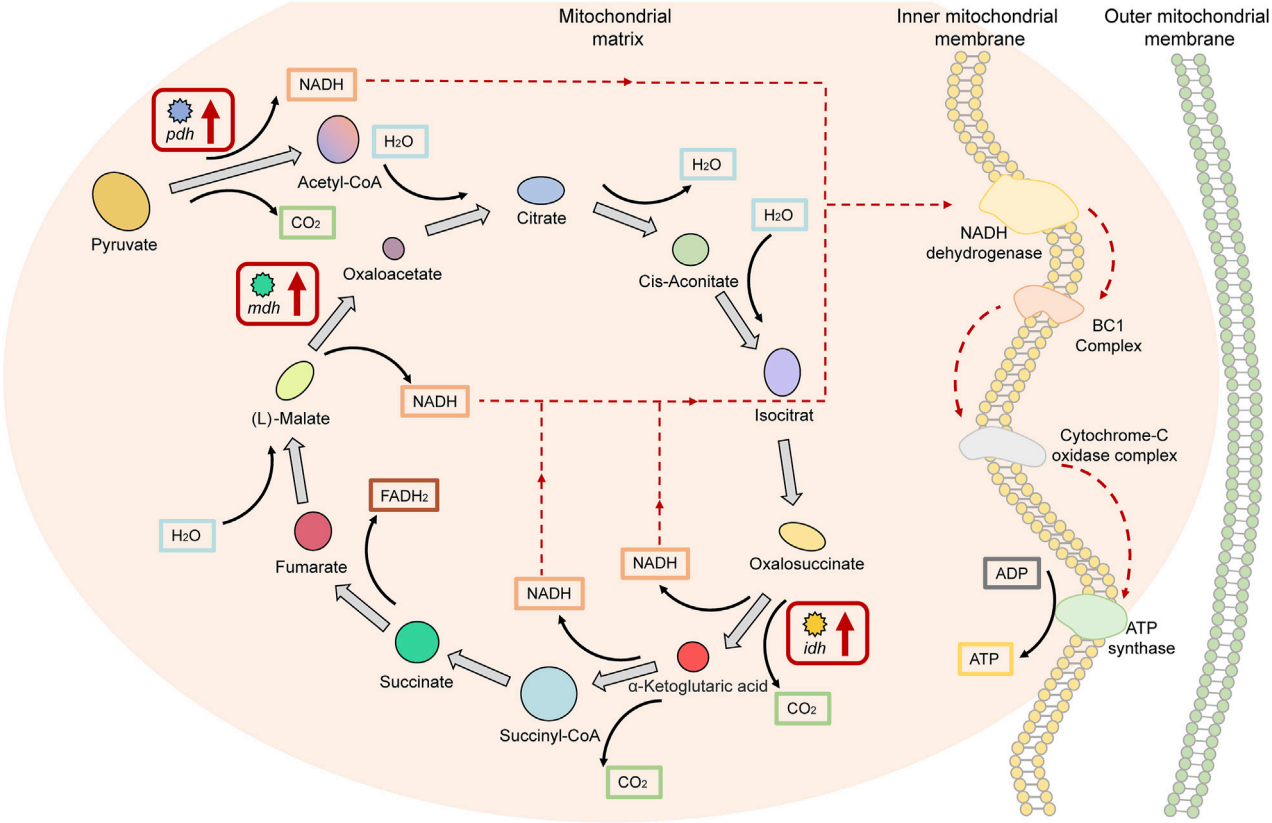


FIGURE 6 | Key factors of the "tricarboxylic acid cycle (TCA cycle) signaling pathway" in superior growth performance Nile tilapia. Red boxes denote upregulated genes. Genes encoding key enzymes (*phd*, *idh*, and *mdh*) in pathways involved in energy metabolism (NADH-ATP) were upregulated.

cholesterol-24S-hydroxylase (*cyp46a1*) are mainly used to promote the biosynthesis of bile acids (**Figure 8**).

Correlation Analysis of Growth Performance, DEGs, and Biochemical Parameters

Correlations between growth performance, DEGs, and biochemical parameters were further investigated. *mdh1* in the TCA cycle signaling pathway, *acly* and *acot1* in the Fatty acid biosynthesis and metabolism pathway, and *idi1*, *ugt*, *dhcr7*, *cyp2r1*, and *cyp46a1* in the cholesterol biosynthesis and metabolism pathway showed significant positive correlations with growth performance, and the correlations were consistent with the results at the transcriptional level, indicating that the expression levels of these genes increased with growth performance of the tilapia. In contrast, *sirt1*, *foxo1*, *foxo3*, *lipe*, *acadvl*, and *cyp46a1* were negatively correlated with growth performance, indicating that the expression levels of these genes decreased with an increase in the growth performance of the tilapia. In the correlation analysis with Glu, the expression of *idh* and *mdh1* showed significant negative correlation with the concentration of Glu, indicating that as the expression of those genes increases, the concentration of Glu decreases, which is consistent with the function of these genes (consumption of Glu in energy metabolism). In the correlation analysis of TC and TG, the correlation trends of the genes were consistent with the transcriptional results, indicating that the above genes significantly influenced lipid metabolism and sterol metabolism in the tilapia. Notably, *srebp1* and *acsls* showed significant positive correlations with both TC and TG, while *cyp8b1* showed significant negative correlations with both TC and TG, which may indicate that these genes play an important role in the metabolic process (**Figure 9**).

Quantitative Real-Time PCR Analysis

The qRT-PCR results showed that the relative expression of each key gene was in line with the trend in the TPM values. Key genes in all three signaling pathways (TCA cycle, Fatty acid biosynthesis and metabolism, and cholesterol biosynthesis and metabolism) showed significant differences between HG and LG tilapia ($p < 0.01$) (**Figure 10**).

DISCUSSION

Recent research on tilapia growth performance has focused on feed and additives (Zhou et al., 2010; Genschick et al., 2021; Yu

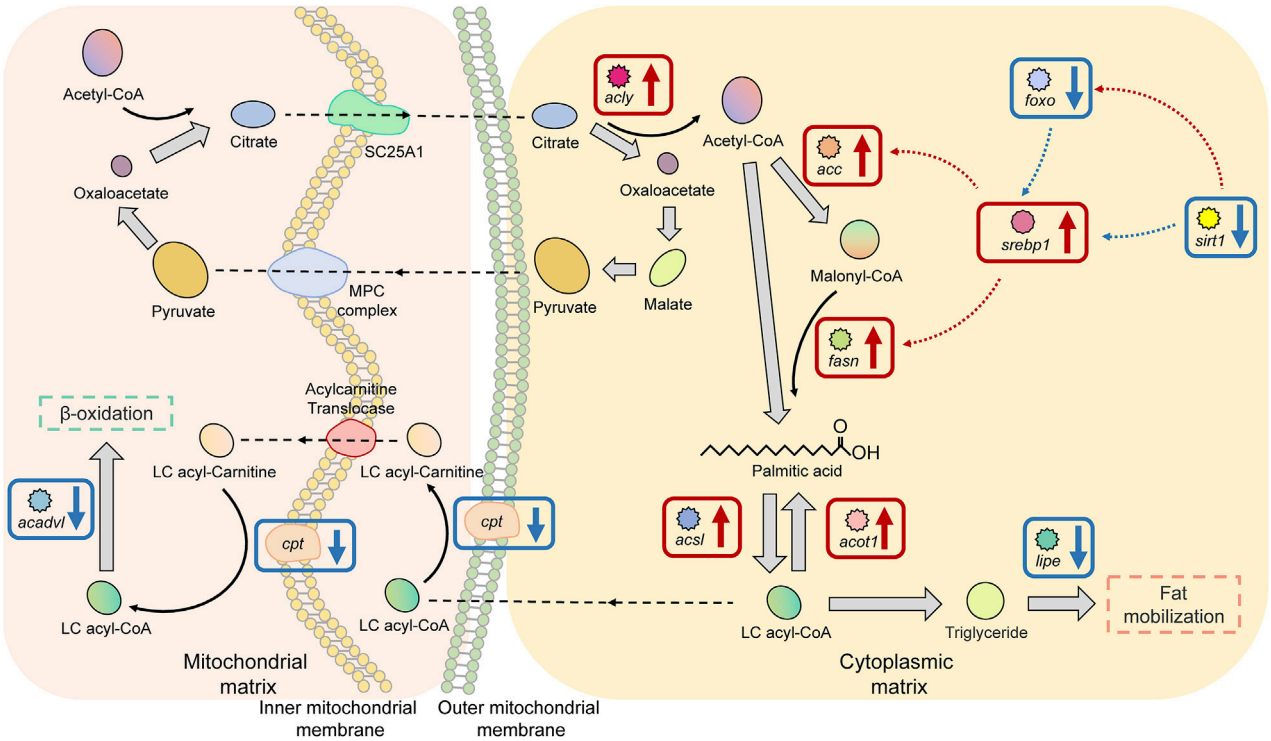


FIGURE 7 | Key factors of the "fatty acid biosynthesis and metabolism pathway" in superior growth performance Nile tilapia. Red boxes mark upregulated genes, and blue boxes mark downregulated genes. Genes encoding key enzymes involved in the *ß*-oxidation pathway (*cpt*, *acadvl*) were downregulated, and genes encoding key enzymes involved in fatty acid biosynthesis and activation (*acly*, *acc*, *srebp1*, *fasn*, *acsl*, and *acot1*) were upregulated.

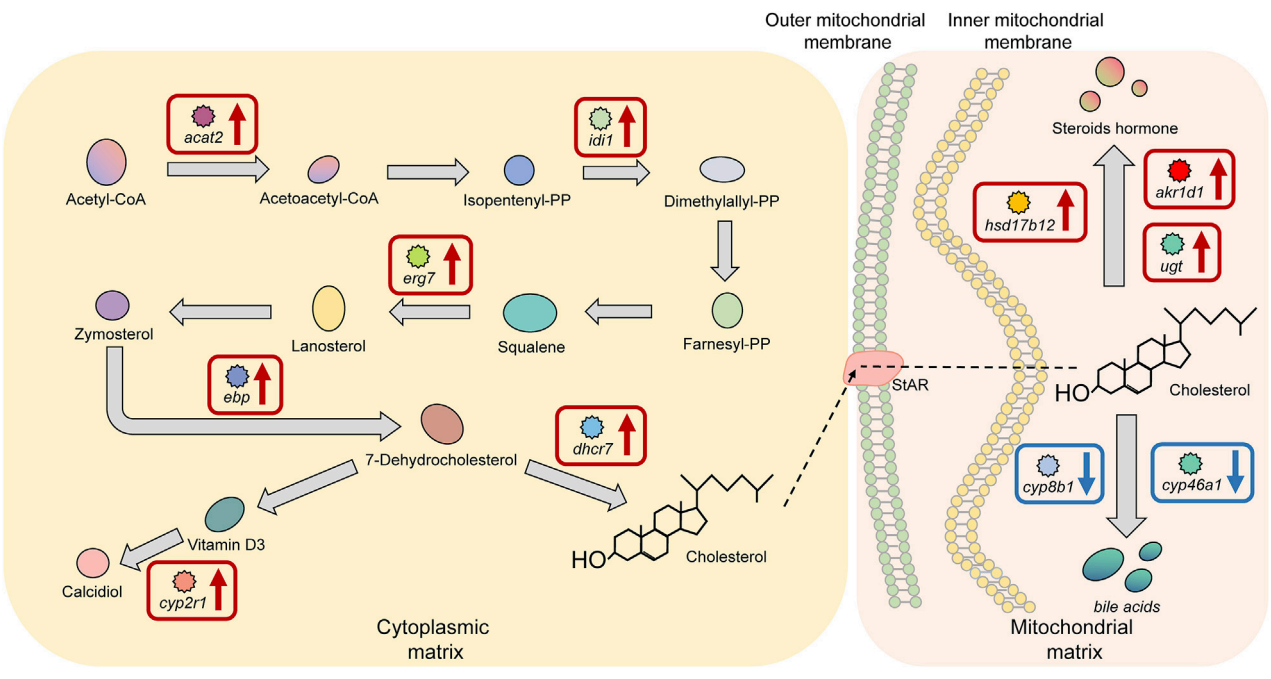


FIGURE 8 | Key factors of the "cholesterol biosynthesis and metabolism pathway" in superior growth performance Nile tilapia. Red boxes mark upregulated genes, and blue boxes mark downregulated genes. Genes encoding key enzymes involved in synthesis of bile acids from cholesterol (*cyp8b1*, *cyp46a1*) were downregulated, and genes encoding key enzymes involved in synthesis of steroid hormones from cholesterol (*hsd17b12*, *akr1d1*, *ugt*) were upregulated.

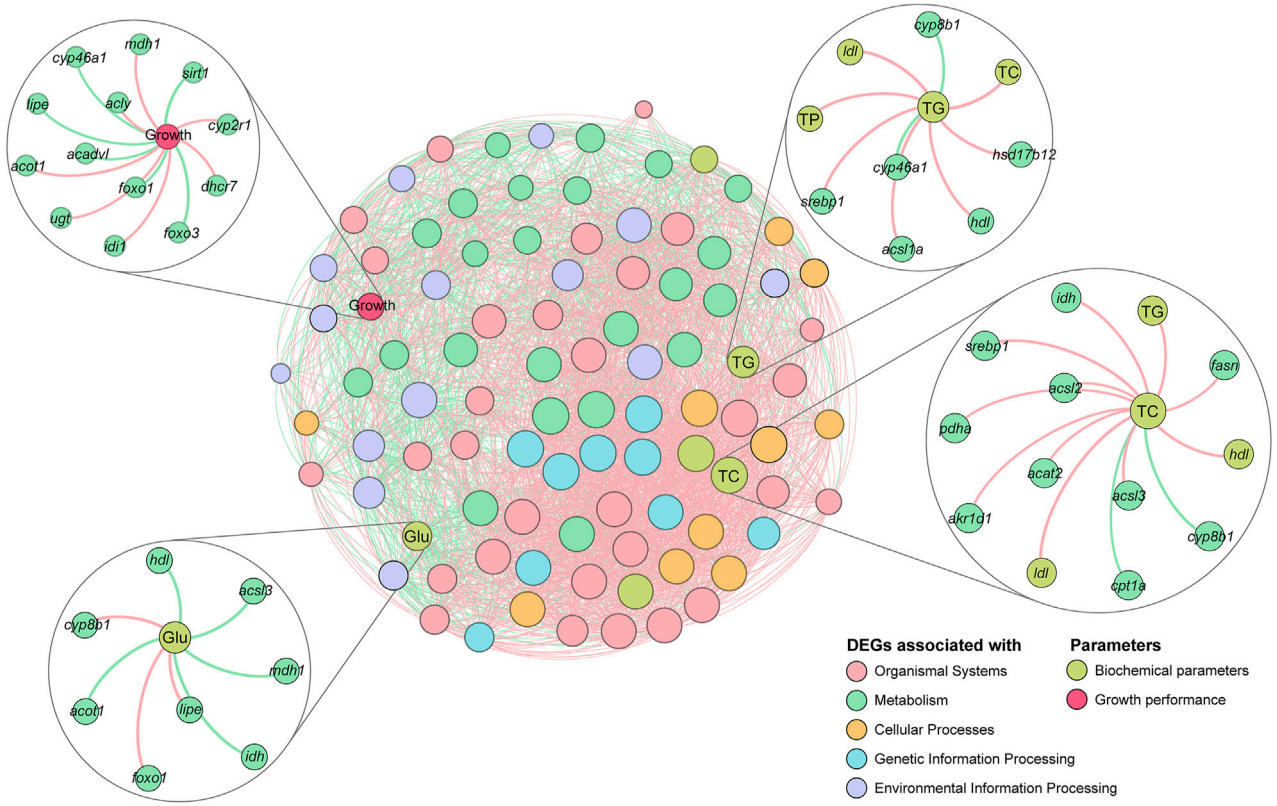


FIGURE 9 | Correlation network analysis of tilapia growth performance, differentially expressed genes (DEGs), and biochemical parameters. The colors of the nodes indicate data types, and the sizes of the nodes indicate the degree in the network. Positive and negative correlations are represented by red and green lines, respectively.

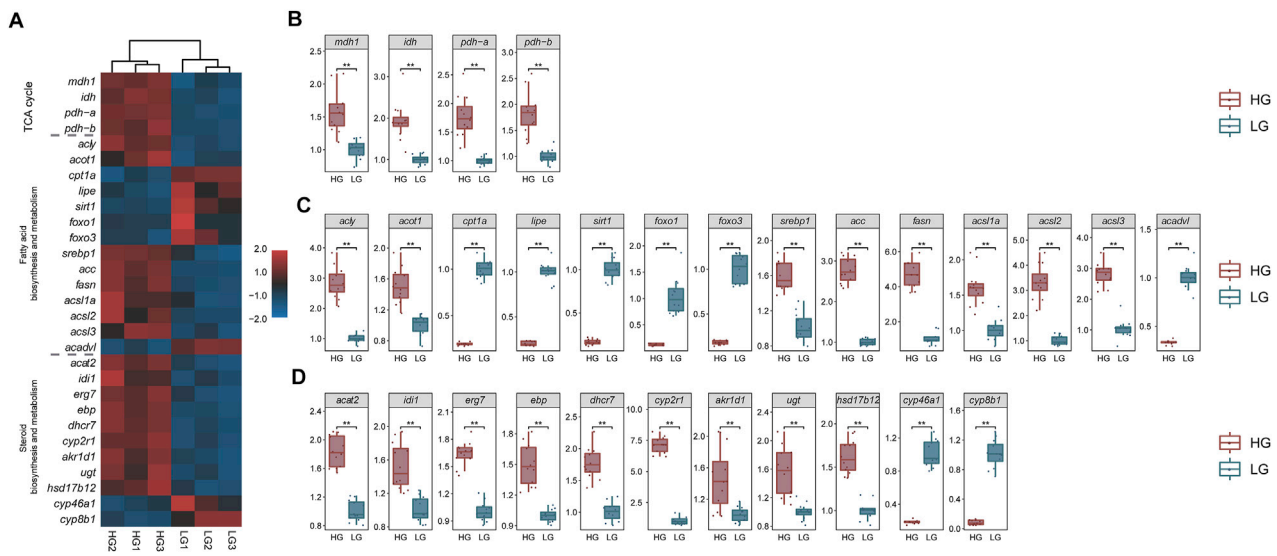


FIGURE 10 | Transcriptome transcripts per million (TPM) values ($n = 3$) and qRT-PCR results of key genes ($n = 12$). **(A)** Heat map of transcriptome TPM values of key genes; mRNA levels of key genes in **(B)** the "TCA cycle pathway", **(C)** the "fatty acid biosynthesis and metabolism pathway", and **(D)** the "cholesterol biosynthesis and metabolism pathway". $^{**}p < 0.01$.

et al., 2021), managing the environment (Charo-Karisa et al., 2006; Turra et al., 2016), culturing models (Zhang et al., 2018; Ani et al., 2021), and growth-related molecular markers (Khatab et al., 2014; Chen et al., 2020). Among these, the more relevant study for growth performance enhancement was the screening of growth-related QTLs in tilapia (Yoshida and Yáñez, 2021), but the 29 metabolism-related candidate genes obtained in the present study were not among them. These different outcomes may be attributable to the exploration levels: identification of the QTLs was on the genomic level in the work of Yoshida and Yáñez (2021), while in our study it was based on the transcriptome level. In addition, the candidate genes obtained from growth-related QTLs were obtained through correlation analysis, whereas the transcriptome results were obtained based on analysis of significant differences. No studies to date have reported on tilapia with the same genetic background but with significant growth differences under the same culture environment and feeding method. In wild-type aquatic animals, gene expression between different growth performance groups has been reported for vannamei shrimp (*Litopenaeus vannamei*) (Santos et al., 2021) and California red abalone (*Haliotis rufescens*) (Valenzuela-Miranda et al., 2015). To some extent, these studies have revealed the relationship between growth performance differences and genes in aquatic animals; however, the available research is insufficient to elucidate the mechanisms of transcriptome-level regulation of growth differences in wild-type bony fishes, especially cichliforms. In a study of non-wild type coho salmon (*Oncorhynchus keta*), McClelland et al. (2020) found that comparative transcriptome results between GH transgenic (T) and non-transgenic (NT) individuals with different growth performance (T_{large}, T_{small}, NT_{large}, and NT_{small}) indicated that there are widespread regulatory influences acting to influence body size and gene expression traits, in addition to the effects of GH transgenesis. Those authors also showed that multiple regulatory loci affecting gene expression were shared between fast-growing and slow-growing fish within T or NT groups, but no such regulatory loci were found to be shared between those two groups (McClelland et al., 2020). This demonstrates the complexity and important research potential of gene expression differences among individuals with different growth performance.

Without a reference genome, DEGs can only be compared by splicing into unigenes or by annotation to databases, such as Non-Redundant Protein Sequence Database (NR), GO, and KEGG, which drastically reduces the number of actual DEGs (Zhang et al., 2019). This implicates the importance of a species' reference genome for comparative transcriptomic analyses. We annotated RNA-seq data into the tilapia reference genome and identified 1,158 DEGs; this availability increases the significance and representativeness of our results. The influence of enzymes (i.e., changes in the activity of the enzyme itself) on the metabolic pathway depends on the changes in the activity of the entire metabolic pathway (Rose, 1968; Bell, 2012). For example, an enzyme with a significant change in activity cannot affect the entire metabolic pathway if the change does not affect other processes in the pathway; conversely, if most enzymes in a metabolic pathway show significant changes in activity, then the entire metabolic pathway is affected (either stimulated or inhibited). Our results show that the significantly upregulated and downregulated DEGs were enriched within the three metabolic pathways, and the expression levels of a proportion of genes were significantly changed within each pathway (**Figures 6–8**). This result is consistent with the above view and further strengthens our speculation that these metabolism-related enzymes or proteins play a key role in body weight gain in Nile tilapia while constituting potential candidate genes at the transcriptional level. Therefore, in the following sections we discuss the three key metabolic pathways in detail, namely the TCA cycle, Fatty acid biosynthesis and metabolism, and cholesterol biosynthesis and metabolism.

Mechanisms of the TCA Cycle Signaling Pathway in Tilapia With Superior Growth Performance

The TCA cycle is the main process for the metabolism of sugars, lipids, and amino acids in the mitochondria and is the metabolic hub in mast eukaryotes (Ryan and O'Neill, 2020). The TCA cycle starts with the oxidative decarboxylation of pyruvate; the resulting acetyl-CoA is regenerated to oxaloacetate and CO_2 after four dehydrogenation reactions; and the whole process takes place within a cycle of eight main reactions (Krebs and Johnson, 1980). NADH and flavin adenine dinucleotide (FADH), which are produced during the TCA cycle, put it at the center of energy metabolism. NADH and FADH in the inner mitochondrial membrane transfer electrons to O_2 *via* an electron transport chain containing NADH dehydrogenase, BC1 complex, and cytochrome-C oxidase complex, and transport protons to the mitochondrial intermembrane space to generate transmembrane proton gradients and electrical potential for the synthesis of ADP to ATP *via* ATP synthase (Saraste, 1999). The ATP produced by the TCA cycle is mainly from NADH, which accounts for approximately 75% of the total (Pelley, 2007). Studies have reported that increasing the concentration of NADH significantly increased the ATP synthesis capacity of cardiomyocytes (Pelzmann et al., 2003); this suggests that the cell's ability to synthesize ATP is positively correlated with its ability to produce NADH in the TCA cycle. The number of molecules of NADH that can be produced in one TCA cycle is 3; hence, the cell's ability to produce ATP is related to the rate, or flux, of the TCA cycle.

Our results showed that the genes encoding Pdh, Idh, and Mdh were significantly upregulated in Nile tilapia with superior growth performance. Pdh has long been considered a key factor in the entry of pyruvate into the TCA cycle, and it plays an important catalytic role in the synthesis of acetyl-CoA from pyruvate, which is the only way for mammals to synthesize pyruvate into acetyl-CoA (Kim et al., 2006; Tennant et al., 2010). As the first enzyme involved in the TCA cycle, its activity directly affects the rate of the cycle (Whitehouse et al., 1974). Idh is one of the key enzymes in the TCA cycle, wherein isocitrate is converted to oxalosuccinate and rapidly

decarboxylated to *a*-ketoglutaric acid; because this reaction is irreversible, Idh is the key rate-limiting enzyme in the TCA cycle (Xu et al., 2004; Leonardi et al., 2012). In the final step of the cycle, the oxaloacetate initially consumed is re-synthesized by (L)-malate and catalyzed by Mdh, which is a key step in making the TCA cycle a closed loop (van der Rest et al., 2000). This suggests that increased Pdh, Idh, and Mdh activity may be the main way by which cells regulate the increased flux of the TCA cycle. Interestingly, we found that the reactions in which Pdh, Idh, and Mdh participate in the TCA cycle all catalyze the production of NADH, the most critical substrate promoting the conversion of ADP to ATP. Our method of providing adequate feed for each experimental fish excluded a decrease in Glu caused by inadequate food intake. As noted above, the concentration of Glu in the HG group of tilapia was significantly lower than in the LG group, combined with the significant upregulation of *pdh*, *idh*, and *mdh* in the superior growth performance tilapia. We suggest that excessive Glu consumption in superior growth performance tilapia may be related to Pdh, Idh, and Mdh and the TCA cycle in which they are located. The specific regulatory processes of which need to be further investigation.

Mechanisms of the Fatty Acid Biosynthesis and Metabolism Pathway in Tilapia With Superior Growth Performance

Fatty acids, the main components of many key substances such as fats, phospholipids, and glycolipids, are the main source of energy for most organisms (Dole and Meinertz, 1960), and their biosynthesis is facilitated when the organism needs to obtain energy or to carry out energy storage (Gibson, 1963). Fatty acids are synthesized from acetyl-CoA in a process that occurs in the cytoplasm and is catalyzed by several enzymes; however, acetyl-CoA is produced in the mitochondria and cannot cross the mitochondrial membrane directly, so it is first bound to oxaloacetate to produce citrate, which can enter the cytoplasm *via* the citrate transporter protein (Slc25a1) and is subsequently reconverted to oxaloacetate and acetyl-CoA catalyzed by Acly (Burke and Huff, 2017). In the initial step of fatty acid biosynthesis, acetyl-CoA is synthesized into malonyl-CoA catalyzed by Acc, which is the main form of acetyl-CoA participating in fatty acid synthesis, and therefore this process is an important rate-limiting step (Park et al., 2002; Saggerson, 2008). In fatty acid biosynthesis in animal cells, acetyl-CoA is used as the starting fragment and may be extended seven times by the addition of two carbon atoms (malonyl-CoA) to the carboxyl terminus to produce palmityl ACP, through a process catalyzed by Fasn (Chavin et al., 2004). Palmityl ACP is catalyzed by palmityl ACP thioesterase to release palmitic acid, which completes the entire fatty acid biosynthesis process (Jones et al., 1995). In our results, the genes encoding Acly, Acc, and Fasn, which are key enzymes involved in fatty acid biosynthesis, were significantly upregulated in tilapia with superior growth performance. This suggests a higher capacity for fatty acid biosynthesis in those fish. In addition, we identified changes in some of the transcription factors that can be involved in fatty acid biosynthesis in superior growth performance tilapia. Among them, *srebp-1* appeared significantly upregulated, whereas *foxos* and *sirt1* were significantly downregulated. The current research suggests that Srebp-1 promotes transcription of *acc* and *fasn* (Wang et al., 1994). Although the role of Foxo in the regulation of lipid metabolism is not fully understood, it has been suggested that it can inhibit the expression of *srebp-1* (Deng et al., 2012). Sirt1 is an important regulator at the center of several lipid metabolism networks (Picard et al., 2004); numerous studies have shown that it can regulate the activity of various transcription factors or enzymes that play an important role in fatty acid synthesis and *ß*-oxidation, such as Srebps and Foxos, by deacetylating target proteins (Wang et al., 2017). There was a significant reduction in cellular beta-oxidation capacity and concomitant hyperlipidemia after specific deletion of Sirt1 in mouse (*Mus musculus*) (Walker et al., 2010). Thus, *via* the upregulation of *srebp-1* and the downregulation of *foxos*, *sirt1* could produce the same final effect (i.e., both promote fatty acid biosynthesis, which provides more support for our hypothesis).

Once fatty acid has been synthesized it must be activated (bind to CoA to synthesize acyl-CoA) to enter the metabolic pathway, which is the first step in the participation of fatty acids in biological reactions (Færgeman and Knudsen, 1997), and this process is catalyzed by the acyl-CoA synthetase family. In particular, Acsl specifically catalyzes the synthesis of long-chain acyl-CoA (LC acyl-CoA) from fatty acids between C12 and C22 and further participates in catabolism or anabolism (Coleman et al., 2000). The main pathway of LC acyl-CoA catabolism is *ß*-oxidation, a process that occurs within the mitochondria. To cross the mitochondrial inner membrane, LC acyl-CoA needs to bind carnitine to synthesize LC acyl-carnitine catalyzed by Cpt, and when it passes through the mitochondrial inner membrane it will release carnitine to convert to LC acyl-CoA catalyzed by Cpt (Townsend et al., 2013). Upon entry into the mitochondria, LC acyl-CoA progressively releases acetyl-CoA, the precursor substance for ATP synthesis in the TCA cycle pathway, and this process is catalyzed by Acadvl (Zhang et al., 2007). The activity of Acot1 is activated when the intracellular ATP concentration is gradually increased. Acot1 reduces the concentration of substrates involved in ATP synthesis by hydrolyzing acyl-CoA in the cytoplasm to free fatty acids and CoA to avoid the waste of energy caused by excessive ATP synthesis (Zhang et al., 2012).

Changes in the activity of the above enzymes constitute the main pathway for bioregulation of *ß*-oxidation. In this study, we found that *acsls* were significantly upregulated but *cpt1a* and *acadvl* were significantly downregulated in superior growth performance tilapia, suggesting that these tilapia have a higher capacity for fatty acid activation, but the *ß*-oxidation process is inhibited. Meanwhile, *acot1* was significantly upregulated, and, since the activity of Acot1 is only activated at high energy levels (Franklin et al., 2017), this implies that superior growth performance tilapia maintain high energy levels despite the inhibition of the *ß*-oxidation pathway, which could suggest that energy acquisition in superior growth performance tilapia is not derived from the catabolism of fatty acids (β-oxidation). In addition, we found that *lipe* was significantly downregulated. Lipe is thought to be a key enzyme in mobilizing TG deposited in adipose tissue, which hydrolyses TG to free fatty

acids, and is also the rate-limiting enzyme in the hydrolysis of diglycerides (Haemmerle et al., 2002; Zimmermann et al., 2004). This finding, combined with the significantly higher TG concentrations in the HG tilapia compared with in the LG fish suggests that the activated fatty acids in superior growth performance tilapia are used more for TG synthesis than for ß-oxidation. Moreover, the inhibition of the catabolic pathway of TG implies that its storage capacity is improved.

An integrated description of the above results suggests that superior growth performance tilapia may have a high capacity for fatty acid synthesis and activation, but that the ß-oxidation pathway is inhibited without a concomitant reduction in cellular energy levels, thus the fish may then be able to synthesize and store TG under the influence of higher energy levels. In short, we suggest that Nile tilapia with superior growth performance may have lower activity in the fatty acid catabolic pathway but higher activity in the anabolic pathway. Consequently, the large amount of synthesized fatty acids may be used more for TG synthesis after activation, to store excess energy.

Mechanisms of the Cholesterol Biosynthesis and Metabolism Pathway in Tilapia With Superior Growth Performance

Cholesterol is the most abundant sterol compound in nature and is one of the most important lipids in the human body (Brown and Goldstein, 1997). The biosynthesis of cholesterol is complex, with nearly 30 enzymatic steps (Stocco and Clark, 1994), which can be divided into three main stages: the synthesis of isopentenyl pyrophosphate (IPP) from acetyl CoA; the conversion of IPP to squalene; and the conversion of squalene to cholesterol. In our study, superior growth performance tilapia showed upregulation of genes encoding key enzymes within all these stages. Combined with the results we reported here, where the concentration of TC was significantly higher in the HG group than in the LG tilapia, this may suggest a higher capacity for cholesterol biosynthesis in superior growth performance tilapia.

Unlike fatty acids, cholesterol cannot be degraded to CO_2 and H_2O but is oxidized to become important bioactive substances, such as bile acids, steroid hormones, and vitamin D3 (Goldstein and Brown, 1990). Bile acids are the main product of cholesterol metabolism in the liver. After being synthesized, they are concentrated and stored in the gallbladder and participate in the digestion and absorption of lipids and fat-soluble vitamins after food intake (Li and Chiang, 2014; Di Ciaula et al., 2017). Bile acids play an important regulatory role in TG and lipid metabolism; in mouse, increasing the number of circulating bile acids improved high-fat diet-induced obesity (Harach et al., 2012; Kohli et al., 2013). Steroid hormones are another important endogenous substance synthesized from cholesterol and can be classified as corticosteroids and sex hormones. Adrenocorticotropic hormones are synthesized from cholesterol in the zona glomerulosa of the adrenal cortex and can promote gluconeogenesis and protein metabolism and increase the synthesis of liver glycogen and muscle glycogen (Campbell, 2011). Sex hormones are synthesized in the zona reticularis of the adrenal cortex and transformed into dihydrotestosterone or estradiol in the testes or ovaries, where they play a physiological role in promoting protein synthesis, especially in the muscle and reproductive organs, as well as bone growth (Gorski and Gannon, 1976). Numerous studies have shown that sex hormones in humans promote the secretion of growth hormone (GH) and mediate the GH-IGF1 signaling pathway to influence the development and growth of bones and muscles in adolescents (Tanner et al., 1976; Holmes and Shalet, 1996; Mauras et al., 1996). In studies on tilapia, it was found that sex hormones may mediate growth hormone-releasing hormone to affect regulation of tilapia growth (Melamed et al., 1995), suggesting the sex hormones that regulate growth in tilapia are similar to those humans. In our study, *cyp8b1* and *cyp46a1* were significantly downregulated, and *hsd17b12*, *akr1d1*, and *ugt* were significantly upregulated in the livers of superior growth performance tilapia. Among these, Cyp8b1 and Cyp46a1 are key regulators of cholesterol synthesis into bile acids (Savolainen et al., 2004; Yang et al., 2004), while Hsd17b12, Akr1d1, and Ugt play key roles in several steps of cholesterol synthesis into steroid hormone precursors (Lima et al., 2013; Chen et al., 2019). Together these results may suggest that superior growth performance Nile tilapia have a lower capacity for cholesterol synthesis for conversion into bile acids and a higher capacity for synthesis towards steroid hormone precursors.

Other Key Factors That May Be Involved in the Regulation of Tilapia Metabolism

The growth hormone-insulin like growth factor1 (GH-IGF1) is a key factor affecting animal growth. Previous studies showed that the growth-promoting effect of Igf1 may depend on its feedback regulation of starvation (Rousseau and Dufour, 2007), such as in rainbow trout (*Oncorhynchus mykiss*) (Niu et al., 1993), channel catfish (*Ictalurus punctatus*) (Small and Peterson, 2005), coho salmon (Moriyama et al., 1994), and tilapia (Uchida et al., 2003). However, in this study, all individuals were fed to apparent satiation (see 2.1 for details), therefore the regulatory role of Igf1 on tilapia growth *via* starvation feedback (to promote food intake) may have been reduced; this might explain why Igf1, and some other appetite-related factors such as *ghrelin* and *prepro-orexin* (Unniappan et al., 2004; Novak et al., 2005), were not screened in this study.

It has been reported that Gh can indirectly participate in metabolic processes in animals by regulating downstream factors, including stimulation of lipid mobilization, and induction of protein synthesis (Sheridan, 1994; Björnsson, 1997). In addition, some key genes located in the hypothalamus, such as pituitary adenylate cyclase activating polypeptide (*pacap*), corticotropin-releasing hormone (*chr*), and thyrotropin-releasing hormone (*trh*), also might be indirectly involved in the metabolic regulation of animals as upstream factors (Gershengorn, 1982; Rousseau and Dufour, 1999; Montero et al., 2000). In our study, many genes related to lipid regulation (*acc*, *srebp1*, *sirt1*, *fasn*, *acsl*, *acot1*, *lipe*) and protein synthesis (*hsd17b12*, *akr1d1*, *ugt*, *cyp8b1*, *cyp46a1*) were identified, and the variability of these factors may be related to regulation of the upstream genes mentioned above. The

metabolism-related genes obtained in this study were mainly in the liver; accordingly, these upstream key factors were not screened, yet their role in the regulation of metabolism cannot be ignored, and the regulatory mechanisms of metabolism-related genes between upstream and downstream need to be further investigated.

CONCLUSION

This study used biochemical and transcriptome analyses to compare individuals of Nile tilapia to reveal significant growth differences under the same culture environment and feeding regime. The results showed that the contents of TP, TC, TG, HDL, and LDL were significantly increased, while Glu was significantly reduced in the superior growth performance individuals. Furthermore, 29 metabolism-related candidate genes were obtained and showed significant differential expression in the superior growth performance tilapia. The possible relationships between these genes were described, and three potential metabolic pathways were mapped. To our knowledge, this is the first study to describe the potential metabolic characterization of superior growth performance Nile tilapia obtained from full-sibling families reared in the same culture environment; our findings provide more information and ideas for improved growth performance of farmed tilapia.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: PRJNA787719.

ETHICS STATEMENT

The animal study was reviewed and approved by Ethics Committee of Freshwater Fisheries Research Center, China Academy of Fisheries Science.

AUTHOR CONTRIBUTIONS

BC, WX, and HY designed the experiments. BC and WX wrote the manuscript. BC performed the experiments and conducted the data analysis. BC, WX, ZZ, JZ, DL, JY, and HY cultured and collected tilapia. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Central Public-interest Scientific Institution Basal Research Fund, CAFS (Nos. 2021JBFM03 and 2020TD37), and supported by China Agriculture Research System of MOF and MARA (CARS-46).

