



UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Departamento de Genética, Ecologia e Evolução
Programa de Pós-graduação em Genética

Daniel Fonseca Teixeira

**ICHTHYOPLANKTON DNA METABARCODING AS A TOOL
FOR LARGE-SCALE INVENTORY AND REPRODUCTIVE
DYNAMICS ANALYSIS OF THE ICHTHYOFAUNA**

Belo Horizonte
2024

Daniel Fonseca Teixeira

**ICHTHYOPLANKTON DNA METABARCODING AS A TOOL FOR LARGE-
SCALE INVENTORY AND REPRODUCTIVE DYNAMICS ANALYSIS OF THE
ICHTHYOFAUNA**

Tese apresentada ao Programa de Pós-Graduação em Genética da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Doutor em Genética.

Orientador: Prof. Dr. Daniel Cardoso de Carvalho

Belo Horizonte

2024

043

Teixeira, Daniel Fonseca.

Ichthyoplankton DNA metabarcoding as a tool for large-scale inventory and reproductive dynamics analysis of the ichthyofauna [manuscrito] / Daniel Fonseca Teixeira. – 2024.

76 f. : il. ; 29,5 cm.

Orientador: Prof. Dr. Daniel Cardoso de Carvalho.

Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Genética. 2. Taxonomia. 3. Plâncton. 4. Código de Barras de DNA Taxonômico. I. Carvalho, Daniel Cardoso de. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 575



UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Genética

ATA DE DEFESA DE TESE

ATA DA DEFESA DE TESE	184/2024 entrada 1º/2019 CPF: 066.805.096-96
Daniel Fonseca Teixeira	

Às nove horas do dia **31 de janeiro de 2024**, reuniu-se a Comissão Examinadora de Tese, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: "**Ichthyoplankton DNA metabarcoding as a tool for large scale inventory and reproductive dynamics analysis of the ichthyofauna**", requisito para obtenção do grau de Doutor em **Genética**. Abrindo a sessão, o Presidente da Comissão, **Daniel Cardoso de Carvalho**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

Prof./Pesq.	Instituição	CPF	Indicação
Daniel Cardoso de Carvalho	PUC Minas	034.375.066-09	APROVADO
Fernanda Antunes Carvalho	UFMG	045.665.426-70	APROVADO
Bruno Francelino de Melo	CUNY/NY	373.764.238-76	APROVADO
Gabriel de Menezes Yazbeck	UFSJ	150.609.338-81	APROVADO
Jonathan Stuart Ready	UFPA	745.826.071-87	APROVADO

Pelas indicações, o candidato foi considerado: APROVADO

O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 31 de janeiro de 2024.

Daniel Cardoso de Carvalho (Orientador)- (PUC-Minas)

Bruno Francelino de Melo (Columbia University e CUNY/NY)

Fernanda Antunes Carvalho (UFMG)

Gabriel de Menezes Yazbeck (UFSJ)

Jonathan Stuart Ready (UFPA)

Assinaturas dos membros da banca:



Documento assinado eletronicamente por **Gabriel de Menezes Yazbeck, Usuário Externo**, em 31/01/2024, às 13:02, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Bruno Francelino de Melo, Usuário Externo**, em 31/01/2024, às 15:58, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Daniel Cardoso de Carvalho, Usuário Externo**, em 31/01/2024, às 16:06, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jonathan Stuart Ready, Usuário Externo**, em 01/02/2024, às 09:44, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Fernanda Antunes Carvalho, Professora do Magistério Superior**, em 01/02/2024, às 15:31, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2999283** e o código CRC **DA92FD41**.



UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Genética

FOLHA DE APROVAÇÃO

"Ichthyoplankton DNA metabarcoding as a tool for large scale inventory and reproductive dynamics analysis of the ichthyofauna "

Daniel Fonseca Teixeira

Tese aprovada pela banca examinadora constituída pelos Professores:

Daniel Cardoso de Carvalho
PUC Minas

Gabriel de Menezes Yazbeck
UFSJ

Fernanda Antunes Carvalho
UFMG

Jonathan Stuart Ready
UFPA

Bruno Francelino de Melo
Columbia University e CUNY/NY

Belo Horizonte, 31 de janeiro de 2024.



Documento assinado eletronicamente por **Gabriel de Menezes Yazbeck, Usuário Externo**, em 31/01/2024, às 13:03, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Bruno Francelino de Melo, Usuário Externo**, em 31/01/2024, às 16:00, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Daniel Cardoso de Carvalho, Usuário Externo**, em 31/01/2024, às 16:06, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Jonathan Stuart Ready, Usuário Externo**, em 01/02/2024, às 09:44, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Fernanda Antunes Carvalho, Professora do Magistério Superior**, em 01/02/2024, às 15:31, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2999962** e o código CRC **020CAD88**.

AGRADECIMENTOS

Ao meu orientador Daniel Cardoso de Carvalho que, desde 2014, tem me ajudado a desenvolver profissional e pessoalmente, e por ter me creditado a oportunidade e confiança para desenvolver este trabalho.

À FAPEMIG pela bolsa de doutorado, e ANEEL e CEMIG (P&D0624) pelo financiamento do projeto.

A todos os colegas, professores e funcionários do Programa de Pós-Graduação em Genética da UFMG. Especialmente aos professores Anderson Myoshi e Frederico Soriani e às secretárias Raíssa Campos e Izabella Gama, sempre disponíveis para ajudar a resolver qualquer dúvida ou problema.

A todos os professores e funcionários do Programa de Pós-Graduação em Biologia de Vertebrados da PUC Minas, que permitiu o desenvolvimento deste trabalho. Em especial ao professor Gilmar Santos, pelos ensinamentos e coordenação do projeto, e à Cledma e ao Rogério, pela solicitude sempre que necessário.

Aos professores Bruno Melo, Fabrício Santos, Fernanda Almeida, Fernanda Carvalho, Gabriel Yazbeck e Jonathan Ready, pela disponibilidade para participar da banca examinadora e pelas contribuições.

À Alyce, Amanda, Laura, Mara, Mari e ao Guilherme e Gustavo, pela amizade, parceria e por fazerem o dia-a-dia mais agradável e animado durante todo o período do desenvolvimento desse projeto, inclusive durante os tempos nebulosos de pandemia.

À Turma da Inbonha que, mesmo nem sempre presentes fisicamente, continuam sendo parte fundamental da minha vida.

A todos os membros atuais e egressos do Laboratório da Genética da Conservação (LGC) e do Grupo de Pesquisa em Ecologia e Peixes de Água Doce (GPEPAD) que fizeram parte do meu caminho e contribuíram para o meu conhecimento e desenvolvimento.

À Bela, pelo apoio, companheirismo, e desenvolvimento conjunto nos últimos 10 anos, e pelo amor e carinho nos últimos quatro. Que nossa admiração e respeito mútuos continuem por muitos novos ciclos.

À família da Bela pelo carinho e acolhimento. Principalmente à Ana, Victor, Adilson, Dani, Laura e Frida, que me permitiram sentir que tenho uma nova casa longe de casa.

À minha família pelo amor e apoio sempre incondicionais. Especialmente aos meus pais Cidinha e Geraldinho, meus principais incentivadores e moldadores do meu caráter, e aos meus irmãos Joel e Estela por me ajudarem a sempre tentar me tornar uma pessoa melhor. Aos meus tios-pais Humberto e Zélia, Luiza e Ronaldo, e Marisa e meus primos-irmãos André, Diêgo, Gabriela, Isabela, Rafael, Raquel e Renato.

Ao monumental Clube Atlético Mineiro, pelas experiências mágicas que me proporcionam tamanha felicidade, e as não tão mágicas, que me proporcionam reflexão.

RESUMO

O estudo do ictioplâncton é fundamental para entender a dinâmica reprodutiva de assembleias de peixes locais, mas a sua identificação morfológica pode ser desafiadora mesmo para taxonomistas experientes. Apesar de ser uma ferramenta poderosa para fornecer identificação taxonômica assertiva, o DNA *barcoding* se torna trabalhoso e caro no inventariamento em larga escala. O DNA *metabarcoding* pode superar essas limitações, sendo uma alternativa rápida, de melhor custo-benefício, abrangente e precisa para a identificação simultânea de múltiplos indivíduos, mas a padronização de protocolos eficientes e acessíveis ainda pode ser um desafio para sua utilização recorrente. Neste trabalho, aplicamos o método de DNA *metabarcoding* para a identificação molecular de amostras de ictioplâncton da região de influência da barragem de Três Marias, no rio São Francisco. Por meio da identificação molecular, temos o objetivo de responder as seguintes perguntas: (a) os marcadores *minibarcodes* MiFish e NeoFish do gene conservado 12S rRNA conseguem detectar e identificar as espécies amostradas tão bem quanto o marcador do gene COI padronizado para o DNA *barcoding* utilizando o sequenciamento de nova geração (NGS)? (b) qual a contribuição relativa de ovos e larvas para a riqueza total encontrada; (c) se existe diferença significativa entre as comunidades de ictioplâncton encontradas à montante e à jusante de Três Marias; (d) se algum dos tributários analisados pode ser considerado como ponto prioritário para conservação. De maneira geral, os resultados indicam que os marcadores minibarcodes podem ser tão eficientes quanto o COI para a detecção de espécies, quando um banco de referência completo e confiável está disponível. Além disso, apesar de a análise estatística evidenciar baixa diferenciação entre as comunidades de ictioplâncton encontradas à jusante e à montante, o rio Paracatu, localizado à jusante, foi o único onde todas as espécies consideradas grandes migradoras foram detectadas, se mostrando como um tributário prioritário para conservação. Assim, este estudo pretende não apenas contribuir para o avanço de técnicas de inventariamento de alta precisão e rendimento, mas também contribuir com a elucidação de questões sobre a dinâmica reprodutiva dos peixes na região do reservatório de Três Marias que são fundamentais para guiar planos de conservação e manejo da ictiofauna.

Palavras-chave: Taxonomia molecular, 12S rRNA, Sequenciamento de DNA de alto rendimento

ABSTRACT

The study of ichthyoplankton is paramount to understand the reproductive dynamics of local fish assemblages, but its morphological identification is often challenging even to experienced taxonomists. Despite being a powerful tool for assertive taxonomic identification, DNA barcoding becomes too expensive and laborious in large-scale inventories. DNA *metabarcoding* can overcome these limitations as a rapid, cost-effective, broad and accurate taxonomy alternative for the identification of multiple individuals simultaneously, but the standardization of accessible and efficient protocols still poses a challenge for its recurrent usage. With that in mind, we applied the DNA metabarcoding method to identify ichthyoplankton samples collected at the area impacted by the Três Marias reservoir, São Francisco River Basin, Southeastern Brazil. Using molecular identification, we aim to elucidate four major questions: (a) can 12S rRNA gene minibarcode markers MiFish and NeoFish detect and identify the sampled species as accurately as the COI gene marker standardized for DNA barcoding, using next generation sequencing (NGS)? (b) what is the relative contribution of eggs and larvae to total observed richness? (c) is there significant difference between upstream and downstream Três Marias ichthyoplankton communities? (d) should any studied tributary be considered as a conservation priority hotspot? Overall, our results suggest that the minibarcode markers can be as efficient as COI for species detection, when a reliable and complete database is available. Furthermore, despite the statistical insignificant and low differentiation found between upstream and downstream ichthyoplankton communities, only in the Paracatu River, located downstream from the reservoir, all the greater migratory species were detected, highlighting this tributary as a primary hotspot for conservation. Therefore, this study aims at contributing not only by enhancing large-scale and precise inventory techniques, but also by elucidating issues about the fish reproductive dynamics in the area impacted by Três Marias reservoir that are pivotal for guiding ichthyofauna conservation and management.

Keywords: Molecular taxonomy, 12S rRNA, High-throughput DNA sequencing

Sumário

1. INTRODUÇÃO GERAL	13
2. JUSTIFICATIVA	17
3. OBJETIVO GERAL	17
4. CAPÍTULO 1: DNA metabarcoding assessment of Neotropical ichthyoplankton communities is marker-dependent (Ecology and Evolution, 2023)	19
4.1. Objetivos Específicos	19
4.2. Introduction	22
4.3. Materials and Methods	24
4.3.1. Sample collection	24
4.3.2. Genomic DNA extraction.....	24
4.3.3. DNA amplification	25
4.3.4. Library preparation and DNA sequencing	25
4.3.5. Bioinformatic analyses	26
4.4. Results	27
4.5. Discussion	32
REFERENCES	35
SUPPORTING INFORMATION	40
5. CAPÍTULO 2: Inventariamento em larga escala do ictioplâncton utilizando DNA metabarcoding para análise da dinâmica reprodutiva da ictiofauna	42
5.1. Objetivos Específicos	42
5.2. Introdução	45
5.3. Materiais e Métodos	46
5.3.1. Coleta e Manuseio das Amostras.....	46
5.3.2. Extração do DNA Genômico.....	48
5.3.3. Amplificação para o DNA <i>barcoding</i>	49
5.3.4. Amplificação para o DNA <i>metabarcoding</i>	49
5.3.5. Preparo de Bibliotecas e Sequenciamento de DNA	51
5.3.6. Análises de Bioinformática	51
5.4. Resultados	53
5.4.1. Triagem.....	53
5.4.2. Identificação molecular	56
5.5. Discussão	63
REFERÊNCIAS	66
6. CONSIDERAÇÕES FINAIS	70

REFERÊNCIAS	72
--------------------------	-----------

1. INTRODUÇÃO GERAL

O estudo da composição, abundância e distribuição do ictioplâncton é fundamental para o entendimento da dinâmica reprodutiva das assembleias de peixes de uma região (MARIAC et al., 2018). A análise desses parâmetros permite a identificação dos locais de desova, dos berçários onde ocorre o recrutamento (i.e., desenvolvimento das larvas), das rotas de migração, dos padrões de variação temporal e espacial e das diferenças na dinâmica reprodutiva de peixes migradores e não-migradores (BAUMGARTNER et al., 2004; BIALETZKI et al., 2005; REYNALTE-TATAJE et al., 2012). Essas informações contribuem para a elucidação da influência das alterações ambientais antropogênicas na reprodução dos peixes e para o delineamento de ações de manejo eficientes para a conservação das espécies e consequente manutenção da reprodução e da pesca (DA SILVA et al., 2015; SILVA et al., 2017).

Para estudar a composição do ictioplâncton, tradicionalmente utiliza-se a identificação morfológica de larvas através da técnica de sequência regressiva de desenvolvimento, baseada na comparação morfológica de indivíduos menores a partir da forma juvenil conhecida (AHLSTROM & MOSER, 1976; NAKATANI et al., 2001). Ou seja, as características conhecidas para o organismo em um estágio mais avançado, como padrão de pigmentação e forma do corpo, sequência de formação e posição das nadadeiras, características merísticas e morfométricas, são utilizadas como referência para identificar indivíduos mais jovens, à medida que essas características vão surgindo. No entanto, o tamanho reduzido dos indivíduos, sua fragilidade e as transformações acentuadas durante o desenvolvimento podem dificultar a identificação nos primeiros estágios de vida (BIALETZKI et al., 2016). Além disso, devido à ausência de características morfológicas diagnósticas na maioria dos ovos, alguns autores optam pela identificação apenas de larvas, cuja taxonomia em estágios iniciais também é limitada, principalmente quando elas foram danificadas durante a coleta ou triagem (BAUMGARTNER et al., 2004; REYNALTE-TATAJE et al., 2012).

Apesar de a morfologia tradicional ser indispensável para a construção de bancos de dado confiáveis, a acurácia da identificação pode variar entre diferentes taxonomistas e laboratórios, de acordo com sua experiência e especialidade (KO et al., 2013). Essas limitações podem comprometer o levantamento de informações essenciais para a conservação das regiões estudadas (NOBILE et al., 2019). Desse modo, para aumentar a precisão e confiabilidade da identificação do ictioplâncton, pesquisadores têm recorrido a técnicas moleculares. Uma das técnicas moleculares mais utilizadas atualmente para a identificação taxonômica é o DNA

barcoding, que se baseia na utilização de um fragmento padronizado do genoma para comparar a sequência de DNA do organismo alvo com a de outros previamente identificados (HEBERT et al., 2003). Em um trabalho englobando 1.088 espécies, WARD (2009) demonstrou a alta confiabilidade da identificação de peixes utilizando o DNA *barcoding*, utilizando um fragmento de aproximadamente 650 pares de base (pb) do gene citocromo c oxidase subunidade I (COI) (WARD et al. 2005). Em se tratando de estudos com ictioplâncton, trabalhos comparativos demonstraram que a taxonomia molecular utilizando DNA *barcoding* é mais eficiente que a identificação morfológica tradicional, sendo capaz de não apenas chegar a níveis taxonômicos inferiores, mas também de corrigir identificações morfológicas equivocadas (KO et al., 2013; BECKER et al., 2015).

Segundo uma revisão recente feita por LIRA et al. (2023), foram publicados 70 artigos científicos utilizando DNA *barcoding* para a identificação do ictioplâncton, dos quais 44 foram realizados com espécies marinhas, entre 2005 e 2022, e apenas 26 com espécies continentais, entre 2011 e 2022. Dentre eles, apenas seis foram realizados no Brasil (BECKER et al., 2015; FRANTINE-SILVA et al., 2015; RODRIGUES et al., 2017; ALMEIDA et al., 2018; LIMA et al., 2020; FREITAS-SOUZA et al., 2022), apesar de sua reconhecida megadiversa ictiofauna.

Em um estudo comparativo, BECKER et al. (2015) identificaram morfológicamente ovos apenas como sendo de espécies migradoras ou não, quando possível, enquanto através do DNA *barcoding* conseguiram identificá-los em nível de espécie e apontar imprecisões na identificação morfológica mesmo nessa análise mais ampla. Além disso, também foi possível identificar molecularmente em nível de espécie larvas danificadas que não puderam ser identificadas através da morfologia. Já FRANTINE-SILVA et al. (2015) conseguiram identificar molecularmente em nível de espécie mais de 99% das 536 amostras de ictioplâncton analisadas, incluindo amostras de ovos, que foram responsáveis por 30% da riqueza de espécies observada. No trabalho desenvolvido por RODRIGUES et al. (2017), O DNA *barcoding* possibilitou a detecção inédita de ovos e larvas de três espécies de peixes-de-bico na costa sudeste do Brasil. Esta técnica foi empregada por ALMEIDA et al. (2018) para demonstrar a sua aplicação na detecção de espécies potencialmente invasoras nos estágios iniciais de vida, permitindo a identificação de ovos e larvas de 13 espécies não nativas. Já LIMA et al. (2020) usou a identificação molecular para identificar ovos coletados em tributários da bacia do rio Paranapanema, para reconhecimento de áreas de desova de espécies nativas e ameaçadas em uma região impactada por barramento. Por fim, FREITAS-SOUZA et al. (2022) utilizaram o

DNA *barcoding* em um estudo que demonstrou a ocorrência da reprodução de híbridos de *Pseudoplatystoma* em regiões naturais do sudeste brasileiro.

Dessa forma, a identificação molecular tem se mostrado útil em diversos estudos envolvendo ictioplâncton. No entanto, apesar da sua grande precisão taxonômica, a técnica do DNA *barcoding* depende do processamento e sequenciamento individual de cada organismo, sendo trabalhoso e caro para o inventariamento em larga escala (TABERLET et al., 2012; YU et al., 2012), como ocorre em muitos trabalhos ecológicos com ovos e larvas de peixes (MARIAC et al., 2018; NOBILE et al., 2019).

Uma abordagem que tem ganhado espaço como alternativa para a identificação molecular de organismos em larga escala é o DNA *metabarcoding*. O método combina a taxonomia baseada em DNA *barcodes* e o sequenciamento de nova geração (NGS) de alta performance para identificar vários espécimes simultaneamente. As análises utilizando *metabarcoding* são econômicas, de rápida produção, abrangentes, pouco dependentes de perícia taxonômica e seus bancos de dados permanecem disponíveis para futuras verificações (TABERLET et al., 2012; YU et al., 2012). Por isso, a técnica vem se destacando por permitir o acesso em massa à biodiversidade (YU et al., 2012). Essa abordagem tem permitido a reconstrução de comunidades ancestrais (JØRGENSEN et al., 2012), o monitoramento da biodiversidade (ANDERSEN et al., 2012), e a detecção de um número maior de unidades taxonômicas operacionais em uma fração do tempo gasto em estudos convencionais baseados em morfologia e DNA *barcoding* (FONSECA et al., 2010).

O DNA *metabarcoding* também tem se mostrado eficiente nos estudos de ictioplâncton, permitindo identificar a composição de múltiplas larvas ou ovos agrupados em uma única amostra (amostra em *bulk*) de maneira precisa e confiável (KIMMERLING et al., 2018; MARIAC et al., 2018). No entanto, sua aplicação nesse contexto ainda está incipiente, e apenas sete trabalhos haviam publicados até a revisão feita por LIRA et al. (2023), sendo que apenas um deles foi realizado com os dois estágios de desenvolvimento, enquanto os demais utilizaram apenas ovos ou larvas. Dentre eles, o único realizado no Brasil fora conduzido utilizando amostras controladas (*mock samples*) e com um número reduzido de larvas (NOBILE et al., 2019).

Mais recentemente, em 2023, foram publicados três novos trabalhos utilizando DNA *metabarcoding* para análise do ictioplâncton em bacias brasileiras (POMPEU et al., 2023; SILVA et al., 2023; TEIXEIRA et al., 2023). No trabalho de POMPEU et al. (2023), os pesquisadores utilizaram o DNA *metabarcoding* para identificar ovos e analisar se a frequência

de coleta utilizada atualmente nos estudos nacionais com ictioplâncton é adequada para detectar todos os eventos de desova de peixes migradores no alto da bacia do rio São Francisco. Na bacia Amazônica, SILVA et al. (2023) utilizaram ovos e larvas para identificar tributários à montante da barragem de Teles Pires importantes para a manutenção da reprodução de espécies migradoras nessa região. Já no estudo de TEIXEIRA et al. (2023), fruto da presente pesquisa, foram analisados possíveis vieses na escolha de marcadores para o DNA *metabarcoding* de ovos e larvas de uma ictiofauna complexa, com grande diversidade e ocorrência várias espécies congêneres, como a do São Francisco.

