



**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**ESTRUTURA GENÉTICA DE POPULAÇÕES DE PIRARUCU (*Arapaima gigas*) NAS
BACIAS DOS RIOS AMAZONAS E TOCANTINS-ARAGUAIA**

ALINE ARAÚJO CAMPELO

DISSERTAÇÃO DE MESTRADO EM CIÊNCIAS ANIMAIS

**BRASÍLIA/DF
MARÇO DE 2024**



**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE AGRONOMIA E MEDICINA VETERINÁRIA**

**ESTRUTURA GENÉTICA DE POPULAÇÕES DE PIRARUCU (*Arapaima gigas*) NAS
BACIAS DOS RIOS AMAZONAS E TOCANTINS-ARAGUAIA**

ALINE ARAÚJO CAMPELO

ORIENTADOR: Dr. ALEXANDRE RODRIGUES CAETANO

COORIENTADORA: Dra. PATRÍCIA IANELLA

DISSERTAÇÃO DE MESTRADO EM CIÊNCIAS ANIMAIS

PUBLICAÇÃO: FEVEREIRO/2024

**BRASÍLIA
MARÇO DE 2024**

REFERÊNCIA BIBLIOGRÁFICA E CATALOGAÇÃO

CAMPELO, A. A. Estrutura Genética de Populações de Pirarucu (*Arapaima gigas*) nas Bacias dos Rios Amazonas e Tocantins-Araguaia. Brasília: Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília, 2024, Xp. Dissertação de Mestrado.

Documento formal, autorizando reprodução desta dissertação de mestrado para empréstimo ou comercialização, exclusivamente para fins acadêmicos, foi passado pelo autor à Universidade de Brasília e acha-se arquivado na Secretaria do Programa. O autor e o seu orientador reservam para si os outros direitos autorais, de publicação. Nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor ou do seu orientador. Citações são estimuladas, desde que citada a fonte.

FICHA CATALOGRÁFICA

CAMPELO, Aline Araújo. Estrutura Genética de Populações de Pirarucu (*Arapaima gigas*) nas Bacias dos Rios Amazonas e Tocantins-Araguaia. Brasília: Faculdade de Agronomia e Medicina Veterinária da Universidade de Brasília, 2024. Xp. Dissertação (Mestrado em Ciências Animais) - Faculdade de Agronomia e Medicina Veterinária da Universidade de Brasília, 2024.

1. Análises genéticas.
2. Chip de genotipagem.
3. Endogamia.
4. Marcadores moleculares.
5. Piscicultura.

**UNIVERSIDADE DE BRASÍLIA. FACULDADE DE AGRONOMIA E
MEDICINA VETERINÁRIA**

**ESTRUTURA GENÉTICA DE POPULAÇÕES DE PIRARUCU (*Arapaima gigas*) NAS
BACIAS DOS RIOS AMAZONAS E TOCANTINS-ARAGUAIA**

ALINE ARAÚJO CAMPELO

**DISSERTAÇÃO DE MESTRADO SUBMETIDA AO
PROGRAMA DE PÓS-GRADUAÇÃO EM
CIÊNCIAS ANIMAIS, COMO PARTE DOS
REQUISITOS NECESSÁRIOS À OBTENÇÃO DO
GRAU DE MESTRE EM CIÊNCIAS
ANIMAIS.**

APROVADA POR:

**Alexandre
Rodrigues Caetano**

Assinado de forma digital por
Alexandre Rodrigues Caetano
Dados: 2024.03.20 08:15:39 -03'00'

**ALEXANDRE RODRIGUES CAETANO, Doutor (Embrapa Recursos Genéticos e
Biotecnologia) (ORIENTADOR)**



Documento assinado digitalmente
CONCEPTA MARGARET MCMANUS PIMENTEL
Data: 20/03/2024 19:09:37-0300
Verifique em <https://validar.iti.gov.br>

**CONCEPTA MARGARET MCMANUS PIMENTEL, Doutora (Universidade de
Brasília) (EXAMINADOR INTERNO)**



Documento assinado digitalmente
PAULO CESAR VENERE
Data: 20/03/2024 11:18:07-0300
Verifique em <https://validar.iti.gov.br>

**PAULO CESAR VENERE, Doutor (Universidade Federal de Mato Grosso)
(EXAMINADOR EXTERNO)**

BRASÍLIA/DF, 19 DE MARÇO DE 2024

AGRADECIMENTOS

Agradeço aos meus pais por todo apoio e oportunidades de aprendizado que me deram, isso me forneceu o conhecimento necessário para chegar até aqui.

Agradeço aos meus amigos e colegas de trabalho Nayelle Meire, Daíza Orth e especialmente Káifer Soares por todo suporte profissional e emocional durante a realização deste trabalho. Ao meu querido amigo Noeliton Araújo, muito obrigada por me ensinar a fazer tantas análises genéticas, seu comprometimento com o trabalho e a sua generosidade para ajudar e ensinar os colegas inspira a todos nós.

Ao meu companheiro Caio Uchôa, muito obrigada por todo apoio, motivação e compreensão durante este processo, seu suporte em diversos momentos, foi essencial para a conclusão dessa etapa.

Aos meus queridos orientadores Dr. Alexandre Rodrigues Caetano e Dra. Patrícia Ianella, muito obrigada por todos os ensinamentos, orientações, paciência e dedicação para me ajudar a concluir este trabalho, sei que muitas vezes foi difícil conciliar suas atividades com as minhas demandas, especialmente por conta das mudanças que ocorreram no último ano e por isso, sempre serei muito grata a vocês.

Aos membros da banca examinadora, Prof. Dr. Paulo César Venere e Prof^a. Dra. Concepta Margaret McManus agradeço por terem aceitado contribuir com seus conhecimentos nesta dissertação.

Meus agradecimentos à Embrapa Recursos Genéticos e Biotecnologia pelo fornecimento da estrutura e apoio profissional para a realização da pesquisa e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pelo suporte financeiro para a realização deste trabalho.

Agradeço também aos colegas dos Institutos Mamirauá e Juruá, pelo fornecimento de amostras de Pirarucu da região amazônica, ao produtor Pedrinho, da piscicultura Boa Esperança, pelo fornecimento de amostras de sua propriedade em Rondônia, e ao colega Prof. Dr. Paulo Venere pelo fornecimento de amostras representativas da bacia Tocantins-Araguaia. Seu apoio foi essencial para a realização deste trabalho.

Finalmente, agradeço a todos que de forma direta ou indireta me ajudaram na trajetória e na conclusão desta importante etapa da minha vida acadêmica e profissional.

SUMÁRIO

CAPÍTULO 1	1
INTRODUÇÃO E REVISÃO BIBLIOGRÁFICA	1
1.1 INTRODUÇÃO	2
1.2 OBJETIVOS	4
1.2.1 Objetivo Geral	4
1.2.2 Objetivos Específicos	4
1.3 REVISÃO DE LITERATURA	5
1.3.1 Aspectos gerais do <i>Arapaima gigas</i>	5
1.3.2 Genética da conservação do <i>Arapaima gigas</i>	9
1.3.3 Marcadores SNPs e Genotipagem Fluidigm.....	12
1.3.4 O cultivo do <i>Arapaima gigas</i> no Brasil	16
REFERENCIAS BIBLIOGRÁFICAS	21
CHAPTER 2	25
GENETIC STRUCTURE OF PIRARUCU (<i>Arapaima gigas</i>) POPULATIONS IN THE AMAZONAS AND TOCANTINS-ARAGUAIA RIVER BASINS	25
2.1 RESUMO	26
2.2 ABSTRACT	28
2.3 INTRODUCTION	30
2.4 MATERIAL AND METHODS	31
2.4.1 Sampling	31
2.4.2 DNA extraction.....	33
2.4.3 SNP mining and assay design	33
2.4.4 SNP genotyping	34
2.4.5 Data quality control	34
2.4.6 Genetic Diversity	34
2.4.7 Genetic Structure	34
2.4.8 Genetic versus geographic distances	35

2.4.9 Kinship analysis	35
2.5 RESULTS	35
2.5.1 Data quality control	35
2.5.2 Genetic diversity	37
2.5.3 Genetic Structure	41
2.5.4 Genetic and geographic distances.....	43
2.5.5 Kinship analysis	45
2.6 DISCUSSION.....	47
2.7 CONCLUSION	53
REREFENCES	54
SUPPLEMENTARY MATERIAL	57
CAPÍTULO 3	61
CONSIDERAÇÕES FINAIS	61
3.1 CONSIDERAÇÕES FINAIS	62

RESUMO

ESTRUTURA GENÉTICA DE POPULAÇÕES DE PIRARUCU (*Arapaima gigas*) NAS BACIAS DOS RIOS AMAZONAS E TOCANTINS-ARAGUAIA

O Pirarucu é o maior peixe de escamas de água doce do mundo e é um importante recurso alimentar e cultural dos indígenas e ribeirinhos na região amazônica. No entanto, a espécie está ameaçada pela pesca excessiva, perda de habitat e poluição. Estudos genéticos têm sido realizados para avaliar a diversidade genética e o estado de conservação das populações de Pirarucu. Esses estudos mostraram que essas populações apresentam baixa diversidade genética, principalmente na bacia hidrográfica do Tocantins-Araguaia. O objetivo deste estudo foi analisar a estrutura genética das populações de Pirarucu selvagens e de cativeiro das bacias Amazônica e Tocantins-Araguaia, utilizando um painel de SNPs (do inglês *Single Nucleotide Polymorphisms*) de baixa densidade, gerando dados úteis para subsidiar a elaboração de estratégias de utilização e conservação do germoplasma da espécie. Foram realizadas análises de diversidade, estrutura, distância genética e parentesco dentro e entre as populações analisadas. Os resultados mostram que as populações da bacia Amazônica apresentaram maior diversidade geral em comparação às populações do Tocantins-Araguaia. As médias de heterozigosidade observada e esperada (H_o e H_e) foram, respectivamente de 0.125 e 0.117 para as populações do Tocantins-Araguaia e de 0.338 e 0.317 para as populações da bacia Amazônica. O F_{st} total foi de 0.424, com estimativas par-a-par variando entre 0.001 e 0.789. Dois clusters principais foram identificados nas análises, separando as populações com base na bacia hidrográfica de origem (Tocantins-Araguaia ou Amazônica), confirmando resultados de estudos semelhantes com populações de ambas as bacias. Análises mais aprofundadas revelaram a existência de mais dois grupos de populações. Na região do sudeste paraense foi detectada uma possível zona de *admixture* entre as populações das bacias Tocantins-Araguaia e Amazônica, e as populações da região de Santarém-PA foram separadas das outras pertencentes a Bacia Amazônica. Foi encontrada uma correlação significativa entre as distâncias genéticas e geográficas, sugerindo que o isolamento por distância desempenha um papel na formação da estrutura genética dessas populações ($r=0.5824$). Os painéis de SNPs de baixa densidade utilizados neste estudo foram eficientes para gerar dados genéticos úteis para a implementação de medidas adequadas à conservação e ao cultivo da espécie, e poderão ser utilizados para melhorar sua produção em cativeiro e atender a demanda dos piscicultores sem causar maiores danos às populações naturais.

PALAVRAS-CHAVE: Conservação, Genotipagem, Piscicultura, SNPs.

ABSTRACT

GENETIC STRUCTURE OF PIRARUCU (*Arapaima gigas*) POPULATIONS IN THE AMAZONAS AND TOCANTINS-ARAGUAIA RIVERS BASINS

The Pirarucu is the largest scaled freshwater fish in the world and is an important food and cultural resource for indigenous people and riverine communities in the Amazon region. However, the species is threatened by overfishing, habitat loss, and pollution. Genetic studies have been conducted to assess the genetic diversity and conservation status of Pirarucu populations. These studies have shown that these populations exhibit low genetic diversity, particularly in the Tocantins-Araguaia watershed. The aim of this study was to analyze the genetic structure of wild and captive Pirarucu populations in the Amazon and Tocantins-Araguaia basins using a panel of low-density SNPs (Single Nucleotide Polymorphisms) to generate data useful for developing strategies for the use and conservation of the species' germplasm. Analyses of diversity, structure, genetic distance, and relatedness within and between the analyzed populations were conducted. The results indicate that populations in the Amazon basin show higher overall diversity compared to Tocantins-Araguaia populations. The observed and expected heterozygosity averages (H_o and H_e) for Tocantins-Araguaia populations were 0.125 and 0.117, respectively, and 0.338 and 0.317 for Amazon basin populations. The total F_{st} was 0.424, with pairwise estimates ranging from 0.001 to 0.789. Two main clusters were identified in the analyses, separating populations based on their watershed of origin (Tocantins-Araguaia or Amazon), confirming similar results from studies with populations in both basins. Further analyses revealed two additional population groups. In the southeast Pará region, a potential admixture zone was detected between populations from the Tocantins-Araguaia and Amazon basins, and populations in the Santarém-PA region were separated from others in the Amazon basin. A significant correlation was found between genetic and geographical distances, suggesting that isolation by distance plays a role in shaping the genetic structure of these populations ($r=0.5824$). The low-density SNP panels used in this study were effective in generating genetic data useful for implementing conservation and cultivation measures for the species, and they could be utilized to enhance captive production while meeting the demands of fish farmers without causing significant harm to natural populations.

KEYWORDS: Conservation, Fish-farming, Genotyping, SNPs.

CAPÍTULO 1
INTRODUÇÃO E REVISÃO BIBLIOGRÁFICA

1.1 INTRODUÇÃO

O Pirarucu (*Arapaima gigas*) é considerado o maior peixe de escamas de água doce do mundo, podendo atingir 2,5m de comprimento e 250kg de peso vivo (NELSON; GRANDE; WILSON, 2016). A espécie pode ser encontrada nas bacias hidrográficas do Brasil, Peru, Colômbia, Equador, Bolívia e Guiana. No território brasileiro, esta espécie habita, principalmente, as bacias Amazônica e Tocantins-Araguaia (CASTELLO; STEWART, 2010).

O Pirarucu foi uma das principais fontes de alimento da dieta dos ameríndios que viviam nas regiões de ocorrência desta espécie. Após a chegada dos europeus à América do Sul a espécie ocupou papel de extrema importância social e econômica na Amazônia e ao longo dos rios na bacia Tocantins-Araguaia (HRBEK et al., 2005; KIRSTEN et al., 2017). Até os dias de hoje, as populações locais consomem e comercializam a carne fresca ou salgada do Pirarucu, utilizando-a em pratos tradicionais. Além disso, seus subprodutos, como couro e carcaça, são utilizados em diversos artesanatos, confecção de vestuário e itens de decoração (FERREIRA; MARCOVITCH; VAL, 2019).

O Pirarucu possui respiração aérea obrigatória e por isso é um alvo fácil para os pescadores regionais. Os animais são capturados com arpões quando vêm à superfície para respirar. Devido aos diversos usos de sua carne, este peixe tornou-se muito explorado comercialmente e, principalmente a partir da década de 70, quando a pesca predatória se intensificou, houve uma drástica redução ou erradicação das populações naturais perto das principais cidades da bacia amazônica (HRBEK et al., 2005; CASTELLO; STEWART; ARANTES, 2011).

A constante perda de habitat natural também influenciou na redução das populações naturais de Pirarucu, colocando-o desde 1986 na lista da Convenção sobre o Comércio Internacional das Espécies Silvestres Ameaçadas de Extinção (CITES) como um alimento ameaçado de extinção (SANTOS et al., 2014). O Pirarucu também está na lista vermelha da União Internacional para a Conservação da Natureza (IUCN) como “*Data Deficient*”, pois, segundo a entidade, ainda não há conhecimento suficiente de sua biologia, ecologia e genética para coordenar com eficácia sua conservação (FARIAS et al., 2019).

O cultivo de Pirarucu em cativeiro oferece uma alternativa importante à pesca extrativista, e, portanto, uma oportunidade de integrar os objetivos de manejo sustentável dos estoques selvagens e redução da pobreza na região amazônica (CAMPOS-SILVA; PERES, 2016). A piscicultura da espécie já contribuiu para a redução de sua superexploração e a

ampliação desse cultivo deve resultar na expansão no mercado nacional e na criação de condições para a conquista de mercados internacionais.

Em cativeiro, existem algumas dificuldades que devem ser superadas. A falta de um manejo genético adequado dos reprodutores e o pequeno número de peixes fundadores nos estoques vindo de populações altamente estruturadas podem causar depressão endogâmica por conta do acasalamento entre indivíduos aparentados (FAZZI-GOMES et al., 2021). Isso leva à diminuição da diversidade genética dos animais e, conseqüentemente, afeta a adaptabilidade das espécies de peixes cultivados, incluindo sua capacidade de resistir a doenças e a variações no ambiente. Por essas razões, populações com baixos níveis de variabilidade genética podem comprometer o desempenho do Pirarucu na aquicultura (WASKO et al., 2004; SANTOS et al., 2014).

Estudos que investiguem a estrutura genética dos plantéis de reprodutores de espécies aquícolas mantidos em cativeiro podem auxiliar na resolução de questões associadas ao manejo genético dos animais para produção de alevinos destinados à engorda ou reposição. Para espécies de grande interesse comercial, como a truta e a tilápia, estudos com marcadores moleculares do tipo SNPs vem sendo amplamente realizados, gerando informações para análises de parentesco, diversidade genética e avaliação e seleção genômica (BOHLING; HAFFRAY; BERREBI, 2016; JOSHI et al., 2018; ARAÚJO et al., 2023). Já para o Pirarucu, a maioria das pesquisas que vem sendo publicadas utilizam marcadores mtDNA e microssatélites. (ARARIPE et al., 2013; HRBEK et al., 2005; VITORINO et al., 2017).

É sabido que a variabilidade genética tem implicações importantes para a organização e domesticação de planteis de peixes reprodutores, podendo alavancar o cultivo e contribuir com o desenvolvimento de estratégias para conservação de populações selvagens. Uma vez que os estoques naturais do Pirarucu estão sendo reduzidos por fatores como a degradação ambiental e exploração comercial extrema, é necessário disponibilizar mais dados de estrutura genética das populações desses peixes para beneficiar a produção em cativeiro e auxiliar na conservação das populações naturais (CARLOS; VERA, 2014; TORATI et al., 2019).

O presente estudo tem como objetivo fornecer informações genéticas adicionais sobre o Pirarucu, que permitem analisar o perfil de conservação das populações naturais e plantéis de reprodutores cativos, utilizando, pela primeira vez para essa espécie, uma tecnologia de genotipagem de SNPs com alta acurácia e de baixo custo da Fluidigm, que provou ser eficaz para o proposto neste trabalho e pode ser utilizada em estudos futuros com espécies aquícolas de interesse comercial emergente, como o Pirarucu.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

- Caracterizar geneticamente populações de Pirarucu (*Arapaima gigas*) naturais e de cativeiro com um painel de SNPs (Single Nucleotide Polymorphisms) de baixa densidade.

1.2.2 Objetivos Específicos

- Analisar a variabilidade genética em populações de Pirarucu naturais e em cativeiro;
- Avaliar as diferenças na diversidade e estruturação genética dentro e entre populações de Pirarucu naturais e de cativeiro;
- Disponibilizar informações úteis para subsidiar a elaboração de estratégias de uso e conservação de germoplasma da espécie.

1.3 REVISÃO DE LITERATURA

1.3.1 Aspectos gerais do Pirarucu

O Pirarucu (*Arapaima gigas*, Schinz, 1822) pertence à ordem Osteoglossiforme, um dos grupos mais primitivos de peixes teleósteos emergentes do período Jurássico, e são caracterizados pela sua língua óssea e áspera e suas escamas grandes e grossas. Esses peixes vêm sendo o foco de vários estudos sobre morfologia, sistemática e evolução, por conta da sua posição ancestral entre os teleósteos atuais (HILTON; LAVOUÉ, 2018).

O Pirarucu prefere ambientes lênticos e águas doces tropicais, com temperaturas entre 25 e 29°C (MOREIRA DA SILVA; LUIZ; DUNCAN, 2016). Ele é nativo das áreas de várzea nas bacias dos rios Solimões-Amazonas e Essequibo, sendo encontrado, no exterior, no Peru, Bolívia e Guiana. No Brasil, ele também pode ser encontrado em outras bacias hidrográficas, além da Amazônica, como a Tocantins-Araguaia (CASTELLO; STEWART, 2010). A Bacia Amazônica é a maior bacia hidrográfica do mundo e possui cerca de 7.500.000 km² de extensão.

O principal rio dessa bacia é o Amazonas, que nasce na Cordilheira dos Andes, no Peru, e desagua no Oceano Atlântico, no Brasil. O rio e seus afluentes passam por vários estados brasileiros, como o Acre, Amapá, Amazonas, Pará, Rondônia, Roraima, Tocantins e Mato Grosso (JUNK, 1984). Algumas localidades, onde são formadas as planícies de inundação, possuem flora altamente adaptada ao regime de inundações característicos da região e compõem, juntamente com toda a diversidade local, o habitat de preferência do Pirarucu, conhecido como várzea ou mata de igapó (JUNK, 1984; VITORINO et al., 2017)..



Figura 1 – Pirarucu (*Arapaima gigas*). Fonte: <http://www.klimanaturali.org/2012/12/Pirarucu-arapaima-gigas.html>

A Bacia do Tocantins-Araguaia também apresenta um regime de cheias e secas em seus rios e formação de lagos ao redor do canal principal. Assim como a Bacia Amazônica, possui alta diversidade biológica e importância para conservação. O principal rio é o Araguaia, que nasce na divisa dos estados Goiás e Mato Grosso no Centro-Oeste brasileiro e tem uma extensão de aproximadamente 2.100 km até se mesclar com o Rio Tocantins, na divisa entre os estados Pará e Maranhão, ao norte do Brasil (AQUINO; LATRUBESSE; SOUZA FILHO, 2009).

As áreas de inundação observadas nessas duas grandes bacias hidrográficas se formam de acordo com o ciclo de precipitações irregulares a montante da área de drenagem da bacia. O fluxo no canal do rio aumenta ou diminui de acordo com o pulso de inundação, formando as planícies. Essa variação é uma das principais forças que regulam o ecossistema de várzea, e, ao longo do tempo, vêm direcionando a adaptação das milhares de espécies que ali vivem (JUNK, 1984; MOREIRA DA SILVA; LUIZ; DUNCAN, 2016).

Os ecossistemas existentes em planícies de inundação são caracterizados por terem águas ricas em minerais e sedimentos, o que resulta em alta produtividade e disponibilidade de alimentos para as espécies aquáticas locais (JUNK, 1984). O Pirarucu, por exemplo, se alimenta de insetos, camarões e caranguejos de água doce nos primeiros estágios de vida, e quando adulto, consome diversas espécies de peixes, disponíveis em abundância nas áreas de várzea.

Esse hábito alimentar permite classificar a espécie como piscívora (DORIA et al., 2019; JUNK, 1984).