ACKNOWLEDGMENTS

We thank Chengliang Wei from the Freshwater Fisheries Research Center, Chinese Academy of Fishery Sciences for his contribution to the germplasm preservation and breeding of Nile tilapia.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.879570/full#supplementary-material

REFERENCES

Ani, J. S., Manyala, J. O., Masese, F. O., and Fitzsimmons, K. (2022). Effect of Stocking Density on Growth Performance of Monosex Nile Tilapia (*Oreochromis niloticus*) in the Aquaponic System Integrated with Lettuce (Lactuca sativa). *Aquac. Fish.* 7, 328–335. doi:10.1016/j.aaf.2021.03.002

Bell, C. (2012). Regulation of Metabolism" in Primer on the Autonomic Nervous System. *elsevier*, 253–255. doi:10.1016/B978-0-12-386525-0.00052-4

Björnsson, B. T. (1997). The Biology of Salmon Growth Hormone: From Daylight to Dominance. *Fish Physiology Biochem.* 17, 9–24. doi:10.1023/a:1007712413908

Brown, M. S., and Goldstein, J. L. (1997). The SREBP Pathway: Regulation of Cholesterol Metabolism by Proteolysis of a Membrane-Bound Transcription Factor. *Cell.* 89, 331–340. doi:10.1016/S0092-8674(00)80213-5

Burke, A. C., and Huff, M. W. (2017). ATP-Citrate Lyase. *Curr. Opin. Lipidol.* 28, 193–200. doi:10.1097/MOL.0000000000000390

Campbell, I. (2011). Adrenocortical Hormones. *Anaesth. Intensive Care Med.* 12, 461–464. doi:10.1016/j.mpaic.2011.07.004

Charo-Karisa, H., Komen, H., Reynolds, S., Rezk, M. A., Ponzoni, R. W., and Bovenhuis, H. (2006). Genetic and Environmental Factors Affecting Growth of Nile tilapia (*Oreochromis niloticus*) Juveniles: Modelling Spatial Correlations between Hapas. *Aquaculture* 255, 586–596. doi:10.1016/j.aquaculture.2005.11.041

Chavin, K. D., Fiorini, R. N., Shafizadeh, S., Cheng, G., Wan, C., Evans, Z., et al. (2004). Fatty Acid Synthase Blockade Protects Steatotic Livers from Warm Ischemia Reperfusion Injury and Transplantation. *Am. J. Transplant.* 4, 1440–1447. doi:10.1111/j.1600-6143.2004.00546.x

Chen, B., Xiao, W., Zou, Z., Zhu, J., Li, D., Yu, J., et al. (2020). Ghrelin Gene Single Nucleotide Polymorphisms and Their Effects on Nile tilapia (*Oreochromis niloticus*) Growth. *Aquac. Rep.* 18, 100469. doi:10.1016/j.aqrep.2020.100469

Chen, B., Xiao, W., Zou, Z., Zhu, J., Li, D., Yu, J., et al. (2021). The Effects of Single Nucleotide Polymorphisms in Neuropeptide Y and Prepro-Orexin on Growth in Nile tilapia (*Oreochromis niloticus*). *Aquaculture* 543, 736974. doi:10.1016/j.aquaculture.2021.736974

Chen, M., Wangtrakuldee, P., Zang, T., Duan, L., Gathercole, L. L., Tomlinson, J. W., et al. (2019). Human and Murine Steroid 5β-Reductases (AKR1D1 and AKR1D4): Insights into the Role of the Catalytic Glutamic Acid. *Chemico-Biological Interact.* 305, 163–170. doi:10.1016/j.cbi.2019.03.025

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an Ultra-fast All-In-One FASTQ Preprocessor. *Bioinformatics* 34, i884–i890. doi:10.1093/bioinformatics/bty560

Circa, A. V., Eknath, A. E., and Taduan, A. G. (1995). Genetic Improvement of Farmed tilapias: the Growth Performance of the Gift Strain of Nile tilapia (*Oreochromis niloticus*) in Rice-Fish Environments. *Aquaculture* 137, 329–330. doi:10.1016/0044-8486(96)83588-5

Coleman, R. A., Lewin, T. M., and Muoio, D. M. (2000). Physiological and Nutritional Regulation of Enzymes of Triacylglycerol Synthesis. *Annu. Rev. Nutr.* 20, 77–103. doi:10.1146/annurev.nutr.20.1.77

Conte, M. A., Gammerdinger, W. J., Bartie, K. L., Penman, D. J., and Kocher, T. D. (2017). A High Quality Assembly of the Nile Tilapia (*Oreochromis niloticus*) Genome Reveals the Structure of Two Sex Determination Regions. *BMC Genomics* 18, 341. doi:10.1186/s12864-017-3723-5

DeBerardinis, R. J., Sayed, N., Ditsworth, D., and Thompson, C. B. (2008). Brick by Brick: Metabolism and Tumor Cell Growth. *Curr. Opin. Genet. Dev.* 18, 54–61. doi:10.1016/j.gde.2008.02.003

Deloffre, L. A. M., Andrade, A., Filipe, A. I., and Canario, A. V. M. (2012). Reference Genes to Quantify Gene Expression during Oogenesis in a Teleost Fish. *Gene* 506, 69–75. doi:10.1016/j.gene.2012.06.047

Deng, X., Zhang, W., O-Sullivan, I., Williams, J. B., Dong, Q., Park, E. A., et al. (2012). FoxO1 Inhibits Sterol Regulatory Element-Binding Protein-1c (SREBP-1c) Gene Expression *via* Transcription Factors Sp1 and SREBP-1c. *J. Biol. Chem.* 287, 20132–20143. doi:10.1074/jbc.M112.347211

Di Ciaula, A., Garruti, G., Lunardi Baccetto, R., Molina-Molina, E., Bonfrate, L., Wang, D. Q.-H., et al. (2017). Bile Acid Physiology. *Ann. Hepatology* 16, S4–S14. doi:10.5604/01.3001.0010.5493

Dole, V. P., and Meinertz, H. (1960). Microdetermination of Long-Chain Fatty Acids in Plasma and Tissues. *J. Biol. Chem.* 235, 2595–2599. doi:10.1016/S0021-9258(19)76920-8

Du, Z., and Turchini, G. M. (2022). Are We Actually Measuring Growth? An Appeal to Use a More Comprehensive Growth Index System for Advancing Aquaculture Research. *Rev. Aquac.* 14, 525–527. doi:10.1111/raq.12604

Enquist, B. J., Economo, E. P., Huxman, T. E., Allen, A. P., Ignace, D. D., and Gillooly, J. F. (2003). Scaling Metabolism from Organisms to Ecosystems. *Nature* 423, 639–642. doi:10.1038/nature01671

Færgeman, N. J., and Knudsen, J. (1997). Role of Long-Chain Fatty Acyl-CoA Esters in the Regulation of Metabolism and in Cell Signalling. *Biochem. J.* 323, 1–12. doi:10.1042/bj3230001

Franklin, M. P., Sathyanarayan, A., and Mashek, D. G. (2017). Acyl-CoA Thioesterase 1 (ACOT1) Regulates PPARα to Couple Fatty Acid Flux with Oxidative Capacity during Fasting. *Diabetes* 66, 2112–2123. doi:10.2337/db16-1519

Genschick, S., Mekkawy, W., Rossignoli, C., and Benzie, J. A. H. (2021). Growth Performance of Three Strains of Nile tilapia (*Oreochromis niloticus*) on Four Different Feeds in Western and Central Kenya. *Aquac. Rep.* 20, 100701. doi:10.1016/j.aqrep.2021.100701

Gershengorn, M. C. (1982). Thyrotropin Releasing Hormone. *Mol. Cell. Biochem.* 45. doi:10.1007/BF00230085

Gibson, D. (1963). Fatty Acid Biosynthesis. *Prog. Chem. Fats other Lipids* 6, 117–136. doi:10.1016/0079-6832(63)90021-1

Ginestet, C. (2011). ggplot2: Elegant Graphics for Data Analysis. *J. R. Stat. Soc. Ser. A (Statistics Soc.* 174, 245–246. doi:10.1111/j.1467-985X.2010.00676_9.x

Goldstein, J. L., and Brown, M. S. (1990). Regulation of the Mevalonate Pathway. *Nature* 343, 425–430. doi:10.1038/343425a0

Gorski, J., and Gannon, F. (1976). Current Models of Steroid Hormone Action: A Critique. *Annu. Rev. Physiology* 38, 425–450. doi:10.1146/annurev.ph.38.030176.002233

Haemmerle, G., Zimmermann, R., Hayn, M., Theussl, C., Waeg, G., Wagner, E., et al. (2002). Hormone-sensitive Lipase Deficiency in Mice Causes Diglyceride Accumulation in Adipose Tissue, Muscle, and Testis. *J. Biol. Chem.* 277, 4806–4815. doi:10.1074/jbc.M110355200

Hansen, J. (2005). Using SPSS for Windows and Macintosh: Analyzing and Understanding Data. *Am. Statistician* 59, 113. doi:10.1198/tas.2005.s139

Harach, T., Pols, T. W. H., Nomura, M., Maida, A., Watanabe, M., Auwerx, J., et al. (2012). TGR5 Potentiates GLP-1 Secretion in Response to Anionic Exchange Resins. *Sci. Rep.* 2, 430. doi:10.1038/srep00430

Holmes, S. J., and Shalet, S. M. (1996). Role of Growth Hormone and Sex Steroids in Achieving and Maintaining Normal Bone Mass. *Hormone Res.* 45, 86–93. doi:10.1159/000184765

Jones, A., Davies, H. M., and Voelker, T. A. (1995). Palmitoyl-acyl Carrier Protein (ACP) Thioesterase and the Evolutionary Origin of Plant Acyl-ACP Thioesterases. *Plant Cell.* 7, 359–371. doi:10.1105/tpc.7.3.359

Khatab, S., Hemeda, S., ElNahas, A., and Naby, W. (2014). Genetic Polymorphism in IGF-II Gene and its Relationship with Growth Rate in tilapia Nilotica. *Alexandria J. Veterinary Sci.* 43, 26. doi:10.5455/ajvs.167827

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: A Fast Spliced Aligner with Low Memory Requirements. *Nat. Methods* 12, 357–360. doi:10.1038/nmeth.3317

Kim, J., Tchernyshyov, I., Semenza, G. L., and Dang, C. V. (2006). HIF-1-mediated Expression of Pyruvate Dehydrogenase Kinase: A Metabolic Switch Required for Cellular Adaptation to Hypoxia. *Cell. Metab.* 3, 177–185. doi:10.1016/j.cmet.2006.02.002

Kohli, R., Setchell, K. D., Kirby, M., Myronovych, A., Ryan, K. K., Ibrahim, S. H., et al. (2013). A Surgical Model in Male Obese Rats Uncovers Protective Effects of Bile Acids Post-bariatric Surgery. *Endocrinology* 154, 2341–2351. doi:10.1210/en.2012-2069

Krebs, H. A., and Johnson, W. A. (1980). The Role of Citric Acid in Intermediate Metabolism in Animal Tissues. *FEBS Lett.* 117, K2–K10. doi:10.1016/0014-5793(80)80564-3

Leonardi, R., Subramanian, C., Jackowski, S., and Rock, C. O. (2012). Cancer-associated Isocitrate Dehydrogenase Mutations Inactivate NADPH-dependent Reductive Carboxylation. *J. Biol. Chem.* 287, 14615–14620. doi:10.1074/jbc.C112.353946

Li, B., and Dewey, C. N. (2011). RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome. *BMC Bioinforma.* 12, 323. doi:10.1186/1471-2105-12-323

Li, T., and Chiang, J. Y. L. (2014). Bile Acid Signaling in Metabolic Disease and Drug Therapy. *Pharmacol. Rev.* 66, 948–983. doi:10.1124/pr.113.008201

Lima, D., Machado, A., Reis-Henriques, M. A., Rocha, E., Santos, M. M., and Castro, L. F. C. (2013). Cloning and Expression Analysis of the 17β Hydroxysteroid Dehydrogenase Type 12 (HSD17B12) in the Neogastropod *Nucella lapillus*. *J. Steroid Biochem. Mol. Biol.* 134, 8–14. doi:10.1016/j.jsbmb.2012.10.005

Lin, G., Zheng, M., Li, S., Xie, J., Fang, W., Gao, D., et al. (2020). Response of Gut Microbiota and Immune Function to Hypoosmotic Stress in the Yellowfin Seabream (*Acanthopagrus Latus*). *Sci. Total Environ.* 745, 140976. doi:10.1016/j.scitotenv.2020.140976

Liu, Y., Li, E., Xu, C., Su, Y., Qin, J. G., Chen, L., et al. (2018). Brain Transcriptome Profiling Analysis of Nile tilapia (*Oreochromis niloticus*) under Long-Term Hypersaline Stress. *Front. Physiology* 9, 219. doi:10.3389/fphys.2018.00219

Love, M. I., Huber, W., and Anders, S. (2014). Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8

Mauras, N., Rogol, A. D., Haymond, M. W., and Veldhuis, J. D. (1996). Sex Steroids, Growth Hormone, Insulin-like Growth Factor-1: Neuroendocrine and Metabolic Regulation in Puberty. *Hormone Res.* 45, 74–80. doi:10.1159/000184763

McClelland, E. K., Chan, M. T. T., Lin, X., Sakhrani, D., Vincelli, F., Kim, J.-H., et al. (2020). Loci Associated with Variation in Gene Expression and Growth in Juvenile Salmon Are Influenced by the Presence of a Growth Hormone Transgene. *BMC Genomics* 21, 185. doi:10.1186/s12864-020-6586-0

Melamed, P., Eliahu, N., Ofir, M., Levavi-Sivan, B., Smal, J., Rentier-Delrue, F., et al. (1995). The Effects of Gonadal Development and Sex Steroids on Growth Hormone Secretion in the Male tilapia Hybrid (*Oreochromis niloticus* X *O. aureus*). *Fish Physiology Biochem.* 14, 267–277. doi:10.1007/BF00004065

Mihaylova, M. M., and Shaw, R. J. (2011). The AMPK Signalling Pathway Coordinates Cell Growth, Autophagy and Metabolism. *Nat. Cell. Biol.* 13, 1016–1023. doi:10.1038/ncb2329

Montero, M., Yon, L., Kikuyama, S., Dufour, S., and Vaudry, H. (2000). Molecular Evolution of the Growth Hormone-Releasing Hormone/pituitary Adenylate Cyclase-Activating Polypeptide Gene Family. Functional Implication in the Regulation of Growth Hormone Secretion. *J. Mol. Endocrinol.* 25, 157–168. doi:10.1677/jme.0.0250157

Moriyama, S., Swanson, P., Nishii, M., Takahashi, A., Kawauchi, H., Diekhoff, W. W., et al. (1994). Development of a Homologous Radioimmunoassay for Coho Salmon Insulin-like Growth Factor-I. *General Comp. Endocrinol.* 96, 149–161. doi:10.1006/gcen.1994.1167

Neira, R., Lhorente, J. P., Araneda, C., Díaz, N., Bustos, E., and Alert, A. (2004). Studies on Carcass Quality Traits in Two Populations of Coho Salmon (*Oncorhynchus kisutch*): Phenotypic and Genetic Parameters. *Aquaculture* 241, 117–131. doi:10.1016/j.aquaculture.2004.08.009

Niu, P. De, Perez-Sanchez, J., and Le Bail, P. Y. (1993). Development of a Protein Binding Assay for Teleost Insulin-like Growth Factor (IGF)-like: Relationships between Growth Hormone (GH) and IGF-like in the Blood of Rainbow Trout (*Oncorhynchus mykiss*). *Fish Physiology Biochem.* 11, 381–391. doi:10.1007/BF00004588

Novak, C. M., Jiang, X., Wang, C., Teske, J. A., Kotz, C. M., and Levine, J. A. (2005). Caloric Restriction and Physical Activity in Zebrafish (*Danio rerio*). *Neurosci. Lett.* 383, 99–104. doi:10.1016/j.neulet.2005.03.048

Park, H., Kaushik, V. K., Constant, S., Prentki, M., Przybytkowski, E., Ruderman, N. B., et al. (2002). Coordinate Regulation of Malonyl-Coa Decarboxylase,sn-Glycerol-3-Phosphate Acyltransferase, and Acetyl-Coa Carboxylase by AMP-Activated Protein Kinase in Rat Tissues in Response to Exercise. *J. Biol. Chem.* 277, 32571–32577. doi:10.1074/jbc.M201692200

Pelley, J. W. (2007). *Citric Acid Cycle, Electron Transport Chain, and Oxidative Phosphorylation" in* Elsevier's Integrated Biochemistry. Netherlands: Elsevier, 55–63. doi:10.1016/B978-0-323-03410-4.50013-4

Pelzmann, B., Hallström, S., Schaffer, P., Lang, P., Nadlinger, K., Birkmayer, G. D., et al. (2003). NADH Supplementation Decreases Pinacidil-Primed I K(ATP) in Ventricular Cardiomyocytes by Increasing Intracellular ATP. *Br. J. Pharmacol.* 139, 749–754. doi:10.1038/sj.bjp.0705300

Picard, F., Kurtev, M., Chung, N., Topark-Ngarm, A., Senawong, T., Machado de Oliveira, R., et al. (2004). Sirt1 Promotes Fat Mobilization in White Adipocytes by Repressing PPAR-γ. *Nature* 429, 771–776. doi:10.1038/nature02583

Ren, Y., Yu, G., Shi, C., Liu, L., Guo, Q., Han, C., et al. (2022). Majorbio Cloud: A One-stop, Comprehensive Bioinformatic Platform for Multiomics Analyses. *iMeta* 1, e12. doi:10.1002/imt2.12

Rose, A. H. (1968). *Regulation of Metabolism" in* Chemical Microbiology. Boston, MA: Springer US, 218–237. doi:10.1007/978-1-4899-6567-7_8

Rousseau, K., and Dufour, S. (1999). Evidence that Corticotropin-Releasing Hormone Acts as a Growth Hormone-releasing Factor in a Primitive Teleost, the European Eel (*Anguilla anguilla*). *J. Neuroendocrinol.* 11, 385–392. doi:10.1046/j.1365-2826.1999.00334.x

Rousseau, K., and Dufour, S. (2007). Comparative Aspects of GH and Metabolic Regulation in Lower Vertebrates. *Neuroendocrinology* 86, 165–174. doi:10.1159/000101029

Rutten, M. J. M., Bovenhuis, H., and Komen, H. (2005). Genetic Parameters for Fillet Traits and Body Measurements in Nile tilapia (*Oreochromis niloticus* L.). *Aquaculture* 246, 125–132. doi:10.1016/j.aquaculture.2005.01.006

Ryan, D. G., and O'Neill, L. A. J. (2020). Krebs Cycle Reborn in Macrophage Immunometabolism. *Annu. Rev. Immunol.* 38, 289–313. doi:10.1146/annurev-immunol-081619-104850

Saggerson, D. (2008). Malonyl-CoA, a Key Signaling Molecule in Mammalian Cells. *Annu. Rev. Nutr.* 28, 253–272. doi:10.1146/annurev.nutr.28.061807.155434

Santos, C. A., Andrade, S. C. S., Teixeira, A. K., Farias, F., Guerrelhas, A. C., Rocha, J. L., et al. (2021). Transcriptome Differential Expression Analysis Reveals the Activated Genes in *Litopenaeus Vannamei* Shrimp Families of Superior Growth Performance. *Aquaculture* 531, 735871. doi:10.1016/j.aquaculture.2020.735871

Saraste, M. (1999). Oxidative phosphorylation at the fin de siècle. *Science* 283, 1488–1493. doi:10.1126/science.283.5407.1488

Savolainen, K., Kotti, T. J., Schmitz, W., Savolainen, T. I., Sormunen, R. T., Ilves, M., et al. (2004). A Mouse Model for α-methylacyl-CoA Racemase Deficiency: Adjustment of Bile Acid Synthesis and Intolerance to Dietary Methyl-Branched Lipids. *Hum. Mol. Genet.* 13, 955–965. doi:10.1093/hmg/ddh107

Sedgwick, P. (2010). Independent Samples T Test. *BMJ* 340, c2673. doi:10.1136/bmj.c2673

Sheridan, M. A. (1994). Regulation of Lipid Metabolism in Poikilothermic Vertebrates. *Comp. Biochem. Physiology Part B Comp. Biochem.* 107, 495–508. doi:10.1016/0305-0491(94)90176-7

Small, B. C., and Peterson, B. C. (2005). Establishment of a Time-Resolved Fluoroimmunoassay for Measuring Plasma Insulin-like Growth Factor I (IGF-I) in Fish: Effect of Fasting on Plasma Concentrations and Tissue mRNA Expression of IGF-I and Growth Hormone (GH) in Channel Catfish (*Ictalurus punctatus*). *Domest. Anim. Endocrinol.* 28, 202–215. doi:10.1016/j.domaniend.2004.09.002

Stocco, D. M., and Clark, B. J. (1994). *Regulation of the Acute Production of Steroids in Steroidogenic Cells" in* Molecular And Cellular Endocrinology Of the Testis. Berlin, Heidelberg: Springer Berlin Heidelberg, 67–98. doi:10.1007/978-3-662-22189-1_4

Tanner, J. M., Whitehouse, R. H., Hughes, P. C. R., and Carter, B. S. (1976). Relative Importance of Growth Hormone and Sex Steroids for the Growth at Puberty of Trunk Length, Limb Length, and Muscle Width in Growth Hormone-Deficient Children. *J. Pediatr.* 89, 1000–1008. doi:10.1016/S0022-3476(76)80620-8

Tao, W., Chen, J., Tan, D., Yang, J., Sun, L., Wei, J., et al. (2018). Transcriptome Display during tilapia Sex Determination and Differentiation as Revealed by RNA-Seq Analysis. *BMC Genomics* 19, 363. doi:10.1186/s12864-018-4756-0

Tennant, D. A., Durán, R. V., and Gottlieb, E. (2010). Targeting Metabolic Transformation for Cancer Therapy. *Nat. Rev. Cancer* 10, 267–277. doi:10.1038/nrc2817

The State of World Fisheries and Aquaculture (2020). *The State of World Fisheries and Aquaculture 2020*. Rome: FAO. doi:10.4060/ca9229en

Townsend, K. L., An, D., Lynes, M. D., Huang, T. L., Zhang, H., Goodyear, L. J., et al. (2013). Increased Mitochondrial Activity in Bmp7-Treated Brown Adipocytes, Due to Increased CPT1- and CD36-Mediated Fatty Acid Uptake. *Antioxidants Redox Signal.* 19, 243–257. doi:10.1089/ars.2012.4536

Turra, E. M., Toral, F. L. B., Alvarenga, É. R. de, Raidan, F. S. S., Fernandes, A. F. de A., Alves, G. F. de O., et al. (2016). Genotype×environment Interaction for Growth Traits of Nile tilapia in Biofloc Technology, Recirculating Water and Cage Systems. *Aquaculture* 460, 98–104. doi:10.1016/j.aquaculture.2016.04.020

Uchida, K., Kajimura, S., Riley, L. G., Hirano, T., Aida, K., and Grau, E. G. (2003). Effects of Fasting on Growth Hormone/insulin-like Growth Factor I axis in the tilapia, *Oreochromis mossambicus*. *Comp. Biochem. Physiology - A Mol. Integr. Physiology* 134, 429–439. doi:10.1016/S1095-6433(02)00318-5

Unniappan, S., Canosa, L. F., and Peter, R. E. (2004). Orexigenic Actions of Ghrelin in Goldfish: Feeding-Induced Changes in Brain and Gut Mrna Expression and Serum Levels, and Responses to Central and Peripheral Injections. *Neuroendocrinology* 79, 100–108. doi:10.1159/000076634

Valenzuela-Miranda, D., Del Río-Portilla, M. A., and Gallardo-Escárate, C. (2015). Characterization of the Growth-Related Transcriptome in California Red Abalone (*Haliotis rufescens*) through RNA-Seq Analysis. *Mar. Genomics* 24, 199–202. doi:10.1016/j.margen.2015.05.009

van der Rest, M. E., Frank, C., and Molenaar, D. (2000). Functions of the Membrane-Associated and Cytoplasmic Malate Dehydrogenases in the Citric Acid Cycle of *Escherichia coli*. *J. Bacteriol.* 182, 6892–6899. doi:10.1128/JB.182.24.6892-6899.2000

Walker, A. K., Yang, F., Jiang, K., Ji, J.-Y., Watts, J. L., Purushotham, A., et al. (2010). Conserved Role of SIRT1 Orthologs in Fasting-dependent Inhibition of the Lipid/cholesterol Regulator SREBP. *Genes Dev.* 24, 1403–1417. doi:10.1101/gad.1901210

Wang, L.-F., Wang, X.-N., Huang, C.-C., Hu, L., Xiao, Y.-F., Guan, X.-H., et al. (2017). Inhibition of NAMPT Aggravates High Fat Diet-Induced Hepatic Steatosis in Mice through Regulating Sirt1/AMPKα/SREBP1 Signaling Pathway. *Lipids Health Dis.* 16, 82. doi:10.1186/s12944-017-0464-z

Wang, L., Liu, P., Wan, Z. Y., Huang, S. Q., Wen, Y. F., Lin, G., et al. (2016). RNA-seq Revealed the Impairment of Immune Defence of tilapia against the Infection of *Streptococcus Agalactiae* with Simulated Climate Warming. *Fish Shellfish Immunol.* 55, 679–689. doi:10.1016/j.fsi.2016.06.058

Wang, L., Zhu, W., Dong, Z., Song, F., Dong, J., and Fu, J. (2018). Comparative microRNA-Seq Analysis Depicts Candidate Mirnas Involved in Skin Color Differentiation in Red Tilapia. *Int. J. Mol. Sci.* 19, 1209. doi:10.3390/ijms19041209

Wang, M.-C., Hsu, M.-T., Lin, C.-C., Hsu, S.-C., Chen, R.-D., Lee, J.-R., et al. (2022a). Adaptive Metabolic Responses in a Thermostabilized Environment: Transgenerational Trade-Off Implications from Tropical tilapia. *Sci. Total Environ.* 806, 150672. doi:10.1016/j.scitotenv.2021.150672

Wang, M., Li, B., Wang, J., Xie, S., and Zhang, L. (2022b). Skin Transcriptome and Physiological Analyses Reveal the Metabolic and Immune Responses of Yellow Catfish (*Pelteobagrus fulvidraco*) to Acute Hypoxia. *Aquaculture* 546, 737277. doi:10.1016/j.aquaculture.2021.737277

Wang, X., Sato, R., Brown, M. S., Hua, X., and Goldstein, J. L. (1994). SREBP-1, a Membrane-Bound Transcription Factor Released by Sterol-Regulated Proteolysis. *Cell.* 77, 53–62. doi:10.1016/0092-8674(94)90234-8

Whitehouse, S., Cooper, R. H., and Randle, P. J. (1974). Mechanism of Activation of Pyruvate Dehydrogenase by Dichloroacetate and Other Halogenated Carboxylic Acids. *Biochem. J.* 141, 761–774. doi:10.1042/bj1410761

Xiao, W., Chen, B., Wang, J., Zou, Z., Wang, C., Li, D., et al. (2022). Integration of mRNA and miRNA Profiling Reveals Heterosis in *Oreochromis niloticus* × *O. aureus* Hybrid Tilapia. *Animals* 12, 640. doi:10.3390/ani12050640

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., et al. (2011). KOBAS 2.0: a Web Server for Annotation and Identification of Enriched Pathways and Diseases. *Nucleic Acids Res.* 39, W316–W322. doi:10.1093/nar/gkr483

Xu, X., Zhao, J., Xu, Z., Peng, B., Huang, Q., Arnold, E., et al. (2004). Structures of Human Cytosolic Nadp-dependent Isocitrate Dehydrogenase Reveal a Novel Self-Regulatory Mechanism of Activity. *J. Biol. Chem.* 279, 33946–33957. doi:10.1074/jbc.M404298200

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

Yang, X., Yang, C., Farberman, A., Rideout, T. C., de Lange, C. F. M., France, J., et al. (2008). The Mammalian Target of Rapamycin-Signaling Pathway in Regulating Metabolism and Growth. *J. Animal Sci.* 86, E36–E50. doi:10.2527/jas.2007-0567

Yang, Y., Eggertsen, G., Gåfvels, M., Andersson, U., Einarsson, C., Björkhem, I., et al. (2004). Mechanisms of Cholesterol and Sterol Regulatory Element Binding Protein Regulation of the Sterol 12α-Hydroxylase Gene (CYP8B1). *Biochem. Biophysical Res. Commun.* 320, 1204–1210. doi:10.1016/j.bbrc.2004.06.069

Yin, X., Zhuang, X., Liao, M., Huang, L., Cui, Q., Liu, C., et al. (2022). Transcriptome Analysis of Pacific White Shrimp (*Litopenaeus Vannamei*) Hepatopancreas Challenged by *Vibrio Alginolyticus* Reveals Lipid Metabolic Disturbance. *Fish Shellfish Immunol.* 123, 238–247. doi:10.1016/j.fsi.2022.03.004

Yoshida, G. M., and Yáñez, J. M. (2021). Multi-trait GWAS Using Imputed High-Density Genotypes from Whole-Genome Sequencing Identifies Genes Associated with Body Traits in Nile tilapia. *BMC Genomics* 22, 57. doi:10.1186/s12864-020-07341-z

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene Ontology Analysis for RNA-Seq: Accounting for Selection Bias. *Genome Biol.* 11, R14. doi:10.1186/gb-2010-11-2-r14

Yu, X., Xin, Y., Cui, L., Jia, J., Yuan, X., Fu, S., et al. (2021). Effects of Neuropeptide Y as a Feed Additive on Stimulating the Growth of tilapia (*Oreochromis niloticus*) Fed Low Fish Meal Diets. *Peptides* 138, 170505. doi:10.1016/j.peptides.2021.170505

Zhang, C., Li, Q., Wu, X., Liu, Q., and Cheng, Y. (2018). Genetic Diversity and Genetic Structure of Farmed and Wild Chinese Mitten Crab (*Eriocheir Sinensis*) Populations from Three Major Basins by Mitochondrial DNA COI and Cyt B Gene Sequences. *Mitochondrial DNA Part A DNA Mapp. Sequencing, Analysis* 29, 1081–1089. doi:10.1080/24701394.2017.1404048

Zhang, D., Liu, Z.-X., Choi, C. S., Tian, L., Kibbey, R., Dong, J., et al. (2007). Mitochondrial Dysfunction Due to Long-Chain Acyl-CoA Dehydrogenase Deficiency Causes Hepatic Steatosis and Hepatic Insulin Resistance. *Proc. Natl. Acad. Sci.* 104, 17075–17080. doi:10.1073/pnas.0707060104

Zhang, X., Yuan, J., Sun, Y., Li, S., Gao, Y., Yu, Y., et al. (2019). Penaeid Shrimp Genome Provides Insights into Benthic Adaptation and Frequent Molting. *Nat. Commun.* 10, 356. doi:10.1038/s41467-018-08197-4

Zhang, Y., Li, Y., Niepel, M. W., Kawano, Y., Han, S., Liu, S., et al. (2012). Targeted Deletion of Thioesterase Superfamily Member 1 Promotes Energy Expenditure and Protects against Obesity and Insulin Resistance. *Proc. Natl. Acad. Sci.* 109, 5417–5422. doi:10.1073/pnas.1116011109

Zhou, X., Tian, Z., Wang, Y., and Li, W. (2010). Effect of Treatment with Probiotics as Water Additives on tilapia (*Oreochromis niloticus*) Growth Performance and Immune Response. *Fish Physiology Biochem.* 36, 501–509. doi:10.1007/s10695-009-9320-z

Zimmermann, R., Strauss, J. G., Haemmerle, G., Schoiswohl, G., Birner-Gruenberger, R., Riederer, M., et al. (2004). Fat Mobilization in Adipose Tissue Is Promoted by Adipose Triglyceride Lipase. *Science* 306, 1383–1386. doi:10.1126/science.1100747

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

BRIEF RESEARCH REPORT
published: 15 July 2022
doi: 10.3389/fgene.2022.928597

The Utility of Small Fishes for the Genetic Study of Human Age-Related Disorders

*Eisuke Dohi and Hideaki Matsui**

Department of Neuroscience of Disease, Brain Research Institute, Niigata University, Niigata, Japan

Animal models have been used to model human diseases, and among them, small fishes have been highlighted for their usefulness in various ways, such as the low cost of maintenance, ease of genetic modification, small size for easy handling, and strength in imaging studies due to their relative transparency. Recently, the use of turquoise killifish, *Nothobranchius furzeri*, which is known to exhibit various aging phenotypes in a short period, has attracted attention in research on aging and age-related diseases. However, when using animal models, it is important to keep their genetic background and interspecies differences in mind for translating them into human diseases. In this article, we obtained the gene symbols of protein-coding genes of turquoise killifish, medaka, zebrafish, and humans from NCBI datasets and extracted common shared genes among four species to explore the potential of interspecies translational research and to apply small fish models for human age-related disorders. Common shared protein-coding genes were analyzed with the Reactome Pathway Database to determine the coverage of these genes in each pathway in humans. We applied common shared genes to the Orphanet database to establish a list of human diseases that contain common shared genes among the four species. As examples, the senescence-related pathways and some pathways of human age-related diseases, such as Alzheimer's disease, Parkinson's disease, frontotemporal dementia, nonalcoholic fatty liver disease, progeria, hepatocellular carcinoma, and renal cell carcinoma, were extracted from the curated pathway and disease list to discuss the further utility of fish models for human age-related disorders.

Keywords: small fishes, zebrafish, medaka, turquoise killifish, age-related disorders, genomes

OPEN ACCESS

Edited by:
Mariana Härter Remião, Federal University of Pelotas, Brazil

Reviewed by:
Tyrone Genade, East Tennessee State University, United States
Frederico Schmitt Kremer, Federal University of Pelotas, Brazil

***Correspondence:**
Hideaki Matsui
hide0729@bri.niigata-u.ac.jp

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

Received: *26 April 2022*
Accepted: *08 June 2022*
Published: *15 July 2022*

Citation:
Dohi E and Matsui H (2022) The Utility of Small Fishes for the Genetic Study of Human Age-Related Disorders. Front. Genet. 13:928597. doi: 10.3389/fgene.2022.928597

INTRODUCTION

Zebrafish produce more than 100 eggs per oviposition, and medaka and turquoise killifish, *Nothobranchius furzeri*, also produce many but fewer eggs than zebrafish (Furutani-Seiki and Wittbrodt, 2004; Skinner and Watt, 2007; Polacik et al., 2016). Because small fish have a larger number of fertilized eggs than mice and because spawning and fertilization take place outside the parent's body, it is very easy to microinject the desired gene-editing factor into the egg. The recent emergence of the CRISPR–Cas9 technique (Hwang et al., 2013; Ablain et al., 2015; Varshney et al., 2015) allows us to quickly knock out or edit specific genes in various organisms compared with zinc-finger nucleases (ZFNs) (Doyon et al., 2008; Meng et al., 2008) or transcription activator-like effectors (TALENs) (Huang et al., 2011; Sander et al., 2011). The CRISPR–Cas9 system also enables

the knock-in of a specific DNA sequence mediated by homology-directed repair (HDR) or other mechanisms (Irion et al., 2014; Prykhozhij et al., 2018). These gene-editing techniques are applicable for zebrafish, medaka, and turquoise killifish. Because of the relatively hard chorion of turquoise killifish compared with zebrafish and medaka, it might be difficult to perform microinjection into the eggs of turquoise killifish. However, the methods of microinjection into the egg of turquoise killifish are being improved (Valenzano et al., 2011; Hartmann and Englert, 2012), and the use of genetic engineering in turquoise killifish has been expanding (Allard et al., 2013; Harel et al., 2016).

Turquoise killifish is a small fish species native to Africa, primarily Mozambique, that lives in ponds, swamps, and puddles (Poeschla and Valenzano, 2020). There is a short rainy season and a long dry season when adult fish cannot survive because the water dries up. Although adult fish cannot survive, turquoise killifish survive as a species through drought-resistant eggs laid in the soil, which can hatch during the future rainy season. In such a life cycle, turquoise killifish may not be subjected to a positive selection pressure to acquire various antiaging properties (Cui et al., 2019). Therefore, the lifespan of turquoise killifish is approximately four to six months, which is very short compared to zebrafish and medaka (Polacik et al., 2016). Around the age of three months, the turquoise killifish exhibits several signs of aging, such as organ atrophy, spine curvature, and increased levels of senescence-associated beta-galactosidase (Genade et al., 2005; Valenzano et al., 2006; Harel et al., 2015). Given their aging phenotypes, we examined the central nervous system of turquoise killifish and found that this fish showed age-dependent degeneration of dopaminergic and noradrenergic neurons, with gradually progressing alpha-synuclein pathology (Matsui et al., 2019). These pathological findings are similar to those of human Parkinson's disease, and very interestingly, genetic depletion of alpha-synuclein with the CRISPR–Cas9 system mitigates neurodegeneration (Matsui et al., 2019). These findings suggest that alpha-synuclein can be a causative protein in the pathogenesis of Parkinson's disease, and turquoise killifish could be a useful tool for unveiling the mechanisms of Parkinson's disease and hopefully other age-related diseases.

To further utilize the potential of translational research of such fish models, it is important to know the genetic background of each small fish compared to those of humans and other small fishes. In this article, we analyzed the genetic backgrounds of turquoise killifish, medaka, zebrafish, and humans (Reichard et al., 2009; Kirchmaier et al., 2015; Valenzano et al., 2015; Delomas and Dabrowski, 2018; Poeschla and Valenzano, 2020) and explored the utility of small fish for translational research of human age-related disorders.

METHODS

Finding the Common Shared Genes

Gene symbols of each species were extracted from protein-coding genes in the NCBI datasets (https://www.ncbi.nlm.nih.gov/datasets/; accessed on 18th January). The gene symbols were capitalized as normalization to extract the common shared genes by generating a Venn diagram in the exact match manner (https://bioinformatics.psb.ugent.be/webtools/Venn/).

Given another round of whole-genome duplication in teleosts, some genes did not correspond between humans and fishes in a 1:1 manner, and two orthologues could be present in the teleost fishes. To extract such duplicated genes in fishes, we first extracted fish genes that did not overlap with human genes in an exact match manner. Then, the gene symbols ending with A or B were picked up. After depleting the last letter, the genes with equal combinations of the remaining strings were considered to be a pair of duplicated genes. These extracted duplicated genes were examined to determine whether orthologues overlapped with human genes. This procedure was repeated through four species to find additional common shared genes.

Coverage of Common Shared Genes in Each Human Pathway

Common shared genes were applied to the Reactome pathway database (Jassal et al., 2020) to determine the coverage of genes in each human pathway.

Extraction of Orphanet Codes That Contain Common Shared Genes

Metascape (Zhou et al., 2019) was utilized by applying the Orphanet database (https://www.orpha.net/consor/cgi-bin/index.php). The list of ORPHAcodes related to each common shared gene was obtained, and the list was organized according to the list with the index of each ORPHAcode.

RESULTS

Common Shared Genes Among Humans and three Fishes

Approximately 500 million years ago, vertebrates, including humans, experienced a whole-genome duplication in which the genome doubled twice in our ancestors (Ohno, 1970; Dehal and Boore, 2005). Another round of whole-genome duplication occurred in teleosts, including zebrafish, medaka, and turquoise killifish (Chiu et al., 2004; Hoegg et al., 2004; Jaillon et al., 2004). This is one of the most significant genetic differences between humans and fishes. Among these three fishes, sex chromosomes have not been identified in zebrafish. Zebrafish sex determinants remain unclear, but environmental factors are known to affect zebrafish sex determination (Baroiller and D'Cotta, 2001; Orn et al., 2003; Abozaid et al., 2012). Similar to humans, medaka or turquoise killifish sex is determined by XX/XY sex chromosomes (Schartl, 2004; Valenzano et al., 2009; Reichwald et al., 2015), which is another major genomic difference among the four species.

We attempted to extract the gene symbols of each species from public databases, such as the UniProt (UniProt, 2021), Ensembl genome browser (Howe et al., 2021), and NCBI datasets, and

TABLE 1 | Lifespans and genomic characteristics of humans, zebrafish, medaka, and turquoise killifish.

	Human	Zebrafish	Medaka	Turquoise killifish
Scientific name	*Homo sapiens*	*Dania rerio*	*Oryzias latipes*	*Nothobranchius furzeri*
Life span	100 years	2–5 years	2–5 years	4–6 months
Genome size	3100 Mbp	1373 Mbp	734 Mbp	1242 Mbp
Number of chromosomes	23 chromosomes (2n = 46)	25 chromosomes (2n-50)	23–24 chromosomes (2n-46–48)	19 chromosomes (2n-38)
Sex determination	XX/XY	Environmental?	XX/XY	XX/XY
Reference genome	**GRCh38.p13**	**GCRz11**	**ASM223467v1**	**Nfu_20140520**
Protein-coding genes	19,671	29,961	22,140	22,207
Small RNAs	1,227	3,899	1,223	89
Pseudogenes	16,570	329	188	324
Noncoding	17,654	9,205	3,270	2,515
Other	8,028	304	105	42

The genomic characteristics of four species were extracted from the NCBI genome datasets (https://www.ncbi.nlm.nih.gov/datasets/; accessed on 18th January).

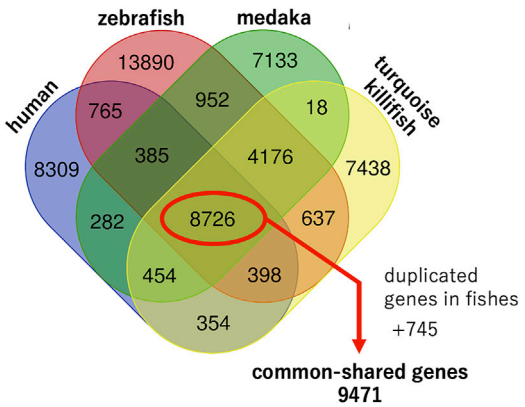


FIGURE 1 | Genomic characteristics of humans, zebrafish, medaka, and turquoise killifish. Venn diagram generated by an exact match with gene symbols normalized as strings. In addition to 8,726 common shared genes, duplicated orthologues in fishes were extracted and an additional 745 common shared genes were found. In total, 9,471 common shared genes were found among the four species.

found that the NCBI datasets contained the most gene symbols of each species. Therefore, we extracted the gene symbols of each species from the NCBI datasets (https://www.ncbi.nlm.nih.gov/datasets/; accessed on 18 January) based on the protein-coding genes. The numbers of protein-coding genes were 19,671; 29,961; 22,140; and 22,207 for human, zebrafish, medaka, and turquoise killifish, respectively (**Table 1**; **Supplementary Table S1**). A Venn diagram was generated by exact matching with gene symbols normalized as strings (https://bioinformatics.psb.ugent.be/webtools/Venn/), and 8,726 genes were found to be common among the four species (**Figure 1**).