Diferentemente de amostras ambientais (e.g. sedimento ou água), cujo material genético normalmente se encontra degradado, amostras de tecidos agrupados resultam em DNA de melhor qualidade, possibilitando a amplificação de sequências de DNA maiores (TABERLET et al., 2012). No entanto, as plataformas de NGS acessíveis à maioria dos laboratórios limitam o sequenciamento de fragmentos a até 600 (2 x 300) pb, impossibilitando a utilização adequada de marcadores tradicionais para o DNA *barcoding*, como o fragmento de 650 pb padronizado para o gene COI. Além disso, a variabilidade do gene COI dificulta o desenho de *primers* (iniciadores) internos para a amplificação de *minibarcodes* e, por isso, ele tem sido preterido em favor de genes mais conservados (DEAGLE et al., 2014), apesar de possuir um banco de referência mais completo. Dentre os genes mais conservados, o 12S rRNA tem ganhado destaque para o *metabarcoding* de peixes (MILAN et al., 2020; MIYA et al., 2020; SALES et al., 2021).

Além da escolha do marcador com especificidade e sensibilidade adequadas, outro desafio é sua utilização para a análise quantitativa do ictioplâncton, ou seja, a abundância relativa de cada espécie na amostra. No entanto, alguns fatores podem enviesar a quantidade de leituras (*reads*) obtidas para cada indivíduo ou espécie, como número de mitocôndrias por célula, espécimes com diferentes tamanhos corporais na mesma amostra, ou diferenças na afinidade dos iniciadores utilizados devido a pareamentos imperfeitos (*mismatches*) (FONSECA, 2018). Ainda assim, alguns trabalhos têm mostrado uma relação positiva da quantidade de ovos ou larvas incluída em cada amostra e o número de leituras obtidas através do *metabarcoding* ao analisar amostras controladas (DUKE & BURTON, 2020; NOBILE et al., 2019).

No presente trabalho, nos propomos a aplicar o DNA *metabarcoding* para análise da composição do ictioplâncton, utilizando como modelo os peixes da bacia do rio São Francisco. Além disso, testaremos a sensibilidade e resolução de dois marcadores para o gene 12S em

comparação às de marcadores tradicionais utilizados para o gene COI. Assim, este estudo contribui para o aprimoramento do método para sua utilização em estudos ecológicos da dinâmica reprodutiva da ictiofauna e para o desenho de estratégias de manejo e conservação adequadas para garantir a preservação de seções das bacias fundamentais para a manutenção da reprodução dos peixes e, conseqüentemente, da pesca.

2. JUSTIFICATIVA

A identificação de ovos e larvas de peixes é fundamental para determinar os locais de desova, recrutamento e entender os padrões de variação temporal e espacial da dinâmica reprodutiva da ictiofauna. No entanto, a identificação morfológica muitas vezes fica comprometida devido à ausência de caracteres diagnósticos nos estágios iniciais da vida desses organismos. Como o estudo do ictioplâncton demanda a análise de um grande número de organismos, o DNA *barcoding* se torna caro e trabalhoso, uma vez que ele depende da manipulação de cada organismo individualmente. Assim, é preciso aprimorar e aplicar de forma prática o método do DNA *metabarcoding*, que permite o agrupamento de vários organismos em uma única amostra para processamento conjunto, visando resolver esses impasses e possibilitar o levantamento de informações adequado para guiar os planos de manejo e conservação da ictiofauna.

O presente trabalho busca avançar o conhecimento e demonstrar aplicações práticas para o uso do DNA *metabarcoding* para a identificação do ictioplâncton em estudos reprodutivos, ao fazer comparações entre diferentes marcadores para analisar uma ictiofauna diversa, e utilizar um grande número de ovos e larvas de diferentes trechos à montante e à jusante da barragem de Três Marias, uma região de grande importância ecológica e econômica. Assim, este estudo busca se juntar aos trabalhos pioneiros de NOBILE et al. (2019), POMPEU et al. (2023) e SILVA et al. (2023) como um dos pilares para a aplicação do DNA *metabarcoding* para estudar a reprodução da megadiversa ictiofauna do Brasil e, conseqüentemente, de outras regiões menos diversas.

3. OBJETIVO GERAL

O objetivo deste trabalho é empregar e aprimorar protocolos para a abordagem do DNA *metabarcoding* na análise da dinâmica reprodutiva da ictiofauna e demonstrar sua aplicação

prática em uma região antropogenicamente impactada e de grande diversidade e importância ecológica e econômica.

4. CAPÍTULO 1: DNA metabarcoding assessment of Neotropical ichthyoplankton communities is marker-dependent (Ecology and Evolution, 2023)

4.1. Objetivos Específicos

- Avaliar a sensibilidade e a resolução de marcadores dos genes 12S e COI para a detecção e identificação de espécies a partir de amostras de larvas ou ovos de peixes agrupados;
- Identificar os marcadores mais adequados para o *metabarcoding* do ictioplâncton;
- Verificar a possibilidade de utilização do DNA *metabarcoding* para análise quantitativa do ictioplâncton;

DNA metabarcoding assessment of Neotropical ichthyoplankton communities is marker-dependent

Daniel Fonseca Teixeira^{1,2}, Heron Oliveira Hilário¹, Gilmar Bastos Santos¹, Daniel Cardoso Carvalho¹

¹ Post-Graduate Program in Vertebrate Biology, Pontifical Catholic University of Minas Gerais, PUC Minas, Belo Horizonte, Brazil

² Post-Graduate Program in Genetics, Federal University of Minas Gerais, Belo Horizonte, Brazil

Correspondence

Daniel Cardoso Carvalho, Post-Graduate Program in Vertebrate Biology, Pontifical Catholic University of Minas Gerais, PUC Minas, Belo Horizonte 30535-610, Brazil. Email: danielcarvalho@pucminas.br and carvalho.lgc@gmail.com

Funding information

ANEEL/CEMIG, Grant/Award Number: P&D0624

ABSTRACT

The study of ichthyoplankton is paramount to understanding fish assemblages' reproductive dynamics. DNA metabarcoding has been applied as a rapid, cost-effective, and accurate taxonomy tool, allowing the identification of multiple individuals simultaneously. However, there remain significant challenges when using DNA metabarcoding, such as molecular marker choice according to the taxonomic resolution and length of the fragment to be sequenced, primer bias, incomplete reference databases, and qualitative inference incongruences. Here, 30 ichthyoplankton pools collected from a Neotropical River were identified at a molecular level using DNA metabarcoding to compare the resolution, sensibility, specificity, and relative read abundance (RRA) recovery of three molecular markers: the standard COI fragment (650 pb, with each end analyzed individually) and two short 12S rRNA genes markers (\cong 200 bp – NeoFish and MiFish markers). The combined use of the three markers increased the genera detection rates by 25%–87.5%, allowing an increased taxonomic coverage and robust taxonomic identification of complex Neotropical ichthyoplankton communities. RRA is marker-dependent, indicating caution is still needed while inferring species abundance based on DNA metabarcoding data when using PCR-dependent protocols.

Keywords: 12S rRNA, COI, high-throughput DNA sequencing, molecular taxonomy

Taxonomy Classification: Genetics

Ecology and Evolution. 2023;13:e10649.

<https://doi.org/10.1002/ece3.10649>

4.2. Introduction

The study of ichthyoplankton composition, abundance, and distribution is pivotal for understanding the reproductive dynamics of local fish assemblages (Mariano et al., 2018). The analysis of these parameters allows the identification of spawning sites, nursery areas where recruitment occurs, migration routes, temporal and spatial pattern variations, and differences in the reproduction patterns of migratory and nonmigratory fish (Baumgartner et al., 2004; Bialecki et al., 2005; Reynalte-Tataje et al., 2012). This information is instrumental in elucidating the influence of anthropogenic environmental alterations on fish reproduction and in the definition of effective management actions for species conservation and, consequently, fishing stock maintenance (Da Silva et al., 2015; Silva et al., 2017).

Traditionally, ichthyoplankton taxonomy has applied the regressive development sequence technique, based on the morphological comparison of younger larvae with previously identified juveniles (Ahlstrom & Moser, 1976; Nakatani et al., 2001). However, due to the absence of morphological diagnostic characters during the egg stage, some authors exclude them from the studies and resort to identifying exclusively larvae, which in the initial stages is also a difficult task (Baumgartner et al., 2008; Reynalte-Tataje et al., 2012). Moreover, the accuracy of the traditional morphological identification can diverge between taxonomists and laboratories, according to their experience and specialty (Ko et al., 2013). These limitations can compromise surveying essential information to conserve the areas of interest (Nobile et al., 2019).

Studies have employed molecular techniques to strengthen the precision and reliability of ichthyoplankton taxonomy. Comparative investigations have demonstrated that molecular taxonomy using DNA barcoding is more efficient than traditional morphological taxonomy, identifying the eggs and larvae to lower taxonomic levels and correcting erroneous morphological identifications (Becker et al., 2015; Ko et al., 2013). Using DNA barcoding, Frantine-Silva et al. (2015) identified over 99% of 536 ichthyoplankton samples at species levels, including eggs, which accounted for 30% of the observed species richness. Morphologically, Becker et al. (2015) identified eggs only as migratory or nonmigratory, when possible, while DNA barcoding allowed the identification of eggs (plus damaged larvae) to species level, and highlighted imprecisions in the morphological taxonomy even with such coarser analysis. Nonetheless, despite its great taxonomic precision, DNA barcoding relies on individual processing and sequencing of each organism and can become expensive and

laborious for large-scale inventories (Taberlet et al., 2012; Yu et al., 2012), such as ichthyoplankton studies (Mariat et al., 2018; Nobile et al., 2019).

The DNA metabarcoding approach, using High-Throughput Sequencing (HTS), has gained prominence for its ability to allow massive biodiversity access and transform ecology (Yu et al., 2012). The method combines DNA barcode-based taxonomy with HTS to simultaneously identify hundreds to thousands of organisms. DNA metabarcoding analyses are economical, quick, broad, minimally dependent on taxonomic expertise, and its data remain available for further verification (Taberlet et al., 2012; Yu et al., 2012). This approach has allowed the reconstruction of ancestral communities (Jørgensen et al., 2012), biodiversity monitoring (Andersen et al., 2012), and detection of larger operational taxonomic units in a fraction of the time spent in conventional studies based on morphology and DNA barcoding (Fonseca et al., 2010). This approach has also shown high efficiency in ecological ichthyoplankton studies, allowing precise and reliable identification of fish egg and larva bulk samples (Kimmerling et al., 2018; Mariat et al., 2018).

Different from environmental samples (for example, soil and water), in which genetic material is often degraded, bulk samples usually provide genomic DNA of better quality, allowing the amplification of markers with longer sequences (Taberlet et al., 2012). However, the HTS platforms accessible to most research laboratories have limited sequencing lengths of up to 600 base pairs (bp). This hampers the usage of the full COI DNA barcoding, since its primers amplified a 650 bp fragment of the mitochondrial cytochrome c oxidase subunit I (COI) gene commonly used for fish (Ward, 2009). Additionally, the variability in COI sequences hinders the design of internal minibarcode primers, taking some researchers to pass this gene over in favor of more conserved ones for metabarcoding (Collins et al., 2019; Deagle et al., 2014). Among these conserved genes, mitochondrial 12S rRNA has been highlighted as a good alternative for fish metabarcoding (Milan et al., 2020; Miya et al., 2020; Sales et al., 2021).

Besides marker selection, another challenge in DNA metabarcoding is quantitative analysis. Some factors can bias the number of read copies obtained for each individual or species, such as the number of mitochondria per cell, different-sized individuals in the same sample, and amplification biases (Carvalho, 2022; Fonseca, 2018). Nonetheless, some studies have shown a positive correlation between the number of eggs or larvae in mock samples and the number of reads obtained for each taxon using DNA metabarcoding with an amplification step (Duke & Burton, 2020; Nobile et al., 2019).

This study used DNA metabarcoding to analyze the composition of ichthyoplankton sampled at the Neotropical megadiverse São Francisco River Basin, in Brazil. Additionally, the sensibility, specificity, and taxonomic resolution of two 12S markers were tested and compared with the traditional COI fragment used for DNA barcoding. The results obtained here will contribute to an improved method for ecological studies focusing on the ichthyofauna reproductive dynamics, and to design management and conservation strategies for the maintenance of fish reproduction locally.

4.3. Materials and Methods

4.3.1. Sample collection

The São Francisco River Basin harbors at least 205 fish species (Barbosa et al., 2017), making it an excellent challenge for testing the sensibility and resolution of different markers within such complex Neotropical ichthyofauna. Fish collection was permitted by governmental agency ICMBio (license 70282-4) and the study was approved by the institutional ethics committee (protocol number 27-2019). Thirty ichthyoplankton samples (22 composed of eggs and eight composed of larvae) were sampled from the upper section of the São Francisco River and were named SF01 to SF30. The larvae samples were assembled using larvae of similar size (approximately 1 millimeter in average) to minimize bias due to biomass difference between species and relative read abundance estimation of input DNA from each pool.

4.3.2. Genomic DNA extraction

Genomic DNA was extracted from samples containing a pool of fish eggs and larvae fixed in ethanol. To ensure the complete evaporation of the alcohol, initially the excess was removed through pipetting, and then the microtubes were kept open for 3 h at 55°C. We then added 600 µL of TNES buffer to each sample and ground the bulk with a plastic pestle until only minuscule tissue fragments were left. Next, 20 µL of proteinase K (20 mg/mL) was added to each microtube. The samples were kept at 55°C until complete tissue digestion. Finally, the genomic DNA was extracted using a low-cost saline protocol adapted from (Aljanabi & Martinez, 1997).

The samples were quantified using a Qubit 4 Fluorometer (Thermo Fisher) with a 1x dsDNA HS Assay Kit (Thermo Fisher) to verify the success of the DNA extraction. The samples were then diluted to 100 ng/ μ L.

4.3.3. DNA amplification

The DNA was amplified via polymerase chain reaction (PCR). For the 12S rRNA gene, NeoFish (193 bp, Milan et al., 2020) and MiFish (185 bp, Miya et al., 2015) markers were amplified. Milan et al. (2020) showed that both these markers are great alternative for fish detection and identification by testing different species delimitations methods. A fragment of the COI gene was amplified using a cocktail of primers targeting the 650 bp COI fragment standardized for fish DNA barcoding (Ward et al., 2005). For each sample, all PCR replicates were amplified with the same index and pooled together for sequencing. These replicates were used to mitigate amplification biases that could affect species detection and RRA estimations. The PCR reaction solution had a final volume of 20 μ L, containing: 8.34 μ L of ultrapure water, 0.16 μ L of BSA (100 μ g/mL), 10 μ L of AmpliTaq™ Gold 360 Master Mix (Thermo Fisher), 0.25 μ L of each primer, and 1.0 μ L of DNA template. An additional 1.0 μ L of ultrapure water was added for the negative control samples instead of the DNA template. For the positive control sample, 1.0 μ L of template DNA from a saltwater fish species, *Prionace glauca*, was added.

PCR conditions consisted of initial denaturation for 10 min at 95°C, followed by 35 cycles of denaturation for 1 min at 95°C, primers annealing for 30 s at 56°C (COI), 60°C (MiFish), or 63°C (NeoFish) and extension for 30 s at 72°C, with a final extension for 7 min at 72°C.

The PCR results were checked using 1.8% agarose gel electrophoresis. Both 12S markers presented bands of the expected size for all samples, including the positive control, whereas for COI, the sample SF08 failed to produce any bands but was also included in the sequencing step. The negative control samples did not produce any bands but were also sequenced with the other samples.

4.3.4. Library preparation and DNA sequencing

According to the manufacturer's protocol, one library for each marker was prepared using the Collibri™ PCR-free Kit (Thermo Fisher). The libraries were quantified at the start of the preparation, after each major step, and at the end, by fluorometry. The 12S libraries were sequenced in a MiniSeq™ Sequencing System (Illumina) using a MiniSeq Mid Output Kit (300-cycle), and the COI library was sequenced in a MiSeq™ Sequencing System (Illumina) using a MiSeq Reagent Kit v3 (600-cycle).

4.3.5. Bioinformatic analyses

The bioinformatics analyses were carried out using a customized pipeline written in R v4.4.0 (R Core Team, 2021). Briefly, primer removal was performed using Cutadapt (Martin, 2011). The DADA2 package (Callahan et al., 2016) was used to filter reads with $Q > 20$, to merge read pairs into Amplicon Sequencing Variants (ASVs) using its quality-aware algorithm to remove chimeric ASVs, and to perform taxonomic assignment against a custom 12S sequence reference database for MiFish and NeoFish markers. This custom database was derived from the 12S library developed by Milan et al. (2020) for 12S markers containing 261 DNA sequences of 129 Neotropical fish species. The taxonomic assignment for the COI marker was performed using local BLASTn (Camacho et al., 2009) against the NCBI nucleotide database (Sayers et al., 2023; NCBIInt). Percentual identity thresholds were applied for identifications at the species level for COI (98%) and 12S (99%). The RRA (relative read abundance) was determined by dividing the absolute counts of each ASV by the sum of the absolute counts of all ASVs in a sample.

To compare species identifications between markers, Venn Diagrams were built using the web application Lucidchart (<https://lucid.app/>). A Permutational Multivariate Analysis of Variance (PERMANOVA) was performed to calculate the dissimilarities between the qualitative and quantitative diversity of fish taxa detected by each marker using the taxon presence/absence-based Jaccard dissimilarity index and the proportional read abundance-based Bray-Curtis dissimilarity index, respectively, with 1000 permutations. We then used a principal coordinate analysis (PCoA) to plot and visualize the differences between the markers in taxa composition. Calculation of the dissimilarity indexes and PCoA analysis were performed using the commands “adonis2” and “betadisper,” respectively (vegan 2.5-2 R package). Correlation analyses were performed to test for similarities between the species identified by each marker, using the R packages ggpubr (Kassambara, 2022) and ggplot2 (Wickham, 2016). For this

purpose, the relative read abundance (RRA) of each identification in each sample was plotted, using one marker per axis. The dispersions were used for determination of correlations, with coefficients and significance values, using the lm method.

Due to the maximum 600 bp length limitation of the sequencing technology available, the forward R1 and reverse R2 COI reads could not be merged by overlap to reconstruct the barcoding amplicon, as each strand covers a different region of the COI gene with possible distinct variations for each taxon. Therefore, reads R1 and R2 were analyzed separately, and each sample's taxonomic assignment results were combined.

The ASVs found in the negative controls were removed from all other samples. Additionally, considering that the high throughput could amplify contaminations not detected by negative controls, and the risk of false positives, but also aiming not to exclude underrepresented taxa, only ASVs with more than 0.01% of relative read abundance (RRA) in each sample were considered.

4.4. Results

All primer sets produced successful sequencing results for most samples. However, one sample (SF08) for COI did not produce any amplification, even after further DNA purification, quantification, and a new PCR adjustment. Nevertheless, SF08 was successfully amplified and sequenced using the 12S markers MiFish and NeoFish. A low number of reads were observed for SF04 and SF08, resulting in only one and three reads, respectively, despite the latter not presenting any problems in the amplification process, resulting in 93.33% (two failed samples out of 30) sequencing success. Notably, both 12S markers resulted in 100% amplification and sequencing success, with at least 38,697 reads in a sample (SF27) for MiFish, and 37,228 reads (SF10) for NeoFish.

The sequencing effort resulted in 4,505,309 reads for all markers after quality filtering. The number of reads showed considerable differences among markers and samples. COI produced 584,309 reads, averaging 19,477 reads per sample, ranging from one (SF04) to 29,485 (SF11). MiFish presented 1,919,545 total reads, averaging 63,985, with a minimum of 38,697 (SF27) and maximum of 83,734 (SF23). Sequencing with NeoFish resulted in 2,001,455 reads, with an average number of 66,715 per sample, varying from 37,228 (SF10) to 96,285 (SF23). After BLASTn searches, 1,699 COI reads remained without taxonomic assignment, and 69

were assigned to Bacteria. On the other hand, all MiFish and NeoFish reads were assigned to fish taxa.

ASVs were assigned to 26 fish taxa, from which 22 were identified at the species level, two at the genus level, and three at the subfamily level. The 12S marker NeoFish was able to detect the highest number of orders, families, genera, and the same number of species as MiFish. In contrast, COI detected fewer species, genera, and families than the other markers and the same number of orders as MiFish (Table S1, Figures 1, S1). The dispersion of relative read abundance values for each identification, plotted for each molecular marker pair (NeoFish vs. MiFish, NeoFish vs. COI, and MiFish vs. COI), shows that there is no correlation between the NeoFish and MiFish markers ($R = .081, p = .11$), nor between the NeoFish and COI markers ($R = -.11, p = .036$; Figure 2). However, a significant positive correlation was observed between the MiFish and COI markers (with $R = .86, p < 2.2^{-16}$).