Em seu habitat natural, o Pirarucu faz somente migrações laterais, do canal do rio para a várzea inundada, para forragear. Diferentemente de outras espécies de peixes, o Pirarucu não nada rio acima no período de reprodução, que ocorre geralmente de novembro a janeiro, coincidindo com a época de cheia (FERNANDES, 1997; HRBEK; CROSSA; FARIAS, 2007). O atingimento da maturidade sexual da espécie pode variar de acordo com o sexo e as condições do ambiente. Geralmente, as fêmeas estão aptas a reproduzir entre os 6 e 7 anos de idade, já os machos, entre 3 e 4 anos (MOREIRA DA SILVA; LUIZ; DUNCAN, 2016; NÚÑEZ et al., 2011).

O Pirarucu forma casais na estação reprodutiva. Os machos constroem ninhos onde as fêmeas depositam os ovos que são fertilizados externamente. Nesse período, os machos podem apresentar uma coloração avermelhada no ventre e na cauda, além de apresentarem comportamentos territorialistas e agressivos com outros machos (CASTELLO, 2008; GODINHO et al., 2005). As fêmeas desovam diversas vezes na estação reprodutiva e, por isso, a fecundidade média da espécie, por ciclo de reprodução, na natureza, pode chegar a 11.000 alevinos durante a fase de cuidado parental (GODINHO et al., 2005). Após a desova, o ninho é protegido pelo macho e pela fêmea até a eclosão dos ovos, quando apenas o macho fica com a guarda dos filhotes, até que eles cresçam o suficiente para sobreviverem sozinhos (CASTELLO; STEWART, 2010).

Apesar das áreas de várzea onde o Pirarucu se reproduz naturalmente serem ambientes propícios para o desenvolvimento da biodiversidade, durante as secas, quando ocorre formação de lagoas ao redor do canal principal do rio, há diminuição da disponibilidade de oxigênio dissolvido nas águas. Espécies dependentes exclusivas do O₂ aquático tem dificuldade para sobreviver nesse novo ambiente. O Pirarucu, porém, possui adaptações evolutivas para lidar com essas condições (DE OLIVEIRA et al., 2012).

As características morfofisiológicas do Pirarucu são peculiares em comparação a outros peixes de água doce. Durante seu desenvolvimento, ele passa por uma transição de respirador aquático exclusivo para aéreo, quando sua bexiga natatória modificada é usada como órgão respiratório. Sendo assim, ele consegue sobreviver em ambientes aquáticos em hipóxia por meio da captação do ar atmosférico, que compõe cerca de 90% do oxigênio que ele necessita. (BRAUNER et al., 2004).

A capacidade de respirar ar atmosférico representa, portanto, uma grande vantagem evolutiva do Pirarucu. Por outro lado, essa adaptação se tornou um problema para a

sobrevivência da espécie, uma vez que para respirar o ar atmosférico, o Pirarucu precisa nadar até a superfície e colocar parte do seu corpo para fora da água a cada 20-30 minutos (MOREIRA DA SILVA; LUIZ; DUNCAN, 2016). Por conta desse comportamento a espécie é um alvo fácil para pescadores profissionais e ribeirinhos, que aproveitam a oportunidade para abater os animais (HRBEK et al., 2005).

A carne de Pirarucu é considerada um alimento de alta qualidade por ser firme, branca, sem ossos e quase sem gordura. Outras partes também podem ser aproveitadas, como a pele do animal, que é usada na fabricação de bolsas e sapatos (HRBEK et al., 2005). Essas características fizeram com que a pesca comercial de Pirarucu crescesse desenfreadamente entre os anos 70 e 80 (HRBEK; CROSSA; FARIAS, 2007). Nesse mesmo período, ocorreu uma exploração desenfreada da floresta amazônica, com o aumento da retirada da cobertura vegetal nativa a cada ano, chegando a mais de 20.000 km² de área desmatada na década de 80. Essa agressão ao meio ambiente atingiu diversas espécies e causou um grande desequilíbrio ambiental, que se perpetua até os dias de hoje (HRBEK; CROSSA; FARIAS, 2007; ILHA; ROSSO; SCHIESARI, 2019). Um dos resultados desse processo foi a extinção do Pirarucu em meados dos anos 80, principalmente perto das grandes cidades amazônicas. Em algumas áreas, a espécie chegou a desaparecer completamente (HRBEK et al., 2005).

A pesca comercial do Pirarucu existe desde o século 18, como alternativa para suprir a demanda por bacalhau, um alimento raro em países fora da Europa (HRBEK et al., 2005; HRBEK; CROSSA; FARIAS, 2007). Ao longo do tempo, o comércio que foi desenvolvido ao redor do pescado de Pirarucu passou a ser de suma importância para a economia das comunidades ao longo dos rios na bacia Amazônica e Tocantins-Araguaia (FARIAS et al., 2019; MENDES FILHO et al., 2021). Devido à alta demanda por este pescado e do declínio nas populações de Pirarucu nos anos 80, especialmente na Bacia Amazônica, medidas foram tomadas para resgatar os estoques selvagens da espécie. No início de 2001, o governo brasileiro proibiu toda a pesca de Pirarucu, exceto na reserva Mamirauá, próximo a Tefé, no rio Amazonas. Desde então, essa atividade é realizada de forma controlada dentro de unidades de conservação como a mencionada (ARARIPE et al., 2013; VITORINO et al., 2017)

Com o passar do tempo, as estratégias de proteção para o Pirarucu apresentaram resultados satisfatórios. As populações voltaram a crescer e a pesca regulada vêm ajudando a suprir a demanda do mercado por esse peixe (ARARIPE et al., 2013). Entretanto, esses esforços não foram suficientes para retirar a espécie da lista da CITES, descrito como um alimento ameaçado de extinção ou da lista vermelha da IUCN, onde ele é listado como “*Data Deficient*”, segundo a entidade, por falta de conhecimento suficiente de sua biologia, ecologia e genética para

gerenciar sua conservação (FARIAS et al., 2019; TORATI et al., 2019), esse cenário, porém, vem mudando ao longo do tempo, uma vez que mais estudos com a espécie estão sendo publicados.

Até o presente a pesca predatória irregular ameaça a existência do Pirarucu e alguns grupos de pesquisa se dedicam a estudar a estrutura genética de suas populações, principalmente na Bacia Amazônica (CATÂNEO et al., 2022; FARIAS et al., 2019; HRBEK et al., 2005; HRBEK; CROSSA; FARIAS, 2007; VITORINO et al., 2017). Esses estudos são essenciais para entender a diversidade dessa espécie e desenvolver e gerenciar estratégias de conservação importantes para o Pirarucu e, também, para a Bacia Amazônica em geral, pois, por ser um predador, ele controla as populações de outras espécies em seu habitat, exercendo um papel essencial no equilíbrio ambiental nesses ecossistemas (CASTELLO, 2008; MOREIRA DA SILVA; LUIZ; DUNCAN, 2016).

1.3.2 Genética da conservação do Pirarucu

As populações de peixes de água doce estão em constante ameaça devido à vários fatores, como perda de habitat, poluição e pesca predatória. Para gerenciar e proteger essas populações, estudos de genética da conservação tornaram-se essenciais, uma vez que a diversidade genética é crucial para a sobrevivência dessas populações a longo prazo (HRBEK et al., 2005a; VITORINO et al., 2017; VRIJENHOEK, 1998). A genética da conservação também ajuda na identificação de populações geneticamente distintas que requerem estratégias de manejo específicas. Além disso, pode fornecer informações sobre a história evolutiva e as relações das populações de peixes, que são críticas na determinação de medidas de conservação apropriadas para cada espécie (CATÂNEO et al., 2022; SAINT-PÉ et al., 2019; VRIJENHOEK, 1998; ZHANG; WANG, 2003).

Em busca de gerar dados para contribuir com a conservação do Pirarucu, diversos autores publicaram estudos a respeito da estrutura genética das populações dessa espécie. Hrbek et al. (2005), analisaram o DNA mitocondrial (mtDNA) de 139 indivíduos de Pirarucu em 8 localidades na Bacia Amazônica. Entre outros achados, os autores relataram maior diversidade de haplótipos em áreas mais distantes de grandes centros urbanos, o que foi associado ao maior distanciamento das áreas mais degradadas pela exploração humana.

Hrbek e colaboradores (2007) fizeram uma análise de autocorrelação espacial de dados moleculares com dados geográficos de Pirarucu utilizando dados de microssatélites e mtDNA de 126 peixes amostrados em sete locais da Bacia Amazônica e observaram que, em distâncias maiores que 2.500 km², a troca genética entre as populações é significativamente baixa. Com

base nesses resultados os autores propuseram a criação de um modelo de gestão e conservação de metapopulações fonte-sumidouro, em que devem ser criadas reservas em áreas de várzea em diferentes pontos da Bacia Amazônica. Nesse modelo de dinâmica populacional, duas populações distintas trocam indivíduos por dispersão: a população fonte e a população sumidouro (DALLOZ, 2013).

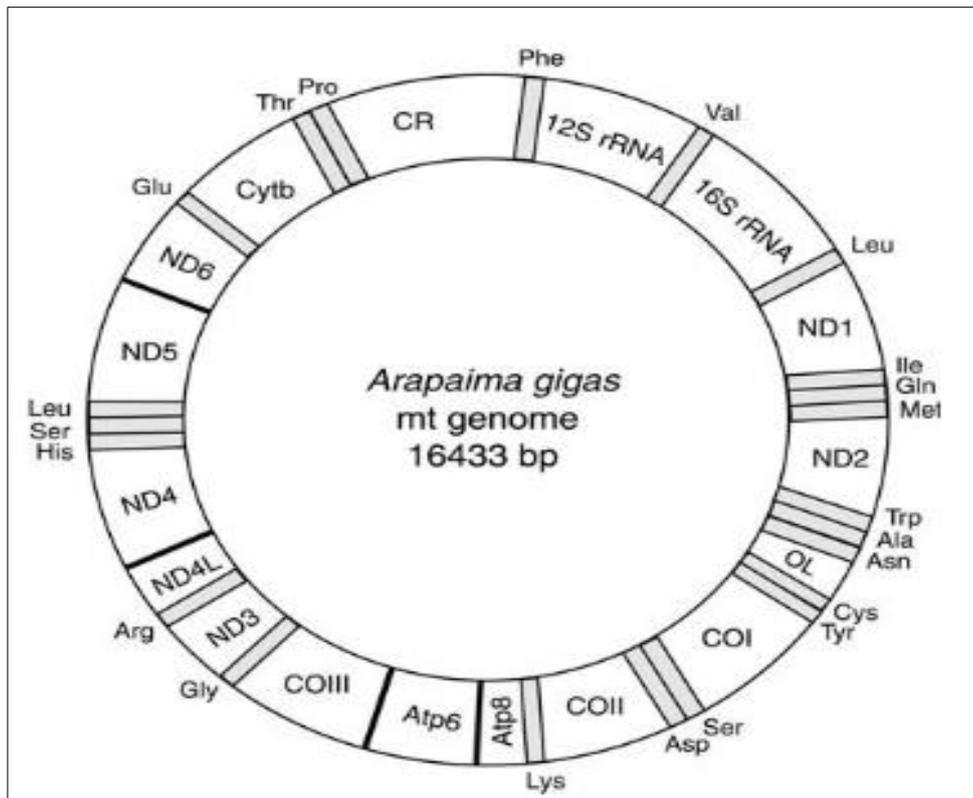


Figura 2 – Mapa esquemático do genoma mitocondrial completo do Pirarucu. Fonte: (HRBEK; FARIAS, 2008).

Os estudos pioneiros de Hrbek e colaboradores mostram como as ferramentas moleculares poderiam contribuir para a formulação de estratégias de conservação. Em 2008, o genoma mitocondrial completo do Pirarucu foi sequenciado, descobrindo seus 16.433 pares de base com 13 genes codificadores de proteínas. Esse trabalho abriu portas para mais estudos moleculares com as populações dessa espécie e de outros Osteoglossiformes (HRBEK; FARIAS, 2008).

Santos et al. (2014) usaram microssatélites para investigar a estrutura genética de 218 animais de duas populações selvagens e cinco de cativeiro de Pirarucu. Os autores constataram que as populações de cativeiro apresentam menor diversidade que as populações selvagens amostradas, e concluíram que em geral as populações têm alta estruturação genética. Eles propuseram, com base em seus resultados, um programa de manejo genético para populações

de cativeiro, com o objetivo de evitar a perda de variabilidade genética e a seleção de características desfavoráveis para fins comerciais e de conservação (CARLOS; VERA, 2014).

No Brasil, os estudos de genética de conservação para o Pirarucu focaram, em grande maioria, nas populações da Bacia Amazônica. Vitorino et al. (2015) realizaram um trabalho que avaliou, pela primeira vez, a estrutura genética das populações de Pirarucu da Bacia do Tocantins-Araguaia. Utilizando uma metodologia baseada em análises de sequências *Inter Simple Sequence Repeat* (ISSR) e do Citocromo Oxidase Subunidade I (COI), os autores descobriram que as populações apresentavam baixa diversidade genética, indicando que também havia a necessidade de criação de estratégias de manejo e conservação para as populações no Tocantins-Araguaia.

A necessidade de estratégias de conservação específicas para cada região de ocorrência ficou mais evidente quando trabalhos publicados em 2019 e 2020 confirmaram a baixa diversidade genética das populações do Tocantins-Araguaia, quando comparadas as populações da Bacia Amazônica. No artigo de Farias et al. (2019), os autores analisaram indivíduos de 19 localidades da Bacia Amazônica, contra apenas três do Tocantins-Araguaia, utilizando microssatélites descritos por Farias et al. (2003). Paralelamente, NOGUEIRA et al. (2020) utilizou fragmentos do Citocromo B (região de controle), COI e NADH desidrogenase subunidade 2, além de lócus microssatélites, para investigar a estrutura genética dos indivíduos de cinco locais da Bacia Amazônica e cinco locais da Bacia Tocantins-Araguaia.

A maioria dos estudos ressalta a importância de estratégias de conservação como a pesca manejada para a recuperação das populações de Pirarucu (ARARIPE et al., 2013; CASTELLO; STEWART, 2010; NOGUEIRA et al., 2020). Até então, a pesca do Pirarucu no Brasil tem sido regulamentada por instruções normativas nacionais e estaduais. A norma federal estabeleceu um comprimento mínimo do corpo de 150 cm para a pesca na Bacia Amazônica e Tocantins-Araguaia. Além disso, também foi estabelecido um tamanho mínimo para comercialização das partes processadas de Pirarucu de 120 cm quando fresco ou 110 cm se salgado (NOGUEIRA et al., 2020). No entanto, essas medidas têm se mostrado ineficazes devido à vasta área ocupada pela espécie, resultando em extrativismo ilegal e ameaça aos estoques, principalmente em áreas com reduzida diversidade genética como a bacia do Tocantins-Araguaia (FARIAS et al., 2019; NOGUEIRA et al., 2020b).

Para conservar o Pirarucu e sua diversidade genética é crucial que existam incentivos para implementar medidas de conservação, especialmente considerando as características ecológicas e comportamentais variáveis da espécie em diferentes ambientes. Além disso, segundo Nogueira et al. (2020), mais estudos ecológicos e genéticos são necessários,

particularmente em áreas-chave como o estuário do Amazonas, rios Negro, Purus e Madeira no Brasil, e o rio Essequibo na Guiana. A avaliação da diversidade genética de outras áreas de gestão pesqueira, como os tributários dos rios Araguaia e Tocantins, também será vital para desenvolver e implementar estratégias de conservação eficazes adaptadas às características da população local (NOGUEIRA et al., 2020).

1.3.3 Marcadores SNPs e Genotipagem Fluidigm

Painéis de SNP de baixa e alta densidade tem sido utilizados em diversos estudos e análises de diversidade genética, parentesco, avaliação e seleção genômica em peixes comerciais como a Truta e a Tilápia (BOHLING; HAFFRAY; BERREBI, 2016; JOSHI et al., 2018; ARAÚJO, 2023). Para o Pirarucu, a maioria das pesquisas publicadas até o momento utilizaram os marcadores mtDNA e microssatélites, além do ISSR (*Inter Simple Sequence Repeat*) (ARARIPE et al., 2013; HRBEK et al., 2005; VITORINO et al., 2015).

Um Polimorfismo de Nucleotídeo Único (SNP – *Single Nucleotide Polymorphism*) é uma diferença em uma única base nitrogenada, em um nucleotídeo individual, observado em uma população. Os SNPs podem ser classificados de acordo com a substituição de nucleotídeos como transições (Citosina/Timina ou Guanina/Adenina) ou transversões (Citosina/Guanina, Adenina/Timina, Citosina/Adenina ou Timina/Guanina). Geralmente, eles ocorrem em espaços intergênicos, com função indeterminada, porém, mais raramente, também podem ocorrer em regiões codificadoras ou com função regulatória (MORIN; LUIKART; WAYNE, 2004; SCHUNTER et al., 2014).

Os SNPs são geralmente bi alélicos e representam o tipo mais frequente de polimorfismo genético. Além disso, o número em potencial de tais marcadores é muito superior ao de outros marcadores na maioria das espécies, possuindo distribuição homogênea pelo genoma. Vários métodos de genotipagem de SNP disponíveis, com baixa, média e alta densidade de marcadores, apresentam baixo custo e alta acurácia (CAETANO, 2009; SCHUNTER et al., 2014). Painéis de SNPs de baixa densidade incluem de centenas a alguns milhares de SNPs e são mais úteis quando há necessidade de genotipar muitas amostras, mesmo com limitação de custos e de tempo, pois eles podem fornecer informações essenciais sobre diversidade e estrutura genética populacional (HABIER; FERNANDO; DEKKERS, 2009).

A evolução dos métodos de sequenciamento permitiu a prospecção dos marcadores SNP com custos mais acessíveis e de forma acelerada, além de aumentar a quantidade de marcadores prospectados (CAETANO, 2009; SCHUNTER et al., 2014). Por conta de suas características, os SNPs se tornaram o marcador mais escolhido para estudos em evolução e genética da

conservação de diversas espécies (MORIN; LUIKART; WAYNE, 2004; SCHUNTER et al., 2014).

Um estudo de genética de populações do Pirarucu envolvendo genotipagem por sequenciamento ddRAD foi realizado por Torati et al. (2019), indicando uma forte estruturação entre as populações de Pirarucu amostradas. A abordagem de Torati et al, 2019, evidenciou a utilidade de marcadores SNPs para a espécie, a exemplo do que já ocorre para outras. Essa técnica mostrou-se eficiente para o proposto, porém, torna-se impeditiva para estudos com grandes números amostrais devido ao seu custo, alto volume de dados gerados e esforços gerados e esforços envolvidos nas análises.

A genotipagem de SNPs com as tecnologias fornecidas pela Fluidigm engloba métodos de alto padrão úteis para analisar variações genéticas em grandes populações. Essa tecnologia utiliza um sistema de circuitos fluídicos integrados (IFC – do inglês *Integrated Fluidics Circuits*) que permitem a análise simultânea de muitas amostras usando amplificação por PCR e detecção fluorescente. A tecnologia é flexível e possibilita a combinação de diferentes SNPs para formar conjuntos de SNPs customizados, minimizando custos operacionais, como no desenvolvimento de ensaios para determinação de haplótipos específicos para característica de interesse em espécies de interesse agrônomo (THOMSON, 2014).

Painéis com sistemas micro fluídicos IFC podem ser fabricados em diferentes formatos (48x48, 96x96, etc.), contendo matrizes com pequenas câmaras de reação, cada uma capaz de amplificar o DNA de uma única célula ou região genômica para testar, simultaneamente, diferentes números de amostras e ensaios de genotipagem de SNPs (Figura 3) (FLUIDIGM CORPORATION, 2018). Os produtos de amplificação são então analisados usando detecção baseada em fluorescência *SNPtype*, com diferentes fluoróforos sendo usados para distinguir os alelos no programa Fluidigm SNP Genotyping Analysis (FLUIDIGM CORPORATION, 2018).

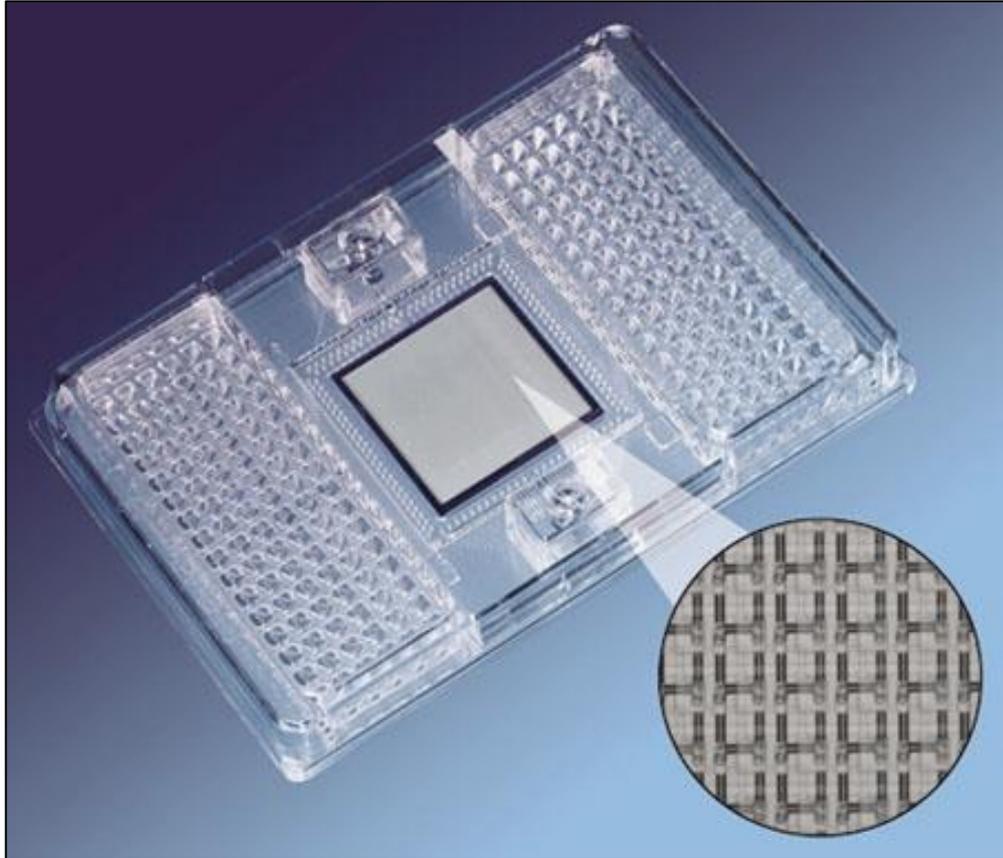


Figura 3. Exemplo de chip com Circuito Fluido Integrado (IFC): *Dynamic Array 96,96* da Fluidigm. Fonte: <https://www.bioprocessonline.com>.