As mentioned previously, teleost fishes experienced an additional round of whole-genome duplication compared to humans, and so, some genes did not correspond one-to-one between humans and fish. In such cases, two orthologues could be present in the teleost fishes. We extracted such duplicated orthologues from fishes and found an additional 745 common shared genes among four species out of a total of 9,471 genes (**Supplementary Table S2**). A study comparing short- and long-lived turquoise killifish strains showed that long-lived killifish acquired specific SNPs in several aging genes (*grn*, *tnfb*, *pdgfr*, *brca1*, *tp53*, *bp*, *ercc6*, *ghr*, *irs4*, *foxo4*, *myc*, *egr1*, *med1*, *ncor1*, *polg*, *gsr*, *mgat5*, *tert*, *hsf1*, and *hspa9*) under positive selection (Valenzano et al., 2015). Among these 19 genes, 16 genes were included in the common shared genes except for *brca1*, *irs4*, and *tnfb*.

Human Pathways and Diseases With Common-Shared Genes

Then, the 9,471 common shared genes were applied to the Reactome Pathway Database (Jassal et al., 2020), a web tool for human pathway analysis, to determine the proportion and number of common shared genes in known human biological pathways. A total of 2,405 pathways were found to be associated with at least one of the common shared genes. Given the list of 2,405 pathways, we extracted some of the senescence-related pathways and noticed that common shared genes covered approximately 30–50% of the genes in the human senescence-related pathways (**Table 2**). Complete gene lists in each pathway are available in **Supplementary Table S3**. Next, we applied the 9,471 common shared genes to the Orphanet database (https://www.orpha.net/consor/cgi-bin/index.php) (Weinreich et al., 2008) to establish the human disease list, which included at least one gene from the common shared genes. From this application, 2,677 Orphacodes were extracted. We extracted several human age-related disorders, such as Alzheimer's disease, Parkinson's disease, frontotemporal dementia, nonalcoholic fatty liver disease, and progeria, from our Orphanet disease list containing common shared genes. We also extracted hepatocellular carcinoma and renal cell carcinoma because turquoise killifish are known to develop hepatocellular carcinoma and renal cell carcinoma at high rates despite their short lifespan (Di Cicco et al., 2011) (**Table 3**; **Supplementary Table S4**).

Based on these lists obtained previously, several age-related disease genes were explored. *APOE* is known to be associated with Alzheimer's disease and is located in the lifespan-related loci of the turquoise killifish (Kirschner et al., 2012). *APOE* was also reported as an aging marker of the short-lived fish *N. guentheri* (Wang et al., 2014). *PARK7* is one of the causative genes of

TABLE 2 | Senescence pathway from the common shared genes.

Pathways identifier	Senescence-related pathways	#Entities found	#Entities total	#Interactors found	#Interactors total
R-MSA-2559583	Cellular senescence	87	200	335	662
R-HSA-2559580	Oxidative stress–induced senescence	40	114	178	357
R-HSA-2559582	Senescence-associated secretory phenotype (SASP)	39	91	23	48
R-H5A-2559586	DNA damage/telomere stress-induced senescence	19	71	135	248
R-H SA-2558585	Oncogene-induced senescence	23	42	79	154
R-HSA-2559584	Formation of senescence-associated heterochromatic foci (SAHF)	5	17	59	116
R-HSA-9630747	Disease of cellular senescence	2	4	22	38

*The pathway searched with "senescence." The search word "senescence" was applied to the pathway list of common shared genes (**Supplementary Table S3**), and part of the results is presented as an example.*

familial Parkinsonism, and polymorphisms in *PARK7* were reported to be related to the lifespan of turquoise killifish (Genade and Wilcox, 2021). *GRN*, one of the causative genes of frontotemporal dementia, is known to be regulated downstream of *PARK7* in the context of neuroprotection (Genade and Wilcox, 2021). A common biological pathway linking *PINK–PARKIN–PARK7* was reported in the pathogenesis of Parkinson's disease (van der Merwe et al., 2015), and these three genes were observed among the common shared genes. A search on Orphanet for "Parkinson" seems to cover the common shared genes well, except for hereditary late-onset Parkinson's disease (ORPHA: 411602). Referring to the gene list of turquoise killifish, *dnaj13*, *eif4g1*, and *vsp35* were found, but the orthologous genes corresponding to human *SNCA* and human *HTRA2* were not found. Although *snca* was not included in the gene list of turquoise killifish constructed with NCBI datasets, *snca* is known to be expressed in turquoise killifish (Matsui et al., 2019) and was found in turquoise killifish in the UniProt database (UniProt, 2021). Thus, turquoise killifish possesses most of the Parkinson's disease-related genes, which could be applied to translational research on Parkinson's disease. Compared to age-related diseases, common shared genes only covered 30–50% of the genes in senescence-related pathways. This difference in gene coverage may be due to the relatively small number of genes in common or to genomic differences that occur during evolution. In interpreting the effects of aging on the pathogenesis of age-related diseases using fish as a model, it will be necessary to keep in mind that the effects may be species dependent.

DISCUSSION

In this article, we explored the genetic background and common shared genes among humans and three fishes by using the available gene symbol data from the NCBI database (https://www.ncbi.nlm.nih.gov/datasets/). We applied common shared genes to the Reactome pathway database (Jassal et al., 2020) to determine the coverage of genes in human pathways. To determine the involvement of commonly shared genes in human diseases, we utilized the Orphanet database (Weinreich et al., 2008). These genetic background data between humans and fishes are important and worth considering before translating the previous findings of fish models in the context of human disease studies, including aging studies. Additionally, these data allow us to design an experiment that applies a fish model including turquoise killifish in aging studies. Although this kind of knowledge-based approach allows us to observe the genomic landscape from a broad perspective in an interspecies manner, several caveats and limitations exist.

First, in the available database, even among protein-coding genes, fairly many genes are still waiting to be annotated in many species. As observed in turquoise killifish, *snca* was not obtained from NCBI datasets. A blastp search (Altschul et al., 1990; States and Gish, 1994; Boratyn et al., 2012) of the amino acid sequence of human *SNCA* based on the turquoise killifish protein dataset yielded the answer "PREDICTED: alpha-synuclein-like" (**Supplementary Table S5**). This might be the reason why the *Snca* protein of turquoise killifish has not been annotated and could not be found in the list of gene symbols in the NCBI dataset. There are many gene symbols beginning with LOC, a notation that indicates that a published symbol is not available for this gene, and orthologues have not yet been determined (https://www.ncbi.nlm.nih.gov/books/NBK3840/#genefaq.Conventions). The numbers of gene symbols beginning with LOC are 266; 4,473; 6,924; and 7,332 for human, zebrafish, medaka, and turquoise killifish, respectively. This caveat of a lack of annotation should be kept in mind when attempting to explore genomic information.

Second, pseudogenes, miRNAs, noncoding RNAs, and others would be worth investigating (Esteller, 2011; Cheetham et al., 2020; Statello et al., 2021); we did not explore such RNAs in an interspecies manner in this study. It is often difficult to obtain sufficient insight by comparing such genes in an interspecies manner. Thus, some meaningful information may be overlooked.

Third, whole-genome alignment is suitable for covering the entire genomes of multiple species. However, whole-genome comparison studies among multiple species require long computation times, complex algorithms, and expensive computational resources and are difficult for researchers not familiar with bioinformatics to repeat in a timely manner (Armstrong et al., 2019). If the target gene has already been determined, blastp (Altschul et al., 1990; States and Gish, 1994; Boratyn et al., 2012) is an effective way to evaluate coverage, e-values, and percent identity by using human sequence data as a reference. It should be noted that a gene may have multiple amino acid sequences, and a single amino acid sequence may be annotated to multiple genes in the protein dataset of the reference organism. For example, medaka and turquoise killifish Tp53 amino acid sequences

TABLE 3 | List of genes related to some age-related disorders.

ORPHAcode	Search with "Alzheimer"	Genes in common shared genes (not in common shared genes)
ORPHA:1020	Early-onset autosomal dominant Alzheimer's disease	*TOMM40*, *ABCA7*, *PSEN2*, *SOL1*, *PSEN1*, *(APP*, *TREM2)*
ORPHA:238616	Non-rare in Europe: Alzheimer's disease	*ABCA7*, *SORL1*, *APOE*, *(TREM2)*
ORPHAcode	**Search with "Parkinson"**	**Genes in common shared genes (not in common shared genes)**
ORPHA:2,828	Young-onset Parkinson disease	*DNAJC6*, *PARK7*, *LRRK2*, *UCHL1*, *PODX1*, *SYNJ1*, *VSP13C*, *PINK1*, *PRKN*, *(HTRA2*, *SNCA)*
ORPHA:53351	X-linked dystonia-Parkinsonism	*TAF1*
ORPHA:71517	Rapid-onset dystonia-Parkinsonism	*ATP1A3*
ORPHA:90020	Parkinson–dementia complex of Guam	*PARK7*, *TPRM7*
ORPHA:98933	Multiple system atrophy Parkinsonian type	*COQ2*
ORPHA:171695	Parkinsonian-pyramidal syndrome	*FBXO7*, *(SNCA)*
ORPHA:199351	Adult-onset dystonia-Parkinsonism	*PLA2G6*
ORPHA:238455	Infantile dystonia-parkinsonism	*SLC6A3*
ORPHA:319705	Non-rare in Europe: Parkinson's disease	*GBA*
ORPHA:363654	X-linked parkinsonism-spasticity syndrome	*ATP6P2*
ORPHA:391411	Atypical juvenile parkinsonism	*DNAJC6*, *PODXL*, *SYNJ1*
ORPHA:411602	Hereditary late-onset Parkinson disease	*LRK2*, *GIGYF2*, *GBA*, (*DNAJC13*, *EIF4G1*, *SNCA*, *VPS3S*)
ORPHA:521406	Dystonia-Parkinsonism-hypermanganesemia syndrome	*SLC39A14*
ORPHAcode	**Search with "frontotemporal dementia"**	**Genes in common shared genes (not in common shared genes)**
ORPHA:52430	Inclusion of body myopathy with Paget disease of bone and frontotemcioral dementia	*VCP*, *(HNRNPA1*, *HNROA2B1)*
ORPHA:275864	A behavioral variant of frontotemporal dementia	*VCP*, *SQSTM1*, *GRN*, *PSEN1*, *(O9ORF72*, *CHMP2B*, *MAPT*, *TMEM106B*, *TPEM2)*
ORPHA:275872	Frontotemporal dementia with motor neuron disease	*VCP*, *SQSTM1*, *TBK1*, *TARDBP (O9ORF72*, *CHCD10*, *FUS)*
ORPHAcode	**Search with "Non-alcoholic fatty liver disease"**	**Genes in common-shared genes (not in common-shared genes)**
ORPHA:33271	Non-rare in Europe: Non-alcoholic fatty liver disease	*PNPLA3*, (*APOC3*)
ORPHAcode	**Search with "Progeria"**	**Genes in common shared genes (not in common shared genes)**
ORPHA:740	Hutchinson–Gilford progeria syndrome	*LMNA*, *ZMPSTE24*
ORPHA:280576	Nestor–Guillermo progeria syndrome	*BANF1*
ORPHA:363618	LMNA-related cardiocutaneous progeria syndrome	*LMNA*
ORPHAcode	**Search with "hepatocellular carcinoma"**	**Genes in common shared genes (not in common shared genes)**
ORPHA:33402	Pediatric hepatocellular carcinoma	*CNNTB1*, *MET*
ORPHA:210159	Adult hepatocellular carcinoma	*PIK3CA*, *PDGFRL*, *AXIN1*, *CTNNB1*, *EGF*, *TSC2*, *CASP8*, *TP53*, (*TSC1*)
ORPHA:435953	Proseroid features—hepatocellular carcinoma predisposition syndrome	*SPRTN*
ORPHAcode	**Search with "renal cell carcinoma"**	**Genes in common shared genes (not in common shared genes)**
ORPHA:47044	Hereditary papillary renal cell carcinoma	*MET*
ORPHA:319294	Papillary renal cell carcinoma	*MET*, *(MITF)*
ORPHA:319393	Chronophage renal cell carcinoma	*HNF1A*, *(MITF*, *PBRW1*, *TMEM127)*
ORPHA:319308	MiT family translocation renal cell carcinoma	*ASPSCRl*, *TFE3*, *SFPQ*, *TFEB*, *PRCC*, *(CLTC*, *NONO*)
ORPHA:404511	Clear-cell papillary renal cell carcinoma	*HNF1F*, *PBRM1*, (*MITF*, *TMEM127*)
ORPHA:422526	Hereditary clear-cell renal cell carcinoma	*OGG1*, *SLC49A4*, *RNF139*, *HSPBAP1*, *FLCN*, (*DIR3*, *FHIT*)

*The list of age-related disorders from Orphanet according to common shared genes. The search words "Alzheimer," "Parkinson," "frontotemporal dementia," "nonalcoholic," "progeria," "hepatocellular carcinoma," and "renal cell carcinoma" were applied to the list of Orphanet codes, which included at least one common shared gene (**Supplementary Table S4**). Search results are presented as age-related disorders. Gene symbols enclosed in parentheses () are disease-related genes that were not found in the common shared gene.*

showed higher homology to human TP63 and human TP73 (TP53 family proteins) (Belyi et al., 2010) than to human TP53 (**Supplementary Table S6**).

Finally, in the disease-related genes, we only focused on Orphanet (Weinreich et al., 2008); genes potentially involved in disease modification were not included. Such regulatory genes would be better to be included when we consider applying an animal model for specific diseases. In addition, there are many variants of uncertain significance (VUS) (Elliott, 2020; Sullivan, 2021) in human genes, and their evaluation is also rapidly progressing (Mahecha et al., 2022; Postel et al., 2022). Thus, the identification of disease-related genes will be increasing, and their significance will be determined in the future.

Given the current situation, timely updates with recent genomics data would be ideal to translate and interpret the data and phenomena with model animals. A goal should be to establish a platform where anyone can easily compare genomes between species, even if they cannot write the code for analysis.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncbi.nlm.nih.gov/datasets/.

AUTHOR CONTRIBUTIONS

ED analyzed data, conducted bioinformatic analysis, and co-wrote the manuscript. HM supervised the project, analyzed results, and co-wrote the manuscript. All authors have approved the final version of the manuscript.

FUNDING

This work was supported by grants from AMED (Grant Numbers JP22gm6110028 (HM)), JSPS KAKENHI (Grant Numbers JP 22484842, JP 18955907 (HM)), and JST (Moonshot R&D) (Grant Number JPMJMS 2024) (HM).

ACKNOWLEDGMENTS

We acknowledge Ms. Shinano Kobayashi and Ms. Noriko Matsui for participating in helpful discussions and providing continuous support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.928597/full#supplementary-material

REFERENCES

Ablain, J., Durand, E. M., Yang, S., Zhou, Y., and Zon, L. I. (2015). A CRISPR/Cas9 Vector System for Tissue-specific Gene Disruption in Zebrafish. *Dev. Cell.* 32 (6), 756–764. doi:10.1016/j.devcel.2015.01.032

Abozaid, H., Wessels, S., and Hörstgen-Schwark, G. (2012). Elevated Temperature Applied during Gonadal Transformation Leads to Male Bias in Zebrafish (*Danio rerio*). *Sex. Dev.* 6 (4), 201–209. doi:10.1159/000336297

Allard, J. B., Kamei, H., and Duan, C. (2013). Inducible Transgenic Expression in the Short-Lived fishNothobranchius Furzeri. *J. Fish. Biol.* 82 (5), 1733–1738. doi:10.1111/jfb.12099

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic Local Alignment Search Tool. *J. Mol. Biol.* 215 (3), 403–410. doi:10.1016/S0022-2836(05)80360-2

Armstrong, J., Fiddes, I. T., Diekhans, M., and Paten, B. (2019). Whole-Genome Alignment and Comparative Annotation. *Annu. Rev. Anim. Biosci.* 7, 41–64. doi:10.1146/annurev-animal-020518-115005

Baroiller, J. F., and D'Cotta, H. (2001). Environment and Sex Determination in Farmed Fish. *Comp. Biochem. Physiology Part C Toxicol. Pharmacol.* 130 (4), 399–409. doi:10.1016/s1532-0456(01)00267-8

Belyi, V. A., Ak, P., Markert, E., Wang, H., Hu, W., Puzio-Kuter, A., et al. (2010). The Origins and Evolution of the P53 Family of Genes. *Cold Spring Harb. Perspect. Biol.* 2 (6), a001198. doi:10.1101/cshperspect.a001198

Boratyn, G. M., Schäffer, A. A., Agarwala, R., Altschul, S. F., Lipman, D. J., and Madden, T. L. (2012). Domain Enhanced Lookup Time Accelerated BLAST. *Biol. Direct* 7, 12. doi:10.1186/1745-6150-7-12

Cheetham, S. W., Faulkner, G. J., and Dinger, M. E. (2020). Overcoming Challenges and Dogmas to Understand the Functions of Pseudogenes. *Nat. Rev. Genet.* 21 (3), 191–201. doi:10.1038/s41576-019-0196-1

Chiu, C.-h., Dewar, K., Wagner, G. P., Takahashi, K., Ruddle, F., Ledje, C., et al. (2004). Bichir HoxA Cluster Sequence Reveals Surprising Trends in Ray-Finned Fish Genomic Evolution. *Genome Res.* 14 (1), 11–17. doi:10.1101/gr.1712904

Cui, R., Medeiros, T., Willemsen, D., Iasi, L. N. M., Collier, G. E., Graef, M., et al. (2019). Relaxed Selection Limits Lifespan by Increasing Mutation Load. *Cell.* 178 (2), 385–399. doi:10.1016/j.cell.2019.06.004

Dehal, P., and Boore, J. L. (2005). Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. *PLoS Biol.* 3 (10), e314. doi:10.1371/journal.pbio.0030314

Delomas, T. A., and Dabrowski, K. (2018). Larval Rearing of Zebrafish at Suboptimal Temperatures. *J. Therm. Biol.* 74, 170–173. doi:10.1016/j.jtherbio.2018.03.017

Di Cicco, E., Tozzini, E. T., Rossi, G., and Cellerino, A. (2011). The Short-Lived Annual Fish Nothobranchius Furzeri Shows a Typical Teleost Aging Process Reinforced by High Incidence of Age-dependent Neoplasias. *Exp. Gerontol.* 46 (4), 249–256. doi:10.1016/j.exger.2010.10.011

Doyon, Y., McCammon, J. M., Miller, J. C., Faraji, F., Ngo, C., Katibah, G. E., et al. (2008). Heritable Targeted Gene Disruption in Zebrafish Using Designed Zinc-Finger Nucleases. *Nat. Biotechnol.* 26 (6), 702–708. doi:10.1038/nbt1409

Elliott, A. M. (2020). Genetic Counseling and Genome Sequencing in Pediatric Rare Disease. *Cold Spring Harb. Perspect. Med.* 10 (3), a036632. doi:10.1101/cshperspect.a036632

Esteller, M. (2011). Non-coding RNAs in Human Disease. *Nat. Rev. Genet.* 12 (12), 861–874. doi:10.1038/nrg3074

Furutani-Seiki, M., and Wittbrodt, J. (2004). Medaka and Zebrafish, an Evolutionary Twin Study. *Mech. Dev.* 121 (7-8), 629–637. doi:10.1016/j.mod.2004.05.010

Genade, T., Benedetti, M., Terzibasi, E., Roncaglia, P., Valenzano, D. R., Cattaneo, A., et al. (2005). Annual Fishes of the Genus Nothobranchius as a Model System for Aging Research. *Aging Cell.* 4 (5), 223–233. doi:10.1111/j.1474-9726.2005.00165.x

Genade, T., and Wilcox, D. A. (2021). Nothobranchius Fish: An Untapped Resource for Studying Aging-Related Neurodegeneration. *Cas* 14 (2), 78–93. doi:10.2174/1874609814666210202091301

Harel, I., Benayoun, B. A., Machado, B., Singh, P. P., Hu, C.-K., Pech, M. F., et al. (2015). A Platform for Rapid Exploration of Aging and Diseases in a Naturally Short-Lived Vertebrate. *Cell.* 160 (5), 1013–1026. doi:10.1016/j.cell.2015.01.038

Harel, I., Valenzano, D. R., and Brunet, A. (2016). Efficient Genome Engineering Approaches for the Short-Lived African Turquoise Killifish. *Nat. Protoc.* 11 (10), 2010–2028. doi:10.1038/nprot.2016.103

Hartmann, N., and Englert, C. (2012). A Microinjection Protocol for the Generation of Transgenic Killifish (Species: Nothobranchius Furzeri). *Dev. Dyn.* 241 (6), 1133–1141. doi:10.1002/dvdy.23789

Hoegg, S., Brinkmann, H., Taylor, J. S., and Meyer, A. (2004). Phylogenetic Timing of the Fish-specific Genome Duplication Correlates with the Diversification of Teleost Fish. *J. Mol. Evol.* 59 (2), 190–203. doi:10.1007/s00239-004-2613-z

Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., et al. (2021). Ensembl 2021. *Nucleic Acids Res.* 49 (D1), D884–D891. doi:10.1093/nar/gkaa942

Huang, P., Xiao, A., Zhou, M., Zhu, Z., Lin, S., and Zhang, B. (2011). Heritable Gene Targeting in Zebrafish Using Customized TALENs. *Nat. Biotechnol.* 29 (8), 699–700. doi:10.1038/nbt.1939

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., et al. (2013). Efficient Genome Editing in Zebrafish Using a CRISPR-Cas System. *Nat. Biotechnol.* 31 (3), 227–229. doi:10.1038/nbt.2501

Irion, U., Krauss, J., and Nüsslein-Volhard, C. (2014). Precise and Efficient Genome Editing in Zebrafish Using the CRISPR/Cas9 System. *Development* 141 (24), 4827–4830. doi:10.1242/dev.115584

Jaillon, O., Aury, J.-M., Brunet, F., Petit, J.-L., Stange-Thomann, N., Mauceli, E., et al. (2004). Genome Duplication in the Teleost Fish *Tetraodon nigroviridis* Reveals the Early Vertebrate Proto-Karyotype. *Nature* 431 (7011), 946–957. doi:10.1038/nature03025

Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., et al. (2020). The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* 48 (D1), D498–D503. doi:10.1093/nar/gkz1031

Kirchmaier, S., Naruse, K., Wittbrodt, J., and Loosli, F. (2015). The Genomic and Genetic Toolbox of the Teleost Medaka (*Oryzias latipes*). *Genetics* 199 (4), 905–918. doi:10.1534/genetics.114.173849

Kirschner, J., Weber, D., Neuschl, C., Franke, A., Böttger, M., Zielke, L., et al. (2012). Mapping of Quantitative Trait Loci Controlling Lifespan in the Short-lived fishNothobranchius Furzeri- a New Vertebrate Model for Age Research. *Aging Cell.* 11 (2), 252–261. doi:10.1111/j.1474-9726.2011.00780.x

Mahecha, D., Nuñez, H., Lattig, M. C., and Duitama, J. (2022). Machine Learning Models for Accurate Prioritization of Variants of Uncertain Significance. *Hum. Mutat.* 43 (4), 449–460. doi:10.1002/humu.24339

Matsui, H., Kenmochi, N., and Namikawa, K. (2019). Age- and α-Synuclein-Dependent Degeneration of Dopamine and Noradrenaline Neurons in the Annual Killifish Nothobranchius Furzeri. *Cell. Rep.* 26 (7), 1727–1733. e1726. doi:10.1016/j.celrep.2019.01.015

Meng, X., Noyes, M. B., Zhu, L. J., Lawson, N. D., and Wolfe, S. A. (2008). Targeted Gene Inactivation in Zebrafish Using Engineered Zinc-Finger Nucleases. *Nat. Biotechnol.* 26 (6), 695–701. doi:10.1038/nbt1398

Ohno, S. (1970). *Evolution by Gene Duplication*. Berlin, Heidelberg: Springer-Verlag.

Örn, S., Holbech, H., Madsen, T. H., Norrgren, L., and Petersen, G. I. (2003). Gonad Development and Vitellogenin Production in Zebrafish (*Danio rerio*) Exposed to Ethinylestradiol and Methyltestosterone. *Aquat. Toxicol.* 65 (4), 397–411. doi:10.1016/s0166-445x(03)00177-2

Poeschla, M., and Valenzano, D. R. (2020). The Turquoise Killifish: a Genetically Tractable Model for the Study of Aging. *J. Exp. Biol.* 223 (Pt Suppl. 1). doi:10.1242/jeb.209296

Polačik, M., Blažek, R., and Reichard, M. (2016). Laboratory Breeding of the Short-Lived Annual Killifish Nothobranchius Furzeri. *Nat. Protoc.* 11 (8), 1396–1413. doi:10.1038/nprot.2016.080

Postel, M. D., Culver, J. O., Ricker, C., and Craig, D. W. (2022). Transcriptome Analysis Provides Critical Answers to the "variants of Uncertain Significance" Conundrum. *Hum. Mutat.* doi:10.1002/humu.24394

Prykhozhij, S. V., Fuller, C., Steele, S. L., Veinotte, C. J., Razaghi, B., Robitaille, J. M., et al. (2018). Optimized Knock-In of Point Mutations in Zebrafish Using CRISPR/Cas9. *Nucleic Acids Res.* 46 (17), e102. doi:10.1093/nar/gky512

Reichard, M., Polačik, M., and Sedláček, O. (2009). Distribution, Colour Polymorphism and Habitat Use of the African killifishNothobranchius Furzeri, the Vertebrate with the Shortest Life Span. *J. Fish. Biol.* 74 (1), 198–212. doi:10.1111/j.1095-8649.2008.02129.x

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 (6), 1527–1538. doi:10.1016/j.cell.2015.10.071

Sander, J. D., Cade, L., Khayter, C., Reyon, D., Peterson, R. T., Joung, J. K., et al. (2011). Targeted Gene Disruption in Somatic Zebrafish Cells Using Engineered TALENs. *Nat. Biotechnol.* 29 (8), 697–698. doi:10.1038/nbt.1934

Schartl, M. (2004). A Comparative View on Sex Determination in Medaka. *Mech. Dev.* 121 (7-8), 639–645. doi:10.1016/j.mod.2004.03.001

Skinner, A. M. J., and Watt, P. J. (2007). Strategic Egg Allocation in the Zebra Fish, *Danio rerio*. *Behav. Ecol.* 18 (5), 905–909. doi:10.1093/beheco/arm059

Statello, L., Guo, C.-J., Chen, L.-L., and Huarte, M. (2021). Gene Regulation by Long Non-coding RNAs and its Biological Functions. *Nat. Rev. Mol. Cell. Biol.* 22 (2), 96–118. doi:10.1038/s41580-020-00315-9

States, D. J., and Gish, W. (1994). QGB: Combined Use of Sequence Similarity and Codon Bias for Coding Region Identification. *J. Comput. Biol.* 1 (1), 39–50. doi:10.1089/cmb.1994.1.39

Sullivan, K. E. (2021). The Scary World of Variants of Uncertain Significance (VUS): A Hitchhiker's Guide to Interpretation. *J. Allergy Clin. Immunol.* 147 (2), 492–494. doi:10.1016/j.jaci.2020.06.011

UniProt, C. (2021). UniProt: the Universal Protein Knowledgebase in 2021. *Nucleic Acids Res.* 49 (D1), D480–D489. doi:10.1093/nar/gkaa1100

Valenzano, D. R., Benayoun, B. A., Singh, P. P., Zhang, E., Etter, P. D., Hu, C.-K., et al. (2015). The African Turquoise Killifish Genome Provides Insights into Evolution and Genetic Architecture of Lifespan. *Cell.* 163 (6), 1539–1554. doi:10.1016/j.cell.2015.11.008

Valenzano, D. R., Kirschner, J., Kamber, R. A., Zhang, E., Weber, D., Cellerino, A., et al. (2009). Mapping Loci Associated with Tail Color and Sex Determination in the Short-Lived Fish Nothobranchius Furzeri. *Genetics* 183 (4), 1385–1395. doi:10.1534/genetics.109.108670

Valenzano, D. R., Sharp, S., and Brunet, A. (2011). Transposon-Mediated Transgenesis in the Short-Lived African KillifishNothobranchius Furzeri, a Vertebrate Model for Aging. *G3 (Bethesda)* 1 (7), 531–538. doi:10.1534/g3.111.001271

Valenzano, D. R., Terzibasi, E., Cattaneo, A., Domenici, L., and Cellerino, A. (2006). Temperature Affects Longevity and Age-Related Locomotor and Cognitive Decay in the Short-Lived Fish Nothobranchius Furzeri. *Aging Cell.* 5 (3), 275–278. doi:10.1111/j.1474-9726.2006.00212.x

van der Merwe, C., Jalali Sefid Dashti, Z., Christoffels, A., Loos, B., and Bardien, S. (2015). Evidence for a Common Biological Pathway Linking Three Parkinson's Disease-Causing genes:parkin,PINK1andDJ-1. *Eur. J. Neurosci.* 41 (9), 1113–1125. doi:10.1111/ejn.12872

Varshney, G. K., Pei, W., LaFave, M. C., Idol, J., Xu, L., Gallardo, V., et al. (2015). High-throughput Gene Targeting and Phenotyping in Zebrafish Using CRISPR/Cas9. *Genome Res.* 25 (7), 1030–1042. doi:10.1101/gr.186379.114

Wang, X., Shang, X., Luan, J., and Zhang, S. (2014). Identification, Expression and Function of Apolipoprotein E in Annual Fish Nothobranchius Guentheri: Implication for an Aging Marker. *Biogerontology* 15 (3), 233–243. doi:10.1007/s10522-014-9493-4

Weinreich, S. S., Mangon, R., Sikkens, J. J., Teeuw, M. E., and Cornel, M. C. (2008). Orphanet: a European Database for Rare Diseases. *Ned. Tijdschr. Geneeskd.* 152 (9), 518–519.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat. Commun.* 10 (1), 1523. doi:10.1038/s41467-019-09234-6

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

TYPE Brief Research Report
PUBLISHED 07 September 2022
DOI 10.3389/fgene.2022.903201

OPEN ACCESS

EDITED BY
Wei Xu,
Texas A&M University Corpus Christi, United States

REVIEWED BY
Huaping Zhu,
Pearl River Fisheries Research Institute (CAFS), China
David Veliz,
Universidad de Chile, Chile

*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 24 March 2022
ACCEPTED 26 July 2022
PUBLISHED 07 September 2022

CITATION
Pagano AD, Barreto BF, Domingues WB, Silveira TLR, Nunes LS, Blodorn EB, Dellagostin EN, Remião MH, Robaldo RB and Campos VF (2022), Modulation of miR-429 during osmotic stress in the silverside Odontesthes humensis. *Front. Genet.* 13:903201. doi: 10.3389/fgene.2022.903201

Modulation of miR-429 during osmotic stress in the silverside *Odontesthes humensis*

Antônio D. Pagano[1†], Bruna F. Barreto[1†], William B. Domingues[1], Tony L. R. Silveira[2], Leandro S. Nunes[1], Eduardo B. Blodorn[1], Eduardo N. Dellagostin[1], Mariana H. Remião[1], Ricardo B. Robaldo[3] and Vinicius F. Campos[1]*

[1]Laboratório de Genômica Estrutural, Programa de Pós-Graduação em Biotecnologia, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, Rio Grande do Sul, Brazil, [2]Instituto de Ciências Biológicas, Universidade Federal do Rio Grande, Rio Grande, Rio Grande do Sul, Brazil, [3]Instituto de Biologia, Universidade Federal de Pelotas, Pelotas, Rio Grande do Sul, Brazil

Silverside fish inhabit marine coastal waters, coastal lagoons, and estuarine regions in southern South America. Although silversides are not fully adapted to freshwater, they can tolerate a wide range of salinity variations. MicroRNAs (miRNAs) are a class of ~22 nucleotide noncoding RNAs, which are crucial regulators of gene expression at post-transcriptional level. Current data indicate that miRNAs biogenesis is altered by situations of environmental stress, thereby altering the expression of target mRNAs. Foremost, the silversides were acutely exposed to 30 $g.L^{-1}$ of salt to reveal in which tissue miR-429 could be differentially expressed. Thus, fish were acclimated to freshwater (0 $g.L^{-1}$) and to brackish water (10 $g.L^{-1}$), and then exposed to opposite salinity treatment. Here, we reveal that miR-429, a gill-enriched miRNA, emerges as a prime osmoregulator in silversides. Taken together, our findings suggest that miR-429 is an endogenous regulator of osmotic stress, which may be developed as a biomarker to assist silverside aquaculture.

KEYWORDS

microRNA, osmoregulation, pejerrey, aquaculture, salinity

Introduction

Salinity is one of the main abiotic properties that determine the distribution of fishes across aquatic environments. When faced with salinity variations, to maintain the osmotic balance, fishes depend on osmotically-induced responses, and systemic endocrine signaling induced by ionic transport in the gills, to salt-stress adaptation (Seale and Breves, 2022). In this sense, *Odontesthes* spp., which are popularly known as pejerreyes or silversides, form the most diverse genus in the Atherinopsidae group, with species inhabiting marine coastal waters, coastal lagoons, and estuarine regions in southern South America (Dyer, 2006). Currently, even though all *Odontesthes* spp. have a common marine origin, some fishes from these species currently occupy freshwater (FW) environments (Campanella et al., 2015). Transitions between FW and marine habitats

are related with increased species richness. Indeed, FW adaptations are shaped by evolutionary patterns, and are essentially involved with osmoregulatory adjustments because FW and diadromous lineages tolerate high salinity (Mank and Avise, 2006).

The silverside *Odontesthes humensis* is an endemic species that inhabits coastal lagoons, especially from southern Brazil, Uruguay, and Argentina (Bemvenuti, 2006). More specifically, fish species that inhabit environments where there are continuous alterations in water salinity are regarded as euryhaline, and are therefore able to adapt to osmotic stress by several efficient mechanisms of osmosensing and osmotic stress signaling (Fiol and Kültz, 2007). Given that silversides occupy both marine and FW environments, they present this interesting ability to tolerate salt. Naturally, they constantly face variations in salinity in the environment. Thus, these fish have developed excellent mechanisms of osmotic adaptation and appear to be excellent bioindicators in environmental studies due to their demand for good water quality, surviving only in a narrow range of water parameters (Zebral et al., 2017; Silveira et al., 2018a; Martins et al., 2021).

In this regard, the FW silversides euryhalinity has a considerable application for the aquaculture of these species in estuarine regions and in brackish environments that are characterized by significant salinity variations (Piedras, 2009). Salinity is a frequent abiotic stressor that restricts fish growth and development, and favors the impairment of the existing macromolecules, such as proteins, mRNAs, DNA, and lipids (Bartel, 2009). Thus, when faced with salinity variations, fishes employ many physiological acclimations to rapidly respond to osmotic stress, such as the induction of molecular chaperones, rapid clearance of damaged macromolecules, growth arrest, and the alteration of gene expression of multiple genes that mediates osmotic stress tolerance (Fiol and Kültz, 2007; Richter and Haslbeck, 2010; Silveira et al., 2018b). This gene expression modulation to maintain the osmotic balance may be directly influenced by miRNAs (Yan et al., 2012a).

In this sense, miRNAs are a class of noncoding RNA molecules, 22–25 nucleotides in length, which negatively regulate gene expression at the post-transcriptional level (Makeyev and Maniatis, 2008). In the cytoplasm, after two-step cleavage of primary miRNA and precursor miRNA, mature miRNA is loaded by Argonaute 2 protein (AGO2), which compounds the miRNA-Induced Silencing Complex (miRISC). The miRNA guides this effector complex to target sites in the 3' untranslated region (UTR) of mRNAs. The silencing mechanism is based on the complementarity levels between miRNA in miRISC and mRNA, and can be cleavage of target mRNA or inhibition of its translation (Gebert and Macrae, 2019).

Based on miRNAs' regulatory function and rapid modulation during stresses, they could act as biomarkers for biotic and abiotic stress because adverse environmental situations, (e.g., osmotic stress) can shape the biogenesis of miRNAs, the expression of target mRNAs, and the miRNA-protein complexes activity (Leung and Sharp, 2010). More specifically, miR-429 is considered to be a biomarker of salinity because it participates in a complex regulatory circuit that is responsible for both osmolality and ionic balance of plasma, as previously shown in tilapia during exposition to salinity of 20 $g.L^{-1}$ (Yan et al., 2012b). Meanwhile, miR-429 has already been established to be a crucial osmotic regulator in euryhaline species. Therefore, we propose to analyze its relative expression during situations of environmental salinity.

This study aims to prospect miR-429 as an epigenetic biomarker of environmental salinity in the silverside *Odontesthes humensis* under osmotic changes. Here, miR-429 expression was evaluated at both acute and chronic exposure to salt. To accomplish this, fish were acclimated to FW and acutely exposed for 24 h to seawater (SW, 30 $g.L^{-1}$) to evaluate the responses of expression of miR-429 by RT-qPCR in the gills, brain, liver, and kidney. Then, after verifying where the differential expression was present in a challenge situation, the fish were submitted to a protocol of milder salinity alterations and with evaluation of miR-429 expression for 15 days. Thus, fish were acclimated to FW and brackish water (BW, 10 $g.L^{-1}$). They were then transferred to the opposite salinity treatment. We mainly discuss the role of miR-429 in stress tolerance and osmoregulation responses triggered by the osmotic changes.

Materials and methods

Animals and conditions

The silversides *O. humensis* used in this study came from eggs collected in nature, from a freshwater lagoon called Mirim Lagoon (Arroio Grande, Brazil 32°09′21.8″S 53°01′27.6″W). The eggs were transported to the laboratory and hatched in tanks. Fish were 1.5 years old and had a mean weight of 23.0 ± 19.3 g at the end of Experiment 1 and 27.3 ± 9.5 g at the end of Experiment 2. Therefore, the silversides were kept within an experimental room, maintained in 1,000 L cylindrical plastic tanks, and fed three times a day with commercial feed (Supra, 38% crude protein). The tank sides were opaque to reduce visual stress and the incidence of the natural luminosity was by the top. A natural photoperiod of 11 h light/13 h dark was applied. The dissolution of non-iodized sea salt in the water was performed to achieve the desired salinity levels. The acclimation period was 4 weeks for silversides of both experiments (described below). The water quality parameters of the experimental times did not differ statistically from those observed in the acclimation period, except for the salinity of the group exposed to different salinities.

The water of Experiment 1 was maintained with pH 7.00 ± 0.6; temperature 13.4 ± 1.8°C; dissolved oxygen 9.7 ± 0.6 $mg.L^{-1}$;

total ammonia levels lower than 0.6 ± 0.2 mg.L^{-1}; and salinity 3.2 ± 0.7 g.L^{-1} (exclusively on acclimation period). The water of Experiment 2 was maintained with pH 7.78 ± 0.03; temperature 20.01 ± 0.34°C; dissolved oxygen 7.8 ± 0.3 mg.L^{-1}; and total ammonia levels lower than 0.6 ± 0.2 mg.L^{-1}. At the experimental time, the animals continued to be fed three times a day and once a week the water of the tanks was renewed. The salinity levels were achieved through the dissolution of non-iodized salt in the water.

Experimental designs

Experiment 1: osmotic challenge and search of tissues where miR-429 is differentially expressed

First, a tissue screening was performed to identify where miR-429 would be modulated by the osmotic challenge. For this purpose, the silversides were separated in two groups and distributed in six tanks with three fish each (two groups; three tanks/group; three fish/tank). Both of the groups were acclimated according to what was already described. In the treated group, the salinity was then increased to 30 g.L^{-1} by adding hypersaturated water containing non-iodized sea salt in dissolution.

The control group had only water added to the tanks in the same amount as the salt-treated tanks. The fish were maintained in these conditions for 24 h until the sample collection. In total, 18 silverside fish were analyzed in Experiment 1.

Experiment 2: monitoring of miR-429 expression in the gills over time

After verifying that the gills were the tissues that showed differential expression of miR-429, other fish were submitted to Experiment 2. The experiment was performed in quadruplicate, totaling two groups and eight tanks with 12 fish each (two groups; four tanks/group; 12 fish/tank). Four tanks contained silversides acclimated to FW and four tanks contained silversides acclimated to BW. Subsequently, both FW and BW groups were transferred to opposite salinity treatment. Fish samples were collected at four different time points by collecting 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. In total, 96 silverside fish were analyzed at distinct times after salt treatment in Experiment 2.