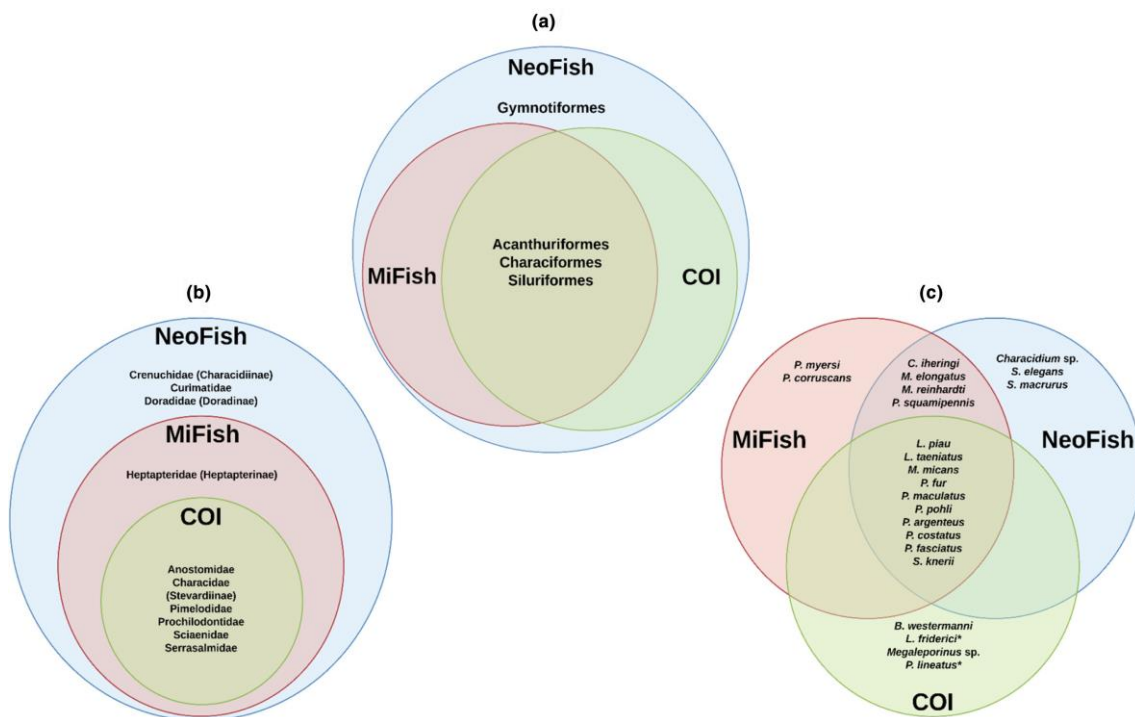


Figure 1. Venn diagrams recovered by each molecular marker MiFish, NeoFish, and COI considering distinct taxonomic levels: (a) order, (b) family/(subfamily), and (c) genus/species. Species marked with an asterisk (*) have not been reported for the São Francisco River Basin.

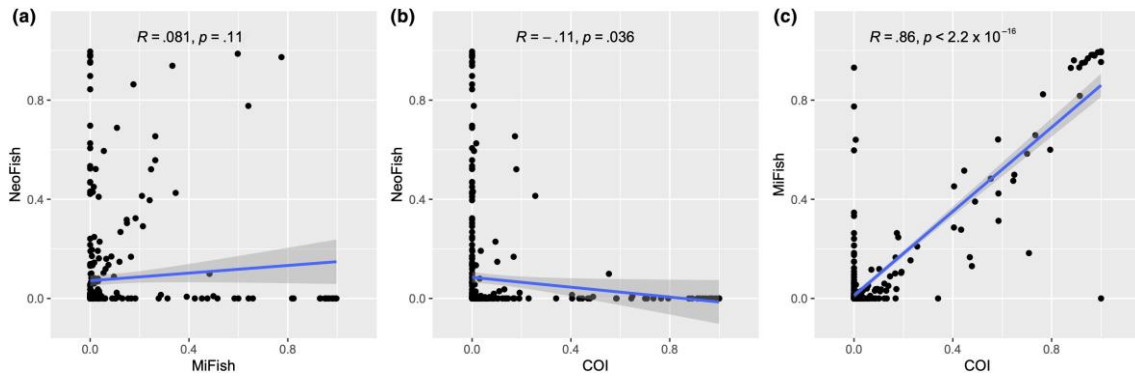


Figure 2. Scatter plots correlation between each molecular marker and the identified species in all ichthyoplankton samples. Each dot represents an identification in a sample, plotted using its respective abundance in the sample as measured by the marker on the corresponding axes. Blue lines show the regression lines with the confidence intervals in dark gray. R and p values are displayed for each correlation.

The combined use of the three markers increased the genera detection rates by 25%–87.5% when considering an initial analysis with only NeoFish or COI, respectively (Figure 3). The improvement in species recovery rates with the use of all three markers combined ranged from 31.25% to 61.54% when considering an initial analysis with either 12S gene markers or COI, respectively (Figure 3).

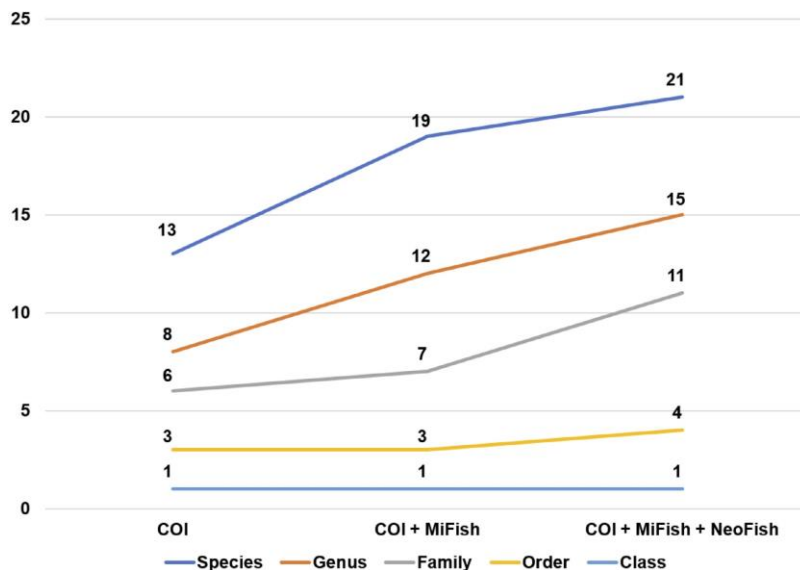


Figure 3. Improvement in detected taxa when using a single marker for the COI gene, combining COI and MiFish and combining COI, MiFish, and NeoFish.

The COI marker detected 16 taxa belonging to 13 species, eight genera, six families, three orders, and one class (Figure S1). Besides the 13 taxa identified at the species level, one was identified at the genus level, one at the subfamily level and another at the family level (Table S1). Among the 13 species, three were detected exclusively by the COI gene (*Bergiaria westermanni*, *Leporinus friderici*, and *Prochilodus lineatus*). The species *B. westermanni* was present in 12 samples, with an average RRA of 2.29%, ranging from 0.13% to 6.16%. The anostomid *L. friderici* was detected in eight samples, with the RRA ranging from 0.13% to 32.64%, and an average of 7.81%. Lastly, *P. lineatus* was present in three samples and had an average RRA of 3.60%, ranging between 1.20% and 7.12% (Figure 4).

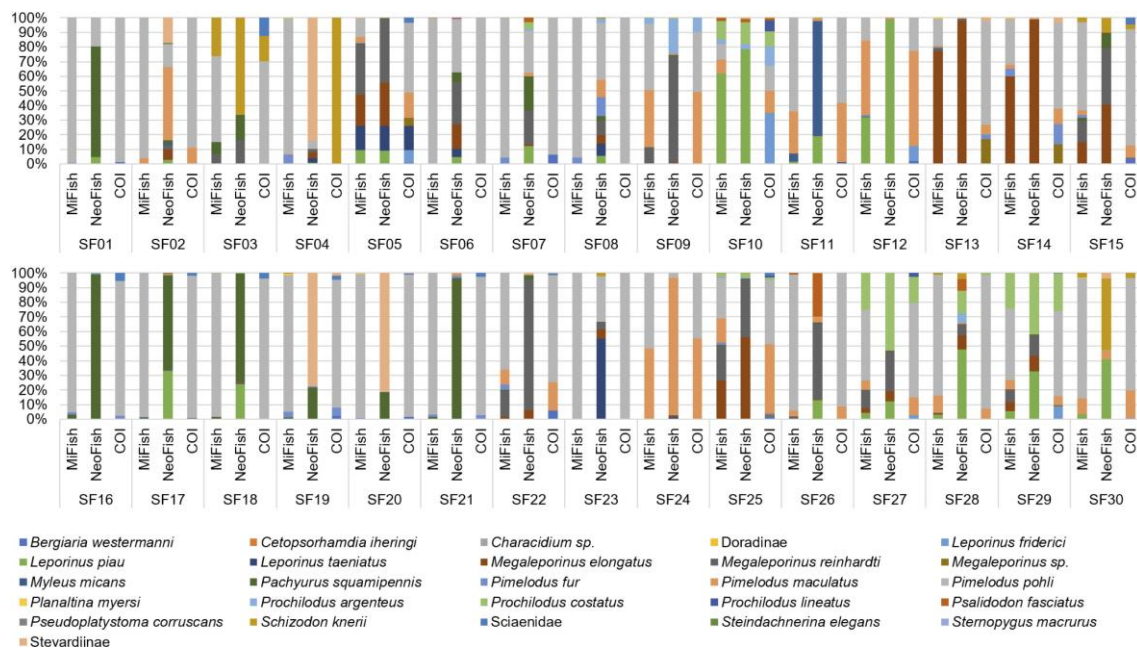


Figure 4. Relative read abundance (RRA) for each taxon, in each sample, with each marker.

With the 12S gene markers, 21 taxa were detected, including 18 species, 14 genera, 11 families, four orders, and one class (Table S1). One of these taxa was identified at the genus level (*Characidium* sp.) and two at the subfamily level (Doradinae and Stevardiinae). Of the 18 identifications at species level, eight were exclusively detected with the 12S gene, all indigenous to the São Francisco Basin: *Cetopsorhamdia iheringi*, *Megaleporinus elongatus*, *M. reinhardti*, *Pachyurus squamipennis*, *Planaltina myersi*, *Pseudoplatystoma corruscans*, *Steindachnerina elegans*, and *Sternopygus macrurus* (Figure 1, Table S1). Four of the eight species detected exclusively with the 12S markers were detected by both markers, but with

some variation in samples and abundance. For example, the species *C. iheringi* was detected by MiFish and NeoFish in sample SF15, with 0.07% and 0.02% RRA, respectively (Figure 4).

Overall, RRA and taxon detection was not consistent between each marker (Figure 4). For instance, the most abundant taxon *P. pohli* had a total of 1,348,589 (70.79% of the total) for MiFish and 424,499 (73.02%) reads recovered for NeoFish. Notably, the highest RRA detected for NeoFish was *M. elongatus*, with 236,332 (20.90% of the total) reads. Additionally, in some samples (e.g., SF05, SF13, SF14 and SF15) where both MiFish and COI detected multiple taxa for Pimelodidae, NeoFish was not able to identify any taxa for this family (Figure 3).

The PERMANOVA evidenced significant differentiation of fish communities using distinct molecular markers (Table 1). The influence of marker choice on taxa recovered in each sample was significant for both presence-absence (Jaccard) and RRA (Bray-Curtis) analyses (Table 1). While there is considerable overlap between both 12S markers considering only presence-absence of each taxon in each sample, as can be seen in the PcoA plot (Figure 5a), the analysis taking taxa abundance into account revealed a slight overlap between MiFish and COI (Figure 5b).

Table 1. Summary of PERMANOVA results (R^2 -effect sizes and significance level) showing the effect of marker choice on taxa recovered. Df, Degree of freedom; Sum of Squares, Value of Test F ; and p -value associated with the F score.

Index	Df	SumSqs	R^2	F	p
Jaccard	2	6.050	.1881	10.079	.001
Bray-Curtis	2	7.898	.2283	12.869	.001

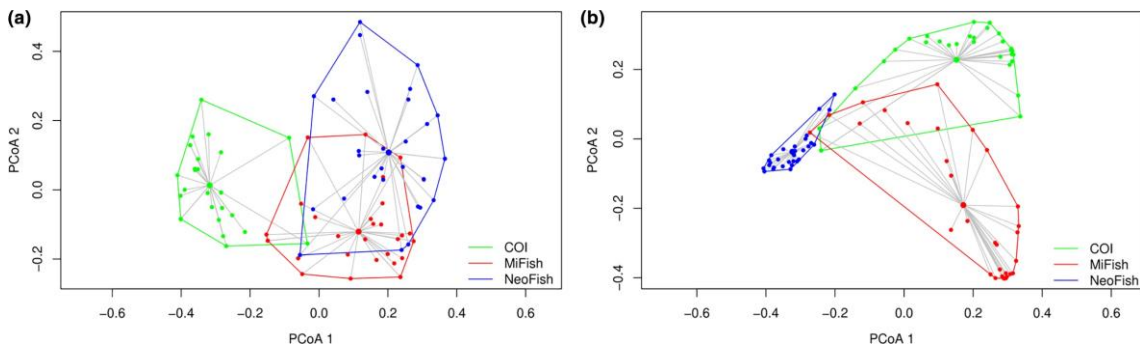


Figure 5. PCoA plots of 30 ichthyoplankton fish communities detected using three different molecular markers: COI, MiFish, and NeoFish. Analyses were conducted using (a) Jaccard coefficient considering

the presence/absence of fish taxa and (b) the Bray-Curtis coefficient, which uses a matrix of abundance based on each taxon.

4.5. Discussion

DNA metabarcoding has become an essential tool for species inventory and monitoring. However, its use in identifying ichthyoplankton is still incipient in the Neotropics, with several methodological challenges and biases still needing to be tackled (Carvalho, 2022). Considering the crescent demand for innovative techniques to unravel the complex reproductive dynamics of fish communities for both research and practical applications, there is an underlying need for the continuous refinement of this methodology. The COI gene has been commonly used as the marker of choice (Mariac et al., 2018; Nobile et al., 2019) because of its well-established primers and complete reference sequences libraries encouraged by the global initiative Fish Barcode of Life. Here, each molecular marker recovers a distinct community structure when considering both quantitative and qualitative analysis.

Using a marker of choice still raises concerns since using several markers is still expensive when using HTS and because each marker has distinct amplification biases and taxonomic resolution (Deagle et al., 2014). The high interspecific variability of the COI gene, when compared to other mitochondrial genes (Hebert et al., 2003), can help differentiate closely related species. However, the same high variability creates the need to use universal degenerate primers with lower specificity than those designed for more conserved genes. Also, it hinders the design of internal minibarcodes for COI (Deagle et al., 2014). In the present study, while COI presented 93.33% sequencing success and some of its sequences remained unassigned or were assigned to Bacteria, both 12S markers were successfully sequenced for all samples, and all their sequences were assigned to fish taxa. Additionally, the technology used for sequencing limits the total fragment size to 600 bp, precluding the merging of both COI strands from forming the full-sized barcode; therefore, each strand was analyzed independently. The loss of resolution power caused by this could explain why COI detected fewer species, genera, and families than the 12S markers and why two of the three exclusive species level identifications were assigned to nonnative fishes closely related to species from São Francisco.

Minibarcodes for the 12S gene have been developed and applied to environmental DNA metabarcoding studies (Milan et al., 2020; Miya et al., 2020; Sales et al., 2021) and, more recently, to ichthyoplankton studies as well (Jiang et al., 2022; Van Nynatten

et al., 2023). One of the main concerns when using these markers is the conserved nature of the gene, which can impact their ability to differentiate closely related species, especially in diverse regions. However, the current study shows that both MiFish and NeoFish were able to successfully identify and distinguish multiple congeneric species, such as *Leporinus piau* and *L. taeniatus*, *Megaleporinus elongatus* and *M. reinhardti*, *Pimelodus fur*, *P. maculatus* and *P. pohli*, and *Prochilodus argenteus* and *P. costatus*. Moreover, the 12S markers have a higher species detection sensibility than COI, considering that the exclusive fishes they retrieved were underrepresented, with low RRA. This could result from low-efficiency primer binding by COI, which can lead to a lack of amplification (Zhang et al., 2020). Moreover, the relatively low number of overall taxa detected ($n = 26$) in comparison with the described regional biodiversity ($n = 205$) is likely not related to the detection power of the DNA metabarcoding method, but rather due to the sampling region, since we sampled only the Upper São Francisco River, combined with the selectiveness of the sampling method that targets only drifting ichthyoplankton.

Database completeness is another variable that directly impacts species detection, as a lack of reference sequences for a given species may hamper accurate taxonomic assignment (Collins et al., 2019). This aspect has affected both COI and 12S markers in this study. For instance, while *Pachyurus squamipennis* is not represented by any COI reference sequence in the public databases and was exclusively detected by 12S, the only native species retrieved solely by COI, *Bergiaria westermanni*, does not have any 12S representative sequence in neither the public nor our custom library. These limitations highlight the importance of continuous sequencing efforts to broaden reference sequence databases, especially for megadiverse regions.

Considering that each marker has advantages and limitations, some studies suggest combining multiple primer sets to increase taxonomic coverage (Liu & Zhang, 2021; Zhang et al., 2020). In a metabarcoding study using multiplexed markers to identify zooplankton mock communities, Zhang et al. (2018) demonstrated that a multi-marker approach can improve species detection and allow the cross-validation of taxa detected by each marker. Our results support this conclusion, as using the three markers combined increased the genera detection by up to 87.5% and species detection by up to 61.54%. Therefore, employing multiple markers, whenever feasible, reduces the likelihood of overlooking species or incorrectly classifying them due to the absence or mislabeling of sequences in the reference database (Locatelli et al., 2020).

Discrepancies between markers were observed in the quantitative analysis using the RRA values. Although some studies with mock samples of eggs (Duke & Burton, 2020) and larvae (Nobile et al., 2019) yielded a positive correlation between input organisms and output reads for each species, the results from this study support the idea that amplification bias is one of the main pitfalls for quantitative metabarcoding analyses, as already reported elsewhere (Carvalho, 2022; Fonseca, 2018). Moreover, while the correlation analyses showed positive correlation between the RRA values found for MiFish and COI ($R = .86, p < .01$), low-efficiency primer binding to Siluriformes and especially Pimelodidae sequences may have led to different abundance patterns for NeoFish when compared to the other two markers. However, despite the significant positive correlation between MiFish and COI, many identified species were detected by only one of these markers (Figure 2). The inefficiency of the NeoFish marker to amplify Siluriformes has already been reported using DNA metabarcoding of Neotropical fish mock communities (Hilário et al., 2023). Therefore, such biases in the amplification and detection of species by only one marker are in agreement with our results using real communities and reinforce the importance of using multiple markers when conducting large-scale monitoring programs using DNA metabarcoding analysis.

In conclusion, using multiple markers from two distinct genes and lengths allowed an increased taxonomic coverage and robust taxonomic classification of complex Neotropical ichthyoplankton communities. Finally, precaution is still needed when inferring species abundance based on DNA metabarcoding data when using PCR-dependent protocols since it is marker dependent. Nonetheless, ichthyoplankton metabarcoding offers superior resolution and feasible scalability compared to traditional techniques, and provides qualitative information, which is paramount for characterizing reproducing species and definition of conservation strategies.

AUTHOR CONTRIBUTIONS

Daniel Fonseca Teixeira: Data curation (equal); formal analysis (equal); investigation (equal); methodology (equal); resources (equal); software (equal); validation (equal); visualization (equal); writing – original draft (equal). Heron Oliveira Hilário: Data curation (equal); formal analysis (equal); methodology (equal); software (equal); validation (equal); visualization (equal). Gilmar Bastos Santos: Funding acquisition (equal); project administration (equal); writing – review and editing (supporting). Daniel Cardoso Carvalho: Conceptualization (equal);

funding acquisition (equal); investigation (equal); project administration (equal); resources (equal); supervision (equal); validation (equal); writing – original draft (equal); writing – review and editing (equal).

ACKNOWLEDGMENTS

We are grateful to Amanda Jéssica Silva Santos for bench work assistance, Guilherme Moreira Santos and Gustavo Ribeiro Rosa for helping with taxonomic assignment curation, and Izabela Santos Mendes for contributing to statistical analysis. This work has been financially supported by FAPEMIG, CAPES, FIP-PUC Minas, and ANEEL/ CEMIG (P&D 0624). We thank CNPq for providing a research fellowship (312102/2022-4) to DCC.

DATA AVAILABILITY STATEMENT

Raw Illumina sequences and corresponding metadata are deposited in the Zenodo repository (<https://doi.org/10.5281/zenodo.8038835>).

ORCID

Heron Oliveira Hilário: <https://orcid.org/0000-0002-2548-1053>

Daniel Cardoso Carvalho: <https://orcid.org/0000-0003-4517-5069>

REFERENCES

- Ahlstrom, E. H., & Moser, H. G. (1976). Eggs and larvae of fishes and their role in systematic investigations and in fisheries. *Revue Des Travaux de L'institut Des Peches Maritimes*, 40(3), 379–398.
- Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high-quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, 25(22), 4692–4693.
- Andersen, K., Bird, K. L., Rasmussen, M., Haile, J., Breuning-Madsen, H., Kjaer, K. H., Orlando, L., Gilbert, M. T. P., & Willerslev, E. (2012). Meta-barcoding of ‘dirt’ DNA from soil reflects vertebrate biodiversity. *Molecular Ecology*, 21(8), 1966–1979.
- Barbosa, J. M., Soares, E. C., Cintra, I. H. A., Hermann, M., & Araújo, A. R. R. (2017). Profile of the fish fauna of São Francisco Basin. *Acta of Fisheries and Aquatic Resources*. 5, 70–90.

