



INSTITUTO DE BIOCÊNCIAS

DEPARTAMENTO DE BIOLOGIA ESTRUTURAL E FUNCIONAL

Laboratório de Biologia e Genética de Peixes

Dissertação de Mestrado

**Diversidade genética de *Parancistrus aurantiacus* (Siluriformes,
Loricariidae) em corredeiras do rio Tocantins**

Fernanda Américo

Botucatu, SP

2022



INSTITUTO DE BIOCIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ESTRUTURAL E FUNCIONAL

Laboratório de Biologia e Genética de Peixes

Dissertação de Mestrado

Diversidade genética de *Parancistrus aurantiacus* (Siluriformes, Loricariidae) em corredeiras do rio Tocantins

Aluna: Fernanda Américo

Orientador: Prof. Dr. Claudio de Oliveira

Co-orientadora: Prof. Dra. Vanessa Paes da Cruz

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Biológicas (Zoologia) do Instituto de Biociências de Botucatu, Universidade Estadual Paulista "Júlio de Mesquita Filho", como parte dos requisitos para obtenção do título de Mestre.

Botucatu, SP

2022

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSEMEIRE APARECIDA VICENTE-CRB 8/5651

Americo, Fernanda.

Diversidade genética de *Parancistrus aurantiacus*
(Siluriformes, Loricariidae) em corredeiras do rio Tocantins
/ Fernanda Americo. - Botucatu, 2022

Dissertação (mestrado) - Universidade Estadual Paulista
"Júlio de Mesquita Filho", Instituto de Biociências de
Botucatu

Orientador: Claudio de Oliveira
Coorientador: Vanessa Paes da Cruz
Capes: 20400004

1. Cascudo (Peixe). 2. Polimorfismo de nucleotídeo único.
3. Dinâmica populacional. 4. Sequenciamento de nucleotídeos
em larga escala. 5. Genômica.

Palavras-chave: Cascudo-borracha; Genômica populacional;
SNPs; Sequenciamento de nova geração.

AGRADECIMENTOS

Agradeço aos meus familiares, especialmente meu avô José, por me incentivar nos estudos com todo suporte necessário, além de toda confiança. Aos meus tios Marcos e Mônica, minhas referências no caminho da educação. Aos meus pais Adriana e André e aos meus irmãos mais novos, Ramon e Mariana, por todo amor e carinho. À minha irmã e amiga Juliana, que vive muitos dos momentos mais felizes comigo. Amo todos vocês e sem apoio não teria concluído esse período incrível.

Aos meus melhores amigos Arlene, Erick, Érik, Cesar, Beatriz, Rafael, Leonardo, Tainá e Igor, obrigada por toda palavra de incentivo, risada, carinho e afeto. Amo muito vocês e cada momento que vivemos. Aos demais, que não conseguirei citar sem deixar tudo muito extenso, agradeço os momentos importantes que já compartilhamos.

Ao meu orientador Claudio, por me acolher desde a graduação. Agradeço todo conselho, conversa, ensinamento e por fazer parte da minha vida científica. Tenho muita admiração por sua pessoa e profissional. Guardarei todo aprendizado com muito carinho.

À minha co-orientadora Vanessa, por toda paciência, disponibilidade, acolhimento e ensinamento passado com a maior dedicação e amor a ciência. Guardarei com muito carinho toda palavra de incentivo e a capacidade de acreditar no potencial de seus alunos.

Aos companheiros de laboratório, mesmo que a pandemia tenha nos retirado algumas experiências presenciais, saibam que os admiro e seus trabalhos. Me senti acolhida nas disciplinas, bancadas, dúvidas e em tantas outras situações. Agradeço principalmente à May, Bia, Gi, Ailton, Sil, Najila, Lari, Mari, Bia D. e Pablo pelo suporte e ensinamentos. E ao Iann, que trouxe os tecidos do projeto pessoalmente e viveu comigo semanas de prática no laboratório. Obrigada a todos.

Agradeço ainda ao Programa de Pós-Graduação em Ciências Biológica (Zoologia), da UNESP de Botucatu e aos funcionários do Instituto de Biociências. E finalmente, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo financiamento recebido, sem o qual não seria possível a realização do meu mestrado.

RESUMO

A região Neotropical possui a mais rica e diversificada ictiofauna de água doce do mundo. Entre a grande diversidade de peixes dessa bacia a ordem Siluriformes se destaca. Nessa ordem a família Loricariidae e a subfamília Hypostominae possuem a maioria dos seus espécimes restritos à América do Sul tropical e subtropical. A espécie *Parancistrus aurantiacus* se encontra presente neste grupo e a sua distribuição ocorre nos rios Araguaia, baixo Tocantins, alto Xingu, Iriri e Ucayali. As informações sobre as populações na região Norte do Brasil são escassas, o que nos impede de avaliar o seu real estado de conservação. Este estudo teve como objetivo caracterizar geneticamente indivíduos da espécie que ocorrem no rio Tocantins, avaliar a distribuição e conexões existentes entre elas, através do uso dos marcadores genéticos de polimorfismos únicos (SNPs), utilizando sequenciamento de nova geração (NGS). Analisamos 36 amostras de *P. aurantiacus* distribuídas em cinco pontos na bacia do rio Tocantins. Após o rastreamento de marcadores, foram obtidos 4.214 *SNPs* para análise genética da espécie. As análises revelaram um valor de F_{ST} mínimo de 0,0058 e máximo de 0,0883 (valor $P < 0,050$), apontando uma única unidade panmítica entre os pontos amostrados. Para todas as localidades, as heterozigosidades foram altas e a heterozigosidade observada (H_o) apresentou valor superior à heterozigosidade esperada (H_e), associado ao coeficiente de endogamia (F_{IS}) negativo, que indica a ausência de endogamia nos espécimes analisados. Através das análises de DAPC e do *software* STRUCTURE identificamos uma só população. Portanto, os resultados obtidos evidenciam que existe estruturação genética nas regiões analisadas e uma única população. Essas informações são importantes para futuras ações de manejo das populações naturais.

ABSTRACT

The Neotropical region has the richest and most diverse freshwater ichthyofauna in the world. Among the great diversity of fish in this basin, the order Siluriformes stands out. In this order, the Loricariidae family and the Hypostominae subfamily have most of their specimens restricted to tropical and subtropical South America. The species *Parancistrus aurantiacus* is present in this group and its distribution occurs in the Araguaia, lower Tocantins, upper Xingu, Iriiri and Ucayali rivers. Information about populations in the North region of Brazil is scarce, which prevents us from assessing their real conservation status. This study aimed to genetically characterize individuals of the species that occur in the Tocantins River, evaluate the distribution and existing connections between them, through the use of genetic markers of unique polymorphisms (SNPs), using next generation sequencing (NGS). We analyzed 36 samples of *P. aurantiacus* distributed in five points in the Tocantins river basin. After marker screening, 4,214 SNPs were obtained for genetic analysis of the species. The analyzes revealed a minimum F_{ST} value of 0.0058 and a maximum of 0.0883 (P value <0.050), indicating a single panmitic unit among the sampled points. For all locations, heterozygosity was high and the observed heterozygosity (H_o) was higher than the expected heterozygosity (H_e), associated with a negative inbreeding coefficient (FIS), which indicates the absence of inbreeding in the analyzed specimens. Through DAPC analysis and STRUCTURE software we identified a single population. Therefore, the results obtained show that there is genetic structure in the analyzed regions and a single population. This information is important for future actions to manage natural populations.

Sumário

1. Introdução	8
1.1. <i>Bacia Tocantins-Araguaia</i>	9
1.2. <i>Marcadores moleculares</i>	10
1.3. <i>Parancistrus aurantiacus</i>	11
1.4. <i>Informações genéticas</i>	12
2. Objetivos	14
3. Materiais e métodos	14
3.1. <i>Obtenção do material genético e extração de DNA</i>	14
3.2. <i>Elaboração das bibliotecas de SNPs</i>	15
3.3. <i>Análises dos dados</i>	17
4. Resultados	19
4.1. <i>Diversidade genética e estrutura populacional</i>	19
4.2. <i>Análises de relacionamento</i>	22
5. Discussão	23
6. Referências	24

1. INTRODUÇÃO

A região Neotropical abriga mais de 5.100 espécies de peixes continentais (Reis *et al.*, 2016), e representa a mais rica e diversificada ictiofauna dulcícola do mundo (Albert *et al.*, 2020). Em território brasileiro, são encontradas mais de 3.100 espécies (Azevedo-Santos *et al.*, 2021a). Este alto número de espécies é um reflexo da complexa história geomorfológica associada com os processos ecológicos/evolutivos que moldam a diversidade de peixes até o presente (Buckup *et al.*, 2016; Reis *et al.*, 2016; Albert *et al.*, 2020). Ainda assim, o número de táxons pode estar subestimado (Levêque *et al.*, 2008). A baixa amostragem parece ser um dos principais motivos por trás desse inventário incompleto.

A ictiofauna neotropical tem uma longa e complexa história evolutiva. Diversos períodos geológicos apresentam evidências de altas taxas de diversificação, sendo que alguns clados se diversificaram no Cretáceo (65-145 m.a) enquanto outros no Paleogeno, por exemplo (23-65 m.a) (Lundberg *et al.*, 1998). Glaciações, movimentação das placas tectônicas e mudanças nos níveis oceânicos que ocorreram ao longo das eras geológicas tem grande impacto sobre as populações de peixes de água doce, mas ainda se sabe pouco sobre como essas alterações influenciaram a ictiofauna neotropical e o quanto esses eventos foram responsáveis por essa diversificação (Moyer *et al.*, 2005). Entre os estudos que investigam essa parcela da biodiversidade, a maioria trata apenas de aspectos morfológicos e osteológicos, deixando de lado aspectos genéticos e biogeográficos.