Sample collection and RNA extraction

When captured, at all-time points, silversides were anesthetized by immersion in water with benzocaine at 50 mg.L^{-1}. While anesthetized, the fish were weighed and euthanized by medullary section and excision of the brain. In Experiment 1, the brain, gills, kidney, and liver from silversides of both groups were collected. In Experiment 2, only the gills of the silversides were collected. All of the experimental procedures were approved by the Ethics Committee on Animal Experimentation of the Federal University of Pelotas (Process no. 23110.007018/2015-85).

The tissues were preserved in liquid nitrogen (N2), and posterior stored at −80°C upon further RNA preparation. Total RNA isolation was performed using the RNeasy™ Mini Kit (Qiagen, United States), as described per the manufacturer's description. RNA samples were treated with DNA-Free™ Kit (Invitrogen™, United States) to remove DNA contamination. Subsequently, the RNA concentration and quality were measured by UV-light spectrophotometry using NanoVue™ Plus equipment (GE Healthcare Life Sciences, United States), and only samples displaying high purity ($OD_{260/280}$ ≥ 2.0 nm) were used in reverse transcription (RT) reactions. Additionally, RNA quality was measured on the 4200 TapeStation (Agilent Technologies, United States) through the TapeStation analysis application.

MicroRNA cDNA synthesis

The generation of cDNA was carried out by using the TaqMan MicroRNA reverse transcription kit (Applied Biosystems; Thermo Fisher Scientific, Inc., United States), following the manufacturer's protocol for microRNA reverse transcription. Each microRNA cDNA synthesis reaction contained 10 ng of extracted total RNA, 50 nM stem-looped RT primer, 1 × RT buffer, 0.25 mM each of dNTPs, 3.33 U/μl Multiscribe reverse transcriptase, and 0.25 U/μl RNase Inhibitor. The thermocycling conditions were as follows: 30 min at 16°C, 30 min at 42°C, and 5 min at 85°C. Thereafter, all samples were stored at −20°C until further use.

Expression of miR-429 analysis by two-step RT-qPCR

For the two-step RT-qPCR assay, amplification was carried out using sequence specific primers for miR-429 on the Agilent Mx3005P system (Agilent Technologies, United States). The 20 μl reaction included 1.33 μl RT product, 1 × TaqMan® Universal PCR Master Mix (catalog number 4324018, Applied Biosystems, United States), and 1 × TaqMan® MicroRNA assays (catalog number 4427975, Applied Biosystems, United States). The reactions were incubated in a 96-well optical plate at 95°C for 10 min, following by 40 cycles of 95°C for 15 s and 60°C for 1 min. Histone h3a mRNA (*h3a,* GenBank accession number KX060037) (Silveira et al., 2018a) was used as a reference

10.3389/fgene.2022.903201

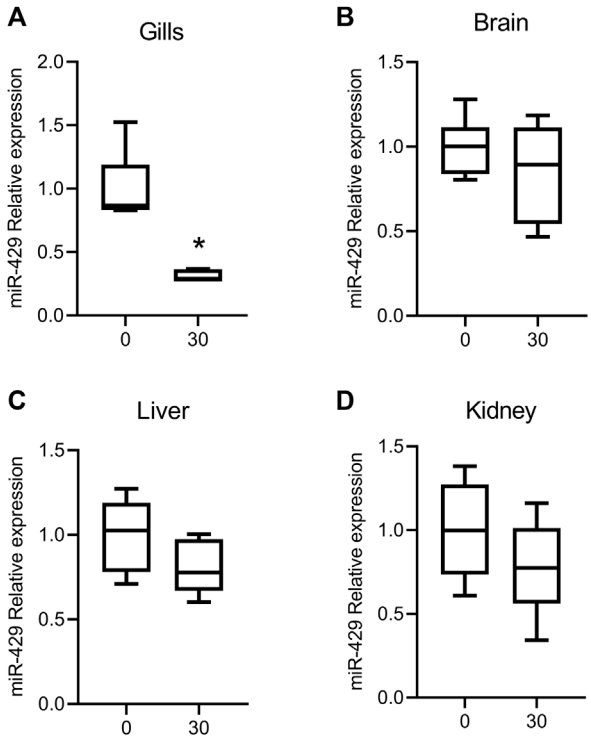


FIGURE 1
Relative expression of miR-429 in gills **(A)**, brain **(B)**, liver **(C)**, and kidney **(D)** from *Odontesthes humensis* in response to osmotic stress. The downregulation of miR-429 in gills was detected 24 h after 30 g.L^{-1} of salt using two-step RT-qPCR. The group exposed to 0 g.L^{-1} was taken as the control group. The box extends from the 25th to 75th percentiles, while the whiskers go down to the smallest value and up to the largest. The data was expressed as the relative change compared with the control group. Asterisk (*) indicates significant difference compared with the control group (*$p < 0.05$).

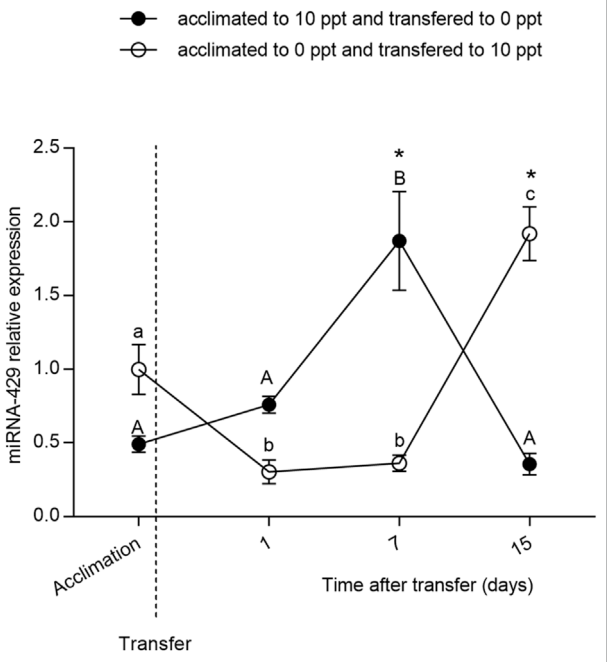


FIGURE 2
Relative expression of miR-429 from *Odontesthes humensis* submitted to salinity transfer. Different letters in each group (lowercase for freshwater (FW) to brackish water (BW) transferred group and uppercase for BW–FW transferred group) represent significant differences between the measurements in the same group. The asterisk (*) represents the significant difference between groups in a same experimental time (*$p < 0.05$). The vertical broken line indicates the transfer from BW to FW and from FW to BW moments.

to obtain the relative fold-change in miR-429 expression in target samples using the $2^{-\Delta\Delta Ct}$ method.

Statistical analysis

The Shapiro-Wilk test was used to test the normality of the data. The homogeneity of variances was evaluated by Levene's test, O'Brien's test, and the Brown and Forsy test. We analyzed the effect of tanks on miR-429 expression results using a factorial analysis of variance test and did not find a significant effect ($p > 0.05$) of tanks in relative expression results. When the data did not show normal distribution or homogeneous variances, data transformation was performed using the Log or Box-Cox methods, with the aim of meeting the requirements for the use of parametric tests, but without success. Therefore, the miR-429 expression in the different tissues from fish used in Experiment 1 was evaluated by Mann-Whitney test. The miR-429 expression in gills from fish used in Experiment 2 were analyzed by two-way analysis of variance (ANOVA), followed by Tukey's posttest. All the applied statistical analyses in the present study used a confidence level of 95%.

Results

Expression of miR-429 after acute exposition to seawater

The relative expression of miR-429 in the brain, liver, and kidney did not differ between SW- and FW-exposed fishes ($p > 0.05$) (Figures 1B–D). However, a notorious downregulation of miR-429 in gills was observed ($p < 0.05$) after 24 h of exposition to 30 g.L^{-1} of salt (Figure 1A). The observation that these tissues were the locality where the miR-429 was differentially modulated after acute osmotic challenge allowed the selection of the gills as a target for long-term evaluation.

Expression of miR-429 after mild salinity changes

The relative expression of miR-429 (Figure 2) did not differ between FW- and BW-acclimated fish at D0 ($p > 0.05$). Nevertheless, a change in miR-429 expression was observed after the FW-BW transfer. In both FW-BW- and BW-FW-transferred groups, miR-429 expression remained without a difference at D1. The miRNA expression significantly increased ($p < 0.05$) at D7 in the BW-FW transfer and remained without a difference at D15 ($p > 0.05$). The miR-429 expression level did not change in the FW-BW-transferred fish ($p > 0.05$) at D1 until D7. However, there was a notable increase ($p < 0.05$) of miR-429 at D15.

Discussion

The silverside *O. humensis* is a teleost that moves from coastal waters near land or in estuaries to near-shore shallow areas. Due to their recent marine origin, and because their currently inhabit both freshwater and brackish waters environments, silversides can tolerate a wide range of environmental salinity and are regarded as euryhaline (Bemvenuti, 2006). The silversides have received attention from researchers who aim to develop the use of fish farming in estuarine regions, where there is a constant variation in the salinity of water available for use in cultivation, and in regions of continental farming, where the available water is brackish (Piedras, 2009; Heras And Roldán, 2011).

As a wild species, *O. humensis* reside directly in water environments and usually experience sudden changes or frequent fluctuation of environmental salinity. Indeed, salinity is a frequent abiotic stressor that restrains fish growth and development by triggering osmotic stress responses. Thus, when fishes are subjected to salinity stress, related genes are activated to induce tolerance to the salinity-mediated stress (Silveira et al., 2018b). Although miRNAs provide fish with elaborate strategies for gene regulation at the post-transcriptional level, the precise mechanisms by which miRNAs employ their regulatory function facing osmotic stress are not fully understood. Therefore, the role of miRNAs in osmoregulation has attracted much attention from researchers.

Similarly to fishes, soil salinity is a major abiotic stress in plant agriculture, where high levels of salt can induce ion imbalance and hyperosmotic stress. Indeed, salinity alters miRNA expression in tobacco in a dose-dependent manner because miR-395 is found to be severely sensitive to salt (Frazier et al., 2011). The expression of the miR-169 family members is thought to be inducible by salinity because they regulate several genes involved with salt tolerance in *Arabidopsis thaliana* and transgenic rice (Zhao et al., 2009). Moreover, oni-miR-30c acts limit the expression of *hsp70* gene, which is regarded as essential in osmoregulation in tilapia (Yan et al., 2012b). In addition, the miR-8 family, expressed mainly in ionocytes, is directly attached to osmoregulation in zebrafish. Furthermore, the miR-8 family acts by regulating the expression of Nherf1, which regulates transmembrane ionic transport (Flynt et al., 2009).

Interestingly, regarding the gills as a crucial regulator of internal osmotic balance, miR-429 seems to be more copious in gills than in any other organ. This indicates that the gills play a pivotal role in regulating environmental salinity tolerance in a post-transcriptional level *via* miRNAs (Bartel, 2009). Furthermore, miR-429 is known to compose a regulatory circuit that allows fish to respond to osmotic stress because it directly regulates the expression of osmotic stress transcription factor 1 (Ostf1) by targeting its 3′-UTR region (Fiol et al., 2006; Yan et al., 2012b).

Given that there are no studies that relate the epigenetic regulation of osmotic stress through miRNAs in silversides, we have pioneered the analysis of miRNAs relative expression under environmental salinity using *O. humensis* as a bioindicator. Here, the expression of miR-429 analysis by Two-step RT-qPCR confirms our hypothesis that this miRNA is directly involved in the osmoregulation acclimation process in silversides and emerges as osmotic stress marker because miR-429 presented significant differential gene expression under environmental salinity.

In this study, we first performed an acute exposure to SW ($30\ g.L^{-1}$) for 24 h to determine in which tissue miR-429 is differentially expressed during hyperosmotic stress. The miR-429 expression pattern was found to be remarkably decreased in the gills after the acute exposure period. These results reveal that miR-429 has a prime role in silversides osmoregulation following acute salt treatment. Similar effects in miR-429 expression were observed in Nile tilapia challenged by salinity stress, which presented lower expression levels of miR-429 in gills after hyperosmotic situations (Yan et al., 2012a).

We subsequently exposed the fish previously acclimated to FW to a salt concentration of $10\ g.L^{-1}$ for 15 days (and *vice versa*) to evaluate the role of miR-429 in osmoregulation. A milder concentration of salt was used because the animals would not be able to stand the SW for such a long period (unpublished personal observations). In BW-FW-transferred fish, miR-429 expression increased at D7 upon salinity stress. Likewise, in FW-BW-transferred fish, miR-429 significantly increased at D15. Taken together, these findings indicate that miR-429 in silversides plays a key regulatory role in regulating salinity tolerance at both acute and chronic exposition.

In addition to being regarded as FW, silversides are often exposed to salt and brackish water on estuaries and coastal lagoons of South America, and have better development and survival in saline environments (Tsuzuki et al., 2001). Previous literature has observed that the BW-FW transfer is more stressful to *O. humensis* than the FW-BW transfer (Silveira et al., 2018b).

In this sense, given that in BW-FW group the expression of miR-429 increased after 1 week of acclimation, more precociously than in FW-BW group, miR-429 directly regulates silversides osmoregulation. In FW-BW-transferred fish, miR-429 expression level significantly increased at D15. This indicates that this miRNA also plays a key role in long-term response to salinity.

The levels of salinity were shown to modulate the energy supply available for growth and reproduction in farmed fishes (Chand et al., 2015), and its optimal adjustment can benefit *O. humensis* production in captivity for aquaculture purposes. Indeed, the brackish medium is ideal for the cultivation of silversides. Cultivating this species in brackish environments, which are near-isosmotic environments, has vast potential in aquaculture to reduce economic losses by mortality due to the handling and transport of the fish, and due to the low quality of the water (Tsuzuki et al., 2001). In the same way, the culture of the silversides in brackish medium makes it possible to increase the survival rate of the embryos, increase the growth rates and feed conversion levels, and improve the efficiency of energy absorption and other parameters of productive interest (Silveira et al., 2018b).

In summary, our results suggests that miR-429 in silversides plays a prime role in osmoregulation and is mainly expressed in the gills during osmotic stress. Our findings suggest that miRNAs appear to regulate salt tolerance at the post-transcriptional level. Therefore, miR-429 has substantial potential to be developed as a novel molecular marker to assist silverside aquaculture by farming this fish in estuarine regions and brackish environments. It can also be used in the selection based on epigenetic markers of salinity-tolerant fish. To understand the integrated responses and regulatory roles of miR-429 during osmotic stress in silversides, further studies that aim to analyze the gene expression of target genes modulated by miR-429 are required. Finally, our study provides evidence for the use of *O. humensis* as a model in scientific research.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The animal study was reviewed and approved by the Ethics Committee on Animal Experimentation of the Federal University of Pelotas (Process no. 23110.007018/2015-85).

Author contributions

VC, WD, AP, and TS conceived and planned the experiments. TS, MR, and RR generated and raised the animals. AP, BB, WD, EB, and ED carried out the laboratorial analyses. LN and WD contributed to miRNA libraries preparation. AP, WD, and VC contributed to the interpretation of the results. AP took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Funding

This study was supported by Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS-FAPESP PqG #19/2551-0000953-3) and was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES) Finance Code 001 and AUXPE #2537/2018. AP, WD, LN, EB, and ED are individually supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. MR, RR, and VC are also individually supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico.

Acknowledgments

We wish to thank the undergraduate students from Laboratório de Genômica Estrutural for helping with the daily routine at the laboratory.

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.

Publisher's note

References

Bartel, D. (2009). MicroRNAs: Target recognition and regulatory functions. *Cell* 136, 215–233. doi:10.1016/j.cell.2009.01.002

Bemvenuti, M. A. (2006). Silversides in south Brazil: morphological and ecological aspects. *Biocell.* 30, 111–118.

Campanella, D., Hughes, L., Unmack, P., Bloom, D., Piller, K., 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

Chand, B., Trivedi, R., Dubey, S., Rout, S., Beg, M., and Das, U. (2015). Effect of salinity on survival and growth of giant freshwater prawn *Macrobrachium rosenbergii* (de Man). *Aquac. Rep.* 2, 26–33. doi:10.1016/j.aqrep.2015.05.002

Dyer, B. (2006). Systematic revision of the South American silversides (teleostei, atheriniformes). *Biocell.* 30, 69–88.

Fiol, D., Chan, S., and Kültz, D. (2006). Regulation of osmotic stress transcription factor 1 (Ostf1) in tilapia (*Oreochromis mossambicus*) gill epithelium during salinity stress. *J. Exp. Biol.* 209, 3257–3265. doi:10.1242/jeb.02352

Fiol, D. F., and Kültz, D. (2007). Osmotic stress sensing and signaling in fishes. *FEBS J.* 274 (22), 5790–5798. doi:10.1111/j.1742-4658.2007.06099.x

Flynt, A., Thatcher, E., Burkewitz, K., Li, N., Liu, Y., and Patton, J. (2009). miR-8 microRNAs regulate the response to osmotic stress in zebrafish embryos. *J. Cell Biol.* 185 (1), 115–127. doi:10.1083/jcb.200807026

Frazier, T., Sun, G., Burklew, C., and Zhang, B. (2011). Salt and drought stresses induce the aberrant expression of microRNA genes in tobacco. *Mol. Biotechnol.* 49, 159–165. doi:10.1007/s12033-011-9387-5

Gebert, L., and Macrae, I. (2019). Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol.* 20, 21–37. doi:10.1038/s41580-018-0045-7

Heras, S., and Roldán, M. (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

Leung, A., and Sharp, P. (2010). MicroRNA functions in stress responses. *Mol. Cell* 40, 205–215. doi:10.1016/j.molcel.2010.09.027

Makeyev, E., and Maniatis, T. (2008). Multilevel regulation of gene expression by microRNAs. *Science* 319, 1789–1790. doi:10.1126/science.1152326

Mank, J., and Avise, J. (2006). Supertree analyses of the roles of viviparity and habitat in the evolution of atherinomorph fishes. *J. Evol. Biol.* 19 (3), 734–740. doi:10.1111/j.1420-9101.2005.01060.x

Martins, A., Silveira, T., Remião, M., Domingues, W., Dellagostin, E., Junior, A., et al. (2021). Acute exposition to Roundup Transorb® induces systemic oxidative stress and alterations in the expression of newly sequenced genes in silverside fish (*Odontesthes humensis*). *Environ. Sci. Pollut. Res. Int.* 28, 65127–65139. doi:10.1007/s11356-021-15239-w

Piedras, S., 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 (3), 235–238. doi:10.5007/2175-7925.2009v22n3p235

Richter, M., Haslbeck, J., and Buchner, J. (2010). The heat shock response: life on the verge of death. *Mol. Cell* 40, 253–266. doi:10.1016/j.molcel.2010.10.006

Seale, A., and Breves, J. (2022). Endocrine and osmoregulatory responses to tidally-changing salinities in fishes. *Gen. Comp. Endocrinol.* 15 (326), 114071. doi:10.1016/j.ygcen.2022.114071

Silveira, T., Domingues, W., Remião, M., Santos, L., Barreto, B., Lessa, I. M., et al. (2018a). 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

Silveira, T., Martins, G., Domingues, W., Remião, M., Barreto, B., Lessa, I., et al. (2018b). Gene and blood analysis reveal that transfer from brackish water to freshwater is more stressful to the silverside *Odontesthes humensis*. *Front. Genet.* 6, 28. doi:10.3389/fgene.2018.00028

Tsuzuki, M., Ogawa, K., Strüssmann, C., 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

Yan, B., Guo, J., Zhao, L., and Zhao, J. (2012a). MiR-30c: a novel regulator of salt tolerance in tilapia. *Biochem. Biophys. Res. Commun.* 425 (2), 315–320. doi:10.1016/j.bbrc.2012.07.088

Yan, B., Zhao, L., Guo, J., and Zhao, J. (2012b). miR-429 regulation of osmotic stress transcription factor 1 (OSTF1) in tilapia during osmotic stress. *Biochem. Biophys. Res. Commun.* 426 (3), 294–298. doi:10.1016/j.bbrc.2012.08.029

Zebral, Y., Costa, P., Knopp, B., Lansini, L., 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

Zhao, L., Ge, R., Liang, W., Li, K., Ruan, H., Lin, Y., et al. (2009). Members of miR-169 family are induced by high salinity and transiently inhibit the NF-YA transcription factor. *BMC Mol. Biol.* 10, 29. doi:10.1186/1471-2199-10-29

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 08 September 2022
DOI 10.3389/fgene.2022.948228

OPEN ACCESS

EDITED BY
Paolo Franchini,
Sapienza University of Rome, Italy

REVIEWED BY
Tereza Manousaki,
Hellenic Centre for Marine Research (HCMR), Greece
Suvra Roy,
Central Inland Fisheries Research Institute (ICAR), India

*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 19 May 2022
ACCEPTED 15 August 2022
PUBLISHED 08 September 2022

CITATION
Martins AWS, Dellagostin EN, Blödorn EB, Silveira TLR, Sampaio LA, Komninou ER, Varela Junior AS, Corcini CD, Nunes LS, Remião MH, Collares GL, Domingues WB and Campos VF (2022), Exposure to salinity induces oxidative damage and changes in the expression of genes related to appetite regulation in Nile tilapia (Oreochromis niloticus). *Front. Genet.* 13:948228. doi: 10.3389/fgene.2022.948228

Exposure to salinity induces oxidative damage and changes in the expression of genes related to appetite regulation in Nile tilapia (*Oreochromis niloticus*)

Amanda W. S. Martins[1], Eduardo N. Dellagostin[1], Eduardo B. Blödorn[1], Tony Leandro R. Silveira[2], Luis A. Sampaio[3], Eliza R. Komninou[1], Antonio S. Varela Junior[4], Carine D. Corcini[5], Leandro S. Nunes[1], Mariana H. Remião[1], Gilberto L. Collares[6], William B. Domingues[1] and Vinicius F. Campos[1]*

[1]Laboratório de Genômica Estrutural, Programa de Pós-Graduação em Biotecnologia, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, RS, Brazil, [2]Instituto de Ciências Biológicas, Universidade Federal do Rio Grande, Rio Grande, RS, Brazil, [3]Laboratório de Piscicultura Estuarina e Marinha, Programa de Pós-graduação em Aquicultura, Instituto de Oceanografia, Universidade Federal do Rio Grande, Rio Grande, RS, Brazil, [4]Laboratório de Reprodução Animal, Programa de Pós-Graduação em Biologia de Ambientes Aquáticos Continentais, Instituto de Ciências Biológicas, Universidade Federal do Rio Grande, Rio Grande, RS, Brazil, [5]ReproPel, Programa de Pós-Graduação em Veterinária, Faculdade de Veterinária, Universidade Federal de Pelotas, Pelotas, RS, Brazil, [6]Agência de Desenvolvimento da Bacia da Lagoa Mirim, Universidade Federal de Pelotas, Pelotas, RS, Brazil

Variations in water salinity and other extrinsic factors have been shown to induce changes in feeding rhythms and growth in fish. However, it is unknown whether appetite-related hormones mediate these changes in Nile tilapia (*Oreochromis niloticus*), an important species for aquaculture in several countries. This study aimed to evaluate the expression of genes responsible for appetite regulation and genes related to metabolic and physiological changes in tilapia exposed to different salinities. Moreover, the study proposed to sequence and to characterize the *cart, cck*, and *pyy* genes, and to quantify their expression in the brain and intestine of the fish by quantitative polymerase chain reaction (qPCR). The animals were exposed to three salinities: 0, 6, and 12 parts per thousand (ppt) of salt for 21 days. Furthermore, lipid peroxidation, reactive oxygen species, DNA damage, and membrane fluidity in blood cells were quantified by flow cytometry. The results indicated an increased expression of *cart, pyy*, and *cck* and a decreased expression of *npy* in the brain, and the same with *cck* and *npy* in the intestine of fish treated with 12 ppt. This modulation and other adaptive responses may have contributed to the decrease in weight gain, specific growth rate, and final weight. In addition, we showed oxidative damage in blood cells resulting from increasing salinity. These results provide essential data on *O. niloticus* when exposed to high salinities that have never been described before and generate

knowledge necessary for developing biotechnologies that may help improve the production of economically important farmed fish.

KEYWORDS

CART, PYY, NPY, CCK, gene expression, food intake

1 Introduction

Nile tilapia (*Oreochromis niloticus*) is a fish species native to African freshwater lakes and rivers. It is currently one of the most widely farmed freshwater fish around the world due to its favorable production characteristics and high economic value and has been used as a biological model in several studies (Beardmore et al., 2001; He et al., 2011; FAO, 2020; Mohamed et al., 2021). Considering the scarcity of fresh water in many countries, studies are being carried out to develop the production of this species in brackish and seawater (El-Sayed et al., 2005; Mohamed et al., 2021).

Changes in water salinity modify the homeostasis of the organisms and biological processes and may even lead to death (Rahmah et al., 2020). Exposure of different tilapia species to increasing salinity is known to affect growth performance (Ninh et al., 2014; Gan et al., 2016), digestive capacity (Tran-Ngoc et al., 2017), blood parameters (Verdegem et al., 1997), reproductive capacity (Cruz Vieira et al., 2019), histopathology and behavior (Hassan et al., 2013), antioxidant status (Gan et al., 2016), and metabolic rate (Iwama et al., 1997).

The homeostatic regulation of food intake depends on a complex network involving signals which promote the release of a wide range of hormones produced by the brain and peripheral organs that can stimulate (orexigenic) or inhibit (anorexigenic) appetite (Conde-Sieira et al., 2018). Among the main hormones involved in appetite regulation, the most important ones are the neuropeptide Y (Npy), cocaine and amphetamine-regulated transcription factor (Cart), cholecystokinin (Cck), and peptide YY (Pyy) (Rønnestad et al., 2017).

Plasma glucose level is one of the most common stress indicators. Furthermore, circulating glucose levels are believed to activate neurocircuitry in regulating food intake (Pankhurst, 2011; Soengas, 2014). In addition, blood oxidative parameters are often used to assess fish health (Bojarski and Witeska, 2020). Thus, evaluating whether salinities caused any damage to tilapia blood cells or whether they modulated plasma glucose levels is extremely interesting.

Living in a marine or freshwater environment, with direct contact with water for gas exchange and secretion of metabolic waste, teleost fish face constant osmotic stress that requires physiological compensation (Mohamed et al., 2021). Under stress conditions, the mechanisms that control food intake in fish are dysregulated. Therefore, understanding the factors underlying the responses to environmental stress observed among species is of great interest, having implications for aquaculture output. Currently, little is known about the influence of salinity on genetic variations related to endocrine factors that regulate appetite in fish, specifically in *O. niloticus*. Therefore, this study aimed to evaluate if exposure to different salinities causes modulation of the expression of genes related to appetite regulation and affects growth rates and the systematic physiology of tilapia.

2 Materials and methods

2.1 Animals and conditions

The fish used in this study were obtained from the Laboratory of Pisciculture of the Barragem do Chasqueiro Fish Farming (Arroio Grande, Brazil-32°14′15 ″S/53°05′13″ W). The specimens, with initial mean weight and length of 200 ± 8.3 g and 18.5 ± 5.3 cm respectively, were fed twice a day with a commercial diet (Supra, 38% crude protein) until apparent satiety. The fish were kept in plastic tanks with a nominal capacity of 1,000 L filled with 650 L water. The tanks used were opaque to reduce visual stress. Approximately 2/3 of the water in each tank was renewed once a week. The water quality parameters during the acclimation period were as follows: temperature of 24.0 ± 0.34°C; dissolved oxygen level of 9.85 ± 0.1 mg O_2. L^{-1}; salinity of 0 ± 0.21 ppt; pH of 7.04 ± 0.19; and total ammonia level lower than 0.3 mg NH_3 L^{-1}.

2.2 Experimental design and sample collection

After 4 weeks of acclimation, fish were randomly distributed into three salinity groups: two experimental groups exposed to 6 and 12 ppt of salt (NaCl) and a control group with fresh water (0 ppt). The salinity levels were obtained by increasing salinity per day until they reached 6 and 12 ppt. The salinity levels mentioned above were maintained throughout the experimental period by regular salt applications, followed by its measurement through a salinometer (Kasvi, Brazil) to ensure the required salinity level daily. The chosen concentrations are within the minimum mortality range (El-Leithy et al., 2019). The experimental groups were exposed for 21 days, and the exposition of all the groups was performed in plastic tanks in triplicate (totaling 9 tanks) with ten animals per tank (n = 90 animals) maintained under the same conditions. The water

quality parameters were measured daily during the experimental period and were held as follows: Temperature of 24.5 ± 0.40°C; dissolved oxygen level of 11 ± 0.1 mg O_2. L^{-1}; pH of 7.27 ± 0.26; and total ammonia level lower than 0.3 mg NH_3 L^{-1}. At the end of the experiment, all fish were anesthetized in a benzocaine bath (50 mg·L^{-1}) and euthanized by cranial spinal section. Their brain and posterior intestine were collected and were used for sequencing, molecular characterization, and gene expression analysis of target genes. The tissues were preserved in liquid nitrogen until their use for molecular biology processing and analyses. Blood samples were collected from the branchial branch artery using 26 G needles attached to heparinized syringes and centrifuged at 2,000×g for 10 min to separate plasma and blood cells (ambient temperature). The plasma osmolality was measured using an automated osmometer (Vapro® Vapor Pressure Osmometer, United States). For flow-cytometry analysis, 20 µl of the blood sample was added to 1 ml of fetal bovine serum (FBS) and stored at 4°C in the dark until use, and the glucose measurement was performed with the Accu-Chek Active Kit (Roche, CH), using 5 µl of blood. The Ethics and Animal Experimentation Committee of Universidade Federal de Pelotas approved the use of animals and all these handling practices (process no 23110.014105/2020-56).

2.3 Performance

Fish were individually weighed at the beginning and end of the experimental period. The weight was used to calculate the survival, weight gain, and specific growth rate (Kang'ombe et al., 2008).

Survival (%) = (final number of fish—initial number of fish)/ initial number of fish x 100

Weight gain (g) (WG) = FW- IW

Specific growth rate (%/day) (SGR) = 100 (ln FW—ln IW)/Δt.

FW is the final weight, IW is the initial weight, and Δt is the number of days between samplings.

2.4 RNA extraction and cDNA synthesis

Total RNA from tissue samples was extracted using TRI Reagent (Sigma, United States), following the manufacturer's instructions with few modifications. RNA concentration and purity were measured by UV-light spectrophotometry using the NanoVue™ equipment (GE Healthcare Life Sciences, United States), and the samples were then standardized by concentration. Only the samples presenting high purity (OD260/280 ≥ 2.0 nm) were used in the following procedures. According to the manufacturer's recommendation, RNA samples were treated with DNase using DNA-free™ Kit (Invitrogen™, United States). First-strand cDNA synthesis was performed with 2 µg of total RNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, United States) according to the manufacturer's protocol. The cDNA samples were stored at −20°C until further use.

2.5 Target genes in appetite regulation and osmoregulation process

The target genes related to the main hormones that play a critical role in appetite regulation were as follows: Neuropeptide Y (*npy*); Peptide YY (*pyy*); Cocaine and Amphetamine Regulated Transcript (*cart*); Cholecystokinin (*cck*). Genes related to osmotic stress were also analyzed: Sodium-Potassium ATPase (*nka*) and Co-transporter Na^+/ K^+/$2Cl^-$ (*nkcc*). Their functions are listed in Table 1.

2.6 Sequencing and molecular characterization of target genes

The gene fragments were amplified by polymerase chain reaction (PCR) using primers designed by the PriFi online tool (https://services.birc.au.dk/prifi/) after the alignment of known sequences deposited for each gene in GenBank. The PCR parameters were as follows: an initial denaturation step for 1 min at 94°C, followed by 35 cycles at 94°C for 30 s, 55.2–65. 5°C (depending on the primer sequence) for 30 s, and 72°C for 1 min, with a final extension of 5 min at 72°C. The fragments were sequenced using the Applied Biosystems 3500 Genetic Analyzer® automatic sequencer (Life Technologies, United States).

2.7 Sequence analysis

The partial sequences of the *cart*, *cck*, and *pyy* genes identified in this study were deposited in GenBank. The translation of sequenced nucleotides to amino acid sequences and the open reading frame (ORF) identification were performed using the ExPASy bioinformatics resource portal (https://www.expasy.org/). The conserved domains and sites were mapped using the UniProt database (https://www.uniprot.org/).

2.8 Gene expression analysis

Gene expression was analyzed by quantitative polymerase chain reaction (qPCR), and the primers used in this study (Table 1) were designed using the Primer3 online tool (https://primer3.ut.ee/http://bioinfo.ut.ee/primer3–0.4.0/) and the sequences of the identified fragments. The qPCR assay was performed using CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad, United States) and GoTaq®

TABLE 1 Target genes in appetite regulation, osmoregulation process, and primer sequences used in this study.

Gene symbol	Function	Primer sequence 5′→ 3′	Efficiency (%)	Objective	References
pyy	Anorexigenic factor	GGACAGTGCTGGTGGCCTTAGTG	—	Cloning	Yan et al. (2017)
		TCTCTGGTCTCTGTTGTTATCGCC			
		AACACTGGCTGATGCCTACC	100.6	qRT-PCR	
		TTCCATACCTCTGCCTGGTG			
cart	Anorexigenic factor	TGGTCTAYCTGTCCGTCTGTCTG	—	Cloning	
		TAGCAGCGCAGGAAGAAGG			
		TGCTGACATCACTCTGTCAAGG	98.6	qRT-PCR	
		AGCCAGCTCACTGGTTGTG			
cck	Anorexigenic factor	TCTCACTCTCACACACTCC	—	Cloning	
		AGGAGTACTCATACTCCTCTG			
		AGAAACTCCACGGCAAACAG	92.5	qRT-PCR	
		ACTCATACTCCTCTGCACTGC			
npy	Orexigenic factor	ACAAGACAGAGGTATGGGAAGA	—	qRT-PCR	
		GGCAGCATCACCACATTG			
nka	Transport of ions and absorption of water	GCTCCAGAGAGGATTTTGGAC	—	qRT-PCR	Velan et al. (2011)
		CTCCAAGACCTCCCAACTCA			
nkcc	Transport of ions and absorption of water	GAGGCAAGATCAACAGGATTG	—	qRT-PCR	Velan et al. (2011)
		AATGTCCGAAAAGTCTATCCTGAACT			
actb	Reference gene	TGGTGGGTATGGGTCAGAAAG	—	qRT-PCR	Yang et al. (2013)
		CTGTTGGCTTTGGGGTTCA			

RT-qPCR Master Mix (Promega, United States). 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 requirements needed to calculate the melting curve. The $2^{-\Delta\Delta CT}$ method was used to normalize the fold change in gene expression using *actb* as a reference gene for normalization (Yang et al., 2013).

2.9 Flow cytometry analysis

Flow cytometry analyses were performed using Attune® (Acoustic Focusing Flow Cytometer, Applied Biosystems, United States) with violet laser (UV 405 nm-450/40, VL^{-1}). To evaluate the effect of different salinities exposure on the systemic physiology of *O. niloticus*, erythrocytes were washed with 500 µl of FBS, stained with Hoechst 33,342 (16.2 mM), and assayed for reactive oxygen species (ROS), lipid peroxidation (LPO), membrane fluidity, and DNA fragmentation index (DFI) (Martinez-Alborcia et al., 2012). Cell debris were discarded by scatter plots of forward scatter × side scatter and negative fluorescence of Hoechst 33,342. To read all parameters, the fluorophore-stained cells were added into calcium- and magnesium-free phosphate-buffered saline (80 $g{\cdot}L^{-1}$ of NaCl, 11.5 $g{\cdot}L^{-1}$ of KCl, 24 $g{\cdot}L^{-1}$ of Na_2HPO_4, 2 $g{\cdot}L^{-1}$ of KH2PO4 dissolved in deionized water). A total of 20,000 cells were measured during each analysis.

2.9.1 ROS evaluation

For the evaluation of ROS produced by the blood cells, 10 µl of previously collected and stored blood sample was added to 20 µl of saline solution containing 2 µM 2′,7′-dichlorofluorescein diacetate, and 5 µM propidium iodide (PI) fluorescent probes (Sigma-Aldrich Co., United States). The samples were analyzed after incubation for 60 min at 22°C in the dark. Only live cells (PI negative) were selected and measured for ROS production by the median intensity of the emitted green fluorescence.

2.9.2 LPO evaluation

The LPO was quantified using 1 µM of the lipid peroxidation sensor 4,4-difluoro-5-(4-phenyl-1,3-butadienyl)-4-bora-3a,4a-diaza-s-indacene-3-undecanoic acid (C11-BODIPY) in 100 µl of the sample. It was incubated for 2 h at room temperature (20°C). The rate of lipoperoxidation was calculated by the median intensity of green fluorescence (peroxidized lipid)/median green fluorescence intensity + median red fluorescence (non-peroxidized lipid) × 100.

2.9.3 Membrane fluidity evaluation

Erythrocyte membrane fluidity was analyzed by hydrophobic merocyanine 540 dye (M540) at a final concentration of 2.7 M (Sigma-Aldrich, United States) and YO-PRO, which fluoresces green, at a final concentration of 0.1 M (Invitrogen,

TABLE 2 Effects of salinity exposure on Nile tilapia growth performance.

	Control	6 ppt	12 ppt
Final weight (g)	295.5 ± 11.42[a]	240.8 ± 14.69[b]	230.9 ± 13.56[b]
Specific growth rate (%/day)	1.70 ± 0.38[a]	−0.12 ± 0.05[b]	−0.16 ± 0.47[b]
Weight gain (g)	72.56 ± 12.1[a]	−5.84 ± 2.68[b]	−9 ± 23.80
Glucose (mg/dl)	60.50 ± 9.13[a]	62.17 ± 5.41[a]	147.3 ± 15.87[b]
Osmolality (mOsm/kg)	325.1 ± 4.17[a]	325.3 ± 3.70[a]	347.8 ± 5.33[b]
Survival (%)	100	100	50

Data are expressed as means ± standard error of the mean. Specific growth rates (%/day) were calculated as SGR = 100 (ln FW—ln IW)/Δt, where IW is the initial weight, FW is the final weight, and Δt is the number of days between samplings. The weight gain (in grams, g) was calculated as WG = final weight—initial weight; and survival (%) = (final number of fish—initial number of fish)/initial number of fish × 100. Different letters indicate significant differences between the experimental groups (one-way analysis of variance; $n = 30$; $p < 0.05$).

United States). Only live cells (YO- PRO negative) were selected and classified into high fluidity cells (high M540 concentration) and low fluidity cells (low M540 concentration).

2.9.4 DNA damage evaluation

To evaluate DNA damage in erythrocytes, 10 µl of blood sample was added to 5 µl of 0.01 M Tris-HCl, 0.15 M NaCl, and 0.001 M EDTA (pH 7.2) and to 10 µl of Triton 1X (Triton X-100, 1%, v/v) 30 s later. Then, 50 µl of acridine orange dye (2 mg·ml^{-1}, #A6014, Sigma-Aldrich, United States) was added to the sample, followed by incubation from 30 s up to 2 min before each reading. The DNA of erythrocytes was classified as undamaged (green fluorescence emission) or damaged (orange/red fluorescence emission). The DNA fragmentation index (DFI) percentage was obtained by the median of red fluorescence intensity/(median of the red + green fluorescence intensities) × 100.

2.10 Statistical analysis

The evaluated parameters, which presented normal distribution and homogeneous variance with or without transformations, were analyzed by one-way analysis of variance (ANOVA), followed by Tukey's post-test with a significance level of 5%. The results from gene expression, ROS, LPO, membrane fluidity, and DFI were expressed as mean ± standard error of the mean (SEM).

3 Results

3.1 Effects of 21 days of different salinity exposure on nile tilapia growth performance

The final weight, weight gain, and specific growth rate were lower ($p < 0.05$) in the groups exposed to 6 and 12 ppt compared to the control group, as shown in Table 2. Furthermore, the plasma osmolality and the blood glucose showed an increase ($p < 0.05$) in the fish exposed to 12 ppt compared to fish from the control and 6 ppt groups. The survival was significantly lower in the group exposed to 12 ppt compared to the other groups.

3.2 Sequencing and molecular characterization

The length of *cart*, *cck*, and *pyy* amplified fragments from *O. niloticus* was 306, 453, and 237 bp, respectively. They were sequenced and deposited under the GenBank accession numbers MW556307.1, MW556308.1, and MW556314.1, respectively.