Nos ensaios *SNPtype* da Fluidigm, os iniciadores de PCR Amplificador de Alvo Específico (*Specific Target Amplification - STA*) e *Primer Locus Específico (Locus-Specific Primer - LSP)* são utilizados em reações de pré-amplificação de cada amostra. Posteriormente, são realizadas reações de amplificação alelo-específicas com *Primers Alelo Específico (ASP1 e ASP2)* e *Primer Locus Específico (LSP)* e sondas com corantes fluorescentes (FAM e HEX), utilizando uma IFC (Figura 4) (THOMSON, 2014; FLUIDIGM CORPORATION, 2018). Após 35 ciclos de PCR a fluorescência se torna mais forte e o sinal pode ser captado pelo módulo EP1 da Fluidigm (Figura 4) (THOMSON, 2014; SMITH; MAUGHAN, 2015; FLUIDIGM CORPORATION, 2018).

O *software* Fluidigm SNP Genotyping Analysis é utilizado para analisar os sinais emitidos por fluorescência e gerar gráficos com os genótipos das amostras investigadas (Figura 5). A partir desta etapa, os dados obtidos pelo sistema podem ser exportados e explorados em programas computacionais de análises genéticas e moleculares, como o SNP & Variation Suite™ (SVS – Golden Helix) e outros.

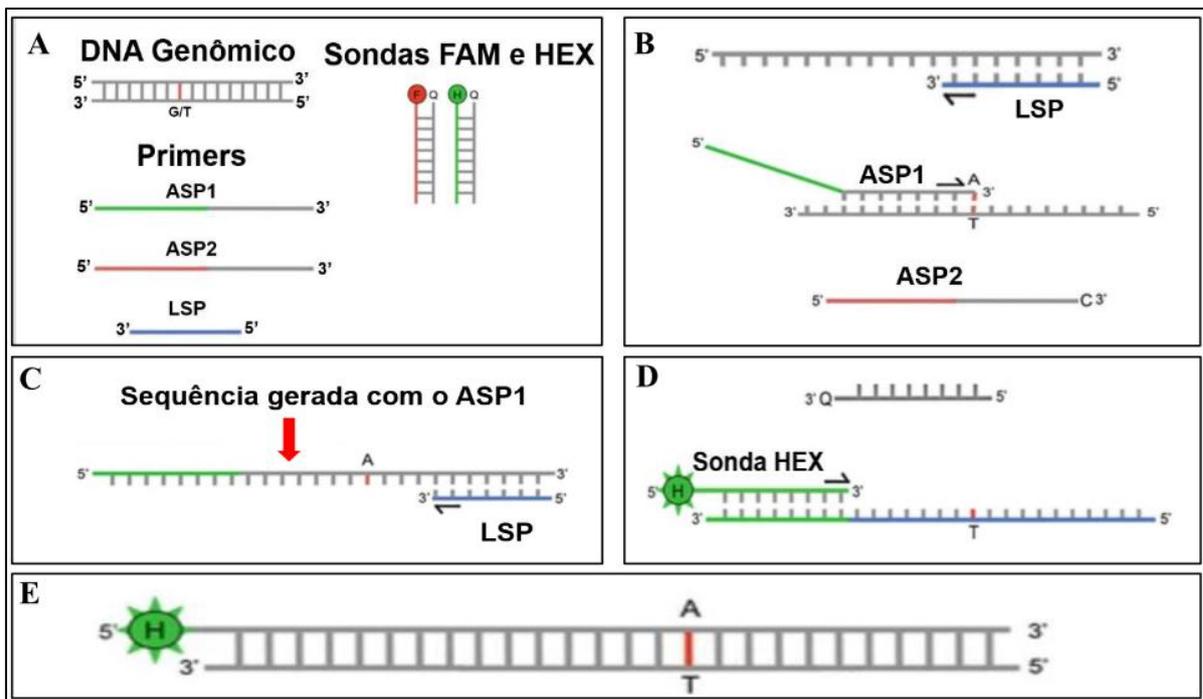


Figura 4. Etapas da genotipagem SNPTYPE na plataforma Fluidigm. A) Elementos necessários para a reação de PCR; B) Passo 1: Desnaturalização da fita de DNA e anelamento dos *primers* ASP e LSP por PCR; C) Passo 2: Sintetização da fita complementar ao *primer* ASP1 a partir do LSP; D) Passo 3: Reação de alinhamento das sondas HEX com a extremidade 5' do *primer* ASP1, ativando o fluoróforo; E) Produto da reação que emite fluorescência. Fonte: (SMITH; MAUGHAN, 2015) com adaptação feita por Araújo, 2022.

O uso de chips SNP de baixa densidade é uma boa estratégia para reduzir os custos de genotipagem em comparação com chips de alta densidade, tornando-os uma opção econômica para estudos que envolvam genotipagem em larga escala. Essa técnica, simplifica o gerenciamento e a análise de dados, pois gera um conjunto de dados menor, requerendo menos recursos computacionais (MAUGHAN et al., 2012). Além disso, os chips SNP de baixa densidade são aplicáveis a uma ampla gama de espécies animais, incluindo espécies não-modelo, que são pouco estudadas e, portanto, não possuem muita informação genética disponível publicada, sendo uma ótima opção de tecnologia disponível para investigação dos

índices de diversidade genética em populações de Pirarucu.

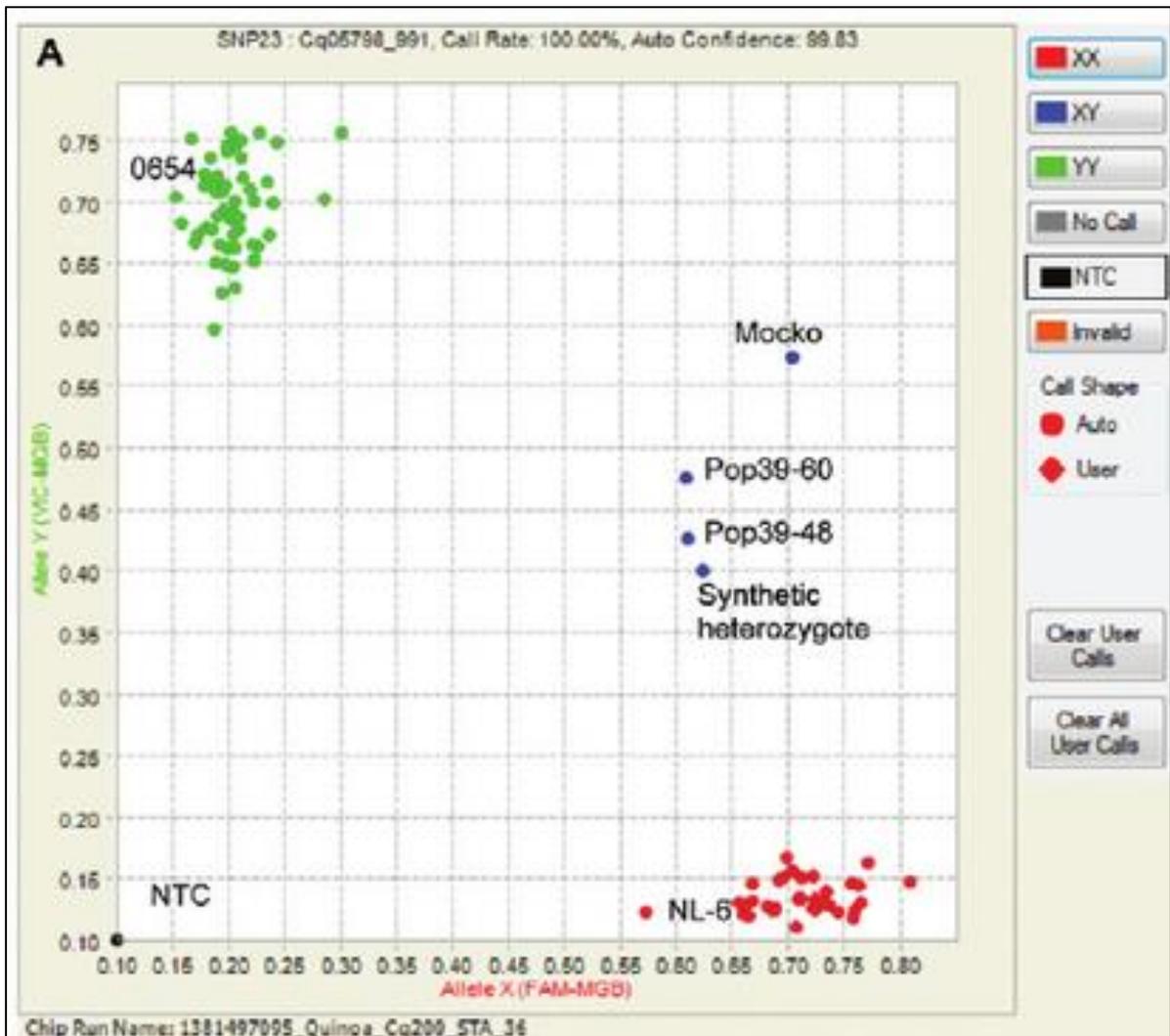


Figura 5. Exemplos de gráficos gerados com o programa Fluidigm SNP Genotyping Analysis. Os pontos representam amostras com diferentes genótipos identificados pela cor. Fonte: (MAUGHAN et al., 2012).

1.3.4 O cultivo do Pirarucu no Brasil

Segundo a FAO (2023), nos últimos dois anos, a aquicultura apresentou um crescimento superior ao da pesca de captura. Espera-se que essa tendência se intensifique na próxima década. Em 2020, a produção global de animais aquáticos em cativeiro alcançou 87,5 milhões de toneladas, um aumento de 6% em relação a 2018. Por outro lado, a produção da pesca de captura diminuiu para 90,3 milhões de toneladas, representando uma redução de 4,0% em relação à média dos três anos anteriores. Essa queda na produção pesqueira de captura foi influenciada principalmente pela pandemia de COVID-19, que teve um impacto significativo nas atividades de pesca, no acesso ao mercado e nas vendas (FAO, 2023).

Em 2021, o Brasil produziu cerca de 1.800 toneladas de Pirarucu, investindo mais de 28 milhões de reais nesse cultivo. Os maiores produtores estão na Região Norte onde o estado de Rondônia lidera a produção com mil toneladas cultivadas em 2018, seguido do Tocantins com 213 toneladas e do Pará com 206 toneladas (IBGE, 2023). A exportação de Pirarucu, ainda que mínima atualmente, é promissora pois, esse pescado possui grande aceitação no mercado por ter uma carne considerada de excelente qualidade para consumo (SOUSA, et al., 2017).

O Pirarucu apresenta ótimas características para o seu cultivo em cativeiro, como o rápido crescimento juvenil, podendo atingir 10 kg de peso vivo nos primeiros 12 meses de vida, sua rusticidade superior em ambientes tropicais (NÚÑEZ-RODRÍGUEZ et al., 2018) e respiração aérea, que possibilita a adaptação a ambientes com baixa disponibilidade de oxigênio dissolvido e altas concentrações de amônia (CAVERO et al., 2004), condições comuns dos tanques de criação.

Segundo Imbiriba (2001), os primeiros estudos realizados para desenvolver o cultivo de Pirarucu no Brasil foram conduzidos por Oliveira (1944) no Pará, e Fontenele (1948) no Ceará, onde ambos conseguiram reproduzir a espécie em cativeiro. Com o passar do tempo, a piscicultura dessa espécie se intensificou, principalmente no norte e nordeste do país. Em meados dos anos 80, a Embrapa Amazônica Oriental, em Belém, iniciou um programa de cultivo de Pirarucu. Os peixes foram cultivados em tanques com uma área inundada de 100 metros quadrados, abastecido com água proveniente de um açude para banho de bubalinos (IMBIRIBA, 2001). Esses estudos pioneiros abriram caminho para o desenvolvimento da criação em larga escala do Pirarucu em sistemas de cultivo controlados. Desde então, a piscicultura do Pirarucu tem se expandido e ganhado importância no país, com a utilização de técnicas e estratégias para o manejo sustentável dessa espécie (IMBIRIBA, 2001; SOUZA, et al., 2017).

Em 2017, a Embrapa Amazônia Oriental, publicou um guia da piscicultura do Pirarucu, onde é possível encontrar informações sobre os sistemas de produção ideais para a espécie: tanques de manutenção dos reprodutores, podendo estes serem de rede, escavados ou de lona; laboratórios para o treinamento alimentar de alevinos e viveiros de recria; viveiros de engorda. O manual também indica a qualidade da água, o tamanho correto dos viveiros, alimentação ideal e reprodução, que é muito delicada e envolve uma combinação de todas as condições ideais de cultivo para ter chances de ocorrer (SOUZA et al., 2017).



Figura 6. Tanque para engorda de Pirarucu. Fonte: <https://panoramadaaquicultura.com.br/a-engorda-de-Pirarucu-em-alta-densidade/>.

Um dos grandes empecilhos a produtividade do Pirarucu em cativeiro é sua reprodução. O empasso começa na identificação do sexo dos reprodutores, que é praticamente impossível por meio de características externas até a maturação sexual, quando o macho pode adquirir uma coloração avermelhada na cauda. Até recentemente, não era possível determinar o sexo por meio de marcadores moleculares e outras técnicas mais sofisticadas, como análises bioquímicas e endoscopia, são muito dispendiosas para o produtor (NÚÑEZ, et al., 2011; SOUZA et al., 2017). Além disso, o ovário da fêmea possui um único lóbulo ou gimnovário que torna a prática da canulação (procedimento para determinar o desenvolvimento e maturação do ovócito, também usado na fertilização artificial) muito complexa (NÚÑEZ, et al., 2011). Outros empecilhos são as disfunções reprodutivas da espécie, como a falha na fase final da maturação oocitária e ovulação e falha na sincronização de fertilização macho-fêmea, pois o espermatozoide não é ejaculado a tempo de encontrar com o óvulo na desova, inibindo a fertilização (NÚÑEZ et al., 2011).

Um exemplo de manejo adequado, proposto do Hrbek e colaboradores (2005), propõem o uso de um grande grupo de matrizes de reprodutores de Pirarucu nas propriedades que cultivam a espécie. Ao criar um plantel de reprodutores com animais de diferentes populações,

independentemente de sua região geográfica, o produtor minimiza os efeitos da perda de diversidade genética devido ao efeito do fundador e à depressão endogâmica, fenômeno que ocorre por conta do acasalamento entre indivíduos geneticamente próximos (HRBEK et al., 2005).

Além das questões relacionadas à reprodução, o cultivo pode ser impactado pela baixa diversidade genética dos estoques de reprodutores. As práticas de manejo utilizadas pelos produtores na aquicultura influenciam diretamente em sua produtividade e, o uso de um pequeno número de peixes reprodutores sem rastreabilidade de origem nos tanques de criação pode levar ao declínio da diversidade genética dos estoques (CARLOS; VERA, 2014). Na aquicultura, diversos fatores como variações ambientais bruscas, alimentação inadequada, desnutrição, elevada densidade de cultivo, transporte, traumas mecânicos, tratamentos com drogas, variações repentinas e abruptas de temperatura e pH, má qualidade da água e/ou poluição do ambiente aquático podem afetar a saúde dos peixes (BANDEIRA JUNIOR et al., 2017). Considerando essas adversidades, manter a variabilidade genética nos estoques é essencial para garantir maior resistência as diferentes condições encontradas nos sistemas de criação (CARLOS; VERA, 2014).

Outro cuidado que deve ser levado em consideração ao montar o estoque de reprodutores de Pirarucu é a origem desses animais. Devido ao declínio nas populações selvagens de Pirarucu nos anos 80 e à dificuldade de recuperação dos estoques naturais desse peixe, a legislação atual ainda proíbe a pesca dessa espécie em alguns períodos do ano e, quando liberada, esta deve ser realizada somente em reservas legais destinadas a este fim (VITORINO et al., 2017). Se os produtores não seguirem a legislação e coletarem ou receberam animais selvagens para criação em cativeiro irregularmente, estão sujeitos à diversas medidas punitivas aplicadas pelos órgãos de fiscalização do meio ambiente.

Entender a estrutura genética das populações de Pirarucu em cativeiro pode ajudar a elucidar problemas relacionados à baixa variabilidade genética nos estoques. Rastrear a origem dos animais contidos nos tanques de criação e realizar o manejo reprodutivo assistido evitam o aumento de problemas causados por endogamia e ajudam o produtor a selecionar melhor suas matrizes (CARLOS; VERA, 2014). A análise detalhada da genética dessas populações pode inclusive identificar animais selvagens provenientes de zonas onde a pesca e coleta são proibidos por lei e evitar maiores prejuízos ao produtor e ao meio ambiente. Diante do exposto, observa-se que é essencial superar as dificuldades remanescentes no cultivo de Pirarucu, para potencializar a piscicultura da espécie e, conseqüentemente, ampliar sua comercialização sem

causar mais danos as populações selvagens da espécie, contribuindo assim, com sua conservação.

A compreensão da estrutura genética das populações de é fundamental para uma gestão eficaz e sustentável desses recursos naturais. O Pirarucu, como uma espécie de importância econômica, cultural e ecológica na Amazônia, enfrenta crescentes pressões de pesca e degradação do habitat. O estudo aqui proposto, traz detalhes da estrutura genética das populações de Pirarucu estudadas, a partir da técnica de genotipagem de SNPs com excelente da Fluidigm, até então não utilizada para estudos com o Pirarucu, e pôde fornecer informações valiosas sobre a conectividade entre os diferentes estoques e populações, padrões de dispersão e fluxo gênico, além de auxiliar na gestão para a conservação da diversidade genética e o manejo sustentável dessas populações.

REFERENCIAS BIBLIOGRÁFICAS

- AQUINO, S.; LATRUBESSE, E. M.; SOUZA FILHO, E. E. DE. Caracterização hidrológica e geomorfológica dos afluentes da Bacia do Rio Araguaia. **Revista Brasileira de Geomorfologia**, v. 10, n. 1, p. 43–54, 2009.
- ARARIPE, J. et al. Dispersal Capacity and Genetic Structure of Arapaima gigas on Different Geographic Scales Using Microsatellite Markers. **PLoS ONE**, v. 8, n. 1, 29 jan. 2013.
- ARAUJO, N.T. et al. Population Structure and Genetic Diversity of Rainbow Trout (*Oncorhynchus Mykiss*) Broodstocks From Brazil Using SNP Markers. **Aquaculture Reports**. v. 31. Ago. 2023.
- BANDEIRA JUNIOR, G., et al. Potential uses of *Ocimum gratissimum* and *Hesperozygis ringens* essential oils in aquaculture. **Industrial Crops and Products** [online]. vol. 97, pp. 484-491. 2017.
- BOHLING, J.; HAFFRAY, P.; BERREBI, P. Genetic diversity and population structure of domestic brown trout (*Salmo trutta*) in France. **Aquaculture**, v. 462, p. 1–9, 1 set. 2016.
- BRAUNER, C. J. et al. Transition in organ function during the evolution of air-breathing; insights from *Arapaima gigas*, an obligate air-breathing teleost from the Amazon. **Journal of Experimental Biology**, v. 207, n. 9, p. 1433–1438, 2004.
- CAETANO, A.R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**. v. 38, n.1, p. 64-71. 2009.
- CAMPOS-SILVA, J.V; PERES, C.A. Community-based management induces rapid recovery of a high-value tropical freshwater fishery. **Scientific Reports**, 6(34745), 1-13. 2016.
- CARLOS, H. DOS A. DOS S.; VERA, M. F. DE A. E V. Genetic relationships between captive and wild subpopulations of *Arapaima gigas* (Schinz, in Cuvier, 1822). **International Journal of Fisheries and Aquaculture**, v. 6, n. 10, p. 108–123, 29 out. 2014
- CASTELLO, L. Lateral migration of *Arapaima gigas* in floodplains of the Amazon. **Ecology of Freshwater Fish**, v. 17, n. 1, p. 38–46, 2008.
- CASTELLO, L.; STEWART, D. J. Assessing CITES non-detriment findings procedures for *Arapaima* in Brazil. **Journal of Applied Ichthyology**, v. 26, n. 1, p. 49–56, fev. 2010.
- CASTELLO, L.; STEWART, D. J. ARANTES, C.C. Modeling Population Dynamics and Conservation of *Arapaima* in The Amazon. **Reviews in Fish Biology and Fisheries**, 21(3): 623-640. 2011.
- CATÂNEO, D. T. B. DOS S. et al. Elucidating a history of invasion: population genetics of Pirarucu (*Arapaima gigas*, Actinopterygii, Arapaimidae) in the Madeira River. **Hydrobiologia**, v. 849, n. 16, p. 3617–3632, 2022.
- CAVERO, B.A.S. et al. Tolerance of Pirarucu juveniles to increasing ammonia concentration in a closed environment. **Pesquisa Agropecuária Brasileira**, v. 39, n. 5, p. 513-516, 2004.
- DALLOZ, M. F. Estrutura populacional de um pequeno mamífero em uma área de mata contínua e em três fragmentos de Mata Atlântica: um teste do modelo fonte-sumidouro. p. 115, 2013.
- DE OLIVEIRA, E. G. et al. Effects of stocking density on the performance of juvenile Pirarucu (*Arapaima gigas*) in cages. **Aquaculture**, v. 370–371, p. 96–101, 11 dez. 2012.
- DORIA, C. R. C. et al. Is there a future for artisanal fishing in the Amazon? **Management of Biological Invasions**, v. In press, n. 1, p. 1–8, 2019.
- DU, K. et al. The genome of the arapaima (*Arapaima gigas*) provides insights into gigantism, fast growth and chromosomal sex determination system. **Scientific Reports**, v. 9, n. 1, 1 dez. 2019.
- EMBRAPA. **Embrapa Pesca e aquicultura**. Palmas, TO. Disponível em: <<https://www.embrapa.br/tema-pesca-e-aquicultura/>>. Acesso em: Jan. 2024.

