

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE  
DEPARTAMENTO DE CIÊNCIAS FISIOLÓGICAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA  
MOLECULAR

LUZ ELENA DE LA OSSA GUERRA

**PERSPECTIVAS SOBRE A BIOLOGIA EVOLUTIVA DE *Psalidodon  
fasciatus* (TELEOSTEI: CHARACIFORMES), COM FOCO NO PARQUE  
ESTADUAL DE VILA VELHA, PARANÁ, BRASIL**

SÃO CARLOS-SP

2025

LUZ ELENA DE LA OSSA GUERRA

**PERSPECTIVAS SOBRE A BIOLOGIA EVOLUTIVA DE *Psalidodon fasciatus*  
(TELEOSTEI: CHARACIFORMES), COM FOCO NO PARQUE ESTADUAL DE  
VILA VELHA, PARANÁ, BRASIL**

Tese apresentada ao Programa de Pós  
Graduação em Genética Evolutiva e Biologia  
Molecular da Universidade Federal de São  
Carlos, para obtenção do título de Doutora em  
Ciências.

Orientador: Prof. Dr. Roberto Ferreira Artoni

SÃO CARLOS-SP

2025



**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

Centro de Ciências Biológicas e da Saúde  
Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular

---

**Folha de Aprovação**

---

Defesa de Tese de Doutorado da candidata Luz Elena de La Ossa Guerra, realizada em 25/07/2025.

**Comissão Julgadora:**

Prof. Dr. Roberto Ferreira Artoni (UEPG)

Prof. Dr. Marcelo de Bello Cioffi (UFSCar)

Prof. Dr. Pedro Manoel Galetti Junior (UFSCar)

Prof. Dr. Igor de Paiva Affonso (UTFPR)

Prof. Dr. Waldir Miron Berbel Filho (UWF)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular.

**DEDICATORIA**

A mi família.

## AGRADECIMENTOS

Agradeço a todos que, direta ou indiretamente, contribuíram para a realização deste projeto de doutorado.

Ao meu orientador, Prof. Dr. Roberto F. Artoni, por seu apoio constante ao longo de toda a jornada, por me oferecer sempre as melhores condições de trabalho e por incentivar meu crescimento pessoal e profissional.

Às agências de fomento, especialmente à CAPES e ao CNPq, pelo financiamento por meio de bolsas de estudo e de estágio no exterior.

Ao Dr. André Marques e a todos os colegas do grupo *Evolution of Meiosis* do *Max-Planck-Institut für Pflanzenzüchtungsforschung*, pela acolhida durante o estágio sanduíche e pela valiosa colaboração com este projeto. Da mesma forma, agradeço ao Dr. Thomas Liehr e sua equipe na *Universitätsklinikum Jena*, onde realizei meu segundo período de estágio no exterior. Neste contexto, deixo um agradecimento especial ao Prof. Dr. Marcelo Cioffi, cujo apoio foi fundamental para que essa oportunidade se concretizasse.

À equipe do Laboratório de Genética Aquática, em especial ao Dr. Diogo T. Hashimoto, por disponibilizar sua infraestrutura para a construção das bibliotecas ddRAD, e à Carolina Borges, pelo acompanhamento técnico e científico neste processo.

Aos colegas e amigos do LabGEv, que caminharam ao meu lado durante esta jornada: Fran, Tiago, Hallana, Jonathan, Jesus, Iasmin, Gabriel, Miguel a prof. Mara e Mateus.

À minha família, base essencial em todos os momentos: minha mãe Luz Mercedes e meu pai José Pedro, por seu amor incondicional; e meus irmãos José, Erika, Lia e Juan, que ocupam um lugar eterno em meu coração.

E, por fim, mas jamais em último lugar, ao meu companheiro de vida, Roger, por seu apoio constante, presença acolhedora e por ser meu porto seguro nos momentos mais desafiadores.

A todos que, de alguma forma, contribuíram para a realização desta tese, meu mais profundo e sincero agradecimento.

## RESUMO

Ambientes cársticos isolados, como furnas, oferecem condições naturais ideais para investigar os efeitos do isolamento ecológico sobre a biodiversidade e os padrões evolutivos de espécies dulcícolas. Nesta tese, investigamos a ictiofauna das formações do Parque Estadual de Vila Velha (PEVV), sul do Brasil, com foco especial na espécie *Psalidodon fasciatus*, integrando dados de DNA ambiental, genômica e citogenética. Por meio do *metabarcoding* de eDNA, identificamos uma baixa diversidade taxonômica nas furnas em comparação com habitats lóticos da Bacia do Rio Tibagi, evidenciando o isolamento desses sistemas. Análises genômicas com SNPs e genomas mitocondriais revelaram uma estrutura genética marcante na população isolada da furna, refletindo ausência de fluxo gênico e divergência evolutiva iniciada no Pleistoceno. A investigação citogenética, por sua vez, demonstrou variações significativas na distribuição de elementos repetitivos como retrotransposons entre populações isoladas e conectadas. Esses resultados destacam o papel dos sistemas cársticos na formação de linhagens geneticamente distintas, contribuindo para o entendimento dos mecanismos de diversificação em ambientes subterrâneos neotropicais.

**Palavras-chave:** Ecossistema cárstico; peixes; DNA ambiental; genética populacional; citogenética; isolamento geográfico; biodiversidade neotropical.

## ABSTRACT

Isolated karstic environments, such as sinkholes, provide ideal natural settings to investigate the ecological and evolutionary consequences of isolation on freshwater fish biodiversity. In this thesis, we examined the ichthyofauna of these formations in Vila Velha State Park (VVSP), southern Brazil, with a particular focus on *Psalidodon fasciatus*, using an integrated approach that combines environmental DNA, genomics, and cytogenetics. eDNA metabarcoding revealed low taxonomic diversity within sinkholes compared to riverine habitats of the Tibagi River Basin, underscoring the ecological isolation of these systems. Genomic analyses based on SNPs and mitochondrial genomes uncovered strong genetic structure and long-term divergence in the sinkhole population, with no evidence of gene flow and divergence beginning in the Pleistocene. Cytogenetic mapping of repetitive elements and satellite DNAs further revealed distinct chromosomal patterns between isolated and connected populations. Collectively, these findings highlight the role of karst systems in shaping genetically distinct lineages and contribute to our understanding of diversification mechanisms in Neotropical subterranean environments.

**Keywords:** karst ecosystems; fish; environmental DNA; population genetics; cytogenetics; geographic isolation; neotropical biodiversity.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Localização do PEVV.....	19
Figura 2- Locais de coleta de água para eDNA.....	25
Figura 3 - Resultados da análise de metabarcoding de eDNA para detecção de peixes nos diferentes sítios amostrados.....	27
Figura 4 - Diversidade genética e distribuição de SNPs entre as populações .....	40
Figura 5- Estrutura genética de <i>Psalidodon fasciatus</i> entre as populações.....	41
Figura 6 - História demográfica inferida das populações de <i>P. fasciatus</i> .....	43
Figura 7 - Filogenia calibrada no tempo de <i>Psalidodon</i> e táxons relacionados, baseada em genomas mitocondriais (13 genes; 11.608 pb) .....	44
Figura 8 - Análise comparativa dos cariótipos analisados .....	53
Figura 9 - Localização cromossômica do rDNA 5s e 18s em quatro populações de <i>Psalidodon fasciatus</i> .....	54
Figura 10 - Distribuição cromossômica dos retroelementos L1 (83 e 108) e L2 (90) em quatro populações de <i>Psalidodon fasciatus</i> .....	55
Figura 11 - Distribuição cromossômica dos retroelementos <i>Rex-Babar</i> em quatro populações de <i>Psalidodon fasciatus</i> .....	56
Figura 12 – Ideograma cromossômico das diferentes populações de <i>P. fasciatus</i> estudadas ..	57
Figura 13 - Mapeamento dos retroelementos Rex.....	58
Figura 14 - Distribuição cromossômica das DNAs satélites a partir da análise comparativa de retroelementos (RE).....	60
Figura 15 - Distribuição cromossômica dos DNAs satélites <i>as51</i> (verde) e <i>sat51</i> (vermelho) em quatro populações de <i>P. fasciatus</i> .....	61

## LISTA DE TABELAS

Tabela 1 - Composição por taxa das sequências únicas identificadas.....	27
Tabela 2 - Identificação taxonômica dos peixes detectados a partir de amostras de eDNA em seis sítios aquáticos utilizando metabarcoding. A tabela apresenta, para cada sítio, os táxons identificados, suas respectivas abundâncias relativas de leituras (RRA, %), o número de variantes únicas de sequência amplificada (ASVs) e o número de Unidades Taxonômicas Operacionais (OTUs). As MOTUs foram atribuídas taxonomicamente ao menor nível possível (por exemplo, espécie, gênero ou família) .....	28
Tabela 3 - Métricas de diversidade a nível populacional de <i>Psalidodon fasciatus</i> . Número de amostras (N), alelos privados ( $P_A$ ), sítios polimórficos ( $Pol_{sites}$ ), porcentagem de loci polimórficos ( $\%Pol_{loci}$ ), heterozigosidade observada ( $Het_o$ ), heterozigosidade esperada ( $Het_e$ ), e porcentagem of genome-wide SNPs no hardy Weinberg disequilibrium.....	39
Tabela 4 - Análise da variância molecular (AMOVA). Df: graus de liberdade; SSD: soma dos desvios quadráticos; MSD: média dos desvios quadráticos; %: porcentagem da variância total atribuída a cada componente. Valores significativos (*). .....	42
Tabela 5 - Primers utilizados para a amplificação dos fragmentos dos elementos transponíveis presentes na biblioteca de <i>P. fasciatus</i> da furna 2 .....	50

## LISTA DE SIGLAS

Ag-NORs	<i>nucleolar organizer regions</i>
AL	população de <i>P. fasciatus</i> de Alagados, Ponta Grossa
AMOVA	análise de variância molecular (do inglês: <i>analysis of molecular variance</i> )
ASV	<i>amplicon sequence variant</i>
ATL	buffer de lise para purificação de ácidos nucleicos
BSA	albumina de soro bovino (do inglês <i>bovine serum albumin</i> )
DAPI	4',6-diamidino-2-phenylindole
ddRADseq	<i>Double digest restriction-site associated DNA sequencing</i>
DL	população de <i>P. fasciatus</i> da lagoa Dourada (do inglês: <i>Dourada lake</i> )
DMEM	<i>Dulbecco's Modified Eagle Medium</i>
DNA	ácido desoxirribonucleico (do inglês <i>desoxirribonucleic acid</i> )
DNBSeq	<i>DNA Nanoballs sequencing</i>
dsDNA	<i>double strand DNA</i>
eDNA	DNA ambiental (do inglês: <i>environmental DNA</i> )
FBS	soro fetal bovino (do inglês: <i>fetal bovine serum</i> )
FISH	<i>fluoresce in situ hybridization</i>
NF	número fundamental
GLGM	ultimo máximo glacial global (do inglês: <i>global last glacial maximum</i> )
HCl	ácido clorídrico (do inglês: <i>hydrochloric acid</i> )
HPD	<i>height posterior density</i>
IA	população de <i>P. fasciatus</i> pertencente a Iapó
IC	intervalo de confiança
IBS	<i>identity by state</i>
KCl	cloreto de potássio
LINE	<i>Long interspersed nuclear element</i>
LTR	retrotransposons com repetições terminais longas (do inglês: <i>long terminal retrotransposon</i> )
m.s.n.m	metros sob o nível do mar

MAF	<i>minor allele frequency</i>
MAFFT	<i>multiple alignment using fast fourier transform</i>
MCC	máxima credibilidade de clado
MCMC	<i>markov monte carlo chain</i>
MHNCI	Museo de história natural Campão de Imbuia
MluCI	enzima de restrição que reconhece a sequência <sup>^</sup> AATT_
MOTU	unidade taxonômica molecular (do inglês: <i>molecular taxonomic unit</i> )
NCBI	<i>national center for biotechnology information</i>
PBS	tampão fosfato (do inglês: <i>phosphate buffered saline</i> )
PCR	reação em cadeia da polimerase (do inglês <i>polymerase chain reaction</i> )
PE	<i>paired ended</i>
PEVV	Parque Estadual de Vila Velha
rDNA	DNA ribossomal (do inglês: <i>ribosomal DNA</i> )
RE	<i>repeat explorer</i>
REs	elementos repetitivos (do ingles: <i>repetitive elements</i> )
RRA	abundâncias relativas de leitura (do inglês: <i>relative read abundance</i> )
satDNA	DNA satélite (do inglês: <i>satellite DNA</i> )
SDS	Dodecilsulfato asódico (do inglês: <i>sodium dodecyl sulfate</i> )
SH	population da furna 2 do PEVV (do inglês: <i>sinkhole 2</i> )
SINE	<i>short interspersed nuclear elements</i>
sNMF	<i>sparse nonnegative matrix factorization</i>
SNPs	polimorfismo de um único nucleotídeo (do inglês: <i>single nucleotide polymorphism</i> )
SphI	enzima de restrição que reconhece a sequência G_CATG <sup>^</sup> C.
SSC	<i>saline-sodium citrate</i>
TE	elemento transponível (do inglês: <i>transposable element</i> )
UpSet	<i>visualization of intersecting sets</i>
VCF	<i>variant calling format</i>

## LISTA DE SIMBOLOS

%	porcentagem
°C	graus Celsius
μm	micrômetro (= $10^{-6}$ m)
<i>K</i>	número de clusters
ka	milhes de anos
km	quilômetros
M	molar
Ma	milhões de anos
Mb	megabase
Mg	miligrama
ml	mililitros
Ne	número efetivo de uma população
ng	nanograma
<i>p</i>	probabilidade estatística
pb	pares de base
rpm	rotações por minuto
U	unidades
μL	microlitro (= $10^{-6}$ L)
$\pi$	diversidade nucleotídica

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>16</b>
1.1	MODELOS NATURAIS DE POPULAÇÕES ISOLADAS DE PEIXES	16
1.2	Formação Furnas de Vila Velha e Lagoa Dourada, Paraná	17
1.3	Bacia hidrográfica do rio Guabiroba	18
1.4	Conservação das populações de peixes no PEVV	19
1.5	contribuições metodológicas ao entendimento da dinâmica nas furnas do pevv e na espécie <i>P. fasciatus</i>	20
<b>2</b>	<b>CAPITULO 1</b>	<b>22</b>
2.1	INTRODUÇÃO	22
2.2	OBJETIVO	24
2.3	MATERIAIS E MÉTODOS	24
<b>2.3.1</b>	<b>Area de estudo e coleta de eDNA</b>	<b>24</b>
<b>2.3.2</b>	<b>Extração e sequenciamento de eDNA</b>	<b>25</b>
<b>2.3.3</b>	<b>Processamento de dados brutos</b>	<b>26</b>
2.4.	<i>RESULTADOS</i>	26
2.4	Resolução taxonômica do gene mitocondrial 12S rDNA em ambientes de furnas	28
<b>2.4.1</b>	<b>Resolução taxonômica do gene mitocondrial 12S rDNA no Rio Tibagi</b>	<b>28</b>
2.5	DISCUSSÃO	29
2.6	CONCLUSÃO	31
<b>3</b>	<b>CAPITULO 2</b>	<b>33</b>
3.1	INTRODUÇÃO	33
3.2	OBJETIVO	34
3.3	MATERIAIS E MÉTODOS	35
<b>3.3.1</b>	<b>Declaração de ética</b>	<b>35</b>
<b>3.3.2</b>	<b>Double digest restriction-site associated DNA (ddRADseq)</b>	<b>35</b>
3.3.2.1	Coleta de Espécimes, Amostragem de Tecidos e Extração de DNA	35
3.3.2.2	Preparação das bibliotecas	35
3.3.2.3	SNPs calling	36
<b>3.3.3</b>	<b>Análise de diversidade genética e posição de SNPs entre populações</b>	<b>36</b>
<b>3.3.4</b>	<b>Análise de genética populacional e inferência demográfica</b>	<b>37</b>
<b>3.3.5</b>	<b>Genomas mitocondriais e estimativa de tempo</b>	<b>37</b>
3.4	RESULTADOS	38

3.4.1	<b>Filtragem e processamento de dados brutos</b>	<b>38</b>
3.4.2	<b>Análise de diversidade genética e posição de SNPs entre populações</b>	<b>39</b>
3.4.3	<b>Estrutura populacional</b>	<b>40</b>
3.4.4	<b>Inferência demográfica</b>	<b>42</b>
3.4.5	<b>Estimativa do tempo de divergência</b>	<b>43</b>
3.5	DISCUSSÃO	44
3.6	CONCLUSÃO	46
<b>4</b>	<b>Capítulo 3</b>	<b>47</b>
4.1	INTRODUÇÃO	47
4.2	OBJETIVOS	48
4.3	MATERIAS E MÉTODOS	48
4.3.1	<b>Amostragem e obtenção de cromossomos</b>	<b>48</b>
4.3.2	<b>Citogenética convencional</b>	<b>49</b>
4.3.3	<b>Obtenção das sondas de elementos repetitivos</b>	<b>50</b>
4.3.3.1	Elementos transponíveis	50
4.3.3.2	DNA satélites	51
4.3.4	<b>Hibridação <i>in situ</i> fluorescente (FISH)</b>	<b>51</b>
4.3.5	<b>Imagens e confirmação dos resultados</b>	<b>52</b>
4.4	RESULTADOS	52
4.4.1	<b>Hibridação comparativa</b>	<b>53</b>
4.4.1.1	Mapeamento dos rDNAs 5s e 18s	53
4.4.2	<b>Mapeamento de alguns LINEs (<i>Long interspersed nuclear elements</i>) nas populações de <i>P. fasciatus</i></b>	<b>54</b>
4.4.2.1	L1 e L2	54
4.4.2.2	Rex-Babar	54
4.4.2.3	Elementos Rex	57
4.4.3	<b>Mapeamento de DNAs Satélites (SatDNA)</b>	<b>59</b>
4.4.3.1	Análise comparativa de satDNAs pelo RepeatExplorer (RE)	59
4.5	DISCUSÃO	61
4.6	CONCLUSÃO	63
<b>5</b>	<b>REFERENCIAS</b>	<b>64</b>

**ANEXO A-** Resultado *process radtags* biblioteca 1 (indivíduos da lagoa Dourada), biblioteca 2 (furna 2), biblioteca 3 (rio Tibagi, e rio Paraopebas).

**ANEXO B-** Pipeline Stacks usada para o *SNPs calling e Upset plot*

**ANEXO C-** Sequencias utilizadas na construção do relógio molecular

**ANEXO D-** *Artigo aceito na Environmental DNA*

**ANEXO E-** Elementos transponiveis escolhidos do resultado de Repeat Explorer para *P. fasciatus*

**ANEXO F-** Análises comparativa RE população furnas vc São Francisco

## APRESENTAÇÃO DA TESE

Atualmente, os estudos voltados às dinâmicas evolutivas, com ênfase na genética, ecologia e comportamento de populações de peixes com algum grau de isolamento, têm ganhado destaque. Espécies troglóbias têm chamado a atenção por habitarem ambientes extremamente específicos e, além disso, por poderem ser utilizadas como espécies modelo.

A ictiofauna de furnas ou dolinas, embora ocorra em ambientes com incidência de luz, está presente em sistemas fechados e ainda é pouco estudada em nível mundial. Nesta tese, utilizamos o modelo das furnas do Parque Estadual de Vila Velha (PEVV), localizado na bacia do rio Tibagi, município de Ponta Grossa, Paraná, para investigar como essa condição ecológica influencia a estrutura populacional de *Psalidodon fasciatus*, especialmente na Furna 2.

O conteúdo da presente tese foi organizado em uma introdução e três capítulos, visando aprofundar e facilitar a discussão dos temas abordados: O Capítulo 1 apresenta um inventário das espécies da bacia do rio Tibagi e de algumas furnas do PEVV, com base em análises de DNA ambiental (eDNA). O Capítulo 2 utiliza dados de SNPs e genomas mitocondriais completos para investigar os efeitos genéticos do isolamento e estimar o tempo de divergência entre populações de *P. fasciatus* do PEVV e de sistemas abertos, como os rios Tibagi e Paraopeba. Finalmente, o Capítulo 3 oferece uma abordagem citogenética das populações de *P. fasciatus*, por meio da análise de elementos repetitivos, destacando as principais diferenças entre populações com e sem fluxo gênico.

## 1 INTRODUÇÃO

*Psalidodon fasciatus*, anteriormente denominada *Astyanax fasciatus*, é uma espécie alocada recentemente na família Acestrorhamphidae (Melo *et al.*, 2024; Terán; Benitez; Mirande, 2020) distribuída na bacia dos rios Paraná, Paraguai e Uruguai e na bacia do rio São Francisco. O estudo desta espécie tem sido complexo desde a primeira identificação por Cuvier (1819) e logo revisada por (Eigenmann, 1921) até sua recente classificação. Já Eigenmann descrevia esta espécie como a mais amplamente distribuída dos caracídeos: “*deu e está dando origem a várias formas distintas por isolamento em diferentes rios*”; e citando Steindachner “*Quase todos os sistemas fluviais possuem uma variedade peculiar dessa espécie; de acordo com a idade, o sexo, a estação do ano; de acordo com a abundância ou escassez de alimento; de acordo com o habitat, em riachos frescos ou claros nas montanhas ou em águas mais profundas e estagnadas, os contornos do corpo variam e, em parte, também o número de fileiras horizontais de escamas e dos raios anais*”.

Estas primeiras impressões acompanharam de alguma forma todos os trabalhos que tem sido realizados com esta espécie. Alguns trabalhos tem apontado *P. fasciatus* como um complexo de espécies (Artoni *et al.*, 2006; Pazza *et al.*, 2008); isto é suportado pela variação cromossômica que existe nos locais de ocorrência. Citótipos com  $2n=46$ , 48, 49, e 50 cromossomos tem sido reportado ao longo da sua distribuição (Ferreira-Neto *et al.*, 2012; Goes *et al.*, 2022; Matoso *et al.*, 2013; Pazza *et al.*, 2008; Pazza; Kavalco; Bertollo, 2006; Penteado; Kavalco; Pazza, 2013).

Alguns trabalhos sugerem que a ampla diversidade de *P. fasciatus* está mais relacionada em nível cromossômico do que em nível molecular (Kavalco *et al.*, 2016). Abordagens de citogenética molecular tem sido realizada, especialmente na de caracterização de DNA satélites têm sido realizados em *P. fasciatus*, descrevendo o satélite *As51* como ancestral deste gênero e *Astyanax* (Goes *et al.*, 2022). Apesar da quantidade de estudos existentes, são poucos ou nulos os que integram a genômica populacional como ferramenta nas análises de estrutura e diversidade genética.

### 1.1 MODELOS NATURAIS DE POPULAÇÕES ISOLADAS DE PEIXES

Relatos de populações de peixes isoladas geograficamente por barreiras naturais, sem qualquer conexão com corpos d’água adjacentes, ainda são escassos. Casos de populações com ausência total de fluxo gênico têm sido observados em espécies troglomórficas que habitam cavernas. O exemplo mais emblemático é o do peixe cavernícola *Astyanax mexicanus*,

amplamente estudado por meio de comparações com sua forma ancestral de superfície, que ainda habita rios próximos (Hoke; Schwartz; Soares, 2012; Warren *et al.*, 2021). Outro caso notável é o de *Garra barreimiae*, cujas populações subterrâneas apresentam clara diferenciação genética em relação às populações de superfície (Kirchner *et al.*, 2017). Algumas variações na morfologia relacionadas ao habitat têm sido relatadas para a espécie *Cyprinodon pecosensis*, pois aquelas populações que habitavam furnas profundas mostravam diferenças, com o corpo mais alto e pedúnculo mais raso que as outras populações (Collyer *et al.*, 2015). Diferenças na morfologia externa tem sido encontrada também na população isolada de *P. fasciatus* do Parque Estadual de Vila Velha, neste caso, tratando-se de malformações causadas por possível efeito de endogamia (Kerniske *et al.*, 2021).

O isolamento permanente entre populações é um fenômeno raramente observado. No entanto, o isolamento temporário pode ser constatado em determinados grupos, como na família Rivulidae, cujos representantes habitam poças temporárias (Lanés; Keppeler; Maltchik, 2014). Durante o verão, essas poças secam completamente, promovendo o isolamento entre as populações nelas presentes. Já no inverno, com o aumento do volume de água, pode haver conexão esporádica entre poças vizinhas. Apesar desse padrão ecológico indicar um potencial impacto na dinâmica genética das populações, os estudos sobre suas implicações genéticas ainda são bastante escassos.

## 1.2 FORMAÇÃO FURNAS DE VILA VELHA E LAGOA DOURADA, PARANÁ

As Furnas de Vila Velha, (localizadas no Parque Estadual de Vila Velha (PEVV)) estão situadas geologicamente no segundo planalto paranaense, no Arenito de Vila Velha que está sob a forma de platôs sub-horizontais, com sobreposição de ritmitos e diamictitos (França; Winter; Assine, 1996). O segundo planalto ou planalto de Ponta Grossa (Figura 1B) constitui uma paisagem suavemente ondulada, constituída por sedimentos paleozoicos do devoniano, carbonífero e do permiano (Maack, 1969).

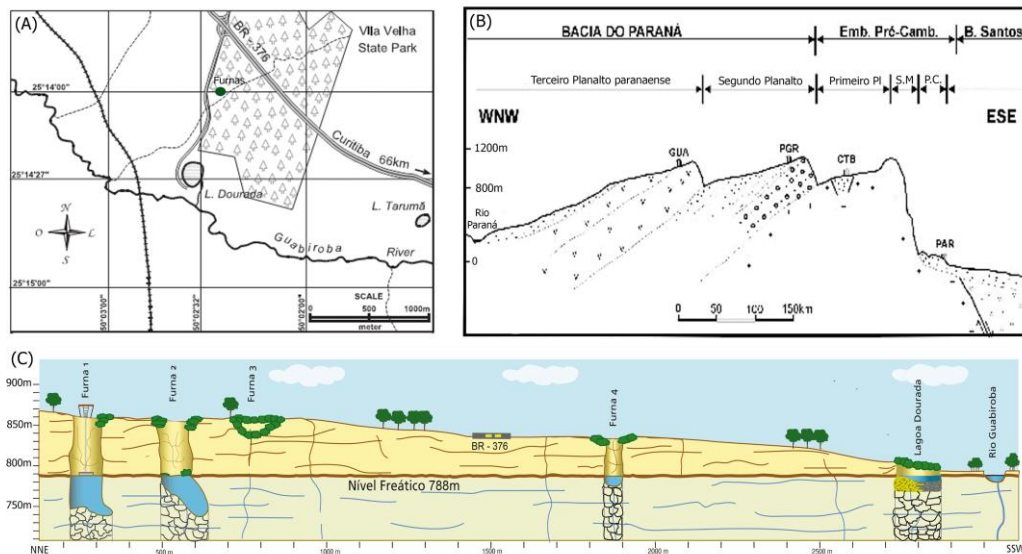
A palavra *Furnas* (que significa “cavernas profundas e íngremes”) são dolinas de abatimento formadas pela existência de cavidades com profundidades superiores a 100m, que podem ser geradas pela erosão dos arenitos sob influência do cruzamento de estruturas tectônicas (falhas e fraturas) ou podem constituir carste subjacente (Maack, 1969; Melo *et al.*, 2011). Brevemente, a história geológica das Furnas de Vila Velha começa desde os processos de formação da bacia do Paraná, desde o Ordoviciano superior até o Cretáceo superior (Assine, 1996). Alguns gráficos estratigráficos do Carbonífero Tardio ao Permiano Superior foram analisados por (Holz *et al.*, 2010), para contribuir ao conhecimento da formação do Arenito de

Vila Velha como resultado do empilhamento de três lobos espalhados em torno de 30 m de espessura cada um e 2 km de largura, depositado provavelmente durante a escavação do vale inciso da Lapa. Por outra parte, processos de intemperismo em Vila Velha responsáveis pelas características pseudo-cársticas começaram há 35Ma, sendo mais intensos entre 17 e 9 Ma (Riffel *et al.*, 2015). Produto destes processos, existem quatro Furnas dentro do PEVV, mas apenas três possuem lagos no interior (Furnas 1, 2 e 4, em sequência, a Furna 3 é seca, Figura 1C). Seguindo esta ordem, encontra-se a Lagoa Dourada, reconhecida como uma depressão formada também no arenito das Furnas com cerca de 200 m de diâmetro e lâmina d'água de até 5,4 m, situada na bacia hidrográfica do Rio Guabiroba, um afluente do alto Rio Tibagi no Estado do Paraná (Figura 1A) (De Melo; Fonseca Giannini; Ruiz Pessenda, 2000).

### 1.3 BACIA HIDROGRÁFICA DO RIO GUABIROBA

A rede de drenagem fluvial do PEVV é formada pelos rios Barrozinho e Quebra Perna que se unem originando o rio Guabiroba (Michels-Souza *et al.*, 2011). A bacia hidrográfica do Rio Guabiroba, à qual pertence a Lagoa Dourada, tem forma aproximadamente triangular, com cerca de 360 km<sup>2</sup> de superfície. O Rio Guabiroba corresponde à base do triângulo. Apresenta direção WNW-ESE, com sentido da drenagem para WNW. Esta direção geral é subparalela a importantes estruturas rúpteis da área, notando-se, nas cabeceiras do Rio Guabiroba, trechos com nítido controle por estas estruturas. Os topos das elevações na bacia do Rio Guabiroba variam de 1.117m.s.n.m. a nordeste até 865 m.s.n.m. a sudoeste, configurando aproximadamente a mesma tendência geral de caimento dos topos para oeste (De Melo; Fonseca Giannini; Ruiz Pessenda, 2000).

Figura 1 - Localização do PEVV



A- Mapa apresentando a localização da lagoa Dourada e sua conexão com o rio Guabiroba (Moro et al., 2004). B - Esquema do Estado do Paraná mostrando a estrutura geológica do relevo (De Melo; Fonseca Giannini; Ruiz Pessenda, 2000). C- Perfil das Furnas do Parque Estadual Vila Velha (Mineropar). Fonte: A autora

#### 1.4 CONSERVAÇÃO DAS POPULAÇÕES DE PEIXES NO PEVV

Os dados genômicos têm tido uma importância crescente na biologia da conservação, principalmente na designação de unidades de conservação, unidades evolutivamente significativas (ESUs), áreas protegidas e espécies ameaçadas (Crandall *et al.*, 2000, Schiebelhut, *et al.*, 2024). Da mesma forma, alguns autores enfatizam no passado evolutivo de populações isoladas como um potencial distinto e adaptado para existir unicamente em tais condições ambientais (Moritz, 1994, McMahon *et al.*, 2014). Devemos considerar que a fisiografia do PEVV, como uma unidade, é uma área importante em termos de eventos históricos evolutivos, e as populações de peixes significam unidades de conservação e subunidades que compõem um antigo evento geológico. Alguns autores (Funk *et al.*, 2012) descrevem a necessidade de incluir abordagens moleculares para entender, definir e proteger áreas naturais. Frankham (2010) também afirma que a falta de manejo genético em algumas populações não se deve à falta de diretrizes científicas, mas devido à falha em não considerar questões genéticas no manejo silvestre. Isto é refletido, por exemplo, na população extintas na Furna 1 do PEVV, o qual significou uma perda da diversidade genética (Artoni; Matiello, 2003) e as características únicas de uma espécie capaz de sobreviver neste ambiente, que se configura *per se* como de especial interesse e de um alto valor histórico.

## 1.5 CONTRIBUIÇÕES METODOLÓGICAS AO ENTENDIMENTO DA DINÂMICA NAS FURNAS DO PEVV E NA ESPÉCIE *P. fasciatus*

A genômica revolucionou a forma como entendemos a evolução, a ecologia e a fisiologia, mas mesmo com os recentes avanços, o estudo do comportamento animal demorou em adotar tecnologias genômicas (Bengston *et al.*, 2018). Algumas dessas tecnologias podem ajudar a responder desde perguntas mais simples até abordagens mais complexas. Ferramentas genômicas incluem sequências conhecidas (sequenciamento completo do exoma, enriquecimento direcionado) ou o direcionamento de sítios de restrição distribuídos aleatoriamente em todo o genoma (sequenciamento de DNA associado ao sítio de restrição (RAD-seq) e genotipagem por sequenciamento (GBS)). Essas abordagens genômicas, embora recentes no estudo do comportamento, já demonstram grande potencial quando aplicadas à compreensão de processos evolutivos em larga escala.

A integração de informações fósseis com filogenias moleculares oferece esforços potenciais junto com métodos de datação molecular para reconstruir tempos de diversificação dos seres vivos (Ho; Duchêne, 2014; Near *et al.*, 2005), e eventos evolutivos históricos (Hipsley; Müller, 2014). Estes dados fornecem algumas informações em torno da dinâmica de manutenção de populações naturais de *P. fasciatus* nas furnas e sua escala de tempo. Usando ferramentas como o relógio molecular como um método probabilístico, é possível por meio da distribuição de *Poisson* indicar padrões de substituição de muitos genes para fornecer *insights* sobre a história de todo o organismo (Bromham; Penny, 2003). As calibrações geralmente são baseadas em evidências fósseis ou geológicas, e sua escolha pode ter um enorme impacto nas estimativas moleculares de escalas de tempo evolutivas (Ho; Duchêne, 2014).

A manutenção da população de *Psalidodon fasciatus* nas furnas do Parque Estadual de Vila Velha (PEVV) permanece como uma questão científica em aberto. Nesse contexto, a aplicação de metodologias de genotipagem por sequenciamento, como o double digest RAD-seq (ddRADseq), representa uma abordagem particularmente apropriada para investigar a diversidade genética e a estrutura dessa população isolada. O ddRADseq permite a geração de milhares de marcadores SNPs distribuídos ao longo do genoma, o que possibilita análises detalhadas de diversidade, endogamia e estrutura populacional, mesmo em organismos não modelo (Peterson *et al.*, 2012). Além disso, os dados gerados podem ser utilizados em estudos de associação genômica ampla (GWAS) ou na identificação de loci de características quantitativas (QTLs) relacionados a comportamentos específicos ou à restrição da plasticidade fenotípica (Bengston *et al.*, 2018). A baixa diversidade genética, por sua vez, pode aumentar o

risco de extinção de populações pequenas, elevando os níveis de endogamia e a probabilidade de depressão endogâmica, o que compromete o potencial adaptativo (Coleman *et al.*, 2018). Portanto, o uso de ddRADseq pode fornecer informações críticas para compreender os mecanismos genéticos subjacentes à persistência das populações de *P. fasciatus* nas furnas do PEVV, contribuindo diretamente para estratégias de conservação.

Os papéis relativos da deriva genética e a seleção divergente constituem questões centrais para orientar esta pesquisa. Neste contexto, é fundamental estabelecer marcadores moleculares associados a fenótipos de interesse e submetê-los a análises de similaridade, agrupamento e padrões de herança. A integração desses dados moleculares com informações morfométricas e marcadores cariotípicos permite elucidar diferentes trajetórias evolutivas, incluindo processos de especiação observados em outros grupos do gênero *Psalidodon* (Castro *et al.*, 2014a, 2014b).

Marcadores citogenéticos derivados de elementos repetitivos representam importantes recursos para a conservação de espécies, possibilitando a detecção de variações genômicas e a definição de prioridades de manejo populacional. As sequências satélites acumulam-se em regiões estruturais vitais dos cromossomos (centrômeros, telômeros e heterocromatina), desempenhando funções essenciais na segregação cromossômica, formação de heterocromatina e estabilidade genômica (Shatskikh *et al.*, 2020). Já os elementos transponíveis (TEs), além de contribuir diretamente para a produção de novos satélites, podem causar rearranjos cromossômicos, como inversões e duplicações, promovendo divergência e isolamento reprodutivo entre populações (Meštrović *et al.*, 2015). A distribuição específica desses elementos em cariótipos permite identificar linhagens cripticamente divergentes, direcionando ações conservacionistas mais precisas e evitando a perda de diversidade genética. Além disso, o aumento da atividade de elementos transponíveis pode fornecer os meios para a diversificação genética em populações naturais, possibilitando o surgimento, a posterior seleção e a fixação de novas variantes adaptativas por meio da seleção natural (Schrader; Schmitz, 2018).

## 2 CAPÍTULO 1

### **Iluminando a escuridão: usando eDNA metabarcoding para estudar a ictiofauna exclusiva da Formação Furnas**

#### 2.1 INTRODUÇÃO

Avaliar e identificar a biodiversidade de peixes de água doce tem sido uma tarefa desafiadora há alguns anos. Novas metodologias têm surgido para começar a preencher essa lacuna. O DNA metabarcoding tem sido utilizado como ferramenta para diversos fins, incluindo a identificação de comunidades de peixes (Tsuji *et al.*, 2022), monitoramento aquático (Yates *et al.*, 2019; Xiong *et al.*, 2022) e identificação de ovos e larvas (Breitbart *et al.*, 2023). O metabarcoding de DNA ambiental (eDNA) também tem sido importante para a descoberta e inventário de espécies ocultas e para a detecção de espécies subestimadas em bacias hidrográficas (Gleason *et al.*, 2023; Sales *et al.*, 2018). Além disso, o eDNA metabarcoding tem se mostrado uma abordagem eficiente para o monitoramento não invasivo de espécies ameaçadas (Beng; Corlett, 2020) e para a detecção precoce de espécies exóticas em reservas naturais (Fernández *et al.*, 2018). Também oferece a vantagem de cobrir áreas de amostragem espacialmente maiores, o que é particularmente crucial para áreas protegidas (Herder *et al.*, 2014).

O monitoramento da diversidade de peixes em pequenos corpos d'água como lagoas usando eDNA metabarcoding não é uma prática comum, pois esses locais são geralmente de fácil acesso e suas características são bem documentadas (Yao *et al.*, 2022). Isso resulta no uso de métodos tradicionais de coleta baseados em captura. No entanto, esse não é o caso de formações geológicas como furnas, especificamente a Formação Furnas, que data do Siluriano-Devoniano (entre 421 e 395 milhões de anos, Milani *et al.*, 2007): a área protegida do Parque Estadual de Vila Velha (PEVV). O gradiente dessas cavernas varia de 15 a 60 metros da superfície até o nível da água. A coluna d'água, representada pelo afloramento do lençol freático, varia de 20 a 45 metros, incrustada na rocha. Embora o lençol freático estabeleça uma conexão entre esses ambientes, esse fluxo ocorre através da porosidade das rochas, sem vias de comunicação efetivas que permitam a passagem de organismos macroscópicos. Na região do parque, existem quatro furnas: três furnas úmidas (1, 2 e 4), onde o lençol freático local forma lagoas na base, e uma fuma seca (3), que possui um fundo mais raso em relação ao lençol freático. Além disso, furnas assoreadas, como as próximas a lagoa Dourada, recebem águas de

enchente carregadas de sedimentos do Rio Quebra-Perna (Melo; Giannini, 2007). Furnas com fundo úmido, como as furnas 1, 2 e 4, podem atingir profundidades de até 60 metros, promovendo microambientes isolados que suportam populações endêmicas (Artoni; Almeida, 2003). Essa configuração única apresenta uma oportunidade intrigante para obter informações sobre a história evolutiva e as dinâmicas das espécies que podem habitar essa área. De fato, estudos anteriores demonstraram a existência de uma única espécie na furna 2 (Kerniske *et al.*, 2021; Kerniske *et al.*, 2023), levantando a questão de se o comportamento dessa espécie corresponde ao de uma única população.

Os processos de formação das furnas na área de estudo permanecem apenas parcialmente compreendidos. No entanto, acredita-se que a dissolução do cimento de caulinita e illita, seguida da arenização do arenito e erosão mecânica, contribua para a gênese complexa dessas formações (de Melo; Giannini, 2007). Até o momento, aproximadamente 30 estruturas geológicas na região foram classificadas como furnas. Essas feições geomorfológicas foram mapeadas e suas características foram sistematicamente descritas em relação à proximidade do sistema de drenagem fluvial. A rede de drenagem no Parque Estadual de Vila Velha (PEVV) consiste principalmente nos rios Barrozinho e Quebra-Perna, que convergem para formar o Rio Guabiroba (Michels-Souza *et al.*, 2011). A bacia do Rio Guabiroba, onde está situada a lagoa Dourada, cobre aproximadamente 360 km<sup>2</sup> e possui formato triangular. O rio em si percorre a base dessa bacia triangular na direção WNW-ESE, com fluxo geralmente em direção ao WNW. Essa direção é subparalela a várias feições estruturais significativas da região. Levantamentos de campo revelaram que trechos do rio em suas nascentes são visivelmente influenciados por essas estruturas geológicas. As medidas de elevação na bacia variaram de 1.117 metros acima do nível do mar (m.s.n.m.) no Nordeste até 865 m.s.n.m. no Sudoeste, refletindo uma tendência de declínio topográfico em direção ao oeste (de Melo *et al.*, 2000).

Estudos morfométricos, citogenéticos e moleculares anteriores documentaram a presença apenas de *Psalidodon fasciatus* (anteriormente designado como *Astyanax aff. fasciatus*) nas furnas 1 e 2 do PEVV (Shibatta; Artoni, 2005; Matoso *et al.*, 2013). No entanto, devido aos desafios inerentes ao processo de coleta na furna 1, os espécimes obtidos nesses estudos apresentaram um tamanho amostral relativamente pequeno. Um caso bem documentado de extinção envolve *Psalidodon fasciatus* na furna 1. Essa extinção, ocorrida no início dos anos 2000, foi atribuída à eutrofização causada por excrementos de aves durante a época reprodutiva (Matoso *et al.*, 2010). Essa espécie representava a única população de peixes

conhecida naquele ambiente. Desde então, apesar do acesso desafiador, *Psalidodon fasciatus* não foi mais observado na furna 1.

Por outro lado, a furna 2 abriga uma única população isolada de *Psalidodon fasciatus*, que tem sido o foco de pesquisas por vários anos (Matoso *et al.*, 2004; Gross *et al.*, 2004; Matoso *et al.*, 2010; Matoso *et al.*, 2013). Até o momento, nenhuma outra espécie de peixe foi registrada nesse ambiente. Considerando que essa formação paisagística existe em sua forma atual há mais de 10 milhões de anos (Riffel *et al.*, 2015), a presença desses peixes é algo intrigante. Nossas descobertas recentes, utilizando genomas mitocondriais completos, indicam que esses peixes estão isolados de seus congêneres da bacia do Rio Tibagi há mais de mil anos (De la Ossa-Guerra *et al.*, em preparação). A furna 4, devido à sua difícil acessibilidade, permanece inexplorada quanto à fauna de peixes. Isso contrasta com a ictiofauna bem documentada do Lago Dourada (Artoni; Shibatta, 2006; Gealh, 2010) e do Rio Tibagi (Shibatta *et al.*, 2002; Shibatta *et al.*, 2007; Garla *et al.*, 2023).

A bacia do Rio Tibagi é uma das mais importantes do estado do Paraná, no sul do Brasil, e a Formação Furnas faz parte dela. Ao longo de seu curso, o rio cruza duas das cinco maiores cidades do estado. Próximo ao canal principal do rio, há importantes empreendimentos industriais. Esses fatores, aliados à alta diversidade espacial e biodiversidade de peixes (mais de 110 espécies), tornam essa região central para a conservação da ictiofauna (Shibatta *et al.*, 2002; Shibatta *et al.*, 2007; Claro-García *et al.*, 2018). A ictiofauna da bacia do Rio Tibagi foi relatada em alguns trabalhos utilizando métodos tradicionais de coleta (Bennemann *et al.*, 1995; Shibatta *et al.*, 2002; Claro-García, 2018; Shibatta *et al.*, 2007), com aumento da diversidade e riqueza da nascente até a foz do rio (Shibatta *et al.*, 2002).

## 2.2 OBJETIVO

O objetivo deste trabalho foi investigar se a população de *P. fasciatus* era a única existente na furna 2, confirmar sua extinção local na furna 1 do Parque Estadual de Vila Velha, bem como detectar a existência dessa e de outras espécies na furna 4.

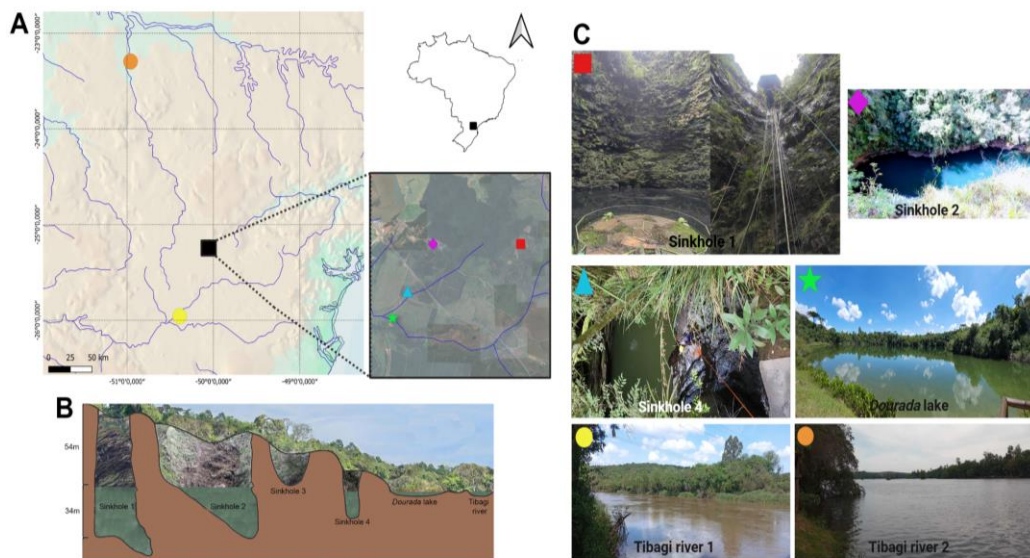
## 2.3 MATERIAIS E MÉTODOS

### 2.3.1 Area de estudo e coleta de eDNA

Seis locais foram amostrados, com dois réplicas em cada uma das quatro furnas do Parque Estadual de Vila Velha (PEVV) e dois no canal principal do Rio Tibagi (Figura 2), utilizando o kit de coleta de eDNA descrito no trabalho de Mendes *et al.* (2024a). As amostras

do Rio Tibagi foram coletadas nos municípios de Ponta Grossa, PR, e Jataizinho, PR (Figura 1). Para cada local, foram utilizados três filtros Sterivex™ de 0,45 µm, com um volume total de água filtrada de 180 ml por filtro e 1080 ml por local. Após a filtração, o líquido remanescente foi expelido por injeção. Em seguida, os filtros foram preenchidos com um buffer ATL e mantidos a 4°C até a extração do eDNA. Antes da coleta em cada local, 180 ml de água mineral foram filtrados no campo de maneira semelhante como amostra controle.

Figura 2- Locais de coleta de água para eDNA



**A** - Localização dos pontos de amostragem na bacia do rio Tibagi, com duas réplicas coletadas em cada local. **B** - Perfis de sumidouros. **C** - As fotografias oferecem uma visão mais de perto as condições ambientais nos locais de amostragem. Fonte: De la Ossa *et al.*, 2025

### 2.3.2 Extração e sequenciamento de eDNA

O DNA ambiental foi extraído usando o kit DNA PowerWater (Qiagen) com modificações (Mendes *et al.*, 2024b). Um fragmento de 200 pb da região mitocondrial 12S foi amplificado utilizando os primers Mifish (Miya *et al.*, 2015), com três amplificações independentes de PCR realizadas por amostra (filtro) para reduzir o viés de amplificação. Esses replicados foram agrupados para o sequenciamento. A reação de PCR utilizou uma mistura de 20 µL contendo 8,34 µL de água, 0,16 µL de BSA (100 µg/mL), 10 µL de AmpliTaq Gold 360 Master Mix (Thermo Fisher), 0,25 µL de cada primer e 1,0 µL de DNA template, com água substituindo o DNA nos controles negativos. Os primers incluíram uma cauda 5' para um segundo PCR, e os produtos de PCR foram purificados utilizando beads magnéticos. Após a purificação, foi realizado um segundo PCR utilizando adaptadores do kit Nextera Index® para marcar e amplificar os amplicons. Adaptadores compatíveis com Illumina (P5 e P7) serviram como primers. O produto da amplificação foi então avaliado em um gel de agarose a 1,5%.

Todas as amostras foram agrupadas, e esse pool foi purificado com o kit ZymoClean™ Large Fragment DNA Recovery (Zymo Research). O pool foi quantificado utilizando PCR em tempo real com o Kit de Quantificação KAPA Biosystems (Illumina). A solução final foi desnaturada e carregada no equipamento NextSeq 2000® (Illumina), utilizando o kit de sequenciamento NextSeq® V2.5 (2x150bp PE).<sup>7</sup>

### 2.3.3 Processamento de dados brutos

As análises bioinformáticas foram realizadas utilizando um pipeline customizado no R 4.3.1 (R Core Team, 2024), com os pacotes DADA2 (Callahan *et al.*, 2016) e Phyloseq (McMurdie; Holmes, 2013). As leituras brutas foram desmultiplexadas pelo Basespace, e as sequências de adaptadores e primers foram removidos utilizando *cutadapt* (Martin, 2011). Leituras de baixa qualidade (abaixo de Q30) ou com bases indeterminadas foram filtradas, e as leituras restantes foram agrupadas em ASVs utilizando o DADA2. As atribuições taxonômicas foram realizadas com BLASTn (Camacho *et al.*, 2009) com limiares mínimos de similaridade e cobertura (80%), utilizando a coleção de nucleotídeos do NCBI (Sayers *et al.*, 2022) como referência. As ASVs foram agrupadas em OTUs utilizando o SWARMv2 (Mahé *et al.*, 2015), e a identificação taxonômica foi atribuída com base nos limiares de similaridade. As abundâncias relativas de leitura (RRA) foram calculadas considerando todas as classes de vertebrados.

Dado que cada local incluía duas réplicas, as leituras foram agrupadas para análises subsequentes, uma vez que a proximidade dos locais de amostragem não justificava tratá-los separadamente. No entanto, essa abordagem não foi aplicada aos locais do Rio Tibagi, onde as maiores distâncias entre os pontos exigiram um tratamento separado.

## 2.4. RESULTADOS

Nosso conjunto final de dados compreendeu um total de 1.990.069 leituras brutas, com 1.497.479 sequências pareadas atribuídas a 90 unidades taxonômicas operacionais moleculares (MOTUs) (Tabela 1).

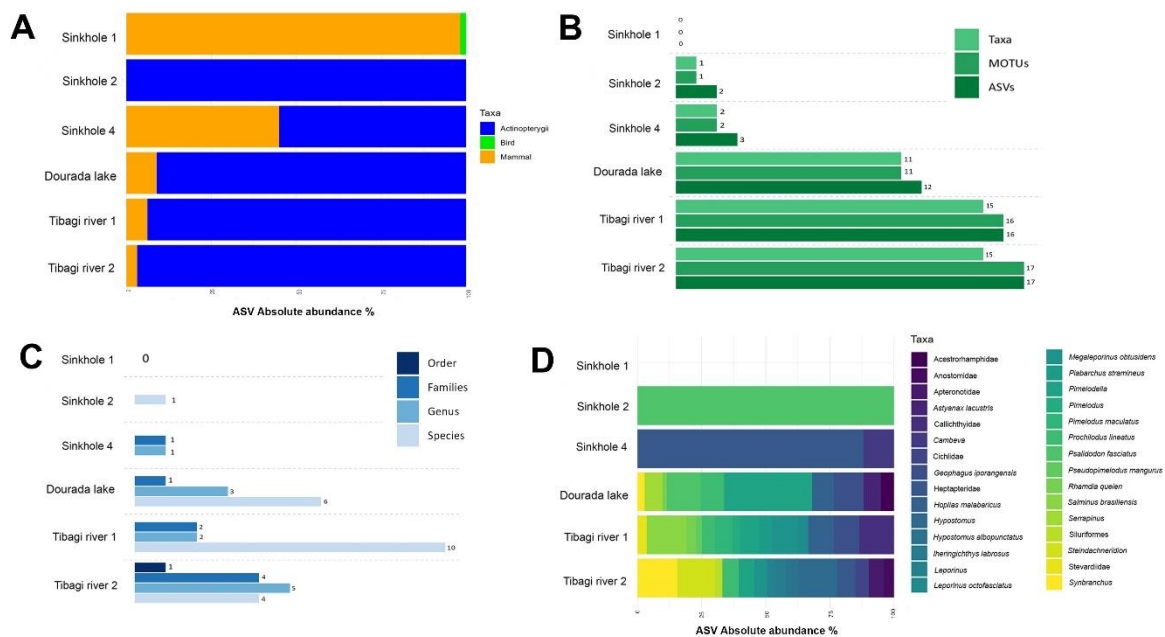
Tabela 1 - Composição por taxa das sequências únicas identificadas

Class	Raw data		Filtering 1%	
	Absolute abundance reads	MOTUs detected	Absolute abundance	MOTUs detected
All	1,470,749	90	1,369,446	37
Teleosts	1,033,854	65	975,197	31
Mammals	431118	21	394,249	5
Birds	5777	4	4844	1

Fonte: De la Ossa *et al.*, 2024

Os táxons da classe Actinopterygii são os mais abundantes na maioria dos locais, particularmente no Rio Tibagi 2, Rio Tibagi 1 e Furna 2, onde chegam a quase 100% (Figura 3A). Essa distribuição sugere que Actinopterygii dominam os sistemas aquáticos nos locais estudados, embora certas furnas, especialmente a Furna 1 e a Furna 4, aparentem sustentar uma gama mais ampla de classes, como aves e mamíferos.

Figura 3 - Resultados da análise de metabarcoding de eDNA para detecção de peixes nos diferentes sítios amostrados.



**A** - Abundância absoluta de ASVs de todas as classes utilizando o gene 12s. **B** - Número de ASVs, MOTUs e táxons atribuídos para todos os sítios. **C** - Número de níveis taxonômicos identificados em cada sítio. **D** - Porcentagem da abundância absoluta de ASVs por táxon em cada sítio. As barras representam os valores para cada categoria. Fonte: De la Ossa *et al.*, 2025

Observamos 49 variantes de sequência (ASVs) com abundâncias relativas de leitura superiores a 1%, considerando apenas as sequências de Actinopterygii, e um conjunto final de dados contendo 975.197 leituras e 31 MOTUs foi utilizado para análises subsequentes. Um total de quatro ordens taxonômicos diferentes foram encontradas: Characiformes, Cichliformes,

Gymnotiformes e Siluriformes (Tabela 2). Um número significativo de MOTUs pôde ser classificado apenas até o nível de ordem ou família (Figura 3D), devido à disponibilidade insuficiente de sequências de referência para a região 12S na bacia do Rio Tibagi.

#### 2.4 RESOLUÇÃO TAXONÔMICA DO GENE MITOCONDRIAL 12S RDNA EM AMBIENTES DE FURNAS

Nenhum ASV pertencente a ordem Actinopterygii foi identificado em qualquer nível taxonômico na furna 1. Na furna 2, *P. fasciatus* apresentou a maior abundância absoluta de ASVs (428.564 reads). Além disso, nenhum outro ASV foi identificado, o que reforça a hipótese de que a população dessa espécie seja a única ocorrendo na furna 2. Para a furna 4, os resultados demonstraram a presença de dois ASVs distintos, com índices de identidade BLASTn de 93% e 97%, respectivamente, correspondendo a Heptapteridae e o gênero *Cambeva* (Figura 3D, Tabela 2).

A amostra da lagoa Dourada resultou em um total de 191.956 leituras, atribuídas a 11 MOTUs distintas, com abundâncias relativas variando de aproximadamente 1% a 38,6%, que puderam ser taxonomicamente atribuídas a 11 táxons. Além disso, apenas duas MOTUs foram identificadas no nível de família (Acestorhamphidae) (Tabela 2).

##### 2.4.1 Resolução taxonômica do gene mitocondrial 12S rDNA no Rio Tibagi

O BLASTn das amostras da bacia do Rio Tibagi revelou a maior diversidade, com 15 e 17 MOTUs identificados nos locais Tibagi 1 e Tibagi 2, respectivamente (Figura 3A). Para alguns MOTUs, a resolução taxonômica mais próxima alcançada foi no nível de família (7 MOTUs). As ordens dominantes identificadas, Characiformes e Siluriformes, estão entre as mais comumente registradas em tributários neotropicais (Claro-García *et al.*, 2018) (Figura 3B). As ASVs mais abundantes no Rio Tibagi foram de *Hypostomus* e *Steindachneridion*.

Tabela 2 - Identificação taxonômica dos peixes detectados a partir de amostras de eDNA em seis sítios aquáticos utilizando metabarcoding. A tabela apresenta, para cada sítio, os táxons identificados, suas respectivas abundâncias relativas de leituras (RRA, %), o número de variantes únicas de sequência amplificada (ASVs) e o número de Unidades Taxonômicas Operacionais (OTUs). As MOTUs foram atribuídas taxonomicamente ao menor nível possível (por exemplo, espécie, gênero ou família)

Sítios	Táxons identificados	RRA (%)	Num ASVs	Num OTU
Sinkhole 1	0	0.00	0	0
Sinkhole 2	<i>Psalidodon fasciatus</i>	98.59	2	1
Sinkhole 4	Heptapteridae	36.67	2	1

Sítios	Táxons identificados	RRA (%)	Num ASVs	Num OTU
	<i>Cambeva</i>	4.99	1	1
<b>Dourada Lake</b>	<i>Pimelodella</i>	28.62	1	1
	<i>Psalidodon fasciatus</i>	10.93	1	1
	<i>Geophagus iporangensis</i>	9.52	1	1
	<i>Hoplias malabaricus</i>	7.10	1	1
	<i>Prochilodus lineatus</i>	7.70	2	1
	<i>Serrapinus</i>	5.84	1	1
	<i>Astyanax lacustris</i>	5.60	1	1
	Acestrorhamphidae	4.34	2	2
	<i>Synbranchus</i>	2.30	1	1
	<i>Rhamdia quelen</i>	1.41	1	1
<b>Tibagi River 1</b>	<i>Salminus brasiliensis</i>	14.16	1	1
	Callichthyidae	12.30	2	2
	<i>Megaleporinus obtusidens</i>	9.37	1	1
	<i>Hoplias malabaricus</i>	9.20	1	1
	<i>Geophagus iporangensis</i>	9.02	2	1
	<i>Pimelodella</i>	6.91	1	1
	<i>Pimelodus maculatus</i>	6.20	1	1
	<i>Prochilodus lineatus</i>	4.63	1	1
	<i>Piabarchus stramineus</i>	4.31	1	1
	<i>Leporinus octofasciatus</i>	3.68	1	1
	<i>Rhamdia quelen</i>	3.45	1	1
	Stevardiidae	3.30	1	1
	<i>Pimelodus</i>	2.60	1	1
	<i>Pseudopimelodus mangurus</i>	2.11	1	1
<b>Tibagi River 2</b>	<i>Steindachneridion</i>	13.92	1	1
	<i>Synbranchus</i>	8.40	1	1
	<i>Hypostomus</i>	14.13	3	3
	Heptapteridae	6.72	1	1
	<i>Prochilodus lineatus</i>	6.02	1	1
	<i>Pimelodella</i>	5.70	1	1
	Apteronotidae	5.51	1	1
	Cichlidae	4.97	1	1
	<i>Megaleporinus obtusidens</i>	4.65	1	1
	<i>Hypostomus albopunctatus</i>	4.64	1	1
	<i>Iheringichthys labrosus</i>	3.99	1	1
	Anostomidae	3.73	1	1
	<i>Leporinus</i>	3.13	1	1
	Siluriformes	2.69	1	1

Fonte: De la Ossa *et al.*, 2025

## 2.5 DISCUSSÃO

Decifrar a distribuição e a ecologia da vida subterrânea é essencial para a conservação das espécies, uma tarefa que apresenta desafios significativos quando abordada por métodos

tradicionais (Boyd *et al.*, 2020). Por exemplo, o *metabarcoding* de eDNA tem sido utilizado como uma ferramenta eficiente para descoberta e monitoramento de formas de vida subterrâneas, como demonstrado na detecção de três espécies de peixes cegos de cavernas do Irã (*Garra typhlops*, *Garra lorestanensis* e *Eidinemacheilus smithi*) por meio de marcadores 12S rRNA MiFish (Hashemzadeh *et al.*, 2022).

No contexto neotropical, ainda não foram aplicados estudos de *eDNA metabarcoding* voltados à avaliação da biodiversidade subterrânea ou de furnas. Neste trabalho, nossos dados apoiam o padrão de distribuição previamente proposto para o peixe *Psalidodon fasciatus*, corroborando a extinção da população na furna 1. Em contraste, confirmamos a presença atual de *P. fasciatus* na furna 2, sendo a única espécie habitante deste local, onde também foi registrada a maior abundância de leituras para essa espécie entre todos os pontos amostrados (Tabela 1). Estudos anteriores já haviam examinado as características únicas dessa população isolada de *P. fasciatus*. Kerniske *et al.* (2022) documentaram malformações morfológicas, provavelmente decorrentes de endocruzamento e competição por nicho, fatores que podem reduzir o desempenho populacional. Em conjunto, esses achados reforçam a importância ecológica da furna 2 como um habitat distinto para essa população isolada.

De forma interessante, relatamos pela primeira vez a detecção da família Heptapteridae e do gênero *Cambeva* na furna 4, ainda inexplorada. Esse registro levanta questões instigantes sobre a biodiversidade local. Apesar do baixo índice de identidade BLASTn (93,73%) para o MOTU atribuído à Heptapteridae, os resultados são considerados confiáveis e reforçam, na verdade, a carência de sequências de referência disponíveis para espécies da região. Devido à dificuldade de acesso a essa furna para fins de inventário biológico, ainda não foi possível realizar esforços de amostragem convencionais. Por exemplo, alguns representantes da família Heptapteridae, como *Rhamdiopsis krugi* (Bockmann; Castro, 2010), já foram documentados em regiões com características semelhantes às formações de furnas do Parque Estadual de Vila Velha (PEVV), como a Chapada Diamantina, no estado da Bahia. Essa área é notável por sua zona freática e paisagem cárstica extensa (mais de 300 km<sup>2</sup>), composta por formações de calcário e quartzito (Bockmann; Castro, 2010). Em relação ao gênero *Cambeva*, três espécies são atualmente conhecidas na bacia do rio Tibagi: *Cambeva castroi* (de Pinna, 1992), *Cambeva davisii* (Haseman, 1911) e *Cambeva diabola* (Bockmann, Casatti, de Pinna, 2004) (Shibatta *et al.*, 2002; Shibatta *et al.*, 2007; Claro-Garcia *et al.*, 2018; Galves *et al.*, 2007; Hoffmann *et al.*, 2015; Shibatta; Cheida, 2003). O registro no PEVV pode corresponder a uma dessas espécies já conhecidas, ou potencialmente representar uma espécie

ainda não descrita — possivelmente troglóbia — considerando que ambientes cársticos como os do PEVV frequentemente abrigam alta diversidade de espécies da subfamília Trichomycterinae, à qual *Cambeva* pertence (Trajano *et al.*, 2009; Rizzato *et al.*, 2012; Rizzato *et al.*, 2014). Essa descoberta abre novos caminhos para pesquisa e pode contribuir para o avanço de iniciativas de conservação na região, mesmo diante da limitação de resolução taxonômica em nível de espécie, além de indicar a possibilidade de espécies ainda não caracterizadas, destacando a necessidade de geração de referências genéticas com base em espécimes locais.

Diversos levantamentos já investigaram a diversidade da bacia do rio Tibagi (Shibatta *et al.*, 2002; Shibatta *et al.*, 2007; Claro-Garcia *et al.*, 2018; Galves *et al.*, 2007; Hoffmann *et al.*, 2015; Shibatta; Cheida, 2003; dos Reis *et al.*, 2020). A maior abundância de MOTUs atribuídos a *Hypostomus* no rio Tibagi pode ser explicada pela elevada riqueza de espécies na bacia, que abriga 13 espécies (Jarduli *et al.*, 2019). No caso de *Steindachneridion*, a ocorrência mais elevada pode ser atribuída ao caráter migratório das espécies desse gênero, que apresentam ampla distribuição nas bacias onde ocorrem (Paixão *et al.*, 2018). É provável que os MOTUs de *Steindachneridion* pertençam à espécie *Steindachneridion scriptum*, a única registrada na bacia (Claro-Garcia *et al.*, 2018). Todas as espécies relatadas neste estudo já haviam sido previamente registradas na bacia.

Esta pesquisa representa o primeiro estudo desse tipo conduzido na bacia do rio Tibagi e demonstra que o gene ribossomal 12S é altamente eficaz para a identificação de espécies em ambientes de dolina. Devido à sua geomorfologia, ambientes fechados e de difícil acesso, como dolinas, impõem restrições ao monitoramento por métodos tradicionais. Apesar da alta resolução taxonômica proporcionada pelo gene 12S, uma base de dados de referência mais abrangente é essencial para estudos de metabarcoding de eDNA em escala ecossistêmica. Essa lacuna, como destacado por Jackman *et al.* (2021), limita o potencial dos métodos e reforça a importância da ampliação das bases de dados, especialmente para espécies neotropicais sub-representadas. No presente estudo, obtivemos 60% dos MOTUs com identidade superior a 98% em relação ao banco de dados NCBI; destes, pouco mais da metade não necessitou de ajustes taxonômicos após a curadoria.

## 2.6 CONCLUSÃO

Nossos resultados destacam a eficácia do marcador de DNA selecionado (MiFish 12S) para a detecção de uma ampla variedade de táxons, mesmo na ausência de uma base de dados de referência local. Ambientes fechados, como as dolinas, apresentam desafios únicos de

monitoramento devido à sua geomorfologia complexa e à acessibilidade limitada. A protegida Formação Furnas, localizada no interior do Parque Estadual de Vila Velha (PEVV) e composta por estruturas antigas de arenito, abriga populações de importância evolutiva significativa. Diante dessas complexidades, a integração de metodologias diversas para a avaliação da biodiversidade, tanto na região das Furnas quanto na bacia do rio Tibagi, contribuiu para a identificação da diversidade de espécies com fins conservacionistas e para subsidiar estratégias futuras de manejo.

### 3 CAPÍTULO 2

#### **Isolamento ancestral inferido a partir da diferenciação genética em populações de *Psalidodon fasciatus* (Teleostei: Characiformes)**

##### 3.1 INTRODUÇÃO

Metodologias inovadoras baseadas em sequenciamento de próxima geração transformaram a pesquisa em biodiversidade dentro da ecologia, biologia evolutiva e genética de populações (Storfer *et al.*, 2018). Métodos de sequenciamento de alto rendimento tem sido utilizados em investigações da ictiofauna de água doce regional (Pritchard *et al.*, 2000; Peterson *et al.*, 2012; Farrell *et al.*, 2016; Bengston *et al.*, 2018). Populações sob condições adversas podem sofrer mudanças genômicas que podem ser elucidadas por abordagens genômicas integrativas. O tetra Neotropical *Astyanax mexicanus* Baird & Girard, 1854 é uma espécie modelo que proporciona populacional em ambientes de superfície e cavernas (Strecker *et al.*, 2004; Wilkens; Strecker, 2017; Warren *et al.*, 2021). Além disso, *Astyanax* é um dos gêneros mais ricos em espécies, com ampla distribuição e pouca diferenciação morfológica e ecológica (Gurgel, 2004; Hirt *et al.*, 2011).

A colonização e adaptação a diversos habitats são características importantes para a existência de diferentes espécies em ambientes de cavernas, subterrâneos ou cenotes. O Escudo Brasileiro possui uma enorme variedade de formações de composição cárstica que constituem o habitat de uma ampla variedade de linhagens hipógenas, frequentemente resultando em isolamento genético e diferenciação fenotípica (Mattox *et al.*, 2008; Mendes *et al.*, 2019; Trajano, 2021). O Brasil chama atenção pela riqueza de cavernas subterrâneas registradas; no entanto, peixes que habitam formações isoladas de cenotes/furnas são especialmente raros em todo o mundo. A formação furnas do Sul do Brasil, localizada no Grupo Paraná (Sedorko *et al.*, 2017), contém quatro furnas, com apenas uma abrigando uma população isolada de *Psalidodon fasciatus* (Teleostei: Characiformes: Acestrorhamphidae: anteriormente *Astyanax*): a fuma 2 do Parque Estadual de Vila Velha (PEVV), Paraná (Gross *et al.*, 2004a; De la Ossa-Guerra *et al.*, submetido). O sistema hidrológico do PEVV corresponde ao rio Tibagi, que flui para o norte em uma extensão de aproximadamente 550 km na bacia do alto Paraná (Shibatta *et al.*, 2007). A diversidade de peixes do Tibagi tem sido mais estudada nos últimos anos, assim como inventários de seus afluentes (Claro-García *et al.*, 2018).

*Psalidodon fasciatus* é reconhecido como um complexo de espécies (Terán *et al.*, 2020) na recém-criada família Acestrorhamphidae (Melo *et al.*, 2024), com a localidade tipo

sendo o rio São Francisco (Cuvier, 1819). Estudos genéticos relataram polimorfismos cromossômicos em diferentes populações de *P. fasciatus* (Artoni *et al.*, 2006; Pazza *et al.*, 2008) e algumas inferências taxonômicas utilizando as regiões mitocondriais 12S rRNA (Matoso *et al.*, 2013) e o gene citocromo b (Mello *et al.*, 2015). Filogenias de tetras neotropicais têm consistentemente encontrado clados bem definidos para *Astyanax*, incluindo a maioria das espécies trans-andinas e da América Central, assim como o clado *Psalidodon*, que inclui a maioria das espécies do Escudo Brasileiro e das ecorregiões de La Plata (Melo *et al.*, 2022, 2024). No entanto, há uma falta de estudos moleculares comparativos para a maioria das espécies de *Psalidodon*, considerando as cerca de 50 espécies no gênero (Toledo-Piza *et al.*, 2024; Melo *et al.*, 2024).

Pesquisas sobre a ictiofauna que habita furnas são limitadas; no entanto, alguns estudos relataram um declínio na variabilidade genética entre populações de *P. fasciatus* na furna 2 do PEVV (Gross *et al.*, 2004b). Pequenas populações isoladas, com migração ou fluxo gênico limitado, frequentemente devido a barreiras físicas, são particularmente vulneráveis à diferenciação populacional e à deriva genética (Coleman *et al.*, 2018). Interrupções na conectividade hidrológica podem impactar significativamente a estrutura genética, o tamanho populacional, a diversidade genética e os padrões filogeográficos (Argentina *et al.*, 2018). Um exemplo bem documentado é *Gambusia hubbsi* nos *blueholes* da ilha de Andros, que exhibe alta estrutura populacional e baixa heterozigosidade devido ao isolamento reprodutivo (Schug *et al.*, 1998). Esses processos podem ser compreendidos de forma mais eficaz no contexto da história natural das populações ao longo de escalas de tempo evolutivas e pela análise das taxas de variabilidade genética. Além disso, a divergência evolutiva, a idade populacional e as mudanças ambientais fornecem *insights* fundamentais sobre os mecanismos subjacentes a esses padrões de isolamento.

### 3.2 OBJETIVO

Neste estudo, utilizamos dados de SNPs e genomas mitocondriais completos para investigar os efeitos genéticos e estimar o tempo de divergência de populações isoladas e não isoladas de *Psalidodon fasciatus* no PEVV e outros sistemas abertos. Nossa hipótese é que a divergência e colonização da população isolada de *P. fasciatus* no PEVV pode estar associada a zonas hiporreicas antigas (ou seja, leitos de rios). Os resultados deste estudo nos ajudam a compreender a base evolutiva da colonização, o tempo e a dinâmica populacional de organismos de água doce adaptados a ambientes de furna no sul Neotropical.

### 3.3 MATERIAIS E MÉTODOS

#### 3.3.1 Declaração de ética

As coletas foram autorizadas pelo Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio — SISBIO — licença número 15115-1), e os procedimentos experimentais foram aprovados pelo Comitê de Ética no Uso de Animais para Pesquisa da UEPG (Processo CEUA número 0769342/2021). O uso dos dados genéticos neste estudo foi autorizado pelo Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado (SISGEN número 6F96AE).

#### 3.3.2 Double digest restriction-site associated DNA (ddRADseq)

##### 3.3.2.1 Coleta de Espécimes, Amostragem de Tecidos e Extração de DNA

Espécimes de *Psalidodon fasciatus* foram coletados com redes de arrasto na furna 2 (-25° 13'28.002"S, -50°2'27.6072"O) (n=29) e na lagoa Dourada (25°14'23"S, 50°03'01"O) (n=29) no Parque Estadual de Vila Velha (PEVV), na cidade de Ponta Grossa; no rio Iapó, bacia do rio Tibagi, no município de Castro (606395S, 7261715O) (n=29), e no rio Paraopeba, São Francisco (20° 2'42.41"S, 44°12'27.20"O) (n=16). Amostras de tecido da nadadeira foram coletadas e armazenadas em álcool 95%. Os espécimes foram depositados no Museu do Capão da Imbuia, Curitiba, Brasil, sob o número de voucher MHNCI 12786. A extração total de DNA foi realizada com o kit *DNeasy Blood and Tissue* (Qiagen), e a qualidade foi verificada por eletroforese em gel de agarose a 1%. A pureza foi avaliada no espectrofotômetro GE NanoVue Plus, e a concentração (ng/μl) foi medida no fluorômetro Qubit com o kit Qubit dsDNA HS Assay (Invitrogen, EUA). Todas as amostras foram padronizadas para ±30 ng/μl.

##### 3.3.2.2 Preparação das bibliotecas

Utilizamos um protocolo modificado para a construção de duas bibliotecas ddRADseq (Peterson *et al.*, 2012). Em resumo, 450 ng de DNA genômico de cada espécime foram digeridos usando a combinação de duas enzimas de restrição, SphI e MluCI (New England Biolabs), e subsequentemente ligados a adaptadores específicos (P1 e P2, 0,25 μM) usando a enzima T4 DNA ligase a 23°C por 4 horas e a 65°C por 10 minutos. Os adaptadores P1 continham cinco nucleotídeos adicionais que funcionavam como etiquetas individuais (*barcode*). Fragmentos de DNA entre 200–300 pb foram selecionados em gel de agarose. Posteriormente, ensaios de PCR foram realizados para incorporar a identificação de cada

biblioteca. O PCR foi realizado sob condições padrão da enzima Platinum SuperFi *DNA Polymerase* (Thermo Fischer Scientific). As reações foram purificadas com o kit *ProNex Size-Selective Purification System* (Promega), e a concentração foi verificada por fluorometria no Qubit 3.0 (Thermo Fisher Scientific). Finalmente, as bibliotecas foram sequenciadas em uma *lane* do NovaSeq XPlus PE150 com 20% de PhiX (Novogene).

### 3.3.2.3 SNPs calling

As leituras brutas do Illumina foram desmultiplexadas usando a ferramenta *process\_radtags* incluída no pipeline de Stacks v2.0 (Catchen *et al.*, 2013), amostras com menos de um milhão de reads foram descartadas (Anexo A). Os *contigs* ddRAD foram alinhados ao genoma de referência de *Astyanax mexicanus* (Número de acesso: GCA\_023375975.1) usando o *bwa* com parâmetros padrão. Em seguida, utilizamos o programa *gstacks* com os parâmetros *gt-alpha=0,01* e *var-alpha=0,05*. Finalmente, filtramos os SNPs usando o programa *populations* com os parâmetros *r=0,8*, *t=8*, *min\_mac=5* e *min\_maf=0,05*. Os detalhes da pipeline usada estão descritos no Anexo B.

### 3.3.3 Análise de diversidade genética e posição de SNPs entre populações

Após a filtragem de SNPs de baixa qualidade em todo o genoma, 22.036 locais de variantes foram utilizados para análises subsequentes. As médias da frequência do alelo menor (MAF) foram calculadas para cada população no software *plink* (Purcell *et al.*, 2007). A heterozigiosidade observada ( $H_o$ ), a heterozigiosidade esperada ( $H_e$ ) e a diversidade nucleotídica ( $\pi$ ) foram extraídas das estatísticas sumárias do programa *populations* no Stacks. O coeficiente de endogamia ( $F$ ) foi calculado usando o software *vcftools* (Danecek *et al.*, 2011).

Para identificar posições de SNPs únicas para cada população e aquelas compartilhadas entre populações, processamos os dados genotípicos de cada amostra no arquivo VCF. Apenas posições com genótipos variantes (não referência) foram incluídas, enquanto genótipos homozigotos de referência (0|0 ou 0/0) foram excluídos para focar exclusivamente em posições variantes. SNPs compartilhados entre as quatro populações foram determinados calculando a interseção das posições de SNPs em cada conjunto populacional, resultando em uma contagem de SNPs comuns a todas as populações. Para visualizar SNPs compartilhados e únicos entre as populações, criamos um diagrama de Venn de quatro conjuntos, com cada conjunto representando uma população distinta. Além disso, um gráfico UpSet foi gerado para quantificar e ilustrar as combinações de posições de SNPs

compartilhadas entre populações. Isso foi feito por meio de uma matriz binária de presença-ausência, que nos permitiu identificar padrões específicos de compartilhamento e exclusividade de SNPs. A análise foi conduzida usando Python. O pacote *vcfpy* foi usado para análise de arquivos VCF, o comando *venn* foi usado para gerar o diagrama de Venn, e o comando *upsetplot* foi usado para produzir o gráfico UpSet. Os detalhes da pipeline usada estão descritos no Anexo B.

### 3.3.4 Análise de genética populacional e inferência demográfica

O índice de fixação ( $F_{ST}$ ) foi calculado no pacote de R *hierfstat*, e a análise de variância molecular (AMOVA) foi calculada para observar a variação genética entre populações usando o pacote de R *poppr*. A análise de Identidade por Estado (IBS) foi obtida no software *plink* (Purcell *et al.*, 2007) e plotada no software R. Estimamos a proporção do genoma de cada indivíduo atribuída a *pools* genéticos ancestrais  $K$  usando a função *sNMF* no pacote LEA, incluído também no R (Frichot; François, 2015). O número de clusters ( $K$ ) foi testado de 1 a 10 com 100 repetições e 1.000 iterações, e o  $K$  ótimo foi determinado por validação cruzada e pelo critério de entropia, que reflete o número de populações ancestrais que melhor explica os dados genéticos com base na teoria da informação. A demografia populacional foi analisada usando o Stairway Plot v2.1.1 (Liu; Fu, 2020), com espectros de frequência de SNPs. Aplicamos a taxa de mutação para espécies de peixes estimada por Bergeron *et al.* (2023) em  $5,97 \times 10^{-9}$  por geração. O tempo de geração foi definido em três anos, representando a vida média de *P. fasciatus*.

### 3.3.5 Genomas mitocondriais e estimativa de tempo

O sequenciamento de próxima geração foi usado para montar três genomas mitocondriais completos de três populações de *Psalidodon fasciatus*: furna 2, lagoa Dourada e bacia do rio Paraopeba. As bibliotecas foram sequenciadas na plataforma DNBSeg para obter leituras pareadas de 2X150 pb, resultando em 13,7 milhões de leituras brutas. As leituras brutas foram avaliadas quanto à qualidade usando FASTQC (Andrews, 2010) e limpas usando Trimmomatic v0.39 (Bolger *et al.*, 2014) para remover regiões de baixa qualidade. O pipeline Getorganelle v1.7.6.1 (Jin *et al.*, 2020) foi usado para montagem de novo, e o MitoAnnotator (Iwasaki *et al.*, 2013) foi usado para anotação de genes. Construímos uma matriz com os 13 genes codificadores de proteínas dos genomas mitocondriais e os alinhamos com 18 outros mitogenomas de Characiformes disponíveis no NCBI (Anexo C) usando MAFFT v.7.0. Esses incluem quatro espécies de Bryconidae (*Brycon* e *Salminus*), uma Acestrorhynchidae

(*Acestrorhynchus*), e nove Acestrorhamphidae (*Grundulus*, *Hasemania*, *Paracheiroduon*, *Petitella* e *Psalidodon*); usamos uma Chalceidae (*Chalceus*) para enraizar as árvores com base na filogenia de Characiformes (Melo *et al.*, 2022).

Os tempos de divergência e a reconstrução filogenética foram estimados no BEAST 2.6.3 (Bouckaert *et al.*, 2019) usando um relógio relaxado não correlacionado (taxa média do relógio = 1,0) sob o modelo evolutivo GTR+G+I (G=4; I=0,5) e um modelo de ramificação *birth-death* para a probabilidade de especiação (taxa de diferença de nascimento = 1,0; taxa de morte relativa = 0,5). Usamos dois fósseis como calibrações primárias e uma restrição de raiz como calibração secundária. O primeiro fóssil foi o †*Brycon avus* articulado da Formação Tremembé, do limite Oligoceno-Mioceno de São Paulo, Brasil (Woodward, 1898). Com esse fóssil, calibramos o nó que engloba todos os *Brycon* e *Salminus*, exceto o *B. henni* trans-andino, hipotetizado como pertencente a um clado inicialmente divergente de Bryconidae (Abe *et al.*, 2014; Melo *et al.*, 2022). Esse *prior* tem uma distribuição *lognormal*, *offset* = 23,0, média = 5,0, desvio padrão = 1,0. O segundo fóssil foi o †*Paleotetra entrecorregos* da Formação Entre-Córregos, do limite Eoceno-Oligoceno de Minas Gerais, Brasil (Weiss *et al.*, 2012). Usamos †*Paleotetra* para calibrar o nó que inclui todas as espécies de Acestrorhamphidae, usando uma distribuição *lognormal*, *offset* = 33,9, média = 5,0, desvio padrão = 1,0. Por fim, usamos uma restrição na raiz com uma distribuição normal (média = 89,28, sigma = 9,0; quantis 95–5% = 104–74,5) representando a divergência entre Chalceidae e o clado contendo Bryconidae, Acestrorhynchidae e Acestrorhamphidae. Essa restrição segue a filogenia calibrada no tempo da ordem Characiformes, que estimou a divergência de Chalceidae e esse clado principal em cerca de 89,2 milhões de anos atrás (104–74 Ma, 95% HPD) (Melo *et al.*, 2022). Realizamos três execuções independentes no BEAST2 com 100.000.000 gerações de MCMC, salvando árvores a cada 10.000 gerações. O Tracer v1.7.1 (Rambaut *et al.*, 2018) foi usado para verificar convergência, tamanho efetivo da amostra (ESS>200) e distribuições de traços. O TreeAnnotator v2.6.2 analisou as 26.403 últimas árvores (12% de *burn-in*) e selecionou a árvore de máxima credibilidade de clado (MCC), que foi subsequentemente visualizada e editada no FigTree v1.4.4 (Rambaut, 2012).

### 3.4 RESULTADOS

#### 3.4.1 Filtragem e processamento de dados brutos

Obtivemos uma média de 164.558.417 leituras retidas por biblioteca. Após o processamento dos dados utilizando o programa *process\_radtags*, 84 indivíduos apresentaram

cobertura de leitura suficiente para análises subsequentes (Anexo 2). Em média, 77.46% das leituras por indivíduo foram mapeadas com sucesso no genoma de referência de *Astyanax mexicanus*. Após a aplicação de um filtro de frequência do alelo menor (MAF) de 0,05 e permitindo até 20% de dados ausentes por loco, identificamos um total de 22,036 SNPs em todo o genoma.

### 3.4.2 Análise de diversidade genética e posição de SNPs entre populações

A frequência média do alelo mais comum por população (P) variou de 0,866 a 0,931, enquanto a heterozigosidade observada variou de 0,085 a 0,117, considerando todos os locais genômicos (Tabela 3, Fig. 4B). A população da lagoa Dourada exibiu a maior diversidade genética, refletida em sua heterozigosidade observada e esperada (Tabela 3, Fig. 4B), diversidade nucleotídica (Fig. 4A) e o maior número de alelos privados (Tabela 3). Em contraste, a população da furna 2 apresentou a menor diversidade genética. Notavelmente, a população do rio Paraopeba não apresentou alelos privados. Tanto as populações do rio Paraopeba quanto do rio Iapó exibiram valores semelhantes de diversidade nucleotídica. A população da furna também apresentou a menor porcentagem de SNPs em todo o genoma desviando do equilíbrio de Hardy-Weinberg ( $p < 0,05$ ) (Tabela 3).

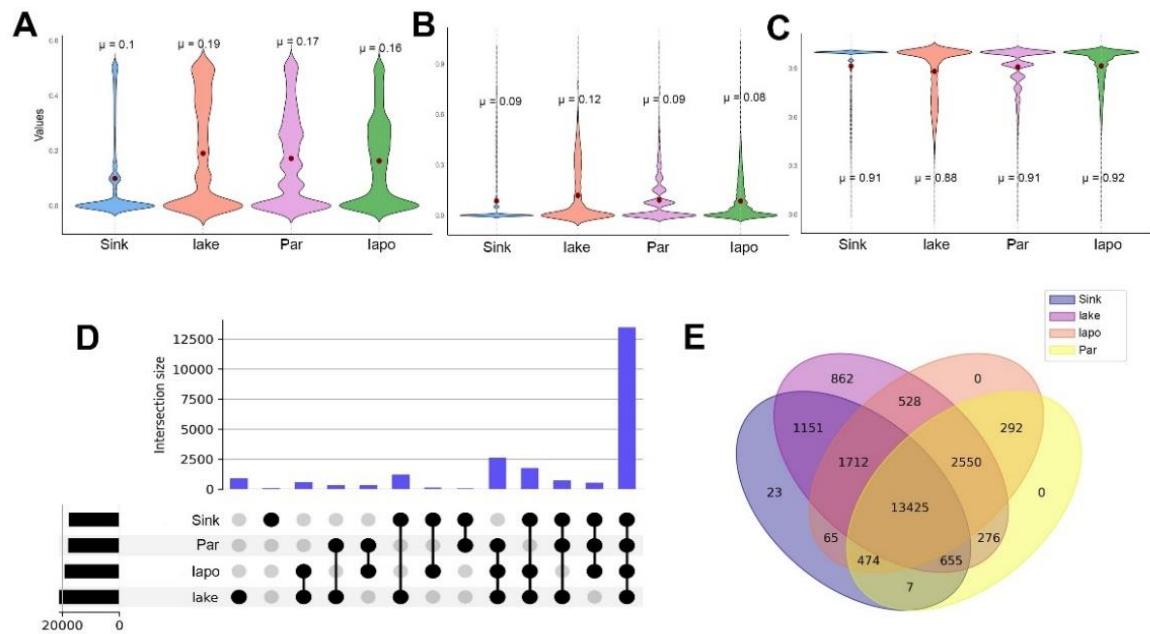
Tabela 3 - Métricas de diversidade a nível populacional de *Psalidodon fasciatus*. Número de amostras (N), alelos privados ( $P_A$ ), sítios polimórficos ( $Pol_{sites}$ ), porcentagem de loci polimórficos ( $\%Pol_{loci}$ ), heterozigosidade observada ( $Het_o$ ), heterozigosidade esperada ( $Het_e$ ), e porcentagem of genome-wide SNPs no hardy Weinberg disequilibrium.

População	N	$P_A$	$Pol_{sites}$	$\%Pol_{loci}$	$Het_o$	$Het_e$	$\%HWE (p < 0.05)$
Furna 2 (PEVV)	20	485	8409	0.11549	0.0856373	0.09683191	10%
Lagoa Dourada (PEVV)	19	4940	13262	0.18214	0.1177429	0.1868107	24%
Iapó river	27	48	14779	0.20298	0.0849937	0.1599085	30%
Paraopeba river	13	0	14492	0.19904	0.09085332	0.1655077	23%

Fonte: A autora

Identificamos 13.425 SNPs compartilhados entre as quatro populações (Fig. 4D, E), o que representa 60,9% de todos os SNPs detectados neste estudo. O maior número de SNPs únicos foi encontrado na população da lagoa Dourada, enquanto nenhum SNP único foi encontrado nas populações do Iapó e do Paraopeba (Figura 4E). As posições variantes estão bem distribuídas entre todas as populações, como evidenciado tanto pelo tamanho das interseções no gráfico UpSet quanto no diagrama de Venn (Figura 4D, E).

Figura 4 - Diversidade genética e distribuição de SNPs entre as populações



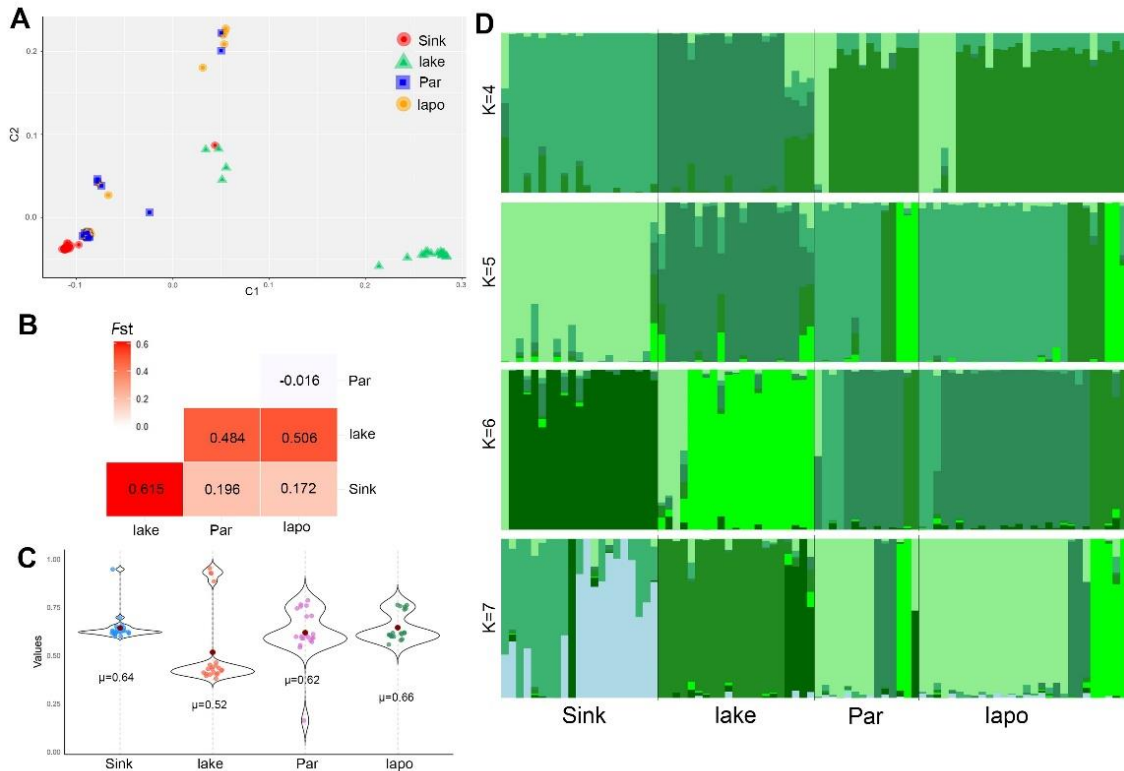
**A** - Diversidade nucleotídica (Pi). **B** - Heterozigosidade observada. **C** - Homozigosidade observada. **D** - Gráfico UpSet e **E**) Diagrama de Venn mostrando os SNPs compartilhados e únicos. Sink = cenote, lake = lago Dourada, Par = rio Paraopeba, Tib = rio Iapó. Os valores de  $\mu$  nos gráficos A, B e C representam o valor médio.

Fonte: A autora

### 3.4.3 Estrutura populacional

Os resultados sugerem uma forte estrutura populacional para a população isolada da furna 2 em comparação com as demais populações, conforme revelado pelas análises de IBS e sNMF (Figura 5A, D). A análise sNMF identificou  $K = 6$  como o número mais provável de clusters, com valores adicionais de  $K$  destacando uma estrutura populacional hierárquica significativa (Figura 5D). As populações dos rios Iapó e Paraopeba apresentaram padrões de agrupamento mais complexos (IBS e sNMF, Figura 5A, D). Em contraste, a população da lagoa Dourada exibiu dois clusters predominantes, um dos quais parece formar um grupo genético distinto. Essa população também mostrou diferenciação genética significativa em relação às outras, conforme indicado pelos valores elevados de  $F_{st}$  (Fig. 5B). As comparações pareadas de  $F_{st}$  revelaram maior diferenciação genética entre a população da lagoa Dourada e as outras populações.

Figura 5- Estrutura genética de *Psalidodon fasciatus* entre as populações



**A** - Identidade por Estado (IBS), **B** - Coeficiente de Diferenciação  $F_{st}$ , **C** - Coeficiente de Endogamia  $F$ , **D**- Padrões de estruturação hierárquica entre as populações de *P. fasciatus*. Estimativas dos coeficientes de mistura para um indivíduo em cada população usando a função ‘sNMF’ no pacote R LEA para diferentes valores de  $K$ . Sink = furna 2, lake = lagoa Dourada, Par = rio Paraopeba, Tib = rio Iapó. Fonte: A autora

A maior diferenciação (38,22%) ocorreu entre as populações (Tabela 4). Os dados indicam que há um nível moderado de diferenciação genética entre as populações. As populações em estudo são geneticamente distintas, conforme evidenciado pela diferenciação observada entre elas. No entanto, elas não estão completamente diversificadas, pois o fluxo gênico ou ancestralidade compartilhada ainda pode ocorrer. Além disso, as amostras individuais coletadas dessas populações apresentaram variabilidade entre si. Apesar dessas distinções, os resultados da AMOVA (Análise de Variância Molecular) não revelaram variação significativa em nenhuma das fontes de variação analisadas.

Tabela 4 - Análise da variância molecular (AMOVA). Df: graus de liberdade; SSD: soma dos desvios quadráticos; MSD: média dos desvios quadráticos; %: porcentagem da variância total atribuída a cada componente. Valores significativos (\*).

Fonte de variação	Df	SSD	MSD	%
Entre populações	3	90566,96	30.188,99	38,22*
Entre amostras dentro das pop	80	131603,81	1.645,0	28,73
Dentro das amostras	84	50452,41	600,62	33,05
<b>Total</b>	167	272623,17	1.632,47	100

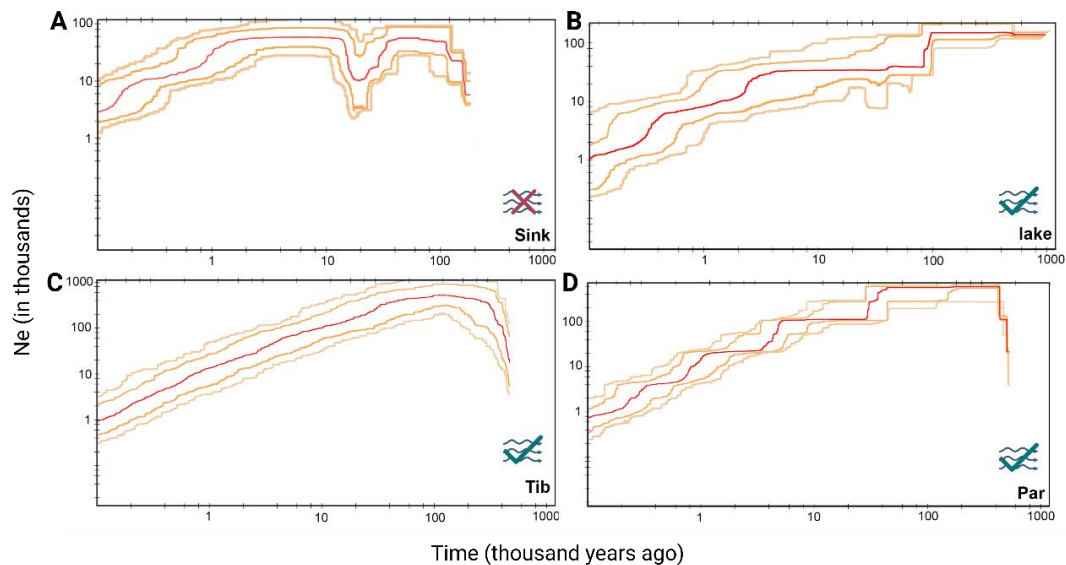
Fonte: A autora

Como estamos comparando diferentes populações, incluindo uma que está completamente isolada, nosso objetivo também é avaliar se a endogamia está contribuindo para as fontes de variação genética e levando a um excesso de homozigotos. Neste sentido, o coeficiente de endogamia  $F$ , de fato, revelou que a população isolada (furna 2) exibiu um nível mais alto de homozigosidade (Figura 5C). Além disso, os dados mostram uma variação mínima, sugerindo que todos os indivíduos têm um grau semelhante de endogamia.

#### 3.4.4 Inferência demográfica

De acordo com os resultados do Stairway Plot (Figura 6), as populações de *P. fasciatus* geralmente declinaram nos últimos 10.000 anos. No entanto, a população da furna passou por uma fase de expansão e estabilização entre 10 ka e 2 ka, seguida por um declínio gradual que persiste até o presente (Figura 6A). Essa população também pode ter experimentado um gargalo pronunciado entre aproximadamente 30 ka e 15 ka, coincidindo com o Último Máximo Glacial Global (GLGM, 26,5–19 ka) (Palacios *et al.*, 2020). Em contraste, as populações da lagoa Dourada, Iapó e Paraopeba mostram uma série de gerações, cada uma começando com um gargalo em torno de 400–100 ka, exceto a população do lago, que começa mais cedo com uma expansão (Figura 6B–D). Essas populações subsequentemente se estabilizaram, com a população do Paraopeba exibindo um período de estabilização mais longo (~360 ka) em comparação com as outras (~100 ka). Os tamanhos efetivos das populações ( $N_e$ ) diminuíram em todas as populações nos últimos 300 anos. Atualmente, a população a furna 2, habitando um ambiente fechado, mostra uma expansão populacional (~3000) em comparação com as outras populações (~1000) (Figura 6).

Figura 6 - História demográfica inferida das populações de *P. fasciatus*

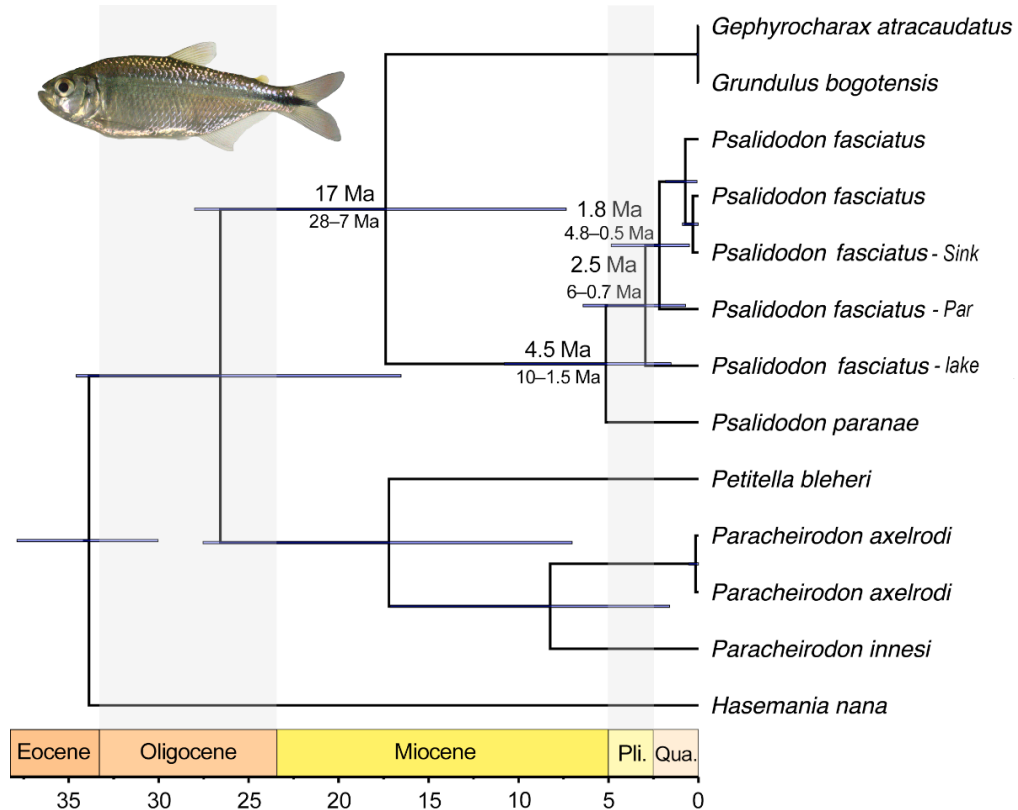


A linha vermelha representa a mediana de 200 inferências baseadas na subamostragem, enquanto as linhas laranja e laranja claro indicam, respectivamente, os intervalos de confiança de 75% e 95% da inferência. Os símbolos em cada gráfico mostram a presença ou ausência de fluxo gênico em cada população: **A** - população da furna 2, **B** - população da lagoa, **C** - população do rio Iapó e **D** - população do rio Paraopeba. Fonte: A autora

### 3.4.5 Estimativa do tempo de divergência

Para contextualizar os padrões genéticos observados nas populações de *P. fasciatus* desde uma perspectiva evolutiva, estimamos os tempos de divergência com base em dados do genoma mitocondrial. A análise incluiu 13 genes codificadores de proteínas (11.608 pb), identificando 6.406 sítios variantes e 5.603 sítios informativos para parcimônia. A composição nucleotídica foi T=30%, C=27,1%, A=27,3% e G=15,4%. A filogenia resultante apresentou altas probabilidades posteriores, com *Acestrorhamphidae* como grupo-irmão de *Acestrorhynchidae*, e esse clado, por sua vez, como grupo-irmão de *Bryconidae*. Os resultados indicam que o clado *Acestrorhamphidae* analisado tem uma idade estimada correspondente à divergência de *Hasemania* e dos demais acestrorhamphídeos durante o Eoceno (48–34 Ma, IC 95%). A separação entre *Grundulus* transandino e *Psalidodon* cisandino ocorreu entre o Oligoceno e o Mioceno (28–7 Ma, IC 95%). Dentro de *Psalidodon*, nossa filogenia calibrada no tempo indica que *P. paranae* divergiu de *P. fasciatus* durante o Plioceno, há aproximadamente 4,5 Ma (10–1,5 Ma, IC 95%). Além disso, a análise sugere uma divergência interna de *P. fasciatus* durante o Pleistoceno, em menos de 2,5 Ma (6,0–0,7 Ma, IC 95%) (Figura 7). A filogenia indica que as populações distintas analisadas de *P. fasciatus* formam um grupo monofilético.

Figura 7 - Filogenia calibrada no tempo de *Psalidodon* e táxons relacionados, baseada em genomas mitocondriais (13 genes; 11.608 pb)



Os números próximos aos nós indicam a idade média dos clados (Ma). Destaca-se a divergência entre o Plioceno e o Quaternário (2,5 Ma) entre os dois espécimes de *Psalidodon fasciatus* isolados (furna) e não isolado (lagoa) do Parque Estadual de Vila Velha. Fonte: A autora

### 3.5 DISCUSSÃO

O *Psalidodon fasciatus* foi originalmente descrito a partir de espécimes da bacia do São Francisco, no leste do Brasil (Cuvier, 1819). *Psalidodon fasciatus* tem sido citado em diversos estudos como um complexo de espécies, o que é particularmente relevante quando se observam agrupamentos genéticos diferenciados em todas as populações — com exceção da Furna 2. Curiosamente, a população da furna mantém um perfil genético semelhante ao da localidade-tipo no rio São Francisco (referido aqui como Paraopeba), conforme indicado pelos valores de  $F_{st}$  (0,196; Figura 5B). Essa similaridade genética é inesperada, considerando o isolamento geográfico da furna, que normalmente promoveria uma maior divergência genética ao longo do tempo devido à ausência de fluxo gênico. No entanto, isso pode ser explicado por ILS (*Incomplete Lineage Sorting*), uma vez que genótipos ancestrais podem persistir ao longo do tempo, mesmo após a separação das linhagens (Combrink *et al.*, 2025).

Várias hipóteses podem explicar a colonização de *P. fasciatus* na furna 2. Uma delas surge da natureza cárstica da região, caracterizada por flutuações na energia e no fluxo de massa influenciados por variáveis externas ao longo do tempo (Gibert *et al.*, 1994). Esse ambiente dinâmico pode ter facilitado trocas faunísticas entre áreas anteriormente conectadas e drenagens fluviais vizinhas. A furna 2 faz parte da Formação Furnas (Melo *et al.*, 2011), conhecida por suas paleocorrentes e arenitos associados. A história da erosão subterrânea nesta formação levou ao desenvolvimento de grandes depressões, incluindo furnas e lagoas (Melo *et al.*, 2004). Além disso, os processos de intemperismo em Vila Velha, responsáveis pelas características pseudo-cársticas, começaram há aproximadamente 35 milhões de anos e foram mais intensos entre 17 e 9 milhões de anos atrás (Riffel *et al.*, 2015). Embora esses processos geológicos tenham contribuído para a formação da furna 2, eles não explicam completamente a presença de *P. fasciatus* nesse habitat isolado. A colonização de *P. fasciatus* na furna 2 permanece sem explicação, uma vez que esses eventos geológicos antecedem a divergência da população, que ocorreu durante o período Plioceno-Quaternário (Figura 7).

O Lagoa Dourada, com aproximadamente 200 metros de diâmetro e cercado por floresta ombrófila mista ripária, está localizado próximo a furna. Estimativas de idade baseadas em restos vegetais sugerem que o lago existe há mais de 11.700 anos (de Melo *et al.*, 2000). Embora tenha sido proposta uma conexão subterrânea entre a lagoa Dourada e a furna 2, não há evidências conclusivas que suportem fluxo gênico ou migração entre os habitats. Os dados genéticos indicam que a população da lagoa Dourada diverge de outras populações, com um tempo de divergência relativamente antigo (~2,5 milhões de anos), o que provavelmente explica os altos valores de  $F_{st}$  observados. Essa maior diversidade pode estar associada a uma história de colonização mais antiga e a um possível fluxo gênico com outros sistemas fluviais em períodos anteriores.

Populações fluviais, como as do Iapó e Paraopeba, apresentam menor estruturação genética, o que pode ser atribuído à maior vagilidade comum em espécies de peixes de rio (Jaisuk; Senanan, 2018). A diversidade genética ( $\pi$ ) dessas populações fluviais é semelhante à da população da lagoa. Embora uma maior diversidade fosse esperada nessas populações fluviais, suas taxas de endogamia não são tão baixas quanto as observadas na população da furna 2, possivelmente devido a declínios populacionais recentes (Figura 7), que podem aumentar os níveis de homozigidade, como também demonstrado aqui (Figura 7C). Populações com fluxo gênico restrito ou ausente são suscetíveis à endogamia (Figura 2C,  $F_{coeff}=0,64$ ), o que pode levar a efeitos adversos no crescimento, fertilidade e desenvolvimento,

conforme relatado em trabalhos anteriores (Nelson *et al.*, 2020; Kerniske *et al.*, 2021). Além disso, evidências recentes de hibridização natural em populações de *P. fasciatus* do alto Paraná (Gavazzoni *et al.*, 2020) podem afetar o fluxo genético das populações fluviais em contraste com as populações isoladas da furna.

Populações geograficamente isoladas são raras e difíceis de identificar, especialmente em ambientes aquáticos (Hastings, 2005). Em pequenas populações de peixes isoladas, a variação genética e os processos de extinção são frequentemente influenciados por eventos estocásticos (Hellmair; Kinziger, 2014). No entanto, o grau de isolamento, a história de vida e a dinâmica populacional afetam como essas populações evoluem e persistem ao longo do tempo (Balloux, 2016). A história demográfica ao longo do tempo (Figura 7) indica que a população da furna apresentou um tamanho populacional maior do que o esperado, considerando a ausência de taxa de migração. Em populações estáveis ou em expansão, uma ampla gama de mutações, incluindo aquelas sob seleção positiva e sobre dominantes, pode contribuir para a adaptação (McDonough; Connallon, 2023); isso pode ser aplicado à manutenção da população do sumidouro ao longo do tempo. Além disso, a heterozigosidade observada, que é semelhante à das populações fluviais, pode facilitar a geração de variantes adaptativas, apoiando a persistência da população (McDonough; Connallon, 2023). Essas descobertas recentes levantam uma questão instigante e ainda em aberto: quais mecanismos genéticos sustentam a capacidade dessas populações isoladas de peixes de contrabalançar a deriva genética e manter altos níveis de heterozigosidade?

### 3.6 CONCLUSÃO

Nossos resultados fornecem insights sobre a manutenção da variabilidade genética na população isolada de *P. fasciatus*. Apesar dos níveis de endogamia, a presença de alelos em equilíbrio de Hardy-Weinberg e alelos privados sugere que essa população persiste há menos de um milhão de anos. Ao longo de inúmeras gerações, é provável que essa população tenha evoluído mecanismos específicos para a sobrevivência. Pesquisas futuras se concentrarão na investigação desses mecanismos, bem como nas estratégias reprodutivas empregadas por *P. fasciatus* na furna 2. Por fim, o método ddRADSeq foi fundamental para mapear SNPs entre as populações, fornecendo resultados robustos sobre a variação genética nos quatro ambientes.

## 4 Capítulo 3

### Mapeamento cromossômico de elementos repetitivos revela padrões microevolutivos em populações isoladas de *Psalidodon fasciatus* no sul do Brasil

#### 4.1 INTRODUÇÃO

As sequências de DNA repetitivo constituem uma grande proporção dos genomas de vertebrados (Reinar *et al.*, 2023). Em peixes, o percentual de elementos repetitivos varia entre as espécies e está relacionado à adaptabilidade e ao ambiente em que vivem (Yuan *et al.*, 2018). Os peixes teleósteos, por exemplo, apresentam tanto o menor (*Tetraodon nigroviridis*, 342,4 Mb) quanto o maior (*Lepidosiren paradoxa*, 91 Gb) genoma conhecido entre os vertebrados, o que pode ser explicado por eventos de duplicação genômica total e, em alguns casos, pela expansão de DNA repetitivo (Jaillon *et al.*, 2004; Scharl *et al.*, 2024). A abundância e diversidade de elementos repetitivos nos genomas de teleósteos têm implicações significativas para a evolução genômica, variações estruturais e adaptabilidade. Esses elementos podem influenciar a expressão gênica, o tamanho do genoma e o surgimento de novas funções genéticas (Yuan, 2018).

Elementos repetitivos como os elementos transponíveis (TEs) podem ser fatores importantes na reorganização genômica, mediada por rearranjos cromossômicos como inversões, translocações e duplicações (Carotti *et al.*, 2021). Os TEs são classificados em cinco classes, com base em sua enzimologia, estrutura e sequência: retrotransposons com repetições terminais longas (LTR), retrotransposons sem LTR (por exemplo, elementos nucleares longos e curtos dispersos – LINEs e SINEs), transposons de DNA tipo "copy paste", transposons de DNA com mecanismo de rolagem (Helitrons) e transposons de DNA autossintetizantes (Polintons) (Jurka, 2000; Jurka *et al.*, 2005). Em *Astyanax mexicanus*, por exemplo, os elementos transponíveis são predominantemente compostos por transposons de DNA, seguidos por LINEs e LTRs, enquanto SINEs são menos abundantes (Shao; Han; Peng, 2019). No entanto, na maioria dos grupos de teleósteos, os mecanismos de atividade desses elementos ainda são pouco compreendidos, especialmente em relação à sua localização cromossômica *in situ*, em grande parte devido à ausência de genomas de referência.

Por outro lado, elementos repetitivos como DNAs satélites (satDNAs), que consistem em sequências não codificadoras repetidas em tandem e localizadas principalmente nas regiões centroméricas, pericentroméricas e subteloméricas, têm se mostrado úteis em estudos de genética populacional devido às suas altas taxas de mutação, variabilidade de sequência e evolução específica de linhagens. Sua rápida evolução os torna particularmente úteis como

marcadores moleculares para avaliar diferenciação genética, estrutura populacional e evolução cromossômica entre táxons proximamente relacionados (Thakur; Packiaraj; Henikoff, 2021; Ugarkovic; Plohl, 2002). Os DNAs satélites vêm sendo mais estudados recentemente devido ao maior acesso a técnicas de sequenciamento de baixa cobertura (Novák; Neumann; Macas, 2020). Estudos de caracterização e comparação de famílias de DNA satélite em Characiformes têm aumentado nos últimos anos (dos Santos *et al.*, 2021; Goes *et al.*, 2022; Utsunomia *et al.*, 2019), permitindo análises comparativas entre espécies não apenas de peixes, mas de diversos outros grupos (Ruiz-Ruano *et al.*, 2016).

## 4.2 OBJETIVOS

O objetivo deste estudo é comparar a distribuição cromossômica e as implicações evolutivas de alguns elementos repetitivos, como LINEs e retroelementos Rex, em diferentes populações de *Psalidodon fasciatus* no Brasil. Além disso, este trabalho busca comparar os achados a fim de compreender a relevância na reorganização genômica e na diferenciação populacional, especialmente em habitats isolados e únicos, como as furnas.

## 4.3 MATERIAS E MÉTODOS

### 4.3.1 Amostragem e obtenção de cromossomos

Espécimes de *Psalidodon fasciatus* foram coletados com redes de arrasto na furna 2 (SH) (-25,224445°S, -50,041002°W) e na Lagoa Dourada (DL) (25°14'23"S, 50°03'01"W), ambos localizados no Parque Estadual de Vila Velha (PEVV), município de Ponta Grossa; no reservatório Alagados, também em Ponta Grossa (AL) (-25,013890, -50,034074), e no rio Iapó, bacia do rio Tibagi, no município de Castro (IA) (606395S, 7261715W). Os peixes foram transportados ao laboratório, e posteriormente foram realizadas preparações mitóticas de acordo com o protocolo de Bertollo *et al.* (1978) e por cultivo de fibroblastos.

Todos os procedimentos seguiram as normas internacionais de experimentação animal, sendo aprovada pelo Comitê de Ética no Uso de Animais da Universidade Estadual de Ponta Grossa (processo CEUA n.º 0769342/2021). A coleta foi autorizada pelo Ministério do Meio Ambiente (MMA/ICMbio n.º 15115-1). O DNA genômico foi acessado com registro no SisGen (Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado n.º A6F96AE).

Os cromossomos mitóticos obtidos a partir de cultura de fibroblastos seguiram o protocolo de Silva *et al.* (2014), com modificações. A nadadeira caudal dos espécimes foi coletada e armazenada em microtubos de 1,5 mL contendo PBS 1× suplementado com penicilina (100 U/mL) e estreptomicina (100 µg/mL) até o processamento para cultivo celular. Sob capela de fluxo laminar, a nadadeira foi removida do tubo, transferida para uma placa de Petri estéril e lavada três vezes com PBS 1×, descartando-se a solução após cada lavagem. Em seguida, a nadadeira foi transferida para uma nova placa de Petri estéril e fragmentada mecanicamente com lâmina de bisturi estéril. Os fragmentos foram transferidos para um tubo contendo 3 mL de colagenase tipo IV, previamente diluída em meio DMEM (Dulbecco's Modified Eagle's Medium) suplementado com penicilina (100U/mL), gentamicina (0,05 mg/mL), estreptomicina (0,05 mg/mL) e anfotericina B (2,5 µg/mL), sendo incubados a 37 °C por 20 minutos.

Após a digestão enzimática, as células foram ressuspensas em 7 mL de DMEM e centrifugadas a 800 rpm por 10 minutos. O sobrenadante foi descartado, e o pellet foi ressuspenso em 5 mL de DMEM suplementado com 20% de soro fetal bovino (FBS). As células foram então transferidas para frascos de cultura de 25 cm<sup>2</sup> e incubadas a 28 °C. O crescimento celular foi monitorado diariamente por microscopia invertida, e o meio foi renovado conforme necessário. Ao atingir cerca de 80% de confluência, foram adicionados 100 µL de colchicina a 0,005% para bloqueio da divisão celular, com incubação durante a noite a 28 °C.

Posteriormente, as células foram destacadas com tripsina a 0,25% e submetidas a tratamento hipotônico com solução de KCl 0,075 M por 5 minutos a 37 °C. Em seguida, as células foram fixadas com três lavagens sequenciais em solução de metanol:ácido acético (3:1). As lâminas foram analisadas por coloração convencional com Giemsa, utilizando microscopia óptica, a fim de verificar a qualidade das metáfases. Todos os procedimentos foram realizados conforme aprovado pelo Comitê de Ética no Uso de Animais da Universidade Estadual de Ponta Grossa (Processo CEUA: 0769342/2021).

#### **4.3.2 Citogenética convencional**

A identificação da heterocromatina constitutiva foi realizada por meio da técnica de bandamento C, conforme descrito por Sumner (1972), e a caracterização das regiões organizadoras de nucléolo (Ag-NORs) seguiu a metodologia proposta por Howell e Black (1980). Os cariótipos foram classificados segundo Levan *et al.* (1964) como metacêntricos (m), submetacêntricos (sm), subtelocêntricos (st) e acrocêntricos (a).

### 4.3.3 Obtenção das sondas de elementos repetitivos

#### 4.3.3.1 Elementos transponíveis

As bibliotecas genômicas de *P. fasciatus* foram sequenciadas com baixa cobertura (*low-pass*) pela BGI Genomics (China). A qualidade das bibliotecas foi avaliada com o software FASTQC (Andrews, 2010). Considerando os potenciais efeitos genéticos do isolamento na população de *P. fasciatus* na fuma 2, foi realizada a identificação e mapeamento dos elementos transponíveis (TEs) mais abundantes presentes em sua biblioteca genômica (Anexo E).

Estas sequências de TEs foram identificadas com base na anotação do RepeatExplorer (RE) (Novák *et al.*, 2020) disponível na plataforma Galaxy (<https://repeatexplorer-elixir.cerit-sc.cz/>). Os *contigs* montados foram extraídos e analisados com o auxílio do visualizador Tablet (Milne *et al.*, 2013). A identificação dos TEs foi realizada com base na primeira anotação dos resultados do RE, sendo posteriormente confirmada no banco de dados FishTEDB (<https://www.fishedb.com/>). Os *contigs* mais completos foram selecionados para o desenho de primers, os quais foram construídos utilizando o software Primer3 v4.1.0 (Kõressaar *et al.*, 2018) (Tabela 5). As reações de PCR para os nove elementos transponíveis foram realizadas com perfil térmico semelhante: 2 min a 95 °C, seguidos por 36 ciclos de 1 min a 94 °C, 45 s entre 55 °C–59 °C, e 2 min a 72 °C; finalizando com uma extensão de 5 min a 72 °C. As sondas dos TEs foram marcadas com Atto550-dUTP, Atto488-dUTP e Atto426-dUTP por Nick-Translation (Jena Biosciences, Jena, Alemanha), conforme as instruções do fabricante, e posteriormente utilizadas nos experimentos de FISH. Em seguida, esses elementos foram comparados com aqueles observados em populações não isoladas (DL, AL, IA). A presença e atividade dos TEs pode influenciar a função genômica e impactar a sobrevivência e a adaptação desse grupo isolado.

Tabela 5 - Primers utilizados para a amplificação dos fragmentos dos elementos transponíveis presentes na biblioteca de *P. fasciatus* da fuma 2

Cluster RE	Identificação	Primer forward	Primer reverse	Tamanho de produto (pb)
CL83	LINE/L1	TTCGTGTGTTTCTCTGCT GC	TGTAGAAGCGGGAGAA GGTG	241
CL84	LINE/Rex	CCACAAAGAGCATACG GACG	ACTGTGACTTGCTACCC GAA	811
CL90	LINE/L2	GCATGGTTTTGGAGAGC ACA	ACACAGTGAAGCCAGA CAGA	583
CL101	LINE/Rex	CGTTCAGCAGAGATGTG TCG	ATGCCAGCTTACAGACC ACT	511
CL108	LINE/L1	CGTGTGTTATCGGCTGA GTG	CGCCATGTGTTCTTTGG GAA	886
CL257	LINE/Rex-Babar	TACCTCCTCCACCGAAC CTA	AGACTGGCTGCAAGGGT AAT	505

<b>CL227</b>	LINE/Rex	CTGGTTTGCTCTGTCCA CAC	AGAGGGCGTTCAAAGAT GGA	213
<b>CL169</b>	LINE/Rex-Babar	TGGTACGGCAACACTAC TGT	TCACACTTCTGTACCTC C	539
<b>CL114</b>	LINE/Rex-Babar	CAGTTTGCCTACCGGGA AAA	CATCTTGCACCACACCA CAA	468

Fonte: A autora

#### 4.3.3.2 DNA satélites

Análises comparativas utilizando a *pipeline* do RepeatExplorer (Novák *et al.*, 2020) foram conduzidas para investigar possíveis diferenças na presença e abundância de elementos repetitivos, com foco especial nos DNAs satélites, nas bibliotecas genômicas disponíveis das populações amostradas neste estudo. Comparações iniciais foram realizadas entre as populações de SH vs DL, não sendo observadas diferenças relevantes na abundância geral dos elementos repetitivos, exceto pela detecção de uma variante do satélite *As5I*. Em seguida, foram realizadas análises comparativas entre a população da fuma 2 e a população do rio São Francisco, esta última escolhida por corresponder à localidade-tipo de *P. fasciatus* (Anexo F). Nessa comparação, foram detectadas diversas diferenças na presença/ausência de elementos satélites, sugerindo possíveis variações estruturais na população da fuma 2. As sequências de satélite selecionadas foram marcadas diretamente com biotina durante sua síntese e posteriormente mapeadas nas quatro populações estudadas.

#### 4.3.4 Hibridação *in situ* fluorescente (FISH)

As misturas de hibridação foram compostas por 100 ng de cada sonda de DNA marcada, além de 50% de formamida, 2× SSC, 10% de SDS, 10% de sulfato de dextrana e tampão Denhardt (pH 7.0), totalizando 20 µL, sob condições de alta stringência para FISH. Para os experimentos de FISH, as lâminas contendo metáfases foram envelhecidas por 1h a 60 °C, seguidas de tratamento a 37 °C por 5 min com solução de pepsina a 0,005% (99 µL de H<sub>2</sub>O, 10 µL de HCl e 2,5 µL de pepsina a 20 mg/mL). Os cromossomos foram desnaturados em formamida 70%/2× SSC a 72 °C por 3 min; as sondas foram desnaturadas separadamente a 85 °C por 10 min e resfriadas a 4 °C por 2 min antes da aplicação. A hibridação ocorreu durante a noite em câmara úmida a 37 °C. As lâminas foram então lavadas por 5 min em SSC 1× a 65 °C, seguidas por lavagem em SSC 4×/Tween em temperatura ambiente e uma lavagem rápida em PBS 1× por 1 min. As lâminas foram desidratadas em uma série alcoólica (70%, 85% e 100%) antes da contra-coloração dos cromossomos com DAPI, montadas com Vectashield (Vector Laboratories, Burlingame, EUA).

### 4.3.5 Imagens e confirmação dos resultados

Para confirmar os resultados obtidos por FISH, foram analisadas metáfases por indivíduo. As imagens foram capturadas com uma câmera CoolSNAP acoplada a um microscópio Axioplan II (Carl Zeiss Jena GmbH, Jena, Alemanha) e processadas com o software ISIS (MetaSystems Hard & Software GmbH, Altlussheim, Alemanha). Todas as imagens foram editadas no software Photoshop.

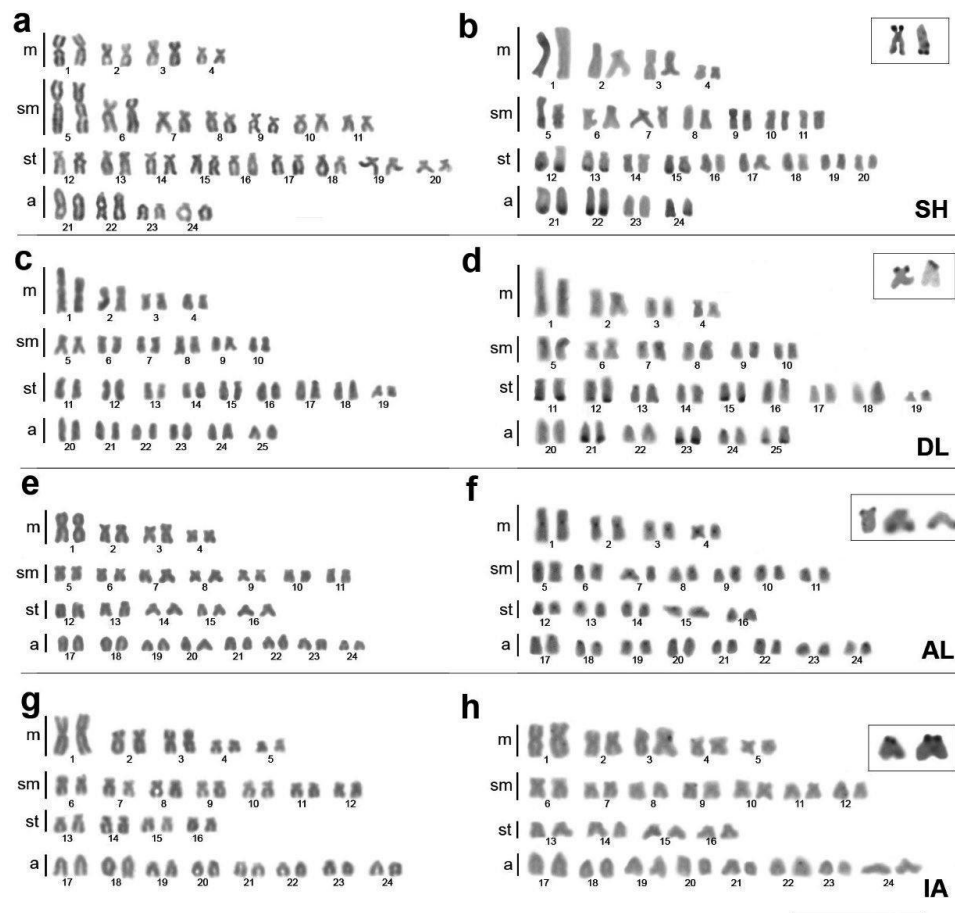
## 4.4 RESULTADOS

As populações SH, AL e IA apresentaram o mesmo número diploide modal de  $2n = 48$  cromossomos, com diferenças em suas fórmulas cariotípicas, conforme segue: SH =  $8m + 14sm + 18st + 8a$  [NF = 88] (Figura 8a); AL =  $8m + 14sm + 10st + 16a$  [NF = 80] (Figura 8e); IA =  $10m + 14sm + 8st + 16a$  [NF = 80] (Figura 8g). A população DL apresentou número diploide de  $2n = 50$  cromossomos, com a fórmula cariotípica DL =  $8m + 12sm + 18st + 6a$  [NF = 82] (Figura 8c).

Nas populações SH e DL, a heterocromatina constitutiva foram predominantemente observadas em blocos localizados nas regiões teloméricas de pares cromossômicos subtelocêntricos e acrocêntricos (Figura 8b, d). Em contraste, a população AL apresentou baixo conteúdo heterocromático, restrito às regiões centroméricas de todos os pares cromossômicos (Figura 8f). Na população IA, a heterocromatina foi extremamente escassa, limitada a poucos sinais intersticiais no cromossomo 1 e a um único sítio telomérico no braço curto do par cromossômico 9 (Figura 8h).

A localização das regiões Ag-NOR não foi conservada entre as populações. No entanto, as populações SH, DL e AL apresentaram uma região NOR ativa em um cromossomo submetacêntrico do par 6, o qual também corresponde ao par que carrega o sinal de rDNA 18S (Figura 9); essa localização tem sido conservada em outras populações de *P. fasciatus* (Pazza *et al.*, 2006). A população AL, em particular, apresentou o maior número de sítios AgNOR, embora apenas três cromossomos tenham sido consistentemente marcados nas metáfases. Em contraste, a população IA apresentou sinais AgNOR em um único par subtelocêntrico.

Figura 8 - Análise comparativa dos cariótipos de *P. fasciatus* analisados.



Cariótipos em giemsa (à esquerda) bandamento C (à direita) e cromossomos portadores de regiões NOR (em destaque) em quatro populações de *Psalidodon fasciatus*: furna 2 (SH) (a, b), Lagoa Dourada (DL) (c, d), Alagados (AL) (e, f) e rio Iapó (IA) (g, h). Barra de escala = 10  $\mu$ m. Fonte: A autora

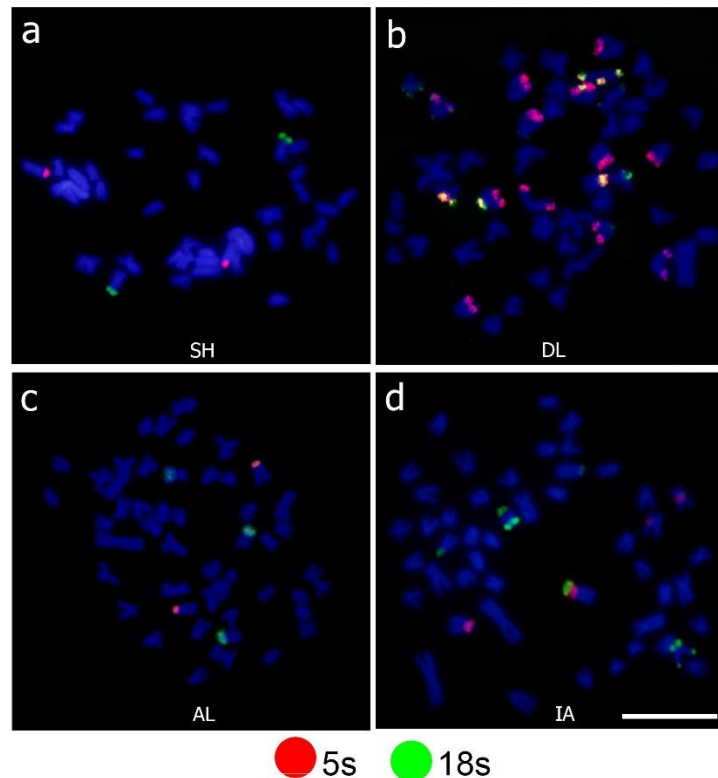
#### 4.4.1 Hibridação comparativa

##### 4.4.1.1 Mapeamento dos rDNAs 5S e 18S

Os padrões de hibridação revelaram variação tanto na localização cromossômica quanto no número de sítios de rDNA entre as populações. O sítio de rDNA 5S (Figura 9) apresentou localização conservada na região centromérica de um par acrocêntrico nas populações SH (par 21), AL (par 17) e IA (par 17) (Figura 9a, c, d). Em contraste, a população DL apresentou múltiplos sinais, distribuídos em sete pares cromossômicos.

Os sítios de rDNA 18S diferiram entre as populações, com co-localização dos rDNAs 18S e 5S observada em quatro cromossomos na população DL. Na população IA, também foi observada co-localização em um cromossomo acrocêntrico.

Figura 9 - Localização cromossômica do rDNA 5s e 18s em quatro populações de *P. fasciatus*



A - Furna 2 (SH). B - Lagoa Dourada (DL). C - Alagados (AL); e D - Rio Iapó (IA). Barra de escala = 10  $\mu$ m.  
Fonte: A autora

#### 4.4.2 Mapeamento de alguns LINEs (*Long interspersed nuclear elements*) nas populações de *P. fasciatus*

##### 4.4.2.1 L1 e L2

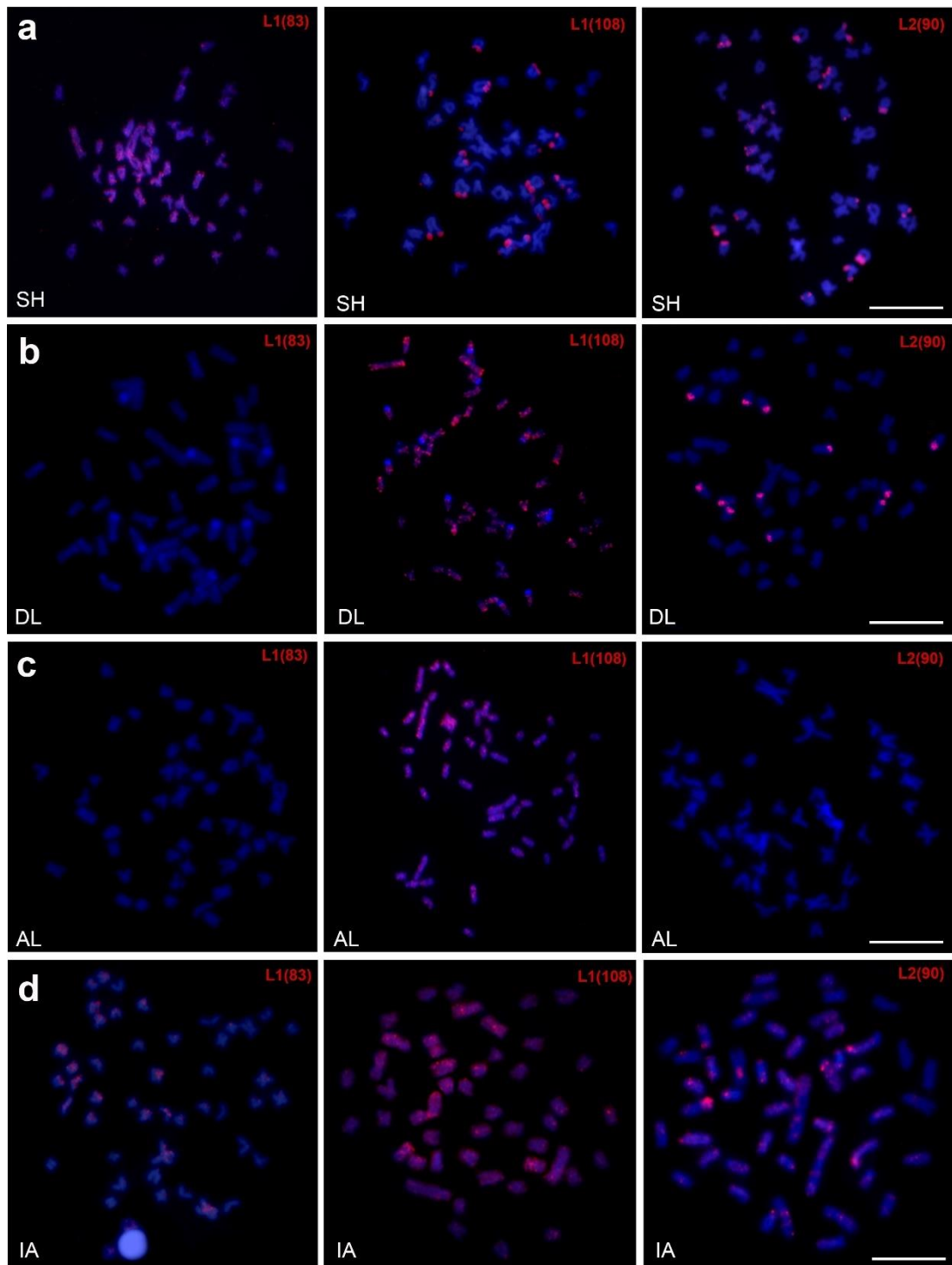
As sequências repetitivas LINE do tipo L1 apresentaram sinais dispersos em *P. fasciatus* (Figura 10) e blocos semelhantes a regiões teloméricas na população SH. A sonda L2 mostrou localização em blocos do tipo telomérico nos pares subtelo-cêntricos e acrocêntricos.

##### 4.4.2.2 Rex-Babar

O resultado mais marcante do mapeamento dos retrotransposons do tipo Rex-Babar foi que a população SH apresentou sinais positivos para todas as sondas, sendo que a sonda Rex-Babar-169 parece ser exclusiva dessa população. Nas populações AL e IA, a sonda Rex-Babar-257 marcou dois pares de cromossomos, com sinais localizados distintamente nas regiões teloméricas (Figura 11c, d). A sonda Rex-Babar-169 hibridou tanto em SH quanto em DL, apresentando-se como blocos teloméricos na primeira e com distribuição dispersa na

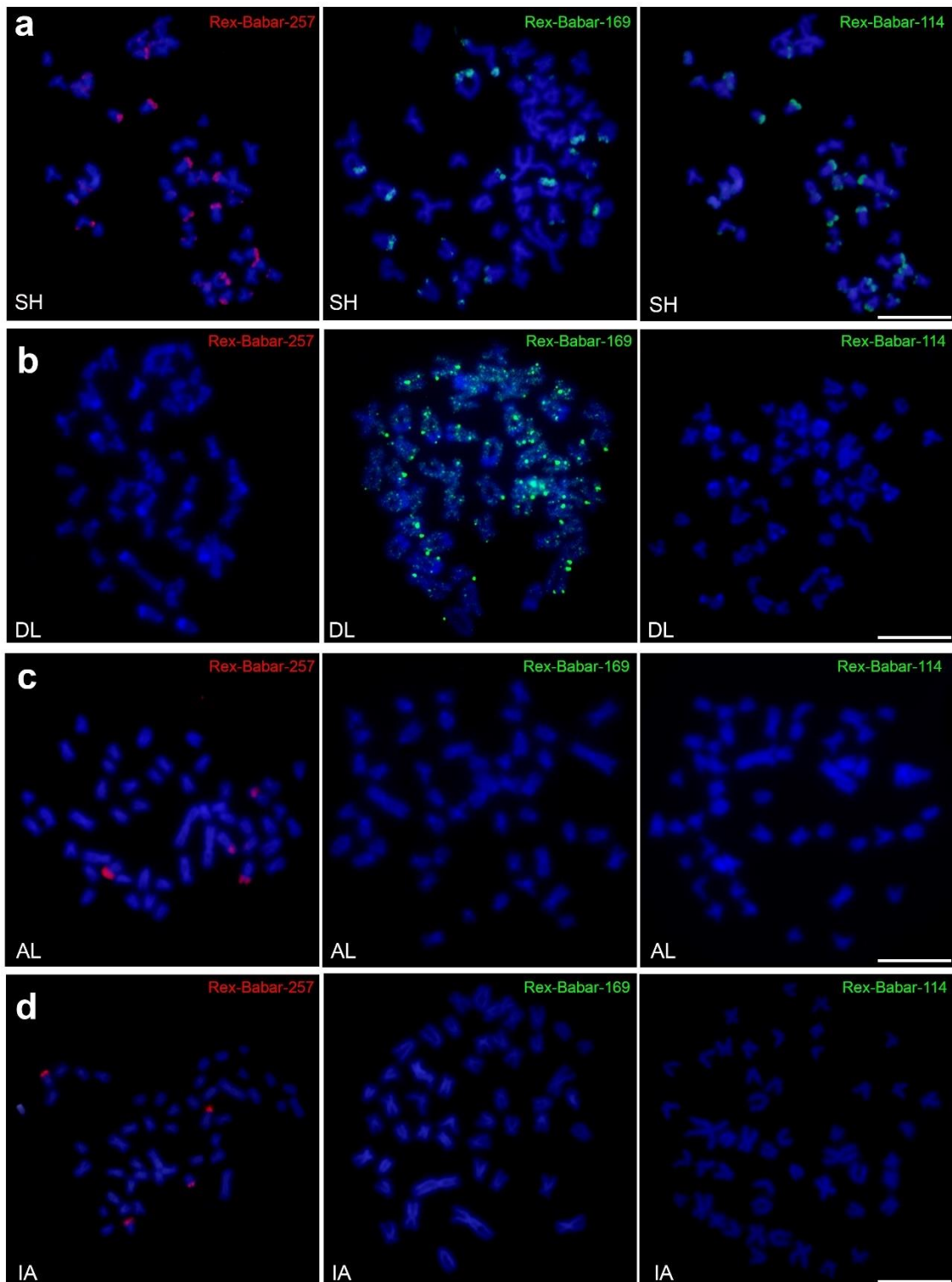
segunda. O elemento babar é pouco estudado em peixes neotropicais; este é o primeiro mapeamento dessas sequências em diferentes populações de *P. fasciatus*.

Figura 10 - Distribuição cromossômica dos retroelementos L1 (83 e 108) e L2 (90) em quatro populações de *P. fasciatus*



**A** - Furna 2 (SH). **B** - Lagoa Dourada (DL). **C** - Alagados (AL); e **D** - Rio Iapó (IA). Barra de escala = 10  $\mu$ m.  
Fonte: A autora

Figura 11 - Distribuição cromossômica dos retroelementos *Rex-Babar* em quatro populações de *P. fasciatus*.

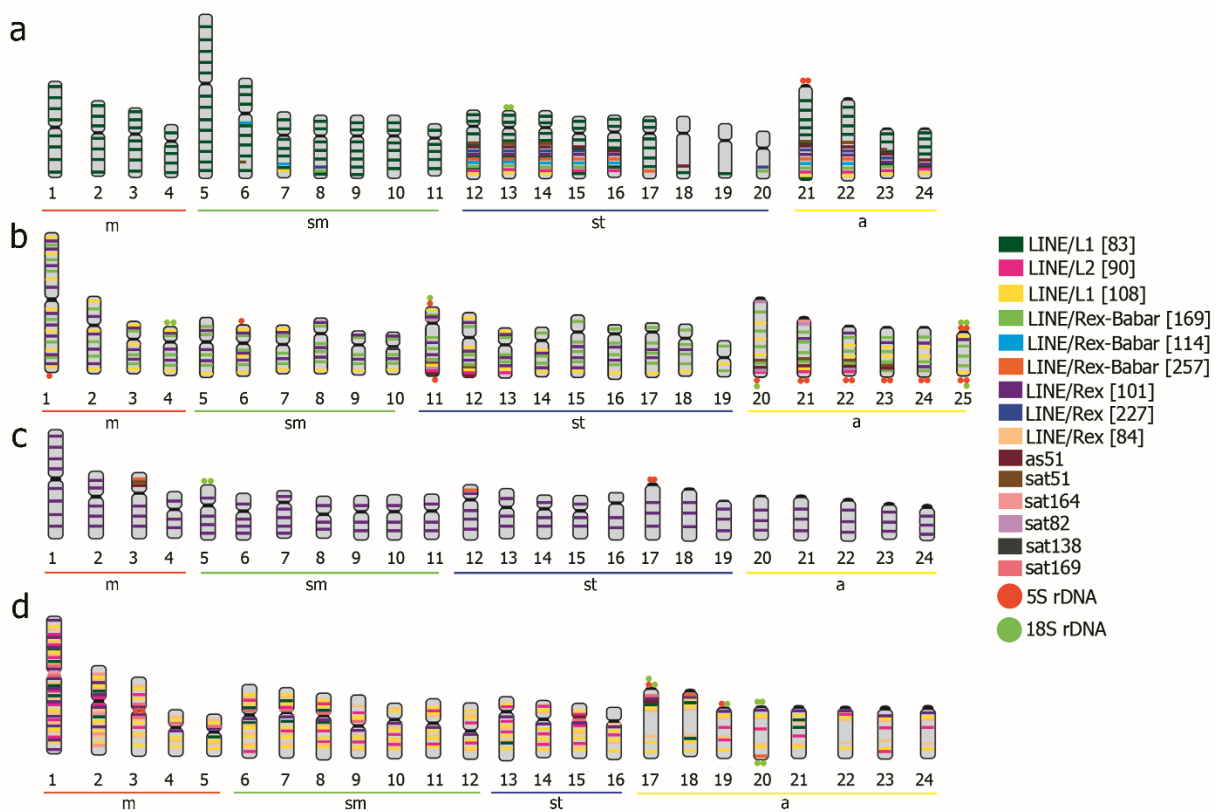


**A** - Furna 2 (SH). **B** - Lagoa Dourada (DL). **C** - Alagados (AL); e **D** - Rio Iapó (IA). Barra de escala = 10  $\mu$ m.  
Fonte: A autora

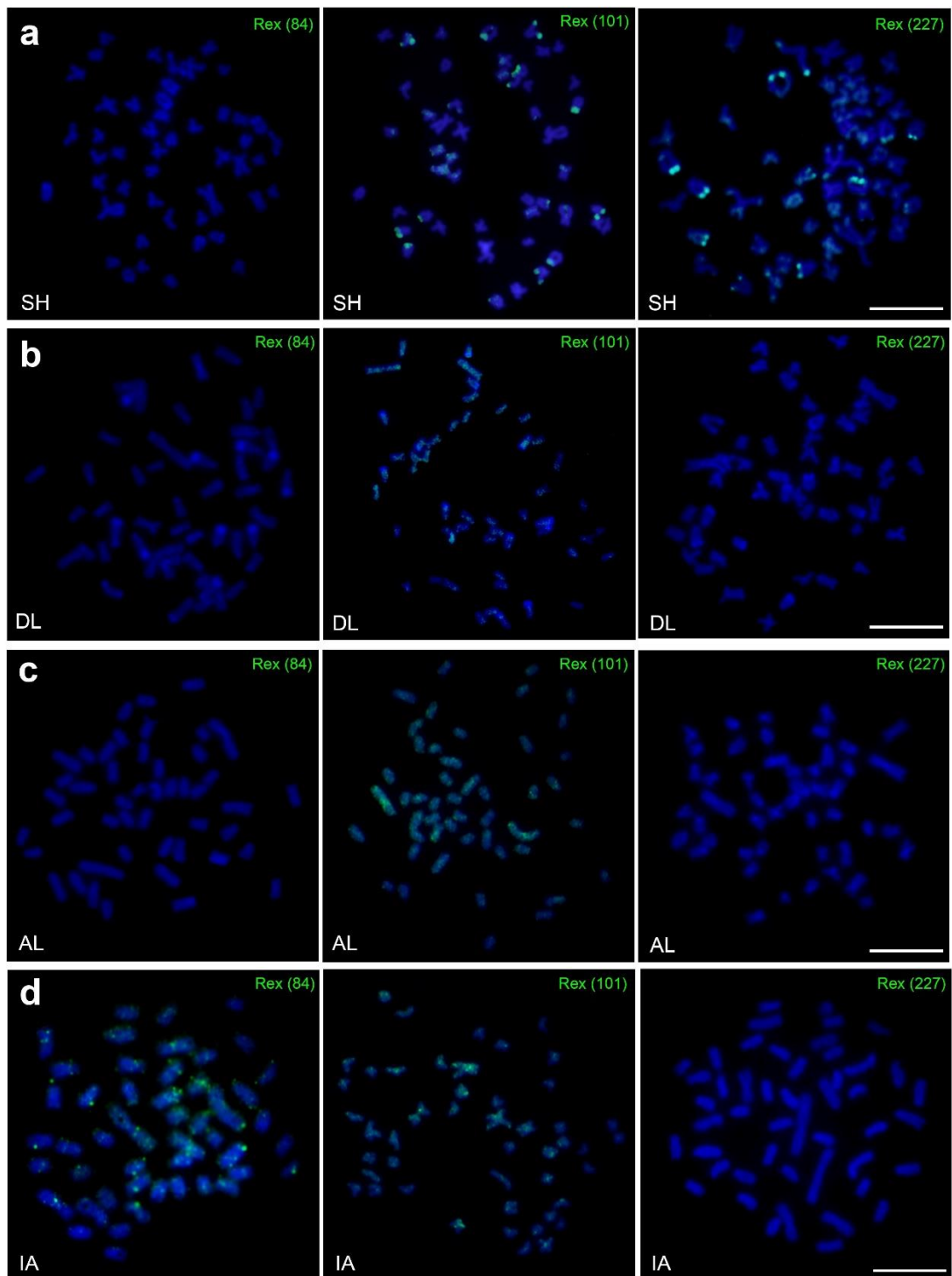
#### 4.4.2.3 Elementos Rex

Nosso estudo identificou diversas sequências com alta similaridade aos genes *Rex*. Dentre os elementos *Rex* analisados em peixes, *Rex1*, *Rex3* e *Rex6* são os mais comumente estudados; no entanto, o conhecimento sobre esses elementos ainda é limitado devido à escassez de pesquisas na área. No presente estudo, apenas o retrotransposon *Rex-101* hibridou em todas as populações, apresentando um padrão de distribuição dispersa em todas, exceto na população SH (Figura 13a). O elemento *Rex -84* parece ser exclusivo da população IA (Figura 5d), enquanto *Rex -227* exibiu um padrão de marcação em blocos do tipo telomérico em pelo menos sete pares cromossômicos, principalmente submetalocêntricos e telocêntricos. Notavelmente, essa sequência *Rex-227* também foi exclusiva da população SH.

Figura 12 – Ideograma cromossômico das diferentes populações de *P. fasciatus* estudadas



**A** - Furna 2. **B** - Lagoa Dourada. **C** - Alagados. **D** - População do Iapó. Cromossomos m=metacêntricos, sm=submetacêntricos, st=subtelocêntricos e a=acrocêntricos. Fonte: A autora

Figura 13 - Mapeamento dos retroelementos Rex em *P. fasciatus*

A - Furna 2 (SH). B - Lagoa Dourada (DL). C - Alagados (AL); e D - Rio Iapó (IA). Barra de escala = 10  $\mu$ m.  
Fonte: A autora

### 4.4.3 Mapeamento de DNAs Satélites (SatDNA)

#### 4.4.3.1 Análise comparativa de satDNAs pelo RepeatExplorer (RE)

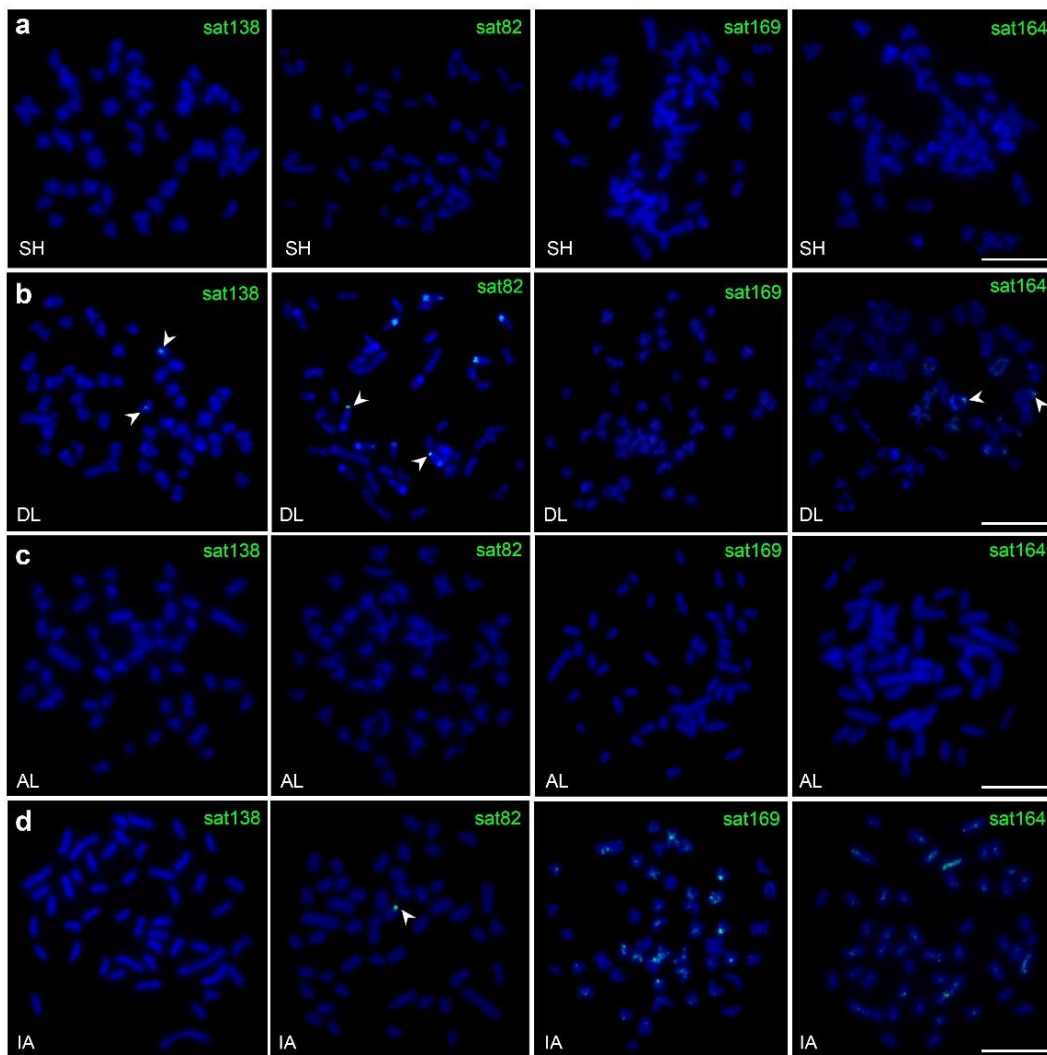
A análise comparativa dos DNAs satélites revelou resolução limitada para a identificação de marcadores exclusivos na população SH, embora tenham aparecido na análise *in silico* (Figura 14, Anexo F). A ausência de sinais de hibridação é, por si só, relevante, uma vez que apenas a população DL apresentou sinais proeminentes e aparente exclusividade para os satDNAs *sat138* e *sat82* (Figura 14b). Nenhum sinal de hibridação foi detectado na população AL para nenhuma das sondas de satélite analisadas (Figura 14c). O elemento *sat169* apresentou hibridação exclusiva na população IA, com localização nas regiões centroméricas e em domínios intersticiais específicos de cromossomos metacêntricos e subtelocêntricos. O satélite *sat164* foi observado na região centromérica de um único par de cromossomos acrocêntricos em DL e apresentou padrão disperso em IA, com marcação centromérica proeminente (Figura 14b, d).

#### *Hibridação de as51 e sat51*

Sinais de hibridação em blocos do tipo telomérico foram detectados para ambos os DNAs satélites (*As51* e *sat51*) nas populações SH e DL, indicando um alto número de cópias desses elementos repetitivos (Figura 15a, b), este resultado também foi encontrado por Mantovani *et al.* (2004) e por Abel; Mantovani; Moreira-filho (2006) em *Astyanax scabripinnis*. A variante *sat51* foi observada exclusivamente nessas duas populações, apresentando de forma consistente um padrão de marcação em blocos teloméricos. Essa distribuição é coerente com o padrão de abundância identificado na análise comparativa pelo RE entre SH vs DL.

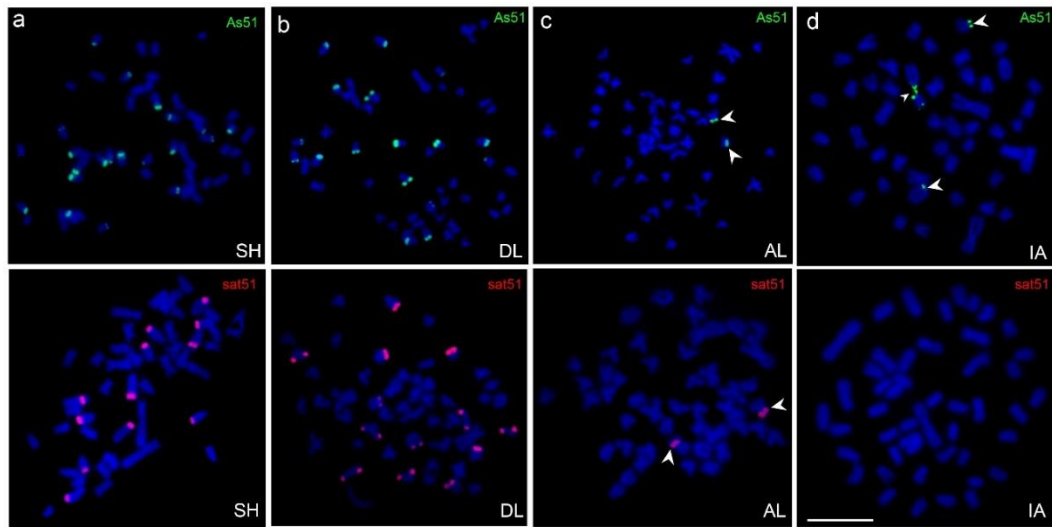
Na população AL, um par de cromossomos submetacêntricos exibiu sinais de hibridação para ambas as sondas, com localização claramente pericentromérica em ambos os casos (Figura 15c). Em contraste, na população IA, *as51* hibridou a apenas dois pares cromossômicos, enquanto *sat51* não apresentou sinal detectável (Figura 15d), sugerindo a ausência ou redução significativa dessa variante nessa população.

Figura 14 - Distribuição cromossômica das DNAs satélites a partir da análise comparativa de retroelementos (RE)



A - Furna 2 (SH). B - Lagoa Dourada (DL). C - Alagados (AL); e D - Rio Iapó (IA). Barra de escala = 10  $\mu$ m.  
Fonte: A autora

Figura 15 - Distribuição cromossômica dos DNAs satélites *As51* (verde) e *sat51* (vermelho) em quatro populações de *P. fasciatus*



**A** - Furna 2 (SH). **B** - Lagoa Dourada (DL). **C** - Alagados (AL); e **D** - Rio Iapó (IA). Barra de escala = 10  $\mu$ m.  
Fonte: A autora

#### 4.5 DISCUSSÃO

Elementos de DNA repetitivo surgem como arquitetos dinâmicos das mudanças microevolutivas em *Psalidodon fasciatus*, moldando a divergência genômica entre populações isoladas e populações de ambientes abertos, por meio de interações complexas entre deriva genética, seleção e rearranjos estruturais (Carotti *et al.*, 2021; Ugarkovic; Plohl, 2002). Nosso estudo sugere que a população da furna 2 (SH) possui uma assinatura genômica distinta, caracterizada pela acumulação exclusiva dos retroelementos *Rex-Babar-169* e *Rex-227*, juntamente com LINEs enriquecidos em regiões teloméricas (Figura 12a). Esse padrão reflete observações feitas em *Astyanax mexicanus* adaptado a cavernas, nas quais elementos transponíveis mediam rearranjos cromossômicos como forma de conter instabilidades genômicas em ambientes isolados (Shao *et al.*, 2019). A fixação desses TEs provavelmente reflete a intensificação da deriva genética em demes de pequeno tamanho populacional, nos quais o fluxo gênico reduzido limita forças homogeneizadoras e acelera expansões específicas (Bourgeois; Boissinot, 2019). A localização telomérica dos LINEs em SH sugere uma possível exaptação para manutenção telomérica, fenômeno já documentado em mamíferos, mas raramente observado em teleósteos, indicando soluções convergentes para o estresse genômico em habitats fragmentados.

Os DNAs satélites (satDNAs) também evidenciam dinâmicas microevolutivas que operam em escalas de tempo notavelmente rápidas (Ruiz-Ruano *et al.*, 2016). A retenção exclusiva dos elementos *As51* e *sat51* nas populações SH e DL, especificamente na região

telomérica, pode estar relacionada à amplificação localizada mediada por recombinação desigual. Novas variantes de sequências satélites, como o *sat51* identificado neste estudo, podem surgir por meio de mutações pontuais ou inserções de elementos transponíveis (TEs), e posteriormente expandir-se em matrizes mais longas por meio de deslizamento de replicação ou recombinação desigual (Flynn; Yamashita, 2024). A ocorrência quase exclusiva desses TEs na população SH pode explicar a coexistência observada entre esses elementos.

O *sat169* ocorre exclusivamente na população IA refletindo os ciclos de nascimento e morte típicos da evolução de satDNAs (Plohl *et al.*, 2012). Essa renovação específica por linhagem reforça o potencial dos satDNAs como marcadores de alta resolução para processos micro evolutivos (Thakur *et al.*, 2021), capazes de detectar divergências antes mesmo da manifestação de diferenciações morfológicas.

A reestruturação dos sítios de rDNA fornece evidências adicionais de divergência micro evolutiva incipiente. Enquanto a localização dos rDNAs 18S/5S, a população SH sugere um estado de conservação (Figura 9a), já que esta marcação (um par de cromossomos) tem sido encontrada em outras populações, inclusive em algumas descritas para a bacia de São Francisco, a qual é definida como a população tipo para *P. fasciatus* (Pazza *et al.*, 2006). Em teleósteos, rearranjos de rDNA estão frequentemente associados à incompatibilidade híbrida posicionando essas alterações como possíveis barreiras genômicas em processos iniciais de especiação (Merlo *et al.*, 2013). No caso da população DL, as múltiplas regiões 5S coincidem com regiões contendo sequências de retro elementos como L1(90), esta co-localização tem sido reportada em algumas espécies de peixe (Rebordinos; Cross; Merlo, 2013; Cioffi *et al.*, 2010).

O habitat de fumaça atua, assim, como um cadinho microevolutivo: o isolamento molda perfis de DNA repetitivo por meio da expansão de heterocromatina, que atua como um amortecedor frente a choques genômicos (Flynn; Yamashita, 2024). Neste último caso, destaque-se a importância na geração de variabilidade por recombinação cromossômica em resposta a erosão genética causada por alguns fatores como endogamia, assim como o caso da população SH.

Apesar das limitações da abordagem de sequenciamento de baixa cobertura, nossos resultados destacam a necessidade urgente de metodologias mais robustas como sequenciamento de cadeias longas ou metodologias mais recentes como CELLO-seq para quantificar frequências de inserção de TEs (Marlow; Deaville; Berrens, 2025). Estudos futuros devem testar se os TEs específicos de SH regulam genes associados à resposta ao estresse.

#### 4.6 CONCLUSÃO

Em conclusão, os elementos repetitivos em *P. fasciatus* transcendem seu papel como ocupantes passivos do genoma e se estabelecem como agentes ativos da microevolução. Populações isoladas acumulam perfis distintos de TEs e satDNAs por meio de processos estocásticos e seletivos, gerando divergências genômicas mensuráveis em escalas de tempo ecológicas no caso de *P. fasciatus* da Furna do Parque estadual de Vila Velha, um isolamento de cerca de 1,5 milhões de anos (De la Ossa-Guerra *et al*, Segundo Capítulo desta tese). Este trabalho posiciona o DNA repetitivo como uma ferramenta essencial para compreender como os genomas se adaptam, em tempo real, à fragmentação de habitats, uma questão urgente em face das rápidas transformações ambientais atuais.

## 5 REFERENCIAS

- ABE, K. T. *et al.* Systematic and historical biogeography of the Bryconidae (Ostariophysi: Characiformes) suggesting a new rearrangement of its genera and an old origin of Mesoamerican ichthyofauna. **BMC Evolutionary Biology**, v. 14, n. 1, p. 1–15, 2014.
- ABEL, D. L. S.; MANTOVANI, M.; MOREIRA-FILHO, O. Chromosomal distribution of the As51 satellite DNA in two species complexes of the genus *Astyanax* (Pisces, Characidae). **Genetics and Molecular Biology**, v. 29, n. 3, p. 448-452, 2006.
- ANDREWS, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. Disponivel online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>, 2010.
- ARGENTINA, J. E. *et al.* Spatial extent of analysis influences observed patterns of population genetic structure in a widespread darter species (Percidae). **Freshwater Biology**, v. 63, n.10, p. 1185–1198, 2018.
- ARTONI, R. F. *et al.* *Astyanax* aff. *fasciatus* Cuvier, 1819 (Teleostei; Characidae): evidences of a species complex in the upper rio Tibagi basin (Paraná, Brazil). **Neotropical Ichthyology** v. 4. 2006.
- ARTONI, R. F.; ALMEIDA MATIELLO M. C. Genética de Peixes Neotropicais. I. Aspectos da Conservação Genética dos Peixes no Parque Estadual de Vila Velha, Paraná, Brasil. **Publ. UEPG Ci. Biol. Saúde**, v. 9, n. 2, p. 7–15, 2003.
- ARTONI, R. F.; SHIBATA, O. A. Catálogo Ilustrado dos Peixes do Parque Estadual de Vila Velha. In: ARTONI, R. F; SHIBATA, O. A. (org.). **Peixes do Parque Estadual de Vila Velha: Aspectos da História Natural, da Biologia Evolutiva e da Conservação**. Ponta Grossa: Editora UEPG, 2006, 92–137.
- ASSINE, M. L. Aspectos de Estratigrafia das sequencias pré-carboníferas da bacia do Paraná no Brasil. **Universidade de São Paulo**, 1996.
- BALLOUX, F. Heterozygote Excess in Small Populations and the Heterozygote-Excess Effective Population Size. **Evolution**. v. 58, n. 9, p. 1891-900, 2016.
- BENG, K. C.; CORLETT, R. T. Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges, and prospects. **Biodiversity and Conservation** v. 29, p. 2089–2121, 2020.
- BENGSTON, S. E. *et al.* Genomic tools for behavioral ecologists to understand repeatable individual differences in behaviour. **Nat Ecol Evol**, v. 2, p. 944–955, 2018.
- BENNEMANN, S. T.; SILVA-SOUZA, A. T.; ROCHA. G. R. A. Composición Ictiofaunística en Cinco Localidades de la Cuenca del Río Tibagi, PR. **Interciencia**, v. 20, n. 1, p. 7–13, 1995.
- BERGERON, L. A. *et al.* Evolution of the germline mutation rate across vertebrates. **Nature**, v. 615, p. 285–291, 2023.
- BLACKMAN, R. *et al.* Environmental DNA: The next chapter. **Molecular ecology**, v. 33, n. 11, e17355, 2024.

- BOCKMANN, F. A.; CASTRO, R. M. C. The Blind Catfish from the Caves of Chapada Diamantina, Bahia, Brazil (Siluriformes: Heptapteridae): Description, Anatomy, Phylogenetic Relationships, Natural History, and Biogeography. **Neotropical Ichthyology**, v. 8, n. 4, p. 625–642, 2010.
- BOLGER, A. M.; LOHSE, M.; USADEL B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, p. 2114, 2014.
- BOUCKAERT, R. *et al.* BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. **PLoS Comput Biol**, v. 15, p. 1–28, 2019.
- BOURGEOIS, Y.; BOISSINOT, S. On the Population Dynamics of Junk: A Review on the Population Genomics of Transposable Elements. **Genes (Basel)**, v. 10, n. 6, 2019.
- BOYD, S. H. *et al.* Using environmental DNA methods to survey for rare groundwater fauna: Detection of an endangered endemic cave crayfish in northern Alabama. **PLoS One**, v. 15, n. 12, e0242741, 2020.
- BREITBART, M. *et al.* Evaluation of DNA Metabarcoding for Identifying Fish Eggs: A Case Study on the West Florida Shelf. **PeerJ**, v. 11, p. e15016, 2023.
- BROMHAM, L.; PENNY, D. The modern molecular clock. **Nature Reviews Genetics**, v. 4, p. 216–224, 2003.
- CALLAHAN, B. J. *et al.* DADA2: High-Resolution Sample Inference from Illumina Amplicon Data. **Nature Methods**, v. 13, n. 7, p. 581–583, 2016.
- CAMACHO, C. *et al.* BLAST+: Architecture and Applications. **BMC Bioinformatics**, v. 10, n. 1, p. 421, 2009.
- CAROTTI, E. *et al.* Transposable elements and teleost migratory behaviour. **International Journal of Molecular Sciences**, v. 22, n. 2, p. 1–12, 2021.
- CASTRO, J. P. *et al.* Diversity of the *Astyanax scabripinnis* species complex (Teleostei: Characidae) in the Atlantic Forest, Brazil: species limits and evolutionary inferences. **Reviews in Fish Biology and Fisheries**, v. 25, p. 231–244, 2014a.
- CASTRO, J. P. *et al.* Evidence of incipient speciation in *Astyanax scabripinnis* species complex (Teleostei: Characidae). **Neotropical Ichthyology**, v. 12, p. 429–438, 2014b.
- CATCHEN, J. *et al.* Stacks: an analysis tool set for population genomics. **Molecular Ecology**, v. 22, p. 3124–3140, 2013.
- CIOFFI, M. B.; MARTINS, C.; BERTOLLO, L. A. C. Chromosome spreading of associated transposable elements and ribosomal DNA in the fish *Erythrinus erythrinus*. Implications for genome change and karyoevolution in fish. **BMC Evol Biol**, v. 10, p. 271–279, 2010.
- CLARO-GARCÍA, A.; ASSEGA, F.M.; SHIBATTA, O. A. Diversity and distribution of ichthyofauna in streams of the middle and lower tibagi river basin, Paraná, Brazil. **Check List**, v. 14, p. 43–53, 2018.

- COLEMAN, R. A. *et al.* Artificial barriers prevent genetic recovery of small isolated populations of a low-mobility freshwater fish. **Heredity (Edinb)**, v. 120, p. 515–532, 2018.
- COLLYER, M. L. *et al.* Habitat-morphotype associations of pecos pupfish (*Cyprinodon pecosensis*) in isolated habitat complexes. **Copeia**, v. 103, n. 1, 181–199, 2015.
- COMBRINK, L. L. *et al.* Population Genomics of Adaptive Radiation. **Molecular ecology**, v. 34, n. 2, 2025.
- CORE NUCLEOTIDE [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004 – [cited 2025, February, 28]. Available from: <https://www.ncbi.nlm.nih.gov/nucleotide/>
- CRANDALL, K. A. *et al.* Considering evolutionary processes in Conservation Biology. **TREE**, v. 15, n. 7, p. 290–295, 2000.
- CUVIER, G. Sur les poissons du sous-genre Hydrocyn [sic], sur deux nouvelles espèces de Chalceus, sur trois nouvelles espèces du Serrasalmes, et sur l'Argentina glossodonta de Forkahl, qui est l'Albula gonorhynchus de Bloch. **Mémoires du Muséum National d'Histoire Naturelle, Paris (N S) (Série A) Zoologie**, v. 5, p. 351–379, 1819.
- DANECEK P. *et al.* The variant call format and VCFtools. **Bioinformatics**, v. 27, p. 2156–2158, 2011.
- DE MELO, M. S. *et al.* Carste em rochas não-carbonáticas: o exemplo dos arenitos da formação furnas, campos gerais do Paraná/Brasil e as implicações para a região. **Espeleo-Tema** 221: 81–97, 2011.
- DE MELO, M. S. *et al.* A geologia no plano de manejo do Parque Estadual de Vila Velha, PR, **Revista Brasileira de Geociências**, v. 34, n. 4, p. 561–570, 2004.
- DE MELO, M. S.; FONSECA, G. P.C.; RUIZ, P. L.C. Gênese e evolução da lagoa Dourada, Ponta Grossa, PR. **Revista do Instituto Geológico**, v. 21, p. 17–31, 2000.
- DE MELO, M. S.; GIANNINI, P. C. F. Sandstone Dissolution Landforms in the Furnas Formation, Southern Brazil. **Earth Surface Processes and Landforms**, v. 32, p. 2149–2164, 2007.
- DOS REIS, R. B. *et al.* Freshwater Fishes from Paraná State, Brazil: An Annotated List, with Comments on Biogeographic Patterns, Threats, and Future Perspectives. **Zootaxa**, v. 4868, n. 4, p. 451–494, 2020.
- DOS SANTOS, R. Z. *et al.* A Long-Term Conserved Satellite DNA That Remains Unexpanded in Several Genomes of Characiformes Fish Is Actively Transcribed. **Genome Biology and Evolution**, v. 13, n. 2, 2021.
- FARRELL, E. D.; CARLSSON, J.E.L.; CARLSSON, J. Next gen pop gen: Implementing a high-throughput approach to population genetics in boarfish (*Capros aper*). **Royal Society Open Science**, v. 3, n. 12, 2016.
- FERNÁNDEZ, S. *et al.* Evaluating Freshwater Macroinvertebrates from eDNA Metabarcoding: A River Nalón Case Study. **PLoS ONE**, v. 13, n. 8, e0201741, 2018.

- FERREIRA-NETO, M. *et al.* Three sympatric karyomorphs in the fish *Astyanax fasciatus* (Teleostei, Characidae) do not seem to hybridize in natural populations. **Comparative Cytogenetics**, v. 6, n. 1, p. 29–40, 2012.
- FLYNN, J. M.; YAMASHITA, Y. M. The implications of satellite DNA instability on cellular function and evolution. **Seminars in Cell and Developmental Biology**, v. 156, pp. 152–159, 2024.
- FRANÇA, A. B. *et al.* Arenitos Lapa-Vila Velha: Um Modelo De Trato De Sistemas Subaquosos Canal-Lobos Sob Influência Glacial, Grupo Itararé (C-P), Bacia Do Paraná. **Revista Brasileira de Geociências**, v. 26, n. 1, p. 43–56, 1996.
- FRANKHAM, R. Challenges and opportunities of genetic approaches to biological conservation. **Biological Conservation**, v. 143, n. 9, p. 1919–1927, 2010.
- FRICHOT, E.; FRANÇOIS, O. LEA: An R package for landscape and ecological association studies. **Methods in Ecology and Evolution**, v. 6, n. 8, p. 925–929, 2015.
- FUNK, W. C. *et al.* Harnessing genomics for delineating conservation units. **Trends in Ecology and Evolution**, v. 27, n. 9, p. 489–496, 2012.
- GALVES, W.; SHIBATTA, O. A.; JEREP, F. C. Fish, Taquara River Basin, Northern of the State of Paraná, Brazil. **Check List**, v. 3, n. 3, p. 253–259, 2007.
- GARLA, A. *et al.* A Ictiologia do Rio Tibagi-PR: Uma Análise Cienciométrica dos Últimos 10 Anos. **Revista Ciência, Tecnologia e Ambiente**, v. 13, e13247, 2023.
- GAVAZZONI, M. *et al.* Detection of natural hybridization and delimitation of two closely related operational taxonomic units of the *Astyanax fasciatus* (Teleostei: Characidae) complex through integrative approaches. **Biological Journal of the Linnean Society**, v. 129, n. 3, p. 687–700, 2020.
- GEALH, A. M. Ictiofauna dos Rios Fortaleza, Iapó, Alto e Médio Tibagi. In *Patrimônio Natural dos Campos Gerais do Paraná*, edited by M. S. de Melo, R. S. Moro, and G. B. Guimarães, 151–162. Ponta Grossa: **Editora UEPG**, 2007.
- GIBERT, J. *et al.* Basic attributes of groundwater ecosystems and prospects for research. In: Danielopol, D.L., Stanford JA (ed) **Groundwater Ecology**, Academic Press: San Diego, CA, USA, p. 7–40, 1994.
- GLEASON, J. E.; HANNER, R. H.; COTTENIE, K. Hidden Diversity: DNA Metabarcoding Reveals Hyper-Diverse Benthic Invertebrate Communities. **BMC Ecology and Evolution**, v. 23, n. 19, 2023.
- GOES, C. A. G. *et al.* Revealing the Satellite DNA History in *Psalidodon* and *Astyanax* Characid Fish by Comparative Satellitomics. **Frontiers in Genetics**, v. 13, 2022.
- GROSS, M. C. *et al.* Population structure, fluctuating asymmetry and genetic variability in an endemic and highly isolated *Astyanax* fish population (Characidae). **Genetics and Molecular Biology**, v. 27, n. 4, p. 529–535, 2004.

- GURGEL, B. H.C. Estrutura populacional e época de reprodução de *Astyanax fasciatus* (Cuvier) (Characidae, Tetragonopterinae) do Rio Ceará Mirim, Poço Branco, Rio Grande do Norte, Brasil. **Revista Brasileira de Zoologia**, v. 21, n. 1, p. 131–135, 2004.
- HASHEMZADEH, S. I. *et al.* eDNA metabarcoding as a means to assess distribution of subterranean fish communities: Iranian blind cave fishes as a case study. **Environmental DNA**, v. 4, n. 2, p. 402–416, 2022.
- HASTINGS, K. Long-term persistence of isolated fish populations in the Alexander Archipelago. **Graduate Student Theses, Dissertations, & Professional Papers**. 9560, 2005.
- HELLMAIR, M.; KINZIGER, A. P. Increased extinction potential of insular fish populations with reduced life history variation and low genetic diversity. **PLoS One**, v. 9, n. 11, 2014.
- HERDER, J. *et al.* Environmental DNA—A Review of the Possible Applications for the Detection of (Invasive) Species. **Stichting RAVON**, Nijmegen. Report 2013-104, 2014.
- HILÁRIO, O. H. BLASr: An R package for parallel taxonomic classification of metabarcoding sequences. Available at GitHub: <https://github.com/heronoh/BLASr>, 2024.
- HILÁRIO, H. O. *et al.* DNA metabarcoding of mock communities highlights potential biases when assessing Neotropical fish diversity. **Environmental DNA**, v. 5, n. 6, p. 1351–1361, 2023.
- HIPSLEY, C. A.; MÜLLER, J. Beyond fossil calibrations: Realities of molecular clock practices in evolutionary biology. **Frontiers in Genetics**, v. 5, p. 1–11, 2014.
- HIRT, L. M.; ARAYA, P. R.; FLORES, S. A. Estrutura populacional, biologia reprodutiva e alimentação de *Astyanax fasciatus* (Cuvier, 1819) em um tributário do alto rio paran, Misiones, Argentina. **Acta Limnologica Brasiliensia**, v. 23, p. 1–12, 2011.
- HO, S. Y. W.; DUCHÊNE, S. Molecular-clock methods for estimating evolutionary rates and timescales. **Molecular Ecology**, v. 23, p. 5947–5965, 2014.
- HOFFMANN, A. C.; NASCIMENTO, C.; SHIBATTA, O. A. Fish Fauna from Tributaries throughout the Tibagi River Basin, Upper Paran Basin, Brazil. **Check List**, v. 11, n. 6, p. 1815, 2015.
- HOKE, K.; SCHWARTZ, A.; SOARES, D. Evolution of the fast start response in the cavefish *Astyanax mexicanus*. **Behavioral Ecology and Sociobiology**, v. 66, n. 8, p. 1157–1164, 2012.
- HOLZ, M. *et al.* A stratigraphic chart of the Late Carboniferous/Permian succession of the eastern border of the Paran Basin, Brazil, South America. **Journal of South American Earth Sciences**, v. 29, n. 2, p. 381–399, 2010.
- IWASAKI, W. *et al.* MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. **Molecular Biology and Evolution**, v. 30, n. 11, p. 2531–2540, 2013.
- JACKMAN, J. M. *et al.* eDNA in a Bottleneck: Obstacles to Fish Metabarcoding Studies in Megadiverse Freshwater Systems. **Environmental DNA**, v. 3, p. 837–849, 2021.

- JAILLON, O., *et al.* Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. **Nature**, v. 431, p. 946-957, 2004
- JAISUK, C.; SENANAN, W. Effects of landscape features on population genetic variation of a tropical stream fish, Stone lapping minnow, *Garra cambodgiensis*, in the upper Nan River drainage basin, northern Thailand. **PeerJ**, 6: e4487, 2018.
- JARDULI, L. R., *et al.* Fish Fauna from the Paranapanema River Basin, Brazil. **Biota Neotropica**, v. 20, 1: e20180707, 2019.
- JIN, J.J. *et al.* GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. **Genome Biology**, v. 21, p. 1–31, 2020.
- JURKA, J. *et al.* Repbase Update, a database of eukaryotic repetitive elements. **Cytogenetic and Genome Research**, v. 110, p. 462–467, 2005.
- JURKA, J. Repbase Update: a database and an electronic journal of repetitive elements. **Trends Genet**, v. 16, p. 418–420, 2000.
- KAVALCO, K. F., *et al.* Chromosomal Diversification Higher Than Molecular Variation in *Astyanax aff. fasciatus* (Teleostei, Characidae). **Zebrafish**, v. 13, n. 4, p. 345–353. 2016.
- KERNISKE, F. F. *et al.* Reproductive Aspects of an Isolated Population of *Psalidodon aff. fasciatus* (Teleostei: Characidae) from Southern Brazil. **Acta Limnologica Brasiliensia**, v. 35: e24, 2023.
- KERNISKE, F. F. *et al.* Spinal malformations in a naturally isolated Neotropical fish population. **PeerJ**, v. 9, 2021.
- KIRCHNER, S. *et al.* Investigating gene flow between the blind cavefish *Garra barreimiae* and its conspecific surface populations. **Scientific Reports**, v. 7, n. 5130, 2017.
- KÖRESSAAR, T. *et al.* Primer3-masker: Integrating masking of template sequence with primer design software. **Bioinformatics**, v. 34, n. 11, p. 1937–1938, 2018.
- LANÉS, L. E. K.; KEPPELER, F. W.; MALTCHIK, L. Abundance variations and life history traits of two sympatric species of Neotropical annual fish (Cyprinodontiformes: Rivulidae) in temporary ponds of southern Brazil. **Journal of Natural History**, v. 48, n. 31–32, p. 1971–1988, 2014.
- LEVAN, A.; FREDGA, K.; SANDBERG, A. A. Nomenclature for centromeric position on chromosomes. **Hereditas**, v. 52, p. 201–220, 1964.
- LIU, X.; FU, Y. X. Stairway Plot 2: demographic history inference with folded SNP frequency spectra. **Genome Biology**, v. 21. 2020.
- LOMOLINO, M. V. Conservation biogeography. In Lomolino M.V. & Heaney L.R. (Eds.). *Frontiers of biogeography: new directions in the geography of nature*. **Sunderland, Massachusetts: Sinauer Associates**; p. 293–296, 2004.
- MAACK, R. Geografia física do Estado de Paraná (4th ed.). **Editores UEPG**, 1969.

- MAHÉ, F. *et al.* Swarm: Robust and Fast Clustering Method for Amplicon-Based Studies. **PeerJ**, v. 3: e1420, 2015.
- MANTOVANI, M. *et al.* Evidence of the differentiated structural arrangement of constitutive heterochromatin between two populations of *Astyanax scabripinnis* (Pisces, Characidae). **Genetics and Molecular Biology**, v. 27, n. 4, p. 536-542, 2004.
- MARLOW, S. A.; DEAVILLE, L. A.; BERRENS, R. V. Long-read RNA sequencing of transposable elements from single cells using CELLO-seq. **Nature Protocols**, 2025.
- MARTIN, M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. **EMB net.journal**, v. 17, n. 1, p. 10–12, 2011.
- MATOSO, D. A. *et al.* Molecular taxonomy and evolutionary hypothesis concerning *Astyanax fasciatus* (Characiformes, Characidae) from Vila Velha State Park and Tibagi and Iguaçú Rivers. **Genetics and Molecular Research**, v. 12, p. 631–638, 2013.
- MATOSO, D. A. *et al.* Preliminary Qualitative Analysis on mtDNA in *Astyanax fasciatus* Populations Cuvier, 1819 (Teleostei; Characidae) Indicate Population Distinctiveness. **Brazilian Archives of Biology and Technology**, v. 53, n. 3, p. 663–667, 2010.
- MATOSO, D. A.; ARTONI, R. F.; GALETTI, P. M. Genetic Diversity of the Small Characid Fish *Astyanax sp.*, and Its Significance for Conservation. **Hydrobiologia**, v. 527, p. 223–225, 2004.
- MATTOX, G. M. T. *et al.* Surface and subterranean ichthyofauna in the Serra do Ramalho karst area, northeastern Brazil, with updated lists of Brazilian troglobitic and troglophilic fishes. **Biota Neotropica**, v. 8, p. 145–152, 2008.
- MCDONOUGH, Y.; CONNALLON, T. Effects of population size change on the genetics of adaptation following an abrupt change in environment. **Evolution**, v. 77, p. 1852–1863, 2023.
- MCMAHON, B. J.; TEELING, E. C.; HÖGLUND, J. How and why should we implement genomics into conservation? **Evolutionary Applications**, v. 7, n. 9, p. 999–1007, 2014.
- MCMURDIE, P. J.; HOLMES, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. **PLoS ONE**, v. 8, n. 4: e61217, 2013.
- MELLO, R. *et al.* Genetic and biogeographical relationships among species of *Astyanax* (Teleostei, Characidae) in Brazilian river basins. **Genetics and Molecular Research**, v. 14, p. 15356–15364, 2015.
- MELO, B. F. *et al.* Accelerated Diversification Explains the Exceptional Species Richness of Tropical Characoid Fishes. **Systematic Biology**, v. 71, p. 78–92, 2022.
- MELO, B. F. *et al.* Phylogenomics of Characidae, a hyper-diverse Neotropical freshwater fish lineage, with a phylogenetic classification including four families (Teleostei: Characiformes). **Zoological Journal of the Linnean Society**, v. 202, n. 1, 2024.
- MELO, M. S. *et al.* Carste em rochas não-carbonáticas: o exemplo dos arenitos da formação furnas, campos gerais do Paraná/Brasil e as implicações para a região. **Espeleo-Tema**, v. 221, p. 81–97, 2011.

- MENDES, G. A. *et al.* eDNA Extraction from MCE Filter and Sterivex Cartridge. **Protocols.io**, 2024b.
- MENDES, G. A. *et al.* Protocolo de Coleta de Amostras para eDNA Metabarcoding Usando Filtros Sterivex. **Protocols.io**, 2024a.
- MENDES, I. S. *et al.* On the evolutionary origin of Neotropical cavefish *Ancistrus cryptophthalmus* (Siluriformes, Loricariidae) based on the mitogenome and genetic structure of cave and surface populations. **Hydrobiologia**, v. 842, p. 157–171, 2019.
- MERLO, M. A. *et al.* The 5S rDNA high dynamism in *Diplodus sargus* is a transposon-mediated mechanism. Comparison with other multigene families and *sparidae* species. **Journal of Molecular Evolution**, v. 76, n. 3, p. 83–97, 2013.
- MEŠTROVIĆ N. *et al.* Structural and functional liaisons between transposable elements and satellite DNAs. **Chromosome Res**, v. 23, n. 3, p. 583–96, 2015.
- MEŠTROVIĆ, N. *et al.* Structural and functional liaisons between transposable elements and satellite DNAs. **Chromosome Research**, v. 23, p. 583–596, 2015.
- MICHELS-SOUZA, M. A. *et al.* A Ictiofauna das Lagoas Adjacentes à Lagoa Dourada e Tarumã no Parque Estadual de Vila Velha, Ponta Grossa, Paraná. **Inst. Ambient. do Paraná, Curitiba**, v. 1, p. 1–374, 2011.
- MILANI, E. J. *et al.* Bacia do Paraná: Estratigrafia e Evolução Tectono-Estratigráfica. **Boletim de Geociências da Petrobras**, v. 15, n 2, p. 265–287, 2007.
- MILNE, I. *et al.* Using tablet for visual exploration of second-generation sequencing data. **Briefings in Bioinformatics**, v. 14, n. 2, p. 193–202, 2013.
- MIYA, M. *et al.* MiFish, a Set of Universal PCR Primers for Metabarcoding Environmental DNA from Fishes: Detection of More than 230 Subtropical Marine Species. **Royal Society Open Science**, v. 2, 7: 150088, 2015.
- MORITZ, C. Defining ‘Evolutionarily Significant Units’ for conservation. **Trends in Ecology & Evolution**, v. 9, n. 10, p. 373–375, 1994.
- NEAR, T. J.; BOLNICK, D. I.; WAINWRIGHT, P. C. Fossil calibrations and molecular divergence time estimates in Centrarchid fishes (Teleostei: Centrarchidae). **Evolution**, v. 59, p. 1768–1782, 2005.
- NELSON, S.L.; TAYLOR, S. A.; REUTER, J. D. An isolated white-tailed deer (*Odocoileus virginianus*) population on St John US Virgin Islands shows low inbreeding and comparable heterozygosity to other larger populations. **Ecol Evol**, v. 11, p. 2775–2781, 2020.
- NOVÁK, P.; NEUMANN, P.; MACAS, J. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. **Nature Protocols**, v. 15, p. 3745–3776, 2020.
- PAIXÃO, R. V.; RIBOLLI, J.; ZANIBONI-FILHO, E. Genetic Variation of the Endangered Neotropical Catfish *Steindachneridion scriptum* (Siluriformes: Pimelodidae). **Frontiers in Genetics**, v. 9, n. 48, 2018.

- PALACIOS, D. *et al.* The deglaciation of the Americas during the Last Glacial Termination. **Earth Science Reviews**, v. 203, 2020.
- PAZZA, R. *et al.* The species complex *Astyanax fasciatus* Cuvier (Teleostei, Characiformes) - A multidisciplinary approach. **Journal of Fish Biology**, v. 72, p. 2002–2010, 2008.
- PAZZA, R.; KAVALCO, K. F.; BERTOLLO, L. A. C. Chromosome polymorphism in *Astyanax fasciatus* (Teleostei, Characidae). 1. Karyotype analysis, Ag-NORs and mapping of the 18S and 5S ribosomal genes in sympatric karyotypes and their possible hybrid forms. **Cytogenetic and Genome Research**, v. 112, n. 3–4, p. 313–319, 2006.
- PENTEADO, P. R.; KAVALCO, K. F.; PAZZA, R. Um terceiro citótipo de *Astyanax* aff. *fasciatus* (Teleostei, Characidae) da bacia do rio São Francisco. **Evolução e Conservação Da Biodiversidade**, v. 4, n. 1, p. 1–7, 2013.
- PETERSON, B. K. *et al.* Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. **PLoS One**, v. 7, n. 5, 2012.
- PLOHL, M.; MEŠTROVIĆ, N. Satellite DNA Evolution. **Genome dynamics**, v. 7, p. 126–152, 2012.
- PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of Population Structure Using Multilocus Genotype Data. **Genetics**, v. 155, n. 2, p. 945–959, 2000.
- PURCELL S, *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. **Am J Hum Genet**, v. 81, n. 3, p. 559–575, 2007.
- R Core Team. R: A Language and Environment for Statistical Computing. **R Foundation for Statistical Computing**. <https://www.r-project.org/>, 2024.
- RAMBAUT, A. *et al.* Posterior summarization in Bayesian phylogenetics using Tracer 1.7. **Syst Biol**, v. 67, p. 901–904, 2018.
- RAMBAUT, A. FigTree v. 1.4.0. <http://tree.bio.ed.ac.uk/software/figtree/>. 2012.
- REBORDINOS, L.; CROSS, I.; MERLO, A. High evolutionary dynamism in 5S rDNA of fish: State of the art. **Cytogenetic and Genome Research**, 141(2–3), 103–113, 2013.
- REINAR, W. B. *et al.* Teleost genomic repeat landscapes in light of diversification rates and ecology. **Mobile DNA**, v. 14, n. 1, 2023.
- RIFFEL, S. B. *et al.* Combined <sup>40</sup>Ar/<sup>39</sup>Ar and (U-Th)/He geochronological constraints on long-term landscape evolution of the Second Paraná Plateau and its ruiniform surface features, Paraná, Brazil. **Geomorphology**, v. 233, p. 52–63, 2015.
- RIZZATO, P. P.; BICHUETTE, M. E. A new species of cave catfish from Brazil, *Trichomycterus rubbioli* sp. n., from Serra do Ramalho karstic area, São Francisco River basin, Bahia State (Siluriformes: Trichomycteridae), **Zootaxa**, v. 3480, n. 1, p. 48–66, 2012.
- RIZZATO, P. P.; BICHUETTE, M. E. *Ituglanis boticario*, a new troglomorphic catfish (Teleostei: Siluriformes: Trichomycteridae) from Mambai karst area, central Brazil. **Zoologia**, v. 31, n. 6, p. 577–598, 2014.

- RUIZ-RUANO, F. J. *et al.* High-throughput analysis of the satellitome illuminates satellite DNA evolution. **Scientific Reports**, v. 6, n. 28333, 2016.
- SALES, N. G. *et al.* Hidden Diversity Hampers Conservation Efforts in a Highly Impacted Neotropical River System. **Frontiers in Genetics**, v. 9, n. 271, 2018.
- SAYERS, E. W. *et al.* GenBank. **Nucleic Acids Research**, 50(D1): D161–D164, 2022.
- SCHARTL, M. *et al.* The genomes of all lungfish inform on genome expansion and tetrapod evolution. **Nature**, v. 634, n. 8032, p. 96–103, 2024.
- SCHIEBELHUT, L. M. *et al.* Genomics and conservation: Guidance from training to analyses and applications. **Molecular Ecology Resources**, v. 24, n. 2, 2024.
- SCHRADER, L.; SCHMITZ, J. The impact of transposable elements in adaptive evolution. **Molecular Ecology**, v. 28, n. 6, p. 1537–1549, 2018.
- SCHUG M.D. *et al.* Isolation and genetic diversity of *Gambusia hubbsi* (mosquitofish) populations in blueholes on Andros Island, Bahamas. **Heredity (Edinb)**, v. 80, p. 336–346, 1998.
- SEDORKO, D. Chronostratigraphy and environment of Furnas Formation by trace fossil analysis: Calibrating the lower Paleozoic Gondwana realm in the Paraná Basin (Brazil). **Palaeogeogr Palaeoclimatol Palaeoecol**, v. 487, p. 307–320, 2017.
- SHAO, F.; HAN, M.; PENG, Z. Evolution and diversity of transposable elements in fish genomes. **Scientific Reports**, v. 9, n. 1, 2019.
- SHATSKIKH, A. S. *et al.* Functional Significance of Satellite DNAs: Insights from Drosophila. **Frontiers in Cell and Developmental Biology**, v. 8, p. 1–19, 2020.
- SHIBATTA, O. A. *et al.* Diversidade e distribuição de peixes na bacia do rio Tibagi. In A bacia do rio Tibagi (M.E. Medri, E. Bianchini, O.A. Shibatta & J.A. Pimenta, eds.). Londrina, M. E. Medri, p.403-423, 2002.
- SHIBATTA, O. A. Ictiofauna dos trechos alto e médio da bacia do rio Tibagi, Paraná, Brasil. **Biota neotropica**, v. 7, n. 2, 2007.
- SHIBATTA, O. A.; ARTONI, R. F. Sobre a identidade das populações alopátricas de *Astyanax* (Characiformes, Characidae) das Formações Furna 1 Ponta Grossa, Paraná, Brasil. **Publicatio UEPG**, v. 11, n. 2, p. 7–12, 2005.
- SHIBATTA, O. A.; CHEIDA, C. C. Composição em tamanho dos peixes (Actinopterygii, Teleostei) de ribeirões da bacia do rio Tibagi, Paraná, Brasil. **Revista Brasileira de Zoologia**, v. 20, n. 3, 2003.
- STORFER, A.; PATTON, A.; FRAIK, A. K. Navigating the interface between landscape genetics and landscape genomics. **Frontier in Genetics**, v. 9, 2018.
- STRECKER, U.; FAÚNDEZ, V. H.; WILKENS, H. Phylogeography of surface and cave *Astyanax* (Teleostei) from Central and North America based on cytochrome b sequence data. **Molecular Phylogenetics and Evolution**, v. 33, p. 469–481, 2004.

- TERÁN, G. E.; BENITEZ, M. F.; MIRANDE, J. M. Opening the Trojan horse: Phylogeny of *Astyanax*, two new genera and resurrection of *Psalidodon* (Teleostei: Characidae). **Zoological Journal of the Linnean Society**, v. 190, p. 1217–1234, 2020.
- THAKUR, J.; PACKIARAJ, J.; HENIKOFF, S. Sequence, chromatin and evolution of satellite DNA. **International Journal of Molecular Sciences**, v. 22, n. 9, 2021.
- TOLEDO-PIZA, M. *et al.* Checklist of the species of the Order Characiformes (Teleostei: Ostariophysi). **Neotropical Ichthyology**, v. 22, 2024.
- TRAJANO, E. Diversity of brazilian troglobitic fishes: Models of colonization and differentiation in subterranean habitats. **Diversity**, v. 13, p. 1–19, 2021.
- TRAJANO, E.; SECUTTI, S.; BICHUETTE, M. E. Natural history and population data of fishes in caves of the Serra do Ramalho karst area, Middle São Francisco basin, northeastern Brazil. **Biota Neotropica**, v. 9, p. 129-133, 2009.
- TSUJI, S. *et al.* Quantitative Environmental DNA Metabarcoding Shows High Potential as a Novel Approach to Quantitatively Assess Fish Community. **Scientific Reports**, v. 12, n. 21524, 2022.
- UGARKOVIC, D.; PLOHL, M. Variation in satellite DNA profiles- causes and effects. **The EMBO Journal**, v. 21, n. 22, p. 5955–5959, 2002.
- UTSUNOMIA, R. *et al.* Satellitome landscape analysis of *Megaleporinus macrocephalus* (Teleostei, Anostomidae) reveals intense accumulation of satellite sequences on the heteromorphic sex chromosome. **Scientific Reports** v. 9, n. 5856, 2019.
- WARREN, W. C, *et al.* A chromosome-level genome of *Astyanax mexicanus* surface fish for comparing population-specific genetic differences contributing to trait evolution. **Nature Communications**, v. 12, n. 1447, 2021.
- WEISS, F. E.; MALABARBA, L. R.; MALABARBA, M. Phylogenetic relationships of Paleotetra, a new characiform fish (Ostariophysi) with two new species from the Eocene-Oligocene of south-eastern Brazil. **J Syst Palaeontol**, v. 10, p. 73–86, 2012.
- WILKENS, H.; STRECKER, U. Evolution in the Dark Darwin's Loss Without Selection (S-V GmbH, Ed.). **Springer Nature**, Berlin. 2017.
- WOODWARD, A. S. Considerações sobre alguns peixes Terciários dos schistos de Taubaté, Estado de S. Paulo. **Revista do Museu Paulista**, v. 3, p. 63–70, 1898.
- XIONG, F. *et al.* Methodology for fish biodiversity monitoring with environmental DNA metabarcoding: The primers, databases and bioinformatic pipelines. **Water Biology and Security**, v. 1, n. 1, 2022.
- YAO, M. *et al.* Fishing for Fish Environmental DNA: Ecological Applications, Methodological Considerations, Surveying Designs, and Ways Forward. **Molecular Ecology**, v. 31, n. 20, p. 5132–5164, 2022.

YATES, M. C.; FRASER, D. J.; DERRY, A. M. Meta-Analysis Supports Further Refinement of eDNA for Monitoring Aquatic Species-Specific Abundance in Nature. **Environmental DNA**, v. 1, p. 5–13, 2019.

YUAN, Z. *et al.* Comparative genome analysis of 52 fish species suggests differential associations of repetitive elements with their living aquatic environments. **BMC Genomics**, v. 19, n. 1, p. 1–10, 2018.

**ANEXO A** - Resultado *process radtags* biblioteca 1 (indivíduos da lagoa Dourada), biblioteca 2 (furna 2), biblioteca 3 (rio Tibagi, e rio Paraopebas). Células em cinza indicam aquelas que foram descartadas.

Biblioteca 1					
Barcode	Filename	Total	NoRadTag	LowQuality	Retained
CTGCG	4L	11363522	94236	276528	10881042,00
GCATG	1L	10523740	115582	255306	10044145
ACACA	8L	10034114	98661	242301	9586663
GACAC	23L	9066562	102131	219725	8648837
AGCTA	2L	8858460	70437	214946	8473934
TGCAT	26L	8026554	66784	195417	7684306
CTTGG	16L	7458848	67004	180812	7141119
TCGAT	20L	7461022	65891	181334	7139578
TTACC	24L	6963096	72207	168557	6650902
CGGCT	22L	6700696	58708	163618	6405410
ATTAC	3L	5739008	65328	138299	5476806
GAGAT	29L	5687286	46738	138510	5445211
CGATC	13L	4945720	127498	119146	4643334
CTGTC	10L	3753338	58150	90238	3566656
AACCA	7L	3670966	71074	88448	3477364
CGGTA	28L	3361348	30377	81783	3217015
GGATA	5L	2944832	29999	71524	2816537
ACTGG	27L	2684402	38405	65084	2555050
CGAAT	15L	2379724	30159	58069	2266160
ACGGT	21L	2099988	35208	51097	1994177
CATAT	9L	1358458	31148	32793	1283082
GGCTC	17L	110124	15654	2366	91177
TACGT	11L	86116	13953	1882	69609
GTCCG	30L	73540	5518	1634	65809
TCACG	19L	70076	8691	1480	59417
TAGTA	6L	49970	1574	1218	46712
TCAGT	25L	47084	3267	1099	42291
TATAC	12L	46762	4592	1057	40794
AATTA	14L	35352	14080	627	20499

Biblioteca 2					
Barcode	Filename	Total	NoRadTag	LowQuality	Retained
AGCTA	2F	9274536	73091	225083	8875465
ACTGG	27F	8416538	75620	205039	8053355
GCATG	1F	8346982	118265	202026	7934259
TTACC	24F	8178378	85779	198651	7813772
GGATA	5F	7712678	64166	187732	7382912
CGATC	14F	7713936	133365	186848	7312796

Biblioteca 2					
ACACA	9F	7538500	81266	183001	7194408
TGCAT	26F	7470686	60615	180615	7154927
CATAT	10F	7002322	75798	169629	6686054
CGGTA	28F	6854562	52442	165953	6563460
CTGCG	4F	6735990	58793	163816	6445765
CGGCT	22F	6660980	54885	161250	6367031
TCGAT	20F	6583128	55547	159019	6302476
GACAC	23F	6111008	74803	148412	5812392
GAGAT	29F	4974454	44946	120684	4757601
ACGGT	21F	4523074	42247	110062	4322844
CTTGG	17F	3517770	36782	85081	3360263
CTGTC	11F	2540348	56043	61090	2393579
ATTAC	3F	2425004	33913	58611	2308129
AACCA	7F	2158040	61226	51771	2020783
CGAAT	16F	1755500	22346	42256	1673494
GGCTC	18F	144494	14166	3270	125886
GTCCG	30F	117516	6160	2653	107794
TACGT	12F	116870	8704	2599	104690
TCACG	19F	108998	8951	2545	96728
TAGTA	6F	94580	3735	2258	87821
TCAGT	25F	80158	4114	1884	73657
TATAC	13F	81166	5492	1899	73226
AATTA	15F	38678	14798	645	23055

Biblioteca 3					
Barcode	Filename	Total	NoRadTag	LowQuality	Retained
AACCA	7T	7727264	95457	186777	7366039
CTGCG	4T	7604968	57162	185441	7286551
AGCTA	2T	7454144	64043	180815	7139407
GGTTG	12MG	7282622	61693	177617	6969304
CGATC	13T	7280094	127762	175823	6906341
ACGGT	20T	7157556	65104	173890	6846814
TGGAA	18MG	7080556	59126	171885	6773676
CGGTA	27T	7063390	53252	171772	6763507
CTTGG	16T	6933804	60217	168960	6633804
ACACA	8T	6879506	72032	167346	6572982
CGTAC	8MG	6766668	69205	164919	6463329
TTACC	23T	6756386	71408	165005	6445917
GACAC	22T	6610104	83369	160074	6295893
CTGTC	10T	6506552	82882	157948	6203251

Biblioteca 3					
CGGCT	21T	6478952	53147	158263	6202595
GTCCG	29T	6257376	51081	152736	5993599
TCGAT	19T	6224872	53996	151265	5957609
CTGAT	21MG	6075594	53857	148489	5811456
TCCGG	1MG	6066448	62329	146681	5799667
CAACC	4MG	6031346	110743	146912	5714183
TATAC	12T	5821838	56132	141999	5566032
TACGT	11T	5615204	50123	136725	5374510
GCCGT	16MG	5544398	46753	135358	5306631
CATAT	9T	5404556	63624	131160	5161088
GAGAT	28T	5165742	43338	126050	4945650
AAGGA	19MG	4880446	44262	118857	4666751
GCATG	1T	4881958	52286	118871	4664065
TACCG	17MG	4800674	62413	116899	4579871
ATGAG	20MG	4626508	38968	112896	4430880
CGAAT	15T	4504010	39651	110094	4310166
GGATA	5T	4449184	41303	108225	4259354
ACTTC	5MG	3395048	54744	82525	3225456
GTCGA	10MG	3361088	27026	81568	3220198
TAGTA	6T	3235978	35726	78995	3089294
TCAGT	24T	3209974	35351	77657	3063600
GGCTC	17T	3188522	34887	77731	3048216
ATACG	14MG	3188132	43242	77136	3035934
TGCAT	25T	3144460	29768	76220	3007915
ACTGG	26T	3045666	35221	74831	2904622
ATTAC	3T	2600466	39970	62458	2470161
GAGTC	9MG	2514920	56419	61504	2372376
TCACG	18T	1796954	29352	43681	1706578
TCTGC	11MG	69560	15717	1464	51924
AATTA	14T	33678	14155	569	18777
CGTCG	15MG	11612	3325	213	7995

## ANEXO B- Pipeline Stacks usada para o *SNPs calling e Upset plot*

### 1. Installing mamba

```
conda create -n myenv
conda activate myenv
conda install -c conda-forge mamba
mamba --version
```

### 2. Installing Stacks

```
mamba install stacks --> you can see documentation bioconda
https://bioconda.github.io/recipes/stacks/README.html
```

**3. Concatenate sequences forward and reverse.** Notice that you need to change the name according to the Catchen website!!

Use the command cat to do that.

### 4. Run process\_radtags

```
process_radtags -P -p ./raw -b./barcodes/barcodes_lane4 -o ./samples/ \-c -q -r --inline-index --renz-1 nlaIII --renz-2 mluCI
```

In my data, the command line should be like this for the sinkhole library:

```
process_radtags -P -p ./Af8 -
b ./raw/B2.txt -o ./raw/samples/ -c -q -r --inline_null --renz_1 sphI --
renz_2 mluCI --adapter_1 AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG --
adapter_2 CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTAAT
2
```

### 5. bwa alignment ## use this bash below, run ./bwa.sh

```
src=/home/lab01/ddradseq2024/Luz/bwa/samples
bwa_db=/home/lab01/ddradseq2024/Luz/genoma_a_mexicanus/mexicanus.fa
output_dir=/home/lab01/ddradseq2024/Luz/bwa/aligned # Define output
directory outside of samples

# Create the output directory if it doesn't exist
mkdir -p $output_dir

# Check if reference genome is indexed; if not, index it
if [ ! -f "${bwa_db}.bwt" ]; then
    echo "Indexing the reference genome..."
    bwa index $bwa_db
fi

files="ID1
ID2
ID3
...
"

# Align paired-end data with BWA, convert to BAM and SORT.
#
for sample in $files
do
    bwa mem -t 8 $bwa_db $src/${sample}.1.fq.gz $src/${sample}.2.fq.gz | \
```

```

samtools view -b | \
samtools sort --threads 4 -o $output_dir/${sample}.bam
done

```

#### 6. Sobre samtools ## install samtools

conda activate bioenv --> para entrar no ambiente onde foi instalado o samtools

I used the

samtools flagstat 1F.bam ## for look at the percentage of alignment of my sequences against the reference genome.

#### 7. gstacks

```

gstacks -I ./aligned_mexicanus -O ./gstacks_mexicanus -M ./popmap -t 8 --gt-
alpha 0.01 --var-alpha 0.05

```

#### 8. Populations

```

populations -P ./gstacks_mexicanus/ --popmap ./gstacks_mexicanus/POPMAP1.txt
--smooth -p 4 -r 0.8 -t 8 --min_mac 5 --hwe --min_maf 0.05 --write_random_snp
--vcf --genepop --structure --plink >>> 22036 snps

```

Filter VCF by population

### To extract a VCF for a specific population, you can use vcftools. First, create a text file that lists the individuals in the population of interest:

```

echo "sample1" > Pop1_samples.txt
echo "sample2" >> Pop1_samples.txt

```

```

vcftools --vcf stacks_output.vcf --keep Pop1_samples.txt --recode --out
Pop1_vcf

```

### Upset plot

```

mport vcfpy # Make sure to install PyVCF with `pip install pyvcf`

```

```

# Open the VCF file

```

```

vcf_reader = vcfpy.Reader.from_path('populations.snps.vcf')

```

```

# Define populations with their sample names (update with actual sample names
in your VCF file)

```

```

populations = {
    "sinkhole": ["1F", "2F", "3F", "4F", "5F", "7F", "9F", "10F", "11F",
"14F", "16F", "17F", "20F", "21F", "22F", "23F", "24F", "26F", "27F", "28F",
"29F"],
    "lake": ["1L", "2L", "3L", "4L", "5L", "7L", "8L", "9L", "10L", "13L",
"15L", "16L", "20L", "21L", "22L", "23L", "24L", "26L", "27L", "28L", "29L"],
    "tibagi": ["1T", "2T", "3T", "4T", "5T", "6T", "7T", "8T", "9T", "10T",
"11T", "12T", "13T", "15T", "16T", "17T", "18T", "19T", "20T", "21T", "22T",
"23T", "24T", "25T", "26T", "27T", "28T", "29T"],
    "minas": ["1MG", "4MG", "5MG", "8MG", "9MG", "10MG", "12MG", "14MG",
"16MG", "17MG", "18MG", "19MG", "20MG", "21MG"],
}

```

```

# Initialize a dictionary to store SNP positions for each population

```

```

snp_positions = {pop: set() for pop in populations.keys()}

```

```

# Loop through each record (SNP) in the VCF file to collect SNP positions
for each population
for record in vcf_reader:
    for pop_name, samples in populations.items():
        for sample_name in samples:
            # Find the correct sample call by matching the name
            sample = next((call for call in record.calls if call.sample ==
sample_name), None)
            if sample is None:
                continue

            # Check if the genotype is a variant
            genotype = sample.data.get("GT")
            if genotype and genotype != "0|0":
                # Add SNP position to the population's set
                snp_positions[pop_name].add(record.POS)

# Step 1: Print SNP positions per population for verification
print("SNP Positions per Population:")
for pop_name, positions in snp_positions.items():
    print(f"{pop_name}: {sorted(positions)}")

# Step 2: Identify SNPs shared by all populations
shared_snps = set.intersection(*snp_positions.values())
print("\nSNPs shared by all populations:", sorted(shared_snps))

# Step 3: Identify unique SNPs per population
unique_snps = {
    pop: positions - set.union(*(snp_positions[p] for p in snp_positions if
p != pop))
    for pop, positions in snp_positions.items()}
print("\nUnique SNPs per Population:")
for pop_name, positions in unique_snps.items():
    print(f"{pop_name}: {sorted(positions)}")

```

### Diagrama de Venn

```

import vcfpy
from itertools import combinations
import matplotlib.pyplot as plt
from venn import venn
import matplotlib.cm as cm

# Path to your VCF file
vcf_path = "populations.snps.vcf" # Replace with the actual path to your VCF
file

# Define the sample names for each population
populations = {
    "sinkhole": ["1F", "2F", "3F", "4F", "5F", "7F", "9F", "10F", "11F",
"14F", "16F", "17F", "20F", "21F", "22F", "23F", "24F", "26F", "27F", "28F",
"29F"],

```

```

    "lake": ["1L", "2L", "3L", "4L", "5L", "7L", "8L", "9L", "10L", "13L",
"15L", "16L", "20L", "21L", "22L", "23L", "24L", "26L", "27L", "28L", "29L"],
    "tibagi": ["1T", "2T", "3T", "4T", "5T", "6T", "7T", "8T", "9T", "10T",
"11T", "12T", "13T", "15T", "16T", "17T", "18T", "19T", "20T", "21T", "22T",
"23T", "24T", "25T", "26T", "27T", "28T", "29T"],
    "minas": ["1MG", "4MG", "5MG", "8MG", "9MG", "10MG", "12MG", "14MG",
"16MG", "17MG", "18MG", "19MG", "20MG", "21MG"],
}

# Initialize a dictionary to store SNP positions for each population
snp_positions = {pop: set() for pop in populations.keys()}

# Open the VCF file and process each record
vcf_reader = vcfpy.Reader.from_path(vcf_path)

for record in vcf_reader:
    # Go through each population and its samples
    for pop_name, sample_names in populations.items():
        for sample_name in sample_names:
            # Find the sample call in the record
            sample = next((call for call in record.calls if call.sample ==
sample_name), None)
            if sample is None:
                continue

            # Check if the sample has a variant (non-reference) genotype
            genotype = sample.data.get("GT")
            if genotype and genotype != "0|0" and genotype != "0/0":
                # Add the SNP position to the population's set
                snp_positions[pop_name].add(record.POS)

# Debugging: Print the number of SNPs in each population
print("Number of SNPs in each population:")
for pop_name, snp_set in snp_positions.items():
    print(f"{pop_name}: {len(snp_set)}")

# Calculate unique and shared SNPs based on the extracted SNP positions
unique_snps = {}
shared_snps = set.intersection(*snp_positions.values())

print("\nShared SNPs across all populations:", len(shared_snps))

for pop_name in populations.keys():
    others = set.union(*(snp_positions[other] for other in populations if
other != pop_name))
    unique_snps[pop_name] = snp_positions[pop_name] - others
    print(f"Unique SNPs in {pop_name}:", len(unique_snps[pop_name]))

# Visualize results with a Venn Diagram
plt.figure(figsize=(10, 8))
venn(snp_positions, cmap=cm.plasma)
plt.title("Venn Diagram of SNPs Across Populations")
plt.show()

```

### Análise de Fst

```

install.packages("vcfR")      # Para importar arquivos VCF
install.packages("adegenet") # Para converter para genind
install.packages("hierfstat") # Para análises genéticas

library(vcfR)
library(adegenet)
library(hierfstat)

setwd("C:/Luz/LEDG/Luz/Doutorado/ddRADseq/2024/Luz/Ref_mexicanus/Fst")

vcf_data <- read.vcfR("populations.snps.vcf")
vcf_data
populations <- c(rep("F", 21), rep("L", 21), rep("T", 28), rep("MG", 14))
genotype_matrix <- vcfR2genind(vcf_data)
genotype_matrix@pop <- as.factor(populations)
genotype_matrix@pop
#####
hierfstat_data <- genind2hierfstat(genotype_matrix)
fst <- basic.stats(hierfstat_data)
fst
#####
fst1 <- basic.stats(genotype_matrix)
fst1
is.genind(genotype_matrix)

# Calcular distância genética (Fst)
pairwise_fst <- genet.dist(genotype_matrix, method = "WC84")

# Visualizar matriz de Fst
print(as.matrix(pairwise_fst))

library(ggplot2)
library(reshape2)

# Transformar matriz em formato longo para ggplot
fst_long <- melt(as.matrix(pairwise_fst))

# Plotar heatmap
ggplot(fst_long, aes(Var1, Var2, fill = value)) +
  geom_tile(color = "white") +
  scale_fill_gradient2(low = "blue", mid = "white", high = "red", midpoint =
0,
                        name = "Fst") +
  theme_minimal() +
  labs(title = "Pairwise Fst", x = "Population", y = "Population") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

# Manter solo la diagonal inferior

```

```
fst_long$value[fst_long$Var1 < fst_long$Var2] <- NA # O cambiar "<" por ">"
para diagonal superior

# Plotar el heatmap
ggplot(fst_long, aes(Var1, Var2, fill = value)) +
  geom_tile(color = "white") +
  scale_fill_gradient2(low = "blue", mid = "white", high = "red", midpoint =
0, na.value = "transparent", # Valores NA como transparentes name = "Fst") +
  theme_minimal() +
  labs(title = "Pairwise Fst (Lower Triangle Only)", x = "Population", y =
"Population") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

## ANEXO C – Sequencias utilizadas na construção do relógio molecular

<b>Species</b>	<b>Accession number</b>	<b>Reference</b>
<i>Salminus brasiliensis</i>	KY825190	Ribeiro,J.M. et al, 2017
<i>Salminus brasiliensis</i>	NC_024941	Brandão-Dias et al. 2016
<i>Brycon nattereri</i>	MT428074	Resende et al. 2020
<i>Brycon orbignyanus</i>	NC_024938	Siqueira et al. 2016
<b>Brycon nattereri</b>	NC051927	Resende et al. 2021
<i>Paracheirodon axelrodi</i>	AB898197	Zhang et al. 2013
<i>Paracheirodon axelrodi</i>	NC043904.1	Huang and Liu. 2019
<i>Paracheirodon innesi</i>	NC_028279	Liu et al. 2019
<i>Gephyrocharax atracaudatus</i>	NC_042882	Huang and Liu. 2019
<i>Grundulus bogotensis</i>	NC_026195	Isaza et al. 2016
<i>Psalidodon paranae</i>	NC_031380	Silva et al. 2016
<i>Psalidodon fasciatus furna</i>	OP210211	This work
<i>Psalidodon fasciatus non-isolated</i>	OP218849	This work
<i>Psalidodon fasciatus Paraopeba</i>	PV700218	This work
<i>Hemigrammus bleheri</i>	NC_027946	Li et al. 2016
<i>Hasemanina nana</i>	NC_022724	Xu et al. 2015
<i>Acestrorhynchus sp</i>	AP011981	Nakatani et al. 2011
<i>Brycon henni</i>	NC_026873	Landínez-García et al. 2016
<i>Chalceus macrolepidotus</i>	NC_004700	Saitoh et al. 2003

## ANEXO D – Artigo publicado na *Environmental DNA*

### Title

Illuminating the Dark: Using eDNA Metabarcoding to Study the Unique Ichthyofauna of the Furnas Formation Sinkholes

### Abstract

Environmental DNA (eDNA) metabarcoding is a powerful tool for uncovering hidden biodiversity and identifying underestimated taxa in aquatic ecosystems. In this study, we applied eDNA metabarcoding to assess fish species in isolated environments of the Tibagi River basin, specifically sinkholes. Water samples were collected from three sinkholes (Sites 1, 2, and 4), one site at Dourada Lake within the Vila Velha State Park (VVSP), and two sites along the main channel of the Tibagi River. A fragment of the mitochondrial 12S rDNA gene was amplified and sequenced, yielding a final dataset of 975,197 reads after filtering. We identified 31 Molecular Operational Taxonomic Units (MOTUs) belonging to Actinopterygii, distributed across four taxonomic orders: Characiformes, Cichliformes, Gymnotiformes, and Siluriformes. No MOTUs were detected in sinkhole 1 indicating local extinction. A single MOTU (*Psalidodon fasciatus*) was identified in sinkhole 2. In sinkhole 4, two distinct MOTUs were found, corresponding to the family Heptapteridae and the genus *Cambeva*, but the potential presence of *P. fasciatus* was not confirmed. As expected, the highest species richness was observed in the Tibagi River. Despite the lack of a local reference database, our findings offer valuable insights into the detection of diverse taxa in isolated aquatic environments using eDNA metabarcoding. This study underscores the potential of eDNA to inform conservation strategies in ecologically unique and underexplored habitats, and highlights the need for developing region-specific genetic reference databases to enhance taxonomic resolution and monitoring efforts in the future.

### Keywords

Ecosystem, Characiformes, Rivers, Genetic Markers, Biodiversity, Water

### Introduction

Identifying and assessing freshwater fish biodiversity is a persistent challenge, particularly in remote or inaccessible regions where the true distribution of many species remains unknown, e.g., the Wallacean shortfall (Lomolino, 2004). Environmental DNA (eDNA) metabarcoding has emerged as a powerful tool to overcome these limitations (Blackman et al., 2024). This technique has been widely used for various applications, including the identification of fish

communities (Tsuji et al., 2022), aquatic ecosystem monitoring (Yates et al., 2019; Xiong et al., 2022), and the survey of rare groundwater fauna (Boyd et al., 2020). Additionally, eDNA metabarcoding has been instrumental in uncovering hidden biodiversity and detecting underestimated taxa in hydrographic basins (Gleason et al., 2023; Sales et al., 2018). This technique has also proven effective for non-invasive monitoring of endangered species (Beng & Corlett, 2020) and the early detection of biological invasions (Fernández et al., 2018). One of the key advantages of eDNA metabarcoding is its ability to monitor biodiversity across extensive and often remote or difficult-to-access regions, making it particularly valuable in protected areas (Herder et al., 2014).

Although eDNA metabarcoding is not commonly used to monitor fish diversity in small water bodies—since such environments are often accessible and well documented through traditional capture-based methods (Yao et al., 2022)—certain geological formations present significant logistical challenges for sampling. A notable example is the sinkholes of the Furnas Formation, located within Vila Velha State Park (VVSP) in Paraná State, southern Brazil. These sinkholes are ancient geological structures dating back to the Silurian-Devonian period (421–395 million years ago; Milani et al., 2007), with depths ranging from 15 to 60 meters from the surface to the groundwater level. The water column, formed by the exposed water table, varies between 20 and 45 meters, embedded within the rock. Although the groundwater system connects these sinkholes via rock porosity, no effective passageways exist that would allow the movement of macroscopic organisms between them.

The formation processes of sinkholes remain only partially understood. However, it is hypothesized that the dissolution of kaolinite and illite cement, followed by the arenization of sandstone and mechanical erosion, contributes to the complex genesis of these features (de Melo & Giannini, 2007). Currently, approximately 30 geological formations in the region have been classified as sinkholes. The VVSP is located within the Tibagi River basin and is drained primarily by the Barrozinho and Quebra-Perna Rivers, which converge to form the Guabiroba River (Michels-Souza et al., 2011). The Guabiroba River basin, where Dourada Lake is situated, covers ~360 km<sup>2</sup> and features a triangular drainage system. The elevation gradient of the basin ranges from 1,117 meters above sea level (m.a.s.l.) in the northeast to 865 m.a.s.l. in the southwest, reflecting a westward decline in topography (de Melo et al., 2000).

The VVSP area contains four major sinkholes: three wet sinkholes (1, 2, and 4), where the groundwater table forms ponds at the bottom, and one dry sinkhole (3), which lacks a permanent

water body. These deep sinkholes create isolated microhabitats, potentially harboring endemic fish populations (Artoni & Almeida, 2003). This unique geological configuration presents an opportunity to investigate the evolutionary history and ecological dynamics of the species inhabiting these environments. Previous morphometric, cytogenetic, and molecular studies have documented the presence of only *Psalidodon fasciatus* (Cuvier 1819) (*formerly Astyanax aff. fasciatus*; Téran et al., 2020) in sinkholes 1 and 2 (Shibatta & Artoni, 2005; Matoso et al., 2013). However, sinkhole 1 has posed significant sampling challenges, leading to relatively small sample sizes in previous studies. Notably, a well-documented local extinction of *P. fasciatus* occurred in sinkhole 1 in the early 2000s, attributed to eutrophication caused by bird guano accumulation during the avian breeding season (Matoso et al., 2010). This species was the only known fish population in this environment. Since then, despite multiple sampling efforts, *P. fasciatus* has not been observed again in sinkhole 1. In contrast, sinkhole 2 hosts a single, isolated population of *P. fasciatus*, which has been the focus of research for several years (Matoso et al., 2004; Gross et al., 2004; Matoso et al., 2010; Matoso et al., 2013; Kerniske et al. 2021; Kerniske et al, 2023). To date, no other fish species have been recorded in this environment. Given that the current configuration of these sinkholes has remained unchanged for over 10 million years (Riffel et al., 2015), the persistence of these isolated fish populations is particularly intriguing. Our recent analyses of whole mitochondrial genomes indicate that these fish have been genetically isolated from their counterparts in the Tibagi River basin for over a thousand years (De la Ossa-Guerra et al., in preparation).

Unlike sinkholes 1 and 2, sinkhole 4 remains unexplored in terms of fish diversity, leaving uncertainty about the possible presence of fish populations. In contrast, the ichthyofauna of Dourada Lake (Artoni & Shibatta, 2006; Gealh, 2007) and the Tibagi River (Shibatta et al., 2002; Shibatta et al., 2007; Garla et al., 2023) has been well documented using traditional collection methods. The Tibagi River basin is one of the most important hydrographic systems in Paraná State, crossing two of the five largest cities in the region. This area supports a high level of fish biodiversity (>110 species) (Shibatta et al., 2002, 2007; Claro-García et al., 2018) and plays a central role in ichthyofaunal conservation. The fish diversity in the basin has been extensively studied (Bennemann et al., 1995; Shibatta et al., 2002, 2007; Claro-García, 2018), with species richness increasing along the river's course from the headwaters to the mouth (Shibatta et al., 2002).

Thus, the current study aims to assess fish biodiversity in the sinkholes of VVSP using eDNA metabarcoding, focusing on determining whether *P. fasciatus* remains the only fish species present in sinkhole 2, confirming its local extinction in sinkhole 1, and investigating the potential presence of *P. fasciatus* and other fish taxa in sinkhole 4, which remains unexplored in terms of ichthyofauna. By applying eDNA metabarcoding, this study seeks to expand knowledge on fish biodiversity within these isolated aquatic systems, providing insights that are relevant to species conservation in hard-to-reach environments.

## **Material and Methods**

### *Study area and eDNA Sampling*

A total of six sampling sites were selected for this study, comprising three sinkholes (Sites 1, 2, and 4), one site at Dourada Lake within the VVSP, and two sites along the main channel of the Tibagi River (Figure 1). Two sample replicates were used at each site, except for the Tibagi River locations, where we sampled at only one site. The sampling site established in the Tibagi River served both as a comparative reference for the isolated environments and as a means to contribute to the 12S reference database for the river system.

At each sample site, surface water was filtered using Sterivex (0.45µm of pore size) closed filters and performed in triplicate using the eDNA sampling kit described elsewhere (Mendes et al., 2024a), with a total of six filters and water volume of 180 ml per filter and 1080 ml per site. After filtering, eDNA was preserved using ATL buffer injected into the filters. Filters were kept at 4°C until further processing. In order to avoid contamination between sample sites, prior to sampling filtration, all collection kits are composed of sterilized discardable plastics. Furthermore, researchers donned fresh gloves and masks at each eDNA sampling site. To serve as a negative control, mineral water (540 ml) was filtered using the same protocol before sampling filtration.

Water sampling at Sinkholes 1 and 4 was conducted via rappelling due to their limited accessibility. In contrast, sampling at more accessible sites—Sinkhole 2, Dourada Lake, and the Tibagi River—was performed on foot, without the need for specialized descent techniques.

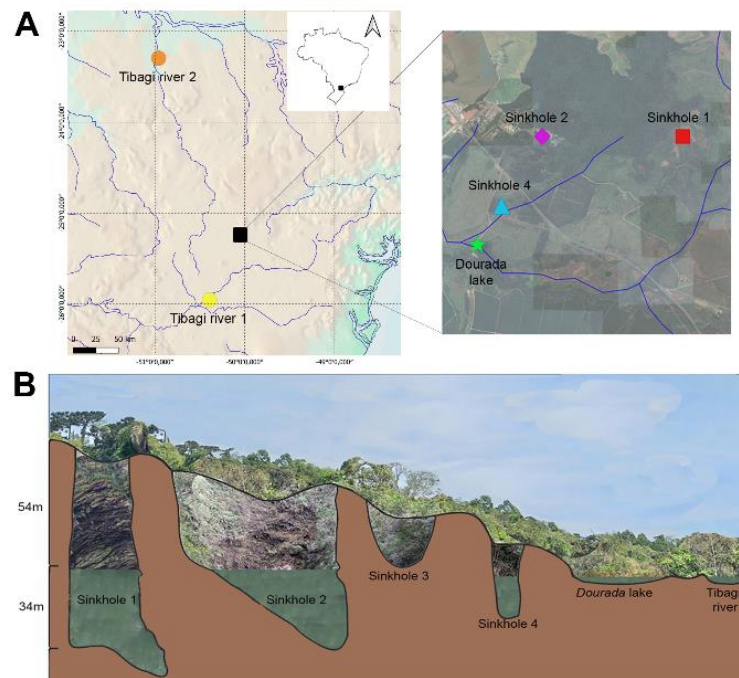


Figure 1. Geographic location and geological context of the study sites. A. Map showing the spatial distribution of the six aquatic environments sampled for eDNA analysis in southern Brazil. A. The left panel indicates the broader region, with the two Tibagi River sites (orange and yellow circles) and their relative positions within the river basin. The right panel provides a detailed satellite view highlighting the locations of the three sinkholes (Sinkholes 1–2–4), and Dourada Lake. B. Conceptual cross-sectional diagram of the karstic landscape, illustrating the relative depth and spatial arrangement of the sinkholes, Dourada Lake, and the Tibagi River.

#### *eDNA extraction and sequencing*

Environmental DNA was extracted using the *DNA PowerWater* kit (Qiagen) with modifications (Mendes et al., 2024b) in a clean lab dedicated to eDNA processing. Sterivex cartridges collected at each sample site was extracted separately to avoid cross contamination between samples. A 200bp fragment of the mitochondrial 12S region was amplified using MiFish primers (Miya *et al.*, 2015), with three independent PCR amplifications performed per sample (filter) to reduce amplification bias. These replicates were pooled for sequencing. The PCR reaction used a 20  $\mu\text{L}$  mix containing 8.34  $\mu\text{L}$  of water, 0.16  $\mu\text{L}$  of BSA (100  $\mu\text{g}/\text{mL}$ ), 10  $\mu\text{L}$  of AmpliTaq Gold 360 Master Mix (Thermo Fisher), 0.25  $\mu\text{L}$  of each primer, and 1.0  $\mu\text{L}$  of DNA template with water replacing the DNA in negative controls. Primers included a 5' tail for a second PCR, and the PCR products were purified using magnetic beads. After purification, a second PCR was conducted using Nextera Index kit® adapters to label and amplify the amplicons. Illumina-compatible adapters (P5 and P7) served as primers. The amplification product was then assessed on a 1.5% agarose gel.

All samples were pooled, and this pool was then purified with the ZymoClean™ Large Fragment DNA Recovery kit (Zymo Research). The pool was quantified using real-time PCR, with the KAPA Biosystems Quantification Kit (Illumina). The final solution was denatured and loaded onto the NextSeq 2000® equipment (Illumina), using the NextSeq® V2.5 sequencing kit (2x150bp PE).

#### *Raw data processing and taxonomic assignment*

Bioinformatic analyses were performed using a custom pipeline (Hilário *et al.*, 2023), organized in R 4.3.1 (R Core Team, 2024), with the DADA2 (Callahan *et al.*, 2016), Phyloseq (McMurdie & Holmes, 2013) and the BLASTr (Hilário, 2024) R packages. Raw reads were demultiplexed by Basespace, and adapter and primer sequences were removed using Cutadapt (Martin, 2011). Low-quality reads (below Q20) or those with indeterminate bases were filtered out, and remaining reads were dereplicated and merged into ASVs using DADA2 with default parameters. ASVs were also clustered into MOTUs using SWARMv2 (Mahé *et al.*, 2015). The complete pipeline is available at [https://github.com/heronoh/fish\\_eDNA](https://github.com/heronoh/fish_eDNA).

The read count of each ASV detected in the negative control was subtracted from the read count detected in each ASV. Thus, to ensure reliable ASV detection, those with <1% RRA (relative read abundance) were excluded (Supplementary material S2). This threshold was chosen to balance sensitivity and specificity, avoiding over-interpretation of low-read ASVs while retaining biologically meaningful taxa.

Currently, no curated 12S reference database exists for the Tibagi River Basin, which limits the resolution of taxonomic assignments. Consequently, we performed the identification of ASVs using as a reference the complete *NCBI core nucleotide* database (NCBI, 2024), combined with a local curated database, generated for the 12S region of freshwater fishes of some Brazilian river basins (Hilário *et al.*, 2023), searched with BLASTn (Camacho *et al.*, 2009, Sayers *et al.*, 2022), with a minimum identity threshold of 80%. All taxonomic assignment results were curated. The ASVs with best hit identity  $\geq 98\%$  were assigned to species level; those with identity between 95% and 98% were assigned to genus level, and those with identity between 90% and 95%, to family level. Only ASVs above 80% were kept for downstream analyses. ASVs that matched to hits of different species of a same genus were assigned to genus level. ASVs with hit sequences from mammals and birds were excluded from analysis among sites.

## Results

The sequencing process generated a total of 1,990,069 raw reads, which, after quality filtering and merging, resulted in 1,497,479 sequences. After bioinformatics processing, 90 Molecular Operational Taxonomic Units (MOTUs) were initially identified across all samples (Table S1). Reads obtained from each Sterivex filter (technical replicates) at each sample site were combined, except for the Tibagi River sites, which were analyzed separately due to their greater geographic distance, resulting in a final number of six sample points (Figure 1).

Actinopterygii (ray-finned fishes) were the most abundant taxa across most locations (Tibagi River 2, Tibagi River 1, Dourada lake, and Sinkhole 2) (Figure 2A). In contrast, Sinkhole 1 and Sinkhole 4 exhibit higher proportions of mammalian ASVs, with Sinkhole 1 also showing a small proportion of bird sequences.

We observed 49 amplicon sequence variants (ASVs) and a final dataset containing 995,775 reads and 31 MOTUs was used for downstream analyses. A total of four different taxonomic orders were assigned: Characiformes, Cichliformes, Gymnotiformes, and Siluriformes (Table 1). A significant number of the ASVs could only be classified to the order or family level (Figure 2C, 2D) due to the insufficient availability of reference sequences in the 12s region for the Tibagi river basin.

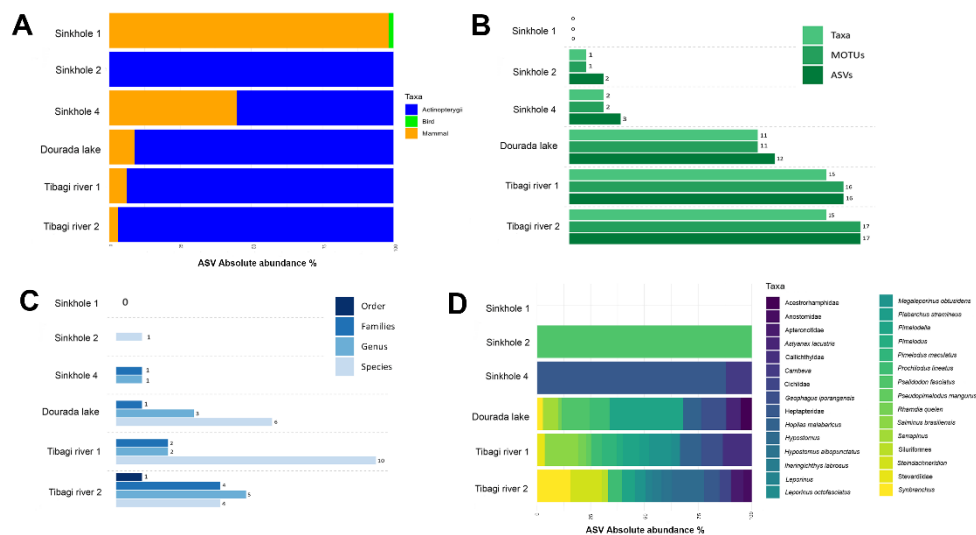


Figure 2. Results of eDNA metabarcoding analysis for fish detection across the different sites. A. ASV absolute abundance of all classes using the 12s gene. B. Numbers of ASVs, MOTUs, and assigned taxa for all sites. C. Number of taxonomic levels identified in each site. D. Percentage of ASV absolute abundance of taxa in each site. Bars represent the values for each category.

*Taxonomic resolution of the 12S rDNA mitochondrial gene in the Vila Velha State Park*

No Actinopterygii ASV were identified at any taxonomic level in sinkhole 1. In sinkhole 2, *P. fasciatus* presented the highest number of ASV absolute abundance (428,564 reads). Moreover, no other ASV were identified, supporting the hypothesis that this species' population is the sole fish inhabitant of sinkhole 2. For sinkhole 4, the results demonstrated the existence of two discrete ASVs, exhibiting BLASTn identity scores of 93% and 97%, respectively, matching Heptapteridae and *Cambeva* (Figure 2D, Table 1).

The Dourada lake sample yielded a total of 191,956 reads, which were attributed to 11 distinct MOTUs, with relative abundance ranging from approximately 1% to 38.6%, which could be taxonomically assigned to 11 taxa. Furthermore, only two MOTUs were identified at the family level (Acestrorhamphidae) (Table 1).

*Taxonomic resolution of the 12S rDNA mitochondrial gene in the Tibagi river*

The BLASTn from the Tibagi River basin samples revealed the highest diversity, with 15 and 17 MOTUs identified in Tibagi 1 and Tibagi 2, respectively (Figure 2a). For some MOTUs, the closest taxonomic resolution reached was at the family level (7 MOTUs). The dominant orders identified, Characiformes and Siluriformes, are among the most commonly recorded in Neotropical tributaries (Claro-García et al., 2018) (Figure 3B). The most abundant ASVs in the Tibagi River were *Hypostomus* and *Steindachneridion*.

Table 1. Taxonomic identification for fish detected from eDNA samples across six aquatic sites using metabarcoding. The table presents, for each site, the identified taxa, their respective relative read abundance (RRA, %), the number of unique amplicon sequence variants (ASVs), and the number of Operational Taxonomic Units (OTUs). MOTUs were taxonomically assigned to the lowest possible level (e.g., species, genus, family).

Sites	Identified taxa	RRA (%)	Num ASVs	Num OTU
<b>Sinkhole 1</b>	0	0.00	0	0
<b>Sinkhole 2</b>	<i>Psalidodon fasciatus</i>	98.59	2	1
<b>Sinkhole 4</b>	Heptapteridae	36.67	2	1
	<i>Cambeva</i>	4.99	1	1
<b>Dourada Lake</b>	<i>Pimelodella</i>	28.62	1	1
	<i>Psalidodon fasciatus</i>	10.93	1	1
	<i>Geophagus iporangensis</i>	9.52	1	1
	<i>Hoplias malabaricus</i>	7.10	1	1
	<i>Prochilodus lineatus</i>	7.70	2	1
	<i>Serrapinus</i>	5.84	1	1
	<i>Astyanax lacustris</i>	5.60	1	1
	Acestrorhamphidae	4.34	2	2
	<i>Synbranchus</i>	2.30	1	1

	<i>Rhamdia quelen</i>	1.41	1	1
<b>Tibagi River 1</b>	<i>Salminus brasiliensis</i>	14.16	1	1
	Callichthyidae	12.30	2	2
	<i>Megaleporinus obtusidens</i>	9.37	1	1
	<i>Hoplias malabaricus</i>	9.20	1	1
	<i>Geophagus iporangensis</i>	9.02	2	1
	<i>Pimelodella</i>	6.91	1	1
	<i>Pimelodus maculatus</i>	6.20	1	1
	<i>Prochilodus lineatus</i>	4.63	1	1
	<i>Piabarchus stramineus</i>	4.31	1	1
	<i>Leporinus octofasciatus</i>	3.68	1	1
	<i>Rhamdia quelen</i>	3.45	1	1
	Stevardiidae	3.30	1	1
	<i>Pimelodus</i>	2.60	1	1
	<i>Pseudopimelodus mangurus</i>	2.11	1	1
<b>Tibagi River 2</b>	<i>Steindachneridion</i>	13.92	1	1
	<i>Synbranchus</i>	8.40	1	1
	<i>Hypostomus</i>	14.13	3	3
	Heptapteridae	6.72	1	1
	<i>Prochilodus lineatus</i>	6.02	1	1
	<i>Pimelodella</i>	5.70	1	1
	<i>Apteronotidae</i>	5.51	1	1
	Cichlidae	4.97	1	1
	<i>Megaleporinus obtusidens</i>	4.65	1	1
	<i>Hypostomus albopunctatus</i>	4.64	1	1
	<i>Iheringichthys labrosus</i>	3.99	1	1
	Anostomidae	3.73	1	1
	<i>Leporinus</i>	3.13	1	1
	Siluriformes	2.69	1	1

## Discussion

Deciphering the distribution and ecology of subterranean life is essential for species conservation, a task that presents significant challenges using traditional approaches (Boyd et al., 2020). For instance, eDNA metabarcoding has been used to provide an efficient means for discovering and monitoring subterranean life forms, such as three Iranian blind cave fish species (blind Iran cave barb *Garra typhlops*, blind Lorestan cave barb *Garra lorestanensis*, and blind cave loach *Eidinemacheilus smithi*) using 12S rRNA MiFish markers (Hashemzadeh et al., 2022).

In the Neotropical realm, no eDNA metabarcoding studies have been yet applied to assess subterranean biodiversity or sinkholes. Here, our data support the previously proposed distribution pattern for the fish *Psalidodon fasciatus*, corroborating the extinction of the

population in sinkhole 1. In contrast, we confirmed the ongoing presence of *P. fasciatus* in sinkhole 2, the only species inhabiting this site, and where we were able to detect the highest read abundance for this species across all sample sites (Table 1). Previous studies have examined the unique characteristics of this isolated population of *P. fasciatus*. Kerniske et al. (2022) documented morphological malformations, likely due to inbreeding and niche competition, both of which may reduce population fitness. Together, these findings underscore the ecological significance of sinkhole 2 as a distinct habitat for this isolated population.

Interestingly, we report the first-time detection of the family Heptapteridae and genus *Cambeva* in the previously unexplored sinkhole 4. This report raises intriguing questions about the site's biodiversity. Despite a relatively low BLASTn identity score (93.73%) for the ASV assigned to Heptapteridae, the findings remain credible and instead emphasize the insufficiency of available reference sequences for regional species. Due to the difficulty in accessing this sinkhole for biodiversity surveys, it has not yet been possible to carry out extensive sampling efforts using traditional methods. For example, some Heptapteridae taxa, such as *Rhamdiopsis krugi* (Bockmann & Castro, 2010), have been documented in various regions with characteristics similar to the VVSP sinkhole formation, including the Chapada Diamantina in Bahia State, Brazil. This area is notable for its phreatic zone and extensive karst landscape (over 300 km<sup>2</sup>), comprising both limestone and quartzite formations (Bockmann & Castro, 2010). Regarding *Cambeva*, three species are currently known from the Tibagi River basin—*Cambeva castroi* (de Pinna, 1992), *Cambeva davisi* (Haseman, 1911), and *Cambeva diabola* (Bockmann, Casatti & de Pinna, 2004) (Shibatta et al., 2002; Shibatta et al., 2007; Claro-Garcia et al., 2018; Galves et al., 2007; Hoffmann et al., 2015; Shibatta & Cheida, 2003). The record from VVSP could correspond to one of these known species, or potentially represent an undescribed species—possibly troglobitic—since karst environments such as VVSP often harbor a high diversity of species within the subfamily Trichomycterinae, to which *Cambeva* belongs (Trajano et al., 2009; Rizzato et al., 2012; Rizzato et al., 2014). This discovery paves the way for new avenues of research and will contribute to the advancement of conservation initiatives in this region despite the lack of resolution at the species level, and also indicate the possibility of uncharacterized species, highlighting the need to generate genetic references using local specimens.

Several surveys have investigated the diversity of the Tibagi River basin (Shibatta et al., 2002; Shibatta et al., 2007; Claro-Garcia et al., 2018; Galves et al., 2007; Hoffmann et al., 2015; Shibatta & Cheida, 2003, Dos Reis, et al., 2020). The higher abundance of *Hypostomus* ASVs

in the Tibagi River can be explained by the considerable species richness in the basin, which harbors 13 species (Jarduli et al., 2019). In the case of *Steindachneridion*, this higher occurrence may be attributed to the migratory nature of the species within this genus, which have a broad distribution in the basins where they occur (Paixão et al., 2018). It is likely that the *Steindachneridion* ASVs belong to *Steindachneridion scriptum*, the only species of the genus recorded in the basin (Claro-Garcia et al., 2018). All species reported in the current work have been reported in the basin.

This research represents the first study of this kind conducted in the Tibagi River basin and demonstrates that the 12s ribosomal gene is very effective for species identification in sinkhole environments. Due to their inherent geomorphology, closed environments with difficult access, such as sinkholes, lead to restrictions in monitoring possibilities. Despite the high taxonomic resolution of the 12S gene, a more extensive reference database is essential for ecosystem-wide eDNA metabarcoding. This gap, as highlighted by Jackman et al. (2021), limits the potential of the methods and underscores the importance of expanding the database, particularly for underrepresented neotropical species. In the current work we obtained 60% of ASVs identified with more than 98% of identity with the NCBI nucleotide database; of these, just over half did not require taxonomic adjustments after curation.

### **Conclusion**

Our results highlight the effectiveness of the selected DNA marker (MiFish 12S) for detecting a wide range of taxa, even in the absence of a local reference database. Enclosed environments, such as sinkholes, present unique monitoring challenges due to their complex geomorphology and limited accessibility. The protected Furnas Formation, located within the Vila Velha State Park (VVSP) and composed of ancient sandstone structures, harbors populations of significant evolutionary importance. Given these complexities, the integration of diverse methodologies for biodiversity assessment, both in the Furnas region and in the Tibagi River basin, has contributed to the identification of species diversity for conservation purposes and to supporting future management strategies.

### **References**

- Artoni, R. F., and M. C. Almeida Matiello. 2003. "Genética de Peixes Neotropicais. I. Aspectos da Conservação Genética dos Peixes no Parque Estadual de Vila Velha, Paraná, Brasil." *Publ. UEPG Ci. Biol. Saúde* 9(2): 7–15.
- Artoni, R. F., and O. A. Shibata. 2006. "Catálogo Ilustrado dos Peixes do Parque Estadual de Vila Velha." In *Peixes do Parque Estadual de Vila Velha: Aspectos da História Natural, da*

*Biologia Evolutiva e da Conservação*, edited by R. F. Artoni and O. A. Shibata, 92–137. Ponta Grossa: Editora UEPG.

Beng, K. C., and R. T. Corlett. 2020. “Applications of Environmental DNA (eDNA) in Ecology and Conservation: Opportunities, Challenges, and Prospects”. *Biodiversity and Conservation* 29: 2089–2121.

Bennemann, S. T., A. T. Silva-Souza, and G. R. A. Rocha. 1995. "Composición Ictiofaunística en Cinco Localidades de la Cuenca del Río Tibagi, PR." *Interciencia* 20(1): 7–13.

Blackman, R., Couton, M., Keck, F., Kirschner, D., Carraro, L., Cereghetti, E., ... & Altermatt, F. 2024. “Environmental DNA: The next chapter”. *Molecular ecology*, 33(11), e17355. <https://doi.org/10.1111/mec.17355>

Bockmann, F.A., and R. M. C. Castro. 2010. "The Blind Catfish from the Caves of Chapada Diamantina, Bahia, Brazil (Siluriformes: Heptapteridae): Description, Anatomy, Phylogenetic Relationships, Natural History, and Biogeography." *Neotropical Ichthyology* 8(4): 625–642.

Boyd, S. H., Niemiller, K. D. K., Dooley, K. E., Nix, J., & Niemiller, M. L. 2020. “Using environmental DNA methods to survey for rare groundwater fauna: Detection of an endangered endemic cave crayfish in northern Alabama”. *PLoS One*, 15(12), e0242741. <https://doi.org/10.1371/journal.pone.0242741>

Breitbart, M., M. Kerr, M. J. Schram, I. Williams, G. Koziol, E. Peebles, and C. D. Stallings. 2023. "Evaluation of DNA Metabarcoding for Identifying Fish Eggs: A Case Study on the West Florida Shelf." *PeerJ* 11: e15016.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. "DADA2: High-Resolution Sample Inference from Illumina Amplicon Data." *Nature Methods* 13(7): 581–583.

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. 2009. "BLAST+: Architecture and Applications." *BMC Bioinformatics* 10(1): 421.

Claro-García, A., F. M. Assega, and O. A. Shibatta. 2018. "Diversity and Distribution of Ichthyofauna in Streams of the Middle and Lower Tibagi River Basin, Paraná, Brazil." *Check List* 14(1): 43–53.

Core Nucleotide [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004 – [cited 2025, February, 28]. Available from: <https://www.ncbi.nlm.nih.gov/nucleotide/>

de Melo, M. S., and P. C. F. Giannini. 2007. "Sandstone Dissolution Landforms in the Furnas Formation, Southern Brazil." *Earth Surface Processes and Landforms* 32: 2149–2164.

- de Melo, M. S., P. C. F. Giannini, and L. C. Ruiz Pessenda. 2000. "Gênese e Evolução da Lagoa Dourada, Ponta Grossa, PR." *Revista do Instituto Geológico* 21(1–2): 17–31
- Dos Reis, R. B., A. Frota, G. de Carvalho Deprá, R. R. Ota, and W. J. da Graça. 2020. "Freshwater Fishes from Paraná State, Brazil: An Annotated List, with Comments on Biogeographic Patterns, Threats, and Future Perspectives." *Zootaxa* 4868(4): zootaxa.4868.4.1.
- Fernández, S., S. Rodríguez, J. L. Martínez, Y. J. Borrell, A. Ardura, et al. 2018. "Evaluating Freshwater Macroinvertebrates from eDNA Metabarcoding: A River Nalón Case Study." *PLOS ONE* 13(8): e0201741.
- Galves, W., O. A. Shibatta, and F. C. Jerep. 2007. "Fish, Taquara River Basin, Northern of the State of Paraná, Brazil." *Check List* 3(3): 253–259.
- Garla, A., D. M. Limeira, F. O. Martins, and R. B. Ruiz. 2023. "A Ictiologia do Rio Tibagi-PR: Uma Análise Cienciométrica dos Últimos 10 Anos." *Revista Ciência, Tecnologia e Ambiente* 13: e13247.
- Gealh, A. M. 2007. "Ictiofauna dos Rios Fortaleza, Iapó, Alto e Médio Tibagi." In *Patrimônio Natural dos Campos Gerais do Paraná*, edited by M. S. de Melo, R. S. Moro, and G. B. Guimarães, 151–162. Ponta Grossa: Editora UEPG.
- Gleason, J. E., R. H. Hanner, and K. Cottenie. 2023. "Hidden Diversity: DNA Metabarcoding Reveals Hyper-Diverse Benthic Invertebrate Communities." *BMC Ecology and Evolution* 23: 19.
- Gross, M. C., C. H. Schneider, M. C. de Almeida Matiello, M. de L. Leite, L. A. C. Bertollo, and R. F. Artoni. 2004. "Population Structure, Fluctuating Asymmetry, and Genetic Variability in an Endemic and Highly Isolated *Astyanax* Fish Population (Characidae)." *Genetics and Molecular Biology* 27(4): 529–535.
- Hashemzadeh, S. I., Tabatabaei, S. N., Abdolahi-Mousavi, E., Hernandez, C., Normandeau, E., Laporte, M., ... & Bernatchez, L. 2022. eDNA metabarcoding as a means to assess distribution of subterranean fish communities: Iranian blind cave fishes as a case study. *Environmental DNA*, 4(2), 402-416.
- Herder, J., A. Valentini, E. Bellemain, T. Dejean, J. J. van Delft, P. Thomsen, and P. Taberlet. 2014. "Environmental DNA—A Review of the Possible Applications for the Detection of (Invasive) Species." Netherlands Food and Consumer Product Safety Authority, Nijmegen, The Netherlands.
- Hilário, OH. 2024. BLASTr: An R package for parallel taxonomic classification of metabarcoding sequences. Available at GitHub: <https://github.com/heronoh/BLASTr>

- Hilário, H. O., Mendes, I. S., Guimarães Sales, N., & Carvalho, D. C. 2023. DNA metabarcoding of mock communities highlights potential biases when assessing Neotropical fish diversity. *Environmental DNA*, 5, 1351–1361. <https://doi.org/10.1002/edn3.456>
- Hoffmann, A. C., R. H. C. Nascimento, and O. A. Shibatta. 2015. "Fish Fauna from Tributaries throughout the Tibagi River Basin, Upper Paraná Basin, Brazil." *Check List* 11(6): 1815.
- Jackman, J. M., C. Benvenuto, I. Coscia, et al. 2021. "eDNA in a Bottleneck: Obstacles to Fish Metabarcoding Studies in Megadiverse Freshwater Systems." *Environmental DNA* 3: 837–849.
- Jarduli, L. R., D. A. Z. Garcia, A. P. Vidotto-Magnoni, A. C. R. Casimiro, N. C. Vianna, F. S. de Almeida, ... and M. L. Orsi. 2019. "Fish Fauna from the Paranapanema River Basin, Brazil." *Biota Neotropica* 20(1): e20180707.
- Kerniske, F. F., J. P. Castro, L. E. De la Ossa-Guerra, B. A. Mayer, V. Abilhoa, I. de P. Affonso, and R. F. Artoni. 2021. "Spinal Malformations in a Naturally Isolated Neotropical Fish Population." *PeerJ* 9: e12239.
- Kerniske, F. F., L. E. De la Ossa-Guerra, R. H. Dalcin, I. de P. Affonso, R. F. Artoni, and V. Abilhoa. 2023. "Reproductive Aspects of an Isolated Population of *Psalidodon aff. fasciatus* (Teleostei: Characidae) from Southern Brazil." *Acta Limnologica Brasiliensia* 35: e24.
- Lomolino, M. V. 2004. Conservation biogeography. In Lomolino M.V. & Heaney L.R. (Eds.). *Frontiers of biogeography: new directions in the geography of nature*. Sunderland, Massachusetts: Sinauer Associates; pp. 293–296.
- Mahé, F., T. Rognes, C. Quince, C. de Vargas, and M. Dunthorn. 2015. "Swarm: Robust and Fast Clustering Method for Amplicon-Based Studies." *PeerJ* 3: e1420.
- Martin, M. 2011. "Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads." *EMBnet.journal* 17(1): 10–12.
- Matoso, D. A., C. Martins, R. F. Artoni, and P. M. Galetti Jr. 2010. "Preliminary Qualitative Analysis on mtDNA in *Astyanax fasciatus* Populations Cuvier, 1819 (Teleostei; Characidae) Indicate Population Distinctiveness." *Brazilian Archives of Biology and Technology* 53(3): 663–667.
- Matoso, D. A., M. da Silva, R. F. Artoni, and R. A. Torres. 2013. "Molecular Taxonomy and Evolutionary Hypothesis Concerning *Astyanax fasciatus* (Characiformes, Characidae) from Vila Velha State Park and Tibagi and Iguaçú Rivers." *Genetics and Molecular Research* 12(1): 631–638.
- Matoso, D. A., R. F. Artoni, and P. M. Galetti Jr. 2004. "Genetic Diversity of the Small Characid Fish *Astyanax sp.*, and Its Significance for Conservation." *Hydrobiologia* 527: 223–225.

- McMurdie, P. J., and S. Holmes. 2013. "Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data." *PLOS ONE* 8(4): e61217.
- Mendes, G. A., G. Costa Berger, J. da Gama Affonseca, H. O. Hilário, and D. C. de Carvalho. 2024a. "Protocolo de Coleta de Amostras para eDNA Metabarcoding Usando Filtros Sterivex." *Protocols.io*.
- Mendes, G. A., G. Costa Berger, J. da Gama Affonseca, I. S. Mendes, H. O. Hilário, and D. C. de Carvalho. 2024b. "eDNA Extraction from MCE Filter and Sterivex Cartridge." *Protocols.io*.
- Michels-Souza, M. A., C. W. Hackradt, R. F. Artoni, and G. Sene-Silva. 2011. *A Ictiofauna das Lagoas Adjacentes à Lagoa Dourada e Tarumã no Parque Estadual de Vila Velha, Ponta Grossa, Paraná*. Curitiba: *Inst. Ambient. do Paraná, Curitiba* 1, 1–374
- Milani, E. J., J. H. G. de Melo, P. A. de Souza, L. A. Fernandes, and A. B. França. 2007. "Bacia do Paraná: Estratigrafia e Evolução Tectono-Estratigráfica." *Boletim de Geociências da Petrobras* 15(2): 265–287.
- Miya, M., Y. Sato, T. Fukunaga, T. Sado, J. Y. Poulsen, K. Sato, T. Minamoto, S. Yamamoto, H. Yamanaka, H. Araki, M. Kondoh, and W. Iwasaki. 2015. "MiFish, a Set of Universal PCR Primers for Metabarcoding Environmental DNA from Fishes: Detection of More than 230 Subtropical Marine Species." *R Soc Open Sci* 2(7): 150088.
- Paixão, R. V., J. Ribolli, and E. Zaniboni-Filho. 2018. "Genetic Variation of the Endangered Neotropical Catfish *Steindachneridion scriptum* (Siluriformes: Pimelodidae)." *Frontiers in Genetics* 9: 48.
- R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>.
- Riffel, S. B., P. M. Vasconcelos, I. O. Carmo, and K. A. Farley. 2015. "Combined  $^{40}\text{Ar}/^{39}\text{Ar}$  and (U-Th)/He Geochronological Constraints on Long-Term Landscape Evolution of the Second Paraná Plateau and Its Ruiniform Surface Features, Paraná, Brazil." *Geomorphology* 233: 52–63.
- Rizzato, P. P., & Bichuette, M. E. 2012. A new species of cave catfish from Brazil, *Trichomycterus rubbioli* sp. n., from Serra do Ramalho karstic area, São Francisco River basin, Bahia State (Siluriformes: Trichomycteridae).
- Rizzato, P. P., & Bichuette, M. E. 2014. *Ituglanis boticario*, a new troglomorphic catfish (Teleostei: Siluriformes: Trichomycteridae) from Mambaí karst area, central Brazil. *Zoologia (Curitiba)*, 31, 577-598.

- Sales, N. G., S. Mariani, G. N. Salvador, T. C. Pessali, and D. C. Carvalho. 2018. "Hidden Diversity Hampers Conservation Efforts in a Highly Impacted Neotropical River System." *Frontiers in Genetics* 9: 271.
- Sayers, E. W., E. E. Bolton, J. R. Brister, K. Canese, J. Chan, D. C. Comeau, ... and J. Ostell. 2022. "GenBank." *Nucleic Acids Research* 50(D1): D161–D164.
- Shatskikh AS, Kotov AA, Adashev VE, Bazylev SS, Olenina LV. Functional Significance of Satellite DNAs: Insights from *Drosophila*. *Front Cell Dev Biol*. 2020 May 5; 8:312. doi: 10.3389/fcell.2020.00312. PMID: 32432114; PMCID: PMC7214746.
- Shibatta, O. A. Cheida, C. C. 2003. Composição em tamanho dos peixes (Actinopterygii, Teleostei) de ribeirões da bacia do ribeirão da bacia do ribeirão da bacia do rio Tibagi, Paraná, Brasil.
- Shibatta, O. A., A. M. Gealh, and S. T. Bennemann. 2007. "Ictiofauna dos Trechos Alto e Médio da Bacia do Rio Tibagi, Paraná, Brasil." *Biota Neotropica* 7(2): 125–134.
- Shibatta, O. A., and R. F. Artoni. 2005. "Sobre a Identidade das Populações Alopátricas de *Astyanax* (Characiformes, Characidae) das Formações Furna 1 Ponta Grossa, Paraná, Brasil." *Publicatio UEPG* 11(2): 7–12.
- Shibatta, O.A., Orsi, M.L., Bennemann, S.T. and Silva-Souza, A.T. 2002. Diversidade e distribuição de peixes na bacia do rio Tibagi. In A bacia do rio Tibagi (M.E. Medri, E. Bianchini, O.A. Shibatta & J.A. Pimenta, eds.). Londrina, M. E. Medri, p.403-423.
- 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). *Zoological Journal of the Linnean Society*, 190(4), 1217–1234.
- Trajano, E., Secutti, S., & Bichuette, M. E. 2009. Natural history and population data of fishes in caves of the Serra do Ramalho karst area, Middle São Francisco basin, northeastern Brazil. *Biota Neotropica*, 9, 129-133.
- Tsuji, S., R. Inui, R. Nakao, et al. 2022. "Quantitative Environmental DNA Metabarcoding Shows High Potential as a Novel Approach to Quantitatively Assess Fish Community." *Scientific Reports* 12: 21524.
- Yao, M., S. Zhang, Q. Lu, X. Chen, S. Y. Zhang, Y. Kong, and J. Zhao. 2022. "Fishing for Fish Environmental DNA: Ecological Applications, Methodological Considerations, Surveying Designs, and Ways Forward." *Molecular Ecology* 31(20): 5132–5164.

Yates, M. C., D. J. Fraser, and A. M. Derry. 2019. "Meta-Analysis Supports Further Refinement of eDNA for Monitoring Aquatic Species-Specific Abundance in Nature." *Environmental DNA* 1: 5–13.

Xiong, F., L. Shu, H. Zeng, X. Gan, S. He, and Z. Peng. 2022. "Methodology for fish biodiversity monitoring with environmental DNA metabarcoding: The primers, databases and bioinformatic pipelines". *Water Biology and Security* 1(1).

### Figure legends

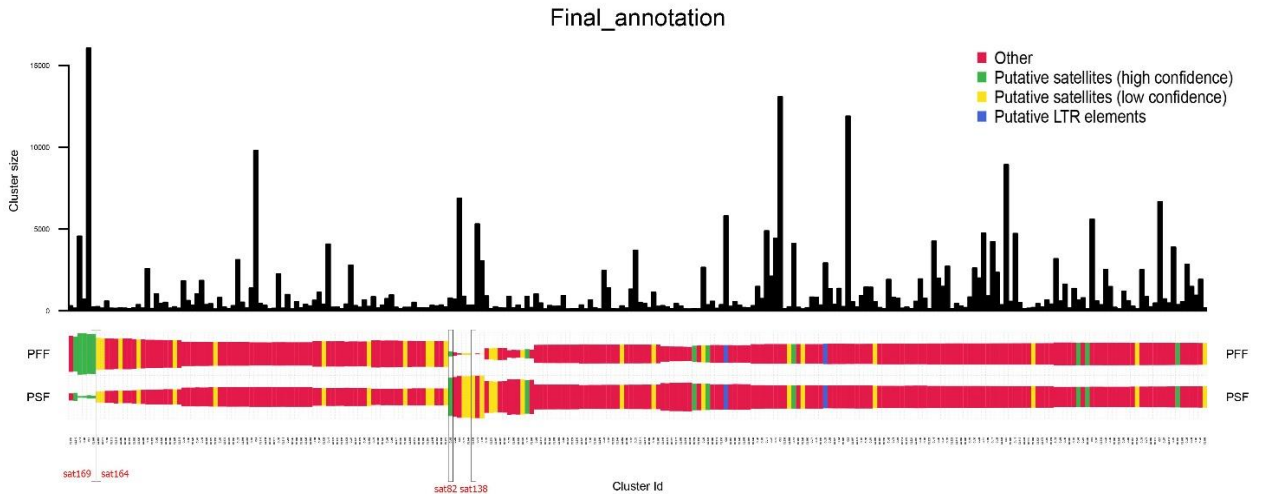
Figure 1. Geographic location and geological context of the study sites. A. Map showing the spatial distribution of the six aquatic environments sampled for eDNA analysis in southern Brazil. A. The left panel indicates the broader region, with the two Tibagi River sites (orange and yellow circles) and their relative positions within the river basin. The right panel provides a detailed satellite view highlighting the locations of three sinkholes (Sinkholes 1–2–4), and Dourada Lake. B. Conceptual cross-sectional diagram of the karstic landscape, illustrating the relative depth and spatial arrangement of the sinkholes, Dourada Lake, and Tibagi River.

Figure 2. Results of eDNA metabarcoding analysis for fish detection across the different sites. A. ASV absolute abundance of all classes using the 12s gene. B. Numbers of ASVs, MOTUs, and assigned taxa for all sites. C. Number of taxonomic levels identified in each site. D. Percentage of ASV absolute abundance of taxa in each site. Bars represent the values for each category.

**ANEXO E – Elementos transponíveis escolhidos do resultado de Repeat Explorer para *P. fasciatus***

<b>Cluster</b>	<b>Similarity based annotation RE</b>	<b>Blast FishTEDB (best match)</b>	<b>Contig</b>	<b>Fragment size</b>	<b>number of reads in the contig</b>
<b>CL83</b>	Class I/DIRS	cavefish_KB872397.1_76277#LINE/L1	19	1486	308
<b>CL84</b>	Class_I/ LINE: LINE-RT	SRSE01000008.1_330692#LINE/Rex	53	1646	392
<b>CL90</b>	Class_I/ LINE: LINE-RT	cavefish_rnd-4_family-1091#LINE/L2	27	2291	569
<b>CL101</b>	31.12% Class_I/ LINE: LINE-RT	sgall_gbLCYQ01S000404.1_274964#LINE/Rex	2	2156	685
<b>CL108</b>	18.42% Class_I/ LINE: LINE-RT	NC_043762.1_732032#LINE/L1	1	3980	816
<b>CL257</b>	85.14% Class_I/ LINE: LINE-RT	cavefish_rnd-5_family-1128#LINE/Rex-Babar	1	1064	113
<b>CL255</b>	59.87% Class_II/Subclass_2/ Helitron: Helitron-HEL2	Zebrafish_Mariner-15_DR#DNA/TcMar	3	1065	91
<b>CL227</b>	45.08% Class_I/ LINE: LINE-ENDO	SRSE01001373.1_53753#LINE/Rex	7	737	87
<b>CL169</b>	11.46% Class_I/ LINE: LINE-RT	cavefish_rnd-1_family-551#LINE/Rex-Babar	2	734	209
<b>CL131</b>	18.52% Class_I/ LINE: LINE-RT	cavefish_KB872250.1_230291#LINE/R2	12	1495	270
<b>CL114</b>	18.18% Class_I/ LINE: LINE-RT	cavefish_rnd-5_family-1128#LINE/Rex-Babar	19	1536	374

## ANEXO F- Análises comparativo Furna vs São Francisco



Os satélites marcados nos quadros foram os que resultaram em marcação positiva na FISH.

PFF: População furna; PSF: População São Francisco