The partial sequence of *cck* cDNA codes for 136 amino acids on ORF +1. The fragment belongs to the gastrin/cholecystokinin family and has a signal peptide between amino acids 1 and 19 and a gastrin domain between amino acids 4 and 136. The partial sequence of *cart* cDNA codes for 102 amino acids on ORF +1. The fragment belongs to the cart family and has a signal peptide between amino acids 1 and 15. The partial sequence of *pyy* cDNA codes for 79 amino acids on ORF +. The fragment belongs to the NPY family and has a signal peptide between amino acids 1 and 18.

3.3 Gene expression analysis

3.3.1 Brain

The mRNA expression of *cart* and *cck* in the brain of the fish exposed to 6 ppt and 12 ppt was higher ($p < 0.05$) than in fish from the control group (Figures 1A,D). The mRNA expression of *npy* in the brain of fish exposed to 12 ppt was lower than in control and 6 ppt groups (Figure 1B). Furthermore, the *pyy* mRNA expression in the brain of fish from the 12 ppt group was higher ($p < 0.05$) than in fish from the other groups (Figure 1C).

3.3.2 Intestine

In the intestine, the mRNA expression of *cck* of the fish exposed to 6 ppt was higher ($p < 0.05$) than that in the

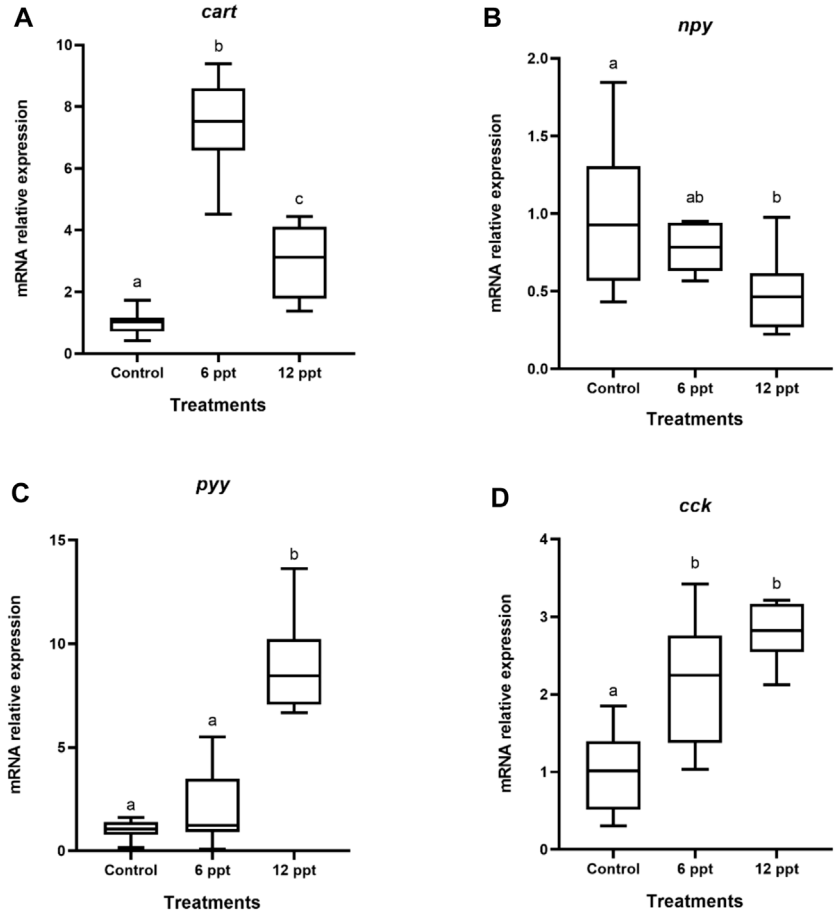


FIGURE 1
Gene expression in the brain of *Oreochromis niloticus* in the control group and those exposed to 6 and 12 parts per thousand (ppt) of salt for 21 days. The relative expression of the *cart* **(A)**, *npy* **(B)**, *pyy* **(C)**, and *cck* **(D)** mRNA was evaluated by quantitative polymerase chain reaction and normalized using the *actb* gene. The values are expressed as mean ± standard error of the mean. Different letters indicate significant differences between the experimental groups (one-way analysis of variance; $n = 12$; $p < 0.05$)

fish from another group (Figure 2A). The mRNA expression of *npy* in the intestine of the fish exposed to 12 ppt was significantly lower than that in the fish from the control group and those exposed to 6 ppt (Figure 2B).

The mRNA expression of *nka* and *nkcc* in the intestine of the fish exposed to 12 ppt was higher ($p < 0.05$) than that in the fish from another group (Figures 3A,B).

3.4 Flow cytometry analysis

The exposure of Nile tilapia for 21 days to12 ppt increased ($p < 0.05$) the ROS (Figure 4A) and LPO levels in erythrocytes (Figure 4B). The membrane fluidity levels in the erythrocytes of fish exposed to 12 ppt showed a significant decrease (Figure 4C). Further, both the concentrations of salinity induced a significant increase ($p < 0.05$) in the DFI of erythrocytes compared with the control group (Figure 4D).

4 Discussion

In the present study, the mRNA partial sequences of *cart*, *cck*, and *pyy* in *O. niloticus* were identified and characterized successfully. Furthermore, this was the first study to evaluate the effects of different salinities on the expression of genes related to appetite in Nile tilapia.

The significant decrease in final body weight, specific growth, and weight found in our study is in line with the report of other authors who observed that increased salinity negatively affected *O. niloticus* growth, as well as feed intake and feed conversion ratio in

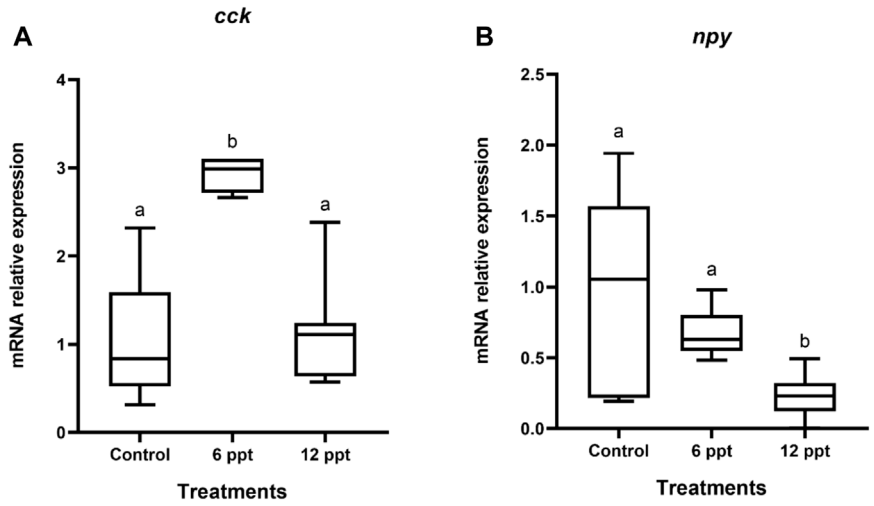


FIGURE 2
Gene expression in the intestine of *Oreochromis niloticus* in the control group and those exposed to 6 and 12 parts per thousand (ppt) of salt for 21 days. The relative expression of *cck* **(A)** and *npy* **(B)** mRNA was evaluated by quantitative polymerase chain reaction and normalized using the *actb* gene. The values are expressed as mean ± standard error of the mean. Different letters indicate significant differences between the experimental groups (one-way analysis of variance; $n = 12$; $p < 0.05$)

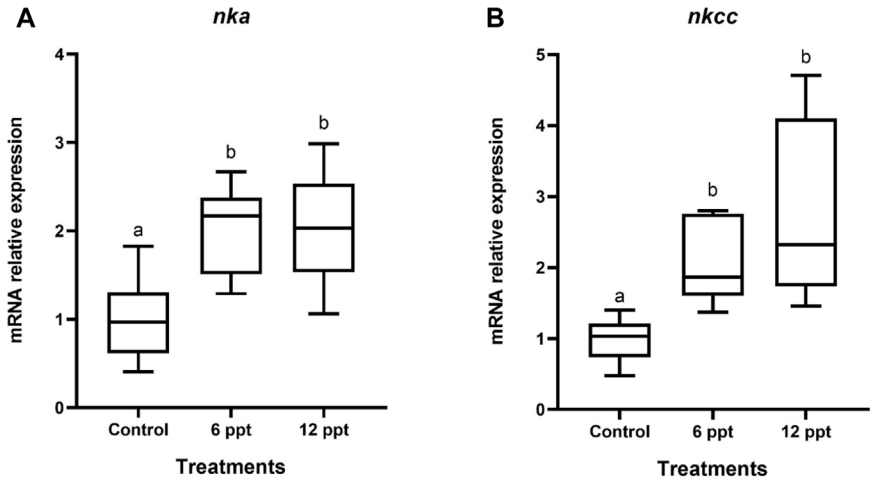


FIGURE 3
Gene expression in the intestine of *Oreochromis niloticus* in the control group and those exposed to 6 and 12 parts per thousand (ppt) of salt for 21 days. The relative expression of nka **(A)** and nkcc **(B)** mRNA was evaluated by quantitative polymerase chain reaction and normalized using the *actb* gene. The values are expressed as mean ± standard error of the mean. Different letters indicate significant differences between the experimental groups (one-way analysis of variance; $n = 12$; $p < 0.05$).

other fish species (Luz et al., 2008; Gan et al., 2016). Researchers suggest that decreased growth in increasing salinity environments is related to reduced food consumption (appetite) (Yan et al., 2004). It is known that feeding in vertebrates is regulated by several factors, including a wide range of orexigenic or anorexigenic hormones, which were analyzed in this study.

Among the anorexigenic hormones, *cart* mRNA is highly expressed in the brain, stomach, and intestine (Ahmadian-Moghadam et al., 2018; Zhang et al., 2018). Here, we observed an increase in the expression of *cart* in the brain, corroborating the report by other authors that this gene may be related to increased appetite. Central administration of Cart inhibits food

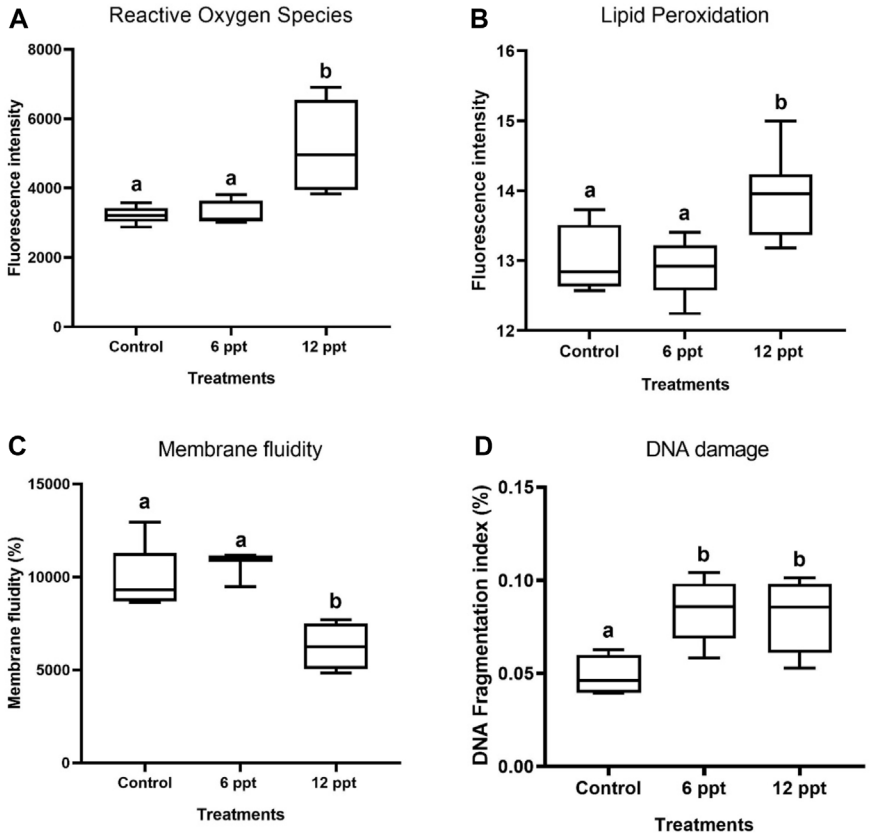


FIGURE 4
The oxidative effects in terms of reactive oxygen species production **(A)**, lipid peroxidation **(B)**, membrane fluidity **(C)**, and DNA fragmentation index (%) **(D)** in the erythrocytes of Oreochromis niloticus in control and in the groups of fish exposed to 6 and 12 parts per thousand (ppt) of salt, as evaluated by flow cytometry.

intake in mice (Kristensen et al., 1998), rats (Stanley et al., 2001) and fish (Volkoff and Peter, 2000).

Npy and Pyy have been extensively studied in several species because of their roles in the mechanism of feeding behavior (Volkoff et al., 2005). Npy injections increase feeding in goldfish (Narnaware and Peter, 2001) and zebrafish (Yokobori et al., 2012). Moreover, Npy treatments have also been shown to stimulate fish growth/growth hormone (GH) secretion both *in vitro* and *in vivo* (Assan et al., 2021). According to what we observed, salinity caused a decrease in the expression of *npy* in the brain and intestine, which could impact appetite modulation in these organs. In turn, Pyy is known to function as an anorexigenic indicator and has been recognized in several fish species, including Atlantic salmon (*Salmo salar*) and goldfish (*Carassius auratus*) (Murashita et al., 2009; Gonzalez and Unniappan, 2010). The increased expression of *pyy* in the brain indicates that this anorexigenic pathway may have been activated by salinity. Cck plays an essential role in food regulation. Here, *cck* had its expression increased by salinity. The literature has already reported that oral Cck administration inhibits feed intake in fish (Rubio et al., 2008). Collectively, these results support the hypothesis that salinity causes modulation in appetite-regulating genes.

Plasma glucose level is one of the most common stress indicators (Pankhurst, 2011; Soengas, 2014). An increase in blood glucose levels in groups exposed to the 12 ppt of salt was observed in the present study. Overall, published data suggest that the fish neural system that controls food intake also responds to stressful situations, increasing plasma glucose and cortisol levels (Conde-Sieira et al., 2018, 2010). It is known that distinct brain regions integrate endocrine and metabolic information to elaborate a coordinated response using neural effector pathways that modulate food intake. These areas contain specialized neurons that use glucose as a signaling molecule. Therefore, glucose-stimulated (GE) neurons increase, while glucose-inhibited (GI) neurons decrease their firing rate as glucose levels increase (Marty et al., 2007). It is already known that neurons in glucose detection areas produce

peptides involved in controlling food intake in mammals. Literature data suggest that the arcuate nucleus neurons in the hypothalamus that produce Npy appear to be GI. In contrast, the neurons that produce Cart appear to be GE, resulting in increased *cart* expression and decreased *npy* expression when glucose levels rise, which is in line with the results shown in this study (Dunn-Meynell et al., 2002; Mobbs et al., 2005; Conde-Sieira et al., 2018).

Another standard stress indicator is the blood cortisol level. In fish, one of the responses to stress caused by high salinities involves activation of the hypothalamic-pituitary-interrenal (HPI) axis, which starts with the production of corticotropin-releasing factor (CRF) and culminates in the production of cortisol. Mohamed et al., 2021 observed that blood cortisol levels were increased in fish exposed to higher salinity while investigating the physiological effects of salt stress on Nile tilapia. Previous studies suggest that this responsiveness to stressors is maintained under chronic stress (Gorissen and Flik, 2016) and that higher doses of cortisol profoundly inhibit the feeding behavior of goldfish (Bernier et al., 2004) and sea bass (Leal et al., 2011).

Oxidative parameters in the blood are frequently used to assess fish health (Bojarski and Witeska, 2020). Therefore, to evaluate whether the chosen salinities caused any damage to Nile tilapia blood cells, cell integrity analyzes were performed by flow cytometry. This was the first study that evaluated ROS, lipid peroxidation, membrane fluidity, and DNA fragmentation in blood cells of *O. niloticus* exposed to different salinities using a flow cytometer. It was possible to observe an increase in ROS levels in the higher salinity group. In normal physiological states, there is a balance between pro-oxidant production and antioxidant defenses. When an imbalance in favor of pro-oxidant output occurs, the antioxidant defenses can no longer neutralize the elevated levels of reactive oxygen species. Studies have already shown that salt stress induces the overproduction of ROS, which can disrupt metabolism and weaken the immune system, eventually leading fish to death (Liu et al., 2007). It is known that an imbalance in ROS and antioxidative enzymes production, leading to a bigger concentration of ROS, can increase lipid peroxidation, and this was observed in the present study. Our result corroborates with other authors who reported the increase in membrane damage caused by salinity (Sinha et al., 2014; Gan et al., 2016; Mohamed et al., 2021).

The change in membrane fluidity observed in the present study may have occurred because cell membranes are susceptible to disturbance from salinity changes. Sodium interacts with membrane lipid components, and increases in the concentration of sodium ions near the membrane lipids can alter membrane geometry and modulate the interaction with other macromolecules nearby (Evans and Kültz, 2020). We also observed an increase in the rate of genomic DNA damage, which could have been due to the increased ROS in the blood, or as a direct response to salinity exposure. DNA is a cellular macromolecule whose structure and function can be affected by changes in intra and extracellular ion concentrations that accompany salt stress (Fiol and Kültz, 2007). Due to its negative charge, DNA becomes susceptible to interaction with cations, such as Na^+, which can increase or decrease in abundance after the change in salinity. However, deviation from the optimal concentration of intracellular cations can result in irregular interactions with nucleic acids and changes in DNA structure and function (Várnai and Zakrzewska, 2004). Consistent with the destabilizing effect of sodium on nucleic acids, high levels of sodium cause DNA strand breaks in various organisms (Evans and Kültz, 2020). Previous studies demonstrated that fish positively regulate transcripts and proteins indicative of DNA damage as part of the salinity stress response. The presence of widespread strand breaks in marine organisms results from living in high sodium environments. These DNA strand breaks persist until salinity is reduced, emphasizing hyperosmotic stress's harmful effects (Dmitrieva et al., 2006; Whitehead et al., 2013).

The intestine displays important functions for absorbing nutrients and, as a barrier to the external environment, it has the additional function of maintaining the necessary absorption of the active fluid in seawater (Grosell, 2010). Increased active fluid absorption in seawater is associated with increased Na^+, K^+-ATPase (NKA) activity throughout the intestinal canal. The co-transporter $Na^+/K^+/2Cl^-$ (NKCC) together with NKA, contributes to the transport of ions and absorption of water. Here, an increase in the expression of *nka* and *nkcc* genes was observed, showing that the fish were under osmotic stress. These results agree with Mohamed et al. (2021), who observed a significant increase in the gene expression of *nka* in *O. niloticus* in groups with salt stress after 10 days of exposure. Similar results were also reported in rainbow trout (*Oncorhynchus mykiss*) and Mozambique tilapia. (Kammerer et al., 2009; Kammerer and Kültz, 2009; Gilmour et al., 2012).

From the analysis performed, it was possible to observe an increase in glucose and oxidative damage in *O. niloticus* blood cells exposed to 12 ppt. Furthermore, after 21 days, we observed modulation in the expression of *cart, npy, pyy, cck, nka,* and *nkcc*. These adaptive salinity responses may have helped to decrease appetite in tilapia since the anorexigenic genes were upregulated while the orexigenic factors were deregulated. Consequently, this modulation may have contributed to the decrease in weight gain, specific growth rate, and final weight.

The first death was observed in the middle of the experimental period. At the end of 21 days, we observed 50% of mortality in 12 ppt group, probably due to the harmful effects on Nile tilapia physiology, reinforcing the damage caused by salinity in this species. It also provides essential data on impact of salinity on Nile tilapia's physiology, which has never been described before. More comprehensive studies are still needed to better understand the effects of salinity on gene expression in Nile tilapia. However, the knowledge generated in this study can serve as a basis for selecting individuals or for the genetic improvement of Nile tilapia through genetic engineering.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Ethics statement

The animal study was reviewed and approved by Ethics Committee of Animal Experimentation of the Federal University of Pelotas.

Author contributions

VC: Conceptualization, Project admnistration, Funding acquisition, Supervision. Writing—Review and Editing. AM: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Writing—Original Draft. EK, EB; ED, TS, and LN: Investigation, Resources, Software, Validation. AJ, CC, MR, GC, LS, and WD: Investigation, Formal analysis, Data Curation, Writing—Review and Editing. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Funding

This study was supported by Fundação de Amparo à pesquisa do Estado do Rio Grande do Sul (FAPERGS-FAPESP #19/2551-0000953-3) and was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES) Finance Code 001 and AUXPE #2537/2018. VFC, LAS, ASVJ, CDC are also individually supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico.

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.

Publisher's note

References

Ahmadian-Moghadam, H., Sadat-Shirazi, M.-S., and Zarrindast, M.-R. (2018). Cocaine- and amphetamine-regulated transcript (cart): A multifaceted neuropeptide. *Peptides* 110, 56–77. doi:10.1016/j.peptides.2018.10.008

Assan, D., Mustapha, U. F., Chen, H., Li, Z., Peng, Y., and Li, G. (2021). The roles of neuropeptide Y (npy) and peptide YY (pyy) in teleost food intake: A mini review. *Life* 11, 547. doi:10.3390/life11060547

Beardmore, J. A., Mair, G. C., and Lewis, R. I. (2001). Monosex male production in finfish as exemplified by tilapia: Applications, problems, and prospects. *Aquac. Reproductive Biotechnol. Finfish Aquac.* 197, 283–301. doi:10.1016/S0044-8486(01)00590-710.1016/b978-0-444-50913-0.50015-1

Bernier, N. J., Bedard, N., and Peter, R. E. (2004). Effects of cortisol on food intake, growth, and forebrain neuropeptide Y and corticotropin-releasing factor gene expression in goldfish. *General Comp. Endocrinol.* 135, 230–240. doi:10.1016/j.ygcen.2003.09.016

Bojarski, B., and Witeska, M. (2020). Blood biomarkers of herbicide, insecticide, and fungicide toxicity to fish-a review. *Environ. Sci. Pollut. Res.* 27, 19236–19250. doi:10.1007/s11356-020-08248-8

Conde-Sieira, M., Agulleiro, M. J., Aguilar, A. J., Míguez, J. M., Cerdá-Reverter, J. M., and Soengas, J. L. (2010). Effect of different glycaemic conditions on gene expression of neuropeptides involved in control of food intake in rainbow trout; interaction with stress. *J. Exp. Biol.* 213, 3858–3865. doi:10.1242/jeb.048439

Conde-Sieira, M., Chivite, M., Míguez, J. M., and Soengas, J. L. (2018). Stress effects on the mechanisms regulating appetite in teleost fish. *Front. Endocrinol.* 9, 631. doi:10.3389/fendo.2018.00631

Cruz Vieira, A. B., Weber, A. A., Ribeiro, Y. M., Luz, R. K., Bazzoli, N., and Rizzo, E. (2019). Influence of salinity on spermatogenesis in adult Nile tilapia (*Oreochromis niloticus*) testis. *Theriogenology* 131, 1–8. doi:10.1016/j.theriogenology.2019.03.013

Dmitrieva, N. I., Ferraris, J. D., Norenburg, J. L., and Burg, M. B. (2006). The saltiness of the sea breaks DNA in marine invertebrates: Possible implications for animal evolution. *Cell Cycle* 5, 1320–1323. doi:10.4161/cc.5.12.2867

Dunn-Meynell, A. A., Routh, V. H., Kang, L., Gaspers, L., and Levin, B. E. (2002). Glucokinase is the likely mediator of glucosensing in both glucose-excited and glucose-inhibited central neurons. *Diabetes* 51, 2056–2065. doi:10.2337/diabetes.51.7.2056

El-Leithy, A. A. A., Hemeda, S. A., El Naby, W. S. H. A., El Nahas, A. F., Hassan, S. A. H., Awad, S. T., et al. (2019). Optimum salinity for Nile tilapia (*Oreochromis niloticus*) growth and mRNA transcripts of ion-regulation, inflammatory, stress- and immune-related genes. *Fish. Physiol. Biochem.* 45, 1217–1232. doi:10.1007/s10695-019-00640-7

El-Sayed, A.-F. M., Mansour, C. R., and Ezzat, A. A. (2005). Effects of dietary lipid source on spawning performance of Nile tilapia (*Oreochromis niloticus*) broodstock reared at different water salinities. *Aquaculture* 248, 187–196. doi:10.1016/j.aquaculture.2005.04.024

Evans, T. G., and Kültz, D. (2020). The cellular stress response in fish exposed to salinity fluctuations. *J. Exp. Zool.* 333, 421–435. doi:10.1002/jez.2350

FAO (2020). *The state of world Fisheries and aquaculture 2020: Sustainability in action, the state of world Fisheries and aquaculture (SOFIA)*. Rome, Italy: FAO. doi:10.4060/ca9229en

Fiol, D. F., and Kültz, D. (2007). Osmotic stress sensing and signaling in fishes. *FEBS J.* 274, 5790–5798. doi:10.1111/j.1742-4658.2007.06099.x

Gan, L., Xu, Z. X., Ma, J. J., Xu, C., Wang, X. D., Chen, K., et al. (2016). Effects of salinity on growth, body composition, muscle fatty acid composition, and antioxidant status of juvenile Nile tilapiaOreochromis niloticus(Linnaeus, 1758). *J. Appl. Ichthyol.* 32, 372–374. doi:10.1111/jai.12997

Gilmour, K. M., Perry, S. F., Esbaugh, A. J., Genz, J., Taylor, J. R., and Grosell, M. (2012). Compensatory regulation of acid-base balance during salinity transfer in rainbow trout (*Oncorhynchus mykiss*). *J. Comp. Physiol. B* 182, 259–274. doi:10.1007/s00360-011-0617-8

Gonzalez, R., and Unniappan, S. (2010). Molecular characterization, appetite regulatory effects and feeding related changes of peptide YY in goldfish. *General Comp. Endocrinol.* 166, 273–279. doi:10.1016/j.ygcen.2009.09.008

Gorissen, M., and Flik, G. (2016). "The endocrinology of the stress response in fish," in *Fish physiology* (Elsevier), 75–111. doi:10.1016/B978-0-12-802728-8.00003-5

Grosell, M. (2010). "The role of the gastrointestinal tract in salt and water balance," in *Fish physiology, the multifunctional gut of fish*. Editors M. Grosell, A. P. Farrell, and C. J. Brauner (Academic Press), 135–164. doi:10.1016/S1546-5098(10)03004-9

He, A., Luo, Y., Yang, H., Liu, L., Li, S., and Wang, C. (2011). Complete mitochondrial DNA sequences of the nile tilapia (*Oreochromis niloticus*) and blue tilapia (*Oreochromis aureus*): Genome characterization and phylogeny applications. *Mol. Biol. Rep.* 38, 2015–2021. doi:10.1007/s11033-010-0324-7

Iwama, G. K., Takemura, A., and Takano, K. (1997). Oxygen consumption rates of tilapia in fresh water, sea water, and hypersaline sea water. *J. Fish Biol.* 51, 886–894. doi:10.1111/j.1095-8649.1997.tb01528.x

Kammerer, B. D., and Kültz, D. (2009). Prolonged apoptosis in mitochondria-rich cells of tilapia (*Oreochromis mossambicus*) exposed to elevated salinity. *J. Comp. Physiol. B* 179, 535–542. doi:10.1007/s00360-008-0333-1

Kammerer, B. D., Sardella, B. A., and Kültz, D. (2009). Salinity stress results in rapid cell cycle changes of tilapia (*Oreochromis mossambicus*) gill epithelial cells. *J. Exp. Zool.* 311A, 80–90. doi:10.1002/jez.498

Kang'ombe, Jeremiah, and Joseph A., Brown. (2008). Effect of salinity on growth, feed utilization, and survival of Tilapia rendalli under laboratory conditions. *Journal of Applied Aquaculture* 20.4, 256–271.

Kristensen, P., Judge, M. E., Thim, L., Ribel, U., Christjansen, K. N., Wulff, B. S., et al. (1998). Hypothalamic CART is a new anorectic peptide regulated by leptin. *Nature* 393, 72–76. doi:10.1038/29993

Leal, E., Fernández-Durán, B., Guillot, R., Ríos, D., and Cerdá-Reverter, J. M. (2011). Stress-induced effects on feeding behavior and growth performance of the sea bass (*Dicentrarchus labrax*): A self-feeding approach. *J. Comp. Physiol. B* 181, 1035–1044. doi:10.1007/s00360-011-0585-z

Liu, Y., Wang, W.-N., Wang, A.-L., Wang, J.-M., and Sun, R.-Y. (2007). Effects of dietary vitamin E supplementation on antioxidant enzyme activities in Litopenaeus vannamei (Boone, 1931) exposed to acute salinity changes. *Aquaculture* 265, 351–358. doi:10.1016/j.aquaculture.2007.02.010

Luz, R. K., Martínez-Álvarez, R. M., De Pedro, N., and Delgado, M. J. (2008). Growth, food intake regulation and metabolic adaptations in goldfish (*Carassius auratus*) exposed to different salinities. *Aquaculture* 276, 171–178. doi:10.1016/j.aquaculture.2008.01.042

Martinez-Alborcia, M. J., Valverde, A., Parrilla, I., Vazquez, J. M., Martinez, E. A., and Roca, J. (2012). Detrimental effects of non-functional spermatozoa on the freezability of functional spermatozoa from boar ejaculate. *PLOS ONE* 7, e36550. doi:10.1371/journal.pone.0036550

Marty, N., Dallaporta, M., and Thorens, B. (2007). Brain glucose sensing, counterregulation, and energy homeostasis. *Physiology* 22, 241–251. doi:10.1152/physiol.00010.2007

Mobbs, C. V., Isoda, F., Makimura, H., Mastaitis, J., Mizuno, T., Shu, I.-W., et al. (2005). Impaired glucose signaling as a cause of obesity and the metabolic syndrome: The glucoadipostatic hypothesis. *Physiology Behav.* 85, 3–23. doi:10.1016/j.physbeh.2005.04.005

Mohamed, N. A., Saad, M. F., Shukry, M., El-Keredy, A. M. S., Nasif, O., Van Doan, H., et al. (2021). Physiological and ion changes of Nile tilapia (*Oreochromis niloticus*) under the effect of salinity stress. *Aquac. Rep.* 19, 100567. doi:10.1016/j.aqrep.2020.100567

Murashita, K., Kurokawa, T., Nilsen, T. O., and Rønnestad, I. (2009). Ghrelin, cholecystokinin, and peptide YY in Atlantic salmon (*Salmo salar*): Molecular cloning and tissue expression. *General Comp. Endocrinol.* 160, 223–235. doi:10.1016/j.ygcen.2008.11.024

Narnaware, Y. K., and Peter, R. E. (2001). Neuropeptide Y stimulates food consumption through multiple receptors in goldfish. *Physiology Behav.* 74, 185–190. doi:10.1016/S0031-9384(01)00556-X

Ninh, N. H., Thoa, N. P., Knibb, W., and Nguyen, N. H. (2014). Selection for enhanced growth performance of Nile tilapia (*Oreochromis niloticus*) in brackish water (15-20 ppt) in Vietnam. *Aquaculture* 428-429, 1–6. doi:10.1016/j.aquaculture.2014.02.024

Pankhurst, N. W. (2011). The endocrinology of stress in fish: An environmental perspective. *General Comp. Endocrinol.* 170, 265–275. doi:10.1016/j.ygcen.2010.07.017

Rahmah, S., Liew, H. J., Napi, N., and Rahmat, S. A. (2020). Metabolic cost of acute and chronic salinity response of hybrid red tilapia Oreochromis sp. larvae. *Aquac. Rep.* 16, 100233. doi:10.1016/j.aqrep.2019.100233

Rønnestad, I., Gomes, A. S., Murashita, K., Angotzi, R., Jönsson, E., and Volkoff, H. (2017). Appetite-controlling endocrine systems in teleosts. *Front. Endocrinol.* 8, 73. doi:10.3389/fendo.2017.00073

Rubio, V. C., Sánchez-Vázquez, F. J., and Madrid, J. A. (2008). Role of cholecystokinin and its antagonist proglumide on macronutrient selection in European sea bass *Dicentrarchus labrax*, L. *Physiology Behav.* 93, 862–869. doi:10.1016/j.physbeh.2007.12.001

Sinha, A. K., AbdElgawad, H., Giblen, T., Zinta, G., De Rop, M. D., Asard, H., et al. (2014). Anti-oxidative defences are modulated differentially in three freshwater teleosts in response to ammonia-induced oxidative stress. *PLOS ONE* 9, e95319. doi:10.1371/journal.pone.0095319

Soengas, J. L. (2014). Contribution of glucose- and fatty acid sensing systems to the regulation of food intake in fish. A review. *General Comp. Endocrinol.* 205, 36–48. doi:10.1016/j.ygcen.2014.01.015

Stanley, S. A., Small, C. J., Murphy, K. G., Rayes, E., Abbott, C. R., Seal, L. J., et al. (2001). Actions of cocaine- and amphetamine-regulated transcript (CART) peptide on regulation of appetite and hypothalamo-pituitary axes *in vitro* and *in vivo* in male rats. *Brain Res.* 893, 186–194. doi:10.1016/S0006-8993(00)03312-6

Tran-Ngoc, K. T., Schrama, J. W., Le, M. T. T., Nguyen, T. H., Roem, A. J., and Verreth, J. A. J. (2017). Salinity and diet composition affect digestibility and intestinal morphology in Nile tilapia (*Oreochromis niloticus*). *Aquaculture* 469, 36–43. doi:10.1016/j.aquaculture.2016.11.037

Varnai, P., and Zakrzewska, K. (2004). DNA and its counterions: A molecular dynamics study. *Nucleic Acids Res.* 32, 4269–4280. doi:10.1093/nar/gkh765

Velan, A., Hulata, G., Ron, M., and Cnaani, A. (2011). Comparative time-course study on pituitary and branchial response to salinity challenge in Mozambique tilapia (*Oreochromis mossambicus*) and Nile tilapia (*O. niloticus*). *Fish. Physiol. Biochem.* 37, 863–873. doi:10.1007/s10695-011-9484-1

Verdegem, M. C. J., Hilbrands, A. D., and Boon, J. H. (1997). Influence of salinity and dietary composition on blood parameter values of hybrid red tilapia, *Oreochromis niloticus* (Linnaeus) x *O. mossambicus* (Peters). *Aquac Res.* 28, 453–459. doi:10.1046/j.1365-2109.1997.00880.x

Volkoff, H., Canosa, L. F., Unniappan, S., Cerdá-Reverter, J. M., Bernier, N. J., Kelly, S. P., et al. (2005). Neuropeptides and the control of food intake in fish. *General Comp. Endocrinol.* 142, 3–19. doi:10.1016/j.ygcen.2004.11.001

Volkoff, H., and Peter, R. E. (2000). Effects of CART peptides on food consumption, feeding and associated behaviors in the goldfish, *Carassius auratus*: Actions on neuropeptide Y- and orexin A-induced feeding. *Brain Res.* 887, 125–133. doi:10.1016/S0006-8993(00)03001-8

Whitehead, A., Zhang, S., Roach, J. L., and Galvez, F. (2013). Common functional targets of adaptive micro- and macro-evolutionary divergence in killifish. *Mol. Ecol.* 22, 3780–3796. doi:10.1111/mec.12316

Yan, M., Li, Z., Xiong, B., and Zhu, J. (2004). Effects of salinity on food intake, growth, and survival of pufferfish (Fugu obscurus). *J. Appl. Ichthyol.* 20, 146–149. doi:10.1046/j.1439-0426.2003.00512.x

Yan, P., Jia, J., Yang, G., Wang, D., Sun, C., and Li, W. (2017). Duplication of neuropeptide Y and peptide YY in Nile tilapia *Oreochromis niloticus* and their roles in food intake regulation. *Peptides* 88, 97–105. doi:10.1016/j.peptides.2016.12.010

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

Yokobori, E., Azuma, M., Nishiguchi, R., Kang, K. S., Kamijo, M., Uchiyama, M., et al. (2012). Neuropeptide Y stimulates food intake in the zebrafish, *Danio rerio*. *J. Neuroendocrinol.* 24, 766–773. doi:10.1111/j.1365-2826.2012.02281.x

Zakariah Mi, M., Wahab, W., and Jasmani, S. (2013). Histopathological and behavioral changes in Oreochromis sp. after exposure to different salinities. *J. Fish. Livest. Prod.* 01. doi:10.4172/2332-2608.1000103

Zhang, X., Gao, Y., Tang, N., Qi, J., Wu, Y., Hao, J., et al. (2018). One evidence of cocaine- and amphetamine-regulated transcript (CART) has the bidirectional effects on appetite in Siberian sturgeon (*Acipenser baerii*). *Fish. Physiol. Biochem.* 44, 411–422. doi:10.1007/s10695-017-0444-2

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 08 September 2022
DOI 10.3389/fgene.2022.925786

OPEN ACCESS

EDITED BY
Tony Silveira,
Federal University of Rio Grande, Brazil

REVIEWED BY
Arndt Telschow,
Osnabrück University, Germany
Alexey Kryukov,
Far Eastern Branch of the Russian Academy of Sciences, Russia

*CORRESPONDENCE
Emily A. Beck,
ebeck8@uoregon.edu
William A. Cresko,
wcresko@uoregon.edu

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

RECEIVED 21 April 2022
ACCEPTED 09 August 2022
PUBLISHED 08 September 2022

CITATION
Beck EA, Bassham S and Cresko WA (2022), Extreme intraspecific divergence in mitochondrial haplotypes makes the threespine stickleback fish an emerging evolutionary mutant model for mito-nuclear interactions.
Front. Genet. 13:925786.
doi: 10.3389/fgene.2022.925786

Extreme intraspecific divergence in mitochondrial haplotypes makes the threespine stickleback fish an emerging evolutionary mutant model for mito-nuclear interactions

Emily A. Beck [1,2†*], Susan Bassham [1] and William A. Cresko [1,2*]

[1]Institute of Ecology and Evolution, University of Oregon, Eugene, OR, United States, [2]Presidential Initiative in Data Science, University of Oregon, Eugene, OR, United States

Mitochondrial DNA is primarily maternally inherited in most animals and evolves about 10 times faster than biparentally inherited nuclear DNA. Mitochondrial dysfunction (mt-dys) arises when interactions between the co-evolving mitochondrial and nuclear genomes are perturbed in essential processes like oxidative phosphorylation (OXPHOS). Over time mt-dys can lead to mitochondrial diseases (mt-diseases), which are surprisingly prevalent and include common diseases such as Alzheimer's, Parkinson's, and diabetes. Unfortunately, the strong impact that intraspecific mitochondrial and nuclear genetic variation has on mt-disease complicates its study and the development of effective treatments. Animal models have advanced our understanding of mt-disease but their relevance to human conditions is often limited by their relatively low nuclear genetic diversity. Many traditional laboratory models also typically have a single mitochondrial haplotype (mitotype), in stark contrast to over 5,000 mitotypes in humans worldwide. The threespine stickleback fish has an evolutionary history that has made it a favorable evolutionary mutant model (EMM) for studying mito-nuclear interactions and possibly mt-diseases. EMMs are species with naturally evolved states that mimic maladaptive human diseases. In threespine stickleback, a period of isolation followed by introgression of the mitochondrial genome from a sister species resulted in the maintenance of two distinct mitochondrial haplotypes which continue to segregate within many populations of wild stickleback. The existence of two mitogenomes segregating in numerous genetically diverse populations provides a unique system for exploring complex mito-nuclear dynamics. Here we provide the first complete coding region analysis of the two threespine stickleback mitotypes, whose mitogenomic divergence exceeds that of other mammalian models for mitochondrial disease and even that between ancient and modern humans. We find that divergence is not uniform across the mitogenome, but primarily impacts protein coding genes, and significantly impacts proteins in Complex I of OXPHOS. The full characterization of these highly divergent intraspecific mitotypes provides a

foundation for the development of threespine stickleback as an EMM for mito-nuclear interactions.