- FAO – FOOD AND AGRICULTURE ORGANIZATION. **Relatório da FAO sobre o Estado Mundial da Pesca e Aquicultura (SOFIA)**. 2023. Disponível em: <<https://www.fao.org/brasil/noticias/detail-events/es/c/1585153/>> Acesso em: Jan .2024.
- FARIAS, I. P. et al. The largest fish in the world's biggest river: Genetic connectivity and conservation of *Arapaima gigas* in the Amazon and Araguaia-Tocantins drainages. **PLoS ONE**, v. 14, n. 8, 1 ago. 2019.
- FAZZI-GOMES, P. et al. Genomic approach for conservation and the sustainable management of endangered species of the Amazon. **PLoS ONE**, v. 16, n. 2 February 2021, 2021.
- FERNANDES, C. C. Lateral migration of fishes in Amazon floodplains. **Ecology of Freshwater Fish**, v. 6, n. 1, p. 36–44, 1997.
- FERREIRA, G.; MARCOVITCH, J.; VAL, A.L. A systematic review of the production chain of the *Arapaima gigas*, the giant fish of the Amazon. **Management of Environmental Quality: An International Journal**. v. 31(2), p. 349-363, 3 jan. 2020.
- FLUIDIGM CORPORATION. **SNP Genotyping User Guide** (PN 68000098 Q1). 2018. Disponível em: <<https://fluidigm.my.salesforce.com/sfc/p/#700000009DAw/a/4u0000019kFV/J5w6q0CKYdITv5ebT5PYx8Ai2T5oGz.CIYtyMxHAWFg>>. Acesso em: 19/01/2024
- GARCIA DE ALMEIDA, I. et al. Estrutura Genética do Pirarucu (*Arapaima gigas*) na Região de Santarém, PA, Brasil. **Anais da 49ª Reunião Anual da Sociedade Brasileira de Zootecnia**. Brasília, DF. 2012.
- GODINHO, H. P. et al. Gonadal morphology and reproductive traits of the Amazonian fish *Arapaima gigas* (Schinz, 1822). **Acta Zoologica**, v. 86, n. 4, p. 289–294, 2005.
- HABIER, D.; FERNANDO, R. L.; DEKKERS, J. C. M. Genomic selection using low-density marker panels. **Genetics**, v. 182, n. 1, p. 343–353, Maio 2009.
- HILTON, E. J.; LAVOUÉ, S. A review of the systematic biology of fossil and living bony-tongue fishes, Osteoglossomorpha (Actinopterygii: Teleostei). **Neotropical Ichthyology**, v. 16, n. 3, 11 out. 2018.
- HRBEK, T. et al. Population genetic analysis of *Arapaima gigas*, one of the largest freshwater fishes of the Amazon basin: Implications for its conservation. **Animal Conservation**, v. 8, n. 3, p. 297–308, 2005.
- HRBEK, T.; CROSSA, M.; FARIAS, I. P. Conservation strategies for *Arapaima gigas* (Schinz, 1822) and the Amazonian várzea ecosystem. **Brazilian Journal of Biology**, v. 67, n. 4 SUPPL., p. 909–917, 2007.
- HRBEK, T.; FARIAS, I. P. The complete mitochondrial genome of the Pirarucu (*Arapaima gigas*, Arapaimidae, Osteoglossiformes). **Genetics and Molecular Biology**, v. 31, n. 1 SUPPL. 1, p. 293–302, 2008.
- IANELLA, P. et al. SNP Discovery in the South American Freshwater Fish Tambaqui (*Colossoma macropomum*) by Deep Sequencing of Reduced Representation Library. In: **Plant and Animal Genome Conference XXIII**. San Diego, CA, p. 9-13, Anais. 2016.
- IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa pecuária municipal**. São Paulo: IBGE, 2023. Disponível em: <<http://www.ibge.gov.br/estatistica/>>. Acesso em: Jan. 2023.
- ILHA, P.; ROSSO, S.; SCHIESARI, L. Effects of deforestation on headwater stream fish assemblages in the Upper Xingu River Basin, Southeastern Amazonia. **Neotropical Ichthyology**, v. 17, n. 1, 31 jan. 2019.
- IMBIRIBA, P. POTENCIAL DE CRIAÇÃO DE PIRARUCU, *Arapaima gigas*. **Acta amazonica**. 31(2): 299-316. 2001.
- JOSHI, R. et al. Development and validation of 58K SNP-array and high-density linkage map in Nile tilapia (*O. niloticus*). **Frontiers in Genetics**, v. 9, n. OCT, 15 out. 2018.
- JUNK, W. J. Water Rivers. p. 8–9, 1984.

- KIRSTEN, I.F. et al. The Pirarucu (*Arapaima* sp) fishery at the Araguaia River basin, state of Mato Grosso – Brazil. **Bol. Inst. Pesca** (Impr.). 38(2): 131-144. 2012.
- LINCK, E.; BATTEY, C. J. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. **Molecular Ecology Resources**, v. 19, n. 3, p. 639–647, 1 maio 2019.
- LÓPEZ-LANDAVERY, E. A. et al. Non-invasive sex genotyping of paiche *Arapaima gigas* by qPCR: An applied bioinformatic approach for identifying sex differences. **Aquaculture**, v. 546, n. February 2021, 2022.
- MAUGHAN, P. J. et al. Single Nucleotide Polymorphism Identification, Characterization, and Linkage Mapping in Quinoa. **The Plant Genome**, v. 5, n. 3, nov. 2012.
- MEANTE, R. E. X.; DÓRIA, C. R. DA C. Caracterização Da Cadeia Produtiva Da Piscicultura No Estado De Rondônia: Desenvolvimento E Fatores Limitantes. **Revista de Administração e Negócios da Amazônia**, v. 9, n. 4, p. 164, 31 jan. 2018.
- MENDES FILHO, O. DA R. et al. Acordos De Pesca No Médio Rio Araguaia, Tocantins, Amazônia, Brasil: a Experiência Do Projeto Pescar. **Engenharia de Pesca: aspectos teóricos e práticos - Volume 2**, p. 13–31, 2021.
- MOREIRA DA SILVA, A.; LUIZ, W.; DUNCAN, P. Aspectos biológicos, ecologia e fisiologia do Pirarucu (*Arapaima gigas*): uma revisão da literatura. **Scientia Amazonia**, v. [s.l: s.n.].
- MORIN, P.A.; LUIKART, G.; WAYNE, R. K. **SNPs in ecology, evolution and conservation. Trends in Ecology and Evolution**, abr. 2004.
- NELSON, J.S.; GRANDE, T.C.; WILSON M.V.H. **Fishes of the world**. 5. ed. Nova Jersey: Wiley, 2016.
- NIU, S. et al. Genetic diversity, linkage disequilibrium, and population structure analysis of the tea plant (*Camellia sinensis*) from an origin center, Guizhou plateau, using genome-wide SNPs developed by genotyping-by-sequencing. **BMC Plant Biology**, v. 19, n. 1, p. 1–12, 2019.
- NOGUEIRA, F. et al. Genetic diversity and structuring in the arapaima (osteoglossiformes, osteoglossidae) population reveal differences between the amazon and the tocantins-araguaia basins. **Anais da Academia Brasileira de Ciências**, v. 92, n. 1, p. 1–16, 2020.
- NÚÑEZ, J. et al. Reproductive success and fry production of the paiche or Pirarucu, *Arapaima gigas* (Schinz), in the region of Iquitos, Perú. **Aquaculture Research**, v. 42, n. 6, p. 815–822, 2011.
- PEDROZA, M.X.F. et al., O mercado de peixes da piscicultura no Brasil: estudo do segmento de supermercados. **Boletim de Pesquisa e Desenvolvimento**. Embrapa Pesca e Aquicultura. v.1. 38p. 2020.
- PICKLER, E.; FILHO, V.; EUSTÁQUIO RIBEIRO, J. **Evolução da piscicultura no Brasil: Diagnóstico e desenvolvimento da cadeia produtiva de tilápia**. Texto para discussão. Disponível em: <<http://hdl.handle.net/10419/177544>> Acesso em: Jan 2024.
- SAINT-PÉ, K. et al. Development of a large SNPs resource and a low-density SNP array for brown trout (*Salmo trutta*) population genetics. **BMC Genomics**, v. 20, n. 1, p. 1–13, 2019.
- SANTOS, C.H.A. et al. Genetic relationships between captive and wild subpopulations of *Arapaima gigas* (Schinz, in Cuvier, 1822). **International Journal of Fisheries and Aquaculture**, v. 6, n. 10, p. 108-123, 29 out. 2014.
- SCHULTER, E. P.; VIEIRA FILHO, J. E. R. Evolução da Piscicultura no Brasil: diagnóstico e desenvolvimento da cadeia produtiva de tilápia. Texto para Discussão. **Instituto de Pesquisa Econômica Aplicada – Ipea**, 2017. Disponível em: <http://repositorio.ipea.gov.br/bitstream/11058/8043/1/td_2328.pdf>. Acesso em: 19 jan. 2023.
- SCHUNTER, C. et al. SNP development from RNA-seq data in a nonmodel fish: How many individuals are needed for accurate allele frequency prediction? **Molecular Ecology Resources**, v. 14, n. 1, p. 157–165, jan. 2014.

- SMITH, S. M.; MAUGHAN, P. J. SNP genotyping using KASPar assays. **Methods in Molecular Biology**, v. 1245, p. 243–256, 2015.
- SNP Genotyping**. Disponível em: <www.fluidigm.com> Acesso em Jan. 2024.
- SOUSA, A. R. B. et al. **Piscicultura de Pirarucu**. Coleção Criar. Embrapa, 2017. V.1 p. 169.
- STEWART, D. J. A new species of arapaima (Osteoglossomorpha: Osteoglossidae) from the solimões river, Amazonas state, Brazil. **Copeia**, n. 3, p. 470–476, 2013.
- TEIXEIRA DE ARAÚJO JÚNIOR, N. et al. Population structure and genetic diversity of rainbow trout (*Oncorhynchus mykiss*) broodstocks from Brazil using SNP markers. **Aquaculture Reports**, v. 31, 1 ago. 2023.
- THOMSON, M. J. High-Throughput SNP Genotyping to Accelerate Crop Improvement. **Plant Breeding and Biotechnology**, v. 2, n. 3, p. 195–212, 30 set. 2014.
- TORATI, L. S. et al. Genetic diversity and structure in Arapaima gigas populations from Amazon and Araguaia-Tocantins River basins. **BMC Genetics**, v. 20, n. 1, 28 jan. 2019.
- VENTURIERI, R.; BERNARDINO, G. Hormônios na reprodução artificial de peixes. **Panorama da Aquicultura**, v. 9, n. 55, p. 39-48, 1999.
- VIANA, J. P. et al. Manejo comunitário do Pirarucu Arapaima gigas na Reserva de Desenvolvimento Sustentável Mamirauá - Amazonas, Brasil. Áreas aquáticas protegidas como instrumento de gestão pesqueira. **Série Áreas Protegidas do Brasil. Brasília: Ministério do Meio Ambiente e IBAMA**, pp. 239-261. 2007.
- VITORINO, C. A. et al. Genetic diversity of arapaima gigas (Schinz, 1822) (Osteoglossiformes: Arapaimidae) in the araguaia-tocantins basin estimated by ISSR marker. **Neotropical Ichthyology**, v. 13, n. 3, p. 557–568, 2015.
- VITORINO, C. A. et al. Low genetic diversity and structuring of the arapaima (Osteoglossiformes, Arapaimidae) population of the Araguaia-Tocantins basin. **Frontiers in Genetics**, v. 8, n. OCT, 24 out. 2017.
- VRIJENHOEK, R. C. Conservation genetics of freshwater fish. **Journal of Fish Biology**, v. 53, n. SUPPL. A, p. 394–412, 1998.
- WASKO, A.P.; MARTINS, C.; OLIVEIRA, C.; FORESTI, F. Genetic conservation of Brazilian fishes - Present state and perspectives. **ARBS – Annual Review of Biomedical Sciences**, v. 6, p. 79-90, 2004.
- ZHANG, H. Y. et al. Population genetic analysis of aquaculture salmonid populations in China using a 57K rainbow trout SNP array. **PLoS ONE**, v. 13, n. 8, p. 1–12, 2019.
- ZHANG, S. M.; WANG, D. Q.; ZHANG, Y. P. Mitochondrial DNA variation, effective female population size and population history of the endangered Chinese sturgeon, *Acipenser sinensis*. **Conservation Genetics**, v. 4, n. 6, p. 673–683, 2003.

CHAPTER 2
**GENETIC STRUCTURE OF PIRARUCU (*Arapaima gigas*) POPULATIONS IN THE
AMAZONAS AND TOCANTINS-ARAGUAIA RIVER BASINS**

2.1 RESUMO

ESTRUTURA GENÉTICA DE POPULAÇÕES DE PIRARUCU (*Arapaima gigas*) NAS BACIAS DOS RIOS AMAZONAS E TOCANTINS-ARAGUAIA

Aline Araújo Campelo¹, Noeliton Teixeira Araújo¹, Káifer José Soares Silva¹, Michel Eduardo Beleza Yamagishi², Leandro Carrijo Cintra², Carla de Andrade Vitorino³, Paulo Cesar Venere³, Alexandre Pucci Hercos⁴, João Vitor Campos-Silva⁵, Patrícia Ianella⁶, Alexandre Rodrigues Caetano⁶

1 PPG Ciências Animais/UnB – Brasília/DF

2 Embrapa Agricultura Digital – Campinas/SP

3 Universidade Federal de Mato Grosso – Cuiabá/MT

4 Instituto de Desenvolvimento Sustentável Mamirauá – Tefé/AM

5 Instituto Juruá – Manaus/AM

6 Embrapa Recursos Genéticos e Biotecnologia – Brasília/DF

O *Arapaima gigas* desempenha um papel vital como recurso alimentar e cultural para indígenas e ribeirinhos na região amazônica. No entanto, a espécie enfrenta sérias ameaças devido à pesca excessiva, perda de habitat e poluição. Estudos genéticos têm sido conduzidos para avaliar a diversidade genética e o estado de conservação dessas populações, revelando preocupante baixa diversidade genética, especialmente na bacia do Tocantins-Araguaia. O cultivo do *Arapaima gigas* pode auxiliar na conservação dos estoques selvagens, porém, este ainda enfrenta diversos desafios relacionados ao manejo inadequado de reprodutores. Superar esses empecilhos é essencial para potencializar a produtividade em cativeiro. O objetivo deste estudo foi analisar a estrutura genética de populações de *Arapaima gigas* selvagens e de cativeiro nas bacias Amazônica e Tocantins-Araguaia, usando um painel de SNPs de baixa densidade. Foram analisadas 360 amostras sendo, 166 da bacia Tocantins-Araguaia e 188 da bacia Amazônica, além de 6 amostras de origem desconhecida, coletadas em um mercado local em Manaus-AM. Do total de amostras, 44 eram provenientes de cativeiro. As análises abordaram diversidade, distância genética, estrutura populacional e parentesco, destacando a maior diversidade geral nas populações da bacia Amazônica. Médias de heterozigosidade observada e esperada indicaram diferenças significativas na diversidade das duas bacias, com as populações do Tocantins-Araguaia (0.050 to 0.204) apresentando valores inferiores em comparação com as da

bacia Amazônica (0.291 to 0.340). O F_{st} total foi de 0.424, indicando que há diferenciação genética entre as populações. A análise de Structure separou as populações em dois principais clusters com base na bacia de origem enquanto a análise de PCA identificou quatro clusters (A - Tocantins-Araguaia, B – Tocantins-Araguaia no sudoeste do Pará, C – Santarém, PA e D – Amazonas, Rondônia e Acre). As análises ainda revelaram potenciais zonas de mistura entre as populações das bacias, especialmente no sudeste paraense (Populações 16 PBP e 17 ITU) e distinções de estruturação nas populações de Santarém-PA. Correlações significativas entre distâncias genéticas e geográficas sugerem a contribuição do isolamento por distância na formação da estrutura genética dessas populações (Teste de Mantel, $r=0.582$). A utilização eficaz de painéis de SNPs de baixa densidade forneceu dados úteis para subsidiar medidas de manejo e conservação da espécie, auxiliando o desenvolvimento de ferramentas genômicas que visam atender à demanda dos piscicultores sem causar maiores prejuízos as populações naturais do *Arapaima gigas*.

PALAVRAS-CHAVE: Conservação, Genotipagem, Piscicultura, SNPs.

2.2 ABSTRACT

GENETIC STRUCTURE OF PIRARUCU (*Arapaima gigas*) POPULATIONS IN THE AMAZONAS AND TOCANTINS-ARAGUAIA RIVERS BASINS

Aline Araújo Campelo¹, Noeliton Teixeira Araújo¹, Káifer José Soares Silva¹, Michel Eduardo Beleza Yamagishi², Leandro Carrijo Cintra², Carla de Andrade Vitorino³, Paulo Cesar Venere³, Alexandre Pucci Hercos⁴, João Vitor Campos-Silva⁵, Patrícia Ianella⁶, Alexandre Rodrigues Caetano⁶

1 PPG Ciências Animais/UnB – Brasília/DF

2 Embrapa Agricultura Digital – Campinas/SP

3 Universidade Federal de Mato Grosso – Cuiabá/MT

4 Instituto de Desenvolvimento Sustentável Mamirauá – Tefé/AM

5 Instituto Juruá – Manaus/AM

6 Embrapa Recursos Genéticos e Biotecnologia – Brasília/DF

The *Arapaima gigas* plays a vital role as a food and cultural resource for indigenous people and riverside dwellers in the Amazon region. However, the species faces serious threats due to overfishing, habitat loss and pollution. Genetic studies have been conducted to assess the genetic diversity and conservation status of these populations, revealing a worryingly low genetic diversity, especially in the Tocantins-Araguaia basin. *Arapaima gigas* cultivation can help conserve wild stocks, however, it still faces several challenges related to inadequate management of breeding stock. Overcoming these obstacles is essential to enhancing productivity in captivity. The aim of this study was to analyze the genetic structure of wild and captive *Arapaima gigas* populations in the Amazon and Tocantins-Araguaia basins, using a low-density SNP panel. A total of 360 samples were analyzed, with 166 from the Tocantins-Araguaia basin and 188 from the Amazon basin, along with 6 samples of unknown origin collected from a local market in Manaus-AM. Out of the total samples, 44 were from captive sources. Comprehensive analyzes addressed genetic diversity and distance, population structure and kinship, highlighting the greater overall diversity in Amazon basin populations. Average observed and expected heterozygosity indicated significant differences in the diversity of the two basins, with the populations of Tocantins-Araguaia (0.050 to 0.204) presenting lower values compared to those of the Amazon basin (0.291 to 0.340). The total F_{st} was 0.424,

indicating that there is genetic differentiation between the populations. The Structure analysis separated the populations into two main clusters based on the basin of origin while the PCA analysis identified four clusters (A - Tocantins-Araguaia, B – Tocantins-Araguaia in southwestern Pará, C – Santarém, PA and D – Amazonas, Rondônia and Acre). The analyzes also revealed potential admixture zones between populations, especially in the southeast of Pará (Populations 16 PBP and 17 ITU) and structuring distinctions in the populations of Santarém-PA. Significant correlations between genetic and geographic distances suggest the contribution of isolation by distance in the formation of the genetic structure of these populations (Mantel test, $r=0.582$). The effective use of low-density SNP panels provided useful data to support management and conservation measures for the species, helping the development of genomic tools that aim to meet the demand of fish farmers without causing greater damage to the natural populations of *Arapaima gigas*.

KEYWORDS: Conservation, Fish-farming, Genotyping, SNPs.

2.3 INTRODUCTION

The Arapaima or Pirarucu (*Arapaima gigas*, Schinz, 1822) is one of the largest freshwater fish species in the world – adults frequently reach up to 2.5 m in length and weigh over 200 kg. The species primarily inhabits calm water bodies such as floodplains and lakes and it is native to the Amazon and Tocantins-Araguaia River basins (CASTELLO, 2008b).

Due to its historical significance as a valuable fishery resource, the Arapaima has been subjected to extensive exploitation. Recorded fishery landings in major ports across the Amazon region show a major decline in wild catches by the late 1970's (HRBEK; CROSSA; FARIAS, 2007). As a result, the *Arapaima gigas* was listed in the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) in 1975 indicating its threatened status and the need for regulation of fishing and trade. In addition, the International Union for Conservation of Nature (IUCN) has categorized the species as data deficient, therefore providing evidence of the existing difficulties to reliably assess the true risks of its extinction (VITORINO et al., 2017; NOGUEIRA et al., 2020).

Previous studies carried out with *Arapaima gigas* populations from the Amazon and Tocantins-Araguaia basins show high levels of genetic structure among populations, pointing to a lower level of diversity in the Tocantins-Araguaia basin (HRBEK et al., 2005; VITORINO et al., 2017; TORATI et al., 2019;). Most studies have focused on characterizing the genetic diversity of populations in the Amazon Basin, while little is known about the real genetic diversity of populations along the Tocantins-Araguaia Basin (VITORINO et al., 2017; NOGUEIRA et al., 2020), which run across the Cerrado, a large biome also known as the Brazilian savanna.

Studies to investigate the genetic diversity of wild *Arapaima gigas* populations are essential to expand the available knowledge base for the species, and to provide reliable information for the establishment of efficient management and conservation policies and activities. Most of the published research has used *Inter Simple Sequence Repeat* (ISSR), mtDNA and microsatellite data to analyze the genetic diversity and structure of *Arapaima gigas* populations (ARARIPE et al., 2013; HRBEK et al., 2005; VITORINO et al., 2017).

The aim of this study is to analyze the genetic structure of wild and captive *Arapaima gigas* populations using a low-density SNP panel, to reveal useful information for subsidizing the elaboration of strategies for the use and conservation of germplasm of the species.

2.4 MATERIAL AND METHODS

2.4.1 Sampling

A total of 360 samples were collected at sites along the Amazon and Tocantins-Araguaia basins (Figure 2.1). A total of 166 samples came from Araguaia River basin (VITORINO et al., 2015) and 188 came from Amazon River basin, most of these samples were collected after slaughter, during fishing carried out by riverside dwellers. Specifically, the wild MLA and SLA populations were collected during an animal rescue operation in two small lakes formed during the dry season in the region of Araguaiana, MT. A total of 46 samples was derived from captive animals, 14 from a fish farm in Parauapebas, PA (PBP) and 30 from a fish farm in Rondônia, RO (POR, BOL, PUR, PER), while six samples of unknown origin were collected in a fish market in the city of Manaus (Table 2.1). Fin and tissue samples were originally fixed in ethanol 95% until further processing. All necessary licenses to access and use this material was obtained from CGEN before the work was carried out.

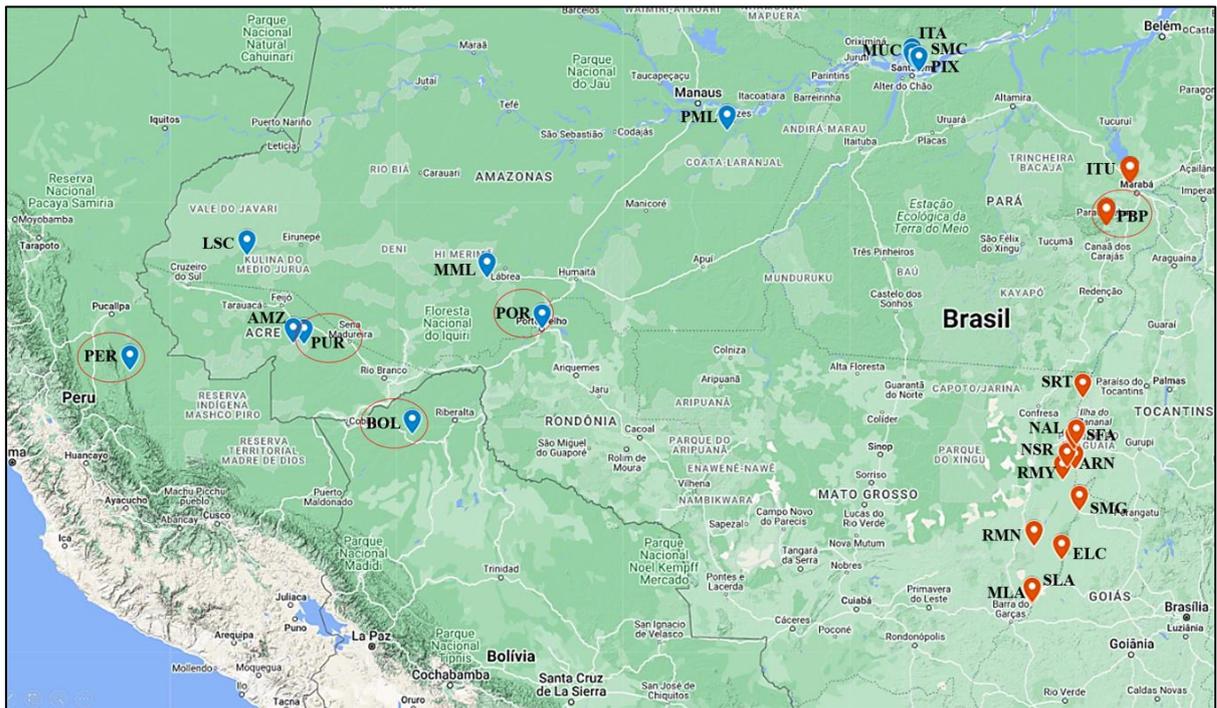


Figure 2.1 – Sites where *Arapaima gigas* samples were collected. Sites indicated with orange markers are from the Tocantins-Araguaia Basin and sites indicated with blue markers are from the Amazon Basin (Table 2.1 for full description). Locations with samples from captivity are marked with a red circle. Source: author-customized map on Google My Maps.

Table 2.1 – Sampling information of all *Arapaima gigas* samples tested.

Researcher/Farmer	Hydrographic basin	Sample location	Pop origin	N° of samples	Pop code
Vitorino et al., 2015	Tocantins-Araguaia	Mimoso Lake, Araguaiana - MT	Wild	45	MLA
Vitorino et al., 2015	Tocantins-Araguaia	Small Lake, Araguaiana - MT	Wild	8	SLA
Vitorino et al., 2015	Tocantins-Araguaia	Égua Lake, Cocalinho - MT	Wild	3	ELC
Vitorino et al., 2015	Tocantins-Araguaia	Rio das Mortes, Nova Nazaré, - MT	Wild	10	RMN
Vitorino et al., 2015	Tocantins-Araguaia	São Miguel - GO	Captive	2	SMG
Vitorino et al., 2015	Tocantins-Araguaia	Rio das Mortes, Yellow Barrier - MT	Wild	2	RMY
Vitorino et al., 2015	Tocantins-Araguaia	Araguaia River, Novo Santo Antônio - MT	Wild	8	ARN
Vitorino et al., 2015	Tocantins-Araguaia	Novo Santo Antônio Region - MT	Wild	30	NSR
Vitorino et al., 2015	Tocantins-Araguaia	Novo Santo Antônio and Luciara Region - MT	Wild	19	NAL
Vitorino et al., 2015	Tocantins-Araguaia	São Félix do Araguaia Region - MT	Wild	11	SFA
Vitorino et al., 2015	Tocantins-Araguaia	Santa Terezinha Region - MT	Wild	3	SRT
Vitorino et al., 2015	Tocantins-Araguaia	Parauapebas Pisciculture - PA	Captive	14	PBP
Vitorino et al., 2015	Tocantins-Araguaia	Itupiranga - PA	Wild	9	ITU
Almeida, 2012	Amazon	Itapera Community, Santarém - PA	Wild	15	ITA
Almeida, 2012	Amazon	Mucurituma Community, Santarém - PA	Wild	9	MUC
Almeida, 2012	Amazon	Pixuna Community, Santarém - PA	Wild	24	PIX
Almeida, 2012	Amazon	Santa Maria Community, Santarém - PA	Wild	41	SMC
Boa Esperança Farm, Rondônia - RO	Amazon	Porto Velho - RO	Captive	6	POR
Boa Esperança Farm, Rondônia - RO	Amazon	Bolivia	Captive	2	BOL
Boa Esperança Farm, Rondônia - RO	Amazon	Purus River - AC	Captive	11	PUR
Vitorino et al., 2015	Amazon	Amazon River - AC	Wild	2	AMZ
Boa Esperança Farm, Rondônia - RO	Amazon	Peru	Captive	11	PER
Mamirauá Institute, Tefé - AM	Amazon	Pantaleão, Mamirauá Lake, Tefé - AM	Undetermined	19	PML
Mamirauá Institute, Tefé - AM	Amazon	Maraã, Mamirauá Lake, Tefé - AM	Wild	18	MML
Juruá Institute, Manaus - AM	Amazon	Lago Serrado Community, Middle Juruá - AM	Wild	32	LSC
Market Samples	Amazon	Undetermined	Undetermined	6	UND*
Total	Total			360	25

***Legend:** hydrographic basin, sampling site, origin, number of samples, acronym and code of *Arapaima gigas* populations.

*Market Samples were attributed on ITA population to conduct the genetics analysis

2.4.2 DNA extraction

DNA was extracted with a modified protocol using CTAB methodology (BOYCE et al., 1989). Tissue and fin samples were cut into 0.25 cm³ blocks and immersed in 500 µl of extraction buffer containing CTAB: 100 mM Tris-HCL (pH 8.0); 1.4 M NaCl; 0.02 M EDTA; 2% CTAB (Hexadecyltrimethylammonium bromide); 0.2% (2-mercaptoethanol). Five µl of proteinase K (10ng/µl) were added, followed by incubation at 60° C for 1 hour in a dry bath. After digestion, the solution was centrifuged for 2 min. at 13,000g. The upper aqueous phase was carefully pipetted to a new tube containing 500µL of chloroform:isoamyl alcohol (24:1), vortexed for 10 sec., and centrifuged at 13,000g for 15 min. The upper aqueous phase was transferred to a new tube containing 250µL of cold Isopropanol, carefully mixed by inversion and kept at -20°C for at least 30 min. DNA was pelleted by centrifugation at 13,000g for 10 min. at room temperature. After removal of supernatants, pellets were washed with 700µL of cold 70% ethanol. Supernatants were removed by decantation and tubes left open to dry at room temperature. Dried pellets were suspended in 100µL of TE (Tris-EDTA pH 8.0) and 6µL of RNase A (10mg/µL) and incubated at 37°C for 30 min.

2.4.3 SNP mining and assay design

STACKS (CATCHEN et al., 2013) was used to demultiplex publicly available ddRAD sequence reads derived from 60 individuals from five distinct sampling locations in the Amazon and Tocantins-Araguaia River basins (TORATI et al., 2019). A total of 33,932,300 raw reads, consisting of 16,966,150 paired-end reads, were generated from the MiSeq run described in Torati et al. (2019). Following sample demultiplexing and quality filtering, 29,133,620 reads (85.8%) were retained. Assembling loci (RAD-tags) across the 60 individuals revealed 12,378 unique RAD-tags. The panel was streamlined to 448 markers with 1–2 SNPs found in over 80% of individuals across five populations. Additional filtering retained only the most polymorphic SNP when two were detected in a RAD locus and eliminated instances of the same SNP at the 3' end of paired reads. Consequently, the initial panel for population genetic analyses was reduced to 393 SNPs. Processed data were subsequently mapped to the *Arapaima gigas* genome (DU et al., 2019) using BWA-MEM (LI, 2013), while maintaining the identification of all individuals included in the original study. SAMtools (DANECEK et al., 2021) was used to order and merge the generated bam files. SNP calls were based on the bam files using FreeBayes (GARRISON et al., 2012). MAF estimates were obtained based on individual genotype calls at observed SNPs using a C++ script. Candidate SNPs were selected using the following criteria:

i) Overall estimated $MAF > 0.10$, ii) minimum of 100 bp on each side of the target SNP with no observed SNPs, iii) SNP observed in a minimum of 10 individuals and iv) 100 reads, v) genome location. A total of 192 evenly spaced SNPs distributed across all chromosomes, were submitted to SNPtype™ assay design and synthesis by Fluidigm (South San Francisco, CA, USA).

2.4.4 SNP genotyping

Genotyping data were produced following standardized protocols provided by Fluidigm for utilization with an EP1 platform. Each sample went through initial amplification using a set of Locus-Specific Primer (LSP) and Specific Target Amplification (STA) oligonucleotides for each of the two sets of 96 SNPs. Subsequently, each sample was subjected to genotyping on 96.96 Integrated Fluidic Circuit Arrays (96.96 Dynamic Array IFC). This involved combining diluted pre-amplified samples with each of the two sets of 96 SNPtype assays which contained Allele-Specific Primers (ASP) and LSP oligonucleotides specific to each SNP. Following amplification, each 96.96 Dynamic Array IFC Array was scanned using Fluidigm EP1 equipment, and genotype calls were determined using Fluidigm SNP 4.1 genotyping analysis software (FLUIDIGM CORPORATION, 2018).

2.4.5 Data quality control

Obtained SNP genotyping data were filtered according to confidence values to exclude genotypes with Call Confidence $< 95\%$ from further analysis. Subsequently, samples and markers with Call Rates $< 85\%$ were also excluded from the dataset.

2.4.6 Genetic Diversity

GenAlEx v6.5 (PEAKALL; SMOUSE, 2012) was used to estimate Observed (H_o) and Expected (H_e) Heterozygosity, F_{IS} (Coefficient of Inbreeding) and pairwise F_{ST} values. UPGMA trees of genetic distances was constructed with the program MEGA11 (TAMURA; STECHER; KUMAR, 2021) based on the pairwise F_{ST} matrix of all studied populations. Analysis of molecular variance (AMOVA) was performed with Arlequin v.3.5.2.2 (EXCOFFIER; LISCHER, 2010).

2.4.7 Genetic Structure

The genetic structure of sampled *Arapaima gigas* populations was evaluated using Structure v2.3.4 (PRITCHARD; STEPHENS; DONNELLY, 2000). All Structure analysis were

carried out considering the admixture model and correlated allele frequencies (FALUSH; STEPHENS; PRITCHARD, 2003), K values between 1 and 19, with 20 repetitions each, and Burn-in of 100,000, with 100,000 Markov-Monte Carlo Chains (MCMC) for each repetition. The most likely number of clusters (Best K) was estimated using Evanno's method (EVANNO; REGNAUT; GOUDET, 2005) implemented by (2012) Structure Harvester (EARL et al., 2012). The Principal Component Analyses (PCA) were conducted with the software SNP & Variation Suite v8.x (Golden Helix, Inc., Bozeman, MT).

2.4.8 Genetic versus geographic distances

Alleles in Space (AIS) 1.0 software (MILLER, 2005) was used to investigate the relationship between estimated genetic distances among populations and their geographic distances. The Mantel Test (MANTEL, 1967) is a widely employed method for assessing the correlation between congruent similarity or dissimilarity matrices. In AIS the matrices were built using data from population genotypes and the geographic coordinates of sample collection sites. We also conducted Spatial Autocorrelation Analyses and Landscape Shape Plot at AIS, to better observe and identify genetic distance patterns across river basins.

2.4.9 Kinship analysis

Identity by descent (IBD) analyses were conducted in SNP & Variation Suite v8.x (Golden Helix, Inc., Bozeman, MT) to estimate the genetic relationships (kinship) between pairs of individuals within group. All populations were first analyzed together followed by individual IBD analysis with each sampled population, considering all river basins and SNP panels. Finally, the IBD analyses were performed considering only the Tocantins-Araguaia populations since initial Structure analyses showed low genetic structuring of the populations from this basin.

2.5 RESULTS

2.5.1 Data quality control

Data quality control resulted in a final dataset of 150 markers and 321 samples in 19 populations. A total of 147 samples were obtained from the Tocantins-Araguaia basin and 174 samples were derived from Amazon River basin. Two samples from Boa Esperança Farm were excluded after data quality control. Populations ELC, SMG, RMY, SRT (Tocantins-Araguaia)

and BOL, AMZ (Amazon) were completely excluded as there were no samples left after the filters.

The MAF quality filter excluded nine monomorphic SNPs in all populations. Markers with low MAF (<0.05) were maintained as they could represent relevant information about rare alleles in specific *Arapaima gigas* populations. The average estimated MAF per population ranged from 0.039 (MLA) to 0.265 (PIX) (Table 2.2, Figure 2.2).

Table 2.2 - Number of markers within estimated Minor Allele Frequency intervals and overall MAF averages for each population.

Population	MAF Interval					MAF Average
	0.00-0.09	0.10-0.19	0.20-0.29	0.30-0.39	0.40-0.50	
MLA	131	4	5	6	4	0.039
SLA	131	6	2	8	3	0.040
RMN	115	12	8	8	7	0.069
ARN	121	8	5	9	7	0.061
NSR	117	10	9	8	6	0.067
NAL	117	12	3	10	8	0.064
SFA	111	9	14	10	6	0.076
PBP	55	9	13	33	40	0.242
ITU	75	19	21	17	18	0.153
ITA	34	32	24	29	31	0.242
MUC	37	32	14	39	28	0.239
PIX	31	23	27	32	37	0.265
SMC	25	31	27	32	35	0.261
POR	49	15	29	25	32	0.227
PUR	46	21	25	26	32	0.225
PER	50	20	22	23	35	0.224
PML	31	25	27	32	35	0.260
MML	34	23	24	27	42	0.258
LSC	37	28	17	30	38	0.248
Total MAF Average						0.172

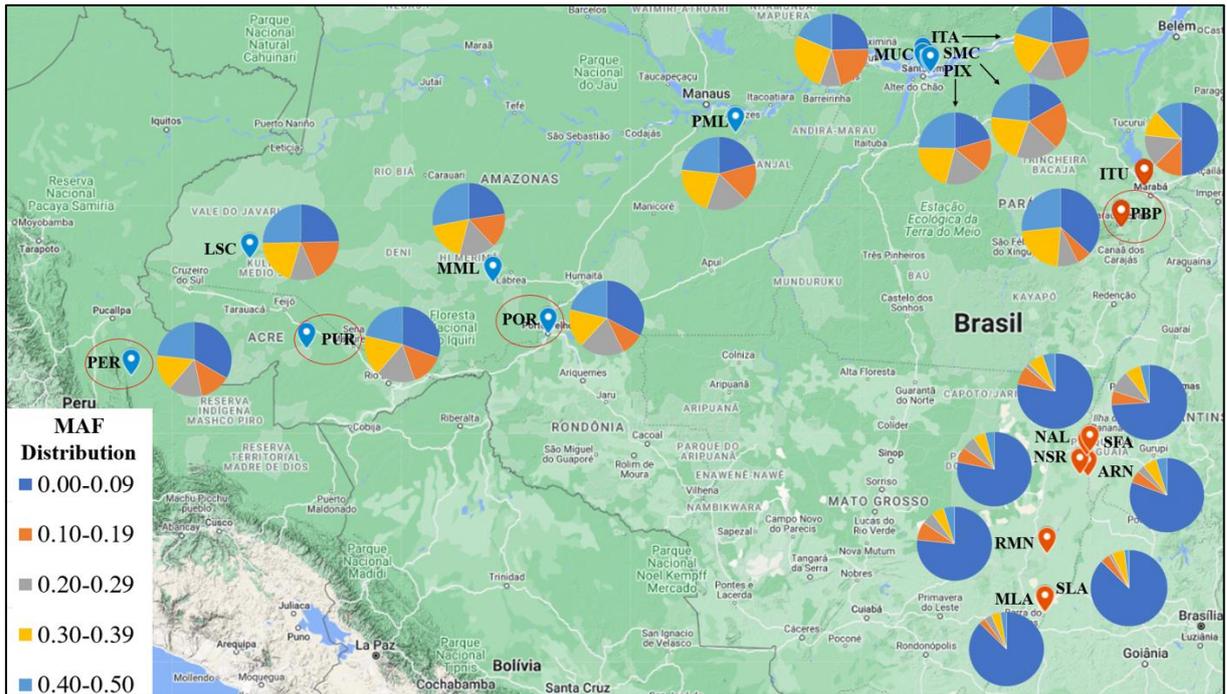


Figure 2.2 – MAF distribution for each *Arapaima gigas* population.

2.5.2 Genetic diversity

The genetic diversity statistics for the 19 studied populations are shown in Table 2.3. The mean of number of alleles ranged from 1.133 to 1.600 in populations from Tocantins-Araguaia basin and from 1.767 to 1.953 in Amazon basin. The expected heterozygosity ranged from 0.050 to 0.204 in Tocantins-Araguaia populations and from 0.291 to 0.340 in Amazon basin. The Wright's inbreeding coefficient (F_{IS}) ranged from 0.030 to 0.334 in Tocantins-Araguaia basin and from 0.034 to 0.220 in Amazon basin.

Table 2.3 – Genetic diversity statistics for the 19 *Arapaima gigas* studied populations.

Population	Na	Ho	He	F _{is}
MLA	1.133	0.068	0.050	-0,334
SLA	1.133	0.071	0.052	-0,327
RMN	1.280	0.105	0.092	-0,083
ARN	1.227	0.096	0.080	-0,107
NSR	1.300	0.091	0.093	0,050
NAL	1.253	0.081	0.086	0.137*
SFA	1.267	0.103	0.101	0.030
PBP	1.767	0.268	0.294	0.118*
ITU	1.600	0.244	0.204	-0.160
ITA	1.880	0.345	0.315	-0.060
MUC	1.860	0.356	0.314	-0.084
PIX	1.913	0.361	0.340	-0.055
SMC	1.953	0.370	0.336	-0.096
POR	1.767	0.281	0.294	0.135*
PUR	1.807	0.298	0.296	0.034
PER	1.793	0.254	0.291	0.179*
PML	1.920	0.356	0.337	-0.037
MML	1.860	0.392	0.328	-0.220
LSC	1.900	0.371	0.318	-0.200
Mean	1.611	0.237	0.222	-

Number of Alleles (Na) per marker, observed heterozygosity (Ho), expected heterozygosity (He) and Wright's inbreeding coefficient (F_{is}). Tocantins-Araguaia: orange shade; Amazon basin: blue shade.
* Indicates significant *P*-value (*P*<0.05).

The overall estimated F-statistics (F_{is}, F_{it}, F_{st}) were -0.044, 0.396 and 0.424, respectively. Overall estimated F_{is} was not significant (*P* > 0.05). Figure 2.3 shows pairwise F_{st} comparisons between all sampled populations. F_{st} values ranged from 0.001 (between PIX and SMC) to 0.789 (between MLA and PER). A table with the pairwise matrix of F_{st} is shown on Supplementary Material S1.

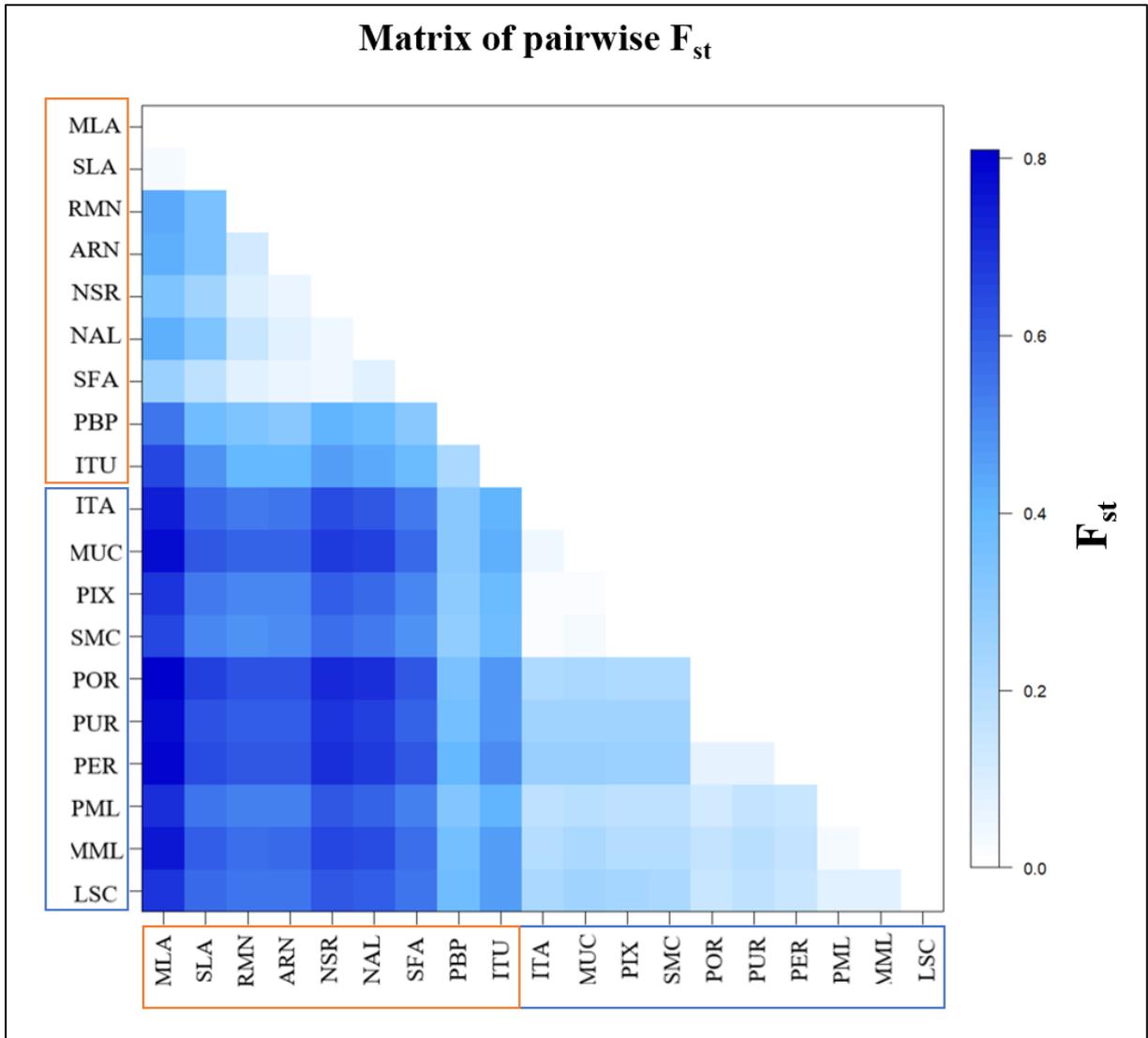


Figure 2.3 – Color diagram graph based on pairwise F_{st} values. The darker shades of blue indicate the highest F_{st} values and the lighter shades indicate the lowest F_{st} values. Population MLA to ITU (orange rectangles) are from Tocantins-Araguaia basin and ITA to LSC (blue rectangles) are from Amazon basin.

The UPGMA tree of genetic distances obtained with the pairwise F_{st} estimates is shown in Figure 2.4. This analysis separated two big clusters within the 19 populations, showing Tocantins-Araguaia populations (MLA to ITU) are genetically distant from the Amazon basin populations (ITA to LSC).

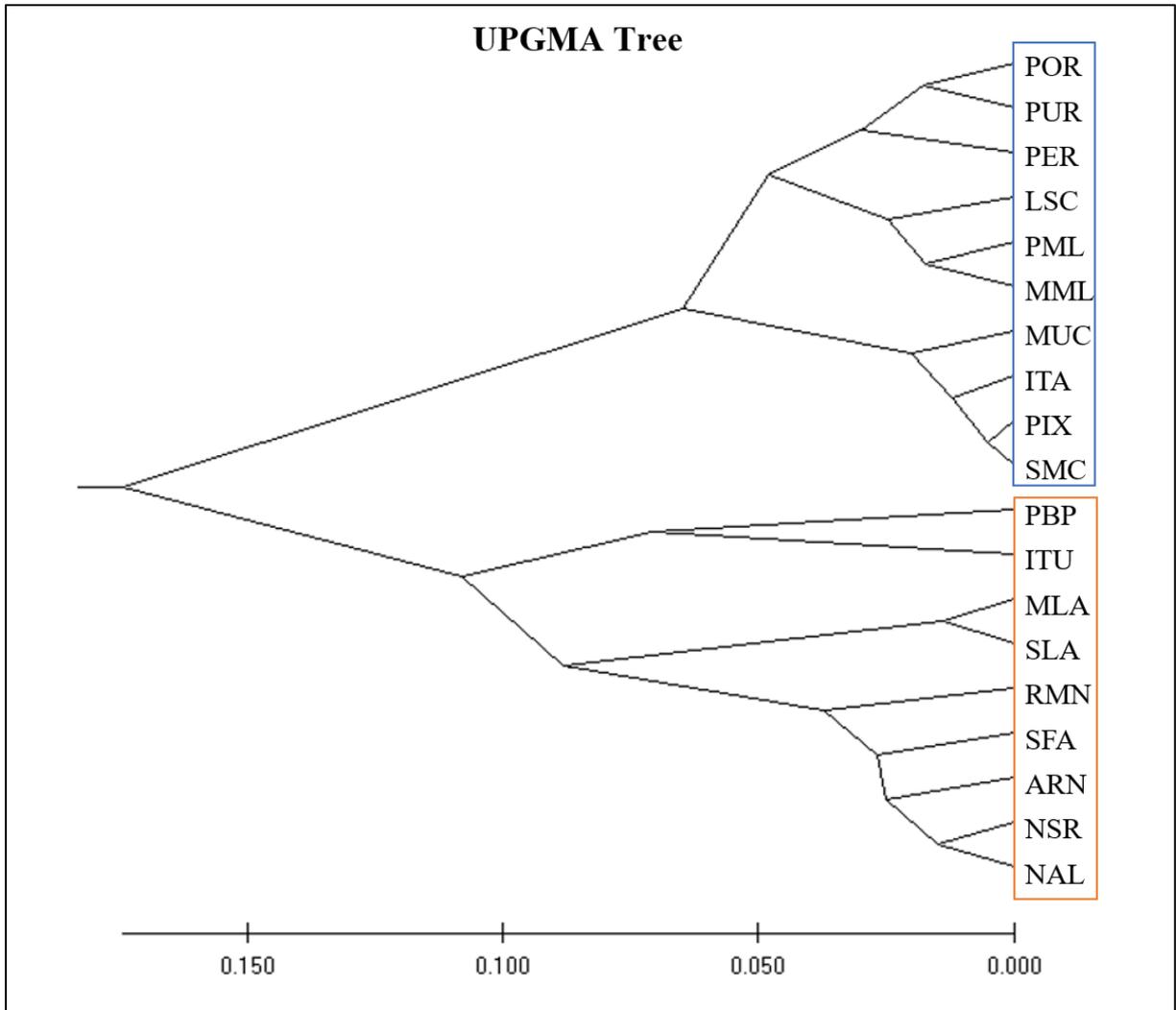


Figure 2.4 – UPGMA tree of genetic distances based on pairwise F_{st} estimates between the 19 populations of *Arapaima gigas*. The orange rectangle indicates the Tocantins-Araguaia basin populations and the blue one indicates the Amazon basin populations.

The AMOVA results show that 45.76% of the observed variation is present among populations ($F_{st} = 0.458$, $P < 0.05$), while -3.71%* and 57.96% is observed among individuals within populations and within individuals, respectively (Table 2.4).

Table 2.4 – AMOVA results for the 19 populations of *Arapaima gigas*.

Source of variation	Variance components	Percentage of variation
Among populations	13.62211	45.76
Among individuals within populations	-1.1052	-3.71*
Within individuals	17.25389	57.96

* $P > 0.05$

2.5.3 Genetic Structure

The genetic structure analyses divided the 19 populations of *Arapaima gigas* in two main clusters, with the higher probability of grouping indicated by Evanno method (K=2). The analysis of the Structure graph based on K=2 (Figure 2.5) shows the division of 19 populations indicating that they have two principal ancestors, one for Tocantins-Araguaia River basin (orange bars) and other for Amazon River basin (blue bars). With K=2 populations PBP and ITU show a noticeable admixture of the two clusters, with orange color (Tocantins-Araguaia) being the one with greatest presence in these two populations. With K=3, an additional cluster containing populations ITA, MUC, PIX and SMC can be observed, with populations PBP and ITU showing an admixture of the orange and the dark purple clusters. Populations ITA, MUC, PIX and SMC were separated as the dark purple cluster. The K=4 graph shows an additional cluster formed with populations MLA and SLA. K=5 shows a greater separation of populations PBP and ITU, forming a new cluster. A similar result is shown in the Principal Component Analysis below. More Structure graphs are shown in Supplementary Material S2.

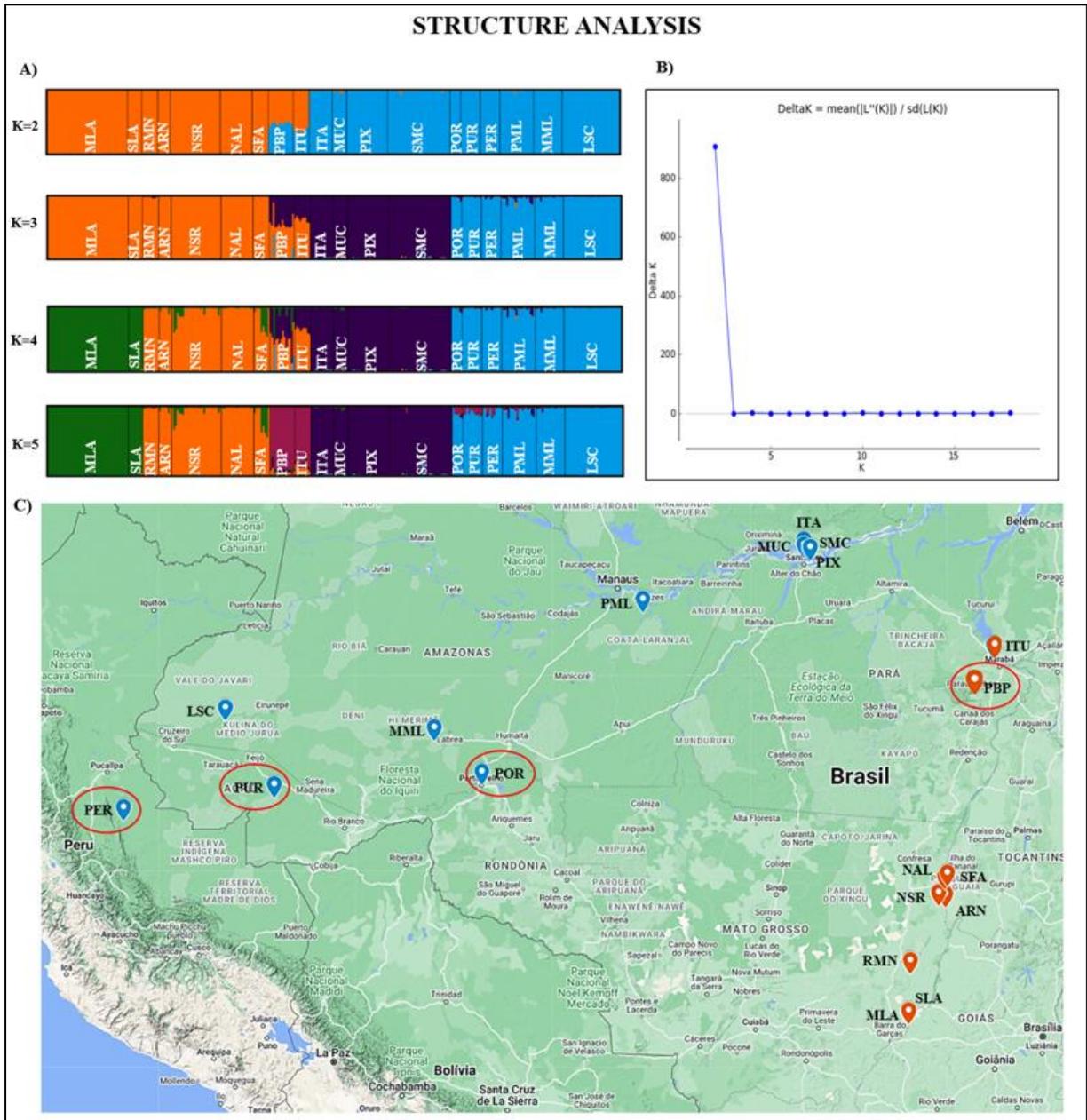


Figure 2.5 – Genetic structure graph of the 19 *Arapaima gigas* populations considering the best $K=2$. (A) Genetic structure graphs. (B) Delta K graph. (C) Sampling locations considering Structure graph ($K=2$) colors (Google My Maps modified by the author). Population PBP is from a commercial farm located in Parauapebas – PA. Populations POR, PUR and PER are broodstock from a commercial farm in Rondônia – RO which kept records of original capture location.

The PCA results show populations separated in four different clusters (Figure 2.6). The first Tocantins-Araguaia populations MLA to SLA (A) formed a distinguishable and well joined cluster. The 16th and 17th (PBP and ITU) populations (B) were separated from the other

Tocantins-Araguaia samples in a distinct cluster. The two additional clusters contain the Amazon basin populations ITA, MUC, PIX and SMC (C) and POR to LSC (D), respectively.

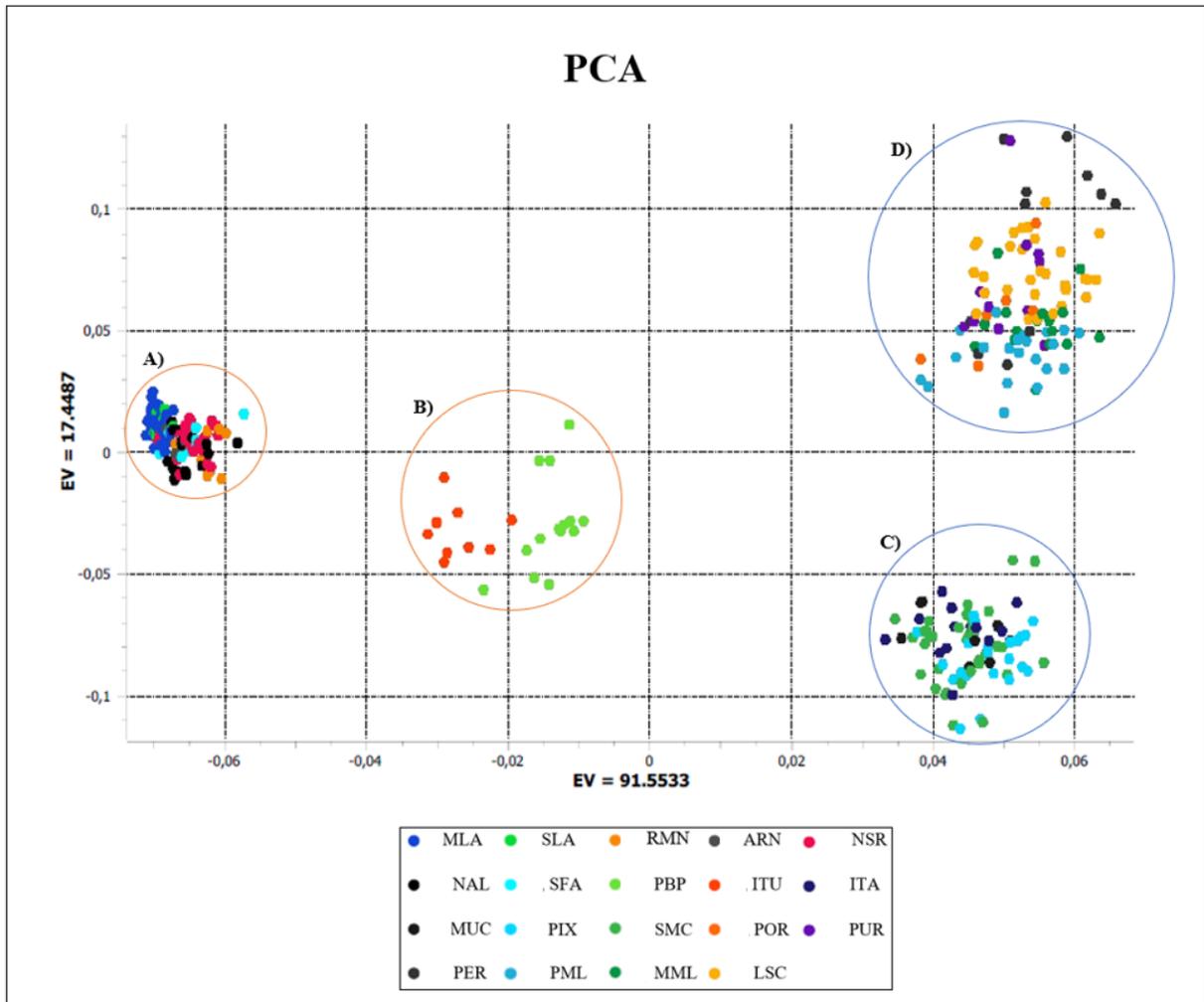


Figure 2.6 - PCA plot showing the four most significant principal components formed for the 19 populations of *Arapaima gigas*. Each individual sample is represented as a dot. (A) Tocantins-Araguaia basin populations along the Araguaia River between Mato Grosso and Tocantins state (MLA; SLA; RMN; ARN; NSR; NAL; SFA). (B) Tocantins-Araguaia basin populations in the southeast of Pará state (PBP; ITU). (C) Amazon basin populations near Santarém-PA (ITA; MUC; PIX; SMC). (D) Amazon basin populations along Amazônia, Rondônia and Acre state and Peru and Bolivia countries.

2.5.4 Genetic and geographic distances

To evaluate the correlation between the geographic and genetic distance of sampled *Arapaima gigas* populations, a Mantel Test were performed (Figure 2.7). The results showed that 58% of genetic variation is related to geographic distance ($r = 0.582$, $P = 0.001$).

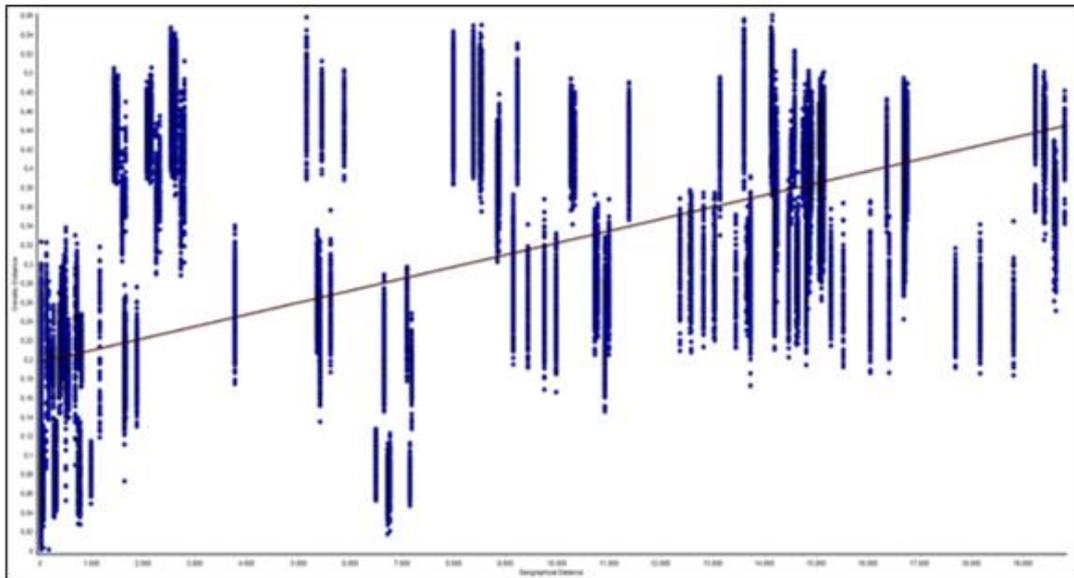
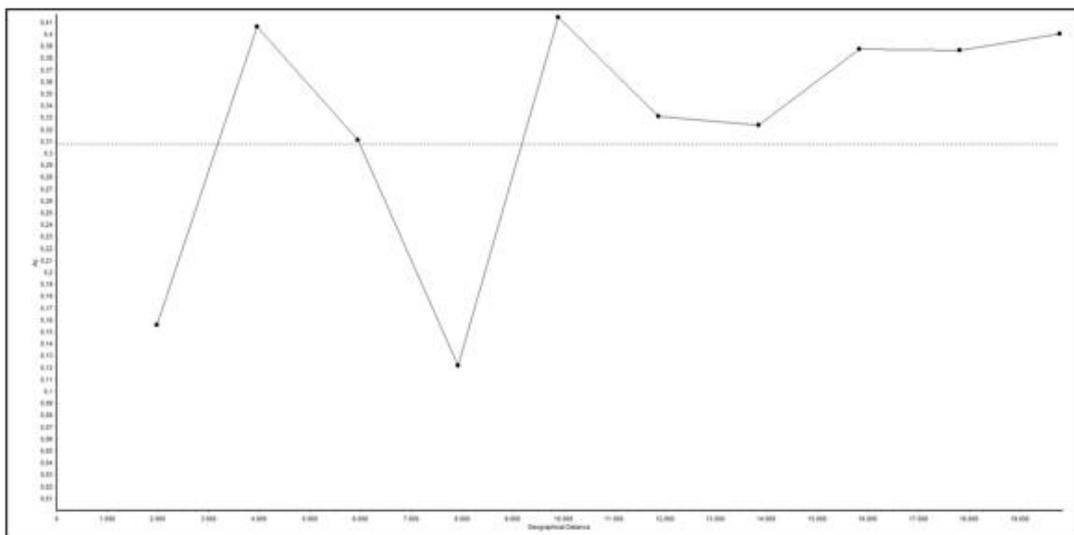
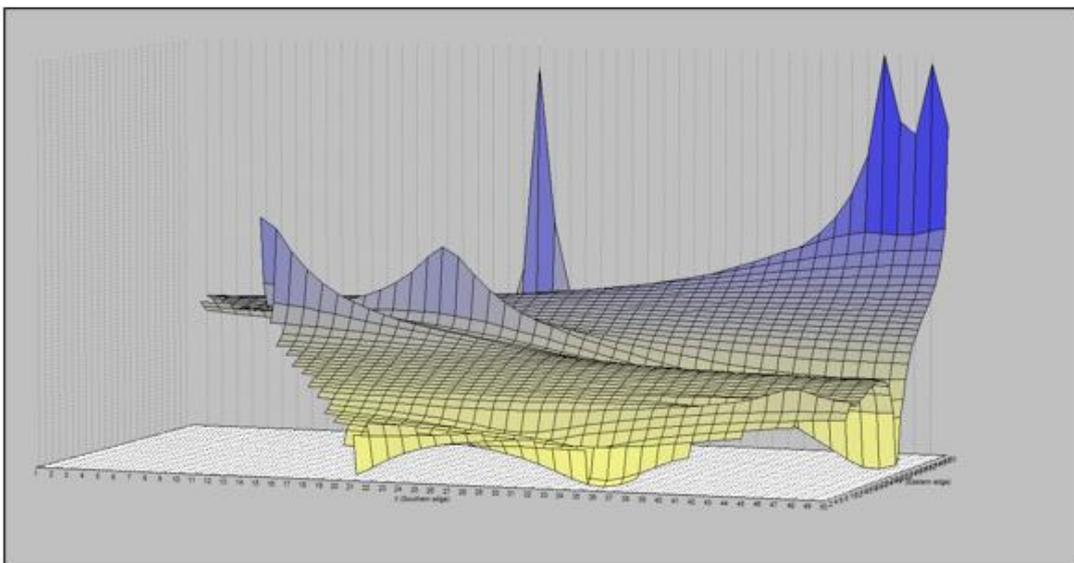
A) Mantel Test**B) Spatial Autocorrelation Analysis****C) Landscape Shape Plot**

Figure 2.7 – (A) Mantel Test showing the relationship between geographical (meters) and estimated genetic distances of *Arapaima gigas* populations (F_{st}). (B) Spatial Autocorrelation Analysis. (C) Landscape Shape Plot.

2.5.5 Kinship analysis

The IBD analysis was made to estimate the genetic relationships (kinship) between pairs of individuals within group considering all 19 *Arapaima gigas* populations. This analysis revealed a noticeable number of samples from population MLA and SLA of Tocantins-Araguaia basin with high levels of kinship (>0.9) therefore, these samples were investigated separately from the others, as shown in the figure 2.8. These analyzes raised the suspicion of a possible bias in SNP markers for these two populations, therefore, a new MAF filter was performed, removing all monomorphic SNPs, leaving only 20 informative markers for these two populations. It could reduce analysis accuracy due to lack of markers so, the ideal solution to this problem would be to mine more SNPs from samples coming of these two populations to have a new set with informative markers. More IBD graphs are available in Supplementary Material S3.