Os peixes exercem funções fundamentais para os ecossistemas, contribuindo na regulação, transporte e ciclagem de nutrientes e energia (Cooke *et al.*, 2016; Atkinson *et al.*, 2019; Lennox *et al.*, 2019). Além disso, representam uma rica fonte de proteína para a sociedade humana, servindo como base alimentar e econômica, especialmente para comunidades ribeirinhas (Khalili Tilami *et al.*, 2018; Tacon *et al.*, 2020). Entretanto, parte considerável desta diversidade se encontra sob algum grau de ameaça, frente à impactos antrópicos diversos, sendo o barramento de rios, como por exemplo o que ocorre na bacia do alto rio Paraná, causando um efeito deletério sobre a ictiofauna (Agostinho *et al.*, 2016; Agostinho *et al.*, 2018; Silveira *et al.*, 2018; Arantes *et al.*, 2019).

Um dos maiores impactos causados pelas usinas hidrelétricas ocorre principalmente no trecho a montante do dique de instalação do empreendimento, com transformação dos rios em grandes lagos, modificando as condições naturais do ecossistema e, quando em associação com despejos e cargas de resíduos decorrentes da ação humana, torna-os

eutrofizados, criando condições aquáticas inadequadas como a proliferação de algas, geração de gases e toxicidade na água (Oliveira *et al.*, 2020; Tomczyk *et al.*, 2020; Rudke *et al.*, 2021; Slaby *et al.*, 2022). Além disso, deve-se ressaltar que os reservatórios recebem grande carga de sedimentos que podem levar ao assoreamento do leito, diminuir o volume da água no reservatório e alterar os habitats de muitas espécies de peixes, especialmente as reofílicas – que necessitam de correntezas para viver (Birnie-Gauvin *et al.*, 2017; Ganassin *et al.*, 2021). Estas alterações do ambiente impedem a movimentação dos peixes, e conseqüentemente podem levar a diminuição dos estoques (Agostinho *et al.*, 2007; Azevedo-Santos *et al.*, 2021b; Bartoň *et al.*, 2022).

Nos últimos anos, com o desenvolvimento de tecnologias genômicas, as informações obtidas estão agora fornecendo *insights* sobre os processos evolutivos e a base molecular das populações em um número expressivo de organismos. Isso é possível através da utilização de marcadores genéticos relacionados à base genética das características envolvidas na adaptação dos organismos (Oomen *et al.*, 2022). Nesse contexto, a realização de levantamentos e a correta identificação das espécies com a utilização de marcadores genéticos, além de permitirem uma visão sobre possíveis mudanças genéticas ocorrendo nas populações, também possibilitam levantar elementos para a compreensão das interações relacionando diversidade e distribuição das espécies, características biogeográficas e ecossistemas.

1.1. *Bacia Tocantins-Araguaia*

A bacia do rio Tocantins-Araguaia é a maior bacia localizada completamente no território brasileiro e como o nome indica, é composta principalmente pelos rios Tocantins e Araguaia, abrangendo 918.273 km² (cerca de 11% do território brasileiro), nos Estados de Goiás, Tocantins, Pará, Maranhão, Mato Grosso e Distrito Federal. Formado pelos rios das Almas e Maranhão, constitui um dos principais rios do Cerrado do Brasil Central. De sua nascente, no Planalto Goiano, a cerca de 1.000 m de altitude, até sua foz no Oceano Atlântico, este rio percorre cerca de 2.400 km (MMA 2006; ANA 2020).

Tradicionalmente, o rio Tocantins é dividido em três trechos: o alto Tocantins, que se estende de suas nascentes até a cachoeira do Lajeado, 1.060 km; o médio Tocantins, entre a cachoeira do Lajeado e a cachoeira de São João do Araguaia, 980 km; e o baixo Tocantins, que vai de São João do Araguaia até sua foz, 360 km. O ambiente do alto curso

é frequentemente composto por corredeiras e cachoeiras, enquanto o baixo curso, especialmente a jusante de Tucuruí, apresenta baixa declividade, o que permite a formação de grandes remansos e planícies inundáveis (Paiva 1982; Agostinho *et al.*, 2009). O rio Araguaia, principal afluente da margem esquerda do rio Tocantins, tem 2.600 km de extensão e nasce no Planalto Central brasileiro. A maior ilha fluvial do mundo está localizada no sistema do rio Araguaia, a Ilha do Bananal (350 km de comprimento por 80 km de largura). O rio Araguaia corre paralelo ao rio Tocantins, até se encontrarem no município de Marabá, Estado do Pará (Goulding *et al.*, 2003; ANA 2020).

O Rio Tocantins originalmente possuía muitos ambientes de corredeiras e cachoeiras, sendo alvo do setor hidrelétrico, com vários projetos deste tipo implantados nas últimas décadas. Já a geografia da bacia do Rio Araguaia, sem grandes desníveis ou acidentes, tem sido preferida para hidrovias e agricultura (Chamon *et al.*, 2022).

Embora com o conhecimento da diversidade de peixes esteja longe de ser completo e compartilhado, o alto grau de endemismo da bacia do rio Tocantins-Araguaia é corroborado em vários estudos (Carvalho *et al.*, 2010; Bertaco *et al.*, 2011; Hales e Petry 2013; Dagosta e de Pinna 2017, 2019). O alto Tocantins detém o maior número absoluto de táxons endêmicos na bacia amazônica (Dagosta e de Pinna 2017, 2019) e o rio Araguaia, a principal artéria fluvial do Brasil central e o ecótono Amazônia-Cerrado, abriga mais espécies de peixes do que qualquer outra bacia do Cerrado (Latrubesse *et al.*, 2019).

1.2. Marcadores moleculares - SNPs

Nos últimos anos, o marcador genético que vem se mostrando mais resolutivo para o entendimento da estrutura das populações é o polimorfismo de nucleotídeo único - *SNPs* (*Single Nucleotide Polymorphism*). Tais polimorfismos são causados por mutações pontuais, que ocorrem em todo o genoma, e a assim a geração de dados de alta qualidade é possível, o que os torna uma ferramenta genética eficiente e viável para o entendimento dos processos de conservação das espécies (Morin *et al.*, 2009).

Uma das vantagens metodológicas mais importantes dos *SNPs* como marcadores para estudos de conservação talvez seja a possibilidade da aplicação em espécies não modelos, propiciando a obtenção de dados a partir de tecnologias avançadas como os sequenciadores de nova geração (*NGS*), além do oferecimento de informações integradas de loci neutros e adaptativos diretamente relacionados com respostas a processos de

mudanças ambientais de genética populacional e de ecologia em várias espécies de peixes (Funk *et al.*, 2012; Zhang *et al.*, 2015; Cruz *et al.*, 2016; Di Battista *et al.*, 2017).

O desenvolvimento de métodos eficazes para identificação de *SNPs* levou a uma revolução no seu uso como marcadores moleculares (Batley *et al.*, 2003; Emerson *et al.*, 2015). A técnica de *ddRADseq* (sequenciamento de DNA associado a dupla digestão) proposta por Peterson *et al.* (2012) foi adaptada ao método de sequenciamento de DNA em associação com uma enzima de restrição (*RADseq*) descrito por Baird *et al.* (2008). A técnica *ddRAD* se baseia no sequenciamento massivo de representações reduzidas do genoma obtidas pela digestão do DNA genômico com duas enzimas de restrição, tendo como finalidade a obtenção de sequências polimórficas como os marcadores *SNPs* e a genotipagem (Andrews *et al.*, 2016).

1.3. *Parancistrus aurantiacus* (Castelnau, 1855)

Dentre a alta diversidade de peixes de água doce existentes na região neotropical, a ordem Siluriformes (Cuvier, 1817) é uma das mais amplamente representadas. Nesta ordem temos a família Loricariidae e a subfamília Hypostominae e a espécie *Parancistrus aurantiacus* (Castelnau, 1855) se encontra presente neste grupo, conforme indicado na **Figura 1**. A maioria dos membros está restrita à América do Sul tropical e subtropical e a sua distribuição ocorre nos rios Araguaia, baixo Tocantins, alto Xingu, Iri e Ucayali (Dagosta e de Pinna, 2019).



Figura 1. Exemplar da espécie *Parancistrus aurantiacus*, presente nas corredeiras do rio Tocantins. Foto: Leandro Sousa, Projeto IctioXingu.

Os indivíduos de *P. aurantiacus*, são uma das espécies mais abundantes do rio Tocantins, popularmente conhecida como “acari-borracha ou cascudo-borracha”. São peixes migratórios e bentônicos, que se encontram em profundidade de aproximadamente 9,5 metros nos espaços entre as rochas, fendas e cavernas submersas. Frequentemente, é possível observar grupos de indivíduos ocupando o mesmo espaço nessas galerias. Espécie noturna, que se alimenta raspando algas da superfície das rochas (Sousa, 2022). Análises de conteúdo estomacal revelaram a presença principalmente de algas, outros conteúdos em menor proporção são: cianobactérias, pedaços de plantas, briozoários, larvas de insetos aquáticos, microcrustáceos, poucos moluscos e alguns grãos de areia (dos Anjos *et al.*, 2009).

Sua coloração é uniformemente marrom escura ou clara, ou manchada, nunca pontilhada (Rapp Py-Daniel *et al.*, 2005). Alguns indivíduos apresentam manchas amareladas por todo o corpo, podendo levar a coloração do animal a ficar completamente amarelo-dourado em alguns períodos do ano. Possui o abdome totalmente coberto de placas, sua nadadeira dorsal é conectada a nadadeira adiposa através de uma membrana, apresenta dentes bucais conspícuos, sua boca ventral e abertura branquial são maiores que a da maioria de outras espécies da região (Akama *et al.*, 2020). Podem atingir até 20 cm de comprimento, aproximadamente, e as principais informações osteológicas observadas incluem ausência de um forte cômulo no etmóide lateral para articulação com o metapterigóide, processos no ângulo-articular mais longos e opérculo exposto com odontódios completamente incluso na pele. Seus olhos são relativamente pequenos quando comparado a outros cascudos. A distância interbranquial é larga (24.9-39.5% no comprimento padrão) e a distância interorbital é estreita (38.5-43.1% comprimento padrão) (Rapp Py-Daniel *et al.*, 2005).