KEYWORDS

mitochondrial disease, mitogenome, co-evolution, dysfunction, outbred

Introduction

Typical metazoan mitochondrial genomes (mitogenomes) are small circular genomes encoding 13 proteins, two rRNAs and 22 tRNAs. Mitogenomes are overwhelmingly maternally inherited in most animals in contrast to the biparentally inherited nuclear genome (Hutchison et al., 1974; Giles et al., 1980), and they evolve rapidly compared to nuclear counterparts due to the less sophisticated prokaryotic DNA repair machinery present in mitochondria (Haag-Liautard et al., 2008; Itsara et al., 2014; Radzvilavicius et al., 2016). Despite these partially independent evolutionary trajectories, components of the mitogenome and nuclear genome must correctly interact to perform essential functions like oxidative phosphorylation (OXPHOS), innate immunity, and regulation of apoptosis (Tait and Green 2010; Cloonan and Choi 2013; Mishra and Chan 2014). Functional incompatibilities between the faster evolving mitogenome and the slower evolving nuclear genome can cause dysfunction (Burton and Barreto 2012; Mishra and Chan 2014; Wolff et al., 2014) which leads to a range of problems including accumulation of mutagenic reactive oxygen species (ROS) and reduction of ATP production, which overtime can lead to mitochondrial disease (mt-disease) (Dautant et al., 2018; Hahn and Zuryn 2019).

Mt-dys and resultant mt-diseases are surprisingly prevalent in humans. Primary mitochondrial diseases (PMD), are defined as diseases caused by a mutation in a protein comprising a subunit of OXPHOS, including diseases such as Leigh's Syndrome, Pearson Syndrome, and Barth Syndrome. The definition of mt-disease has further expanded to include secondary mitochondrial diseases (SMD). These include any disease exhibiting mt-dys as a symptom, such as Parkinson's Disease, Alzheimer's Disease, diabetes, cancer, muscular dystrophy, and amyotrophic lateral sclerosis (ALS) (Guo et al., 2013; Nicholson 2014; Niyazov et al., 2016). The identification of mt-dys as a pervasive issue underlying many common diseases has intensified interest in mt-dys. However, understanding mt-dys disease implications remains complicated by the strong role that mitochondrial and nuclear genetic variation has on symptom severity and treatment efficacy (Benit et al., 2010).

Clinical studies have been extremely valuable in identifying disease-causing genetic variants (Claussnitzer et al., 2020), and model organism research has helped improve mechanistic understanding of disease-causing variants via forward and reverse genetic approaches (Wangler et al., 2017; Cano-Gomez and Trynka 2020; Baldridge et al., 2021). Unfortunately, for several reasons forward- and reverse-genetic screens in model organisms often fail to identify subtle phenotypes that lead to disease later in life (Albertson et al., 2009; Beck et al., 2021) and are unable to fully recapitulate the roles nuclear and mitogenomic variation play in mt-disease (Benit et al., 2010). Most laboratory lines are inbred by design and contain a single mitochondrial haplotype (mitotype), poorly modeling the over 5,000 human mitotypes worldwide (Pipek et al., 2019). An important addition to human and traditional model organism research therefore comes from evolutionary genetics and the use of evolutionary mutant models (EMMs) (Albertson et al., 2009; Beck et al., 2021). EMMs do not exhibit a disease phenotype, but instead have adaptative modifications of homologous gene pathways that would cause disease in humans. EMMs therefore provide models for positive phenotypic outcomes in the context of disease mutations, guiding exploration of genetic compensation and discovery of novel therapeutics (Albertson et al., 2009; Beck et al., 2021). EMMs arise via similar mechanisms as disease mutations in human populations in organisms with diverse genetic backgrounds, making them excellent bridges between human and model organism research (Albertson et al., 2009; Beck et al., 2021).

Despite recent advances using EMMs to better understand several diseases (Albertson et al., 2009; Beck et al., 2021), a missing component has been an EMM with natural mitogenomic and nuclear genetic variation that can be used to understand mito-nuclear dynamics. Threespine stickleback fish (*Gasterosteus aculeatus*) have an evolutionary history that makes them well-suited EMMs for mito-nuclear interactions (Orti et al., 1994; Yamada et al., 2001; Lescak et al., 2014; Ravinet et al., 2018). During the Pleistocene Glacial Maximum, the Sea of Japan was largely enclosed, trapping a subset of *G. aculeatus* from the larger global population spread across the North Arctic (Orti et al., 1994). During a prolonged period of isolation, the trapped population of *G. aculeatus* hybridized with its sister species, *G. nipponicus*—the Japan Sea stickleback. Through hybridization the highly divergent mitochondrial genome from *G. nipponicus* introgressed to *G. aculeatus* along with parts of the nuclear genome resulting in a population of *G. aculeatus* hosting a mitogenome now called the Trans-North-Pacific (TNP) mitotype that had accumulated many mutations (Orti et al., 1994; Yamada et al., 2001; Lescak et al., 2014; Ravinet et al., 2018). When glacial ice retreated, the TNP threespine stickleback rejoined the oceanic threespine stickleback where the original Euro-North American (ENA) mitotype was still segregating. This mingling resulted in

interbreeding *G. aculeatus* with distinct mitotypes (Orti et al., 1994; Yamada et al., 2001; Lescak et al., 2014; Ravinet et al., 2018).

The existence of two independently evolved mitotypes within a natural population is promising for an EMM of mt-dys on its own. What makes the threespine stickleback system particularly unique is that the segregation of these two distinct mitotypes is happening in many regional freshwater and marine populations in the North Arctic, all with varying ratios of TNP/ENA mitotype individuals (Lescak et al., 2014). As a consequence, there are many natural 'experiments' generating countless mito-nuclear genetic combinations in the context of multiple environments. Still, there is little known about the extent of divergence and physiological impact of these mitotypes. Fortunately, threespine stickleback is a genetically amenable organism easily maintained in the laboratory and has been used for decades as a model for population genetics, behavioral studies, and developmental biology (Bell and Foster 1994; Colosimo et al., 2004; Cresko et al., 2004; Cresko et al., 2007; Hohenloe et al., 2010; Kimmel et al., 2012; Miller et al., 2014; Glazer et al., 2015; Greenwood et al., 2015; Peichel and Marques 2017). It is also an emerging model in immunology (Milligan-Mhyre et al., 2016; Small et al., 2017; Weber et al., 2017; Beck et al., 2020; Fuess et al., 2021). A high-quality reference genome is available (Jones et al., 2012; Peichel et al., 2017) and tools for genetic manipulation have been developed, making threespine stickleback an even more promising EMM for disease research.

Until now, the TNP and ENA threespine stickleback mitotypes have been distinguished solely based on divergence of the *cytb* gene for which a restriction digest assay was developed for rapid genotyping of individuals within a population (Orti et al., 1994; Lescak et al., 2014). While whole *Gasterosteus* mitogenomes have previously been assembled for studies of introgression (Ravinet et al., 2018), a complete analysis of the mitogenomes has been lacking. Here we present the first complete assessment of the ENA mitogenome coding sequence. We identify marked levels of nucleotide divergence exceeding that of other vertebrate sister species used to study mito-nuclear discordance, as well as surpassing divergence between modern and ancient humans. Interestingly, the observed divergence is primarily within protein coding genes and is disproportionately present in proteins involved in Complex I of OXPHOS, which could prove to be particularly useful in the study of mitochondrial diseases.

Materials and methods

Sample collection and data curation

We assembled mitogenomes for 34 threespine stickleback to be compared to the reference genome (TNP mitotype from Bear Paw Lake, Alaska) from Ensembl; total n = 36 mitogenomes. For population data, we extracted mitochondrial sequences from genomic data on the Sequence Read Archive (SRA) for six TNP fish from the Akkeshi River System in Japan and six ENA fish from Limfjord, Denmark. Accession numbers are summarized in Supplementary Table S1. We also generated new mitochondrial sequencing for 22 fish (ENA mitotype) collected from North America, including one ENA fish from Alaska in a region of TNP/ENA sympatry and for 21 allopatric fish from Oregon (ENA only) with population selection informed by Catchen et al., 2013 and Currey et al., 2019 (Supplementary Table S1). All collections were performed following University of Oregon Institutional Animal Care and Use Committee (IACUC) protocols. Oregon collections were approved under Oregon Department of Fish and Wildlife scientific taking permit numbers 19122 and 20,770. Alaska collections were approved under permit number SF-2011-153. For outgroup species comparisons, we obtained mouse (*Mus musculus* and *M m domesticus*), modern human (*Homo sapiens*), and ancient human (*H. s. denisova*, *H. s. neanderthalensis*, *H. heidelbergensis*) mitogenomes (GenBank numbers DQ874614.2, FJ374649.1, NC_012920.1, KX663333.1, MT677921.1, NC_023100.1 respectively).

Sequencing library preparation for mitogenomes

We isolated DNA from somatic tissue using standard methods and amplified the mitogenome from these genomic templates via PCR. To do so we designed five primer pairs (Supplementary Table S2) to amplify overlapping fragments that covered the mitogenome and that were optimized for long-range PCR using Phusion High-Fidelity DNA Polymerase. DNA concentrations post-amplification for each fragment were assessed using fluorometry via High Sensitivity Qubit (Supplementary Table S3). When possible, we standardized DNA to 200 ng and pooled fragments for each sample prior to library preparation. We sheared the pooled amplicons to a range of 300–400 base pairs (bp) using sonication (Covaris, Woburn, United States). Libraries were prepared from these fragments using end repair via NEB Quick blunting reactions, followed by addition of A-overhangs, T4 ligation to uniquely barcoded Illumina-compatible adapters, and total library amplification. We then multiplexed libraries and sequenced on the HiSeq4000 in the UO Genomics and Cell Characterization Core Facility (GC3F) generating paired-end 150 bp reads (Supplementary Table S1).

Mitogenome assemblies and annotations

We aligned demultiplexed 150 bp reads to the Bear Paw TNP reference mitogenome from Ensembl using the Burrows-Wheeler Aligner (BWA) (Li and Durbin 2009) and filtered using Samtools view (Li et al., 2009) to a minimum MAPQ score of 20. We called

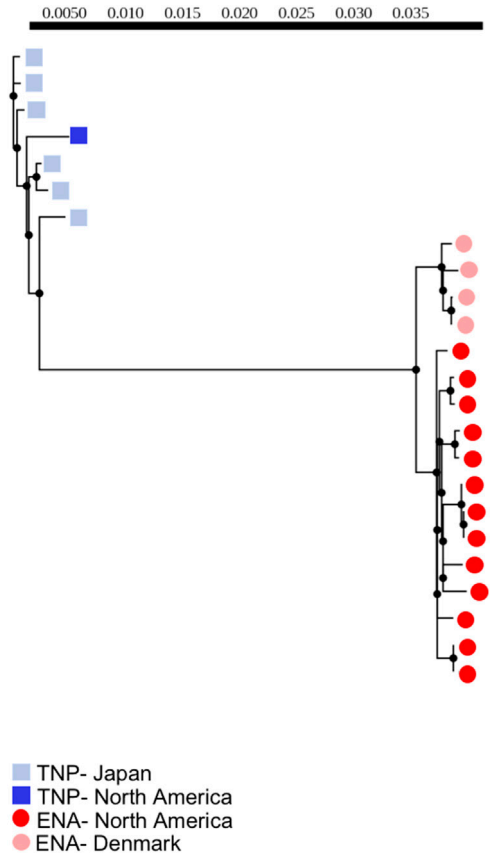


FIGURE 1
Maximum-likelihood phylogeny of threespine stickleback (*Gasterosteus aculeatus*) whole mitogenomes (excluding the control region) of individuals with no evidence of heteroplasmy. Branch lengths denote nucleotide divergence. Dark red circles indicate fish from North America with the Euro-North American (ENA) mitotype, light red circles indicate fish from Denmark with the ENA mitotype, dark blue squares indicate fish from North American with the Trans-North-Pacific (TNP) mitotype, light blue squares indicate fish from Japan with the TNP mitotype. Scale bar correlates to branch lengths and indicates level of divergence.

consensus sequence on filtered bam files using Geneious (version 10.2.2) and generated a multiple alignment of all consensus sequences in Geneious (version 11.0.12) (https://www.geneious.com) using MUSCLE. We then trimmed the multiple alignment to the limits of the coding region to exclude the partially assembled control region. Therefore, we included annotated 22 tRNAs, two rRNAs, and 13 coding genes based on the reference annotation from Ensembl. Because tRNAs were not specifically annotated in the reference, we confirmed the identity of each tRNA using tRNAscan-SE v2.0 (Chan et al., 2021). For external species comparisons (mouse and primate), we additionally performed multiple alignments within each group in Geneious (version 11.0.12) (https://www.geneious.com) via MUSCLE and annotated genes based on reference annotations from Ensembl.

Statistical analyses and accounting for heteroplasmy

We calculated pairwise percent sequence identity using Geneious (version 11.0.12) (https://www.geneious.com) and calculated population genetic statistics using DnaSP v6 (Rozas et al., 2003; Rozas et al., 2017). We also generated a whole mitogenome maximum-likelihood phylogeny using PhyML (Guindon et al., 2010). Due to the presence of heteroplasmy in some samples we had to account for ambiguous sites in our consensus sequences. Because we did not have population-level data for each of our samples we were unable to consistently PHASE genotypes. Instead, we opted for conservative approaches for estimating divergence and generating our phylogeny. For all analyses performed in DnaSP, all SNPs in all samples at locations containing ambiguous sites were removed from analyses, therefore lowering estimates of divergence. In Geneious, sites of heteroplasmy were treated as heterozygous sites using IUPAC standard coding. Percent identity was then adjusted based on the relationship between the heterozygous site and the other alleles to which it was being compared. For example, Y indicating a C/T heterozygous site is considered to have more identity with a C or T than a G or A. For the phylogeny generated in PhyML we also took a conservative approach, with heterozygous sites sharing an allele in the other populations being treated as non-divergent. For example, a Y is considered to be identical to a C or a T. In our phylogenies, we generated maximum-likelihood trees including all individuals (Supplementary Figure S1) and with only individuals showing no evidence of heteroplasmy (Figure 1). Finally, to assess the impact of genetic variants on tRNA structures we used tRNAscan-SE v 2.0 (Chan et al., 2021).

Results

Mitogenomic variation exists primarily between mitotypes irrespective of population

We assembled and annotated 34 mitogenomes from 24 populations to be compared to the reference genome (TNP) in order to assess divergence among and within populations, as well as between ENA and TNP mitotypes. These included six genomes from a single population in Denmark (ENA), six genomes from a single population in Japan (TNP), and 22 individuals from different populations in North America (ENA). To standardize comparisons to other models of mitochondrial divergence, we trimmed the assemblies to exclude the highly variable control region and included the full suite of 13 protein coding, two rRNA, and 22 tRNA genes.

We used maximum-likelihood phylogenetic analyses to assess divergence. We found that the majority of mitogenomic divergence was between mitotypes irrespective of geographic origin of

population, as indicated by a clear phylogenetic clustering of all ENA mitotype fish to the exclusion of TNP fish with branch lengths between mitotypes greatly exceeding those separating other groups (Figure 1). We additionally identified small levels of within-individual mitogenomic variation but did not identify any individual mitogenome sequence outliers exhibiting extreme patterns of divergence (Figure 1), suggesting that within-individual divergence is low compared to between-mitotype divergence, at least in the context of the single tissue sampled from each fish which prevents us from testing for tissue specific mutations within individuals. We also identified some population-level variation, particularly between allopatric Denmark and North American populations, with Denmark ENA fish clustering separately from North American ENA fish. Importantly, however, we found that the North American TNP fish - from a geographic region of admixture with the ENA mitotype - clustered with the TNP fish from Japan and not ENA North American individuals (Figure 1).

Mitogenomic variation is consistent with the life history of stickleback

We found that divergence between mitotypes was not due to positive selection by calculating Tajima's D. We did not identify any significant Tajima's D values—positive or negative—but did find largely negative Tajima's D values in the Japan populations consistent with a population expansion (Supplementary Table S4) (Tajima 1989). The finding of no selection is consistent with introgression, and not rapid evolution, as the presence of a second mitotype. Importantly, a population expansion in the TNP population is consistent with the previous hypothesis about the history of these mitotypes because the small population of TNP *G. aculeatus* would have undergone a population expansion at the end of the glacial maximum once they rejoined the Pacific Ocean populations (Orti et al., 1994; Yamada et al., 2001; Lescak et al., 2014; Ravinet et al., 2018).

Intraspecific stickleback mitotype variation exceeds mitogenomic variation between subspecies of mouse and between modern and ancient humans

We identified rates of intraspecific mitogenomic divergence between threespine stickleback mitotypes that exceeded that of previously identified notable intersubspecific mitogenomic divergence for two sets of subspecies pairs (Table 1). We compared ranges of whole stickleback mitogenome divergence (excluding the highly variable control region) with the sister species pair *Mus musculus* and *M.m. domesticus*, which are used to study mito-nuclear dynamics because of the high mitogenomic divergence between the taxa (Ma et al., 2016). We also made comparisons with the divergence between modern (*Homo sapiens*) and ancient humans (*H.s. neanderthalensis; H.s. denisova*; and *H. heidelbergensis*). These results confirm that what we document in the single stickleback species is an unusual amount of mitotype divergence, as it exceeds that of mitogenomic divergence between subspecies.

Mitochondrial heteroplasmy detected in some populations of threespine stickleback

In several populations we identified fish that had polymorphic nucleotide sites in their mitogenome assemblies, indicating either paternal leakage of mitochondria or spontaneous mutations that led to more than one mitochondrial haplotype within single individuals. Such a condition is called mitochondrial heteroplasmy (Supplementary Table S5). In all cases, evidence of heteroplasmy was found in allopatric ENA populations and we saw no evidence of both mitotypes present in individuals from regions of mitotype admixture. Unfortunately, given the limited sample size and lack of parentage data from our wild caught fish, we are unable to disentangle paternal leakage from spontaneous convergent mutations, but the abundance of heteroplasmy in specific populations could provide avenues to study genetic mechanisms underlying the maintenance of mitochondrial heteroplasmy.

Stickleback mitotype divergence lies primarily in protein coding genes

By assessing between mitotype divergence in each protein-coding and non-coding gene (tRNAs and rRNAs) through comparisons of North American ENA fish to the North American TNP reference genome (Figure 2), we found that between mitotype divergence is not uniform across the mitogenome. We did not assess intergenic content as mitogenomes contain little intergenic material—65 bp total in stickleback—outside of the excluded control region. Generally, we found that protein-coding gene divergence fell below genome-wide average and non-coding genes trended above average (Figure 2; Tables 2, 3) with a few exceptions: *trna-cys* exhibited below genome-wide average sequence identity and the protein-coding gene mt-*coII* had above average identity. For protein-coding genes there was no correlation between gene size and average between mitotype percent sequence identity (Adjusted R^2 = -0.052, p = 0.53) (Figure 2; Table 2). For non-coding genes there was also no correlation between size and average percent sequence identity (Adjusted R^2 = -0.006, p = 0.36) (Figure 2; Table 3), though it is likely that the observation of low sequence identity in *trna-cys* could be in part because it is the shortest length of any element in the mitogenome and therefore any change would have a proportionally large impact on its

TABLE 1 Mitogenomic coding divergence in threespine stickleback (*Gasterosteus aculeatus*) compared to mammalian models for mitochondrial divergence. TNP indicates the Trans-North-Pacific mitotype and ENA indicates the Euro-North American mitotype.

Comparison	Nucleotide divergence
Threespine stickleback (*G. aculeatus*) TNP vs. ENA	0.032–0.034
Mouse (*Mus musculus* vs. *M m domesticus*)	0.023
Human (*Homo sapien sapiens* vs. *H. s. neanderthalensis*)	0.011
Human (*Homo sapien sapiens* vs. *H. s. denisova*)	0.02
Human (*Homo sapien sapiens* vs. *H. s. heidelbergensis*)	0.026

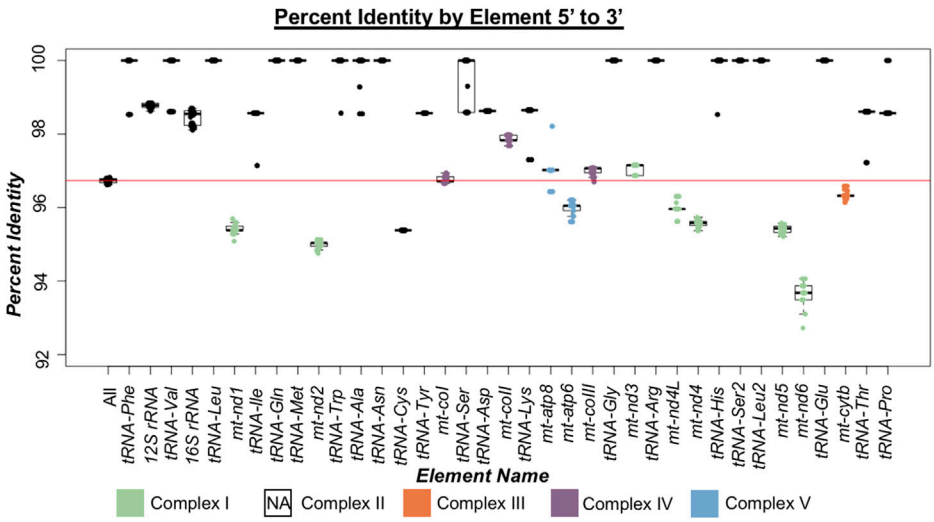


FIGURE 2
Percent identity of mitogenomic elements 5′ to 3′ in threespine stickleback (*Gasterosteus aculeatus*). Box plots indicate the range of percent identity between the Trans-North-Pacific (TNP) reference genome from North America and Euro-North American (ENA) mitogenomes from North America. The first box plot represents the percent identity across the whole mitogenome (excluding the control region). Box plots for each genic element are ordered based on their 5′ to 3′ orientation. Protein coding genes are colored based on which oxidative phosphorylation (OXPHOS) complex they comprise. Complex II of OXPHOS does not contain mitochondrially encoded proteins and therefore not represented indicated by NA (not applicable). The red horizontal line indicates average mitogenomic percent identity.

percent identity (Table 3; Supplementary Table S6). Analysis of Variance (ANOVA) across genic elements supports significant deviations from genome-wide averages in most elements (Tukey post hoc test summarized in Supplementary Table S7) including strikingly low sequence identity in several protein-coding genes especially mt-*nd6*, which exhibited the lowest sequence identity of any element (Figure 2).

Protein coding divergence is primarily in complex I of OXPHOS

We observed the highest rates of divergence in proteins involved in Complex I of OXPHOS comprised of ND1, ND2, ND3, ND4, ND4L, ND5, and ND6 (Figure 2) suggesting sequence divergence correlates with some aspect of protein function. ANOVA indicated significant deviations from genome wide average across all OXPHOS complexes, but Complex I exhibited the most divergence (Figure 3; Supplementary Table S8). This result was not surprising given that five of the seven proteins comprising Complex I exhibited the lowest sequence identity of any of the coding genes (ND1, ND2, ND4, ND5, and ND6). Complex II does not contain any mitochondrially encoded proteins and was therefore excluded from analysis. The single Complex III mitochondrial protein exhibited slightly below average sequence identity (Figure 3), while proteins from Complexes IV and V showed a wide range of values both above and below genome wide averages of sequence identity (Figures 2, 3). Combined, these data reveal a pattern of increased divergence in proteins functioning early in the OXPHOS Pathway.

TABLE 2 Between mitotype divergence in proteins in threespine stickleback (*Gasterosteus aculeatus*), excluding ambiguous sites.

Protein	Length (bp)[a]	D_N[b]	D_S[c]	P_N[d]	P_S[e]	N.I.[f]	P_{Yates}[g]
ND1	975	1	35	9	28	11.25	0.015
ND2	1047	3	38	10	27	4.69	0.04
ND3	351	0	6	2	6	NA	ns
ND4	1381	3	38	8	34	2.98	ns
ND4L	297	0	8	1	5	NA	ns
ND5	1839	5	58	19	47	4.69	0.004
ND6	522	3	18	2	17	0.71	ns
CytB	1141	1	26	5	28	4.64	ns
COI	1551	0	39	2	41	NA	ns
COII	691	0	12	3	11	NA	ns
COIII	786	0	18	3	19	NA	ns
ATPase6	684	0	17	4	26	NA	ns
ATPase8	168	0	3	2	4	NA	ns

[a]length in base pairs of the related protein coding gene based on the Trans-North-Pacific (TNP) reference genome;
[b]number of fixed nonsynonymous changes between mitotypes;
[c]number of fixed synonymous changes between mitotypes;
[d]number of polymorphic nonsynonymous changes within Euro-North American (ENA) fish from North America;
[e]number of polymorphic synonymous changes within ENA, North American fish;
[f]McDonald-Kreitman Neutrality Index;
[g]probability with Yates' correction. Horizontal lines denote protein groups by OXPHOS, Complex (I, III, IV, and V respectively).

We found evidence of strong purifying selection in some protein-coding genes when we combined all TNP samples from Japan and North America to compare with the North American ENA sample. Using a McDonald-Kreitman 2x2 test for selection (McDonald and Kreitman 1991), we assessed rates of synonymous and nonsynonymous polymorphisms within each population, and synonymous and nonsynonymous fixed differences between mitotypes. We observed an increase in synonymous changes compared to nonsynonymous changes indicating significant purifying selection (Table 2) and found no evidence of positive selection. A lack of positive selection was expected due to the introgression of the mitogenome (Orti et al., 1994; Yamada et al., 2001; Lescak et al., 2014; Ravinet et al., 2018), and strong purifying selection was expected based on previous evidence of consistently strong purifying selection on mitogenomic mutations (Stewart et al., 2008). There were, however, fixed nonsynonymous sites only in proteins from Complexes I and III as well as higher levels of nonsynonymous polymorphisms in these two complexes compared to late acting complexes IV and V, again supporting increased divergence early in OXPHOS (Table 2).

Discussion

Existing models for mt-dys and mt-disease lack some desirable attributes, such as natural genetic variation in nuclear and mitochondrial genomes. This deficit has hindered the study of mt-dys, which has been repeatedly shown to be heavily influenced by both mitochondrial, nuclear, and gene-by-environment variation (Benit et al., 2010; Zanon et al., 2018). There are important examples of hybridizable pairs of taxa with mitochondrial variation, including murine models *M m musculus* and *M.m domesticus* (Ma et al., 2016) and mummichog fish subspecies *Fundulus heteroclitus* and *F. heteroclitus macrolepidotus* (Flight et al., 2011; Baris et al., 2017). These existing outbred models, however, are commonly studied in single populations or at discrete, rare hybrid zones. What has been needed is a widespread, outbred model for mito-nuclear dynamics maintaining mitogenomic variation in multiple environments. Additional animal models exhibit mitochondrial-nuclear introgression with the mitogenome of the species with the larger effective population size introgressing and replacing that of the smaller effective population size species. A good example is the laboratory amenable system *Drosophila yakuba* and *D. santomea* (Llopart et al., 2005; Llopart et al., 2014; Beck et al., 2015) in which the *D. santomea* mitogenome was replaced by *D. yakuba*. In this case, no mitotype admixture is possible as only the *D. yakuba* mitotype exists. One possible explanation for the maintenance of divergence and admixture in the threespine stickleback system is the large global effective population size and widely dispersed populations of *G. aculeatus* (originally ENA mitotype) throughout the Northern Hemisphere. As such, there has not been enough time for the *G. nipponicus* TNP-mitogenome to replace the endogenous ENA-mitogenome (Ravinet et al., 2018). We are therefore studying this system at a very fortuitous point in evolutionary time where studies of mito-nuclear interactions can occur during a window of admixture.

TABLE 3 Polymorphisms in non-coding mitogenomic elements between and within mitotype in threespine stickleback (*Gasterosteus aculeatus*).

Gene	Length (bp)[a]	P_{TNP}[b]	P_{ENA}[c]	P_{shared}[d]	F[e]
12S-rRNA	946	5	6	0	9
16S-rRNA	1690	8	29	1	15
tRNA-asp	73	1	0	0	0
tRNA-cys	65	1	1	0	1
tRNA-his	68	0	1	0	0
tRNA-ile	70	1	1	0	1
tRNA-lys	74	2	1	0	0
tRNA-phe	68	0	2	0	0
tRNA-pro	70	0	1	0	0
*tRNA-ser-2**	68	1	0	0	0
tRNA-thr	72	0	2	0	1
tRNA-trp	70	0	1	0	0
tRNA-tyr	70	1	0	0	0
tRNA-val	72	0	1	0	0

Asterisk after number 2 indicates it is the 3' copy of *tRNA-Ser*.
[a]length in base pairs based on the Trans-North-Pacific (TNP) reference genome;
[b]polymorphisms exclusive to TNP mitotype;
[c]polymorphisms exclusive to the Euro-North American (ENA) mitotype;
[d]polymorphisms shared between TNP and ENA mitotypes;
[e]fixed differences between TNP and ENA mitotypes.

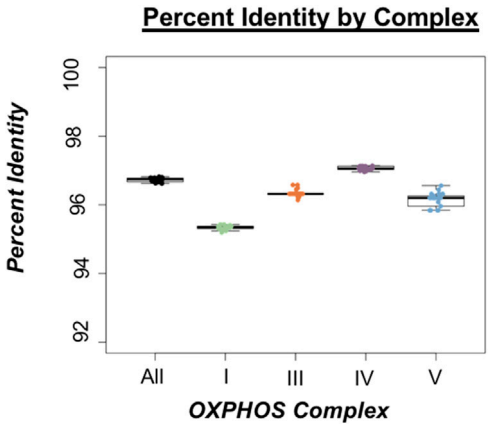


FIGURE 3
Percent identity of oxidative phosphorylation (OXPHOS) complexes in threespine stickleback (*Gasterosteus aculeatus*). Box plots indicate the spread of percent identity between the Trans-North-Pacific (TNP) reference genome from North America and Euro-North American (ENA) mitogenomes from North America. Color choices are consistent throughout the manuscript.

Our quantification of intraspecific mitotype variation in threespine stickleback makes a strong case for its use as an outbred model for mito-nuclear interactions. First, mitogenomic variation in threespine stickleback is primarily between two highly divergent mitotypes (TNP and ENA) rather than among populations or within individuals. Second, the observed level of divergence also exceeds that of existing outbred models with whole mitogenome assemblies including *M m musculus* and *M.m domesticus* and exceeds that of modern and ancient humans. Third, hybridization and admixture of stickleback mitotypes has occurred repeatedly in many different freshwater and oceanic environments, generating a multitude of natural experiments for the study of mito-nuclear dynamics in the context of environmental variation.

These strengths are extremely valuable in addressing questions concerning mito-nuclear dynamics in the context of environmental and genetic heterogeneity. Fortunately, while threespine stickleback are a promising EMM due to their repeated mitotype admixture in nature, they are also amenable to laboratory studies. As in zebrafish, developmental and physiological studies are facilitated by external fertilization in stickleback (Bell and Foster 1994). Threespine stickleback can also be crossed and reared for QTL/eQTL mapping studies (Greenwood et al., 2015; Peichel and Marques 2017; Beck et al., 2020), and they can be genetically manipulated using technology like CRISPR-Cas9 genome editing (Wucherpfennig et al., 2019; Xi et al., 2019; Kingman et al., 2021). Recent advances in mitogenome editing could theoretically be applied in this system to further enhance stickleback as an EMM for mitochondrial physiology (Hussain et al., 2021).

Another strength of stickleback as an EMM is the heterogeneity of divergence across the stickleback mitogenome. We observed divergence primarily in coding genes irrespective of their physical location in the mitogenome, with *mt-nd6* being a clear outlier for elevated levels of nucleotide divergence. Importantly, we discovered that much of the protein divergence is in those functioning early in OXPHOS in Complex I, even though Complex I proteins are spread throughout the mitogenome. More work is needed to understand what, if any, are the physiological consequences that exist because of divergence in Complex I proteins within populations. However, these findings are suggestive and intriguing, as Complex I is the primary producer of ROS, which, if not cleared from mitochondria by antioxidants and other proteins in Complex I, can lead to mt-dys and cellular damage (Hirst et al., 2008). As the largest complex in OXPHOS, Complex I presents a large target for mutation and is the source of many disease variants in humans (Carroll et al., 2003; Koopman et al., 2012; Barshad et al., 2019). Divergence in specific regions of the mitogenome is something that remains unexplored in many natural systems, in which studies have primarily used restriction digestion assays or targeted identification of SNPs to characterize mitotypes. Until now, threespine stickleback researchers have relied on a restriction digest to identify SNPs in a PCR amplicon within *mt-cytb* (Orti et al., 1994; Lescak et al., 2014). From our analyses we can

conclude that *mt-cytb* was a serendipitous choice as a representative of average mitotype divergence, but it missed the more extreme divergence identified in other areas of the mitogenome. We propose that full mitogenome sequencing is therefore required for the advancement of EMMs for mito-nuclear interactions.

Data availability statement

The data presented in the study are deposited in the Sequencing Read Archive (SRA) repository, accession numbers SAMN30467736 -SAMN30467757.

Ethics statement

The animal study was reviewed and approved by UO IACUC.

Author contributions

Experiments were designed by all authors (EAB; SB; WAC) and executed by EAB. The manuscript was written by EAB with contributions from SB and WAC.

Funding

This work was supported by the National Institutes of Health grant P50-DA04875602S2 from the National Institute on Drug Abuse (WC), the National Institutes of Health grant R24RR032670 from the National Institute of General Medical Sciences (WC), the UO Office of the Vice President of Research and Innovation (OVPRI) Faculty Research Award (EB) and the UO-OVPRI Incubating Interdisciplinary Initiatives (I3) grant (WC and EB), and the National Institute of Health NRSA fellowship F32GM122419 from the National Institute of General Medical Sciences (EB).

Acknowledgments

We would like to thank Carrie E. McCurdy and Byron Hetrick for helpful conversations surrounding mitochondrial physiology and Clayton M. Small for help with statistics and experimental design. We would also like to thank Emily Lescak, Ann Peterson, and Mark C. Currey for helping obtain samples.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.925786/full#supplementary-material

References

Albertson, R. C., Cresko, W., Detrcih, H. W., III, and Postlethwait, J. H. (2009). Evolutionary mutant models for human disease. *Trends Genet.* 25, 74–81. doi:10.1016/j.tig.2008.11.006

Baldridge, D., Wangler, M. F., Bowman, A. N., Yamamoto, S., Undiagnosed Disease NetworkSchedl, T., et al. (2021). Model organisms contribute to diagnosis and discovery in the undiagnosed diseases network: Current state and a future vision. *Orphanet J. Rare Dis.* 16, 206. doi:10.1186/s13023-021-01839-9

Baris, T. Z., Wagner, D. N., Dayan, D. I., Du, X., Blier, P. U., Pichaus, N., et al. (2017). Evolved genetic and phenotypic differences due to mitochondrial-nuclear interactions. *PLoS Genet.* 13, e1006517. doi:10.1371/journal.pgen.1006517

Barshad, G., Zlotnikov-Pozianski, N., Gal, L., Schildiner, M., and Mishmar, D. (2019). Disease-causing mutations in subunits of OXPHOS complex I affect certain physical interactions. *Sci. Rep.* 9, 9987. doi:10.1038/s41598-019-46446-8

Beck, E. A., Currey, M. C., Small, C. M., and Cresko, W. A. (2020). QTL mapping of intestinal neutrophil variation in threespine stickleback reveals possible gene targets connecting intestinal inflammation and systemic health. *G3 (Bethesda)* 10, 613–622. doi:10.1534/g3.119.400685

Beck, E. A., Healey, H. M., Small, C. M., Currey, M. C., Desvignes, T., Cresko, W. A., et al. (2021). Advancing human disease research with fish evolutionary mutant models. *Trends Genet.* 1, 22–44. doi:10.1016/j.tig.2021.07.002

Beck, E. A., Thompson, A. C., Sharbrough, J., Brud, E., and Llopart, A. (2015). Gene flow between *Drosophila yakuba* and *Drosophila santomea* in subunit V of cytochrome *c* oxidase: A potential case of cytonuclear cointrogression. *Evolution* 69, 1973–1986. doi:10.1111/evo.12718

Bell, M. A., and Foster, S. A. (1994). *The evolutionary biology of the threespine stickleback*. Oxford, UK: Oxford University Press.

Benit, P., El-Khoury, R., Schiff, M., Sainsard-Chanet, A., and Rustin, P. (2010). Genetic background influences mitochondrial function: Modeling mitochondrial disease for therapeutic development. *Trends Mol. Med.* 16, 210–217. doi:10.1016/j.molmed.2010.03.001

Burton, R. S., and Barreto, F. S. (2012). A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities? *Mol. Ecol.* 21, 4942–4957. doi:10.1111/mec.12006

Cano-Gomez, E., and Trynka, G. (2020). From GWAS to function: Using functional genomics to identify the mechanisms underlying complex diseases. *Front. Genet.* 11, 424. doi:10.3389/fgene.2020.00424

Carroll, J., Fearnley, I. M., Shannon, R. J., Hirst, J., and Walker, J. E. (2003). Analysis of the subunit composition of complex I from bovine heart mitochondria. *Mol. Cell. Proteomics* 2, 117–126. doi:10.1074/mcp.M300014-MCP200

Catchen, J., Bassham, S., Wilson, T., Currey, M., O'Brien, C., Yeates, Q., et al. (2013). The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Mol. Ecol.* 22, 2864–2883. doi:10.1111/mec.12330

Chan, P. P., Lin, B. Y., Mak, A. J., and Lowe, T. M. (2021). tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* 49, 9077–9096. doi:10.1093/nar/gkab688

Claussnitzer, M., Cho, J. H., Collins, R., Cox, N. J., Dermitzakis, E. T., Hurles, M. E., et al. (2020). A brief history of human disease genetics. *Nature* 577, 179–189. doi:10.1038/s41586-019-1879-7

Cloonan, S. M., and Choi, A. M. (2013). Mitochondria: Sensors and mediators of innate immune receptor signaling. *Curr. Opin. Microbiol.* 16, 327–338. doi:10.1016/j.mib.2013.05.005

Colosimo, P. F., Peichel, C. L., Nereng, K., Blackman, B. K., Shapiro, M. D., Schluter, D., et al. (2004). The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol.* 2, e109. doi:10.1371/journal.pbio.0020109

Cresko, W. A., Amores, A., Wilson, C., Murphy, J., Curry, M., Phillips, P., et al. (2004). Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc. Natl. Acad. Sci. U. S. A.* 101, 6050–6055. doi:10.1073/pnas.0308479101

Cresko, W. A., McGuigan, K. L., Phillips, P. C., and Postlethwait, J. H. (2007). Studies of threespine stickleback developmental evolution: Progress and promise. *Genetica* 129, 105–126. doi:10.1007/s10709-006-0036-z

Currey, M. C., Bassham, S. L., and Cresko, W. A. (2019). Genetic divergence outpaces phenotypic evolution among threespine stickleback populations in old freshwater habitats. *Biol. J. Linn. Soc. Lond.* 128, 415–434. doi:10.1093/biolinnean/blz106

Dautant, A., Meier, T., Hahn, A., Tribouillard-Tanvier, D., di Rago, J-P., and Kucharczyk, R. (2018). ATP synthase diseases of mitochondrial genetic origin. *Front. Physiol.* 9, 329. doi:10.3389/fphys.2018.00329

Flight, P. A., Nacci, D., Champlin, D., Whitehead, A., and Rand, D. M. (2011). The effects of mitochondrial genotype of hypoxic survival and gene expression in a hybrid population of the killifish, *Fundulus heteroclitis*. *Mol. Ecol.* 20, 4503–4520. doi:10.1111/j.1365-294X.2011.05290.x

Fuess, L. E., Weber, J. N., de Haan, S., Steinel, N. C., Shim, K. C., and Bolnick, D. I. (2021). Between-population differences in constitutive and infection-induced gene expression in threespine stickleback. *Mol. Ecol.* 30, 6791–6805. doi:10.1111/mec.16197

Giles, R. E., Blanc, H., Cann, H. M., and Wallace, D. C. (1980). Maternal inheritance of human mitochondrial DNA. *Proc. Natl. Acad. Sci. U. S. A.* 77, 6715–6719. doi:10.1073/pnas.77.11.6715

Glazer, A. M., Killingbeck, E. E., Mitros, T., Rokhsar, D. S., and Miller, C. T. (2015). Genome assembly improvement and mapping convergently evolved skeletal traits in sticklebacks with genotyping-by-sequencing. *G3 (Bethesda)* 5, 1463–1472. doi:10.1534/g3.115.017905

Greenwood, A. K., Ardekani, R., McCann, S. R., Dubin, M. E., Sullivan, A., Bensussen, S., et al. (2015). Genetic mapping of natural variation in schooling tendency in the threespine stickleback. *G3 (Bethesda)* 5, 761–769. doi:10.1534/g3.114.016519

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

Guo, C., Sun, L., ChenXueping, C., and Zhang, D. (2013). Oxidative stress, mitochondrial damage and neurodegenerative diseases. *Neural Regen. Res.* 8, 2003–2014. doi:10.3969/j.issn.1673-5374.2013.21.009

Haag-Liautard, C., Coffey, N., Houle, D., Lynch, M., Charlesworth, B., and Keightley, P. D. (2008). Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*. *PLoS Biol.* 6, e204. doi:10.1371/journal.pbio.0060204

Hahn, A., and Zuryn, S. (2019). Mitochondrial genome (mtDNA) mutations that generate reactive oxygen species. *Antioxidants* 8, 392. doi:10.3390/antiox8090392

Hirst, J., King, M. S., and Pryde, K. R. (2008). The production of reactive oxygen species by complex I. *Biochem. Soc. Trans.* 36, 976–980. doi:10.1042/BST0360976

Hohenloe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., and Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* 6, e1000862. doi:10.1371/journal.pgen.1000862

Hussain, S. A., Yalvac, M. E., Khoo, B., Eckardt, S., and McLaughlin, K. J. (2021). Adapting CRISPR/Cas9 system for targeting mitochondrial genome. *Front. Genet.* 12, 627050. doi:10.3389/fgene.2021.627050

Hutchison, C. A., 3rd, Newbold, J. E., Potter, S. S., and Edgell, M. H. (1974). Maternal inheritance of mammalian mitochondrial DNA. *Nature* 251, 536–538. doi:10.1038/251536a0

Itsara, L. S., Kennedy, S. R., Fox, E. J., Yu, S., Hewitt, J. J., Sanchez-Contreras, M., et al. (2014). Oxidative stress is not a major contributor to somatic mitochondrial DNA mutations. *PLoS Genet.* 10, e1003974. doi:10.1371/journal.pgen.1003974

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russel, P., Mauceli, E., Johnson, J., et al. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484, 55–61. doi:10.1038/nature10944

Kimmel, C. B., Ullmann, B., Currey, M., Hohenloe, P. A., and Cresko, W. A. (2012). Developmental dissociation in morphological evolution of the stickleback opercle. *Evol. Dev.* 14, 326–337. doi:10.1111/j.1525-142X.2012.00551.x

Kingman, G. A. R., Lee, D., Jones, F. C., Kingsley, D. M., and Bell, M. A. (2021). Longer or shorter spines: Reciprocal trait evolution in stickleback via triallelic regulatory changes in *Stanniocalcin2a*. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2100694118. doi:10.1073/pnas.2100694118

Koopman, W. J., Willems, P. H., and Smeitink, J. A. (2012). Monogenic mitochondrial disorders. *N. Engl. J. Med.* 366, 1132–1141. doi:10.1056/NEJMra1012478

Lescak, E. A., Marcotte, R. W., Kenney, L. A., von Hippel, F. A., Cresko, W. A., Sherbick, M. L., et al. (2014). Admixture of ancient mitochondrial lineages in three-spined stickleback populations from the North Pacific. *J. Biogeogr.* 42, 532–539. doi:10.1111/jbi.12426

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* 25, 1754–1760. doi:10.1093/bioinformatics/btp324

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

Llopart, A., Herrig, D., Brud, E., and Stecklein, Z. (2014). Sequential adaptive introgression of the mitochondrial genome in *Drosophila yakuba* and. *Mol. Ecol.* 23, 1124–1136. doi:10.1111/mec.12678

Llopart, A., Lachaise, D., and Coyne, J. A. (2005). Multilocus analysis of introgression between two sympatric sister species of *Drosophila: Drosophila yakuba* and *D. santomea*. *Genetics* 171, 197–210. doi:10.1534/genetics.104.033597

Ma, H., Marti Gutierrez, N., Morey, R., Van Dyken, C., Kang, E., Hayama, T., et al. (2016). Incompatibility between nuclear and mitochondrial genomes contributes to an interspecies reproductive barrier. *Cell. Metab.* 24, 283–294. doi:10.1016/j.cmet.2016.06.012

McDonald, J. H., and Kreitman, M. (1991). Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351, 652–654. doi:10.1038/351652a0

Miller, C. T., Glazer, A. M., Summers, B. R., Blackman, B. K., Norman, A. R., Shapiro, M. D., et al. (2014). Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait loci. *Genetics* 1, 405–420. doi:10.1534/genetics.114.162420

Milligan-Mhyre, K. C., Small, C. M., MIttge, E., Agarwal, M., Currey, M. C., Cresko, W. A., et al. (2016). Innate immune responses to gut microbiota differ between oceanic and freshwater threespine stickleback populations. *Dis. Model. Mech.* 9, 187–198. doi:10.1242/dmm.021881

Mishra, P., and Chan, D. C. (2014). Mitochondrial dynamics and inheritance during cell division, development and disease. *Nat. Rev. Mol. Cell. Biol.* 15, 634–646. doi:10.1038/nrm3877

Nicholson, G. L. (2014). Mitochondrial dysfunction and chronic disease: Treatment with natural supplements. *Integr. Med.* 13, 35–43.