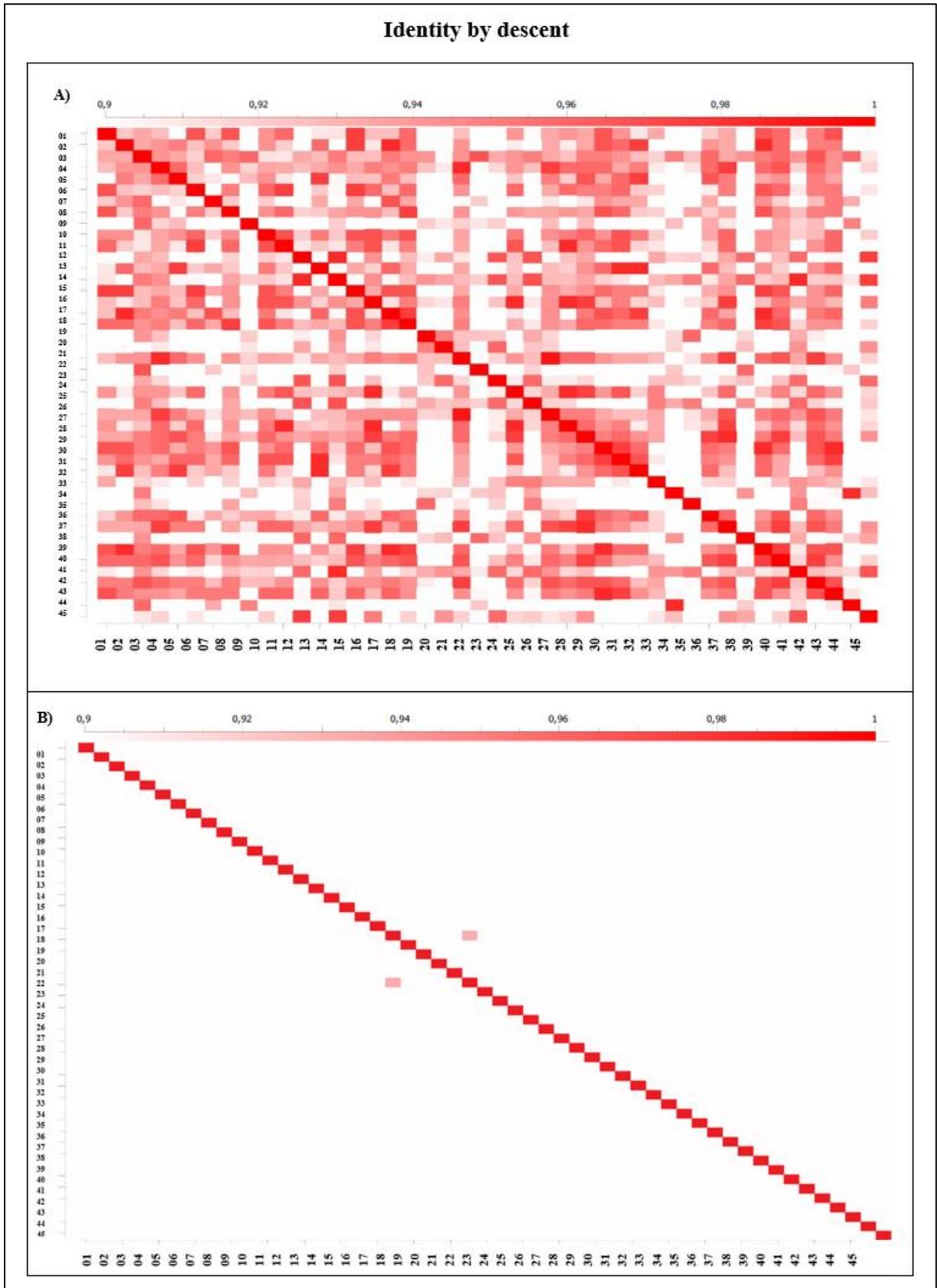


Figure 2.8 – Identity by descent analysis of Tocantins-Araguaia *Arapaima gigas* samples. A) Populations MLA and SLA before MAF filtering and B) Populations MLA e SLA after MAF filtering.

2.6 DISCUSSION

This represents the first comprehensive study using SNP markers to investigate the genetic structure and diversity of 19 *Arapaima gigas* populations sampled throughout the Amazon and Tocantins-Araguaia River basins. Results showed high levels of genetic differentiation among all populations and specially between the Amazon and Tocantins-Araguaia populations. These outcomes demonstrate that the set of SNP markers designed with Amazon and Tocantins-Araguaia samples can be used with other *Arapaima gigas* populations, maintaining good data quality. This is an excellent low-cost tool to explore genome data in captive and wild samples and help in conservation and broodstock management (HABIER; FERNANDO; DEKKERS, 2009).

Several technologies are available for genotyping SNP (Single Nucleotide Polymorphisms) markers at different densities. Depending on the application, lower-density SNP panels can be more cost-effective. SNP data has been used for analyzing kinship, genetic diversity and structure, and even genomic evaluation and selection of aquaculture species such as Tilapia (*Oreochromis niloticus*), Rainbow Trout (*Oncorhynchus mykiss*) and Tambaqui (*Colossoma macropomum*) (IANELLA, 2016; JOSHI et al., 2018; ZHANG et al., 2019; ARAÚJO et al., 2023). The advancement of sequencing methods made it possible to prospect SNP markers in a more accessible and faster way, in addition to increasing the number of analyzed markers (CAETANO, 2009; SCHUNTER et al., 2014). Due to these characteristics, SNPs are cited as favorite markers in studies of evolution and conservation genetics (SCHUNTER et al., 2014; DELORD et al., 2018) and for that were used to investigate *Arapaima gigas* population structure in this study.

This was the first study of genetic structure for *Arapaima gigas* populations using a low-density SNP panel performed with Fluidigm genotyping. In addition, the high number of samples from several sampling sites in the two river basins it is also a difference between the genetic studies of *Arapaima gigas* already published. The genetic diversity statistics for the 19 populations of *Arapaima gigas* corroborate the findings of other studies with the species in the investigated river basins, using, in most of them, molecular markers as ISSR, mtDNA and Microsatellites (HRBEK et al., 2005; TORATI et al., 2019; VITORINO et al., 2017b; NOGUEIRA, et al., 2020). In the Tocantins-Araguaia basin, populations exhibited a mean number of alleles ranging from 1.133 to 1.600, reflecting a relatively narrower spectrum of genetic diversity. Conversely, populations within the Amazon basin displayed a broader range, with mean allele numbers spanning from 1.767 to 1.953, indicative of a higher genetic richness.

In general, populations of endangered species which have experienced strong bottlenecks consistently exhibit low values of expected heterozygosity as a consequence of the decline in genetic diversity (VITORINO et al., 2015; FARIAS et al., 2019). Furthermore, subsequent inbreeding events resulting from the reduction in the number of available individuals in reproductive age can intensify genetic erosion processes, leading to diminished reproductive success and affecting the overall evolutionary potential of a species. In the Tocantins-Araguaia populations, expected heterozygosity ranged from 0.050 to 0.204, suggesting a moderate level of diversity and a possible bottleneck effect. In contrast, the Amazon basin populations exhibited wider heterozygosity ranges, varying from 0.291 to 0.340. Higher genetic diversity levels observed in populations from the Amazon basin have also been reported by TORATI et al., (2019), who proposed this could be a consequence from the larger population sizes naturally found in less urbanized environments in this basin, in comparison to the Tocantins-Araguaia basin. Wild populations in Amazon basin were captured in sites with less urbanization, for example in Mamirauá and Juruá reserve, while Tocantins-Araguaia populations were captured near country cities along Mato Grosso, Goiás e Pará states, consequently, this river basin passes through more urbanized areas than Amazon basin.

Large within-population genetic diversity differences observed between the two river basins may be explained by large differences observed in the morphodynamical characteristics of the Amazon and Tocantins-Araguaia basins. Unlike in the Amazon basin, flood periods are shorter in the Tocantins-Araguaia and therefore, during the dry season, larger numbers of small lakes are formed along the river (AQUINO; LATRUBESSE; SOUZA FILHO, 2009). Given the close connection between the dispersal of juvenile and adult *Arapaima gigas* in floodplain environments and the annual flooding cycle, alterations in the river's hydrological dynamics could impact the migration patterns of this fish as the *Arapaima gigas* has a known sedentary behavior, that is, it does not tend to migrate long distances in the main river channel, as other wild species (CASTELLO, 2008; DE OLIVEIRA et al., 2012), potentially enhancing genetic drift and population bottlenecks effects and causing the decrease in the overall genetic diversity of Tocantins-Araguaia populations (AQUINO; LATRUBESSE; SOUZA FILHO, 2009; NOGUEIRA et al., 2020). Population 16 (PBP) was collected in a production farm in Parauapebas - PA, contradicting the perception that captive samples have less genetic diversity (CARLOS; VERA, 2014). This population showed a higher level of genetic diversity than other populations from the same river basin. In some circumstances, recent genetic bottleneck events may result in a temporary increase in the number of heterozygotes in a population. During a population bottleneck, there is a substantial decrease in population size, resulting in

diminished genetic diversity as alleles are depleted from the gene pool. Nonetheless, if this bottleneck precedes a phase of population expansion or recovery, individuals with different genetic diversity might interbreed, causing a surge in heterozygosity. This phenomenon arises from the fact that survivors of the bottleneck may carry various alleles that existed in the ancestral population but were scarce or absent prior to the bottleneck occurrence (ZHANG et al., 2019). This could explain why it had a high level of genetic diversity and the distinct genetic structure found, compared to the other Tocantins-Araguaia samples (VITORINO et al., 2017; TORATI et al., 2019).

Wright's inbreeding coefficient (F_{is}) provides insights into the mating patterns within populations. In the Tocantins-Araguaia basin, F_{is} values ranged from 0.030 to 0.334, with high values in populations MLA and SLA. Compared to Tocantins-Araguaia, populations from the Amazon basin demonstrated a similar, but smaller F_{is} range (0.034 to 0.220), providing further evidence that these populations are being subjected to lower inbreeding levels, as previously suggested by TORATI et al. (2019) and VITORINO et al. (2015).

The global F_{st} value of 0.424, indicates a significant level of genetic differentiation among populations which was corroborated by the AMOVA results, as a large proportion of the observed variation was allocated among the different studied populations (45.76%). The range of F_{st} values demonstrates varying degrees of genetic divergence among population pairs. Notably, the highest F_{st} value (0.789) between populations MLA and PER suggests a substantial genetic differentiation between the populations of the two river basins.

The observed genetic differentiation among *Arapaima gigas* populations at certain locations along the species' distribution may be also attributed to the fact that the regions where *Arapaima gigas* samples were collected have experienced significant impacts from fishing activities and the conversion of untouched land to agriculture and livestock production. These activities disrupt natural habitats, degrade water quality, reduce oxygen levels, and alter ecosystem dynamics. As a result, aquatic biodiversity, ecological integrity, and ecosystem services are jeopardized, causing a huge impacting in aquatic organisms (ALMEIDA., 2012; VITORINO et al., 2015).

The presented analysis results support previously reported findings which suggested that limited levels of gene flow between the studied populations of *Arapaima gigas*, in addition to high levels of inbreeding, especially in the Tocantins-Araguaia basin, have had prominent impacts over the historical and current genetic structure of the species (NOGUEIRA et al., 2020; VITORINO et al., 2017). Farias et al. (2019) found that the genetic structure of the Amazonas population of *Arapaima gigas* revealed significant isolation-by-distance at the

largest scale. At smaller scales of sampling, it is probable that sub-populations of *Arapaima gigas* operate independently in demographic terms, challenging the presumptions of compensation for exploitation or replenishment over fishery-relevant timescales (FARIAS et al., 2019; HRBEK et al., 2005). Regardless of the determining factors, genetic fragility is a cause for concern, as future environmental impacts, whether natural or not, could further reduce the diversity of these populations and pose a threat to their long-term viability (VITORINO et al., 2017).

The UPGMA tree analysis revealed a clear separation between the Amazon and Tocantins-Araguaia River basin, supporting the hypothesis that their ancestors became sufficiently genetically distant to form distinct large clusters in the studied hydrographic basins (CASTELLO, 2008; CATÂNEO et al., 2022). Vitorino et al. (2017) discussed that the Tocantins-Araguaia samples may have a distinct demographic history because of the differences related to the individual's maturity, sex ratio, offspring survival rates and availability and quality of habitats. For the Itupiranga samples (ITU), for example, a genetic bottleneck was described as the responsible factor for the genetic structure of this population, possibly caused by the influence of a hydroelectric dam construction nearby (VITORINO et al., 2017). Further studies with wider sampling in the region, integrating genetic, ecological and environmental factors could help to further elucidate the evolutionary history of *Arapaima gigas* populations in the two basins.

The genetic structure analysis also separated populations into two large clusters, as the best K was two, corroborating previous findings and recent studies of genetic structure in the two river basins (FARIAS et al., 2019; TORATI et al., 2019). The PCA generated four different clusters (Figure 2.6). Cluster A is composed mostly of populations from the upper Tocantins-Araguaia, while Cluster B contains populations PBP and ITU, from the lower basin. In addition, these populations have a mix of ancestry considering the Structure analysis. Clusters C and D contain populations from the Amazon basin which are genetically distant most likely because of the large geographic distance, resulting in reduced gene flow between populations. Wild fish populations may encounter natural barriers like rivers, mountains, or unsuitable habitats that hinder individual movement between them. Furthermore, fish species often have specific habitat preferences, including particular water temperatures, currents, or substrate types, which further restrict their ability to disperse over long distances. (ARARIPE et al., 2013). Samples from populations POR, PUR and PER, were collected in a farm with the original capture location informed by farm management. *Arapaima gigas* production farms typically use animals captured from the wild as broodstock, overlooking the population or region of origin

as well as disregarding eventual close kinship relations between the animals (HRBEK et al., 2005). Without adequate breeding management, cultivated stocks may accumulate inbreeding and suffer from losses in genetic variability, becoming more susceptible to diseases and environmental variations, ultimately leading to losses in productivity (HRBEK et al., 2005b). A small difference was observed between the genetic diversity of studied wild and captive populations. In the Amazon basin, captive populations had H_e levels between 0.291 and 0.294 while the remaining free-living populations in the basin had H_e above 0.300. Although, in general, the evolutionary potential of the wild individuals is significantly greater than that of the captive individuals (CARLOS; VERA, 2014), in this case, the genetic diversity of captive populations may have been maintained by the right choice of the stock's founding fish. The farmer collected broodstock from different rivers and locations along the Amazon River basin, thus decreasing the inbreeding effects and ensuring greater variability in their stocks.

The IBD results demonstrated a high level of kinship (>0.9) in populations MLA and SLA of the Tocantins-Araguaia River basins. These samples were captured from small lakes in Araguaiana Region and therefore have a high probability of being full-sibs (VITORINO et al., 2015), or even to be duplicates of the same individual. The first hypothesis could corroborate previous findings as these populations have lower genetic diversity and notable genetic structure compared to others in the same basin. However, further investigation is needed to determine the cause of this unexpected Kinship level in these populations.

SNP ascertainment bias refers to the systematic deviation of population genetic statistics from theoretical expectations, which may arise from the nonrandom sampling of individuals or biased SNP discovery protocols. Although this issue can impact the accuracy of data generation and may lead to misleading analysis and results, ascertainment bias is inherently present in many SNP panels widely used in studies of different plant and animal species (LACHANCE; TISHKOFF, 2013). The original publicly available SNP data used for mining the two Panels of 96 SNPs used for genotyping all 19 *Arapaima gigas* populations samples came from animals originally from five different locations throughout the Tocantins-Araguaia and Amazon River Basin (Taipas – TO, Lagoa da Confusão – TO, Tucuruí – PA, Solimões River – AM and Amazon River in Iquitos, Peru) (TORATI et al., 2019). Different strategies have been proposed to deal with this issue (BAYE et al., 2009), but the most efficient strategy would include the use of additional sequence data from *Arapaima gigas* populations from a wider geographical sampling range for mining additional SNPs for generation of genotyping data.

Proper genetic management practices to maintain genetic diversity in broodstocks are essential to avoid inbreeding and its subsequent negative effects on fish growth and overall

productivity. Hrbek et al. (2005) suggested the use of a diverse set of *Arapaima gigas* founders' matrices in fish farms, independently of their geographical origins. This approach aims to mitigate potential loss of genetic diversity resulting from inbreeding depression and founder effect. Low-density SNP panels can provide valuable low-cost solutions for population genetics studies and offer essential support for conservation and broodstock management. Four of the 19 analyzed populations were sampled from aquaculture farms in Parauapebas – PA (PBP) and Rondônia – AM (POR, PUR, PER). Genetic diversity estimates obtained for these populations did not notably diverge from wild populations, suggesting correct genetic management practices have been used. Nonetheless, additional procedures using SNP data could be used in these and other farms to direct mating based on estimated kinship, etc., as in other aquaculture species (BOHLING; HAFFRAY; BERREBI, 2016; JOSHI et al., 2018; ARAÚJO et al., 2023).

In this study, for the first time, genetic data for *Arapaima gigas* populations were generated from a panel of low-density SNPs genotyped with Fluidigm technology. This approach proved to be effective for collecting highly accurate information on the structure and genetic diversity of wild and captive populations of *Arapaima gigas* and can be used to assist producers in the efficient management of captive breeders, identifying diverse genetic origins and possible kinship relationships among the founding animals of stocks. It is known that fish broodstocks for aquaculture production should be derived from wild-caught individuals from different selected areas to maximize genetic diversity in confined systems, which may limit domestication and genetic improvement processes (CARLOS; VERA., 2014). Torati et al. (2019) discussed that genetic analysis can be used to optimize broodstock formation and use, minimizing translocations of animals across geographic regions, which could inadvertently lead to unwanted genetic contamination of local wild populations as a consequence of unwanted breeding with escapees from aquaculture operations. Genetic conservation or restoration programs should consider genetic origin of broodstock used for alevin production for release in nature to avoid unwanted intermixing of germplasm, which may consequently affect survival wild populations. Considering that *Arapaima gigas* wild populations are highly structured, as shown, the unwanted introduction of individuals from different genetic backgrounds could lead to major impacts to wild populations, ranging from changes in local ecosystems to erosion and loss of local genetic resources (HRBEK et al., 2005; DE OLIVEIRA, et al., 2012; CARLOS; VERA, 2014).

Due to the historic decline of *Arapaima gigas* populations, prohibition or strict regulation of fishing is common in several areas of the Amazon Basin, where the species is naturally found. Currently, fishing is allowed under close monitoring and restrictions within

conservation units such as the Mamirauá and Juruá Institutes (ARARIPE et al., 2013; FARIAS et al., 2019). Markets close to intensive fishing sites are end points in the supply chain for legitimately caught fish, and may also receive and process fish illegitimately in unregulated areas. Genomic data have been analyzed to provide powerful insights for surveillance and regulation of this trade (BAKER, 2008; MEIKLEJOHN et al., 2021). An important discussion is the origin of the animals used as breeding in captive stocks and sold in regional markets. According to OGDEN, LINACRE (2015), in cases where illegally obtained timber or fish are integrated into legal commercial transactions, establishment of geographic origin becomes a pivotal aspect of any investigation. Alternatively, for species subject to a complete trade ban, geographic assignment may not be the primary focus for prosecution. However, it can still play a crucial role in pinpointing illegal fishing hotspots, comprehending criminal networks, or facilitating the repatriation of live seizures (OGDEN & LINACRE, 2015). Obtained results show that low-density SNP panels can identify genetic origin of samples. As seen in the results and analyzes of Structure and PCA, it is possible to trace the original collection site of the samples along the river basins studied. These data can be used in forensic investigations, comparing the animal genetic origin provided by the analyzes with the location of illegal trade happens as local markets, predatory fishing operations or unregulated captivity and therefore, could be used as an addition tool to monitor and regulate the fisheries.

2.7 CONCLUSION

SNP markers designed with Tocantins-Araguaia and Amazon River basins samples provided valuable insights into genome data for both captive and wild samples, offering essential support for broodstock management and genetic conservation of *Arapaima gigas* wild populations. Obtained results highlighted significant genetic differentiation and structuring between populations in the Amazon and Tocantins-Araguaia basins, revealing aspects of their diversity, structure, and the influence of environmental and anthropogenic factors. Obtained findings can help develop and maintain germplasm conservation strategies for *Arapaima gigas* populations to allow a proper balance between the preservation of wild genetic diversity and the demands arising from human activities as fishing and captive production and breeding programs.

REREFENCES

- ALMEIDA, I.G. **Prospecção de Marcadores Moleculares Sexo-Específicos e Análise de Estrutura Populacional de Pirarucu (*Arapaima gigas*) na região de Santarém, Pará.** Dissertação (Mestrado). Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília, Brasília, DF. 93p. 2012.
- ARARIPE J, RÊGO PS, QUEIROZ H, SAMPAIO I, SCHNEIDER H. Dispersal capacity and genetic structure of *Arapaima gigas* on different geographic scales using microsatellite marker. **PLoS One**. 8(1):1–7. 2013.
- AQUINO, S.; LATRUBESSE, E. M.; SOUZA FILHO, E. E. DE. Caracterização hidrológica e geomorfológica dos afluentes da Bacia do Rio Araguaia. **Revista Brasileira de Geomorfologia**, v. 10, n. 1, p. 43–54. 2009.
- ARAUJO, N.T. et al. Population Structure and Genetic Diversity of Rainbow Trout (*Oncorhynchus Mykiss*) Broodstocks From Brazil Using SNP Markers. **Aquaculture Reports**. v. 31. 2023.
- BAKER CS. A truer measure of the market: the molecular ecology of fisheries and wildlife trade. **Molecular Ecology**. 17(18):3985-3998. 2008.
- BAYE T, TIWARI H, ALLISON D, GO R. Database mining for selection of SNP markers useful in admixture mapping. **BioData Min**. 2:1. 2009.
- BOHLING, J.; HAFFRAY, P.; BERREBI, P. Genetic diversity and population structure of domestic brown trout (*Salmo trutta*) in France. **Aquaculture**, v. 462, p. 1–9, 1 set. 2016.
- BOYCE TM, ZWICK ME AND AQUADRO CF. Mitochondrial DNA in the bark weevils: Size, structure and heteroplasmy. **Genetics**. 123(4): 825-836. 1989.
- CAETANO, A.R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**. v. 38, n.1, p. 64-71. 2009.
- CARLOS, H. DOS A. DOS S.; VERA, M. F. DE A. E V. Genetic relationships between captive and wild subpopulations of *Arapaima gigas* (Schinz, in Cuvier, 1822). **International Journal of Fisheries and Aquaculture**, v. 6, n. 10, p. 108–123, 29 out. 2014.
- CASTELLO, L. Lateral migration of *Arapaima gigas* in floodplains of the Amazon. *Ecology of Freshwater Fish*, v. 17, n. 1, p. 38–46, 2008.
- CATÂNEO, D. T. B. DOS S. et al. Elucidating a history of invasion: population genetics of Pirarucu (*Arapaima gigas*, Actinopterygii, Arapaimidae) in the Madeira River. **Hydrobiologia**, v. 849, n. 16, p. 3617–3632, 2022.
- CATCHEN J, HOHENLOHE PA, BASSHAM S, AMORES A, CRESKO WA. Stacks: an analysis tool set for population genomics. **Mol Ecol**. 22(11):3124-40. 2013.
- DELORD, C. et al. A cost-and-time effective procedure to develop SNP markers for multiple species: A support for community genetics. **Methods in Ecology and Evolution**, 9, p. 1959 - 1974. 2018.
- DANECEK P. et al. Twelve years of SAMtools and BCFtools, **GigaScience**. 10(2). 2021.
- DE OLIVEIRA, E. G. et al. Effects of stocking density on the performance of juvenile Pirarucu (*Arapaima gigas*) in cages. **Aquaculture**, v. 370–371, p. 96–101, 11 dez. 2012.
- DU, K., WUERTZ, S., ADOLFI, M. et al. The genome of the arapaima (*Arapaima gigas*) provides insights into gigantism, fast growth and chromosomal sex determination system. **Sci Rep** 9, 5293. 2019.
- EARL, DENT A. AND VONHOLDT, BRIDGETT M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. **Conservation Genetics Resources** vol. 4 (2) pp. 359-361. 2012.

- EMBRAPA. **Embrapa Recursos Genéticos e Biotecnologia**. Brasília, DF. Available in: <<https://www.embrapa.br/busca-de-solucoes-tecnicas/-/produto-servico/8370/plataforma-aquaplus>> Access in Jan. 2024.
- EVANNO, G., REGNAUT, S., GOUDET, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. **Mol. Ecol.** 14 (8), 2611–2620. 2005.
- EXCOFFIER, L., LISCHER, H.E.L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under linux and windows. **Mol. Ecol. Resour.** 10 (3), 564–567. 2010.
- FALUSH, D.; STEPHENS, M.; PRITCHARD, J. K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. **Genetics Society of America**, v. 164, n. 4, p. 1567–1587, 2003.
- FARIAS, I. P. et al. The largest fish in the world's biggest river: Genetic connectivity and conservation of *Arapaima gigas* in the Amazon and Araguaia-Tocantins drainages. **PLoS ONE**, v. 14, n. 8, 1 ago. 2019.
- FLUIDIGM CORPORATION. **SNP Genotyping User Guide** (PN 68000098 Q1). 2018. <https://fluidigm.my.salesforce.com/sfc/p/#700000009DAw/a/4u0000019kFV/J5w6q0CKYdITv5ebT5PYx8Ai2T5oGz.CIYtyMxHAWFg>.
- GARRISON E, MARTH G. Haplotype-based variant detection from short-read sequencing. arXiv preprint arXiv:1207.3907 [q-bio.GN]. 2012.
- GOLDEN HELIX, INC., BOZEMAN, M. **SNP & Variation Suite v8.x**, [s.d.]. Available in: <www.goldenhelix.com> Access in Jan. 2024.
- HABIER D, FERNANDO R, DEKKERS J. Genomic selection using low-density marker panels. **Genetics**. v. 182(1), p. 343-353. 2009.
- HRBEK, T. et al. Population genetic analysis of *Arapaima gigas*, one of the largest freshwater fishes of the Amazon basin: Implications for its conservation. **Animal Conservation**, v. 8, n. 3, p. 297–308, 2005.
- HRBEK, T.; CROSSA, M.; FARIAS, I. P. Conservation strategies for *Arapaima gigas* (Schinz, 1822) and the Amazonian várzea ecosystem. **Brazilian Journal of Biology**, v. 67, n. 4 SUPPL., p. 909–917, 2007.
- IANELLA, P. et al. SNP Discovery in the South American Freshwater Fish Tambaqui (*Colossoma macropomum*) by Deep Sequencing of Reduced Representation Library. In: **Plant and Animal Genome Conference XXIII**. San Diego, CA. p. 9-13, Anais. 2016.
- JOSHI, R. et al. Development and validation of 58K SNP-array and high-density linkage map in Nile tilapia (*O. niloticus*). **Frontiers in Genetics**, v. 9, n. OCT, 15 out. 2018.
- LACHANCE, J., TISHKOFF, S. SNP ascertainment bias in population genetic analyses: Why it is important, and how to correct it. **BioEssays** 780-786, 35(9). 2013.
- Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN]. 2013.
- MANTEL, N. The detection of disease clustering and a generalized regression approach. **Cancer Res** 27:209–220. 1967.
- MEIKLEJOHN, Kelly A. et al. Current methods, future directions and considerations of DNA-based taxonomic identification in wildlife forensics. **Forensic Science International: Animals and Environments**. v. 1, p. 100030. 2021.
- MILLER. M.P. Alleles in Space (AIS): Computer Software for the Joint Analysis of Interindividual Spatial and Genetic Information. **Journal of Heredity**. v. 96(6), p. 722-724. 2005.
- NOGUEIRA, F. et al. Genetic diversity and structuring in the arapaima (osteoglossiformes, osteoglossidae) population reveal differences between the amazon and the tocantins-araguaia basins. **Anais da Academia Brasileira de Ciências**, v. 92, n. 1, p. 1–16, 2020.

- OGDEN, R.; LINACRE, A. Wildlife forensic science: a review of genetic geographic origin assignment. **Forensic Science International: Genetics**, [S.L.], v. 18, p. 152-159, set. 2015.
- PEAKALL, R.; SMOUSE, P. E. GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. **Bioinformatics**, v. 28, n. 19, p. 2537–2539, out. 2012.
- PRITCHARD, J.K., STEPHENS, M., & DONNELLY, P. Inference of Population Structure Using Multilocus **Genotype Data**. **Genetics**, 155, 945-959. 2000.
- SCHUNTER, C. et al. SNP development from RNA-seq data in a nonmodel fish: How many individuals are needed for accurate allele frequency prediction? **Molecular Ecology Resources**, v. 14, n. 1, p. 157–165, jan. 2014.
- SLATKIN, M. Rare alleles as indicators of gene flow. **Evolution**, 39, 53-65. 1985.
- TAMURA, K., STECHER, G., KUMAR, S. MEGA11: molecular evolutionary genetics analysis version 11. **Mol. Biol. Evol.** 38 (7), 3022–3027. 2021.
- TORATI, L.S., TAGGART, J.B., VARELA, E.S. *et al.* Genetic diversity and structure in *Arapaima gigas* populations from Amazon and Araguaia-Tocantins River basins. **BMC Genet** 20, 13. 2019.
- VITORINO, C. A. et al. Genetic diversity of arapaima gigas (Schinz, 1822) (Osteoglossiformes: Arapaimidae) in the araguaia-tocantins basin estimated by ISSR marker. **Neotropical Ichthyology**, v. 13, n. 3, p. 557–568, 2015.
- VITORINO, C. A. et al. Low genetic diversity and structuring of the arapaima (Osteoglossiformes, Arapaimidae) population of the Araguaia-Tocantins basin. **Frontiers in Genetics**, v. 8, n. OCT, 24 out. 2017.
- ZHANG, H. Y. et al. Population genetic analysis of aquaculture salmonid populations in China using a 57K rainbow trout SNP array. **PLoS ONE**, v. 13, n. 8, p. 1–12, 2019.

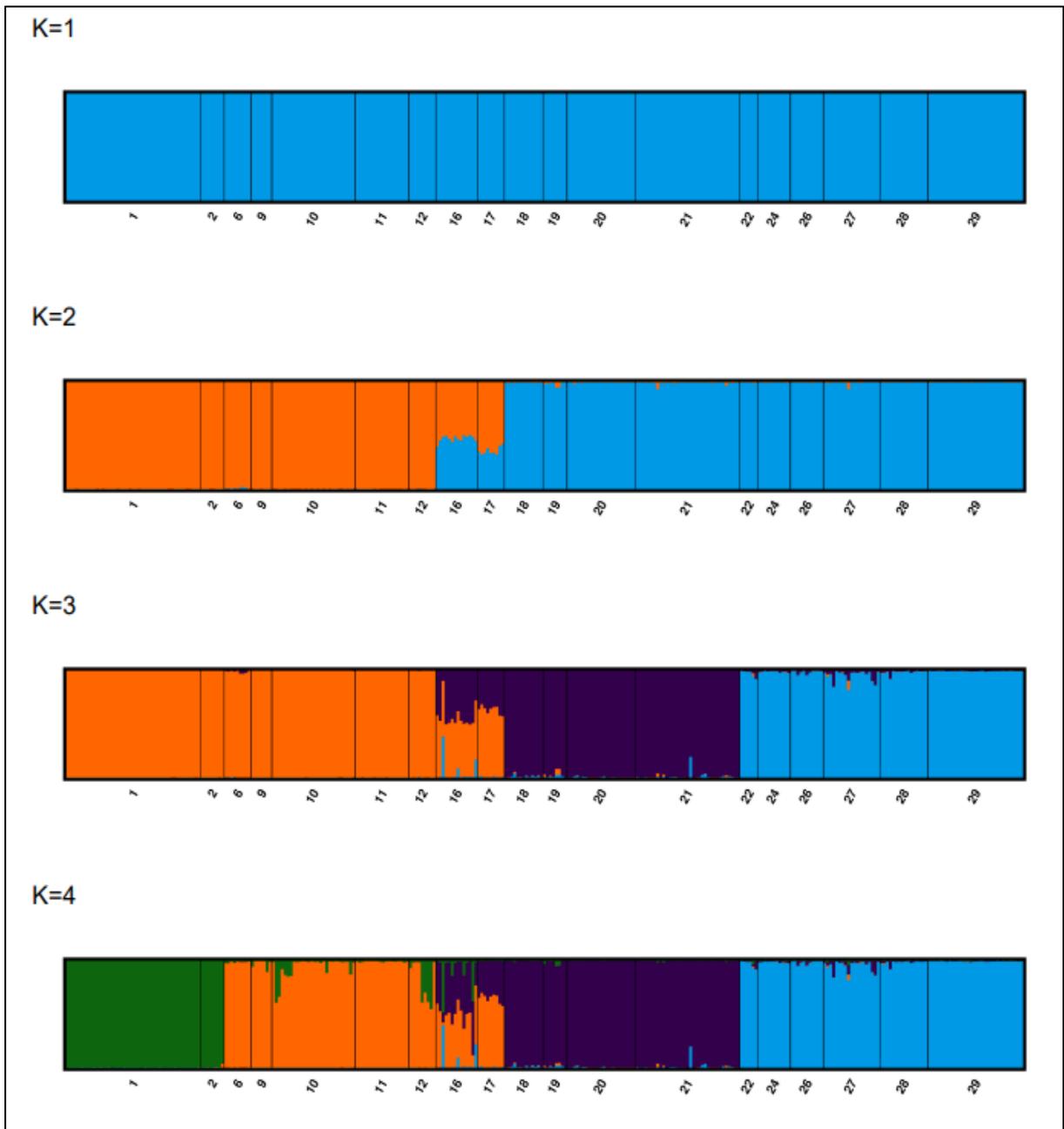
SUPPLEMENTARY MATERIAL

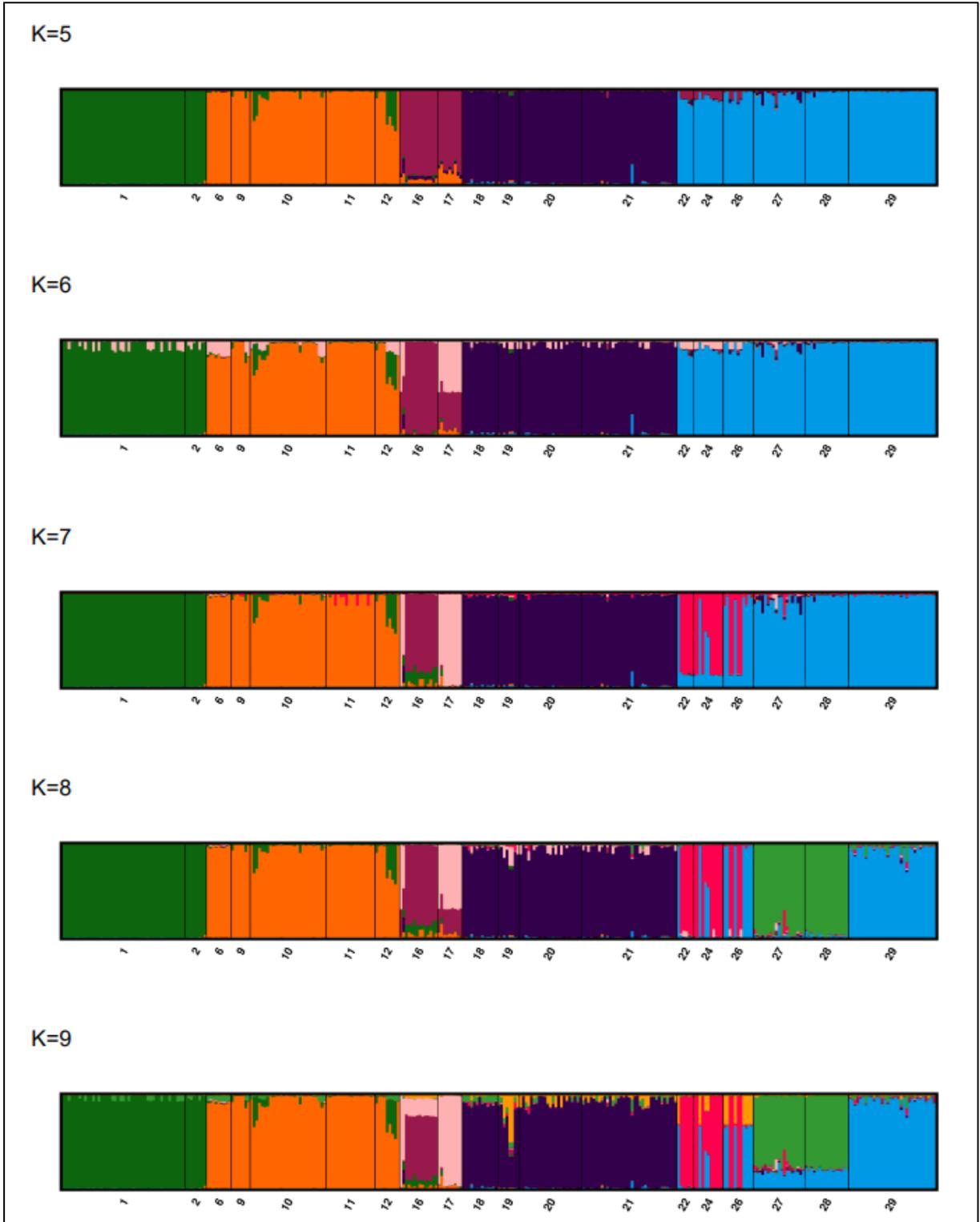
Supplementary material S1. Pairwise F_{st} values of the 19 *Arapaima gigas* populations.

Populations	MLA	SLA	RMN	ARN	NSR	NAL	SFA	PBP	ITU	ITA	MUC	PIX	SMC	POR	PUR	PER	PML	MML	LSC
	0,000																		
SLA	0,027	0,000																	
RMN	0,436	0,353	0,000																
ARN	0,421	0,351	0,121	0,000															
NSR	0,332	0,252	0,099	0,053	0,000														
NAL	0,426	0,339	0,149	0,080	0,042	0,000													
SFA	0,257	0,176	0,088	0,052	0,049	0,088	0,000												
PBP	0,551	0,373	0,329	0,306	0,415	0,390	0,310	0,000											
ITU	0,656	0,486	0,398	0,401	0,467	0,439	0,383	0,226	0,000										
ITA	0,744	0,580	0,543	0,548	0,640	0,617	0,542	0,307	0,408	0,000									
MUC	0,783	0,618	0,582	0,584	0,682	0,657	0,577	0,311	0,429	0,045	0,000								
PIX	0,691	0,541	0,511	0,515	0,598	0,571	0,512	0,298	0,387	0,018	0,024	0,000							
SMC	0,645	0,517	0,490	0,497	0,562	0,541	0,492	0,290	0,379	0,020	0,028	0,001*	0,000						
POR	0,809	0,658	0,625	0,624	0,712	0,699	0,615	0,341	0,475	0,208	0,221	0,211	0,209	0,000					
PUR	0,777	0,622	0,600	0,601	0,688	0,670	0,593	0,364	0,476	0,246	0,249	0,244	0,243	0,004*	0,000				
PER	0,789	0,640	0,615	0,613	0,700	0,680	0,609	0,396	0,494	0,268	0,275	0,259	0,263	0,067*	0,074	0,000			
PML	0,708	0,548	0,523	0,522	0,613	0,590	0,519	0,322	0,410	0,168	0,180	0,174	0,165	0,124	0,159	0,139	0,000		
MML	0,747	0,596	0,568	0,572	0,657	0,637	0,566	0,358	0,461	0,197	0,218	0,197	0,194	0,153	0,177	0,152	0,037	0,000	
LSC	0,695	0,570	0,550	0,552	0,613	0,598	0,546	0,374	0,456	0,226	0,243	0,230	0,225	0,147	0,164	0,140	0,084	0,076	0,000

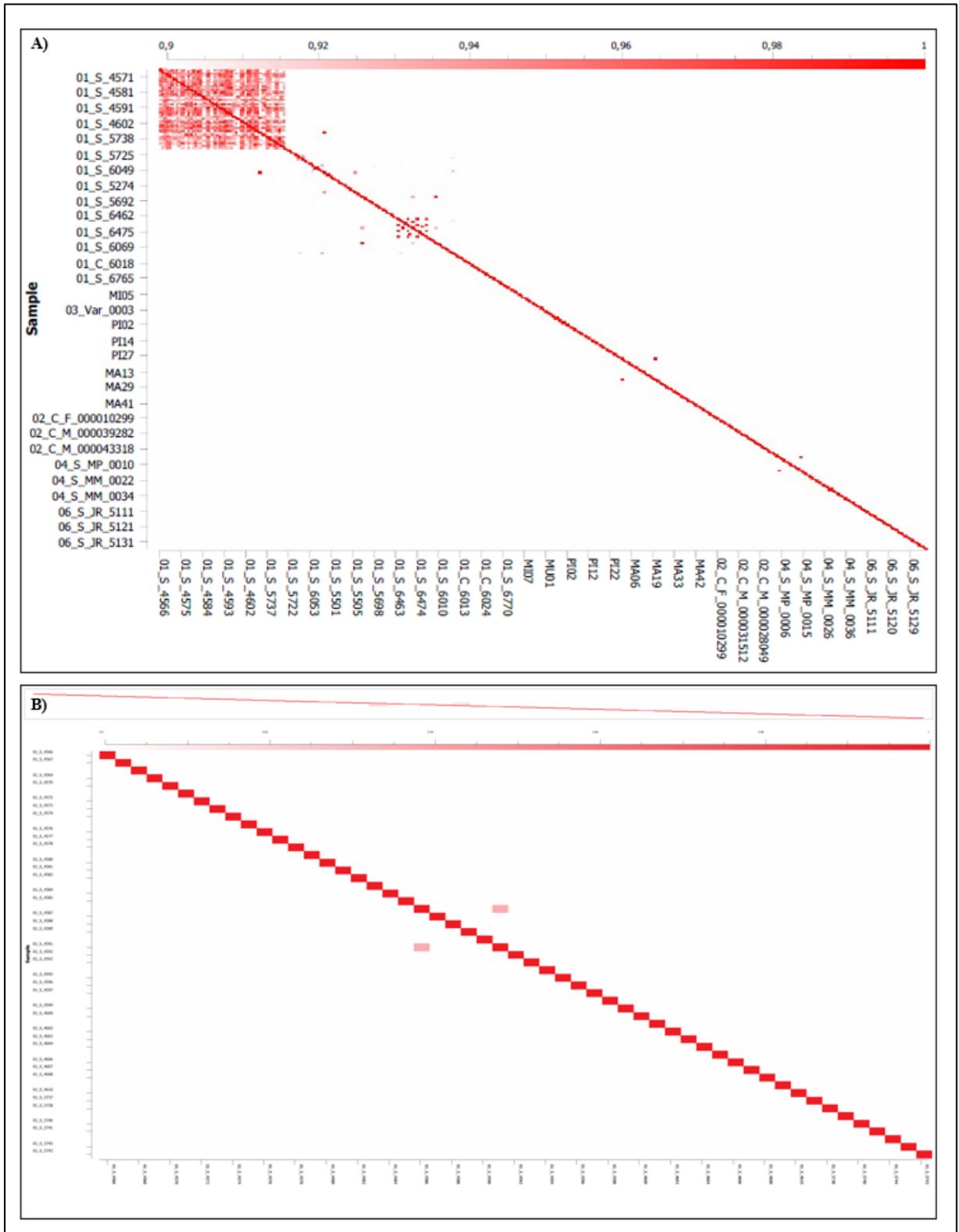
*Indicates P -values >0.05

Supplementary material S2. Structure plots for $K=2$ to $k=9$ for all 19 *Arapaima gigas* analyzed. Populations code legend: 1 MLA, 2 SLA, 6 RMN, 9 ARN, 10 NSR, 11 NAL, 12 SFA, 16 PBP, 17 ITU, 18 ITA, 19 MUC, 20 PIX, 21 SMC, 22 POR, 24 PUR, 26 PER, 27 PML, 28 MML, 29 LSC





Supplementary material S3. (A) IBD graphs for all *Arapaima gigas* populations analyzed in this study. (B) IBD graph of population MLA and SLA after MAF filter removing all monomorphic markers.



CAPÍTULO 3
CONSIDERAÇÕES FINAIS

3.1 CONSIDERAÇÕES FINAIS

O Pirarucu (*Arapaima gigas*) está listado como espécie ameaçada de extinção pela CITES e categorizado como "Data Deficient" pela IUCN devido a exploração comercial intensa e perda de habitat por degradação humana. Embora a piscicultura tenha reduzido a superexploração da espécie, os desafios na reprodução em cativeiro, como o acasalamento consanguíneo, limitam o potencial de seu cultivo. Estudos sobre a estrutura populacional e diversidade genética são cruciais para suportar estratégias de conservação e cultivo da espécie em cativeiro, visando a manutenção dos estoques naturais e suprimindo a demanda deste importante recurso alimentar no mercado, especialmente na região norte do Brasil.

Neste trabalho, identificamos que as populações de Pirarucu ao longo das bacias Tocantins-Araguaia e Amazônica têm diferenciação genética considerável, que pode ser explicada por fatores como o isolamento por distância, causando baixo fluxo gênico entre as populações e maior estruturação das subpopulações. A estrutura genética das populações analisadas indica a existência de ao menos duas populações ancestrais em comum, um para a bacia Tocantins-Araguaia e outro para a bacia Amazônica. O maior nível de diversidade genética foi observado nas populações da Bacia Amazônica, o que já era esperado, já que esta região possui maior área conservada em relação a bacia Tocantins-Araguaia, que vem sendo forte e intermutavelmente afetada pelas alterações antrópicas no Cerrado.

Os dados genéticos gerados neste estudo se mostraram úteis para auxiliar em estratégias de manejo genético de peixes reprodutores em cativeiro e para a conservação dos estoques naturais de Pirarucu nas bacias hidrográficas estudadas. Essa espécie é essencial para a manutenção do equilíbrio do ecossistema e para as comunidades tradicionais locais ao longo das bacias Amazônica e Tocantins-Araguaia que dependem desse recurso pesqueiro para subsistência. A prioridade em programas de conservação é manter os estoques selvagens mitigando o risco de extinção da espécie. Para esse propósito, pesquisas que tragam mais informações robustas das características genéticas das populações de Pirarucu são fundamentais para compreender a fundo a estrutura dessas populações e entender os fatores que as colocam em risco.