Atualmente a espécie *P. aurantiacus*, está categorizada como “Quase ameaçada” no Livro Vermelho da Fauna Brasileira Ameaçada de Extinção (ICMBio, 2019b). A maioria das espécies sofre ameaça por alteração de habitat e são muito procuradas no comércio aquarista internacional. Essa espécie não está presente na Lista Vermelha da União Internacional para a Conservação da Natureza e dos Recursos Naturais (IUCN, 2022).

1.4. Informações genéticas

Dois principais estudos foram utilizados, com base na relação direta com a espécie estudada de indivíduos, sendo um deles menos recente e outro que se encontra em análise. Existem diversas publicações muito ricas em informações sobre a genética de cascudos de água doce e de igual importância para a compreensão das populações, no entanto, para os indivíduos da região do rio Tocantins analisadas, os referidos estudos são interessantes devido a seus resultados significativos sobre a conectividade genética e sua relação com gênero e subfamília da espécie.

Com o objetivo de identificar a história biogeográfica na diversidade de ictiofauna tropical, Montoya-Burgos *et al.* (2003), avaliaram sequências da região de controle do DNA mitocondrial de *Hypostomus*. Utilizando a região D-loop e transcrição interna completa de sequências de região espaçadora (ITS), encontraram as principais inter-relações em um subconjunto de 21 espécies. Nas análises filogenéticas o gênero pode ser dividido em quatro clados monofiléticos, e na análise biogeográfica de cada clado identificaram sete eventos cladogenéticos principais. As estimativas de data foram atribuídas a cada um dos sete eventos, permitindo uma ligação entre quatro deles com alterações hidrogeológicas documentadas. Comparações com padrões de distribuição de peixes não relacionados indicaram que várias estruturas hidrogeológica-cladogenéticas reconstruídas e datadas podem ter atuado na diversificação dos peixes neotropicais de água doce durante o final do Terciário.

Em estudos envolvendo genética, Pereira e Akama (2022), descrevem os resultados, de uma tese em desenvolvimento, com o objetivo de identificar a diversidade da subfamília Hypostominae em corredeiras do rio Tocantins para além de sua morfologia. Através da técnica de *DNA barcode*, que utiliza um segmento padronizado do gene da subunidade I do citocromo c oxidase (COI), constatou-se divergências significativas, sendo possível estabelecer que indivíduos aparentados morfologicamente podem estar isolados na sua forma reprodutiva. Foram identificadas seis espécies, morfologicamente, incluindo entre elas *P. aurantiacus*, que se agrupou com indivíduos de *Scobinancistrus pariolispos*, e *Lamontichthys sp.*, com alto apoio estatístico segundo a filogenia gerada por Inferência Bayesiana. Já os espécimes restantes formaram três grupos distintos, onde também foram feitos testes de distância entre os grupos, obtendo uma diferença média de 3,34% entre eles.

2. OBJETIVOS

Considerando a falta de informações acerca da diversidade genética de *Parancistrus aurantiacus*, o presente estudo tem como objetivo principal:

- a caracterização genética dos indivíduos que ocorrem no rio Tocantins;
- avaliar a distribuição e as conexões existentes entre eles;

3. MATERIAIS E MÉTODOS

3.1. Obtenção do material genético e extração de DNA

Para realização do presente estudo, foram utilizadas 36 amostras do espécime de *Parancistrus aurantiacus*, distribuídas em cinco pontos de coleta, provenientes do rio Tocantins, no estado Pará, conforme indicado na **Figura 2**.

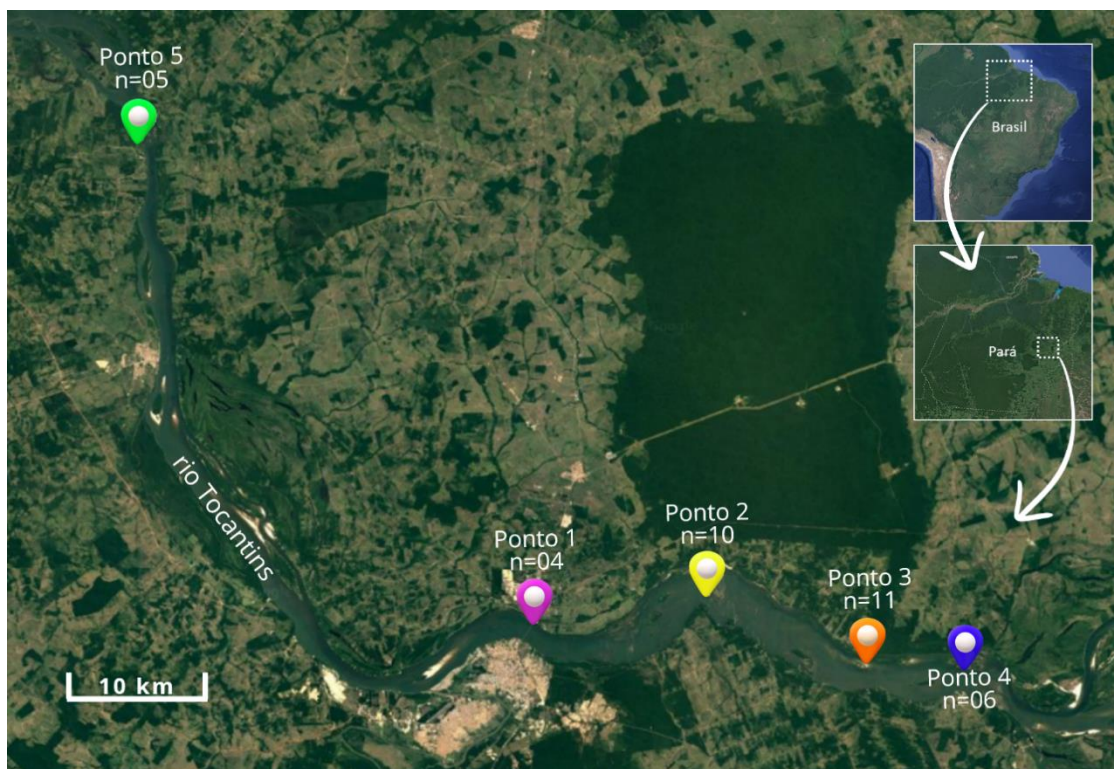


Figura 2. Mapa ilustrando o território brasileiro, evidenciando o estado do Pará, em destaque ilustrando a região do alto rio Tocantins e os pontos amostrados de *Parancistrus aurantiacus*. Ponto 1 = Ponte Mista de Marabá (n=04), Ponto 2: Ilha das Cabras (n=10), Ponto 3: Ilha do Jaú (n=11), Ponto 4: região do Caju (n=06), Ponto 5: Pedral do Lorenção (n=05).

As amostras foram coletadas e obtidas através da parceria com o Dr. Alberto Akama, do Museu Paraense Emílio Goeldi (MPEG). Os fragmentos dos tecidos de cada

amostra foram recebidos preservados em etanol 95%. Estes tecidos foram depositados na coleção do Laboratório de Biologia e Genética de Peixes (LBGP), credenciada no Ministério do Meio Ambiente como Fiel Depositária de Amostras do Patrimônio Genético. Para extração de DNA das amostras utilizamos o *Wizard® Genomic DNA Purification Kit* (Promega, EUA), seguindo o protocolo do fabricante e para determinar a quantidade de DNA genômico de cada amostra, utilizamos o *Qubit® TM dsDNA BR Assay Kit*.

3.2. Elaboração das bibliotecas de *SNPs*

ddRAD (double digest Restriction Site Association)

A elaboração e desenvolvimento das bibliotecas de *SNPs* seguiu o protocolo *ddRad*, descrito por Peterson *et al.* (2012), com adaptação em Campos *et al.* (2017). A digestão enzimática foi realizada de modo simultâneo, com duas enzimas de restrição, a *EcoRI* (adaptador P1), com reconhecimento de corte raro e a *MspI* (adaptador P2), com reconhecimento de corte mais frequente. Essas enzimas criam pontas coesivas em ambas as extremidades das moléculas de DNA digeridas permitindo assim a sua ligação aos adaptadores. Para a digestão utilizamos 4µl *TANGO buffer (Cut Smart)*, em um volume final de 40µl, que permaneceu incubado por três horas a uma temperatura de 37°C. A purificação do material digerido foi feita com uso do *beads Agencourt AMPure XP Kit (Beckman Coulter, USA)*, seguindo o protocolo do fabricante. Após a purificação, as amostras foram quantificadas para obter a concentração de DNA presente, novamente através do *Qubit 4.0 Fluorometer (Invitrogen)*.

Ligação dos adaptadores

A reação de ligação com os adaptadores foi realizada em um volume final de 10 µl com 0,2 µl da enzima T4 ligase, 1 µl do tampão 10X T4 *DNA ligase Reaction Buffer*, 1,0 µl do adaptador P1 4,04 µM (referente a *EcoRI*), e 1,0 µl do adaptador P2 0,07 µM (referente a *MspI*) e 8,27 µl do material digerido. Este processo de ligação ocorre com as amostras incubadas a de 23°C por 30 minutos, 65°C por 10 minutos, 63°C por 90 segundos, no qual, a temperatura diminui 2°C a cada 90 segundos até alcançar 23°C. Após a reação de ligação dos adaptadores, as amostras foram purificadas novamente com o uso do *beads Agencourt AMPure XP Kit (Beckman Coulter, USA)*, seguindo o protocolo do fabricante.

Indexação

Para indexação, sintetizamos primers com os index da Illumina, I5 (502 ao 511) e I7 (701 ao 707). Estes *primers* foram inseridos em cada amostra e combinados de diferentes formas, nos insertos ligados aos adaptadores. A reação de indexação continha: 15µl do produto de ligação, 25µl de Master Mix (GoTaq), 5µl de cada index I5 e I7, para cada amostra apresentando uma combinação única de index. A reação de PCR para indexação foi realizada de acordo com as seguintes condições de ciclo: 1 ciclo com um passo inicial de 72°C por 3 minutos seguido de 16 ciclos de desnaturação (95°C por 30 segundos e 95°C por 10 segundos), anelamento (55°C por 30 segundos), extensão (72°C por 30 segundos) e extensão final a 72°C por 5 minutos. Após a reação de PCR, os fragmentos foram purificados utilizando *beads* Agencourt AMPure XP (Beckman Coulter, USA), seguindo o protocolo do fabricante. Após a purificação da indexação, o material foi aplicado em um gel de agarose 1% a fim de verificar se as amostras haviam indexado de acordo com o tamanho esperado. Posteriormente, as amostras foram quantificadas em um aparelho *Qubit 4.0 Fluorometer* (Invitrogen), sendo normalizadas a 6ng/µl.