Niyazov, D. M., Kahler, S. G., and Frye, R. E. (2016). Primary mitochondrial disease and secondary mitochondrial dysfunction: Importance of distinction for diagnosis and treatment. *Mol. Syndromol.* 3, 122–137. doi:10.1159/000446586

Orti, G., Bell, M. A., Reimchen, T. E., and Meyer, A. (1994). Global survey of mitochondrial DNA sequences in the threespine stickleback: Evidence for recent migrations. *Evolution* 43, 608–622. doi:10.1111/j.1558-5646.1994.tb01348.x

Peichel, C. L., and Marques, D. A. (2017). The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372, 20150486. doi:10.1098/rstb.2015.0486

Peichel, C. L., Sullivan, S. T., Liachko, I., and White, M. A. (2017). Improvement of the threespine stickleback genome using a hi-C-based proximity-guided assembly. *J. Hered.* 108, 693–700. doi:10.1093/jhered/esx058

Pipek, O. A., Medgyes-Horvath, A., Dobos, L., Steger, J., Szalai-Gindl, J., Visontai, D., et al. (2019). Worldwide human mitochondrial haplogroup distribution from urban sewage. *Sci. Rep.* 9, 11624. doi:10.1038/s41598-019-48093-5

Radzvilavicius, A. L., Hadjivasiliou, Z., Pomiankowski, A., and Lane, N. (2016). Selection for mitochondrial quality drives evolution of the germline. *PLoS Biol.* 14, e2000410. doi:10.1371/journal.pbio.2000410

Ravinet, M., Yoshida, K., Shigenobu, S., Toyoda, A., Fujiyama, A., and Kitano, J. (2018). The genomic landscape at a late stage of stickleback speciation: High genomic divergence interspersed by small localized regions of introgression. *PLoS Genet.* 14, e1007358. doi:10.1371/journal.pgen.1007358

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Mol. Biol. Evol.* 34, 3299–3302. doi:10.1093/molbev/msx248

Rozas, J., Sanchez-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

Small, C. M., Milligan-Mhyre, K., Bassham, S., Guillemin, K., and Cresko, W. A. (2017). Host genotype and microbiota contribute asymmetrically to transcriptional variation in the threespine stickleback gut. *Genome Biol. Evol.* 9, 504–520. doi:10.1093/gbe/evx014

Stewart, J. B., Freyer, C., Elson, J. L., Wrendenberg, A., Cansu, Z., Trifunovic, A., et al. (2008). Strong purifying selection in transmission of mammalian mitochondrial DNA. *PLoS Biol.* 6, e10. doi:10.1371/journal.pbio.0060010

Tait, S. W. G., and Green, D. R. (2010). Mitochondria and cell death: Outer membrane permeabilization and beyond. *Nat. Rev. Mol. Cell. Biol.* 11, 621–632. doi:10.1038/nrm2952

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595. doi:10.1093/genetics/123.3.585

Wangler, M. F., Yamamoto, S., Chao, H-T., Posey, J., Westerfield, M., Postlethwait, J., et al. (2017). Model organisms facilitate rare disease diagnosis and therapeutic research. *Genetics* 1, 9–27. doi:10.1534/genetics.117.203067

Weber, J. N., Steinel, N. C., Shim, K. C., and Bolnick, D. I. (2017). Recent evolution of extreme cestode growth suppression by a vertebrate host. *Proc. Natl. Acad. Sci. U. S. A.* 114, 6575–6580. doi:10.1073/pnas.1620095114

Wolff, J. N., Ladoukakis, E. D., Enriquez, J. A., and Dowling, D. K. (2014). Mitonuclear interactions: Evolutionary consequences over multiple biological scales. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20130443. doi:10.1098/rstb.2013.0443

Wucherpfennig, J. I., Miller, C. T., and Kingsley, D. M. (2019). Efficient CRISPR-Cas9 editing of major evolutionary loci in sticklebacks. *Evol. Ecol. Res.* 20, 107–132.

Xi, K. T., Wang, G., Thompson, A. C., Wucherpfennig, J. I., Reimchen, T. E., Maccoll, A. D. C., et al. (2019). DNA fragility in the parallel evolution of pelvic reduction in stickleback fish. *Science* 363, 81–84. doi:10.1126/science.aan1425

Yamada, M., Higuchi, M., and Goto, A. (2001). Extensive introgression of mitochondrial DNA found between two genetically divergent forms of threespine stickleback, *Gasterosteus aculeatus*, around Japan. *Environ. Biol. Fishes* 61, 269–284. doi:10.1023/A:1010874117127

Zanon, A., Pramstaller, P. P., Hicks, A. A., and Pichler, I. (2018). Environmental and genetic variables influencing mitochondrial health and Parkinson's Disease penetrance. *Park. Dis.* 2018, 8684906. doi:10.1155/2018/8684906

frontiers | Frontiers in Genetics

TYPE Mini Review
PUBLISHED 30 November 2022
DOI 10.3389/fgene.2022.903990

OPEN ACCESS

EDITED BY
Tony Silveira,
Federal University of Rio Grande, Brazil

REVIEWED BY
Izani Acosta,
Federal University of Pelotas, Brazil
Vinicius Farias Campos,
Federal University of Pelotas, Brazil

*CORRESPONDENCE
George Shigueki Yasui,
yasui@usp.br

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

RECEIVED 25 March 2022
ACCEPTED 09 September 2022
PUBLISHED 30 November 2022

CITATION
Yasui GS, Ferreira do Nascimento N, Pereira-Santos M, Santos Silva APd, Coelho GCZ, Visintin JA, Porto-Foresti F, Okada Nakaghi LS, Vianna NC, Carvalho GB, Monzani PS, López LS and Senhorini JA (2022), Establishing a model fish for the Neotropical region: The case of the yellowtail tetra Astyanax altiparanae in advanced biotechnology. *Front. Genet.* 13:903990. doi: 10.3389/fgene.2022.903990

Establishing a model fish for the Neotropical region: The case of the yellowtail tetra *Astyanax altiparanae* in advanced biotechnology

George Shigueki Yasui[1,2,3,4]*, Nivaldo Ferreira do Nascimento[5], Matheus Pereira-Santos[6], Amanda Pereira dos Santos Silva[1,4], Geovanna Carla Zacheo Coelho[1,4], José Antônio Visintin[2], Fábio Porto-Foresti[7], Laura Satiko Okada Nakaghi[8], Norberto Castro Vianna[9], Gabriela Braga Carvalho[1,2], Paulo Sérgio Monzani[1,4], Lucia Suárez López[1,4] and José Augusto Senhorini[1,3,4]

[1]Laboratory of Fish Biotechnology, National Center for Research and Conservation of Continental Fish, Chico Mendes Institute of Biodiversity Conservation, Brasília, Brazil, [2]Department of Animal Reproduction, Faculty of Veterinary Medicine, University of São Paulo, São Paulo, Brazil, [3]Peixetec Biotecnologia Em Organismos Aquáticos LTDA, São Paulo, Brazil, [4]Graduate Course of Biological Sciences (Zoology), São Paulo State University, São Paulo, Brazil, [5]Academic Unit, Federal Rural University of Pernambuco (UFRPE), Serra Talhada, Brazil, [6]Federal Rural University of Rio de Janeiro, Animal Science Graduate Program, Seropédica, Brazil, [7]Department of Biological Sciences, São Paulo State University, São Paulo, Brazil, [8]Aquaculture Center, São Paulo State University, São Paulo, Brazil, [9]China Three Gorges Corporation (CTG), Beijing, China

The use of model organisms is important for basic and applied sciences. Several laboratory species of fishes are used to develop advanced technologies, such as the zebrafish (*Danio rerio*), the medaka (*Oryzias latipes*), and loach species (*Misgurnus* spp.). However, the application of these exotic species in the Neotropical region is limited due to differences in environmental conditions and phylogenetic distances. This situation emphasizes the establishment of a model organism specifically for the Neotropical region with the development of techniques that may be applicable to other Neotropical fish species. In this work, the previous research efforts are described in order to establish the yellowtail tetra *Astyanax altiparanae* as a model laboratory species for both laboratory and aquaculture purposes. Over the last decade, starting with artificial fertilization, the yellowtail tetra has become a laboratory organism for advanced biotechnology, such as germ cell transplantation, chromosome set manipulation, and other technologies, with applications in aquaculture and conservation of genetic resources. Nowadays, the yellowtail tetra is considered the most advanced fish with respect to fish biotechnology within the Neotropical region. The techniques developed for this species are being used in other related species, especially within the characins class.

KEYWORDS

chromosome set manipulation, experimental fish, germline chimera, laboratory fish, micromanipulation, polyploids

Introduction

Biological models are important to develop technologies in basic and applied sciences. In fish studies, the main application of model organisms focuses on small laboratory species and species for aquaculture production. Several fish species arose as laboratory species worldwide, such as the zebrafish (*Danio rerio*) (Westerfield, 2007; Feitsma and Cuppen, 2008; Dai et al., 2014), the medaka (*Oryzias latipes*) (Wittbrodt et al., 2002), and the loach (*Misgurnus* spp.) (Kostomarova, 1991; Arm, 2003), among other species (see Table 1). For the Neotropical region, there has been no model organism established for the laboratory work. A model organism specifically for the Neotropical region may improve the technologies in aquaculture, and the data are directly applicable for local conditions, as well as for other related species. Thus, the yellowtail tetra is considered a candidate for the model organism for the Neotropical region (Yasui et al., 2020c).

Several biological characteristics make this species a prime candidate for laboratory studies, including: 1) small size; 2) domestication into artificial conditions (aquaria and dry food); 3) early sex maturation (4–5 months); 4) easy breeding management for *in vitro* fertilization; and 5) external sexual dimorphism (Arai, 2001; Wittbrodt et al., 2002; Westerfield, 2007; Yasui et al., 2020c).

Although the use of laboratory species is interesting and may accelerate several technologies in the field of genetics and biotechnology, it is necessary to first establish the basic information for the successful application of a laboratory fish. Determination of characteristics, such as feeding and maintenance in the laboratory, environmental conditions (photoperiod, aquarium size, temperature for maintenance, and reproduction, etc.), reproduction, larvae culture, and disease treatment and prevention, is first necessary to then advance into other techniques such as transgenesis and chromosome manipulation. Most of the research topics in the field of biotechnology, such as transgenesis (Stahl et al., 2019), chromosome set manipulation (Dunham, 2004), primordial germ cell (PGC) transplantation (Yamaha et al., 2010), intracytoplasmic sperm injection (ICSI) (Yasui et al., 2018), and other biotechnological approaches, require the manipulation of embryos during the early stages, and therefore, knowledge of fertilization timing is necessary.

This review shows the main techniques developed in the field of biotechnology to establish the yellowtail tetra as the most advanced laboratory fish native to the Neotropical region.

The yellowtail tetra *Astyanax altiparanae*

The generic name "tetra" denotes several small-bodied species of fishes belonging to Characidae from the Neotropical region, although an African group also exists (subfamily Alestidae). The name "tetra" was originated from the genus *Tetragonopterus*, an important genus in this group. In this group, the genus *Astyanax* is widely distributed across America, from the south of Argentina to North America. Although tetra species are commonly associated with aquarium fish trade, some aquaculture species intended for production also exist, such as the yellowtail tetra *Astyanax altiparanae* (Garutti and Britski, 2000). This fish is a small-bodied species (12–15 cm) and largely distributed throughout the Neotropical region. *Astyanax altiparanae* is considered to be a junior synonym of *Astyanax lacustris* by Lucena and Soares, (2016). However, this recent classification is still not unanimously agreed upon by ichthyologists. Therefore, in the present review, the traditional classification will be used, and the name *Astyanax altiparanae* (Garutti and Britski, 2000) will be adopted.

The yellowtail tetra adapts easily into aquaria, aquaculture tanks, and ponds, and it can be fed with artificial commercial pellets. The intertidal spawning pattern allows it to be bred year-round, given that the temperature and photoperiod are manipulated (Machado-Evangelista et al., 2019). The yellowtail tetra presents sexual dimorphism as shown in our recent study (Siqueira-Silva et al., 2020), noting that the male presents bony hooks in the anal and ventral fins that are not present in the females (Figure 1).

The yellowtail tetra is an interesting model species to develop biotechniques because experiments may be conducted in aquaria year-round. In addition, the results are applicable to other species, especially other characin species that include more than 1,150 species (Nelson et al., 2016).

The first step in fish biotechnology: Sperm storage and *in vitro* fertilization

For biotechnological studies, such as chromosome set manipulation and germ cell transplantation, it is necessary to manipulate the fertilization timing. In the case of the yellowtail tetra, studies regarding reproduction were previously conducted using natural spawning (without

TABLE 1 Biological characteristics of some model fish species.

Species	Size	Sex maturation	Fecundity (egg/female)	Blood sampling	*In vitro* fertilization
Yellowtail tetra (*Astyanax altiparanae*)	4–15 cm [1]	4 months [1]	11,086–31,720 [2]	Feasible [3]	Easy
Loach (*Misgurnus anguillicaudatus*)	13–17 cm [4]	1–2 years [5]	1,800–15,500 [6]	Feasible [7]	Easy
Medaka (*Oryzias latipes*)	3–4 cm [8]	2 months [9]	30–50/day [8]	Feasible [10]	Moderate
Stickleback (*Gasterosteus aculeatus*)	2.5–8 cm [11] < 10 cm [12]	1–2 years [13]	161–4,130 [14]	Feasible [15]	Difficult
Zebrafish (*Danio rerio*)	2.5–4.5 cm [16]	2.5–3 months [17]	300/week [18]	Feasible [19]	Moderate

1: Yasui et al. (2020c); 2: Sato et al. (2006); 3: Nascimento et al. (2020b); 4: Gao et al. (2014); 5: Lei and Sinica, (2020); 6: Suzuki (1983); 7: Gao et al. (2007); 8: Wittbrodt et al. (2002); 9: Wakamatsu et al. (2001); 10: Niimi and Imada, (2008); 11: Olsson et al. (2019); 12: Cresko et al. (2007); 13: Mehlis and Bakker, (2013); 14: Patimar et al. (2010); 15: Wirzinger et al. (2007); 16: Clark et al. (2018); 17: Nasiadka and Clark, (2012); 18: Hsu et al. (2007); 19: Carradice and Lieschke, (2008).

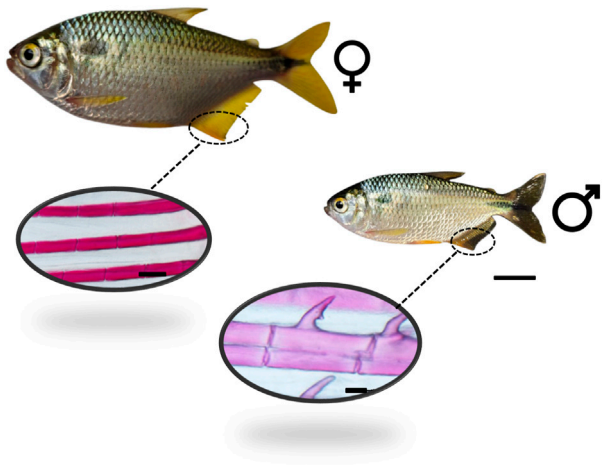


FIGURE 1
Adult male and female yellowtail tetra *Astyanax altiparanae*. The male presents bony hooks in the anal fin, which are not present in the females. Bar scale is 1 cm for the fish, 50 μm for the male fin, and 100 μm for the female fin.

hormonal treatment) and semi-natural spawning (hormonal treatment followed by spontaneous spawning) (Garutti, 1989, 2003; Veloso-Júnior et al., 2009; Weber et al., 2013). Artificial insemination was conducted in *A. bimaculatus* (Sato et al., 2006); however, the spermatozoa were not immobilized, and then, the timing of gamete activation was not controlled.

The first step to establish the yellowtail tetra in genetic and reproductive studies was to collect gametes and succeed with *in vitro* fertilization. Based on the loach *Misgurnus anguillicaudatus* procedures (Yasui et al., 2010, 2011), Yasui et al. (2015) evaluated several hormonal treatments for gamete maturation and then established sperm sampling and refrigerated storage in an extender. The extender allows one to immobilize and control the timing of sperm activation and fertilization. Oocyte sampling using Petri dishes was also important for laboratorial management for subsequent gamete and embryo manipulation. The same authors showed that oocyte storage was not possible in the species. After the studies mentioned previously, other advances in the reproduction of the yellowtail tetra have also been published (Brambila-Souza et al., 2021; Roza de Abreu et al., 2021).

Knowledge of gamete characteristics

After artificial propagation (Yasui et al., 2015), the next step in this line of research was to investigate basic characteristics of the gametes and embryo development. Pereira-Santos et al. (2016) analyzed the gametes, including ultrastructural analysis, second polar body extrusion, pronucleus fusion, and embryonic development at different temperatures. The spermatozoa of *A. altiparanae* have a typical morphology of the teleost fish, presenting a spherical head (1.88 μm), a midpiece (0.75 μm), and a single flagellum (18.67 μm). Temperature significantly influenced the development, where hatching occurred at 25 h post-fertilization (hpf) at 22°C, 16 hpf at 26°C, and 11 hpf at 30°C. At 22°C, extrusion of the second polar body occurred at 6 min post-fertilization (mpf), (Figure 2) and pronucleus fusion occurred at 10 mpf. This basic information gave important support for later works with chromosome manipulation and germ cell transplantation.

The first attempt to use computer-assisted sperm analysis (CASA) (Sperm Class Analyzer, Microptic, Barcelona, Spain) evaluated motility (%), linearity (LIN), beat cross-frequency (BCF), amplitude of lateral head (ALH), curvilinear velocity (VCL), straight line velocity (VSL), average path velocity (VAP), wobble (WOB), and straightness (STR). At 15 and 45 s post-activation, motility percentages were 83.9 ± 3.1% and 54.5 ± 5.5%, respectively, and the mean motility duration was 75 s. Sperm were maintained at 2.5°C in modified

Ringer's solution (128.3 mM NaCl, 23.6 mM KCl, 3.6 mM $CaCl_2$, and 2.1 mM $MgCl_2$) with good viability up to 3 days and later. Other studies implemented CASA with open-source software and successfully used it in experimentation on yellowtail tetra with similar results, using simpler parameters such as total motility, curvilinear velocity (VCL), average path velocity (VAP), and straight line velocity (VSL) (Gonçalves et al., 2018; Rocha et al., 2020).

The study of oocytes and fertilization success rates was also conducted (Pereira-Santos et al., 2017), and the researchers concluded that *A. altiparanae* has one of the lowest insemination doses among teleosts (2,390 spermatozoa. oocyte^{-1} ml^{-1}). Those conclusions were attributed to the small oocyte diameter (695.119 μm), large micropyle (7.57 μm), long motility duration (>75 s), and also the grooves in the oocytes surface that can guide the spermatozoa into the micropyle to optimize fertilization efficacy. This set of information was an important database to initiate advanced studies described in the following sections. The information obtained in the aforementioned studies also opened up new possibilities for approaches such as dispermic fertilization, since the micropyle diameter is greater than that of two sperm heads.

Chromosome set manipulation

Chromosome manipulation in fish refers basically to polyploidy (triploids and tetraploids, etc.) and uniparental inheritance induced by gynogenesis and androgenesis. Artificially induced polyploids focus on the production of triploids and tetraploids for large-scale production of sterile fish. Inhibition of second polar body extrusion, achieved by heat, cold, pressure, or chemical treatments, gives rise to triploid progenies (Dunham, 2004). The inhibition of second mitotic division by similar treatments may induce tetraploids (Zhang and Onozato, 2004). The main procedures for chromosome manipulation in the yellowtail tetra are listed in the following sections.

Triploids and hybrid triploids: Searching for sterile fish

Sterile yellowtail tetra is important for aquaculture because sterile fish present increased growth performance. In the field of conservation, sterile fish are important to avoid negative environmental impact from escaping, since in the Neotropical region, the introduction of exotic species is the second major cause of species endangerment (ICMBio, 2018). In addition, sterile fish may be a good recipient for cells of endangered species for subsequent surrogate propagation (Yamaha et al., 2001, 2007; Takeuchi et al., 2003), later serving as a repository gene bank. Previous attempts to obtain sterile yellowtail tetra were conducted by the depletion of germ cells, but the approach did not succeed a hundred percent in producing all sterile fish (Siqueira-Silva et al., 2015), so our group focused on polyploidization.

Initially, basic cytology information for chromosome set manipulation, including timing for second polar body release and fusion of male and female pronucleus, was studied (Santos et al., 2016). Based on such information, a more precise timing for diploidization and second polar body retention was established. This generated high percentages of triploids produced by using heat shock (40°C for 2 min) at 2 mpf, which guarantees 97.44% of triploids at the larvae stage (Adamov et al., 2017). The growth and reproductive performance of triploids were then studied (Nascimento et al., 2017a), showing that triploid females are sterile (Nascimento et al., 2017a) and present an increased carcass yield (%) (Nascimento et al., 2017b) when compared with diploids. On the other hand, triploid males were not sterile (Nascimento et al., 2017a; 2017b), limiting their application in aquaculture. In a later study, 100% sterile fish were achieved using triploid hybrids (Piva et al., 2018) in a special crossing of *A. altiparanae* and *A. fasciatus*. In this set of experiments, oocytes from yellowtail tetra *Astyanax altiparanae* were inseminated with sperm from five males (*A. altiparanae*, *A. fasciatus*, *A. schubarti*, *Hyphessobrycon anisitsi*, and *Oligosarcus pintoi*) in order to produce several interspecies hybrids and triploid hybrids. Surprisingly, only one cross (*A. altiparanae* x *A. fasciatus*) generated sterile offspring, and the progenies did not present germ lineage in the gonads.

The rise of spontaneously occurring triploids

Surprisingly, spontaneously occurring triploids arose in some progenies, even without any treatment for second polar body retention. This phenomenon led our group to investigate the rise of these triploids, and aged oocytes stored *in vivo* (Nascimento et al., 2018) and *in vitro* (Pereira-Santos et al., 2018) gave rise to these triploids, indicating that oocytes must be fertilized just after ovulation in order to prevent the rise of triploids.

Tetraploids

Based on the timing of pronucleus fusion (within 10 min) studied by Pereira-Santos et al. (2016) and the temperature (40°C) for heat shock obtained by Adamov et al. (2017), a

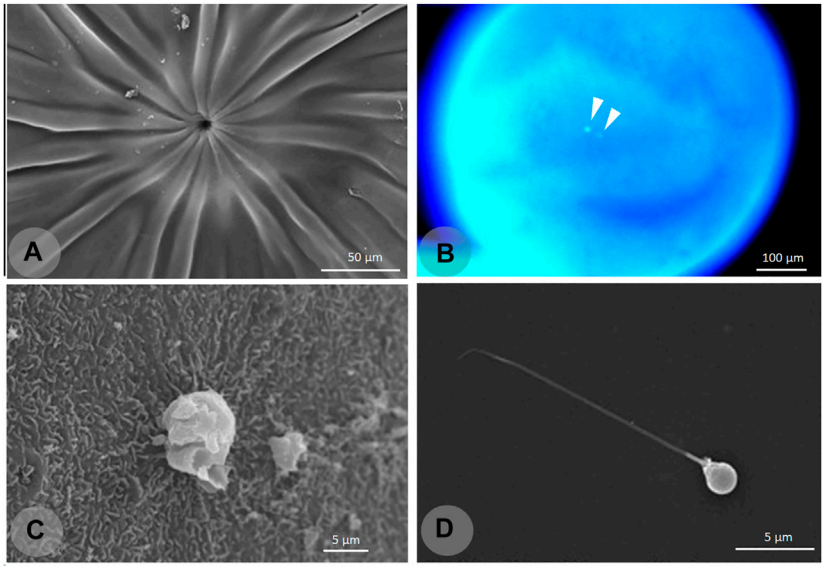


FIGURE 2
Cytological and ultrastructural images from the oocytes and spermatozoa of the yellowtail tetra *Astyanax altiparanae*. Ultrastructural analysis (scanning electron microscopy,SEM) of the oocyte micropyle **(A)** showing grooves that guide the sperm entry. DAPI staining of the fertilized oocyte showing male and female pronucleus **(B)**. Second polar body extrusion **(C)**. Spermatozoon of the yellowtail tetra **(D)**. This information was important to develop chromosome set manipulation techniques.

more precise tetraploidization procedure was established, improving the success rate of tetraploidization (Nascimento et al., 2020b). The procedures were optimized using temperature shock at 26 mpf (40°C for 2 min), followed by incubation at 26°C, and this resulted in 94.55% tetraploids at the larvae stage. This was the first reported tetraploids within the characin group. In addition, these protocols were recently improved (Martins et al., 2021), where it was observed that post-shock temperature (22°C, 26°C, and 28°C) affects tetraploid production in *A. altiparanae* and must be considered in future protocols. Tetraploid males and females were also able to produce viable and diploid spermatozoa and oocytes, respectively, which are capable of mass production (100%) of triploid fish (Nascimento et al., 2020b; Alves et al., 2022). As tetraploidization is difficult to achieve and viable lines are limited to a few species (Piferrer et al., 2009), the current protocol makes the yellowtail tetra one of the most successful species with respect to tetraploidization. The tetraploids are fertile, and they are being used to produce 100% triploid progenies.

Gynogenesis: Searching for monosex female progenies

Female yellowtail tetra fish are large and present increased growth performance when compared to males (Nascimento et al., 2017b), emphasizing the need for the establishment of a monosex female population in aquaculture. Studies on sex chromosomes using induced gynogenesis were then conducted (Nascimento et al., 2020a), indicating that the species presents an XX sex-determining system. This was the first attempt in gynogenesis within a Neotropical species, and most of the progenies reached 100% females (three out of four crosses). The resultant males derived from gynogenetic progenies were studied, concluding they were functional males (Lázaro et al., 2021).

Development of flow cytometric procedures

In order to assess the success of chromosome manipulation, flow cytometry is a valuable tool to confirm the ploidy status of the polyploids and the efficacy of chromosome inactivation and doubling in uniparental progenies (i.e., gynogenesis and androgenesis). An important step for chromosome manipulation in yellowtail tetra was to develop flow cytometric analysis using dorsal fin samples (Xavier et al., 2017). This technique is based on a two-step procedure with cell lysing and nuclear staining for the subsequent analysis. In addition, cold storage of the fin samples for the subsequent flow cytometric analysis was developed (Yasui et al., 2020a). Regarding chromosome

manipulation in a Neotropical species, the yellowtail tetra was the first to be analyzed using flow cytometry.

Larvae feeding under laboratory conditions

In aquaculture, the general procedure for larvae feeding and raising them into a juvenile stage consists of releasing the larvae into a fertilized pond containing plankton. On the other hand, it is important to feed the fish under laboratory conditions in small containers (Petri dishes, aquaria, or plastic containers). It was necessary then to understand in detail the adequate food for the yellowtail tetra. In order to fill this gap, Bertolini et al. (2018) evaluated different diets and concluded that artemia and dry food optimized the growth and survival under laboratory conditions, giving them a routine for dependable growth of larvae into the juvenile stage.

Prevention of sperm activation by urine

In small-bodied fish like the yellowtail tetra, sperm sampling is difficult to achieve because of small size. In the yellowtail tetra, the sperm is collected using a 1,000-μL micropipette. However, urine and sperm are released during sperm sampling, and the urine activates sperm motility and decreases the fertilization ability. The problem was partially solved using an immobilizing solution to re-immobilize the sperm (Yasui et al., 2015), although some activation still occurs. In a recent study (Rocha et al., 2020), the problem of activation was solved by maintaining males in a hyperosmotic environment (1% NaCl) for a few hours. The males were induced to spermatization and maintained in 1% NaCl for 6 h. The hyperosmotic environment concentrates the urine and does not activate the sperm motility at sampling. This procedure is now used for sperm sampling under laboratory conditions.

Surrogate propagation using yellowtail tetra

Surrogate propagation denotes a fish producing gametes from other fish. This approach is interesting for aquaculture and conservation of genetic resources. Considering the yellowtail tetra as a model organism, this species may be used to produce gametes from endangered fish species. In addition, cells from endangered species may be cryopreserved in liquid nitrogen, serving later as a repository gene bank. The yellowtail tetra is being used as a model for other endangered characin species, such as *Brycon orbignyanus*. Other aquaculture characins, such as the streaked prochilod (*Prochilodus lineatus*) and pacu (*Piaractus mesopotamicus*) (Coelho et al., 2019, 2021, respectively), are also being studied to serve as cell donors, using yellowtail tetra to produce gametes from those species. As the sterile host was already established (Piva et al., 2018), transplantation procedures are now being established for several characin species.

The PGCs of the yellowtail tetra were identified *in vivo* using a GFP-nos1 3′UTR mRNA from *Danio rerio*. The injection of this artificial mRNA resulted in the expression of GFP in the PGCs (Figure 3). In addition, the spermatogonial stem cells from endangered *B. orbygnianus* were successfully transplanted into sterile adults of yellowtail tetra (Figure 3). Those procedures are being established as repository procedures for endangered species.

The use of yellowtail tetra in other studies

The yellowtail tetra is also being used in other studies. For instance, based on the first attempt to analyze sperm motility by CASA (Pereira-Santos et al., 2016), motility parameters are being used in toxicological studies about pollutants, such as herbicides (Gonçalves et al., 2018) and aluminum (Pinheiro et al., 2020). The embryos and adults of yellowtail tetra species were used to evaluate toxicity of cyanopeptides (Fernandes et al., 2019) and aflatoxins (Michelin et al., 2017). An immunological study using triploids was conducted by Levy-Pereira et al. (2021), who observed that the cell counts of erythrocytes, leukocytes, and neutrophils were lower in triploid fish than diploids. Triploid erythrocytes were also larger with higher frequencies of abnormalities. Differences in gene expression related to immune response were also observed, reporting the lower expression of cytokine IL-1 (in the head, kidney, liver, and spleen) and TGF-β (in the spleen) in triploids. These results indicate that triploid fish present impaired immune systems and probably lower resistance to diseases. However, future studies involving pathogen challenges are necessary to confirm these assumptions.

Application of technologies in other species

Based on the results obtained for yellowtail tetra, several advances within other Neotropical species were established. Using the same temperature (40°C) to induce triploids in the yellowtail tetra, triploid progenies were obtained for the streaked prochilod *Prochilodus lineatus* (Yasui et al., 2020b) and *Brycon amazonicus* (Nascimento et al., 2021). With some minor

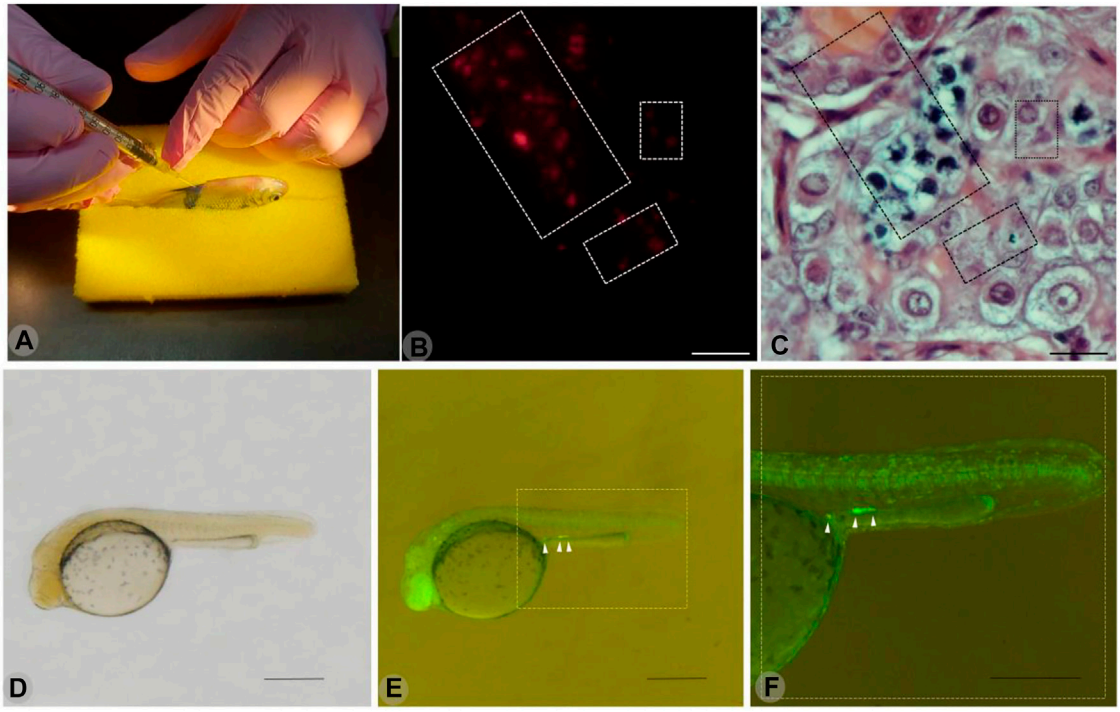


FIGURE 3
Transplantation of spermatogonial stem cells from endangered *Brycon orbignyanus* through the papillae of sterile juvenile yellowtail tetra *Astyanax altiparanae* **(A)**. Detached area indicates the testis containing transplanted cells stained with PKH26 and observed under fluorescence microscopy **(B)**. Same histological section visualized under normal light, and hematoxylin-eosin staining shows spermatocytes and spermatogonia from *B. orbignyanus*. Scale **(B–C)**: 50 μm. Below is a yellowtail tetra embryo injected with GFP-nos1 3′UTR mRNA from *Danio rerio*, observed under normal light **(D)**.Same embryo, observed under fluorescence microscopy, shows the presumptive primordial germ cells (PGCs) with GFP expression **(E)**. Gonadal ridge area in detail **(F)**.

modifications, triploids and tetraploids were also generated when the temperature was set at 38°C for *Pimelodus maculatus* (Bertolini et al., 2020) and *Rhamdia quelen* (Garcia et al., 2017; García et al., 2017). *B. amazonicus* (Da Silva et al., 2017), *Pimelodus maculatus* and *Pseudopimelodus mangurus* (Arashiro et al., 2018), and *P. lineatus* (Coelho et al., 2019) were also studied in the field of embryology using the same procedures and temperatures (22°C, 26°C, and 30°C) as used for the yellowtail tetra. The same protocol for cytometric analysis established in the yellowtail tetra (Xavier et al., 2017) was used to identify polyploids in *B. amazonicus* (Nascimento et al., 2021), *P. lineatus* (Yasui et al., 2020b), *P. maculatus* (Bertolini et al., 2020), and *R. quelen* (Garcia et al., 2017).

Future directions

Development of molecular markers is a priority for the yellowtail tetra, since it is the main confirmation tool for paternity within androgenesis, gynogenesis, and surrogate propagation. Techniques regarding cryopreservation of unusual fish developed recently are also being developed, such as gynogenetic, androgenetic, and polyploid genebanking. Cryopreservation of X spermatozoa obtained from spontaneously occurring gynogenetic males will be important to obtain monosex female populations. Regarding aquaculture, the establishment of sterile female populations will increase growth performance and also avoid the negative environmental impact caused by fish escaping. Transgenesis, single and multiple ICSI, and germ cell transplantation are among the ongoing studies in our group. However, several biotechniques currently used in other fish species have potential for *A. altiparanae*, such as the use of single nucleotide polymorphisms (SNPs), CRISPR/Cas9, transgenic, and microRNAs (miRNAs), and are discussed in the following paragraphs.

Clustered regularly interspaced palindromic repeats (CRISPR/Cas9) are DNA sequences that are used in genome editing technology. This technology, unprecedented in native fish species, has been increasingly used in aquaculture to manipulate reproduction and omega-3 content, growth, and metabolism (Straume et al., 2020; Sun et al., 2020), and the main model species for CRISPR/Cas9 in fish is the zebrafish (*Danio rerio*) (Chaudhary et al., 2020).