- Baumgartner, G., Nakatani, K., Gomes, L. C., Bialezki, A., & Sanches, P. V. (2004). Identification of spawning sites and natural nurseries of fishes in the upper Paraná River, Brazil. *Environmental Biology of Fishes*, *71*(2), 115–125.
- Baumgartner, G., Nakatani, K., Gomes, L. C., Bialezki, A., Sanches, P. V., & Makrakis, M. C. (2008). Fish larvae from the upper Paraná River: Do abiotic factors affect larval density? *Neotropical Ichthyology*, *6*, 551–558.
- Becker, R. A., Sales, N. G., Santos, G. M., Santos, G. B., & Carvalho, D. C. (2015). DNA barcoding and morphological identification of neotropical ichthyoplankton from the Upper Paraná and São Francisco. *Journal of Fish Biology*, *87*(1), 159–168.
- Bialezki, A., Nakatani, K., Sanches, P. V., Baumgartner, G., & Gomes, L. C. (2005). Larval fish assemblage in the Baía River (Mato Grosso do Sul State, Brazil): Temporal and spatial patterns. *Environmental Biology of Fishes*, *73*, 37–47.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: *Architecture and applications*. *BMC Bioinformatics*, *10*, 1–9.
- Carvalho, D. C. (2022). Ichthyoplankton DNA metabarcoding: Challenges and perspectives. *Molecular Ecology*, *31*(6), 1612–1614.
- Collins, R. A., Bakker, J., Wangensteen, O. S., Soto, A. Z., Corrigan, L., Sims, D. W., Genner, M. J., & Mariani, S. (2019). Non-specific amplification compromises environmental DNA metabarcoding with COI. *Methods in Ecology and Evolution*, *10*(11), 1985–2001.
- Da Silva, P. S., Makrakis, M. C., Miranda, L. E., Makrakis, S., Assumpção, L., Paula, S., Dias, J. H. P., & Marques, H. (2015). Importance of reservoir tributaries to spawning of migratory fish in the upper Paraná River. *River Research and Applications*, *31*(3), 313–322.
- Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters*, *10*(9), 20140562.
- Duke, E. M., & Burton, R. S. (2020). Efficacy of metabarcoding for identification of fish eggs evaluated with mock communities. *Ecology and Evolution*, *10*(7), 3463–3476.
- Fonseca, V. G. (2018). Pitfalls in relative abundance estimation using eDNA metabarcoding. *Molecular Ecology Resources*, *18*(5), 923–926.
- Fonseca, V. G., Carvalho, G. R., Sung, W., Johnson, H. F., Power, D. M., Neill, S. P., Packer, M., Blaxter, M. L., Lambshead, P. J. D., Thomas, W. K., & Creer, S. (2010). Second-generation environmental sequencing unmasking marine metazoan biodiversity. *Nature Communications*, *1*(1), 98.

- Frantine-Silva, W., Sofia, S. H., Orsi, M. L., & Almeida, F. S. (2015). DNA barcoding of freshwater ichthyoplankton in the Neotropics as a tool for ecological monitoring. *Molecular Ecology Resources*, *15*(5), 1226–1237.
- Hebert, P. D., Ratnasingham, S., & De Waard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(suppl_1), S96–S99.
- Hilário, H. O., Mendes, I. S., Guimaraes Sales, N., & Carvalho, D. C. (2023). DNA metabarcoding of mock communities highlights potential biases when assessing Neotropical fish diversity. *Environmental DNA* in press. <https://doi.org/10.1002/edn3.456>
- Jiang, R., Lusana, J. L., & Chen, Y. (2022). High-throughput DNA metabarcoding as an approach for Ichthyoplankton survey in Oujiang River estuary, China. *Diversity*, *14*(12), 1111.
- Jørgensen, T., Haile, J., Möller, P. E. R., Andreev, A., Boessenkool, S., Rasmussen, M., Kienast, F., Coissac, E., Taberlet, P., Brochmann, C., Bigelow, N. H., Andersen, K., Orlando, L., Gilbert, M. T. P., & Willerslev, E. (2012). A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. *Molecular Ecology*, *21*(8), 1989–2003.
- Kassambara, A. (2022). *ggpubr: 'ggplot2' based publication ready plots*. R package version 0.5.0. <https://CRAN.R-project.org/package=ggpubr>
- Kimmerling, N., Zuqert, O., Amitai, G., Gurevich, T., Armoza-Zvuloni, R., Kolesnikov, I., Berenshtein, I., Melamed, S., Gilad, S., Benjamin, S., Rivlin, A., Ohavia, M., Paris, C. B., Holzman, R., Kiflawi, M., & Sorek, R. (2018). Quantitative species-level ecology of reef fish larvae via metabarcoding. *Nature Ecology & Evolution*, *2*(2), 306–316.
- Ko, H. L., Wang, Y. T., Chiu, T. S., Lee, M. A., Leu, M. Y., Chang, K. Z., Chen, W. Y., & Shao, K. T. (2013). Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *PLoS One*, *8*(1), e53451.
- Liu, J., & Zhang, H. (2021). Combining multiple markers in environmental DNA metabarcoding to assess deep-sea benthic biodiversity. *Frontiers in Marine Science*, *8*, 684955.
- Locatelli, N. S., McIntyre, P. B., Therkildsen, N. O., & Baetscher, D. S. (2020). GenBank's reliability is uncertain for biodiversity researchers seeking species-level assignment for eDNA. *Proceedings of the National Academy of Sciences*, *117*(51), 32211–32212.
- Mariac, C., Vigouroux, Y., Duponchelle, F., García-Dávila, C., Núñez, J., Desmarais, E., & Renno, J. F. (2018). Metabarcoding by capture using a single COI probe (MCSP) to identify and quantify fish species in ichthyoplankton swarms. *PLoS One*, *13*(9), e0202976.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal*, *17*(1), 10–12.
- Milan, D. T., Mendes, I. S., Damasceno, J. S., Teixeira, D. F., Sales, N. G., & Carvalho, D. C. (2020). New 12S metabarcoding primers for enhanced Neotropical freshwater fish biodiversity

assessment. *Scientific Reports*, 10(1), 1–12.

Miya, M., Gotoh, R. O., & Sado, T. (2020). MiFish metabarcoding: A high-throughput approach for simultaneous detection of multiple fish species from environmental DNA and other samples. *Fisheries Science*, 86(6), 939–970.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., Minamoto, T., Yamamoto, S., Yamanaka, H., Araki, H., Kondoh, M., & Iwasaki, W. (2015). 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*, 2(7), 150088.

Nakatani, K., Agostinho, A. A., Baumgartner, G., Bialletzki, A., Sanches, P. V., Makrakis, M. C., & Pavanelli, C. S. (2001). *Ovos e larvas de peixes de água doce: Desenvolvimento e manual de identificação* (first ed.). EDUEM.

Nobile, A. B., Freitas-Souza, D., Ruiz-Ruano, F. J., Nobile, M. L. M., Costa, G. O., De Lima, F. P., Camacho, J. P. M., Foresti, F., & Oliveira, C. (2019). DNA metabarcoding of Neotropical ichthyoplankton: Enabling high accuracy with lower cost. *Metabarcoding and Metagenomics*, 3, e35060.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Reynalte-Tataje, D. A., Agostinho, A. A., Bialletzki, A., Hermes-Silva, S., Fernandes, R., & Zaniboni-Filho, E. (2012). Spatial and temporal variation of the ichthyoplankton in a subtropical river in Brazil. *Environmental Biology of Fishes*, 94, 403–419.

Sales, N. G., Wangenstein, O. S., Carvalho, D. C., Deiner, K., Præbel, K., Coscia, I., McDevitt, A. D., & Mariani, S. (2021). Space-time dynamics in monitoring neotropical fish communities using eDNA metabarcoding. *Science of the Total Environment*, 754, 142096.

Sayers, E. W., Bolton, E. E., Brister, J. R., Canese, K., Chan, J., Comeau, D. C., Funk, K., Kim, S., Klimke, W., Marchler-Bauer, A., Landrum, M., Lathrop, S., Lu, Z., Madden, T. L., O'Leary, N., Phan, L., Rangwala, S. H., Schneider, V. A., Skripchenk, Y., & Sherry, S. T. (2023). Database resources of the National Center for biotechnology information in 2023. *Nucleic Acids Research*, 51(D1), D29–D38.

Silva, C. B., Dias, J. D., & Bialletzki, A. (2017). Fish larvae diversity in a conservation area of a neotropical floodplain: Influence of temporal and spatial scales. *Hydrobiologia*, 787, 141–152.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, 21(8), 2045–2050.

Van Nynatten, A., Gallage, K. S., Lujan, N. K., Mandrak, N. E., & Lovejoy, N. R. (2023). Ichthyoplankton metabarcoding: An efficient tool for early detection of invasive species establishment. *Molecular Ecology Resources*, 23, 1319–1333.

Ward, R. D. (2009). DNA barcode divergence among species and genera of birds and fishes. *Molecular Ecology Resources*, 9(4), 1077–1085.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 360(1462), 1847–1857.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, 3(4), 613–623.

Zhang, G. K., Chain, F. J., Abbott, C. L., & Cristescu, M. E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. *Evolutionary Applications*, 11(10), 1901–1914.

Zhang, S., Zhao, J., & Yao, M. (2020). A comprehensive and comparative evaluation of primers for metabarcoding eDNA from fish. *Methods in Ecology and Evolution*, 11(12), 1609–1625.

How to cite this article: Teixeira, D. F., Hilário, H. O., Santos, G. B., & Carvalho, D. C. (2023). DNA metabarcoding assessment of Neotropical ichthyoplankton communities is marker-dependent. *Ecology and Evolution*, 13, e10649. <https://doi.org/10.1002/ece3.10649>

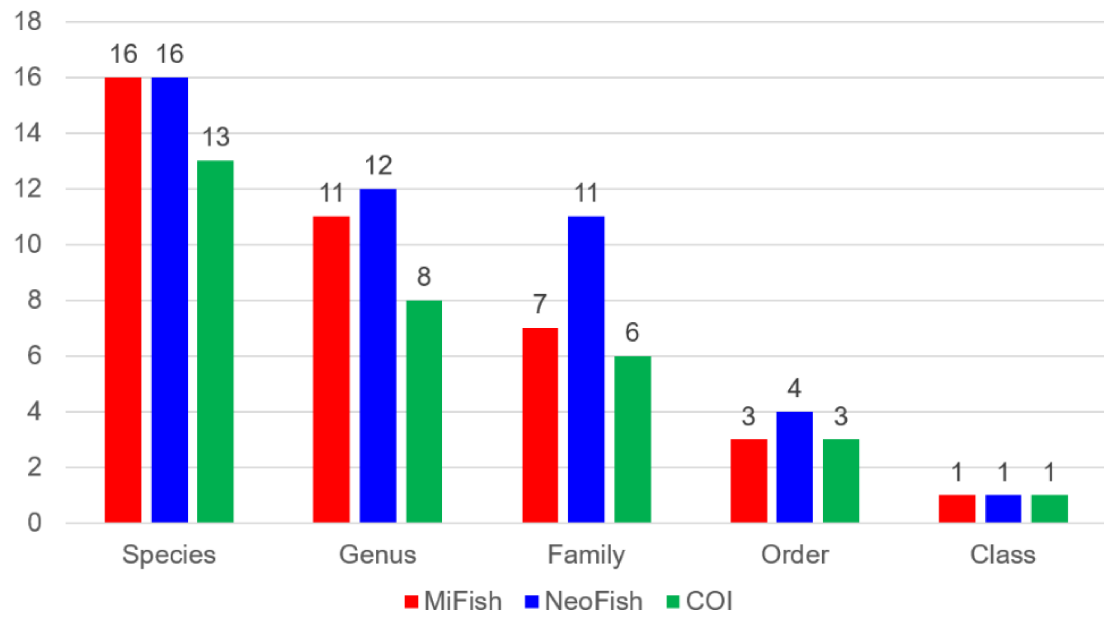
SUPPORTING INFORMATION

Figure S1. Number of species, genera, families, order, and classes detected by each marker.

Table S1. Twenty-six taxa were detected within all samples by the three markers MiFish, NeoFish and COI. An asterisk (*) highlights the species not reported for the São Francisco River Basin.

Order	Family	Subfamily	Genus/Species	MiFish	NeoFish	COI
Acanthuriformes	Sciaenidae	-	-	-	-	X
			<i>Pachyurus squamipennis</i>	X	X	-
Characiformes	Anostomidae	-	<i>Leporinus friderici*</i>	-	-	X
			<i>Leporinus piau</i>	X	X	X
			<i>Leporinus taeniatus</i>	X	X	X
			<i>Megaleporinus elongatus</i>	X	X	-
			<i>Megaleporinus reinhardti</i>	X	X	-
			<i>Megaleporinus</i> sp.	-	-	X
			<i>Schizodon knerii</i>	X	X	X
	Characidae	Stethaproninae	<i>Psalidodon fasciatus</i>	X	X	X
		Stevardiinae	-	X	X	X
			<i>Planaltina myersi</i>	X	-	-
	Crenuchidae	Characidiinae	<i>Characidium</i> sp.	-	X	-
	Curimatidae	-	<i>Steindachnerina elegans</i>	-	X	-
	Prochilodontidae	-	<i>Prochilodus argenteus</i>	X	X	X
			<i>Prochilodus costatus</i>	X	X	X
			<i>Prochilodus lineatus*</i>	-	-	X
Serrasalminidae	Serrasalminae	<i>Myleus micans</i>	X	X	X	
Gymnotiformes	Sternopygidae	Sternopyginae	<i>Sternopygus macrurus</i>	-	X	-
Siluriformes	Doradidae	Doradinae	-	-	X	-
	Heptapteridae	Heptapterinae	<i>Cetopsorhamdia iheringi</i>	X	X	-
	Pimelodidae	-	<i>Bergiaria westermanni</i>	-	-	X
			<i>Pimelodus fur</i>	X	X	X
			<i>Pimelodus maculatus</i>	X	X	X
<i>Pimelodus pohli</i>	X	X	X			
<i>Pseudoplatystoma corruscans</i>	X	-	-			

5. CAPÍTULO 2: Inventariamento em larga escala do ictioplâncton utilizando DNA *metabarcoding* para análise da dinâmica reprodutiva da ictiofauna

5.1. Objetivos Específicos

- Identificar a composição do ictioplâncton em diversos pontos da bacia do rio São Francisco;
- Avaliar a contribuição de organismos de diferentes estágios de desenvolvimento (ovos ou larvas) para a riqueza detectada;
- Analisar o impacto do isolamento causado pelo barramento de Três Marias na reprodução dos peixes;
- Identificar pontos prioritários para a conservação visando a manutenção da reprodução de espécies consideradas grandes migradoras.

RESUMO

A utilização de ferramentas moleculares, como o DNA *barcoding* e *metabarcoding*, tem se mostrado uma alternativa atraente para realizar identificações confiáveis e assertivas do ictioplâncton. Além disso, a utilização do sequenciamento de DNA de alto rendimento para identificar múltiplos organismos com o DNA *metabarcoding* tem demonstrado a grande utilidade dessa técnica para o inventariamento em larga escala, como normalmente demandam os estudos de ictioplâncton. No entanto, a maior parte dos trabalhos com ictioplâncton utilizando o DNA *metabarcoding* até a presente data são metodológicos, e sua aplicação prática ainda foi pouco explorada. Durante o presente estudo, foram identificados 12.539 ovos e 15.225 larvas de peixes provenientes de 11 tributários, de diferentes ordens de grandeza, à jusante e à montante da barragem de Três Marias, no rio São Francisco. O DNA *barcoding* foi utilizado em amostras contendo cinco ou menos organismos, enquanto o *metabarcoding* foi empregado para processar simultaneamente todos os indivíduos de amostras com maior abundância. A identificação do ictioplâncton visou responder três principais questões: (a) qual a contribuição relativa de ovos e larvas para a riqueza total encontrada; (b) se existe diferença significativa entre as comunidades de ictioplâncton encontradas à montante e à jusante de Três Marias; (c) se algum dos tributários analisados pode ser considerado como ponto prioritário para conservação. Ao todo, foram detectados 54 táxons, sendo 51 em nível de espécie. Apesar de no geral haver baixa diferenciação entre táxons detectados por ovos ou larvas ($R^2 = 0,0448$, $p = 0,51$), ao analisar cada tributário isoladamente, ovos foram responsáveis exclusivos por até 61,90% da riqueza observada, enquanto larvas contribuíram com até 60%. Também foi observada grande sobreposição entre os táxons encontrados à montante e à jusante do reservatório ($R^2 = 0,0599$, $p = 0,115$). No entanto, o rio Paracatu, localizado à jusante, apresentou a maior riqueza (28 táxons) e foi o único tributário onde todas as espécies consideradas grandes migradoras foram detectadas, se destacando como um ponto prioritário para a conservação. Dessa forma, este trabalho demonstrou a aplicabilidade do DNA *metabarcoding* para processar e identificar uma grande quantidade de ictioplâncton em estudos práticos, além de contribuir para o conhecimento da dinâmica reprodutiva da ictiofauna na área de influência de Três Marias, podendo ser utilizado para subsidiar estratégias de conservação que visem a manutenção da reprodução e da pesca na região.

Palavras-chave: 12S rRNA, berçário, sequenciamento de nova geração, sítio de desova.

ABSTRACT

Molecular tools such as DNA barcoding and metabarcoding have proven to be great alternatives for assertive and reliable ichthyoplankton identification. Furthermore, the use of high-throughput DNA sequencing for identifying multiple organisms through DNA metabarcoding has shown great utility for large-scale inventory, as is often necessary in ichthyoplankton studies. However, most works to date using DNA metabarcoding to identify fish eggs and larvae are methodological, and its practical application remains incipient. During the present study, we analyzed 12,539 eggs and 15,225 larvae from 11 different-sized tributaries upstream of and downstream from the Três Marias reservoir, São Francisco River, Southeastern Brazil. DNA barcoding was applied to samples containing five or less individuals, while metabarcoding was used to simultaneously process all organisms in samples with larger sizes. Ichthyoplankton identification aimed at answering three major questions: (a) what is the relative contribution of eggs and larvae to total observed richness; (b) if there are significant differentiation between the ichthyoplankton communities found upstream of and downstream from Três Marias reservoir; (c) if a tributary can be considered a primary conservation site. Overall, we detected 54 taxa, 51 of which were identified at species level. Despite the low statistical differentiation in taxon composition between eggs and larvae ($R^2 = 0,0448$, $p = 0,51$), when analyzing each tributary separately, eggs were exclusively responsible for up to 61,90% of observed richness, while larvae contributed with up to 60%. We also observed great overlap between taxa found in the upstream and downstream regions ($R^2 = 0,0599$, $p = 0,115$). However, the Paracatu River, located downstream from the reservoir, presented the highest richness (28 taxa), and was the only tributary where all the great migratory species were detected, highlighting it as a primary site for conservation. Therefore, this work demonstrated the applicability of DNA metabarcoding for processing and identifying large numbers of ichthyoplankton in practical studies, while also contributing to the ichthyofauna reproduction dynamics knowledge in the area impacted by the Três Marias reservoir, and may be used to subsidize conservation strategies aimed at maintaining fish reproduction and, consequently, fishing stocks in the region.

Keywords: 12S rRNA, nursery, next generation sequencing, spawning site.

5.2. Introdução

Devido à alta demanda de dados taxonômicos e rápida resposta necessárias para embasar os esforços atuais de conservação e manejo da biodiversidade, em muitos casos os métodos baseados em caracteres morfológicos são insuficientes quando empregados isoladamente (HARVEY et al., 2017). Além disso, inventariamentos baseados na identificação morfológica são normalmente dependentes de especialistas em diferentes táxons, o que pode levar a inconsistências nos dados taxonômicos entre estudos conduzidos por diferentes pesquisadores (KO et al., 2013).

Especialmente para o ictioplâncton, métodos morfológicos exigem experiência aprofundada em anatomia e ontogênese (RICHARDSON et al., 2007). A identificação de vários grupos taxonômicos de peixes de forma precisa nos estágios larvais pode ser desafiadora, mesmo para taxonomistas experientes (KOCHZIUS, 2009; FRANTINE-SILVA et al., 2015; KO et al., 2013). Nos seus primeiros estágios de vida, peixes possuem tamanho diminuto, são frágeis e passam por grandes alterações na forma e coloração (BIALETZKI et al., 2016). Dessa forma, a utilização de técnicas alternativas é fundamental para o estudo adequado dos ovos e larvas de peixes.

Para aumentar a confiabilidade da identificação do ictioplâncton, pesquisadores têm recorrido a técnicas moleculares baseadas em DNA *barcodes*, que são regiões do DNA utilizadas para comparação e identificação de organismos de acordo com sua similaridade. Estudos comparativos demonstraram que a identificação molecular utilizando DNA *barcodes* é mais eficiente que a identificação morfológica tradicional, sendo capaz de chegar a níveis taxonômicos menores e corrigir identificações morfológicas errôneas (KO et al., 2013; BECKER et al., 2015).

Estudos com ictioplâncton, no entanto, envolvem um grande número de organismos, demandando a utilização de métodos de alta performance, como o DNA *metabarcoding* (MARIAC et al., 2018; NOBILE et al., 2019). Essa técnica combina a identificação baseada em DNA *barcodes* com o sequenciamento de nova geração (NGS) para analisar vários organismos simultaneamente. Além disso, o método demanda pouca perícia taxonômica, sua execução é rápida e econômica quando comparada a técnicas que analisam cada organismo individualmente, e produz bancos de dados verificáveis (TABERLET et al., 2012; YU et al., 2012). Essa abordagem tem se mostrado eficiente nos estudos de ictioplâncton, permitindo identificar sua composição de maneira precisa e confiável, e sua distribuição temporal e espacial

através da análise de *bulks* de ovos e larvas (KIMMERLING et al., 2018; MARIAC et al., 2018).

Informações precisas sobre a distribuição do ictioplâncton são fundamentais para identificar locais de desova e recrutamento, rotas de migração e padrões na dinâmica reprodutiva de peixes migradores e não-migradores (BAUMGARTNER et al., 2004; BIALETZKI et al., 2005; REYNALTE-TATAJE et al., 2012). Esses dados contribuem para a investigação do impacto de alterações ambientais antropogênicas, como a construção de barramentos, na reprodução dos peixes e para o delineamento de ações de manejo eficientes para a conservação das espécies e consequente manutenção da reprodução e da pesca (DA SILVA et al., 2015; SILVA et al., 2017). No entanto, a maioria dos estudos utilizando DNA *metabarcoding* para analisar o ictioplâncton até o momento são apenas metodológicos e foram realizados apenas com ovos ou larvas (LIRA et al., 2023), não demonstrando sua aplicação prática ou gerando informações incompletas sobre a dinâmica reprodutiva da ictiofauna.

No presente trabalho, utilizamos o DNA *barcoding* para analisar, individualmente, os organismos de amostras com pouco ictioplâncton, e o DNA *metabarcoding* para identificar amostras com várias larvas ou ovos agrupados, provenientes de diferentes cursos d'água à jusante e à montante da barragem de Três Marias, no rio São Francisco. Além de contribuir para a ecologia e economia ao apontar locais prioritários para a manutenção da reprodução das espécies de peixes da bacia, esta pesquisa também ajuda a demonstrar a eficácia do DNA *metabarcoding* para estudos de ictioplâncton em larga escala, utilizando o maior número de organismos, considerando ovos e larvas, dentre os trabalhos realizados na região neotropical até o momento.

5.3. Materiais e Métodos

5.3.1. Coleta e Manuseio das Amostras

A bacia do rio São Francisco foi escolhida como local amostral por possuir pelo menos 205 espécies de peixes (BARBOSA et al., 2017), sendo um ótimo modelo para testar a sensibilidade do método de DNA *metabarcoding* em regiões com alta biodiversidade. Além disso, nela se situa a Usina Hidrelétrica de Três Marias, na região central do estado de Minas Gerais, inaugurada em 1962 e cujo reservatório possui grande importância econômica e ecológica, regulando a vazão de água na região. Considerando o impacto negativo na conexão

de populações ictiológicas normalmente acarretado por barramentos, a instalação de um mecanismo de transposição de peixes na área é cogitada. Assim, foram selecionados 23 pontos amostrais nas regiões à montante e à jusante da barragem de Três Marias, em rios classificados, de acordo com a ordem de grandeza, como grandes, médios ou pequenos, a fim de verificar a ocorrência de sítios de desova e recrutamento acima e abaixo da barragem. Na região à montante, foram estudados dois rios de grande porte (Pará e São Francisco), dois médios (Bambuí e Samburá), e dois pequenos (Marmelada e Picão). Já à jusante, foram analisados dois rios grandes (Paracatu e Velhas), dois médios (Abaeté e Pacuí) e um pequeno (Roncador). Nos rios de grande porte, foram amostrados os trechos alto, médio e baixo, nos rios médios, foram amostrados os trechos alto e médio, enquanto nos rios de pequeno porte, foi amostrado apenas um ponto no trecho baixo. As classificações como trecho alto, médio ou baixo foram definidas de acordo com a área de amostragem, não de acordo com toda a área da bacia (Figura 1). As coletas foram realizadas quinzenalmente, de novembro a março, nos períodos reprodutivos consecutivos de 2019/2020 e 2020/2021, visando diminuir o impacto da aleatoriedade na representatividade das espécies que ocorrem em cada ponto.

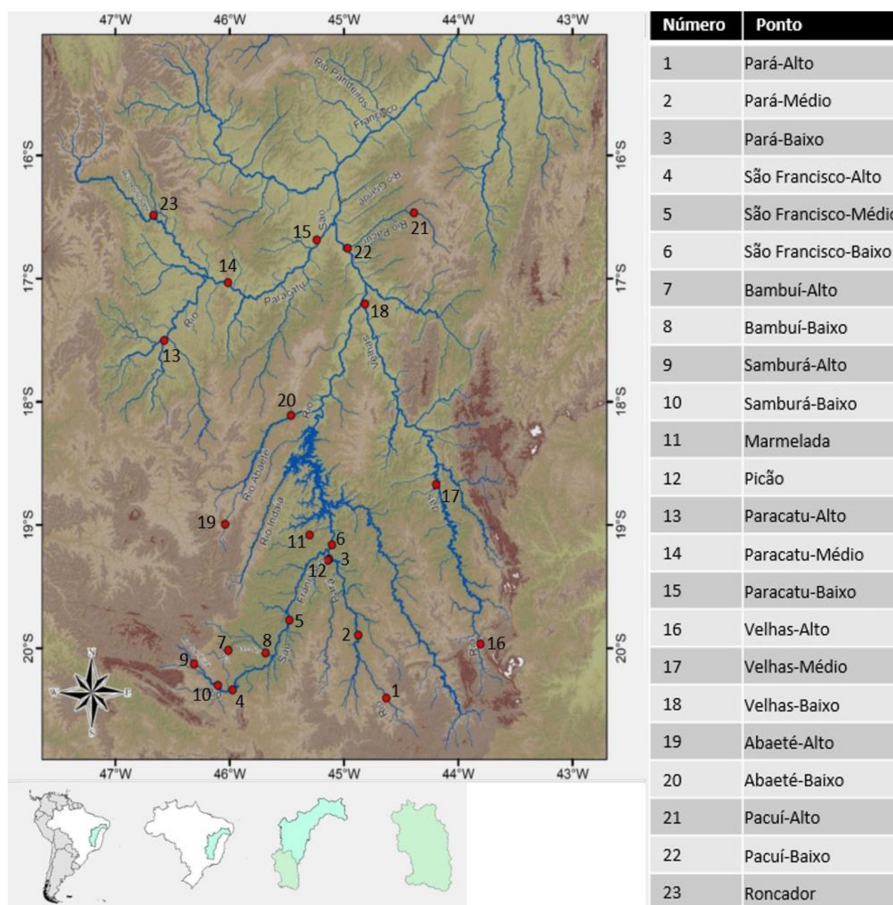


Figura 1. Mapa da área de estudo contendo os pontos amostrais, numerados de 1 a 23.

As amostras foram coletadas utilizando uma rede de plâncton cônico-cilíndrica com malha de 0,5 mm em áreas livres de vegetação. As amostragens foram realizadas na superfície das margens e do talvegue, por 10 min cada, entre o fim da tarde e o início da noite. As amostras foram acondicionadas em potes de polietileno de 500 ml, com álcool absoluto e, posteriormente, armazenadas em freezer a -4°C. Cada amostra foi triada individualmente, utilizando microscópio estereoscópico, para separar o ictioplâncton do material em suspensão. Após serem quantificados, ovos e larvas de cada amostra foram acomodados separadamente em tubos Falcon de 15 ml com álcool absoluto até a extração de DNA.

5.3.2. Extração do DNA Genômico

Para a extração genômica, amostras de um mesmo ponto, estágio de desenvolvimento (ovo ou larva) e período reprodutivo foram combinadas em um único tubo. Por exemplo, todos os ovos coletados no ponto Pará-Alto durante todas as campanhas do primeiro período reprodutivo (2019/2020) foram agrupados em uma única amostra. A ausência ou baixa quantidade de ovos ou larvas em algumas campanhas resultaram em amostras finais com poucos indivíduos para determinados pontos. Dessa forma, espécimes de amostras contendo cinco ou menos indivíduos foram processados individualmente utilizando um protocolo de extração salina adaptada de ALJANABI & MARTINEZ (1997) para posterior sequenciamento de DNA individualizado (i.e., DNA *barcoding*).

As demais amostras (i.e., agrupamentos com mais de cinco ovos ou larvas) passaram por uma etapa de preparo adicional antes de serem submetidas ao protocolo de extração, uma vez todos os organismos de cada ponto, estágio e período foram processados simultaneamente. Procurando normalizar a quantidade de DNA a ser extraída de cada organismo, larvas de maior porte foram cortadas e apenas um fragmento de tamanho similar ao tamanho total das demais foi agrupado na amostra a ser analisada. Os ovos e larvas foram transferidos para microtubos de 1,5 ml. Amostras contendo grande quantidade de organismos ou biomassa foram divididas em subamostras de volumes similares para facilitar a digestão tecidual.

Para garantir a evaporação completa do álcool, que poderia inibir reações posteriores, os tubos foram mantidos em um banho seco por 3 h a 55°C após a remoção do excesso por pipetagem. Depois de secarem completamente, adicionamos 600 µl de solução tampão TNES em cada microtubo e maceramos os tecidos agrupados com um pistilo plástico até restarem apenas fragmentos minúsculos. Em seguida, adicionamos 20 µl de proteinase K (20mg/ml) em

cada microtubo. As amostras foram mantidas a 55°C no banho seco até que, por inspeção visual, fosse constatada a completa digestão tecidual e, então, foram submetidas à extração de DNA salina. Após o fim do procedimento, os produtos de extração das amostras subdivididas foram agrupados em um único microtubo.

A fim de verificar o sucesso da extração, as amostras foram quantificadas em um fluorômetro Qubit 4 (Thermo Fisher) utilizando um kit de ensaio 1x dsDNA HS (Thermo Fisher). Em seguida, as amostras foram diluídas para uma concentração de 100 ng/μl.

5.3.3. Amplificação para o DNA *barcoding*

Organismos de amostras contendo cinco ou menos espécimes foram processados individualmente utilizando o DNA *barcoding*. Para isso, o DNA foi amplificado por reação em cadeia da polimerase (PCR), utilizando os iniciadores Fish_F1 e Fish_R1 (WARD et al., 2005) visando amplificar um fragmento de aproximadamente 650 pb do gene COI. A reação foi realizada com volume total de 25μl, contendo os seguintes reagentes: 2,5 μl de Buffer (10x); 0,75 μl de MgCl₂ (50 mM); 0,5 μl dNTP (10 mM); 0,5 μl de cada iniciador (10 mM); 0,3 μl de BSA (100 mg/mL); 0,1 μl de Taq Platinum™ (5 U/μM); 18,85 μl de água ultrapura e 1,0 μl de DNA molde. A PCR seguiu os seguintes estágios: desnaturação inicial a 94°C (2 min), seguida de 35 ciclos desnaturação a 94°C (30 seg), anelamento dos iniciadores a 56°C (30 seg) e extensão a 72°C (30 seg), com uma extensão final a 72°C (5 min). O resultado da PCR foi conferido em gel de agarose 1,8%.

5.3.4. Amplificação para o DNA *metabarcoding*

Para o DNA *metabarcoding*, amplificamos diferentes fragmentos do gene 12S rRNA utilizando os marcadores moleculares NeoFish (193 pb – MILAN et al., 2020) e MiFish (185 pb, MIYA et al., 2015). O marcador NeoFish foi desenvolvido utilizando um banco de sequências majoritariamente provenientes da bacia do rio São Francisco, enquanto o MiFish é um dos marcadores 12S mais utilizados em estudos de peixe do mundo. Segundo MILAN et al. (2020), ambos possuem resolução adequada para a identificação precisa da ictiofauna da bacia do rio São Francisco. Os iniciadores foram modificados com a inclusão de caudas (*tags*) específicas para cada amostra, possibilitando o agrupamento e sequenciamento de todas as

amostras simultaneamente em uma plataforma e NGS, e posterior separação por bioinformática (Figura 2).

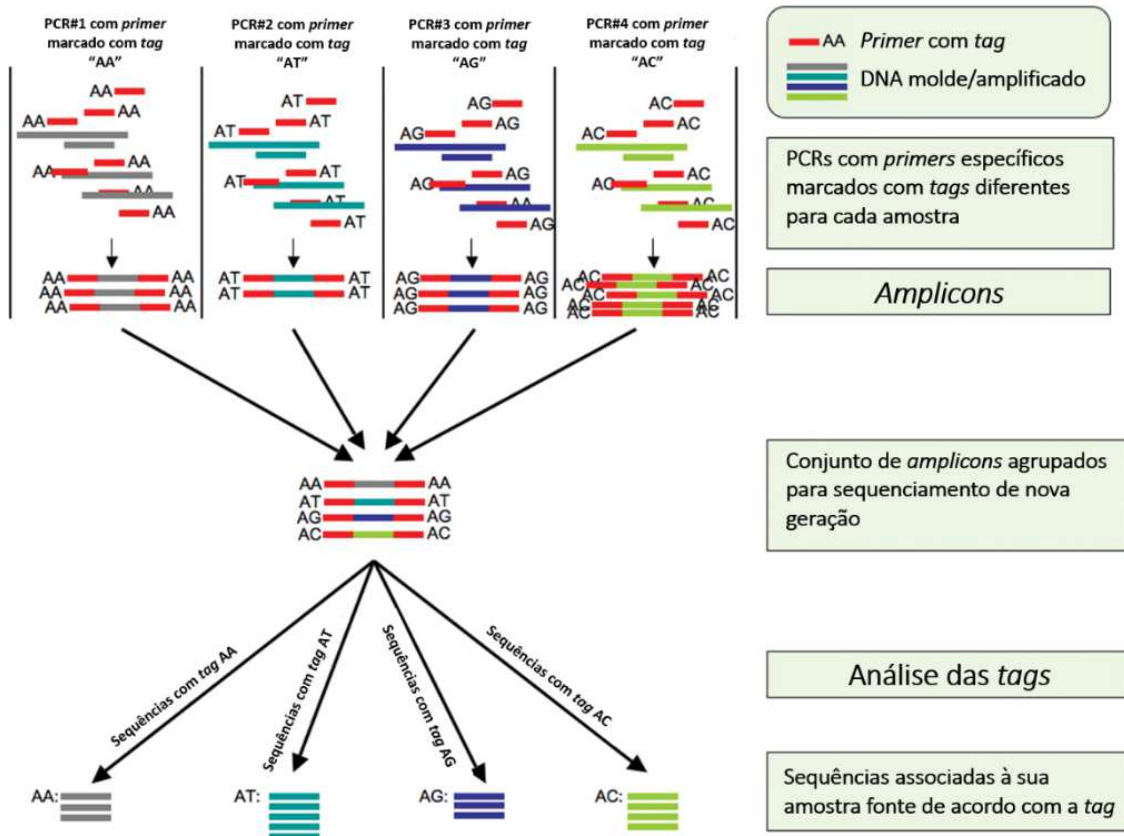


Figura 2. Utilização de iniciadores modificados para permitir a associação de cada sequência à sua amostra fonte. Adaptado de BINLADEN et al. (2007).

Para cada amostra, foram preparadas três réplicas, visando mitigar os impactos dos vieses de amplificação que a PCR pode causar na representatividade de cada organismo. Além disso, em cada rodada de PCR, foi incluída uma amostra de controle negativo, com adição de água ultrapura ao invés de DNA molde, no intuito de monitorar possíveis contaminações. Adicionalmente, foi realizada uma PCR com uma amostra da espécie marinha *Epinephelus itajara*, utilizada como controle positivo e para detectar a ocorrência de *tag jumping*, quando sequências de uma amostra recebem a *tag* de outra e causam detecções positivas falsas (SCHNELL et al., 2015).

Para cada reação, utilizamos uma solução com volume total de 20 µl contendo: 8,34 µl de água ultrapura, 0,16 µl de BSA, 10 µl de Master Mix AmpliTaq GOLD 360 (ThermoFisher),

0,25 µl de cada iniciador (10 µM), e 1,0 µl de DNA molde. Para as amostras de controle negativo, adicionamos mais 1,0 µl de água ultrapura ao invés de DNA molde.

As condições da PCRs consistiram em uma desnaturação inicial a 95°C (10 min), seguidos de 35 ciclos de desnaturação a 95°C (1 min), anelamento de iniciadores a 63°C (NeoFish) ou 60°C (MiFish) (30 seg), e extensão a 72 °C (30 seg), com extensão final a 72 °C (5 min). O resultado da PCR foi conferido em gel de agarose 1,8%. A amplificação do DNA foi bem-sucedida para todas amostras, exceto para a amostra contendo 181 ovos do segundo período reprodutivo do rio Marmelada, e nenhum controle negativo apresentou banda, mas ainda assim foram submetidos ao sequenciamento junto com as demais amostras.

5.3.5. Preparo de Bibliotecas e Sequenciamento de DNA

Para a realização do DNA *barcoding*, as reações de sequenciamento foram realizadas para ambas as fitas de DNA, de maneira bidirecional, utilizando o Kit comercial Big Dye® Terminator Cycle Sequencing Kit v3.1 de acordo com as instruções do fabricante, pela empresa ACTGene Análises Moleculares. As amostras foram sequenciadas individualmente seguindo o método de Sanger em um sequenciador automático.

Para o DNA *metabarcoding*, os produtos de amplificação com iniciadores marcados com *tags* exclusivas para cada amostra foram combinados na construção de duas bibliotecas (MiFish ou NeoFish). As bibliotecas foram preparadas para o sequenciamento utilizando o kit livre de PCR Collibri™ PS DNA Library Prep Kit (Thermo Fisher) seguindo as instruções do fabricante. A seleção de tamanho dos fragmentos foi realizada usando esferas magnéticas 1.8x Agencourt AMPure XP (Beckman Coulter) e as bibliotecas foram quantificadas utilizando o Qubit 4. O sequenciamento foi realizado utilizando a plataforma Illumina MiniSeq, com um kit Mid Output (300 ciclos) seguindo o protocolo do fabricante, no Laboratório de Genética da Conservação da PUC Minas.

5.3.6. Análises de Bioinformática

Para analisar as amostras submetidas ao DNA *barcoding*, foram montadas sequências consenso utilizando as duas fitas complementares no programa DNA Baser v5.08 (DNA Sequence Assembler v4 2013). A conferência visual e edição manual das sequências para remover extremidades ambíguas foram realizadas no programa MEGA7 (KUMAR et al. 2016),

que também foi empregado para o alinhamento das sequências com o algoritmo ClustalW (THOMPSON et al., 1994). As sequências obtidas foram comparadas com sequências disponíveis nos bancos de dados públicos GenBank e BOLD (Barcode of Life Data System). O limiar utilizado para a identificação em nível de espécie foi de 98% de similaridade, baseado no trabalho de WARD (2009).

Para o DNA *metabarcoding*, as análises de bioinformática foram realizadas utilizando um *pipeline* (sequência de etapas e comandos) customizado escrito no R v4.4.0 (R Core Team, 2021). A remoção de iniciadores e adaptadores foi feita usando o pacote Cutadapt (MARTIN, 2011). O pacote de análises de *metabarcoding* DADA2 (CALLAHAN et al., 2016) foi utilizado para filtrar leituras com baixa qualidade, para juntar as fitas em ASVs (*Amplicon Sequence Variants*) usando um algoritmo que remove ASVs quiméricas. As ASVs são todas as variantes dos *amplicons* (produtos da PCR) sequenciados que, posteriormente, são associadas a determinada unidade taxonômica operacional (OTU) de acordo com a similaridade com a sequência daquela OTU usada como referência. O pacote DADA2 também foi empregado para realizar a atribuição taxonômica, com um banco de sequência referência customizado para o gene 12S, baseado no desenvolvido por MILAN et al. (2020), contendo 261 sequências de DNA de 129 espécies de peixes neotropicais, e o banco de dados público GenBank. O banco customizado conta, majoritariamente, com sequências de espécies provenientes das bacias dos rios São Francisco e Jequitinhonha. Para identificações em nível de espécie, foram consideradas similaridades acima de 99%, baseado em valores encontrados por MILAN et al. (2020). Amostras com similaridades inferiores ou com alta similaridade com mais de uma espécie ou gênero, foram identificadas em nível de gênero ou subfamília, respectivamente, e assim sucessivamente.

Todas as ASVs encontradas nos controles negativos foram completamente removidas das demais amostras, exceto quando ASVs com identificações compatíveis também foram detectadas pelo outro marcador na mesma amostra. Além disso, devido ao alto rendimento do processo de amplificação, mesmo pequenas contaminações podem influenciar nos resultados de sequenciamento. Para diminuir o risco de analisarmos sequências provenientes de contaminações ou falsos positivos, mas também visando não deixar de incluir espécies subrepresentadas nas amostras, nós consideramos válidas apenas as ASVs frequência maior que 0,01%. Todos os resultados passaram por uma etapa de curadoria após a identificação por bioinformática. A curadoria foi realizada manualmente para cada ASV a fim de verificar se elas

havia sido associadas a uma ou mais identificações com porcentagens de similaridade próximas, e adequar a identificação ao nível taxonômico hierárquico mais adequado.

Para comparar o número de táxons encontrados entre região ou estágio de desenvolvimento, diagramas de Venn foram construídos com o auxílio do aplicativo de web Lucidchart (<https://lucid.app/>). Uma PERMANOVA (*Permutational Multivariate Analysis of Variance*) foi realizada para calcular as dissimilaridades entre a diversidade de peixes detectada em cada região ou estágio de desenvolvimento, usando o índice de dissimilaridade Jaccard baseado na presença/ausência de táxons, com 1.000 permutações. As diferenças na composição de táxons foram plotadas em uma PCoA (*Principal Coordinate Analysis*) para facilitar a visualização. O cálculo do índice de dissimilaridade e a PCoA foram realizados utilizando os comandos “adonis2” e “betadisper”, respectivamente, com o pacote vegan (OKSANEN et al., 2018).

5.4. Resultados

5.4.1. Triagem

Considerando os dois períodos reprodutivos, foram triados 12.539 ovos, sendo 11.025 provenientes de amostras coletadas à montante e 1.514 daquelas à jusante (Tabela 1), representando 87,93% e 12,07% do total, respectivamente (Figura 3a). Em ambas regiões, a soma das amostras dos rios de maior grandeza apresentou a maior quantidade de ovos, seguida pelos rios médios e, por fim, pelos rios de pequenos (Tabela 1).

Tabela 1. Número de ovos encontrados por ponto, rio, grandeza e região.

Região	Total Região	Grandeza	Total Grandeza	Rio	Total Rio	Ponto	Total Ponto		
Montante	11025	Grande	7450	Pará	4911	Alto	4833		
						Médio	66		
						Baixo	12		
				São Francisco	2539	Alto	2359		
						Médio	161		
						Baixo	19		
		Médio	3313			Bambuí	502	Alto	110
								Baixo	392
						Samburá	2811	Alto	18
								Baixo	2793
Pequeno	262			Marmelada	195	-	195		
				Picão	67	-	67		
Jusante	1514	Grande	819	Paracatu	468	Alto	42		
						Médio	32		

			Baixo	394
			Alto	0
		Velhas	Médio	236
			Baixo	115
Médio	563	Abaeté	Alto	90
			Baixo	463
		Pacuí	Alto	4
			Baixo	6
Pequeno	132	Roncador	-	132

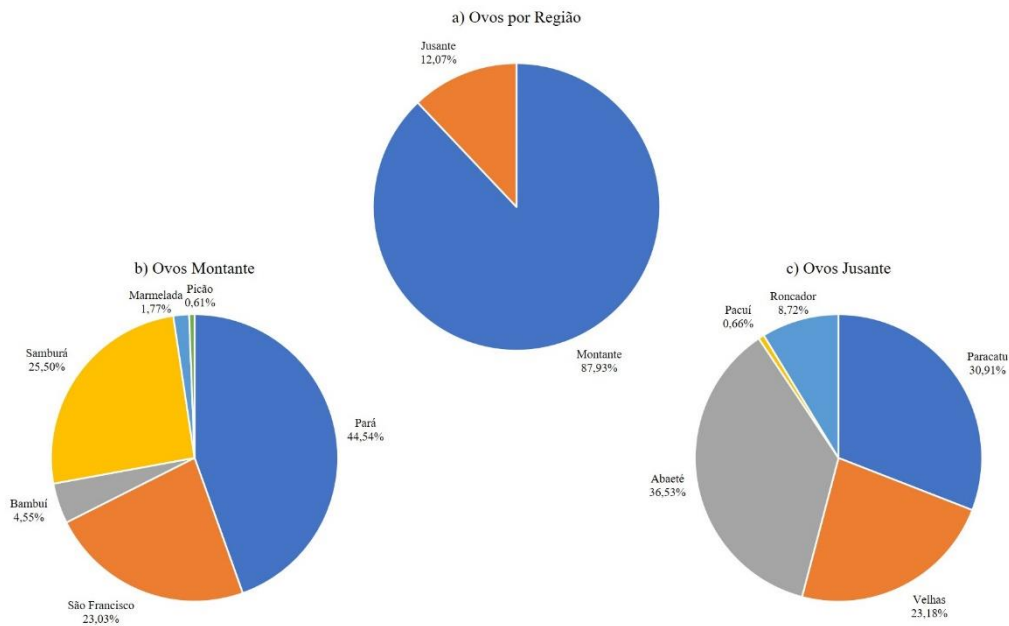


Figura 3. Distribuição de ovos coletados (a) por região, (b) por rio à montante e (c) por rio à jusante.

A maior quantidade de ovos (4.911) foi coletada no rio Pará, de grande porte, correspondendo a 44,54% do total de ovos encontrados à montante. Ainda à montante, o rio Samburá, de médio porte, apresentou mais ovos que o São Francisco, de grande porte (Tabela 1, Figura 3b). À jusante, o rio onde foi coletado o maior número de ovos foi o Abaeté (553), de médio porte, representando 36,53% de todos ovos da região. O rio Pacuí, de médio porte, apresentou a menor contribuição (10) para o total de ovos coletados à jusante, ficando atrás do Roncador, de pequeno porte (132) (Tabela 1, Figura 3c).

Diferentemente do que foi observado para os ovos, a maior quantidade de larvas foi encontrada na região à jusante, com 13.041 organismos que representaram 85,66% do total. Já para a região à montante, foram triadas 2.184 larvas, correspondendo a apenas 14,34% do total (Tabela 2, Figura 4a). Para os organismos nesse estágio, os rios de grande porte também foram os mais representativos, mas enquanto na região à jusante, os rios pequenos apresentaram

maior número de larvas que os médios, à jusante os médios contribuíram mais para o quantitativo total que o pequeno (Tabela 2).

Tabela 2. Número de larvas encontradas por ponto, rio, grandeza e região.

Região	Total Região	Grandeza	Total Grandeza	Rio	Total Rio	Ponto	Total Ponto		
Montante	2184	Grande	1924	Pará	112	Alto	50		
						Médio	38		
						Baixo	24		
				São Francisco	1812	Alto	10		
						Médio	287		
		Médio	61	Pequeno	199	Bambuí	48	Alto	6
						Samburá	13	Baixo	42
						Marmelada	194	Alto	0
						Picão	5	Baixo	13
								-	194
Jusante	13041	Grande	12812	Paracatu	11120	Alto	13		
						Médio	10060		
						Baixo	1047		
				Velhas	1692	Alto	1		
						Médio	1474		
		Médio	200	Pequeno	29	Abaeté	44	Baixo	217
						Pacuí	156	Alto	2
						Roncador	29	Baixo	42
								Alto	0
								Baixo	156

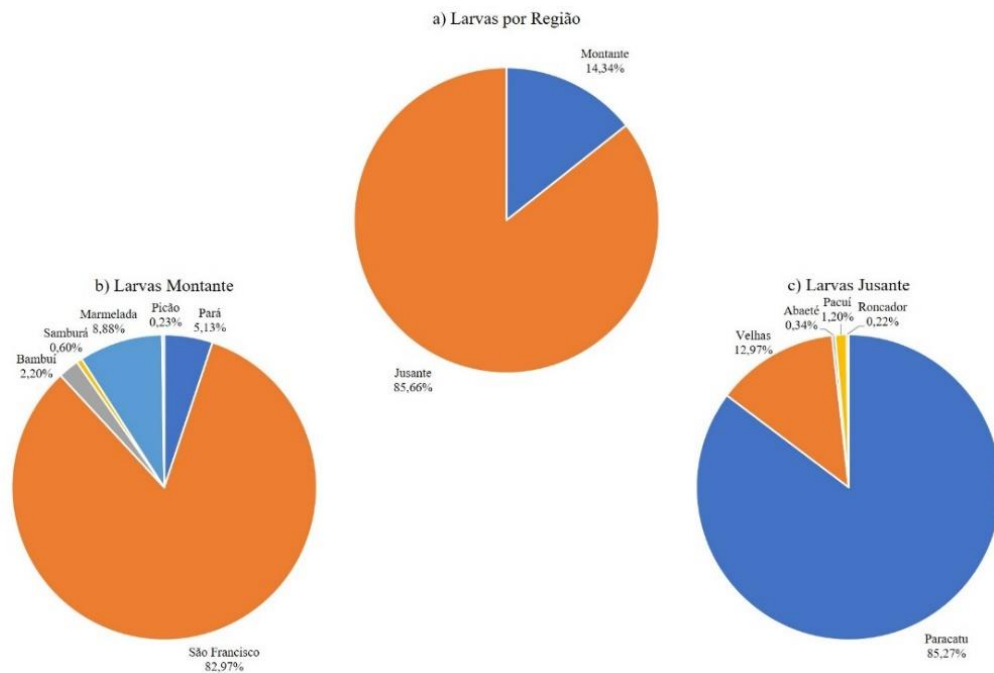


Figura 4. Distribuição de larvas coletadas no primeiro ciclo reprodutivo por região (a), por rio à montante (b) e por rio à jusante (c).

A grande maioria das larvas foi coletada no rio de grande porte Paracatu, 11.448, compondo 85,27% daquelas coletadas à jusante. Nessa região, rios de maior grandeza apresentaram maior número de larvas que os menores, e o rio Roncador foi aquele com menor número de organismos coletados (Tabela 2, Figura 4c). Já à montante, a maior contribuição veio do rio São Francisco (1.812), com 82,97% do total de larvas coletadas na região. Diferentemente do observado à jusante, porém, o rio com a segunda maior quantidade de organismos nesse estágio foi o Marmelada (194), de pequeno porte (Tabela 2, Figura 4b).

5.4.2. Identificação molecular

O DNA *barcoding* foi utilizado para identificar 39 organismos (24 ovos e 15 larvas) de 15 amostras. Ao todo, foram detectados 19 táxons, sendo 16 em nível de espécie, dois em nível de gênero e um em nível de família. Quatro desses táxons foram detectados exclusivamente com o DNA *barcoding*: três em nível de espécie (*Bergiaria westermanni*, *Characidium fasciatum* e *Sternopygus macrurus*) e um em nível de família (Trichomycteridae) (Tabela 3). Das três espécies detectadas exclusivamente utilizando o DNA *barcoding*, *B. westermanni* e *C. fasciatum* não possuem sequências de 12S disponíveis nos bancos de dados de referência empregados.

Tabela 3. Lista de famílias e táxons identificados por rio de acordo com o estágio. 1: rio Pará, 2: rio São Francisco, 3: rio BambuÍ, 4: rio Samburá, 5: rio Marmelada, 6: rio Picão, 7: rio Paracatu, 8: rio das Velhas, 9: rio Abaeté, 10: rio Pacuí, 11: rio Roncador, O: táxon detectado em amostras de ovos, L: detectado em amostra de larvas, OL: detectado em amostras de ovos e larvas, ^M: detectado exclusivamente com MiFish, ^N: detectado exclusivamente com NeoFish, ^B: detectado exclusivamente com DNA *barcoding*, *: espécies consideradas grandes migradoras, **: espécie com grande similaridade genética com *Piaractus mesopotamicus*, cujos híbridos são comuns na piscicultura.

Família	Táxon	Rios											
		1	2	3	4	5	6	7	8	9	10	11	
Acestrorhynchidae	<i>Acestrorhynchus lacustris</i>			L									
Anostomidae	<i>Leporinus piau</i>	OL	OL	OL	OL	OL	O	OL			L	L	
	<i>Leporinus taeniatus</i>	OL	OL	O	OL		L	OL	OL	O	L	OL	
	<i>Megaleporinus obtusidens</i>	O	OL	OL	O	OL	O	OL	OL	O	L	O	
	<i>Megaleporinus macrocephalus</i> ^M										L		
	<i>Megaleporinus reinhardti</i>	OL	OL	OL	OL				OL	OL	O	L	O
	<i>Schizodon knerii</i>	OL	OL	OL	L		L		OL	OL	O	L	O
Auchenipteridae	<i>Glanidium albescens</i>	L	O					L		O			
Bryconidae	<i>Brycon orthotaenia</i> *							OL		O			
	<i>Salminus franciscanus</i> *				O			L	OL	OL			
	<i>Salminus hilarii</i> *			L				L					
Characidae	<i>Astyanax lacustris</i>			L						L			
	<i>Lepidocharax burnsi</i> ^N		L	L						O			
	<i>Moenkhausia costae</i> ^M			L									
	<i>Oligosarcus sp.</i> ^M	L											
	<i>Piabarchus stramineus</i> ^N				O								
	<i>Piabina argentea</i>	OL								L			
	<i>Planaltina myersi</i> ^M			L	L						O		
	<i>Psalidodon fasciatus</i>	OL	OL	L		L	L	OL	OL	O		O	
	<i>Psalidodon rivularis</i>	OL	O							L			
<i>Tetragonopterus franciscoensis</i>	L								OL				
Cichlidae	<i>Oreochromis niloticus</i> ^M	OL	OL							OL			
Crenuchidae	<i>Characidium fasciatum</i> ^B											O	
	<i>Characidium zebra</i>	O	O										
Curimatidae	<i>Curimatella lepidura</i>	L	OL	OL	L			OL	OL		O	O	
	<i>Cyphocharax gilbert</i>								L				
Engraulidae	<i>Anchoviella vaillanti</i>							L					
Erythrinidae	<i>Hoplias intermedius</i> ^N		L	L		L							
	<i>Hoplias malabaricus</i>		L										
Heptapteridae	<i>Cetopsorhamdia iheringi</i>					O							
	<i>Pimelodella vittata</i>		L										
	<i>Rhamdia quelen</i>		O	L						L			
<i>Incertae sedis</i>	<i>Conorhynchos conirostris</i> ^{M*}							L					
Loricariidae	<i>Pterygoplichthys etentaculatus</i>									L			
Parodontidae	<i>Apareiodon piracicabae</i>									L			
	<i>Parodon hilarii</i>	O						L			O	L	
Pimelodidae	<i>Bergiaria westermanni</i> ^B							O			O		
	<i>Iheringichthys labrosus</i> ^M	O	O	O				OL	O	O			
	<i>Pimelodus fur</i>	OL	OL	OL				OL	OL	OL	O	OL	
	<i>Pimelodus maculatus</i>	OL	OL	OL	OL		L	OL	OL	OL	O	OL	
	<i>Pimelodus pohli</i> ^M	OL	OL	OL	O			OL	L	O		L	
Prochilodontidae	<i>Pseudoplatystoma corruscans</i> *	O						OL		O		L	
	<i>Prochilodus argenteus</i> *	OL	OL	L		OL	OL	OL	OL	OL	L	OL	
	<i>Prochilodus costatus</i> *	OL	OL	OL	OL	OL	O	OL	OL	OL	L	O	

Pseudopimelodidae	<i>Pseudopimelodus</i> sp.				L					
Sciaenidae	<i>Pachyurus squamipennis</i>	OL	OL	L	L		OL	OL	O	L
Serrasalminidae	<i>Colossoma macropomum</i> **						OL	L		
	<i>Myleus micans</i>	L	OL	O			OL	L	OL	L
	<i>Pygocentrus piraya</i>	L	OL	L	OL		L	OL	OL	L
	<i>Serrasalmus brandtii</i> ^M			L						
Sternopygidae	<i>Eigenmannia virescens</i>			L						
	<i>Sternopygus macrurus</i> ^B	L								
Trichomycteridae	Trichomycteridae ^B				O					
Triporthidae	<i>Triportheus guentheri</i>	OL	OL		L		L	OL	OL	O

O esforço de sequenciamento para o DNA *metabarcoding* resultou em 4.506.120 leituras analisáveis somando os dois marcadores, após as etapas de controle de qualidade e curadoria. Para o MiFish, foram obtidas 2.362.312 leituras ao todo, com uma média de 39.372 leituras por amostra, variando de 10.881 leituras para a amostra de ovos do ponto baixo do rio Samburá no primeiro período reprodutivo a 458.180 leituras para a amostra de ovos do ponto médio do São Francisco no segundo período. Já para o NeoFish, o total de leituras foi 2.143.808, resultando em 35.730 leituras em média por amostra, com a amostra de larvas do primeiro período reprodutivo do alto São Francisco apresentando apenas 937 leituras e a amostra de larvas do primeiro período do baixo rio Pará somando 72.992 leituras.

Ao todo, foram identificados 50 táxons de peixes utilizando o DNA *metabarcoding*, sendo 48 em nível de espécie e dois em nível de gênero (*Oligosarcus* sp. e *Pseudopimelodus* sp.), considerando todos os rios (Tabela 3). Identificações em nível de gênero ou em níveis hierárquicos superiores só foram mantidas caso não fosse identificado nenhum táxon em nível mais específico que as representasse. Dentre as identificações em nível de espécie, oito táxons foram detectados exclusivamente com o MiFish (*Conorhynchos conirostris*, *Iheringichthys labrosus*, *Megaleporinus megacephalus*, *Moenkhausia costae*, *Oreochromis niloticus*, *Pimelodus pohli*, *Planaltina myersi* e *Serrasalmus brandtii*), enquanto três foram detectados apenas pelo NeoFish (*Hoplias intermedius*, *Lepidocharax burnsi* e *Piabarchus stramineus*) (Tabela 3).

Considerando ovos e larvas, 54 táxons detectados somando DNA *barcoding* e *metabarcoding*, foram identificadas 51 espécies, 43 gêneros e 22 famílias (contabilizando *incertae sedis* como a família de *C. conirostris* para fins de comparação) (Tabela 3). Dentre os ovos analisados, foram encontrados 36 táxons, representados por 35 espécies, 29 gêneros e 15 famílias. Destes, seis táxons (11,11% do total), cinco espécies (9,80%), quatro gêneros (9,30%) e duas famílias (9,09%) foram detectados exclusivamente nas amostras de ovos (Tabela 3, Figura 5). Já entre as larvas, foram encontrados 48 táxons, compostos por 46 espécies, 39

gêneros e 19 famílias. Exclusivamente, as larvas foram responsáveis pela detecção de 18 táxons (33,33% do total), 16 espécies (31,37%), 15 gêneros (34,88%) e seis famílias (31,82%) (Tabela 3, Figura 5). Levando em conta os rios isoladamente, os ovos contribuíram com um aumento de até 61,90% táxons detectados (13 dos 21 táxons encontrados no rio Abaeté) e as larvas, com um aumento de até 60% (nove dos 15 táxons detectados no rio Pacuí) (Tabela 3).

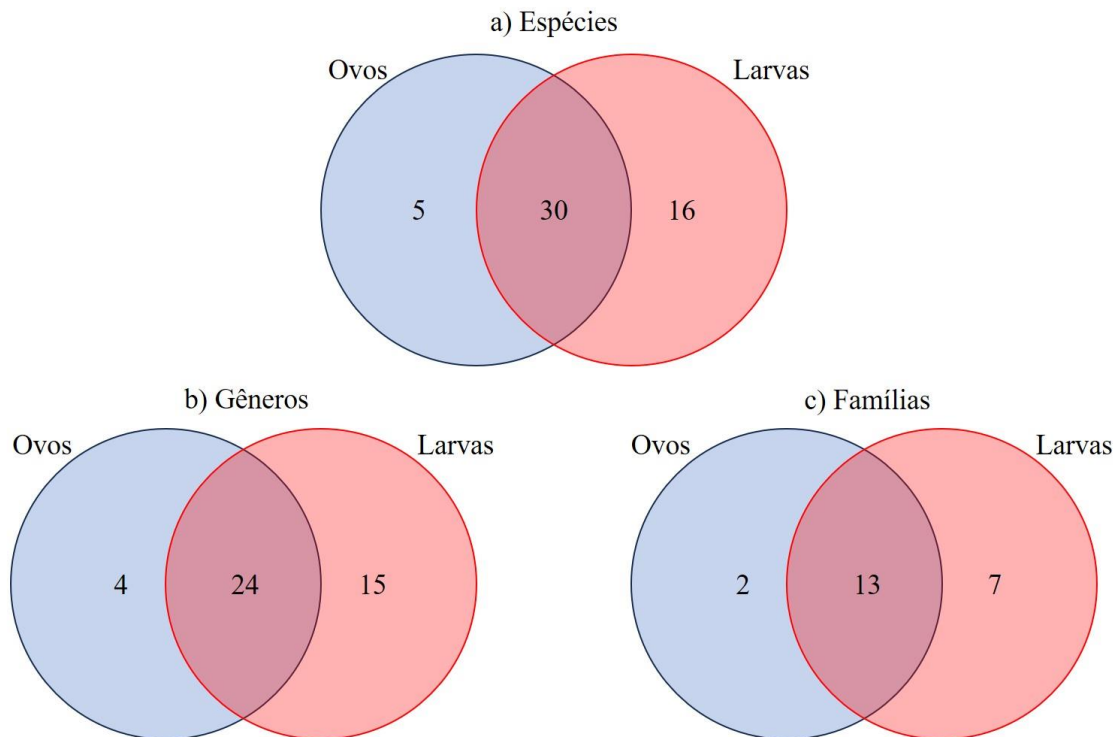


Figura 5. Diagramas de Venn representando o número de táxons identificados entre ovos e larvas por (a) espécie, (b) gênero e (c) família.

Apesar da grande contribuição de táxons detectados exclusivamente em cada estágio de desenvolvimento dependendo do caso analisado, a PERMANOVA evidenciou diferenciação baixa e estatisticamente não significativa ($R^2 = 0,0448$, $p = 0,51$) entre as comunidades de ovos ou larvas de cada rio (Tabela 4). Concordantemente, na PCoA é possível observar grande sobreposição entre os dois estágios, mas demonstrando um maior número de táxons detectados entre as larvas que não figuraram entre os ovos (Figura 6).

Tabela 4. Resumo dos resultados da PERMANOVA (R^2 -tamanho do efeito e níveis de significância) mostrando o efeito do estágio de desenvolvimento nos táxons detectados. Df: grau de liberdade, SumSqs: soma dos quadrados, valor do teste F , e valor de p associado ao teste F .

Índice	Df	SumSqs	R^2	F	p
Jaccard	1	0,3154	0,0448	0,9373	0,51

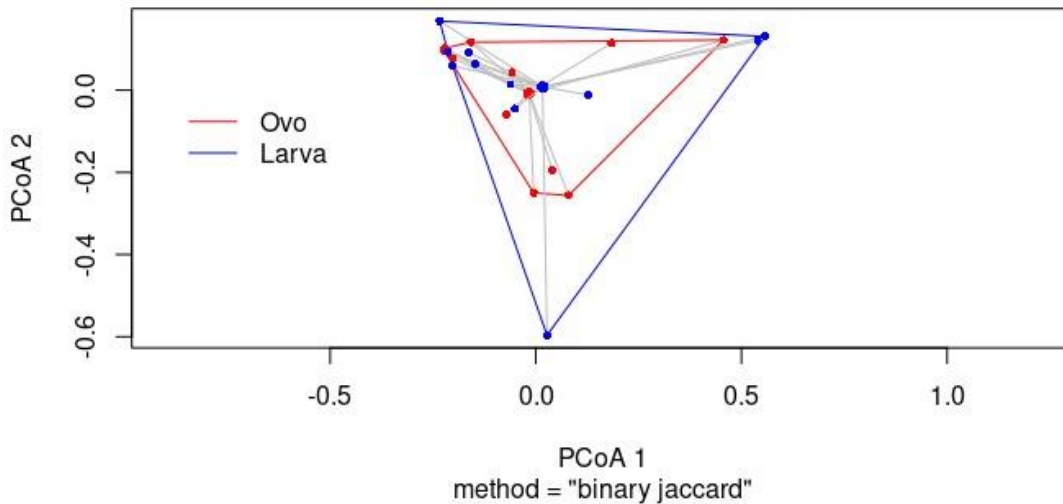


Figura 6. PCoA com as comunidades de peixes encontradas em cada rio analisado, de acordo com o estágio de desenvolvimento: ovo ou larva.

Na comparação entre os rios, o Paracatu, com 28 táxons, foi aquele que apresentou maior riqueza somando todos os organismos, seguido por Pará, São Francisco, Bambuí e Velhas, com 27 táxons (Figura 7). Já os rios Picão e Marmelada foram os que apresentaram o menor número de táxons detectados: oito e sete, respectivamente. Considerando apenas organismos em estágio larval, o Paracatu se manteve à frente, com 27 táxons, seguido por Velhas com 26, Bambuí com 23, e Pará e São Francisco com 22. Para organismos nesse estágio, o Picão e o Marmelada também apresentaram menor riqueza, com cinco e seis táxons, respectivamente. Já entre os ovos analisados, o maior número de táxons foi identificado no rio São Francisco, com 22, seguido por Abaeté, com 21, Pará com 20, e Paracatu, com 19, enquanto os menores foram Pacuí, Marmelada e Picão, com seis, cinco e quatro, respectivamente (Figura 7).

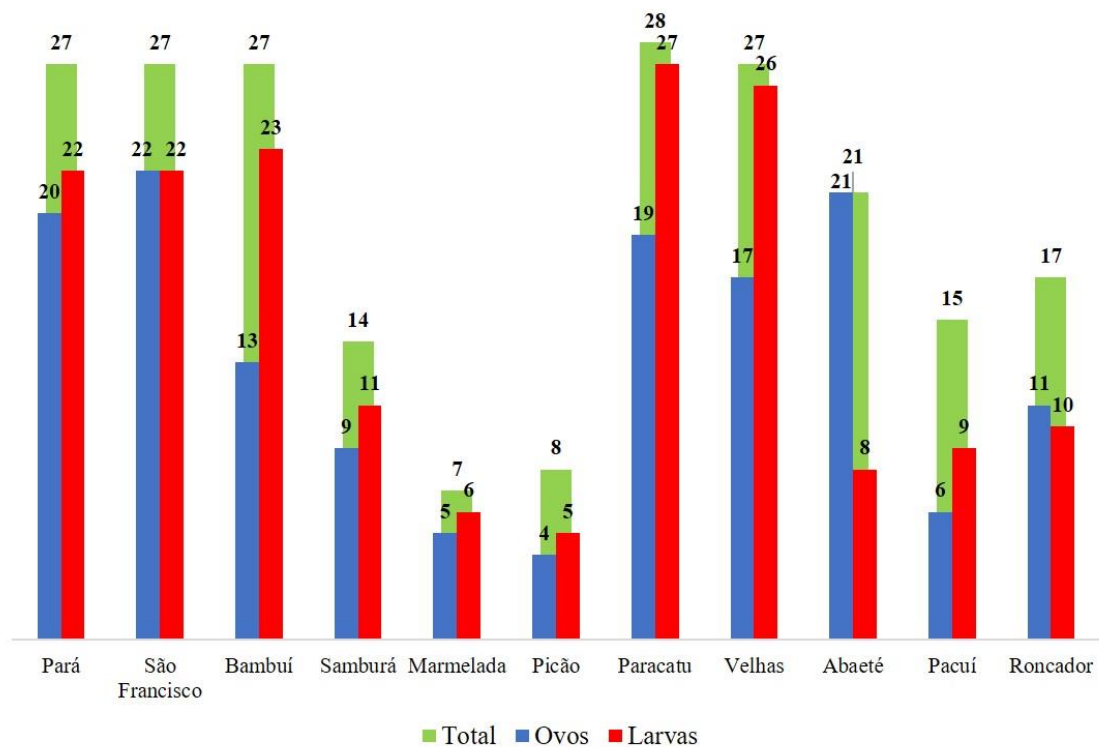


Figura 7. Número de táxons identificados por rio, nas amostras de ovos, de larvas e total.

Das 51 espécies encontradas, sete foram consideradas grandes migradoras (*Brycon orthotaenia*, *C. conirostris*, *Prochilodus argenteus*, *P. costatus*, *Pseudoplatystoma corruscans*, *Salminus franciscanus* e *S. hilarii*), com importância ecológica e econômica destacadas (Tabela 3). O rio Paracatu foi o único onde todas as sete foram encontradas (as sete em estágio larval, mas apenas *B. orthotaenia*, *P. argenteus*, *P. costatus* e *P. corruscans* dentre os ovos). Além disso, a espécie *C. conirostris* foi identificada apenas neste rio, considerando ambas as regiões. Dentre os outros rios da região à jusante, a espécie *B. orthotaenia* foi encontrada no Abaeté como ovo, *P. argenteus* em todos como larva e exceto no Pacuí como ovo, *P. costatus* exceto no Roncador como larva e exceto no Pacuí como ovo, *P. corruscans* no Roncador como larva e no Abeté como ovo, *S. franciscanus* no Velhas e Abaeté como ovo e larva, e *S. hilarii* não foi encontrado em nenhum outro rio. À montante, *P. argenteus* foi detectada em todos os rios, exceto Samburá, como larva, e exceto Bambuí e Samburá como ovo, *P. costatus* em todos como ovo e exceto no Picão como larva, *P. corruscans* apenas no Pará como ovo, *S. franciscanus* apenas no Samburá como ovo, e *S. hilarii* apenas no Bambuí como larva (Tabela 3).

Considerando cada região separadamente, à montante da barragem de Três Marias foram detectados 44 táxons, compostos por 41 espécies, 36 gêneros e 19 famílias (Tabela 3,

Figura 8). Dentre estes, 14 táxons (25,93% do total), 12 espécies (23,53%), 12 gêneros (27,91%) e cinco famílias (22,73%) foram detectados exclusivamente à montante. Já à jusante foram detectados 40 táxons, 40 espécies, 32 gêneros e 17 famílias. Exclusivamente, a região à jusante contribuiu com 10 táxons (18,52%), 10 espécies (19,61%), oito gêneros (18,60%) e três famílias (13,64%) (Tabela 3, Figura 8). Dentre as sete espécies consideradas grandes migradoras, duas (*B. orthotaenia* e *C. conirostris*) foram encontradas exclusivamente à jusante, mas nenhuma foi detectada exclusivamente à montante (Tabela 3).

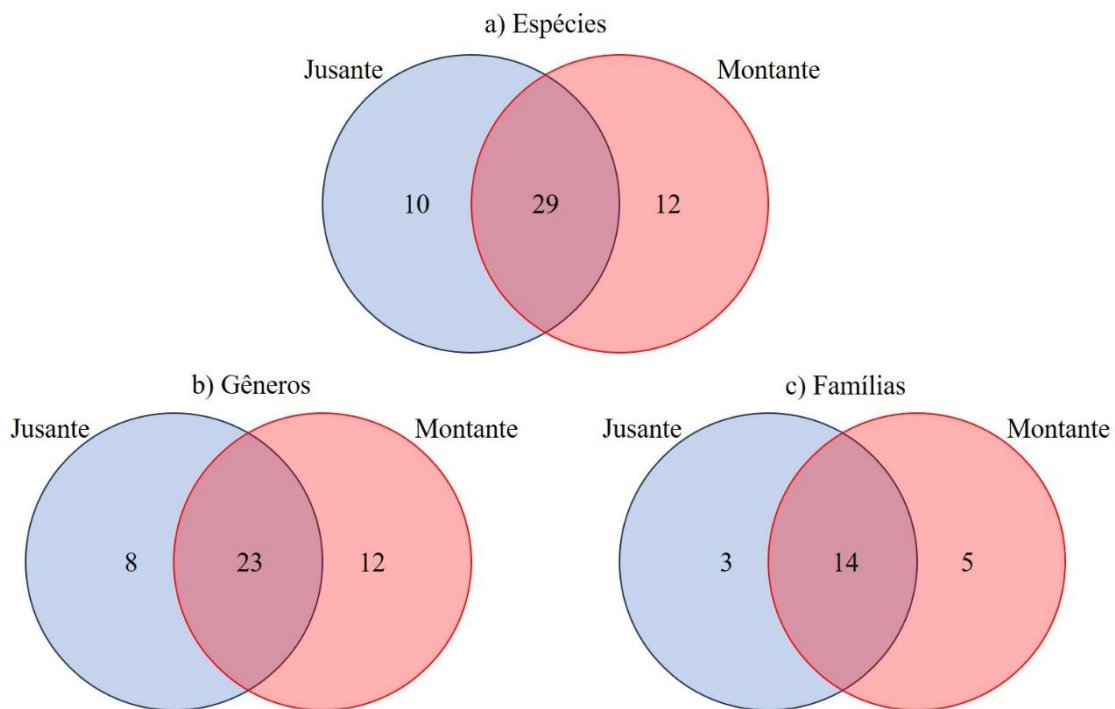


Figura 8. Diagramas de Venn representando o número de táxons identificados entre amostras da jusante e montante por (a) espécie, (b) gênero e (c) família.

Assim como na análise considerando os diferentes estágios de desenvolvimento, a PERMANOVA evidenciou diferenciação baixa e não significativa ($R^2 = 0,0599$, $p = 0,115$) na comparação entre as regiões montante e jusante (Tabela 5). Em concordância, a maioria dos pontos das regiões na PCoA ficaram próximos ou sobrepostos (Figura 9).

Tabela 5. Resumo dos resultados da PERMANOVA (R^2 -tamanho do efeito e níveis de significância) mostrando o efeito da região nos táxons detectados. Df: grau de liberdade, SumSqs: soma dos quadrados, valor do teste F , e valor de p associado ao teste F .

Índice	Df	SumSqs	R^2	F	p
Jaccard	1	0,4218	0,0599	1,2735	0,51

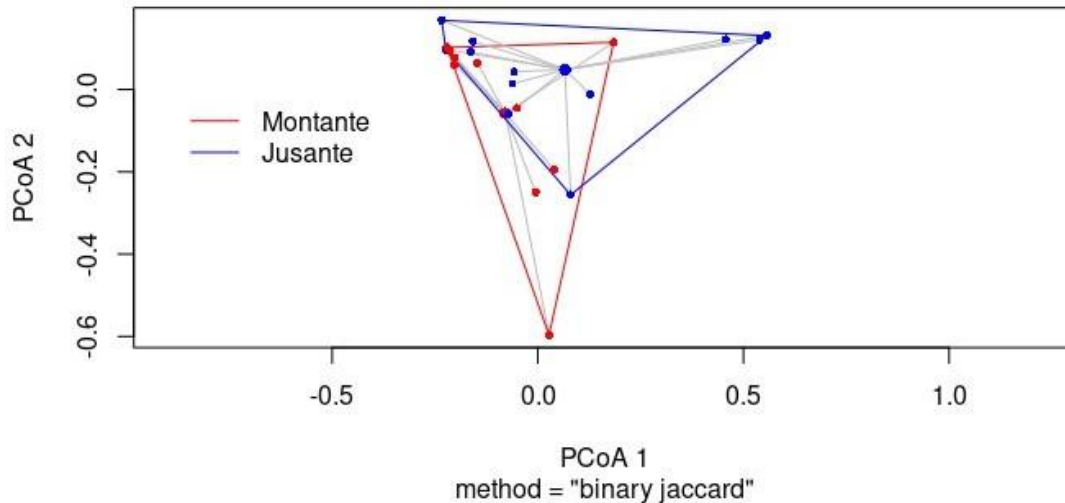


Figura 9. PCoA com as comunidades de peixes encontradas em cada rio analisado, de acordo com a região: montante ou jusante.

5.5. Discussão

Com a crescente demanda por técnicas inovadoras para estudar a complexa dinâmica reprodutiva da ictiofauna, tanto na pesquisa quanto em demandas comerciais para avaliar o impacto das alterações antropogênicas, o DNA *metabarcoding* tem se mostrado uma ferramenta valiosa para analisar o ictioplâncton (CARVALHO, 2022). Entretanto, a maior parte dos trabalhos até o momento focaram em aspectos metodológicos e poucos exploraram sua aplicação prática, ou o fizeram utilizando apenas um dos estágios de desenvolvimento: ovo ou larva (LIRA et al., 2023). Apesar de o refinamento contante da técnica ser de suma importância para que ela possa produzir resultados cada vez mais confiáveis, avaliar a sua aplicação para responder perguntas do mundo real é fundamental, principalmente em regiões amplamente impactadas por ações antropogênicas, como a construção de barramentos. A investigação do ictioplâncton é fundamental para compreender diversos aspectos da história de vida dos peixes,

como seus locais de reprodução e desova, o uso de habitats, contribuindo para a conservação e manutenção da pesca (MILLER & KENDALL, 2009).

O presente estudo empregou o DNA *metabarcoding* para identificar o maior número de ovos e larvas de peixes dentre os trabalhos publicados na região neotropical até o momento, utilizando como modelo a área de influência da barragem de Três Marias, localizada no rio São Francisco, com ictiofauna megadiversa. Ao todo, foram processados mais de 12 mil ovos e 15 mil larvas, e quase todos os 54 táxons detectados foram identificados em nível de espécie (51), reforçando a eficiência e praticidade do método já destacadas em trabalhos anteriores (MAGGIA et al., 2017; KIMMERLING et al., 2018; MARIAC et al., 2018; SILVA et al., 2023). Apesar do grande número de táxons detectados, esse número representa 26,34% das 205 espécies registradas na bacia do São Francisco (BARBOSA et al., 2017). Duas possíveis explicações são a seletividade do método de coleta, que prioriza a amostragem de organismos à deriva na superfície, e a não ocorrência de determinadas espécies na área amostrada. No entanto, a efemeridade do processo de reprodução da ictiofauna pode ter impactado negativamente o número de organismos coletados em determinados rios, fazendo com que a reprodução local possa ter sido subestimada. Dois exemplos são o rio Pacuí, de médio porte e localizado à jusante da barragem de Três Marias, que contou com apenas 10 ovos, e o rio Picão, de pequeno porte e localizado à montante, representado por apenas cinco larvas. Buscando evitar esse tipo de gargalo amostral, POMPEU et al. (2023) sugeriram reduzir o intervalo de amostragem, uma vez que encontraram grande diferença entre amostragens semanais e quinzenais.

Ainda assim, os resultados aqui encontrados permitiram fazer algumas inferências sobre a distribuição espacial do ictioplâncton na região da barragem de Três Marias. Apesar de a análise estatística ter demonstrado grande sobreposição entre as comunidades encontradas à montante e à jusante do barramento, as espécies detectadas exclusivamente nessas regiões representaram 23,53% e 19,61% do total, respectivamente. Mesmo que a maior parte desse montante seja composto por espécies sedentárias, duas das sete espécies consideradas como grandes migradoras e, portanto, dentre as principais afetadas pelo barramento, foram encontradas apenas na região à jusante: *B. orthotaenia* e *C. conirostris*. A espécie *B. orthotaenia* foi detectada nos ovos coletados no rio Abaeté e entre os ovos e larvas coletados no rio Paracatu. O rio Paracatu foi, ainda, o único local onde foram encontradas larvas da espécie ameaçada *C. conirostris* (ICMBio, 2018). Esse é, possivelmente, o primeiro registro de ictioplâncton da

espécie na natureza, uma vez que seus locais de desova e recrutamento são desconhecidos (GODINHO & GODINHO, 2018).

Adicionalmente, apenas no Paracatu foram encontradas todas as sete grandes migradoras, além de ser o rio com a maior riqueza, ressaltando sua importância como local prioritário para a conservação e manutenção da reprodução das principais espécies de peixes da bacia, principalmente sabendo-se que tributários são habitats importantes para a reprodução de peixes migratório (LIMA et al., 2020). Já à montante, três das cinco grandes migradoras identificadas na região foram detectadas em apenas um tributário: *S. franciscanus* entre os ovos coletados no rio Samburá, *S. hilarii* entre as larvas do Bambuí, e *P. corruscans* entre os ovos do rio Pará. As espécies de curimba, outras duas grandes migradoras, *P. argenteus* e *P. costatus*, foram coletadas como ovos e larvas em vários tributários nas duas regiões. Apesar de a presença de ovos e larvas serem bons indicadores de possíveis locais de desova, o ictioplâncton pode se manter à deriva por vários dias (NAKATANI et al., 2001; VASCONCELOS et al., 2022). Dessa forma, experimentos aprofundados são recomendados para a delimitação mais precisa das regiões de desova e recrutamento, tanto para o Paracatu como para os demais rios aqui estudados (e. g., MIRANDA-CHUMACERO et al., 2020).

Das 48 espécies detectadas utilizando o DNA *metabarcoding*, oito foram detectadas exclusivamente com o marcador MiFish, enquanto três foram detectadas apenas com o NeoFish. Apesar de ser uma proporção pequena, esses resultados reforçam achados anteriores de que o emprego de um único marcador pode acarretar subestimação da diversidade amostrada (ZHANG et al., 2020; LIU & ZHANG, 2021; TEIXEIRA et al., 2023). Além disso, das três espécies identificadas exclusivamente pelo DNA *barcoding* utilizando o gene COI, duas não possuem sequências de 12S referência, corroborando como bancos incompletos podem atrapalhar a detecção de espécies (COLLINS et al., 2019).

Outro fator que pode impactar o número de espécies detectadas é a análise de organismos em diferentes estágios de desenvolvimento. Devido à ausência de caracteres morfológicos durante o estágio de ovo, muitas vezes os trabalhos sobre reprodução optam pela identificação apenas de larvas (BAUMGARTNER et al., 2004; REYNALTE-TATAJE et al., 2012). No entanto, FRANTINE-SILVA et al. (2015) demonstraram, utilizando identificação molecular, a importância da inclusão de ovos nessas análises, uma vez que eles contribuíram com 30% da riqueza observada. No presente trabalho, apesar de a análise estatística apontar diferenciação baixa e não significativa entre os táxons detectados por ovos e larvas, a inclusão de ovos na análise foi responsável por até 61,90% da riqueza observada em determinado

tributário. Além disso, devido ao movimento de deriva do ictioplâncton de acordo com a velocidade da água, a análise dos ovos pode indicar maior proximidade do sítio de desova em relação ao ponto de coleta.

Em conclusão, o DNA *metabarcoding* se mostrou uma ferramenta prática e eficiente para analisar a dinâmica reprodutiva de uma comunidade de peixes megadiversa, permitindo o processamento de múltiplos indivíduos simultaneamente. Além disso, o emprego dessa técnica permitiu a inclusão de ovos e de larvas danificadas na análise, contribuindo para uma representação real das espécies que estão reproduzindo na região estudada. Apesar de estudos mais aprofundados serem recomendados, os resultados aqui encontrados permitiram apontar um tributário de fundamental importância para a manutenção da reprodução e, conseqüentemente, da pesca de espécies de grande interesse ecológico e econômico.

REFERÊNCIAS

- ALJANABI, S. M.; MARTINEZ, I. Universal and rapid salt-extraction of high-quality genomic DNA for PCR-based techniques. **Nucleic Acids Research**, v. 25, n. 22, p. 4692–4693, 1997.
- BARBOSA, J. M. et al. Profile of the fish fauna of São Francisco basin. **Acta of Fisheries and Aquatic Resources**, v. 5, p. 70-90, 2017.
- BAUMGARTNER, G. et al. Identification of spawning sites and natural nurseries of fishes in the upper Paraná River, Brazil. **Environmental Biology of Fishes**, v. 72, n. 2, p. 115-125, 2004.
- BECKER, R. A. et al. DNA barcoding and morphological identification of neotropical ichthyoplankton from the Upper Paraná and São Francisco. **Journal of Fish Biology**, v. 87, n. 1, p. 159–168, 2015.
- BIALETZKI, A. et al. Larval fish assemblage in the Baía River (Mato Grosso do Sul State, Brazil): Temporal and spatial patterns. **Environmental Biology of Fishes**, v. 73, p. 37-47, 2005.
- BIALETZKI, A. et al. O estudo de ovos e larvas de peixes. In: ORSI, M. L. et al. (Orgs.). **Ovos, larvas e juvenis dos peixes da Bacia do Rio Parapanema uma avaliação para a conservação**. Assis: Triunfal Gráfica e Editora, p. 17-24, 2016. 136 p.
- BINLADEN, J. et al. The Use of Coded PCR Primers Enables High-Throughput Sequencing of Multiple Homolog Amplification Products by 454 Parallel Sequencing. **PLoS ONE**, v. 2, n. 2, e197, 2007.
- CALLAHAN, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. **Nature Methods**, v. 13, n. 7, p. 581-583, 2016.

CARVALHO, D. C. Ichthyoplankton DNA metabarcoding: Challenges and perspectives. **Molecular Ecology**, v. 31, n. 6, p. 1612-1614, 2022.

COLLINS, R. A. et al. Non-specific amplification compromises environmental DNA metabarcoding with COI. **Methods in Ecology and Evolution**, v. 10, n. 11, p. 1985-2001, 2019.

DA SILVA, P. S. et al. Importance of Reservoir Tributaries to Spawning of Migratory Fish in the Upper Paraná River. **River Research and Applications**, v. 31, n. 3, p. 313-322, 2015.

FRANTINE-SILVA, W. et al. DNA barcoding of freshwater ichthyoplankton in the Neotropics as a tool for ecological monitoring. **Molecular Ecology Resources**, v. 15, n. 5, p. 1226–1237, 2015.

GODINHO, A. L.; GODINHO, H. P. Fisheries, life-history and conservation status of the catfish pirá *Conorhynchus conirostris* (Ostariophysi: Siluriformes) in Brazil. **Neotropical Ichthyology**, v. 16, n. 1, e170131, 2018.

HARVEY, J. B. J. et al. Comparison of morphological and next generation DNA sequencing methods for assessing zooplankton assemblages. **Journal of Experimental Marine Biology and Ecology**, v. 487, p. 113-126, 2017.

ICMBIO. INSTITUTO CHICO MENDES DE CONSERVAÇÃO DA BIODIVERSIDADE. Livro Vermelho da Fauna Brasileira Ameaçada de Extinção: Volume VI - Peixes. In: Instituto Chico Mendes de Conservação da Biodiversidade (Org.). **Livro Vermelho da Fauna Brasileira Ameaçada de Extinção**. Brasília: ICMBio. 2018. 1232p.

KIMMERLING, N. et al. Quantitative species-level ecology of reef fish larvae via metabarcoding. **Nature Ecology and Evolution**, v. 2, n. 2, p. 306–316, 2018.

KO, H. L. et al. Evaluating the Accuracy of Morphological Identification of Larval Fishes by Applying DNA Barcoding. **PLoS ONE**, v. 8, n. 1, e53451, 2013.

KOCHZIUS, M. Trends in fishery genetics. In: BEAMISH, R. J.; ROTHSCILD, B. J. (Orgs.). 'The Future of Fishery Science in North America'. **Springer Science + Business Media B.V.**, p. 453–493, 2009.

KUMAR, S., et al. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. **Molecular Biology and Evolution**, v. 33, n. 7, p. 1870-1874, 2016.

LIMA, M. C. C. et al. Use of DNA barcode in the identification of fish eggs in tributaries of the Paranapanema River Basin. **Genetics and Molecular Biology**, v. 43, n. 3, p. 1–9, 2020.

LIRA, N. L. et al. Identifying fish eggs and larvae: from classic methodologies to DNA metabarcoding. **Molecular Biology Reports**, v. 50, n. 2, p. 1713-1726, 2023.

LIU, J.; ZHANG, H. Combining multiple markers in environmental DNA metabarcoding to assess deep-sea benthic biodiversity. **Frontiers in Marine Science**, 8, 684955, 2021.

MAGGIA, M. E. et al. DNA Metabarcoding of Amazonian Ichthyoplankton Swarms. **PLoS ONE**, v. 12, n. 1, e0170009, 2017.

MARIAC, C. et al. Metabarcoding by capture using a single COI probe (MCSP) to identify and quantify fish species in ichthyoplankton swarms. **PLoS ONE**, v. 13, n. 9, e0202976, 2018.

MARTIN, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. **EMBnet. Journal**, v. 17, n. 1, p. 10–12, 2011.

MILAN, D. T. et al. New 12S metabarcoding primers for enhanced Neotropical freshwater fish biodiversity assessment. **Scientific Reports**, v. 10, n. 1, 17966, 2020.

MILLER, B. S.; KENDALL, A. W. **Early Life History of Marine Fishes**. Bekerley, California: University of California Press, 2009. 364 p.

MIRANDA-CHUMACERO, G. et al. Threatened fish spawning area revealed by specific metabarcoding identification of eggs and larvae in the Beni River, Upper Amazon. **Global Ecology and Conservation**, v. 24, e01309, 2020.

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, n. 7, 150088, 2015.

NAKATANI, K. et al. **Ovos e Larvas de peixes de água doce: Desenvolvimento e manual de identificação**. Maringá, Brasil: EDUEM, 2001. 378p.

NOBILE, A. B. et al. DNA metabarcoding of neotropical ichthyoplankton: Enabling high accuracy with lower cost. **Metabarcoding and Metagenomics**, v. 3, e35060, 2019.

OKSANEN, J. et al. vegan: community ecology package. **R package version 2.5-2**. 2018. Disponível em: <https://CRAN.R-project.org/package=vegan>.

POMPEU, P. S. et al. Inadequate Sampling Frequency and Imprecise Taxonomic Identification Mask Results in Studies of Migratory Freshwater Fish Ichthyoplankton. **Fishes**, v. 8, n. 10, 2023.

R CORE TEAM. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, 2021. Disponível em: <https://www.R-project.org/>.

REYNALTE-TATAJE, D. A. et al. Spatial and temporal variation of the ichthyoplankton in a subtropical river in Brazil. **Environmental Biology of Fishes**, v. 94, p. 403-419, 2012.

RICHARDSON, D. E.; VANWYE, J. D.; EXUM, A. M.; COWEN, R. K.; CRAWFORD, D. L. High-throughput species identification: from DNA isolation to bioinformatics. **Molecular Ecology Notes**, v. 7, n. 2, p. 199–207, 2007.

SCHNELL, I. B. et al. Tag jumps illuminated – reducing sequence-to-sample misidentifications in metabarcoding studies. **Molecular Ecology Resources**, v. 15, n. 6, p. 1289-1303, 2015.

- SILVA, C. B. et al. Fish larvae diversity in a conservation area of a neotropical floodplain: influence of temporal and spatial scales. **Hydrobiologia**, 2017.
- SILVA, L. E. et al. Amazonian ichthyoplankton assessment via DNA metabarcoding: A baseline for detecting spawning sites of migratory fishes. **Biological Conservation**, v. 284, 110180, 2023.
- TABERLET, P. et al. Towards next-generation biodiversity assessment using DNA *metabarcoding*. **Molecular Ecology**, v. 21, n. 8, p. 2045–2050, 2012.
- TEIXEIRA, D. F. et al. DNA metabarcoding assessment of Neotropical ichthyoplankton communities is marker-dependent. **Ecology and Evolution**, v. 13, e10649, 2023.
- THOMPSON, J. D. et al. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v. 22, p. 4673–4680, 1994.
- VASCONCELOS, L. P. et al. Fish eggs and larvae drifting through hydropower reservoirs: a case study in the Brazilian Amazon. **Hydrobiologia**, v. 849, p. 357-372, 2022.
- WARD, R. D. et al. DNA barcoding Australia's fish species. **Philosophical Transactions of the Royal Society, B: Biological Sciences**, v. 360, n. 1462, p. 1847–1857, 2005.
- WARD, R. D. DNA barcode divergence among species and genera of birds and fishes. **Molecular Ecology Resources**, v. 9, n. 4, p. 1077-1085, 2009.
- YU, D. W. et al. Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. **Methods in Ecology and Evolution**, v. 3, n. 4, p. 613–623, 2012.
- ZHANG, G. K. et al. A comprehensive and comparative evaluation of primers for metabarcoding eDNA from fish. **Methods in Ecology and Evolution**, v. 11, n. 12, p. 1609-1625, 2020.

6. CONSIDERAÇÕES FINAIS

Em um mundo cada vez mais desenvolvido, modificações antropogênicas da natureza são frequentes para atender a demanda energética crescente. No Brasil, onde a energia hidrelétrica corresponde a 56,80% da matriz de eletricidade (EPE, 2023), o planejamento da construção de mais hidrelétricas costuma ser bem visto por ser considerada uma opção mais barata e segura (PRADO et al., 2016). Os barramentos para a construção de reservatórios fragmentam o hábitat, prejudicando a migração e o acesso dos peixes a sítios de desova e recrutamento (ANTONIO et al., 2007; PELICICE et al. 2014). Em uma região onde a maior parte das bacias hidrográfica já se encontra fragmentada (NILSSON et al., 2005), a conservação de tributários com fluxo livre se torna ainda mais importante, uma vez que eles são fundamentais para o abrigo e reprodução de espécies migradoras (DA SILVA, et al. 2015; ROSA et al., 2017; LIMA et al., 2020; VASCONCELOS et al., 2021).

A identificação dos tributários prioritários para a conservação visando garantir a manutenção da reprodução dos peixes de uma região, principalmente migradores, depende fundamentalmente da detecção dos sítios de desova e recrutamento (POMPEU et al., 2011). Nesse contexto, o interesse e a demanda de estudos acadêmicos e práticos do ictioplâncton tem crescido. No entanto, sem a identificação precisa dos organismos, locais importantes para a reprodução dos peixes podem passar despercebidos (POMPEU et al., 2023).

O DNA *metabarcoding* vem se destacando em estudos de ictioplâncton, por ser uma ferramenta poderosa para a identificação de vários organismos simultaneamente de maneira prática, rápida e acessível (KIMMERLING et al., 2018; MARIAC et al., 2018; POMPEU et al., 2023; SILVA et al., 2023). Dessa forma, ele permite atender à crescente demanda de dados taxonômicos e rápida resposta necessária para embasar os esforços atuais de conservação e manejo da biodiversidade. Com isso em mente, vários estudos metodológicos têm sido feitos visando aumentar a confiabilidade dos resultados do DNA *metabarcoding* para o estudo do ictioplâncton (MAGGIA et al., 2017; MARIAC et al., 2018; NOBILE et al. 2019; DUKE & BURTON, 2020). Aqui, nós demonstramos a importância da utilização de múltiplos marcadores em um estudo, reduzindo os vieses causados por bancos incompletos (LOCATELLI et al., 2020) ou afinidade dos iniciadores com as sequências das espécies analisadas. Além disso, visando aumentar a acessibilidade do método, utilizamos um protocolo de extração salina de baixo custo e um protocolo de preparo de biblioteca livre de PCR, durante

o qual todas as amostras amplificadas utilizando *tags* específicas podem ser agrupadas, eliminando vieses de amplificação adicionais.

Apesar do grande potencial da técnica, ela ainda tem sido pouco aplicada aos ecossistemas megadiversos brasileiros. O nosso trabalho se junta ao de POMPEU et al. (2023) e SILVA et al. (2023) como um dos primeiros a utilizar o DNA *metabarcoding* para responder questões ecológicas sobre a dinâmica reprodutiva da ictiofauna brasileira. Aqui, nós comparamos diferentes tributários na região de influência da barragem de Três Marias e demonstramos que, apesar de as comunidades de ictioplâncton encontradas à jusante e à montante do barramento não apresentarem diferenciação estatisticamente significativa, os locais analisados podem contribuir de maneira desigual para a reprodução de espécies consideradas grandes migradoras. Em especial, o rio Paracatu se mostrou um ponto prioritário para a conservação. Além de apresentar a maior riqueza, nele todas as sete espécies consideradas grandes migradoras foram detectadas, inclusive *C. conirostris*, espécie ameaçada, com distribuição restrita, e cujos locais de reprodução permaneciam desconhecidos (GODINHO & GODINHO, 2018). Dessa forma, este é um tributário elegível a ser preservado livre de barragens, a fim de se manter como ponto para abrigo e reprodução, principalmente de espécies migradoras.

As técnicas moleculares podem ser escaladas mais facilmente que as tradicionais para gerar maior quantidade de dados e ainda permitem a identificação de organismos que muitas vezes seriam ignorados, como larvas danificadas e ovos, com poucos caracteres morfológicos. No estudo de FRANTINE-SILVA et al. (2015), por exemplo, os ovos contribuíram exclusivamente com 30% das espécies detectadas. Aqui, eles compuseram até 61,90% da riqueza observada quando analisados os tributários isoladamente. Outra vantagem da inclusão dos ovos nos estudos de reprodução é que eles podem indicar com maior precisão os sítios de desova, uma vez que o ictioplâncton pode se manter à deriva por vários dias (NAKATANI et al., 2001; VASCONCELOS et al., 2022).

Dessa forma, o avanço observado para o DNA *metabarcoding* no estudo do ictioplâncton fazem dele uma ferramenta valiosa para o monitoramento, principalmente em regiões antropizadas. Trabalhos como o de MIRANDA-CHUMACERO et al. (2020) devem ser incentivados para a delimitação mais precisa de sítios de desova e recrutamento, principalmente de espécies de grande interesse ecológico e econômico, como as grandes migradoras, ou ameaçadas, como *C. conirostris*. A técnica ainda tende a se beneficiar do crescente interesse

pelos estudos com DNA ambiental (BENG & CORLETT, 2020), que tem estimulado a ampliação dos bancos de referência.

REFERÊNCIAS

AHLSTROM, E. H.; MOSER, H. G. Eggs and larvae of fishes and their role in systematic investigations and in fisheries. **Revue Des Travaux de L'institut Des Peches Maritimes**, v. 40, n. 3, p. 379–398, 1976.

ALMEIDA, F. S. et al. DNA barcoding as a useful tool for identifying non-native species of freshwater ichthyoplankton in the neotropics. **Hydrobiologia**, v. 817, p. 111-119, 2018.

ANDERSEN, K. et al. Meta-barcoding of ‘dirt’ DNA from soil reflects vertebrate biodiversity. **Molecular Ecology**, v. 21, n. 8, p. 1966–1979, 2012.

ANTONIO, R. R. et al. Blockage of migration routes by dam construction: can migratory fish find alternative routes? **Neotropical Ichthyology**, v. 5, n. 2, p. 177-184, 2007.

BAUMGARTNER, G. et al. Identification of spawning sites and natural nurseries of fishes in the upper Paraná River, Brazil. **Environmental Biology of Fishes**, v. 72, n. 2, p. 115-125, 2004.

BECKER, R. A. et al. DNA barcoding and morphological identification of neotropical ichthyoplankton from the Upper Paraná and São Francisco. **Journal of Fish Biology**, v. 87, n. 1, p. 159–168, 2015.

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.

BIALETZKI, A. et al. Larval fish assemblage in the Baía River (Mato Grosso do Sul State, Brazil): Temporal and spatial patterns. **Environmental Biology of Fishes**, v. 73, p. 37-47, 2005.

BIALETZKI, A. et al. O estudo de ovos e larvas de peixes. In: ORSI, M. L. et al. (Orgs.). **Ovos, larvas e juvenis dos peixes da Bacia do Rio Parapanema uma avaliação para a conservação**. Assis: Triunfal Gráfica e Editora, p. 17-24, 2016. 136 p.

BINLADEN, J. et al. The Use of Coded PCR Primers Enables High-Throughput Sequencing of Multiple Homolog Amplification Products by 454 Parallel Sequencing. **PLoS ONE**, v. 2, n. 2, e197, 2007.

DA SILVA, P. S. et al. Importance of Reservoir Tributaries to Spawning of Migratory Fish in the Upper Paraná River. **River Research and Applications**, v. 31, n. 3, p. 313-322, 2015.

DEAGLE, B. E. et al. DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. **Biology Letters**, v. 10, n. 9, 20140562, 2014.

DUKE, E. M.; BURTON, R. S. Efficacy of metabarcoding for identification of fish eggs evaluated with mock communities. **Ecology and Evolution**, v. 10, n. 7, p. 3463–3476, 2020.

EPE. EMPRESA DE PESQUISA ENERGÉTICA. **Balço Energético Nacional 2023: Ano base 2022**. Empresa de Pesquisa Energética, Rio de Janeiro: EPE, 2023.

FONSECA, V. G. Pitfalls in relative abundance estimation using eDNA metabarcoding. **Molecular Ecology Resources**, v. 18, n. 5, p. 923–926, 2018.

FONSECA, V. G. et al. Second-generation environmental sequencing unmask marine metazoan biodiversity. **Nature Communications**, v. 1, n. 1, p. 98, 2010.

FRANTINE-SILVA, W. et al. DNA barcoding of freshwater ichthyoplankton in the Neotropics as a tool for ecological monitoring. **Molecular Ecology Resources**, v. 15, n. 5, p. 1226–1237, 2015.

FREITAS-SOUZA, D. et al. Genetic markers indicate that hybrids of *Pseudoplatystoma* (Siluriformes, Pimelodidae) are reproducing in natural environments in southeastern Brazil. **Biological Invasions**, v. 24, p. 1049-1058, 2022.

GODINHO, A. L.; GODINHO, H. P. Fisheries, life-history and conservation status of the catfish pirá *Conorhynchos conirostris* (Ostariophysi: Siluriformes) in Brazil. **Neotropical Ichthyology**, v. 16, n. 1, e170131, 2018.

HEBERT, P. D. et al. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. **Proceedings of the Royal Society of London. Series B: Biological Sciences**, v. 270, (suppl_1), p. S96–S99, 2003.

JØRGENSEN, T. et al. A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. **Molecular Ecology**, v. 21, n. 8, p. 1989–2003, 2012.

KIMMERLING, N. et al. Quantitative species-level ecology of reef fish larvae via metabarcoding. **Nature Ecology and Evolution**, v. 2, n. 2, p. 306–316, 2018.

KO, H. L. et al. Evaluating the Accuracy of Morphological Identification of Larval Fishes by Applying DNA Barcoding. **PLoS ONE**, v. 8, n. 1, e53451, 2013.

LIMA, M. C. C. et al. Use of DNA barcode in the identification of fish eggs in tributaries of the Paranapanema River Basin. **Genetics and Molecular Biology**, v. 43, n. 3, p. 1–9, 2020.

LIRA, N. L. et al. Identifying fish eggs and larvae: from classic methodologies to DNA metabarcoding. **Molecular Biology Reports**, v. 50, n. 2, p. 1713-1726, 2023.

LOCATELLI, N. S. et al. GenBank's reliability is uncertain for biodiversity researchers seeking species-level assignment for eDNA. **Proceedings of the National Academy of Sciences**, v. 117, n. 51, p. 32211–32212, 2020.

- MAGGIA, M. E. et al. DNA Metabarcoding of Amazonian Ichthyoplankton Swarms. **PLoS ONE**, v. 12, n. 1, e0170009, 2017.
- MARIAC, C. et al. Metabarcoding by capture using a single COI probe (MCSP) to identify and quantify fish species in ichthyoplankton swarms. **PLoS ONE**, v. 13, n. 9, e0202976, 2018.
- MILAN, D. T. et al. New 12S metabarcoding primers for enhanced Neotropical freshwater fish biodiversity assessment. **Scientific Reports**, v. 10, n. 1, 17966, 2020.
- MIRANDA-CHUMACERO, G. et al. Threatened fish spawning area revealed by specific metabarcoding identification of eggs and larvae in the Beni River, Upper Amazon. **Global Ecology and Conservation**, v. 24, e01309, 2020.
- 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, n. 7, 150088, 2015.
- NAKATANI, K. et al. **Ovos e Larvas de peixes de água doce: Desenvolvimento e manual de identificação**. Maringá, Brasil: EDUEM, 2001. 378p.
- NOBILE, A. B. et al. DNA metabarcoding of neotropical ichthyoplankton: Enabling high accuracy with lower cost. **Metabarcoding and Metagenomics**, v. 3, e35060, 2019.
- PELICICE, F. M. et al. Large reservoirs as ecological barriers to downstream movements of neotropical migratory fish. **Fish and Fisheries**, v. 16, n. 4, p. 697-715, 2014.
- POMPEU, P. S. et al. Existing and future challenges: the concept of successful fish passage in South America. **River Research and Applications**, v. 28, n. 4, p. 504-512, 2011.
- POMPEU, P. S. et al. Inadequate Sampling Frequency and Imprecise Taxonomic Identification Mask Results in Studies of Migratory Freshwater Fish Ichthyoplankton. **Fishes**, v. 8, n. 10, 2023.
- PRADO, A. P. et al. How much is enough? An integrated examination of energy security, economic growth and climate change related to hydropower expansion in Brazil. **Renewable and Sustainable Energy Reviews**, v. 53, p.1132–1136, 2016.
- REYNALTE-TATAJE, D. A. et al. Spatial and temporal variation of the ichthyoplankton in a subtropical river in Brazil. **Environmental Biology of Fishes**, v. 94, p. 403-419, 2012.
- RODRIGUES, T. et al. Occurrence and identification of Istiophoridae larvae and Xiphiidae eggs off the southeastern Brazilian Coast. **Boletim do Instituto de Pesca**, v. 43, n. 1, p. 78-86, 2017.
- ROSA, G. R. et al. Spatial and temporal distribution of ichthyoplankton during an unusual period of low flow in a tributary of the São Francisco River, Brazil. **River Research and Applications**, v. 34, n. 1, p. 69-82, 2017.

- SALES, N. G. et al. Space-time dynamics in monitoring neotropical fish communities using eDNA metabarcoding. **Science of the Total Environment**, v. 754, 142096, 2021.
- SILVA, C. B. et al. Fish larvae diversity in a conservation area of a neotropical floodplain: influence of temporal and spatial scales. **Hydrobiologia**, 2017.
- SILVA, L. E. et al. Amazonian ichthyoplankton assessment via DNA metabarcoding: A baseline for detecting spawning sites of migratory fishes. **Biological Conservation**, v. 284, 110180, 2023.
- TABERLET, P. et al. Towards next-generation biodiversity assessment using DNA *metabarcoding*. **Molecular Ecology**, v. 21, n. 8, p. 2045–2050, 2012.
- TEIXEIRA, D. F. et al. DNA metabarcoding assessment of Neotropical ichthyoplankton communities is marker-dependent. **Ecology and Evolution**, v. 13, e10649, 2023.
- WARD, R. D. et al. DNA barcoding Australia's fish species. **Philosophical Transactions of the Royal Society, B: Biological Sciences**, v. 360, n. 1462, p. 1847–1857, 2005.
- WARD, R. D. DNA barcode divergence among species and genera of birds and fishes. **Molecular Ecology Resources**, v. 9, n. 4, p. 1077–1085, 2009.
- YU, D. W. et al. Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. **Methods in Ecology and Evolution**, v. 3, n. 4, p. 613–623, 2012.