Pool das amostras

Após os ensaios de *PCR* e a quantificação das amostras, realizamos um *pool* composto com a concentração de 6ng/µl de DNA baseado na quantificação anterior. O *pool* foi composto de 5µl de cada amostra obtendo um volume final de 200µl. A purificação foi feita com o uso de *AMPure XP beads* (Beckman Coulter, DE), seguindo o protocolo do fabricante.

Size selection e sequenciamento

No *size selection*, que seleciona o tamanho dos fragmentos de interesse, no qual aplicamos o *pool* em gel de agarose a 1%. Ao final da eletroforese, cortamos um fragmento do gel com um tamanho entre 300 e 500pb, em seguida, foi realizada a purificação utilizando o *kit Wizard® SV Gel and PCR Clean-Up System* (Promega, EUA), seguindo o protocolo do fabricante. Após a purificação do gel, realizamos a quantificação do *pool* em um aparelho *Qubit 4.0 Fluorometer* (Invitrogen) obtendo a concentração final de 1,84ng/µl. Para verificar a precisão da biblioteca, realizamos a *PCR em tempo real* (*qPCR*) e, posteriormente, a sua normalização. A biblioteca foi sequenciada com leituras *single-end* de 150pb, em uma *flow cell High Output* e então sequenciados na plataforma *NGS Illumina NEXTSEQ500*, no Laboratório de Pesquisa em Análise Genéticas no

Instituto de Biotecnologia (IBTEC– UNESP, Botucatu/SP), com auxílio colaborativo do Prof. Dr. Paulo Eduardo M. Ribolla.

3.3 Análise dos dados

Deteção e filtragem de SNPs

Após a etapa de sequenciamento as amostras foram disponibilizadas no site da plataforma *NGS Illumina* (Basespace) e através do seu acesso foi possível baixar e verificar as *reads* brutas (*raw data*) obtidas, para então dar início às análises.

Os dados gerados foram concatenados e, para verificação da qualidade, utilizamos dois *softwares*. O *FastQC*, que fornece o controle de qualidade das *reads* e a impressão de cada sequência (Andrews, 2010) (disponível em bioinformatics.babraham.ac.uk/projects/fastqc/) e o *MultiQC*, que resume os dados do *FastQC*, onde todas as sequências das amostras ficam organizadas em apenas um relatório (Ewels *et al.*, 2016) (disponível em github.com/ewels/MultiQC). Os resultados gerados permitem visualizar, em forma de gráficos, se existe a presença de sequências dos adaptadores que foram utilizados na construção da biblioteca, *reads* que apresentam um *Phred Quality Score* de qualidade (10-60) e os tamanhos das *reads* em pares de base (pb) de cada amostra. As filtrações para a remoção dos adaptadores e as *reads* com qualidade abaixo de *Phred Quality* <30 (Q30) foram realizadas utilizando o *software TRIMMOMATIC v.o.32* (Bolger *et al.*, 2014) e para remoção de *reads* que ainda apresentavam regiões de sítios enzimáticos, a digestão *in silico* (Driller *et al.* 2020). As *reads* foram padronizadas com 146 pb, novamente através do *software TRIMMOMATIC v.o.32* (Bolger *et al.*, 2014).

Ainda não existe um genoma de referência da espécie, portanto na montagem do catálogo de marcadores e o rastreamento de *SNPs* utilizamos a abordagem *de novo* e o *refmap*, a partir da *pipeline* STACKS (Catchen *et al.* 2011). A montagem do catálogo em *P. aurantiacus* foi desenvolvida com indivíduos bem representados das amostras, com maior número de *reads*. Selecionamos cinco amostras, sendo uma do Ponto 1: Ponte Mista de Marabá (voucher: 61P1), uma do Ponto 4: região do Caju (voucher: 394P4) e três do Ponto 5: Pedral do Lorenção (voucher: 399P5, 400P5 e 462P5). No rastreamento de *SNPs*, utilizamos o parâmetro M=2 (número de incompatibilidades permitidas entre os dois alelos de um determinado loci) e o parâmetro m=3 (número mínimo de leituras idênticas necessárias para indicar um possível alelo).

Após desenvolver o catálogo de marcadores, construímos a referência através do *software Bowtie2* (v2.2.4) e em seguida utilizamos a *pipeline Refmap* do STACKS. Ao final, utilizamos o programa *Populations*, implementado no STACKS que selecionou os filtros de *SNPs* que ocorrem em pelo menos 90% dos indivíduos ($r = 0,9$) e um filtro excluiu os que apresentaram valor de MAF (*Minor allele frequency*) $<0,05$.

Análise de diversidade genética populacional

Para análise da diversidade genética, utilizamos o *software* GENEPOP v4.1.0 (Rousset, 2008), que calcula os parâmetros do coeficiente de endogamia (F_{IS}) e a probabilidade de desvio do equilíbrio de *Hardy-Weinberg* (*p-value*). Os alelos privados (A_p), a Heterozigosidade observada (H_o) e Heterozigosidade esperada (H_e) foram calculados utilizando o *software* ARLEQUIN v3.5.2.2 (Excoffier *et al.*, 2010).

Para a análise da diferenciação genética entre as populações, foram calculados os valores de F_{ST} *pairwise*, utilizando também o *software* ARLEQUIN v3.5.2.2 (Excoffier *et al.*, 2010).

As análises de estruturação genética da espécie, calculada pela análise bayesiana, através do *software* STRUCTURE (Pritchard *et al.*, 2000), onde o $k = 1$ a 7, com 10 interações para cada k , e parâmetros *burn-in* de 10% em 1.000.000 de MCMC (Markov Chain Monte Carlo). A identificação do número de *clusters* genéticos (k), através do método Evanno (Δk) (Evanno *et al.*, 2005) e de Puechmaille (Puechmaille, 2016), com o uso do STRUCTURE Selector (Li & Liu, 2018) e seus gráficos foram gerados para visualização no CLUMPAK (Kopelman *et al.*, 2015).

A análise discriminante de componentes principais (DAPC) (Jombart *et al.*, 2010), realizou a estimativa dos *clusters* genéticos, com o auxílio do pacote *adegenet* do RStudio (Jombart, Ahmed, 2011) e o número apropriado de *clusters* genéticos foi determinado através do critério de informação bayesiana (BIC). Os números retidos das PCs estão ilustrados nas **Figura S1 e S2**.

A análise de relacionamento das populações entre os 36 indivíduos calculou a probabilidade de dois alelos serem idênticos por descendência (IBD), utilizando o coeficiente de parentesco ϕ (Manichaikul *et al.* 2010) com a função “*Relatedness2*” do *vcftools* (Danecek *et al.* 2011) e a partir disso, criamos o dendrograma, seguindo o script *relatedness R* (disponível em github.com/davemcg/R_play/blob/master/relatedness.R) em R.

4. RESULTADOS

4.1 Diversidade genética e estrutura populacional

A diversidade genética foi obtida através de 4.214 *SNPs* que foram utilizados em todos os cálculos, entre os quais foram detectadas 2.366 transições e 1.848 transversões (**Figura 3**). Esses marcadores foram submetidos à análise estatística em 36 amostras de *Parancistrus aurantiacus*, distribuídas nas localidades do Rio Tocantins, na região de Marabá, Pará.

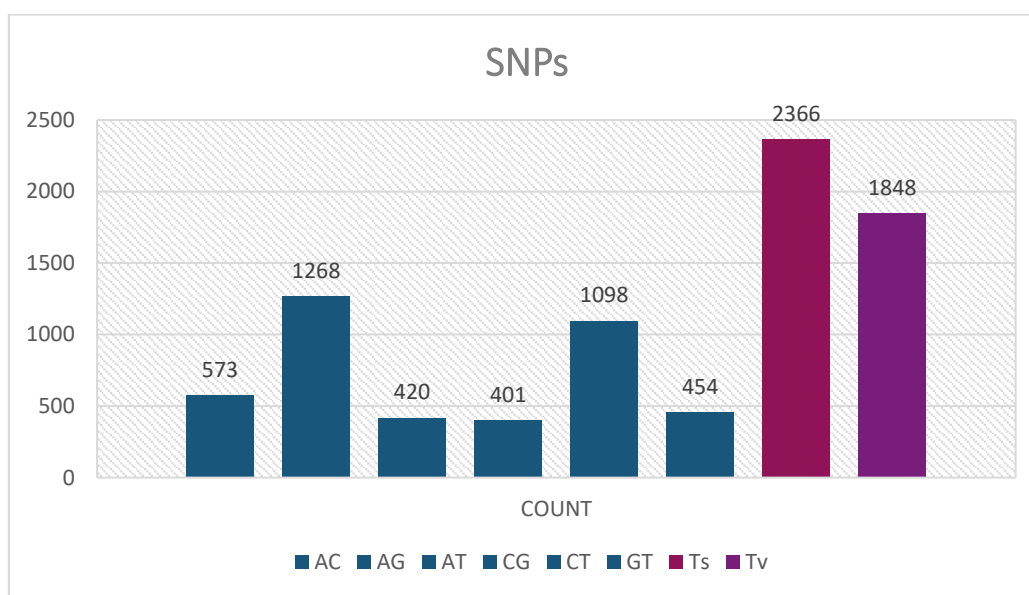


Figura 3. Identificação da distribuição 4.214 *SNPs* em azul, e a quantidade de *SNPs* com transições (vinho) e transversões (roxo) obtidas nas amostras de *P. aurantiacus*.

Para obtenção dos dados populacionais, calculamos o número de alelos privados (A_p) em cada localidade, com resultados que revelaram 11 alelos privados para amostras do P1; 15 para o P2; 155 para o P3; 391 para o P4 e 277 para o P5. Para todas as localidades, a heterozigiosidade observada (H_o) apresentou valor superior à heterozigiosidade esperada (H_e), sendo esses valores máximos em H_o (0,614) e H_e (0,440) para o P1 e o mínimo em H_o (0,338) e H_e (0,246) para o P3, o que indica excesso de heterozigotos. O índice de fixação ou coeficiente de endogamia (F_{IS}) foi negativo para os locais amostrados, indicando a ausência de endogamia nestes grupos. A região P1 teve um valor de (-0,501); P2 (-0,549); P3 (-0,378); P4 (-0,260) e P5 (-0,207). A diversidade de nucleotídeos (π) foi de 0,114 (+/-0,05) no P3; 0,155 (+/-0,08) no P5. Os dados estão ilustrados conforme a descrição na **Tabela 2**.

Tabela 2. Estatísticas dos valores de diversidade genética encontrados para *Parancistrus aurantiacus* utilizando 4.214 SNPs. N: número de indivíduos por localidade; Ap: número de alelos privados; Ho: heterozigosidade observada; He: heterozigosidade esperada; Fis: coeficiente de endogamia; (π): diversidade de nucleotídeos.

Localidades	N	Ap	Ho	He	Fis	(π)
Ponto 1	04	11	0.614	0.440	-0.501	0.128 (+/-0.07)
Ponto 2	10	15	0.530	0.327	-0.549	0.125 (+/-0.06)
Ponto 3	11	155	0.338	0.246	-0.378	0.114 (+/-0.05)
Ponto 4	06	391	0.403	0.329	-0.260	0.141 (+/-0.07)
Ponto 5	05	277	0.443	0.372	-0.207	0.155 (+/-0.08)

A análise *pairwise* com base nos valores de diferenciação genética (F_{ST}) entre os cinco locais revelou um valor de mínimo de 0,0058 e máximo de 0,0883 (**Tabela 3**, valor $P < 0,050$).

Tabela 3. Estatística da diferenciação genética *pairwise* encontrada para *Parancistrus aurantiacus* utilizando 4.214 SNPs.

Localidades	1	2	3	4	5
Ponto 1	0.0000				
Ponto 2	0.0778	0.0000			
Ponto 3	0.0647	0.0366	0.0000		
Ponto 4	0.0883	0.0484	0.0492	0.0000	
Ponto 5	0.0269	0.0221	0.0058	0.0065	0.0000

A análise bayesiana realizada no STRUCTURE (Figura 4A) com base em dados de cluster, demonstrou que a maior probabilidade de clusters foi $k = 1$, de acordo com a estimativa do método de Evanno ($\Delta k = 20608.48993$) (**Figura S3**), evidenciando panmixia entre os espécimes de diferentes regiões amostradas. O número de clusters genéticos via DAPC, identificou cinco clusters em *P. aurantiacus* distribuídos nas

localidades do Rio Tocantins, Pará (*Figura 4B*) e a probabilidade de atribuição de membros (*Figura 4C*).

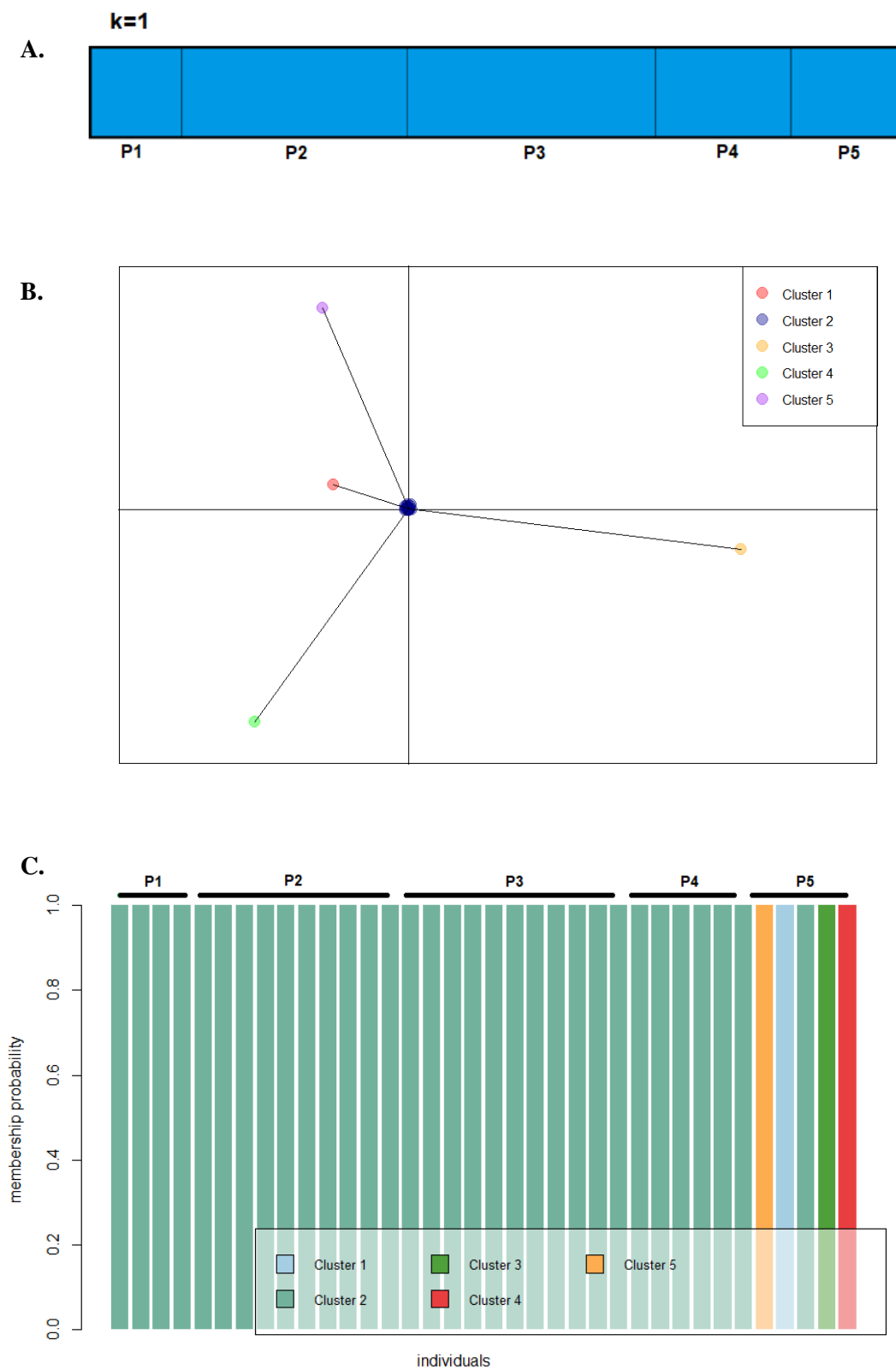


Figura 4. Análise de *clusters* realizada com 4.214 *SNPs* em *P. aurantiacus*. (A) Resultado do STRUCTURE com valores de $k = 1$ considerando as localidades do Rio Tocantins, na região de Marabá, Pará (B) O gráfico DAPC demonstrando os cinco clusters, indicados de acordo com as cores e símbolos da legenda, realizado adotando um número ótimo de componentes principais (PC = 70) calculado com o *software Adegnet*. (C) Representação gráfica das probabilidades de atribuição de associação (eixo x) para 36 indivíduos (eixo y) conforme calculado por meio de análise DAPC. Ponto 1 = Ponte Mista de Marabá (n=04), Ponto 2: Ilha das Cabras (n=10), Ponto 3: Ilha do Jaú (n=11), Ponto 4: região do Caju (n=06), Ponto 5: Pedral do Lorenção (n=05).

4.2. Análise de relacionamento

O dendrograma de relacionamento (**Figura 5**), baseado na estatística Φ , quando plotado, identificou a correlação de genótipo nas 36 amostra, usando 4.214 *SNPs* com 100% dos indivíduos genotipados, um total de cinco grupos, sendo eles representados por 04 indivíduos na região P1; 10 indivíduos na região P2; 11 indivíduos na região P3; 05 indivíduos na região P4 e 06 indivíduos na região P5.

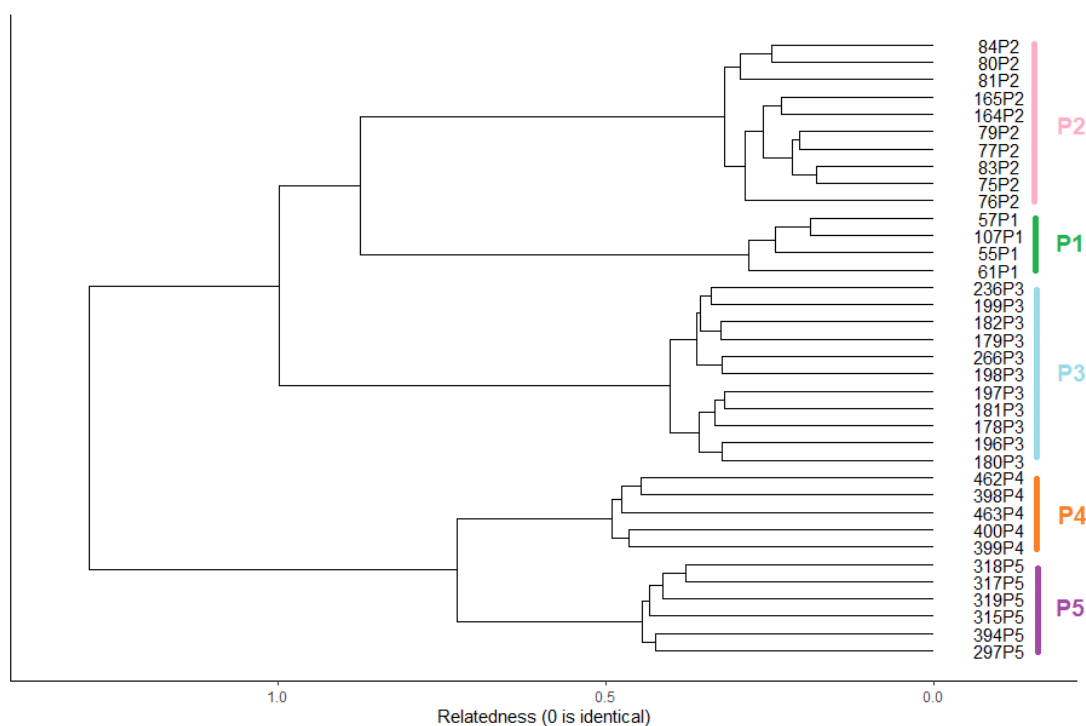


Figura 5. Dendrograma baseado na união hierárquica do coeficiente de relacionamento. As cores representam cada grupo (1 a 5).

5. DISCUSSÃO

No presente estudo, as amostras de exemplares provenientes de cinco pontos, na bacia do Tocantins próximo à região da usina hidrelétrica de Tucuruí, em Marabá, estado do Pará, foram submetidas ao uso dos marcadores *SNPs*, sendo possível identificar as características a uma população. Os resultados obtidos nos permitem observar que existe diversidade e que não há estruturação genética entre as amostras analisadas. É possível que as regiões amostradas sejam consideradas um único estoque genético, e os planos de gestão e manejo devem ser realizados de forma conjunta. Além disso, indicam que é necessário dar continuidade ao trabalho, reforçando as políticas de conservação em relação ao período de pesca – sendo que a reprodução desta espécie ocorre nos períodos de cheia e chuva - neste grupo. O monitoramento da espécie pode ajudar a redefinir as políticas públicas e socioambientais para proteção dos estoques explorados comercialmente e assim promover a atualização das listas de espécies ameaçadas. É necessário ainda observar mais adequadamente os impactos antrópicos diversos - como as barragens e retirada das rochas para hidrovias – que podem contribuir para acelerar as ameaças. Apesar da espécie estar em panmixia, e conseqüentemente ser migratória, o represamento dos rios de corredeira adquire uma nova característica no seu fluxo e vazão, modificando este ambiente essencial para a viabilidade, dinâmica e interações das espécies presentes nestas regiões.

O conhecimento das características genéticas nos auxilia na identificação das espécies e sua estrutura populacional e a utilização dos marcadores genéticos possibilita delimitar as barreiras geográficas, seu fluxo gênico (Zolet, 2017) e nos ajudam a determinar quais devem ser os cuidados e as ações a serem realizados. Segundo Fitzgerald, *et al.*, (2018), em estudo onde avaliou a ameaça a biodiversidade aquática devido os efeitos das barragens no rio Xingu, encontrou diferenças na estrutura populacional de três regiões próximas a Usina de Belo Monte. A diversidade das espécies mostrou uma pequena variação entre os segmentos, mas houve diferença significativa na estrutura das assembleias quando considerados os segmentos e as estações. A maioria das espécies consideradas altamente dependentes do habitat de corredeiras se restringiram ou ficaram bem mais abundantes nas seções acima e intermediária do rio. A complexidade dos habitats é positiva para a diversidade de peixes de água doce, porém o aumento na sedimentação, causado por uma redução na velocidade da correnteza, preenche os espaços dos substratos rochosos reduzindo a divisão em nichos, abrigo contra predadores e a disponibilidade de locais adequados para reprodução, podendo levar a uma provável diminuição da diversidade local.

Existe um trabalho de acompanhamento das comunidades de peixes (Barbosa et al., 2015), no entanto, sempre que possível, é preciso implantar estratégias de monitoramento não-invasivo (Schmid et al., 2017). Os programas de reprodução em cativeiro para várias espécies, como por exemplo, o cascudo zebra (em estado crítico de ameaça), também são importantes, no entanto ainda falta incluir a preservação de habitats críticos para as espécies remanescentes (Fitzgerald, *et al.*, 2018).

6. REFERÊNCIAS

Agostinho, A. A., Gomes, L. C., Pelicice, F. M. (2007). Estocagem. Ecologia e manejo de recursos pesqueiros em reservatórios do Brasil, p. 253–274.

Agostinho, A. A., Thomaz, S. M., Minte-Vera, C. V., & Winemiller, K. O. (2018). Biodiversity in the high Paraná River floodplain.

Akama, A. Sousa, L. (2020). IctioXingu: Expedição Tocantins - Peixes das profundezas do rio Tocantins (disponível em: https://www.youtube.com/watch?v=kjxoLB-2N7c&ab_channel=MuseuParaenseEm%C3%ADlioGoeldi). Acesso em dez2022.

Albert, J. S., Tagliacollo, V. A., & Dagosta, F. (2020). Diversification of Neotropical freshwater fishes. *Annual Review of Ecology, Evolution, and Systematics*, 51, 27-53.

Andrews S (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>

Andrews KR, Good JM, Miller MR, Luikart G & Hohenlohe PA (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, 17(2), 81.

dos ANJOS, H. D. B., de Souza AMORIM, R. M., SIQUEIRA, J. A., & dos ANJOS, C. R. (2009). Exportação de peixes ornamentais do estado do Amazonas, Bacia Amazônica, Brasil. *Boletim do instituto de Pesca*, 35(2), 259-274.

Arantes, C. C., Fitzgerald, D. B., Hoeinghaus, D. J., & Winemiller, K. O. (2019). Impacts of hydroelectric dams on fishes and fisheries in tropical rivers through the lens of functional traits. *Current Opinion in Environmental Sustainability*, 37, 28-40.

- Atkinson, C. L., Alexiades, A. V., MacNeill, K. L., Encalada, A. C., Thomas, S. A., & Flecker, A. S. (2019). Nutrient recycling by insect and fish communities in high-elevation tropical streams. *Hydrobiologia*, 838(1), 13-28.
- Awise JC, Haig SM, Ryder OA, Lynch M & Geyer CJ (1995). Descriptive genetic studies: applications in population management and conservation biology. In: Ballou, J.D. & Fose, T.J. (Ed.) Population management for survival and recovery. Columbia University Press, New York, p.183-244.
- Azevedo-Santos, V. M., Daga, V. S., Pelicice, F. M., & Henry, R. (2021b). Drifting in a free-flowing river: Distribution of fish eggs and larvae in a small tributary of a Neotropical reservoir. *Biota Neotropica*, 21.
- Azevedo-Santos, V. M., Rodrigues-Filho, J. L., Fearnside, P. M., Lovejoy, T. E., & Brito, M. F. (2021a). Conservation of Brazilian freshwater biodiversity: Thinking about the next 10 years and beyond. *Biodiversity and Conservation*, 30(1), 235-241.
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA & Johnson EA (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one*, 3(10), e3376.
- Bartoň, D., Brabec, M., Sajdlová, Z., Souza, A. T., Duras, J., Kortan, D., ... & Šmejkal, M. (2022). Hydropeaking causes spatial shifts in a reproducing rheophilic fish. *Science of The Total Environment*, 806, 150649.
- Baxter SW, Davey JW, Johnston JS, Shelton AM, Heckel DG, Jiggins CD & Blaxter ML (2011). Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PloS one*, 6(4), e19315.
- Berra, MT (2001) *Freshwater fish distribution*. Academic Press.
- Birnie-Gauvin, K., Aarestrup, K., Riis, T. M., Jepsen, N., & Koed, A. (2017). Shining a light on the loss of rheophilic fish habitat in lowland rivers as a forgotten consequence of barriers, and its implications for management. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 27(6), 1345-1349.
- Bolger AM, Lohse M, Usael B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.

Bruno MC, Casciotta JR, Almirón AE, Ricillio FL, Lizarrade MS (2015) Quaternary refugia and secondary contact in the southern boundary of the Brazilian subregion: comparative phylogeography of freshwater fish. *Vertebr Zool* 65:45-55.

Buckup PA, Menezes NA e Ghazzi MS (2007) Catálogo das Espécies de Peixes de Água Doce do Brasil. Museu Nacional, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 290pp.

Buckup, P. A., Sant, M., Ghazzi, A. Chapter · January 2007. n. January 2007, 2016.

Catchen JM, Amores A, Hohenlohe P, Cresko W & Postlethwait JH (2011). Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, 1(3), 171-182.

Chamon, C. C., Serra, J. P., Camelier, P., Zanata, A. M., Fichberg, I., & Marinho, M. M. F. (2022). Building knowledge to save species: 20 years of ichthyological studies in the Tocantins-Araguaia River basin. *Biota Neotropica*, 22.

Cooke, S. J., Martins, E. G., Struthers, D. P., Gutowsky, L. F., Power, M., Doka, S. E., & Krueger, C. C. (2016). A moving target - incorporating knowledge of the spatial ecology of fish into the assessment and management of freshwater fish populations. *Environmental Monitoring and Assessment*, 188(4), 1-18.

Cruz VP, Vera M, Pardo BG, Taggart J, Martinez P, Oliveira C & Foresti F (2016). Identification and validation of single nucleotide polymorphisms as tools to detect hybridization and population structure in freshwater stingrays. *Molecular Ecology Resources*. Doi 10.1111/1755-0998.12564.

DiBattista JD, Saenz-Agudelo P, Piatek MJ, Wang X, Aranda M & Berumen ML (2017). Using a butterflyfish genome as a general tool for RAD Seq studies in specialized reef fish. *Molecular Ecology Resources*. Doi 10.1111/1755-0998.12662

Driller M, Vilaca ST, Arantes LS, Carrasco-Valenzuela T, Heeger F, Chevallier D, ... & Mazzoni CJ (2020). Optimization of ddRAD-like data leads to high quality sets of reduced representation single copy orthologs (R2SCOs) in a sea turtle multi-species analysis. *bioRxiv*. Available online <https://doi.org/10.1101/2020.04.03.024331>.

Drummond AJ (2009). *Genetics* 4.8.5 (www.geneious.com).

- Drummond AJ & Rambaut A (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:1. 214. doi: 10.1186/1471-2148-7-214.
- Elmer KR & Meyer A (2011). Adaptation in the age of ecological genomics: Insights from parallelism and convergence. *Trends in Ecology & Evolution* 26(6), 298-306.
- Etter PD, Bassham S, Hohenlohe PA, Johnson EA & Cresko WA (2011). SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Molecular methods for evolutionary genetics*, 157-178.
- Ewels P, Magnusson M, Lundin S, Käller M (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report, *Bioinformatics*, 32(19), 3047–3048. Available online <https://doi.org/10.1093/bioinformatics/btw354>.
- Excoffier L, Laval G, & Schneider S (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, 1, 47-50.
- Fitzgerald, D. B., Perez, M. H. S., Sousa, L. M., Gonçalves, A. P., Py-Daniel, L. R., Lujan, N. K., ... & Lundberg, J. G. (2018). Diversidade e estrutura das comunidades de peixes que habitam as corredeiras do Rio Xingu: implicações para a conservação diante do desenvolvimento de uma hidrelétrica de grande escala.
- Foll M & Gaggiotti O (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180(2), 977–993.
- Fu YX (1996). New statistical tests of neutrality for DNA samples from a population. *Genetics*. 43,557–570.
- Funk WC, McKay JK, Hohenlohe PA & Allendorf FW (2012). Harnessing genomics for delineating conservation units. *Trends in ecology & evolution*, 27 (9), 489-96.
- Ganassin, M. J. M., Muñoz-Mas, R., de Oliveira, F. J. M., Muniz, C. M., Dos Santos, N. C. L., García-Berthou, E., & Gomes, L. C. (2021). Effects of reservoir cascades on diversity, distribution, and abundance of fish assemblages in three Neotropical basins. *Science of the Total Environment*, 778, 146246.

- Houston RD, Davey JW, Bishop SC, Lowe NR, Mota-Velasco JC, Hamilton A., ... & Gharbi K (2012). Characterization of QTL-linked and genome-wide restriction site-associated DNA (RAD) markers in farmed Atlantic salmon. *BMC genomics*, 13(1), 244.
- Khalili Tilami, S., & Sampels, S. (2018). Nutritional value of fish: lipids, proteins, vitamins, and minerals. *Reviews in Fisheries Science & Aquaculture*, 26(2), 243-253.
- Kimura M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution*, 16(2), 111-120
- Lennox, R. J., Crook, D. A., Moyle, P. B., Struthers, D. P., & Cooke, S. J. (2019). Toward a better understanding of freshwater fish responses to an increasingly drought-stricken world. *Reviews in fish biology and fisheries*, 29(1), 71-92.
- Lewinsonh TM & Prado PI (2005). How many species are there in Brazil? *Conservation Biology* 19, 619-624.
- Miller OS (1995). Selective breeding programs for rare alleles: examples from the Przewalski's horse and California condor pedigrees. *Conserv. Biology*. 9: 1262-1273.
- Montoya-Burgos JI (2003) Historical biogeography of the catfish genus *Hypostomus* (Siluriformes: Loricariidae), with implications on the diversification of Neotropical ichthyofauna. *Mol Ecol* 12:1855-1867.
- Morin PA, Matien KK & Taylor BL (2009). Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources*, 9 (1) 66-73.
- Myers N, Mittermeier RA, Mittermeier CG, Fonseca GAB & Kent J (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403:853-858.
- Ning Z, Cox AJ & Mullikin JC (2001). SSAHA: A fast search method for large DNA databases. *Genome Research*, 11, 1725–1729.
- Oliveira, J. F., de Oliveira, J. C. D., Nascimento, L., Fernandes, R. T. V., Novaes, J. L. C., & Peretti, D. (2020). Feeding patterns, trophic structure and damming rivers effect: studies applied in freshwater environments in the Brazilian semi-arid region. *Journal of Animal Behaviour and Biometeorology*, 6(2), 41-47.

- Oomen, R. A., & Hutchings, J. A. (2022). Genomic reaction norms inform predictions of plastic and adaptive responses to climate change. *Journal of Animal Ecology*, 91(6), 1073-1087.
- Paris JR, Stevens JR, & Catchen JM (2017). Lost in parameter space: a road map for stacks. *Methods in Ecology and Evolution*, 8(10), 1360-1373.
- Pereira A.R., Akama A. (2022) Diversidade dos cascudos da subfamília Hypostominae das corredeiras do baixo rio Tocantins com a utilização de DNA barcode. *Biblioteca do Museu Domingos Soares Ferreira Penna. Caderno de resumos e a pluralidade da pesquisa no Museu Paraense Emílio Goeldi*, 67.
- Peterson BK, Weber JN, Kay EH, Fisher HS & Hoekstra HE (2012). Double digest RADseq: an inexpensive method for de novo SNPs discovery and genotyping in model and non-model species. *PloS one*, 7(5), e37135.
- Primmer CR (2009). From conservation genetics to conservation genomics. *Annals of the New York Academy of Sciences*, (1162) 357-68.
- Pritchard JK, Stephens M & Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945-959.
- Rambaut A, Suchard MA Xie D, Drummond AJ (2014). Tracer v1.6. Available at: <http://beast.bio.ed.ac.uk/Tracer>
- Reis, R. E., Albert, J. S., Di Dario, F., Mincarone, M. M., Petry, P., & Rocha, L. A. (2016). Fish biodiversity and conservation in South America. *Journal of fish biology*, 89(1), 12-47.
- Rudke, A. P., Xavier, A. C. F., Fujita, T., Abou Rafee, S. A., Martins, L. D., Morais, M. V. B., ... & Martins, J. A. (2021). Mapping past landscapes using landsat data: Upper Paraná River Basin in 1985. *Remote Sensing Applications: Society and Environment*, 21, 100436.
- Saitou N & Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*, 4(4), 406-425.
- Silveira, E. L. D., Ballester, E. L. C., Costa, K. A. D., Scheffer, E. W. D. O., & Vaz-dos-Santos, A. M. (2018). Fish community response to environmental variations in an impacted Neotropical basin. *Ecology of Freshwater Fish*, 27(4), 1126-1139.

Slaby, S., Le Cor, F., Dufour, V., Auger, L., Pasquini, L., Cardoso, O., ... & Banas, D. (2022). Distribution of pesticides and some of their transformation products in a small lentic waterbody: Fish, water, and sediment contamination in an agricultural watershed. *Environmental Pollution*, 292, 118403.

Sousa, L. (2022). IctioXingu: Um Parancistrus LINDO do rio Araguaia (disponível em: https://www.youtube.com/watch?v=Z8Rtj16mlbs&ab_channel=LeandroSousa). Acesso: dez2022.

Tacon, A. G., Lemos, D., & Metian, M. (2020). Fish for health: improved nutritional quality of cultured fish for human consumption. *Reviews in Fisheries Science & Aquaculture*, 28(4), 449-458.

Tamura K, Stecher G, Peterson D, Filipinski A & Kumar S (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution*, 30(12), 2725-2729.

Tang J, Vosman B, Voorrips RE, Van der Linden CG & Leunissen JA (2006). QualitySNP: a pipeline for detecting single nucleotide polymorphisms and insertions/deletions in EST data from diploid and polyploid species. *BMC bioinformatics*, 7(1), 438.

Tomczyk, P., & Wiatkowski, M. (2020). Shaping changes in the ecological status of watercourses within barrages with hydropower schemes: literature review. *Archives of Environmental Protection*, 46(4).

Xia X & Xie Z (2001). DAMBE: software package for data analysis in molecular biology and evolution. *Journal of Heredity*, 92(4), 371-373.

Walpole JM, Morgan-Davis M, Bett P & Leader-Williams N (2001). Population dynamics and future conservation of a free-ranging black rhinoceros (*Diceros bicornis*) population in Kenya. *Biology. Conserv.* 99, 237-243.

Weir BS & Cockerham CC (1984). Estimating F statistics for the analysis of population structure. *Evolution*, 38(6), 1358-1370.

Wright S (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, 19(3), 395-420.

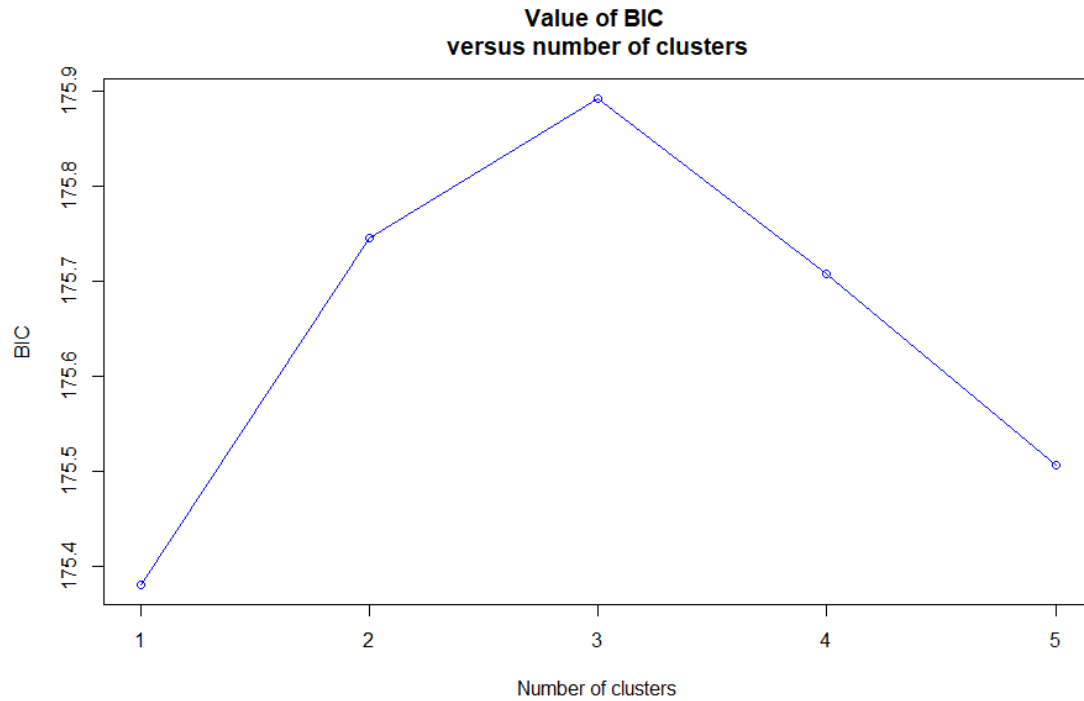
Yeh FC & Boyle T (1996). Popgene v1.32 Microsoft Windows based software for population genetic analysis. Edmonton: University of Alberta.

Zhang HW, Yin SW, Zhang LJ, Hou XY, Wang YY (2015). Development and validation of single nucleotide polymorphism markers in *Odontobutis potamophila* from transcriptomic sequencing. *Genetics and Molecular Research*, 14 (1), 2.

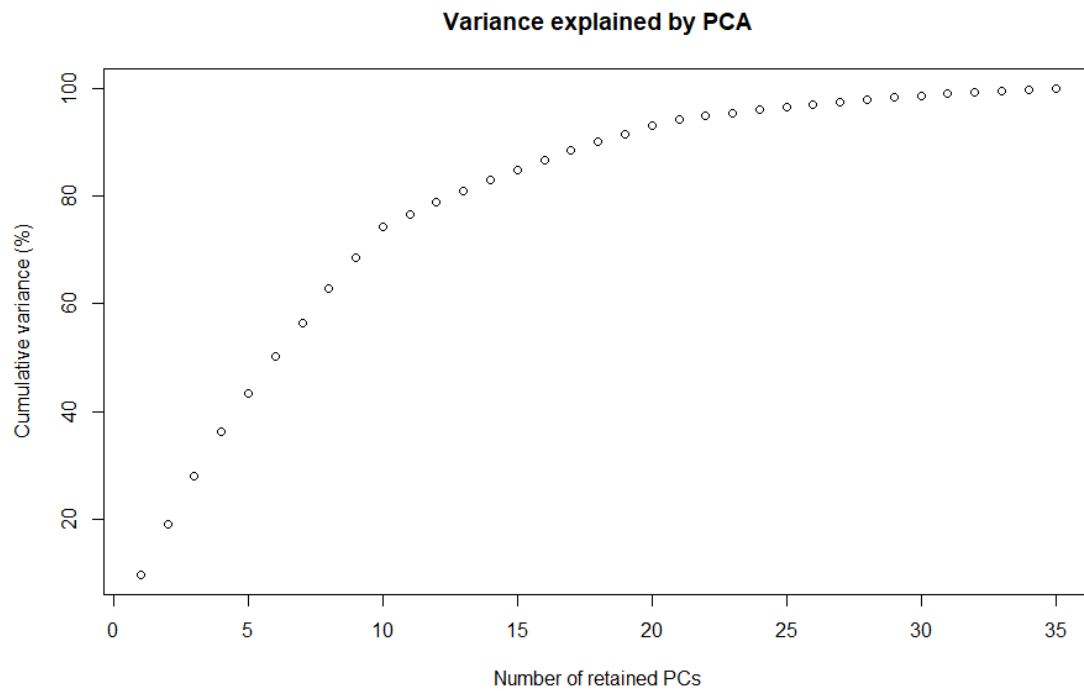
Zolet, A. C. T., Turchetto, C., Zanella, C. M., & Passaia, G. (2017). Marcadores moleculares na era genômica: metodologias e aplicações.

MATERIAL SUPLEMENTAR

Material Suplementar S1. Cálculo das estatísticas associadas resumidas (por padrão, BIC) com número máximo de 5 clusters.



Material Suplementar S2. Cálculo das variâncias cumulativas de 70% (por padrão, PCA) com 10-12 PCs retidas.



Material Suplementar S3. Tabela com resultado da estrutura populacional de K 1-7 com base nas análises do STRUCTURE.

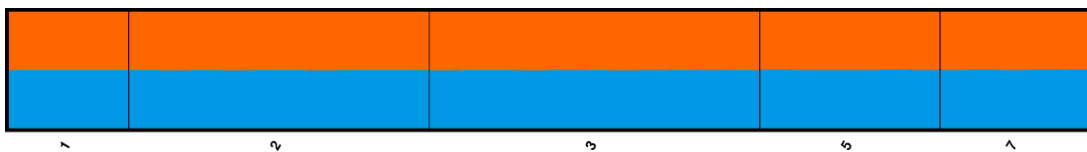
K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	Ln''(K)	Delta K
1	10	-43280.12000	18.20988	NA	NA	NA
2	10	-43294.87000	9.14027	-14.75000	14.84000	1.62358
3	10	-43294.78000	11.91095	0.09000	0.80000	0.06717
4	10	-43293.89000	17.95417	0.89000	2.53000	0.14091
5	10	-43295.53000	12.40798	-1.64000	5.73000	0.46180
6	10	-43291.44000	9.03957	4.09000	12.13000	1.34188
7	10	-43299.48000	15.94573	-8.04000	NA	NA

Material Suplementar S4. Estrutura populacional de K 1-7. Cada linha vertical representa um indivíduo. Os pontos e as localidades pertencentes foram colocados abaixo da figuras. As cores representam a probabilidade de atribuição de cada indivíduo a cada cluster. K=1 foi o que melhor se ajustou ao nosso resultado baseado no método Delta K.

$k=1$



$k=2$



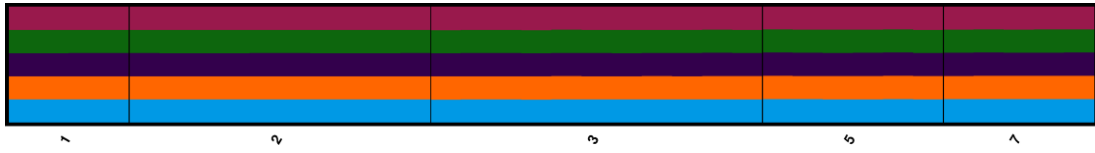
$k=3$



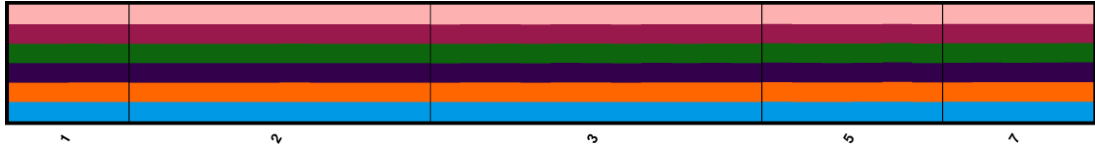
$k=4$



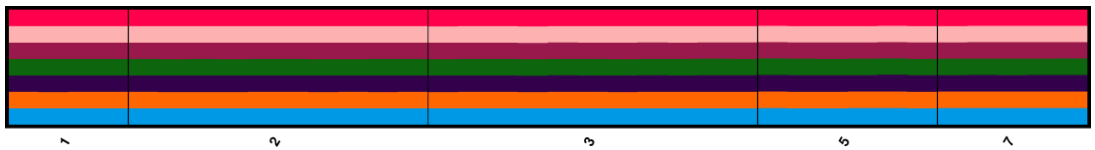
$k=5$



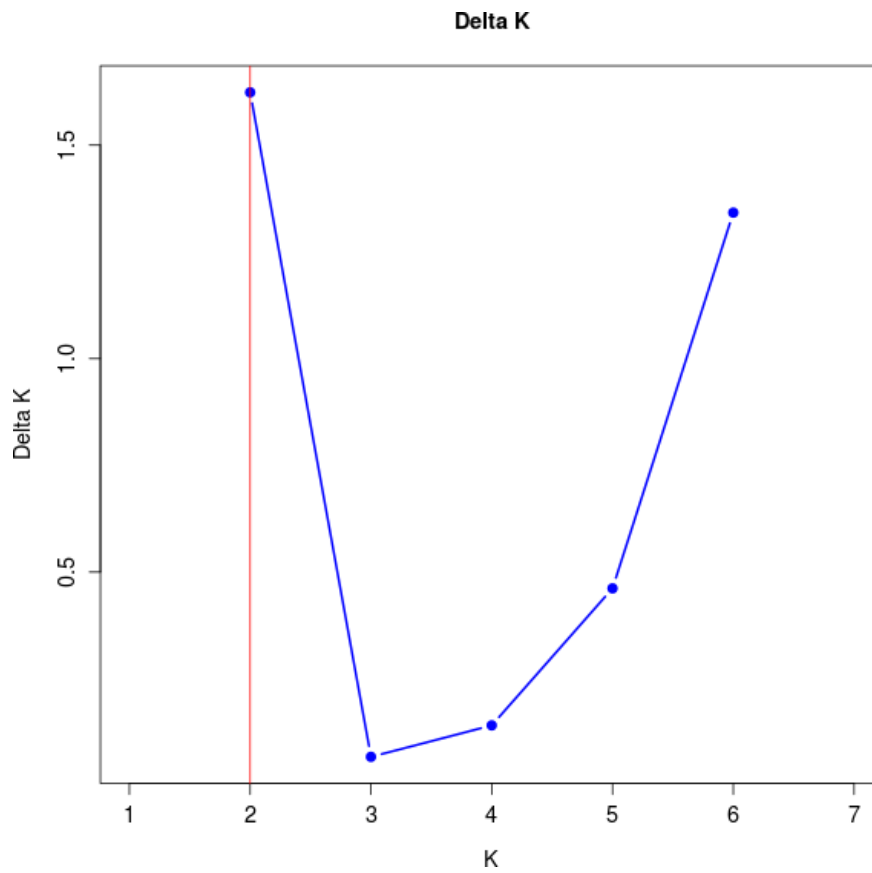
$k=6$



$k=7$



Material Suplementar S5. A distribuição de ΔK estimada a partir da análise Bayesiana implementada pelo STRUCTURE.



Material Suplementar S6. A probabilidade posterior média de cada K (7 repetições) estimada a partir da análise Bayesiana, conforme implementado pelo STRUCTURE.