Transgenic fish, on the other hand, have already been developed for many fish species, such as the medaka (*Oryzias latipes*), zebrafish (*Danio rerio*), rainbow trout (*Oncorhynchus mykiss*), and loach (*M. anguillicaudatus*). Transgenic organisms receive DNA sequencing by artificial methods, incorporating one or more sequences into their chromosomal DNA. This type of animal is generally produced by microinjection or electroporation of newly fertilized oocytes or unfertilized gametes (egg or spermatozoa) (Maclean and Laight, 2000; Chen and Chen, 2020). Among the potential applications of transgenics in aquatic organisms, the increase in growth (Nam et al., 2001), tolerance to temperature (Cortemeglia and Beitinger, 2005) and salinity (El-Zaeem et al., 2014; Bystriansky et al., 2017), resistance to diseases (Dunham, 2009), and induction of sterility (Yu et al., 2011) are highlighted.

The SNPs are produced by mutations that occur in the genome, and its analysis has offered several applications in fish biology and aquaculture. An SNP marker was developed by liver transcriptome sequencing in pacu (*Piaractus mesopotamicus*), enabling a broad understanding of the population structure in the species and the possibility of elucidating adaptive mechanisms and manipulation of assisted reproduction. The establishment of SNPs also enabled the selection of groups with better genetic variability for storage and production (Mastrochirico-Filho et al., 2016).

The use of miRNAs, which are small and non-protein-coding RNA sequences, has increased the knowledge about several aspects of biological regulation mechanisms in animals and plants (Bizuayehu and Babiak, 2014; Mennigen, 2016). In fish, miRNAs are involved in several biological functions, such as regulation of development, organogenesis, growth, immune response, and reproduction (Mennigen, 2016; Andreassen and Høyheim, 2017; Tang et al., 2019; Zayed et al., 2019). As the study with miRNAs in fish are relatively recent and focused on a few species, more efforts are necessary.

In light of this, *A. altiparanae* can be considered the perfect model for Neotropical fish mainly because it is used in both basic and applied studies, such as aquaculture.

Discussion

The establishment of a model fish is strategic for basic and applied sciences. The yellowtail tetra is being successfully used for this purpose. Triploids, tetraploids, gynogenetic, and chimeric fish were successfully developed by means of advanced biotechnologies (Adamov et al., 2017; Nascimento et al., 2020a, 2020b). In the Neotropical region, the yellowtail tetra became the most advanced with regards to such technologies, having important implications in embryology, genetics, reproduction, cryobiology, and even in medical sciences such as flow cytometry (Santos et al., 2016; Xavier et al., 2017; Yasui et al., 2020a). Several characin species present critical reproduction challenges or are considered endangered (Silveira and Straube, 2008). Some of the migratory aquaculture species such as *Prochilodus lineatus*, *Piaractus mesopotamicus*, *Colossoma macropomum*, *Brycon amazonicus*, and *Salminus brasiliensis* present large size (2–15 kg), and sex maturation occurs within 2–3 years (Hainfellner et al., 2012; Pardo-Carrasco et al., 2006; da Costa and Mateus, 2009; Almeida et al., 2016; Barzotto and Mateus, 2017). The yellowtail tetra, notably also a characin species, may then be used to produce gametes from these characins. This would facilitate the reproductive management and accelerate techniques for genetic improvement, which requires several successive generations (Arai, 2001; Du et al., 2021).

Despite their importance to aquaculture and inland fisheries, the development of biotechniques is difficult to achieve because spawning management is more difficult. However, some techniques developed for the yellowtail tetra may be promptly used in other species. For instance, triploidization of *Brycon amazonicus* (Nascimento et al., 2021) and *Prochilodus lineatus* (Yasui et al., 2020b) was successful when using the same procedures of heat shocking for second polar body retention (2 mpf for 2 min, 40°C). In addition, confirmation of the ploidy status using flow cytometry was achieved using the protocol established for yellowtail tetra (Xavier et al., 2017; Yasui et al., 2020a).

In the case of chromosome set manipulation, cytological observation of post-insemination events (second polar body extrusion and pronucleus fusion) gave a precise timing for the successful manipulation of the reproduction cycle. The observation of second polar body release by histological sections and scanning electron microscopy improved the production of triploids and gynogenetic progenies. The use of fluorescent dyes such as DAPI (4′,6-diamidino-2-phenylindole) to observe pronucleus fusion (Itono et al., 2006) gave a precise timing for diploidization. Most of the previous protocols to obtain triploids and tetraploids used trial-and-error procedures (Piferrer et al., 2009) without cytological observations, which is time-consuming and more difficult to achieve.

Toxicological studies conducted on the yellowtail tetra are being used to establish safe concentrations for the environment (Gonçalves et al., 2018; Fernandes et al., 2019) and food products (Michelin et al., 2017), which also emphasize the importance of a model species.

Establishment of a model organism requires long-term research efforts for not only in the field of genetics and reproduction, but it also lays the foundation for complementary studies in the field of nutrition, physiology, animal behavior, water quality, immunology, pathology, and many other research fields. In addition, several confirmation tools become necessary for each of them (flow cytometry and water analysis, etc). In order to develop advanced biotechnologies in the yellowtail tetra, several initial stages were developed using trial-and-error experiments with most resulting in unpublished data, focusing on parameters that include container volume to maintain fish, materials (plastic and glass, etc.), type of food, substrate (as plants or tubes to avoid

aggressive behavior), and many other steps that are time-consuming. Although it is simple to evaluate, in general, these kinds of bottlenecks are not published in the literature, and the successful procedures and parameters take long periods to be developed. As seen previously, establishing a model organism is thus a multidisciplinary task, for which teamwork is extremely necessary. In the case of the yellowtail tetra, a part of the laboratorial procedures were adapted and transferred from the loach *Misgurnus anguillicaudatus*, but several biological differences did not allow for advancement in some research fields. In the case of gamete sampling, transfer of the procedures was successful, but the transplantation of embryonic cells was not possible in the yellowtail tetra because of the enzymes for chorion digestion and solutions to maintain the denuded embryos. For example, the hatching period from fertilization to hatching, depending on the temperature, takes 2–3 days in the loach (Fujimoto et al., 2006) and approximately 60 days in salmonids (Danner, 2008), but in the yellowtail tetra, hatching occurs in only 11 h (Santos et al., 2016). This difference makes PGC transplantation much more difficult in the yellowtail tetra, since transplantation is possible only within a few minutes. On the other hand, other biological features of the yellowtail tetra are interesting, such as the possibility of blood sampling and a subsequent serum analysis. Notably, this is difficult to achieve in some small-bodied fish like the medaka and the zebrafish.

In conclusion, the yellowtail tetra is an emerging experimental fish model for several research fields, especially with regards to genetics and reproduction. Basic and advanced studies were carried out during the last decade in order to establish this species as a model organism, and it is currently considered the most advanced model organism in techniques such as germ cell transplantation, micromanipulation, and chromosome set manipulation. Other research efforts are still ongoing.

Author contributions

GY, NN, JS and PM conceived and planned the experiments; NN, CC, GC, LL, and MP-S executed the experiments. GY, NN and JS wrote the first version. All authors contributed to the interpretation of the results and contributed to the final version of the manuscript. NV, F-PF. F, and NF performed all analyses. All authors discussed the results and contributed to the conclusion of the manuscript.

Funding

This study was supported by FAPESP (Young Investigator Award #2010/17429-1) and China Three Gorges Brasil (CTG Brasil, Project ANEEL #PD-00387-0418/2019). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Acknowledgments

Authors are grateful to FAPESP (Young Investigator Award #2010/17429-1) which permitted the initial works in the yellowtail tetra and China Three Gorges Brasil (CTG Brasil, Project ANEEL #PD-00387-0418/2019) for funding subsequent partial studies on the species. GYasui is grateful to the Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT), especially to Katsutoshi Arai and Takafumi Fujimoto for all the teachings at Hokkaido University, Japan, which were applied to develop most of the technologies in the yellowtail tetra in Brazil.

Conflict of interest

GY and JS were employed by Peixetec Biotecnologia Em Organismos Aquáticos LTDA. NV was employed by China Three Gorges Corp.

The remaining 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.

Publisher's note

References

Adamov, N. S. de M., Nascimento, N. F. do, Maciel, E. C. S., Pereira-Santos, M., Senhorini, J. A., Calado, L. L., et al. (2017). Triploid induction in the yellowtail tetra, *Astyanax altiparanae*, using temperature shock: Tools for conservation and aquaculture. *J. World Aquac. Soc.* 48, 741–750. doi:10.1111/jwas.12390

Almeida, F. L., Lopes, J. S., Crescencio, R., Izel, A. C. U., Chagas, E. C., and Boijink, C. (2016). Early puberty of farmed tambaqui (*Colossoma macropomum*): Possible influence of male sexual maturation on harvest weight. *Aquaculture* 452, 224–232. doi:10.1016/j.aquaculture.2015.10.031

Alves, A. C., Yasui, G. S., Nascimento, N. F., Monzani, P. S., Senhorini, J. A., and Pereira-Santos, M. (2022). *All-triploid offspring in the yellowtail tetra* Astyanax altiparanae *Garutti e Britski 2000 (Teleostei, Characideae) derived from female tetraploid x male diploid crosses*. Zygote: Press.

Andreassen, R., and Høyheim, B. (2017). miRNAs associated with immune response in teleost fish. *Dev. Comp. Immunol.* 75, 77–85. doi:10.1016/j.dci.2017.02.023

Arai, K. (2001). Genetic improvement of aquaculture finfish species by chromosome manipulation techniques in Japan. *Aquaculture* 197, 205–228. doi:10.1016/S0044-8486(01)00588-9

Arashiro, D. R., Yasui, G. S., Calado, L. L., do Nascimento, N. F., dos Santos, M. P., do Santos, S. C. A., et al. (2018). Synchronizing developmental stages in Neotropical catfishes for application in germ cell transplantation. *Zygote* 26, 135–148. doi:10.1017/S0967199418000035

Arm, K. (2003). Genetics of the loach, *Misgurnus anguillicaudatus*: Recent progress and perspective. *Folia Biol.* 51, 107–117.

Barzotto, E., and Mateus, L. (2017). Reproductive biology of the migratory freshwater fish *Salminus brasiliensis* (Cuvier, 1816) in the Cuiabá River basin, Brazil. *J. Appl. Ichthyol.* 33, 415–422. doi:10.1111/jai.13262

Bertolini, R. M., Lopez, L. S., do Nascimento, N. F., Arashiro, D. R., de Siqueira Silva, D. H., dos Santos, S. C. A., et al. (2020). Strategies for aquaculture and conservation of Neotropical catfishes based on the production of triploid *Pimelodus maculatus*. *Aquac. Int.* 28, 127–137. doi:10.1007/s10499-019-00449-4

Bertolini, R. M., Senhorini, J. A., Nascimento, N. F., Pereira-Santos, M., Nakaghi, L. S. O., Peres, W. A. M., et al. (2018). First feeding of diploid and triploid yellowtail tetra *Astyanax altiparanae*: An initial stage for application in laboratory studies. *Aquac. Res.* 49, 68–74. doi:10.1111/are.13433

Bizuayehu, T. T., and Babiak, I. (2014). MicroRNA in teleost fish. *Genome Biol. Evol.* 6, 1911–1937. doi:10.1093/gbe/evu151

Brambila-Souza, G., Mylonas, C. C., de Mello, P. H., Kuradomi, R. Y., Batlouni, S. R., Tolussi, C. E., et al. (2021). GnRHa implants and temperature modulate Lambari-do-rabo-amarelo, *Astyanax altiparanae* (Characiformes: Characidae) induced reproduction out of the reproductive season. *Aquac. Res.* 52, 5595–5605. doi:10.1111/are.15435

Bystriansky, J. S., Clarke, W. C., Alonge, M. M., Judd, S. M., Schulte, P. M., and Devlin, R. H. (2017). Salinity acclimation and advanced parr–smolt transformation in growth-hormone transgenic coho salmon (*Oncorhynchus kisutch*). *Can. J. Zool.* 95, 633–643. doi:10.1139/CJZ-2016-0201

Carradice, D., and Lieschke, G. J. (2008). Zebrafish in hematology: Sushi or science? *Blood* 111, 3331–3342. doi:10.1182/blood-2007-10-052761

Chaudhary, D. K., Singh, S. K., Gohil, N., and Bhattacharjee, G. (2020). "Recent progress of CRISPR-Cas9 in zebra fish"," in *Genome engineering via CRISPR-cas9 system* (Massachusetts, United States: Academic Press), 251–261. doi:10.1016/b978-0-12-818140-9.00019-2

Chen, T. T., and Chen, M. J. (2020). "Transgenic fish technology," in *Encyclopedia of marine biotechnology* (New Jersey, United States: Wiley), 3245–3282. doi:10.1002/9781119143802.ch145

Clark, T. S., Pandolfo, L. M., Marshall, C. M., Mitra, A. K., and Schech, J. M. (2018). Body condition scoring for adult zebrafish (*Danio rerio*). *J. Am. Assoc. Lab. Anim. Sci.* 57, 698–702. doi:10.30802/AALAS-JAALAS-18-000045

Coelho, G. C. Z., Arashiro, D. R., Disselli, T., Pereira-Santos, M., Mira-López, T. M., Monzani, P. S., et al. (2021). Developmental stages, incubation temperature, and *in vivo* traceability of primordial germ cell in an important aquaculture species *Piaractus mesopotamicus*. *Aquaculture* 535, 736381. doi:10.1016/j.aquaculture.2021.736381

Coelho, G. C. Z. Z., Yo, I. S., Mira-López, T. M., Monzani, P. S., Arashiro, D. R., Fujimoto, T., et al. (2019). Preparation of a fish embryo for micromanipulation: Staging of development, removal of the chorion and traceability of PGCs in *Prochilodus lineatus*. *Int. J. Dev. Biol.* 63, 57–65. doi:10.1387/ijdb.180348gc

Cortemeglia, C., and Beitinger, T. L. (2005). Temperature tolerances of wild-type and red transgenic zebra danios. *Trans. Am. Fish. Soc.* 134, 1431–1437. doi:10.1577/t04-197.1

Cresko, W. A., McGuigan, K. L., Phillips, P. C., and Postlethwait, J. H. (2007). Studies of threespine stickleback developmental evolution: Progress and promise. *Genetica* 129, 105–126. doi:10.1007/s10709-006-0036-z

da Costa, R. M. R., and Mateus, L. A. de F. (2009). Reproductive biology of pacu *Piaractus mesopotamicus* (holmberg, 1887) (teleostei: Characidae) in the cuiabá river basin, mato grosso, Brazil. *Neotrop. Ichthyol.* 7, 447–458. doi:10.1590/s1679-62252009000300012

Da Silva, R. C., Pereira Dos Santos, M., Senhorini, J. A., Paes, M. D. C. F., Valentin, F. N., Fujimoto, T., et al. (2017). The effect of temperature on the initial development of *Brycon amazonicus* Spix & Agassiz, 1829 as tool for micromanipulation of embryos. *Zygote* 25, 637–651. doi:10.1017/S096719941700051X

Dai, Y.-J., Jia, Y.-F., Chen, N., Bian, W.-P., Li, Q.-K., Ma, Y.-B., et al. (2014). Zebrafish as a model system to study toxicology. *Environ. Toxicol. Chem.* 33, 11–17. doi:10.1002/etc.2406

Danner, G. R. (2008). Salmonid embryo development and pathology. *Am. Fish. Soc. Symp.* 65, 37–58.

Du, Z., Nie, P., and Liu, J. (2021). Genetic improvement for aquaculture species: A promising approach for aquaculture challenges and development. *Rev. Aquac.* 13, 1756–1757. doi:10.1111/raq.12600

Dunham, R. A. (2004). *Aquaculture and fisheries biotechnology: Genetic approaches*. Oxford: GABI Publishing.

Dunham, R. A. (2009). Transgenic fish resistant to infectious diseases, their risk and prevention of escape into the environment and future candidate genes for disease transgene manipulation. *Comp. Immunol. Microbiol. Infect. Dis.* 32, 139–161. doi:10.1016/j.cimid.2007.11.006

El-Zaeem, Y., Khouriaba, H., El-Sherif, M., and Shahin, M. (2014). Production of salinity tolerant nile Tilapia, *Oreochromis niloticus* through introducing foreign DNA into fish gonads. *J. Animal Poult. Fish Prod.* 2, 1–10. doi:10.21608/japfp.2014.7430

Feitsma, H., and Cuppen, E. (2008). Zebrafish as a cancer model. *Mol. Cancer Res.* 6, 685–694. doi:10.1158/1541-7786.mcr-07-2167

Fernandes, K., Gomes, A., Calado, L., Yasui, G., Assis, D., Henry, T., et al. (2019). Toxicity of cyanopeptides from two microcystis strains on larval development of *Astyanax altiparanae*. *Toxins* 11, 220. doi:10.3390/toxins11040220

Fujimoto, T., Kataoka, T., Sakao, S., Saito, T., Yamaha, E., and Arai, K. (2006). Developmental stages and germ cell lineage of the loach (*Misgurnus anguillicaudatus*). *Zool. Sci.* 23, 977–989. doi:10.2108/zsj.23.977

Gao, L., Duan, M., Cheng, F., and Xie, S. (2014). Ontogenetic development in the morphology and behavior of loach (*Misgurnus anguillicaudatus*) during early life stages. *Chin. J. Ocean. Limnol.* 32, 973–981. doi:10.1007/s00343-014-3302-4

Gao, Z., Wang, W., Abbas, K., Zhou, X., Yang, Y., Diana, J. S., et al. (2007). Haematological characterization of loach *Misgurnus anguillicaudatus*: Comparison among diploid, triploid and tetraploid specimens. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 147, 1001–1008. doi:10.1016/j.cbpa.2007.03.006

Garcia, S., Júnior, H. A., Yasui, G. S., Liebl, F., Souto, L. I. M., and Zaniboni-Filho, E. (2017). Tetraploidia em *Rhamdia quelen* (Quoy E Gaimard, 1824) por choque térmico duplo (quente e frio). *Bol. Inst. Pesca* 43, 257–265. doi:10.20950/1678-2305.2017v43n2p257

García, S., Yasui, G. S., Bernardes-Júnior, J. J., Corrêa da Silva, B., Amaral-Júnior, H., and Zaniboni-Filho, E. (2017). Induction of triploidy in *Rhamdia quelen* (Siluriformes, Heptapteridae) by double-temperature shock. *Lat. Am. J. Aquat. Res.* 45, 209–212. doi:10.3856/vol45-issue1-fulltext-22

Garutti, V., and Britski, H. A. (2000). Descrição de uma espécie nova de *Astyanax* (Teleostei: Characidae) da bacia do alto rio Paraná e considerações sobre as demais espécies do gênero na bacia. *Comun. do Mus. Ciências Tecnol. PUCRS, Série Zool.* 13, 65–88.

Garutti, V. (1989). Contribution to the knowledge of the reproduction of *Astyanax bimaculatus* (ostariophysi, Characidae), in waters of the paraná river basin. *Rev. Bras. Biol.* 49, 489–495.

Garutti, V. (2003). *Piscicultura ecológica*. 1st ed. São Paulo: Editora Unesp.

Gonçalves, B. B., Nascimento, N. F., Santos, M. P., Bertolini, R. M., Yasui, G. S., and Giaquinto, P. C. (2018). Low concentrations of glyphosate-based herbicide cause complete loss of sperm motility of yellowtail tetra fish *Astyanax lacustris*. *J. Fish. Biol.* 92, 1218–1224. doi:10.1111/jfb.13571

Hainfellner, P., Souza, T. G. D., Moreira, R. G., Nakaghi, L. S. O., and Batlouni, S. R. (2012). Gonadal steroids levels and vitellogenesis in the formation of oocytes in *Prochilodus lineatus* (Valenciennes)(Teleostei: Characiformes). *Neotrop. Ichthyol.* 10, 601–612. doi:10.1590/s1679-62252012005000021

Hsu, C.-H., Wen, Z.-H., Lin, C.-S., and Chakraborty, C. (2007). The zebrafish model: Use in studying cellular mechanisms for a spectrum of clinical disease entities. *Curr. Neurovasc. Res.* 4, 111–120. doi:10.2174/156720207780637234

ICMBio (2018). "Livro Vermelho da Fauna Brasileira Ameaçada de Extinção - Peixes," in *Livro Vermelho da Fauna Brasileira Ameaçada de Extinção*. Editor I. C. M. de C. da Biodiversidade (Brasília: Instituto Chico Mendes de Conservação da Biodiversidade).

Itono, M., Morishima, K., Fujimoto, T., Bando, E., Yamaha, E., and Arai, K. (2006). Premeiotic endomitosis produces diploid eggs in the natural clone loach, *Misgurnus anguillicaudatus* (Teleostei: Cobitidae). *J. Exp. Zool. A Comp. Exp. Biol.* 305, 513–523. doi:10.1002/jez.a.283

Kostomarova, A. A. (1991). "The loach *Misgurnus fossilis*," in *Animal species for developmental studies* (Berlin, Germany: Springer), 125–144.

Lázaro, T. M., Rocha, N. R. de A., Monzani, P. S., Nakaghi, L. S. O., Senhorini, J. A., Yasui, G. S., et al. (2021). Sperm quality of spontaneously occurring gynogenetic males in the yellowtail tetra *Astyanax altiparanae*. *J. Appl. Ichthyol.* 37, 909–915. doi:10.1111/jai.14244

Lei, F., and Sinica, B. (2020). Studies on reproduction and growth of loach. *Acta Hydrobiol.* 14, 60–67.

Levy-Pereira, N., Carriero, M. M., Yasui, G. S., Meira, C. M., de Sousa, R. L. M., Maia, A. A. M., et al. (2021). Effects of triploid induction on innate immunity and hematology in Astyanax altiparanae. *Fish. Shellfish Immunol.* 116, 12–18. doi:10.1016/j.fsi.2021.05.004

Lucena, C. A. S. de, and Soares, H. G. (2016). Review of species of the *Astyanax bimaculatus* "caudal peduncle spot" subgroup sensu Garutti & Langeani (Characiformes, Characidae) from the rio La Plata and rio São Francisco drainages and coastal systems of southern Brazil and Uruguay. *Zootaxa* 4072, 101–125. doi:10.11646/zootaxa.4072.1.5

Machado-Evangelista, M., Sussel, F. R., and Romagosa, E. (2019). Environmental manipulation on *Astyanax altiparanae* out-of-season spawning. *Lat. Am. J. Aquat. Res.* 47, 292–302. doi:10.3856/vol47-issue2-fulltext-9

Maclean, N., and Laight, R. J. (2000). Transgenic fish: An evaluation of benefits and risks. *Fish. Fish. (Oxf).* 1, 146–172. doi:10.1046/j.1467-2979.2000.00014.x

Martins, L. F., Hilbig, C. C., Yasui, G. S., Monzani, P. S., Senhorini, J. A., Nakaghi, L. S. O., et al. (2021). Return temperature after heat shock affects the production of tetraploids in the yellowtail tetra *Astyanax altiparanae*. *Zygote* 29, 82–86. doi:10.1017/S096719942000043X

Mastrochirico-Filho, V. A., Hata, M. E., Sato, L. S., Jorge, P. H., Foresti, F., Rodriguez, M. V., et al. (2016). SNP discovery from liver transcriptome in the fish *Piaractus mesopotamicus*. *Conserv. Genet. Resour.* 8, 109–114. doi:10.1007/s12686-016-0521-3

Mehlis, M., and Bakker, T. C. M. (2013). Male reproductive traits of full-sibs of different age classes in three-spined sticklebacks (*Gasterosteus*). *Springerplus* 2, 175–211. doi:10.1186/2193-1801-2-175

Mennigen, J. A. (2016). Micromanaging metabolism - a role for miRNAs in teleost energy metabolism. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 199, 115–125. doi:10.1016/j.cbpb.2015.09.001

Michelin, E. C., Massocco, M. M., Godoy, S. H. S., Baldin, J. C., Yasui, G. S., Lima, C. G., et al. (2017). Carryover of aflatoxins from feed to lambari fish (*Astyanax altiparanae*) tissues. *Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess.* 34, 265–272. doi:10.1080/19440049.2016.1266097

Nam, Y. K., Noh, J. K., Cho, Y. S., Cho, H. J., Cho, K. N., Kim, C. G., et al. (2001). Dramatically accelerated growth and extraordinary gigantism of transgenic mud loach *Misgurnus mizolepis*. *Transgenic Res.* 10, 353–362. doi:10.1023/A:1016696104185

Nascimento, N. F., Bertolini, R. M., Lopez, L. S., Nakaghi, L. S. O., Monzani, P. S., Senhorini, J. A., et al. (2021). Heat-induced triploids in *Brycon amazonicus*: A strategic fish species for aquaculture and conservation. *Zygote* 29, 372–376. doi:10.1017/S0967199421000125

Nascimento, N. F., De Siqueira-Silva, D. H., Pereira-Santos, M., Fujimoto, T., Senhorini, J. A., Nakaghi, L. S. O., et al. (2017a). Stereological analysis of gonads from diploid and triploid fish yellowtail tetra *Astyanax altiparanae* (Garutti & Britski) in laboratory conditions. *Zygote* 25, 537–544. doi:10.1017/S0967199417000399

Nascimento, N. F., Lázaro, T. M., Alcântara, N. R., Senhorini, J. A., Santos, S. C. A., Nakaghi, L. S. O., et al. (2018). *In vivo* storage of oocytes leads to lower survival, increased abnormalities and may affect the ploidy status in the yellowtail tetra *Astyanax altiparanae*. *Zygote* 26, 471–475. doi:10.1017/S0967199418000527

Nascimento, N. F., Monzani, P. S., Pereira-Santos, M., Niedzielski, D., Senhorini, J. A., Silva, L. A., et al. (2020a). The first case of induced gynogenesis in Neotropical fishes using the yellowtail tetra (*Astyanax altiparanae*) as a model organism. *Aquaculture* 514, 734432. doi:10.1016/j.aquaculture.2019.734432

Nascimento, N. F., Pereira-Santos, M., Levy-Pereira, N., Monzani, P. S., Niedzielski, D., Fujimoto, T., et al. (2020b). High percentages of larval tetraploids in the yellowtail tetra *Astyanax altiparanae* induced by heat-shock: The first case in Neotropical characins. *Aquaculture* 520, 734938. doi:10.1016/j.aquaculture.2020.734938

Nascimento, N. F., Pereira-Santos, M., Piva, L. H., Manzini, B., Fujimoto, T., Senhorini, J. A., et al. (2017b). 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

Nasiadka, A., and Clark, M. D. (2012). *Zebrafish* breeding in the laboratory environment. *ILAR J.* 53, 161–168. doi:10.1093/ilar.53.2.161

Nelson, J. S., Grande, T. C., and Wilson, M. V. H. (2016). *Fishes of the world*. New Jersey, United States: John Wiley & Sons.

Niimi, G., and Imada, H. (2008). Observation of non-nucleated erythrocytes in the peripheral blood of medaka, *Oryzias latipes*. *Okajimas Folia Anat. Jpn.* 85, 1–3. doi:10.2535/ofaj.85.1

Olsson, J., Jakubavičiūte, E., Kaljuste, O., Larsson, N., Bergström, U., Casini, M., et al. (2019). The first large-scale assessment of three-spined stickleback (*Gasterosteus aculeatus*) biomass and spatial distribution in the Baltic Sea. *ICES J. Mar. Sci.* 76, 1653–1665. doi:10.1093/icesjms/fsz078

Pardo-Carrasco, S. C., Arias-Castellanos, J. A., Suárez-Mahecha, H., Cruz-Casallas, P. E., Vásquez-Torres, W., Atencio-García, V., et al. (2006). Induction to maturation and ovulation of yamú *Brycon amazonicus* with CPE and mGnRHa. *Rev. Colomb. Ciencias Pecu.* 19, 160–166.

Patimar, R., Najafabadi, M. H., and Souraki, M. G. (2010). Life history features of the nonindigenous three-spined stickleback (*Gasterosteus aculeatus* Linnaeus, 1758) in the Gomishan wetland (southeast Caspian Sea, Iran). *Turkish J. Zoology* 34, 461–470. doi:10.3906/zoo-0903-25

Pereira-Santos, M., Nascimento, N. F., Yasui, G. S., Pereira, N. L., Fujimoto, T., Senhorini, J. A., et al. (2018). Short-term storage of the oocytes affects the ploidy status in the yellowtail tetra *Astyanax altiparanae*. *Zygote* 26, 89–98. doi:10.1017/S0967199417000739

Pereira-Santos, M., Shimoda, E., de Andrade, A. F. C., Silva, L. A., Fujimoto, T., Senhorini, J. A., et al. (2017). Grooves surrounding the micropyle decrease the inseminating dose in fish. *Zygote* 25, 731–739. doi:10.1017/S0967199417000624

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

Pinheiro, J. P. S., de Assis, C. B., Sanches, E. A., and Moreira, R. G. (2020). Aluminum, at an environmental concentration, associated with acidic pH and high water temperature, causes impairment of sperm quality in the freshwater teleost *Astyanax altiparanae* (Teleostei: Characidae). *Environ. Pollut.* 262, 114252. doi:10.1016/j.envpol.2020.114252

Piva, L. H., de Siqueira-Silva, D. H., Goes, C. A. G., Fujimoto, T., Saito, T., Dragone, L. V., et al. (2018). Triploid or hybrid tetra: Which is the ideal sterile host for surrogate technology? *Theriogenology* 108, 239–244. doi:10.1016/j.theriogenology.2017.12.013

Rocha, N. A., Schade, G. M., de Miranda Alves, Á., de Souza Silva, C., Nakirimoto, J. M., Lauri, L. S., et al. (2020). Acute exposure to hyperosmotic conditions reduces sperm activation by urine in the yellowtail tetra *Astyanax altiparanae*, a freshwater teleost fish. *Braz. J. Vet. Res. Anim. Sci.* 57, e166205. doi:10.11606/issn.1678-4456.bjvras.2020.166205

Roza de Abreu, M., Silva, L. M. de J. L. M. de J., Figueiredo-Ariki, D. G., Sato, R. T., Kuradomi, R. Y., Batlouni, S. R., et al. (2021). Reproductive performance of lambari (*Astyanax altiparanae*) in a seminatural system using different protocols. *Aquac. Res.* 52, 471–483. doi:10.1111/are.14905

Santos, M. Dos, Yasui, G. S. G. S. G., Xavier, P. P. L. P. P. L. P., Dos Santos, M. P., Yasui, G. S. G. S. G., Xavier, P. P. L. P. P. L. P., 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, 795–807. doi:10.1017/S0967199416000101

Sato, Y., Sampaio, E. V., Fenerich-Verani, N., and Verani, J. R. (2006). Biologia reprodutiva e reprodução induzida de duas espécies de Characidae (Osteichthyes, Characiformes) da bacia do São Francisco, Minas Gerais, Brasil. *Rev. Bras. Zool.* 23, 267–273. doi:10.1590/S0101-81752006000100021

Silveira, L. F., and Straube, F. C. (2008). Livro Vermelho da Fauna Brasileira Ameaçada de Extinção. Livro vermelho da fauna Bras. *ameaçada extinção* VI, 1420. doi:10.1590/S0102-76382009000500027

Siqueira-Silva, D. H., Bertolini, R. M., Levy-Pereira, N., Nascimento, N. F., Senhorini, J. A., Piva, L. H., et al. (2020). Factors affecting secondary sex characteristics in the yellowtail tetra, *Astyanax altiparanae*. *Fish. Physiol. Biochem.* 47, 737–746. doi:10.1007/s10695-020-00832-6

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

Stahl, B. A., Peuss, R., McDole, B., Kenzior, A., Jaggard, J. B., Gaudenz, K., et al. (2019). Stable transgenesis in *Astyanax mexicanus* using the Tol2 transposase system. *Dev. Dyn.* 248, 679–687. doi:10.1002/dvdy.32

Straume, A. H., Kjærner-Semb, E., Ove Skaftnesmo, K., Güralp, H., Kleppe, L., Wargelius, A., et al. (2020). Indel locations are determined by template polarity in highly efficient *in vivo* CRISPR/Cas9-mediated HDR in *Atlantic salmon*. *Sci. Rep.* 10, 409. doi:10.1038/s41598-019-57295-w

Sun, Y., Zheng, G. D., Nissa, M., Chen, J., and Zou, S. M. (2020). Disruption of mstna and mstnb gene through CRISPR/Cas9 leads to elevated muscle mass in blunt snout bream (*Megalobrama amblycephala*). *Aquaculture* 528, 735597. doi:10.1016/j.aquaculture.2020.735597

Suzuki, R. (1983). Multiple spawning of the cyprinid loach, *Misgurnus anguillicaudatus*. *Aquaculture* 31, 233–243. doi:10.1016/0044-8486(83)90315-0

Takeuchi, Y., Yoshizaki, G., and Takeuchi, T. (2003). Generation of live fry from intraperitoneally transplanted primordial germ cells in rainbow trout. *Biol. Reprod.* 69, 1142–1149. doi:10.1095/biolreprod.103.017624

Tang, T., Hu, Y., Peng, M., Chu, W., Hu, Y., and Zhong, L. (2019). Effects of high-fat diet on growth performance, lipid accumulation and lipid metabolism-related MicroRNA/gene expression in the liver of grass carp (*Ctenopharyngodon idella*). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 234, 34–40. doi:10.1016/j.cbpb.2019.04.006

Veloso-Júnior, V. C., Guimarães-Cruz, R. J., Barros, M. D. M., Barata, R. S. L., and Santos, J. E. (2009). Reproduction of the lambari *Astyanax scabripinnis* (jenyns, 1842)(pisces: Characidae) in a small stream in southeastern Brazil. *J. Appl. Ichthyology* 25, 314–320. doi:10.1111/j.1439-0426.2008.01152.x

Wakamatsu, Y., Pristyazhnyuk, S., Kinoshita, M., Tanaka, M., and Ozato, K. (2001). The see-through medaka: A fish model that is transparent throughout life. *Proc. Natl. Acad. Sci. U. S. A.* 98, 10046–10050. doi:10.1073/pnas.181204298

Weber, A. A., Arantes, F. P., Sato, Y., Rizzo, E., and Bazzoli, N. (2013). Oocyte adhesiveness and embryonic development of *Astyanax bimaculatus* (linnaeus, 1758) (pisces: Characidae). *Zygote* 21, 198–202. doi:10.1017/S096719941200007X

Westerfield, M. (2007). *The zebrafish book: A Guide for the laboratory Use of zebrafish Danio ("Brachydanio rerio")*. Eugene, United States: University of Oregon.

Wirzinger, G., Weltje, L., Gercken, J., and Sordyl, H. (2007). Genotoxic damage in field-collected three-spined sticklebacks (*Gasterosteus aculeatus* L.): A suitable biomonitoring tool? *Mutat. Res.* 628, 19–30. doi:10.1016/j.mrgentox.2006.11.011

Wittbrodt, J., Shima, A., and Schartl, M. (2002). Medaka—A model organism from the far east. *Nat. Rev. Genet.* 3, 53–64. doi:10.1038/nrg704

Xavier, P. L. P., Senhorini, J. A., Pereira-Santos, M., Fujimoto, T., Shimoda, E., Silva, L. A., et al. (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

Yamaha, E., Goto-Kazeto, R., Saito, T., Kawakami, Y., Fujimoto, T., Adachi, S., et al. (2010). Primordial germ cell in teleost fish with special references to its specification and migration. *J. Appl. Ichthyol.* 26, 816–822. doi:10.1111/j.1439-0426.2010.01548.x

Yamaha, E., Kazama-Wakabayashi, M., Otani, S., Fujimoto, T., and Arai, K. (2001). Germ-line chimera by lower-part blastoderm transplantation between diploid goldfish and triploid crucian carp. *Genetica* 111, 227–236. doi:10.1023/a:1013780423986

Yamaha, E., Saito, T., Goto-Kazeto, R., 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

Yasui, G. S., Bertolini, R. M., Suárez-López, L., Xavier, P. P., Monzani, P. S., do Nascimento, N. F., et al. (2020a). Flow cytometric analysis from fish samples stored at low, ultra-low and cryogenic temperatures. *Cryobiology* 95, 68–71. doi:10.1016/j.cryobiol.2020.06.004

Yasui, G. S., Fujimoto, T., and Arai, K. (2010). Restoration of the loach, *Misgurnus anguillicaudatus*, from cryopreserved diploid sperm and induced androgenesis. *Aquaculture* 308, S140–S144. doi:10.1016/j.aquaculture.2010.05.041

Yasui, G. S., Fujimoto, T., Sakao, S., Yamaha, E., and Arai, K. (2011). Production of loach (*Misgurnus anguillicaudatus*) germ-line chimera using transplantation of primordial germ cells isolated from cryopreserved blastomeres. *J. Anim. Sci.* 89, 2380–2388. doi:10.2527/jas.2010-3633

Yasui, G. S., Nakaghi, L. S. O., Monzani, P. S., Nascimento, N. F., Pereira dos Santos, M., Goes, C. A. G., et al. (2020b). Triploidization in the streaked prochilod *Prochilodus lineatus* inferred by flow cytometry, blood smears and karyological approaches. *J. Appl. Ichthyol.* 36, 336–341. doi:10.1111/jai.14025

Yasui, G. S., Porto-Foresti, F., Castilho-Almeida, R. B., Senhorini, J. A., and Foresti, F. (2020c). Biologia e criação do lambari-do-rabo-amarelo (*Astyanax altiparanae*). *Espécies Nativ. Piscic. no Bras.* 3, 101–116.

Yasui, G. S., Saito, T., Zhao, Y., Fujimoto, T., Yamaha, E., and Arai, K. (2018). Intra-ooplasmic injection of a multiple number of sperm to induce androgenesis and polyploidy in the dojo loach *Misgurnus anguillicaudatus* (Teleostei: Cobitidae). *Zygote* 26, 408–416. doi:10.1017/S0967199418000448

Yasui, G. S., Senhorini, J. A., Shimoda, E., Pereira-Santos, M., Nakaghi, L. S. O., Fujimoto, T., et al. (2015). Improvement of gamete quality and its short-term storage: An approach for biotechnology in laboratory fish. *Animal* 9, 464–470. doi:10.1017/S1751731114002511

Yu, F., Xiao, J., Liang, X., Liu, S., Zhou, G., Luo, K., et al. (2011). Rapid growth and sterility of growth hormone gene transgenic triploid carp. *Chin. Sci. Bull.* 56, 1679–1684. doi:10.1007/s11434-011-4446-7

Zayed, Y., Qi, X., and Peng, C. (2019). Identification of novel MicroRNAs and characterization of MicroRNA expression profiles in *zebrafish* ovarian follicular cells. *Front. Endocrinol. (Lausanne)* 10, 518. doi:10.3389/fendo.2019.00518

Zhang, X., and Onozato, H. (2004). Hydrostatic pressure treatment during the first mitosis does not suppress the first cleavage but the second one. *Aquaculture* 240, 101–113. doi:10.1016/j.aquaculture.2004.07.004

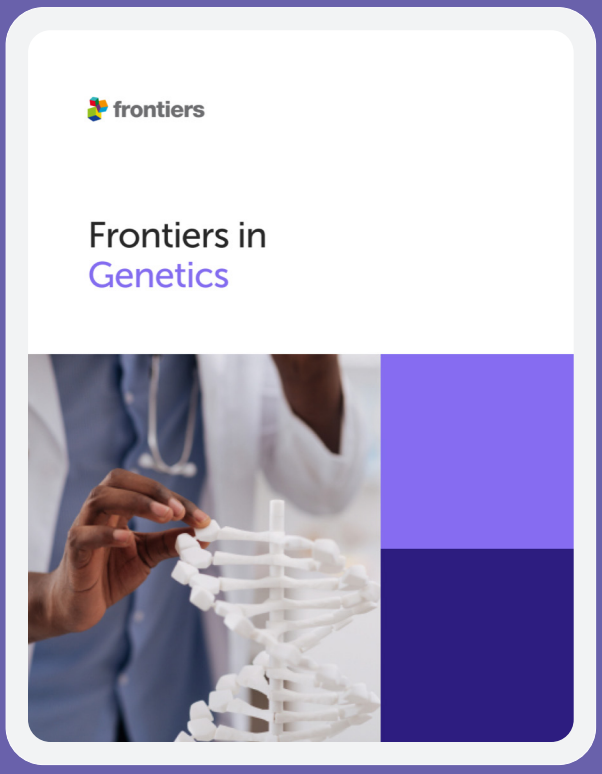
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact