



**Universidade Federal de São Carlos  
Centro de Ciências Biológicas e da Saúde  
Departamento de Genética e Evolução**

# **O IMPACTO DA PRESENÇA DE SISTEMAS DE CROMOSSOMOS SEXUAIS MÚLTIPLOS NA DIFERENCIAÇÃO GENÉTICA**

---

A família Erythrinidae (Teleostei, Characiformes)  
como modelo

**Dissertação de Mestrado**

**Fernando Henrique Santos de Souza  
São Carlos - 2022**

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE  
DEPARTAMENTO DE GENÉTICA E EVOLUÇÃO

FERNANDO HENRIQUE SANTOS DE SOUZA

O IMPACTO DA PRESENÇA DE SISTEMAS DE  
CROMOSSOMOS SEXUAIS MÚLTIPLOS NA  
DIFERENCIAÇÃO GENÉTICA: A família Erythrinidae  
(Teleostei, Characiformes) como modelo.

SÃO CARLOS -SP  
2022

FERNANDO HENRIQUE SANTOS DE SOUZA

O IMPACTO DA PRESENÇA DE SISTEMAS DE CROMOSSOMOS SEXUAIS  
MÚLTIPLOS NA DIFERENCIAÇÃO GENÉTICA: A família Erythrinidae (Teleostei,  
Characiformes) como modelo.

Dissertação apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular ao Departamento de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos, para obtenção do título de mestre em Genética e Evolução.

Orientador: Marcelo de Belo Cioffi  
Coorientador: Manolo Fernandes Perez

São Carlos-SP  
2022

Souza, Fernando Henrique Santos de

O impacto da presença de sistemas de cromossomos sexuais múltiplos na diferenciação genética : A família Erythrinidae (Teleostei, Characiformes) como modelo / Fernando Henrique Santos de Souza -- 2022.  
94f.

Dissertação (Mestrado) - Universidade Federal de São Carlos, campus São Carlos, São Carlos

Orientador (a): Marcelo de Bello Cioffi

Banca Examinadora: Marcelo de Bello Cioffi, Manolo Fernandes Perez, Isabel Aparecida da Silva Bonatelli, Ricardo Utsunomia

Bibliografia

1. Evolução. 2. Diversidade genética. 3. Cromossomos sexuais. I. Souza, Fernando Henrique Santos de. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática  
(SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Ronildo Santos Prado - CRB/8 7325



# UNIVERSIDADE FEDERAL DE SÃO CARLOS

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

---

## Folha de Aprovação

---

Defesa de Dissertação de Mestrado do candidato Fernando Henrique Santos de Souza, realizada em 24/02/2022.

### Comissão Julgadora:

Prof. Dr. Marcelo de Bello Cioffi (UFSCar)

Prof. Dr. Manolo Fernandez Perez (UFSCar)

Profa. Dra. Isabel Aparecida da Silva Bonatelli (UNIFESP)

Prof. Dr. Ricardo Utsunomia (UFRRJ)

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

## **DEDICATÓRIA**

Dedico este trabalho aos meus pais, Marlene e José e à minha querida Daniele.

## AGRADECIMENTO

Agradeço aos meus pais Marlene e José pelo amor, carinho, incentivo constante na busca de mais conhecimento e apoio nos projetos que decidi realizar.

A todos os docentes que contribuíram de maneira direta ou indireta para minha formação, fornecendo uma base sólida de conhecimento que permitiu a execução deste e de outros trabalhos, especialmente aos docentes do Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular.

Ao meu prezado orientador Prof. Dr. Marcelo, pelo apoio constante durante minha trajetória desde a IC até a pós-graduação, por sua organização e planejamento, por toda dedicação em sua orientação, pelos seus ensinamentos e pelo seu grande exemplo em ensino e pesquisa.

Ao querido Dr. Manolo, por sua exímia coorientação, pelo seu empenho em transmitir seus conhecimentos de bioinformática a mim desde minha iniciação científica, por sua capacidade de resolver qualquer problema, mesmo que não possua solução e pela paciência de ensinar e corrigir quantas vezes forem necessárias.

À minha namorada Daniele, pelo apoio ao longo da pós-graduação, pelas sugestões e ajuda na revisão deste trabalho, e por sempre estar ao meu lado me incentivando a alcançar todos os meus objetivos.

A meus amigos que me apoiaram durante a pós-graduação, principalmente ao Pedro, Alessandro e Renata.

A toda equipe do Laboratório de Citogenética de Peixes, por todo o apoio, pelas importantes discussões e sugestões ao longo do projeto.

À Dra. Deborah Charlesworth por sua gigantesca contribuição desde a recomendação de determinadas análises até a discussão dos resultados.

À Universidade Federal de São Carlos e ao CCBS pela oportunidade de realizar este trabalho e pela estrutura necessária.

À Alexandra Elbakyan e a equipe do Sci-Hub pela democratização do conhecimento.

À Fapesp pelo processo nº 2019/25009-7, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

E ao Conselho Nacional de Desenvolvimento Científico e Tecnológico pelo financiamento do projeto.

“Quando servimos ao nosso propósito, somos postos de lado. Mas os genes são habitantes do tempo geológico: os genes são para sempre.”

Richard Dawkins

## RESUMO

A biodiversidade da região Neotropical ainda permanece pouco compreendida, e, embora a alopatria seja um dos principais fatores que favorecem a especiação, os rearranjos cromossômicos e a presença de cromossomos sexuais heteromórficos também podem desempenhar papéis importantes. Os rearranjos cromossômicos envolvendo cromossomos sexuais tem o potencial de reduzir eventos de introgressão e facilitar o isolamento reprodutivo. Desta forma, o objetivo deste trabalho foi analisar a influência da alopatria, dos rearranjos cromossômicos e da presença de sistemas de cromossomos sexuais múltiplos na diversificação genética das espécies de peixes Neotropicais *Hoplias malabaricus* e *Erythrinus erythrinus* por meio de análises citogenéticas e genômicas. Analisamos populações simpátricas e alopátricas que diferem por suas i) características cromossômicas (denominadas cariomorfos), ii) distribuições simpátricas versus alopátricas e iii) presença e ausência de cromossomos sexuais heteromórficos. Ao todo, foram analisadas 10 localidades de *H. malabaricus* que apresentaram 6 cariomorfos distintos e 3 localidades de *E. erythrinus* com 2 cariomorfos distintos. Os diferentes cariomorfos apresentaram variações tanto em número de cromossomos, quanto em presença ou ausência de sistemas de cromossomos sexuais simples e múltiplos. Os resultados indicaram que apesar dos cariomorfos representarem grupos geneticamente divergentes, eles podem ser melhor compreendidos quando analisados em conjunto com sua distribuição geográfica e condições ambientais. O sistema de cromossomos sexuais múltiplos  $X_1X_2Y$ , presente em um dos cariomorfos de ambas as espécies, se mostrou associado a uma maior diferenciação genética interpopulacional. Para *H. malabaricus*, a presença de tal sistema sexual pode ter contribuído para níveis de diferenciação genética entre os cariomorfos A e D em sintopia equivalentes aos níveis de diferenciação de populações alopátricas do cariomorfo A. De maneira geral, os resultados de ambas as espécies sugerem que a presença de sistemas de cromossomos neo-sexuais pode ter contribuído para o surgimento e fixação de diferenças interpopulacionais, que podem ter contribuído de forma relevante para o surgimento e manutenção do isolamento reprodutivo. Desta forma, uma das hipóteses é de que a ocorrência de rearranjos cromossômicos, especialmente os relacionados aos cromossomos sexuais, podem ter contribuído para o processo de especiação para estas espécies e amplificado a diferenciação advinda da alopatria.

**Palavras-chave:** Evolução. Diversidade genética. Peixes. Cromossomos sexuais. DArTSeq.

## ABSTRACT

The biodiversity of the Neotropics remains poorly understood, and although allopatry is one of the main factors that promote speciation, chromosomal rearrangements and the evolution of sex chromosomes can also play important roles. Chromosomal rearrangements involving sex chromosomes have the potential to reduce introgression events and facilitate reproductive isolation. Thus, the objective of this work was to analyze the influence of allopatry, chromosomal rearrangements and the presence of multiple sex chromosome systems on the genetic diversification of the Neotropical fish species *Hoplias malabaricus* and *Erythrinus erythrinus* through cytogenetic and genomic analyses. We analyzed sympatric and allopatric populations that differ by i) chromosomal characteristics (named karyomorphs), ii) sympatric versus allopatric distributions, and iii) presence and absence of heteromorphic sex chromosomes. Altogether, 10 localities of *H. malabaricus* were analyzed that presented six distinct karyomorphs and three localities of *E. erythrinus* with two distinct karyomorphs. The different karyomorphs showed variations both in the number of chromosomes and in the presence or absence of multiple sex chromosome systems. The results indicated that although karyomorphs represent genetically divergent groups, they can be better understood when analyzed in conjunction with their geographic distribution and environmental conditions. The multiple  $X_1X_2Y$  sex chromosome system, present in one of the karyomorphs of both species, was associated with greater interpopulation genetic differentiation compared to the effect of allopatry. For *H. malabaricus*, the presence of such a multiple system may have contributed to levels of genetic differentiation between karyomorphs A and D in syntopy equivalent to the levels of differentiation of allopatric populations of karyomorph A. In general, the results for both species suggest that the presence of neo sex chromosome systems may have contributed to the emergence and fixation of interpopulational differences, which may have contributed significantly to the emergence and fixation of reproductive isolation. Thus, one of the hypotheses is that the occurrence of chromosomal rearrangements, especially those related to sex chromosomes, may have accelerated the speciation process for these species and amplified the differentiation resulting from allopatry.

**Keywords:** Evolution. Genetic diversity. Fish. Sex chromosomes. DART-Seq.

## LISTA DE ILUSTRAÇÕES

- Figura 1** - Potencial distribuição de *Hoplias malabaricus*. Modificado de <https://www.gbif.org/pt/species/2352230> utilizando apenas dados de coleta. ....22
- Figura 2** - Espécime de *Hoplias malabaricus* .....23
- Figura 3** – Idiograma parcial dos cariomorfos de *Hoplias malabaricus*. Fonte: Bertollo et al. 2000. ....26
- Figura 4** - Potencial distribuição de *Hoplias malabaricus*. Modificado de <https://www.gbif.org/pt/species/2352207> utilizando apenas dados de coleta. ....27
- Figura 5** - Espécime de *Erythrinus erythrinus*. Fonte: <https://www.gbif.org/pt/occurrence/624180106>. ....28
- Figura 6** - Modelo evolutivo dos cariomorfos A-D em *E. erythrinus* proposto por Bertollo et al. 2004. Fonte: Bertollo et al. 2004. ....29
- Figura 7** - Mapa da América do Sul indicando as localidades com amostras de *Hoplias malabaricus* analisadas. As cores indicam as diferentes bacias hidrográficas no Brasil, listadas na legenda, e os círculos coloridos indicam os pontos de coleta, com os cariomorfos .....43
- Figura 8** - Análise de Coordenadas Principais. Indivíduos de diferentes cariomorfos estão representados por cores distintas de acordo com a Figura 7. ....47
- Figura 9** - Resultado do FastStructure para K=6. Cada barra vertical representa um indivíduo, os cariomorfos e locais de coleta estão representados abaixo; a cor das barras representa a população na qual a análise classificou cada indivíduo. ....47
- Figura 10** - Valores de  $D_{XY}$  par-a-par por ponto de coleta são representados na diagonal superior, e os valores de  $D_A$  par-a-par na diagonal inferior. Valores mais altos são verdes, e conforme os valores diminuem a cor muda para amarelo, laranja e então vermelho para os menores valores. ....49
- Figura 11** - Gráfico apresentando dois eixos principais da RDA. Os círculos representam os indivíduos, com cores indicando seu local de amostragem e cariomorfo, conforme a legenda à direita. Os fatores ambientais considerados significativos são representados por linhas azuis, com setas indicando a localização na RDA em que os dados estão concentrados. ....50
- Figura 12** - Mapa da América do Sul indicando os pontos de coleta para *E. erythrinus*. As cores indicam as diferentes bacias hidrográficas brasileiras, coloridas de acordo com a legenda. Os círculos coloridos indicam os pontos de coleta, com os respectivos cariomorfos coloridos da seguinte forma: A em verde e D em vermelho. O idiograma à direita indica os cariótipos parciais de cada cariomorfo e o seu sistema de cromossomos sexuais, quando presente. ....61

- Figura 13** - Cariótipos de *E. erythrinus* do cariomorfo A da localidade A2 analisados pelo método de coloração com Giemsa (A) e FISH com sequências de DNAr 18S (verde) e 5S (vermelho) como sondas (B). Tamaho da barra: 5  $\mu$ m..... 65
- Figura 14** - Cromossomos mitóticos de machos de *E. erythrinus* após o CGH – comparação entre cariomorfos. (a-d) A sonda genômica derivada de machos do cariomorfo A1 (b) e A2 (c) mapeadas contra cromossomos de macho do cariomorfo A1. (e-h) Sondas genômicas derivadas de machos d (f) e D (g) mapeadas contra cromossomos de machos do cariomorfo A1. (i-l) Sondas genômicas de macho do cariomorfo de D (j) e A2 (k) mapeadas contra cromossomos de macho do cariomorfo D. Na primeira coluna (a,e,i): imagens de DAPI (azul); Segunda e terceira colunas (b,c,f,g,j,k): padrão de hibridização utilizando as sondas derivadas de macho (vermelho) de cada cariomorfo analisado; Quarta coluna (d,h,l): imagens sobrepostas de ambas as sondas genômicas e coloração com DAPI. As regiões genômicas comuns de ambos os cariomorfos comparados são representadas em amarelo. Tamano da barra = 10  $\mu$ m. .... 66
- Figura 15** - (A) Análise de Coordenadas Principais. Indivíduos estão representados em cores diferentes: A1 em verde escuro; A2 em verde claro; D em vermelho. (B) Resultado do fastStructure para K=2. Cada barra vertical representa um indivíduo, cariomorfos e locais de coleta estão indicados abaixo das barras, a cor das barras indica a população na qual cada indivíduo foi classificado. (C) Resultado do ConStruct para K=2. Cada barra vertical representa um indivíduo, cariomorfos e locais de coleta estão indicados abaixo das barras..... 67
- Figura 16** - Idiogramas representativos dos cariomorfos A1, A2, e D de *E. erythrinus* evidenciando a distribuição do DNAr 18S (verde) e DNAr 5S (vermelho). Os cromossomos sexuais encontram-se representados no retângulo..... 70
- Figura 17** - Resultado da análise ConStruct para o valor de K=3. Fonte: Autor (2022). ..... 90
- Figura 18** - Resultado da análise ConStruct para o valor de K=5. Fonte: Autor (2022). ..... 90

## LISTA DE TABELAS

<b>Tabela 1</b> - Indivíduos analisados, seus cariomorfos, número de cromossomos diploides (2n), Sistema de cromossomos sexuais, Código de identificação dos grupos (indicando o local de coleta e o cariomorfo), número de indivíduos analisados citogeneticamente e número de indivíduos sequenciados.....	42
<b>Tabela 2</b> - Diversidade genética estimada para <i>H. malabaricus</i> por cariomorfo. Os tamanhos amostrais são indicados, bem como a diversidade haplotípica ( $H_d$ ), Diversidade nucleotídica ( $\pi$ ), Theta de Watterson por sitio ( $\theta_w$ ) e D de Tajima ( $D$ ). .....	48
<b>Tabela 3</b> - Diversidade genética estimada para <i>H. malabaricus</i> por ponto de coleta. Os tamanhos amostrais são indicados, bem como a diversidade haplotípica ( $H_d$ ), Diversidade nucleotídica ( $\pi$ ), Theta de Watterson por sitio ( $\theta_w$ ) e D de Tajima ( $D$ ). .....	48
<b>Tabela 4</b> - Lista de indivíduos de <i>E. erythrinus</i> analisados, evidenciando cariomorfo, número de cromossomos (2n), sistema de cromossomos sexual, código do grupo (indicando o ponto de coleta e o cariomorfo), número de indivíduos analisados citogeneticamente, número de indivíduos sequenciados e referências.....	60
<b>Tabela 5</b> - Diversidade genética estimada por ponto de coleta. Estão evidenciados os tamanhos amostrais, a diversidade haplotípica ( $H_d$ ), diversidade nucleotídica ( $\pi$ ), Theta de Watterson por sitio ( $\theta_w$ ), e D de Tajima ( $D$ ). .....	68
<b>Tabela 6</b> - Valores de $D_{XY}$ par-a-par por ponto de coleta indicados na diagonal superior, e valores de $D_A$ par-a-par na diagonal inferior. ....	68
<b>Quadro 1</b> - Valores de $D_{XY}$ par-a-par por cariomorfo.....	91
<b>Quadro 2</b> - Valores de $D_A$ par-a-par por cariomorfo .....	91
<b>Quadro 3</b> - Lista de marcadores potencialmente sob seleção.....	92
<b>Quadro 4</b> - $D_{XY}$ entre cariomorfos .....	94
<b>Quadro 5</b> - $D_a$ entre cariomorfos .....	94
<b>Quadro 6</b> - $F_{ST}$ entre cariomorfos .....	94

## SUMÁRIO

1. INTRODUÇÃO GERAL .....	14
1.1 BIODIVERSIDADE NEOTROPICAL .....	14
1.2 ICTIOFAUNA NEOTROPICAL.....	18
1.3 PROCESSOS QUE FAVORECEM A ESPECIAÇÃO.....	19
1.4 A FAMÍLIA ERYTHRINIDAE .....	21
1.4.1 <i>Hoplías malabaricus</i> .....	22
1.4.2 <i>Erythrinus erythrinus</i> .....	27
1.5 OBSERVAÇÕES ACERCA DA METODOLOGIA .....	29
1.5.1 Métodos de redução de complexidade do genoma.....	29
1.5.2 O uso de SNPs na genética de populações .....	30
1.6 PANORAMA GERAL .....	31
2. OBJETIVOS.....	33
2.1 OBJETIVOS GERAIS .....	33
2.2 OBJETIVOS ESPECÍFICOS.....	34
3. MATERIAIS E MÉTODOS.....	35
3.1 ANÁLISES DE ESTRUTURA POPULACIONAL .....	35
4. RESULTADOS E DISCUSSÃO .....	37
CAPÍTULO 1 .....	38
RESUMO .....	39
INTRODUÇÃO.....	40
MÉTODOS.....	41
AMOSTRAGEM DE ESPÉCIMES.....	41
SEQUENCIAMENTO E FILTRAGEM.....	43
DETECÇÃO E EXCLUSÃO DOS MARCADORES POTENCIALMENTE SOB SELEÇÃO.....	44
DIVERSIDADE GENÉTICA.....	44
ANÁLISES DE ESTRUTURA POPULACIONAL.....	44
RESULTADOS .....	46
DADOS CITOGENÉTICOS, SEQUENCIAMENTO, FILTRAGEM E DETECÇÃO DE	

MARCADORES SOB SELEÇÃO .....	46
ESTRUTURA POPULACIONAL.....	46
DIVERSIDADE GENÉTICA E DIFERENCIAÇÃO GENÉTICA .....	47
ASSOCIAÇÃO GENÓTIPO-AMBIENTE.....	49
DISCUSSÃO .....	50
DIVERSIDADE GENÉTICA E ESTRUTURA POPULACIONAL.....	51
CROMOSSOMOS SEXUAIS MÚLTIPLOS E O PROCESSO DE ESPECIAÇÃO.....	53
CONSIDERAÇÕES FINAIS .....	55
REFERÊNCIAS BIBLIOGRÁFICAS .....	55
CAPÍTULO 2 .....	56
RESUMO .....	57
INTRODUÇÃO.....	58
MATERIAIS E MÉTODOS.....	60
AMOSTRAGEM DOS ESPÉCIMES .....	60
PREPARAÇÃO CROMOSSÔMICA E HIBRIDIZAÇÃO IN SITU FLUORESCENTE (FISH) PARA MAPEAMENTO DO DNAR.....	61
HIBRIDIZAÇÃO GENÔMICA COMPARATIVA (CGH).....	61
MICROSCOPIA E PROCESSAMENTO DE IMAGEM.....	62
SEQUENCIAMENTO E FILTRAGEM DOS DADOS.....	62
DETECÇÃO DE MARCADORES POTENCIALMENTE SOB SELEÇÃO.....	63
DIVERSIDADE GENÉTICA.....	63
ANÁLISES DE ESTRUTURA POPULACIONAL.....	64
RESULTADOS .....	65
CARIÓTIPOS E MAPEAMENTO DE DNA RIBOSSÔMICO .....	65
HIBRIDIZAÇÃO GENÔMICA COMPARATIVA.....	66
SEQUENCIAMENTO, PREPARO DOS DADOS E BUSCA DE MARCADORES SOB SELEÇÃO .....	67
ESTRUTURA POPULACIONAL.....	67
DIVERSIDADE GENÉTICA.....	68
DISCUSSÃO .....	69
DIVERSIDADE CROMOSSÔMICA ENTRE POPULAÇÕES DE <i>E. ERYTHRINUS</i> .....	69
DIVERSIDADE GENÉTICA E ESTRUTURA POPULACIONAL.....	71
SISTEMA DE CROMOSSOMOS SEXUAIS MÚLTIPLOS E O PROCESSO DE	

ESPECIAÇÃO.....	73
CONCLUSÕES.....	74
REFERÊNCIAS BIBLIOGRÁFICAS .....	74
5. CONSIDERAÇÕES FINAIS .....	75
REFERÊNCIAS .....	77
APÊNDICE 1 – Resultados da análise ConStruct para <i>H. malabaricus</i> .....	90
APÊNDICE 2 – Valores de $D_{XY}$ e $D_a$ par-a-par por cariomorfo para <i>H. malabaricus</i> .....	91
APÊNDICE 3 – Lista de marcadores potencialmente sob seleção para <i>H. malabaricus</i> .....	92
APÊNDICE 4 – Resultados de $D_{XY}$ e $D_a$ e $F_{ST}$ entre cariomorfos de <i>E. erythrinus</i> .....	94

# 1. INTRODUÇÃO GERAL

## 1.1 BIODIVERSIDADE NEOTROPICAL

A região Neotropical é uma das seis grandes zonas biogeográficas definidas inicialmente com base em distribuição e abundância de aves por Sclater (1858) e posteriormente validada por Wallace (1876) com base em outros grupos animais. A definição mais ampla da região Neotropical engloba a América Central e a América do Sul até a região sul do México (COX, 2001). A região apresenta um histórico de eventos de vicariância e dispersão que resultam de uma complexa história geológica e climática (RULL e CARNAVAL, 2020). Grandes eventos como a inversão do curso de grandes rios, incursões marinhas e eventos de formação de montanhas ajudaram a moldar a biodiversidade neotropical (HUBERT e RENNO, 2006). Estes eventos foram lentamente moldando as paisagens da região Neotropical. As incursões marinhas ocorreram desde o Paleoceno até o Mioceno, os eventos que geraram os Andes iniciaram-se durante a transição entre o Oligoceno e o Mioceno (RULL e CARNAVAL, 2020). Adicionalmente, houve o fechamento do Istmo do Panamá há cerca de 3,5 milhões de anos, e glaciações que ocorreram do fim do Plioceno ao Pleistoceno (RULL e CARNAVAL, 2020). Desta forma, estes eventos contribuíram para a composição da grande biodiversidade de organismos presentes na região.

A biodiversidade da região Neotropical, que possui sua maior parte localizada entre os trópicos, é notável, tanto para espécies vegetais quanto para microrganismos e animais (ANTONELLI et al. 2018). Entretanto, as causas responsáveis por este alto nível de biodiversidade são alvo de debate há longa data, sem que haja total consenso. Duas hipóteses clássicas versam sobre esta problemática: i) a hipótese dos trópicos como berço de diversidade, primeiramente apresentada por Dobzhansky (1950) e popularizada por Axelrod (1970) e ii) a hipótese dos trópicos como um museu de diversidade, apresentada por Stebbins (1974). Apesar da complexidade das ideias iniciais, atualmente ao referir-se ao termo museu de diversidade o principal objetivo é ressaltar baixas taxas de extinção em determinado local, e ao referir-se a berço de diversidade o intuito é de indicar que determinada região apresenta taxas mais elevadas de especiação.

O principal padrão de riqueza de espécies que está relacionado à biodiversidade das regiões tropicais é o Gradiente Latitudinal de Biodiversidade (LDG). O LDG é o padrão biogeográfico que prevê que quanto mais próximo dos trópicos maior o número de espécies, e de forma complementar, quanto mais próximo dos polos, menor a biodiversidade encontrada

(MESEGUER e CONDAMINE, 2019). O LDG foi evidenciado para diversos organismos como plantas, microrganismos, insetos, vertebrados terrestres, animais de água doce e até mesmo animais marinhos (HILLEBRAND, 2004). Muitos autores tentam até a atualidade compreender o LDG sob a luz das hipóteses de berço e museu.

A dicotomia dos trópicos como berço ou museu foi amplamente discutida por Vasconcelos, O'meara e Beaulieu (2021). Conforme revisado pelos autores, a hipótese dos trópicos como berço afirma que a região Neotropical possui grande capacidade de criar diversidade, baseando-se na premissa de que as características de determinada espécie são uma combinação de adaptações otimizadas para o ambiente em que ela evoluiu. De acordo com as variações do ambiente ao longo do tempo, que podem estar associadas principalmente a eventos tectônicos ou climáticos, surge uma pressão que leva os organismos a migrarem ou a se adaptarem às novas condições ambientais. Ademais, estes eventos que alteram a paisagem possuem a capacidade de separar grupos de indivíduos de determinada espécie, promovendo a diversificação. Para que uma região possa desempenhar o papel de berço de diversidade, existem dois principais requisitos que precisam ser cumpridos. O primeiro é que o ambiente no qual determinada espécie habita seja espacialmente instável, aumentando a possibilidade de separação da população em grupos menores, favorecendo desta forma o isolamento reprodutivo (VASCONCELOS; O'MEARA; BEAULIEU, 2021). A segunda característica necessária é que a região apresente uma diversidade de ambientes, com diferentes características ecológicas, dessa forma existe pressão seletiva selecionando os indivíduos a adaptarem-se a ambientes distintos (VASCONCELOS; O'MEARA; BEAULIEU, 2021). Adicionalmente, devido a extensão da região tropical e sua diversidade de ambientes, regiões menores dentro do território tropical podem desempenhar papéis específicos. Para borboletas da tribo Brassolini (Nymphalidae, Satyrinae) a região da Mesoamérica e da parte noroeste dos Andes desempenharam papel de berço, com taxas de especiação crescentes nos últimos 2 milhões de anos (MATOS-MARAVÍ et al. 2021).

De forma mais conservadora, como proposto inicialmente por Stebbins (1974), os berços apresentam um conjunto de fatores que propiciam o aumento tanto das taxas de especiação quanto de extinção. Isso ocorre pois as populações menores e isoladas que surgem têm mais dificuldade para se manter. Estas populações estão sujeitas a maior perda de diversidade por deriva genética e conseqüentemente estão mais suscetíveis à extinção. De acordo com Stebbins (1974) as regiões classificadas com berço apresentam maior surgimento de novas espécies e maior diversificação, entretanto apresentam taxas mais baixas de imigração

e de preservação de fósseis (VASCONCELOS; O'MEARA; BEAULIEU, 2021). A hipótese de berço é apoiada pela riqueza atual de espécies, pelos padrões de distribuição dessas espécies e pelo endemismo de certos táxons na região tropical (MCKENNA e FARRELL, 2006).

Em contraponto, a hipótese dos trópicos como museu discutida por Stebbins (1974) sugere que as espécies presentes em ambientes com características de berço de diversidade, que estão sob constante processo de diversificação irão em determinado momento adaptar-se a áreas de estabilidade. As áreas de estabilidade apresentam características opostas às áreas berço, desta forma a diversificação nessas áreas é mais difícil de ocorrer (VASCONCELOS; O'MEARA; BEAULIEU, 2021). O que reforça a hipótese dos trópicos como museus é principalmente a presença de evidências paleontológicas que evidenciam uma radiação bastante antiga para diversos clados (MCKENNA e FARRELL, 2006). Adicionalmente, regiões biogeográficas dentro da região tropical podem desempenhar o papel de museu. Para as borboletas da tribo Brassolini, a região da Mata Atlântica desempenhou o papel de museu de diversidade, as espécies ali presentes evoluíram sob baixas taxas de especiação desde 12 milhões de anos atrás (MATOS-MARAVÍ et al. 2021).

Os museus de diversidade são então classificados como regiões que apresentam baixas taxas de extinção, mas também baixas taxas de especiação, principalmente devido à ausência de pressão seletiva para adaptação a novos ambientes e situações, devido a estabilidade dessas áreas. Desta forma, os museus de diversidade não apresentam altas taxas de diversificação das espécies, porém apresentam uma maior preservação de fósseis e maiores taxas de migração (VASCONCELOS; O'MEARA; BEAULIEU, 2021).

Tendo em vista os conceitos modernos de berço e museu, existe atualmente uma ampla discussão acerca da classificação da região Neotropical em um desses conceitos. Com o advento das tecnologias de sequenciamento de DNA e o avanço das análises biogeográficas, existem mais dados disponíveis para inferir se uma ou ambas as hipóteses são mais adequadas para a biodiversidade da região Neotropical. Entretanto, alguns autores consideram que esta dicotomia é apenas uma simplificação da problemática acerca da maior biodiversidade localizada nos trópicos, enquanto outros autores chegam à conclusão de que os trópicos são tanto berço quanto museu (VASCONCELOS; O'MEARA; BEAULIEU, 2021).

Com o objetivo de investigar as hipóteses de berço e museu, um amplo estudo de filogeografia utilizou dados moleculares de 295 espécimes de todas as 21 subfamílias existentes de formigas (Hymenoptera, Formicidae) e dados de 45 fósseis combinados. Os autores chegaram à conclusão de que os Neotrópicos desempenharam tanto papel de berço, com alto

surgimento de novas espécies, quanto de museu, onde espécies antigas persistem ao longo do tempo (MOREAU e BELL, 2013). Nesta mesma linha de raciocínio, Jablonski, Roy e Valentine (2006) sugeriram que os trópicos são simultaneamente um berço e museu de espécies ao pontuar que o (LDG) pode ser descrito de forma matemática, de uma maneira onde diversas combinações de diversos fatores resultam no padrão do LDG. De acordo com Jablonski, Roy e Valentine (2006) o LDG pode ser resumido matematicamente como a relação entre  $D_T$  que é a diversidade nos trópicos e  $D_F$ , a diversidade fora dos trópicos. Inicialmente, os autores definem  $D_T = O_T - E_T + I_T$  sendo  $O_T$  a taxa de origem nos trópicos,  $E_T$  a taxa de extinção nos trópicos e  $I_T$  a taxa de imigração para os trópicos, e  $D_F = O_F - E_F + I_F$  onde  $O_F$  é a taxa de origem fora dos trópicos,  $E_F$  a taxa de extinção fora dos trópicos e  $I_F$  a taxa de imigração para fora dos trópicos. A partir destas equações é possível observar que várias combinações possíveis destas variáveis podem resultar no padrão esperado de  $D_T > D_F$ , portanto é uma tarefa bastante complexa definir qual combinação exata de características e fatores é realmente responsável pelo LDG.

Considerando as dificuldades em estimar as taxas de extinção com base no registro fóssil, além do desafio causado pela variação dessas taxas ao longo do tempo, os autores então propõem o modelo “*Out of the tropics*” (JABLONSKI; ROY; VALENTINE, 2006). Este modelo considera que  $O_T > O_F$ ,  $E_T \leq E_F$  e  $I_T < I_F$ , ou seja, existem maiores taxas de origem nos trópicos, porém estes táxons expandem-se para regiões fora dos trópicos enquanto também mantém sua distribuição inicial. De acordo com a hipótese, este processo teria sido causado principalmente devido a mudanças de temperatura.

Realmente, a temperatura média no planeta antes do Eoceno era mais alta, posteriormente, com o processo de resfriamento global que ocorreu durante o período do Eoceno-Oligoceno as espécies adaptadas ao clima quente foram negativamente afetadas (MESEGUER e CONDAMINE, 2019). Nas regiões mais distantes dos polos, as espécies adaptadas ao clima quente sofreram pressão seletiva, levando à perda de biodiversidade nessas regiões (MESEGUER e CONDAMINE, 2019). Adicionalmente, essas regiões receberam menos eventos de dispersão partindo dos trópicos (MESEGUER e CONDAMINE, 2019).

Ademais, o evento do fechamento do Istmo do Panamá permitiu o Grande Intercâmbio Americano (evento de colonização da América do Sul por táxons originários da América do Norte) que em conjunto com o soerguimento dos Andes permitiu que os táxons que migraram pudessem permanecer nessa região devido ao surgimento de habitats montanhosos (RULL e CARNAVAL, 2020). Dessa forma, tanto os eventos climáticos, geológicos e de dispersão

contribuíram para a composição da biodiversidade Neotropical.

Apesar da riqueza de teorias e dos diversos estudos acerca das causas da biodiversidade tropical, diversos autores focaram em apenas um aspecto ou em uma teoria principal para diversificação nos trópicos, e essa abordagem não é a mais recomendada para elucidar a biodiversidade Neotropical. É necessário abordar a temática através de diferentes pontos de vista, com equipes multidisciplinares que quebrem paradigmas e explorem a temática de maneira ampla (RULL e CARNAVAL, 2020).

## 1.2 ICTIOFAUNA NEOTROPICAL

Os peixes representam um grupo de vertebrados extremamente biodiverso, que ocupam uma grande variedade de ambientes como os oceanos, rios e lagos. Estima-se que o grupo é formado por mais de 30.000 espécies, sendo o grupo de vertebrados mais especioso (NELSON et al. 2016). Dentre os peixes, os teleósteos representam aproximadamente 96% das espécies conhecidas, cerca de 27.000 espécies. Dados empíricos apontam que um evento de duplicação genômica ocorreu nos teleósteos entre 300 a 320 milhões de anos, este evento antecede à grande radiação observada neste grupo (CHRISTOFFELS et al. 2004; JAILLON et al. 2004; VANDEPOELE et al. 2004). Deste modo, as evidências apontam que esse evento de duplicação pode ter contribuído para a enorme quantidade de espécies dos teleósteos (GLASAUER e NEUHAUSS, 2014). Após uma duplicação, uma das cópias do gene evolui sob pressão seletiva relaxada, aumentando assim a taxa de substituição de aminoácidos, e posteriormente uma das cópias poderá eventualmente resultar em uma nova função ou na perda de função, se transformando num pseudogene (WALSH, 1995; LYNCH e CONERY, 2000). Deste modo, sabe-se que a redundância gênica pode fornecer uma vantagem adaptativa pela possibilidade de que uma das cópias, sob determinadas condições, possa adquirir uma nova função (PERRY et al. 2007). Se o efeito da duplicação de um ou poucos genes tem o potencial de resultar em mudanças importantes para o processo evolutivo da espécie, a duplicação de um genoma completo tem o potencial de gerar grandes mudanças em termos de evolutivos. Apesar de muitos genes perderem sua função rapidamente após um evento de duplicação completa do genoma (INOUE et al. 2015), alguns serão mantidos ativos e podem atuar como matéria prima para a geração de novas funções (GLASAUER e NEUHAUSS, 2014). A retenção da função de genes duplicados ocorre de forma enviesada, favorecida por genes chave em vias de sinalização e regulação da expressão gênica (MCGRATH *et al.* 2014). Portanto, é muito plausível que este evento de duplicação tenha contribuído grandemente para a especiação dos teleósteos.

A ictiofauna neotropical apresenta alta biodiversidade, contendo cerca de 5.200 espécies de peixes (FRICKE; ESCHMEYER; VAN DER LAAN, 2021). Entretanto, a vasta biodiversidade ictiológica da região Neotropical ainda permanece em grande parte inexplorada. Sabe-se que eventos geológicos na região Neotropical possivelmente contribuíram para a diversificação dos peixes. Os rios que localizam-se sobre o Escudo Brasileiro (uma região de formação rochosa antiga e tectonicamente estável) passaram por eventos neotectônicos que incluem eventos de captura de cabeceira (ROSSETTI e VALERIANO, 2007). Estes eventos de captura de cabeceira podem alterar a composição da fauna dos rios, e permitir a expansão e o isolamento de populações (ALBERT, 2011). Adicionalmente, processos como o recuo erosivo que ocorreu na margem leste do sudeste brasileiro e as deformações tectônicas permitiram intercâmbio entre espécies de drenagens costeiras e de planaltos (RIBEIRO, 2006).

Ainda que estes eventos geológicos sejam conhecidos, o uso de dados moleculares e os avanços na avaliação dos mecanismos genéticos que ocorreram em conjunto com esses eventos permitirão maior compreensão acerca das causas da biodiversidade da ictiofauna neotropical.

### 1.3 PROCESSOS QUE FAVORECEM A ESPECIAÇÃO

A diminuição de fluxo gênico entre espécies emergentes resulta no processo evolutivo conhecido como especiação. Entretanto, identificar o processo de especiação em alguns grupos pode ser extremamente desafiador, o que pode ser percebido quando observa-se a grande quantidade de conceitos de espécies existentes, com base em características e considerações distintas acerca dos organismos. De acordo com o conceito unificado de espécies, estas podem ser definidas como linhagens de populações que evoluem de maneira independente (DE QUEIROZ, 2007). Deste modo, compreender os mecanismos pelos quais o isolamento reprodutivo é estabelecido torna-se muito importante para que os limites entre as espécies possam ser inferidos de maneira mais robusta.

Dentre os diferentes processos de especiação, a especiação alopátrica é geralmente considerada a forma mais comum. De acordo com o mecanismo sugerido, o isolamento reprodutivo surge como um produto da divergência entre as populações. Dessa forma, a especiação alopátrica pode ser induzida por reconfigurações da paisagem, como a separação de massas continentais através do movimento das placas tectônicas, formação de cadeias montanhosas, ou da emergência de rios que isolam populações, sendo assim uma das principais causas da diversificação biológica (DE QUEIROZ, 2007). Neste aspecto, a região Neotropical apresenta uma história climática e geológica bastante complexa, marcada por diversos eventos tectônicos e climáticos que ocorreram principalmente durante o Neógeno e o Pleistoceno

(HOORN et al. 2010; GARZON-ORDUÑA; BENETTI-LONGHINI; BROWER, 2014) que podem ter contribuído de forma relevante para a grande biodiversidade encontrada nessa região.

Nas últimas décadas houve diversos avanços na compreensão dos tipos mais comuns de isolamento reprodutivo, suas causas e seus impactos na evolução. O isolamento pós-zigótico foi um dos tipos de isolamento reprodutivos mais investigados. Em animais, a maior causa do isolamento pós-zigótico são as incompatibilidades genéticas, dentre estas, os rearranjos cromossômicos e o desenvolvimento de cromossomos sexuais podem desempenhar um papel importante (BATESON, 1909; DOBZHANSKY, 1933; MULLER, 1942; COYNE e ORR, 2004). Os rearranjos cromossômicos apresentam o potencial de limitar a introgressão, facilitando desta forma a origem e a manutenção do isolamento reprodutivo através da supressão de recombinação (FARIA e NAVARRO, 2010; SICHOVÁ et al. 2015). A importância do surgimento de sistemas de cromossomos sexuais múltiplos na formação de barreiras reprodutivas entre espécies também já foi evidenciada (KITANO e PEICHEL, 2012).

Alterações estruturais, como fusões, translocações entre cromossomos autossomos e cromossomos sexuais (X, Y, Z, W) criam um sistema de cromossomos neo-sexuais. Contrário ao processo de recombinação lento e gradual que ocorre em sistemas de cromossomos sexuais simples (XY, ZW, XO), o surgimento de um sistema de cromossomos neo-sexuais pode acarretar uma rápida redução ou mesmo a supressão de recombinação em regiões próximas aos pontos de quebra (SEMBER et al. 2021). Adicionalmente, além da formação de grupos de ligação entre genes que antes encontravam-se em diferentes cromossomos, estes rearranjos tem a capacidade de aumentar a quantidade de genes ligados ao sexo e de acelerar a acumulação de incompatibilidades genéticas entre as populações (VIEIRA; COELHO; VIEIRA, 2003). Os genes que desempenham um papel importante no isolamento reprodutivo entre populações podem acumular-se em cromossomos neo-sexuais, podendo levar à especiação (KITANO e PEICHEL, 2012). Como exemplo, o surgimento de um sistema de cromossomos sexuais em subespécies de bicho-da-seda *Samia cynthia ssp.* (Saturniidae), contribuíram para a formação de barreiras reprodutivas que estão diretamente relacionadas a especiação deste grupo (YOSHIDO et al. 2011). Isso é suportado pela instabilidade intraespecífica de sistemas de cromossomos sexuais múltiplos, que contrasta fortemente com a dinâmica evolutiva dos autossomos. De certa forma, talvez o aspecto mais relevante dos rearranjos que envolvem cromossomos sexuais são as consequências meióticas geradas por esse processo. Estes eventos acarretam em uma barreira pós-zigótica muito eficiente para o indivíduo que carrega o

rearranjo, devido às particularidades mecânicas durante o processo meiótico. Ao contrário do que ocorre em sistemas simples, nos quais a segregação meiótica envolve apenas um par de cromossomos, trivalentes meióticos característicos são formados em sistemas múltiplos como resultado de fusões, fissões ou translocações. Isso leva a uma alta probabilidade de formação de gametas desbalanceados (ZHANG et al. 2021).

#### 1.4 A FAMÍLIA ERYTHRINIDAE

A ordem dos Characiformes é uma das maiores ordens de peixes de água doce do mundo, com cerca de 2.000 espécies (NELSON et al. 2006). Dentro desta ordem, a família Erythrinidae é composta por apenas três gêneros que são endêmicos da América do Sul e América Central: *Hoplias* Gill, 1903; *Hoplerythrinus* Gill, 1895 e *Erythrinus* Scopoli, 1977 (OYAKAWA, 2003). O gênero *Hoplias* é dividido em dois grupos principais, o grupo das traíras gigantes (grupo *Hoplias lacerdae*) que engloba *H. australis*, *H. brasiliensis*, *H. curupira*, *H. intermedius*, e *H. lacerdae*, e o grupo das traíras comuns (grupo *Hoplias malabaricus*) formado por *H. malabaricus* (OYAKAWA e MATTOX, 2009). Ademais, existe uma outra espécie de traíra de grande porte, *H. aimara*, que não é incluída no grupo das grandes traíras devido a características morfológicas (OYAKAWA e MATTOX, 2009). Ao contrário dos grupos *H. lacerdae* e *H. aimara*, que passaram recentemente por revisões taxonômicas (MATTOX; TOLEDO-PIZA; OYAKAWA, 2006; OYAKAWA e MATTOX, 2009), a taxonomia do grupo *H. malabaricus* é bastante imprecisa e necessita de revisão. O gênero *Erythrinus*, por sua vez, apresenta duas espécies: *E. kessleri* e *E. erythrinus* (OYAKAWA, 2003) sendo *E. erythrinus* uma das espécies menos estudadas da família.

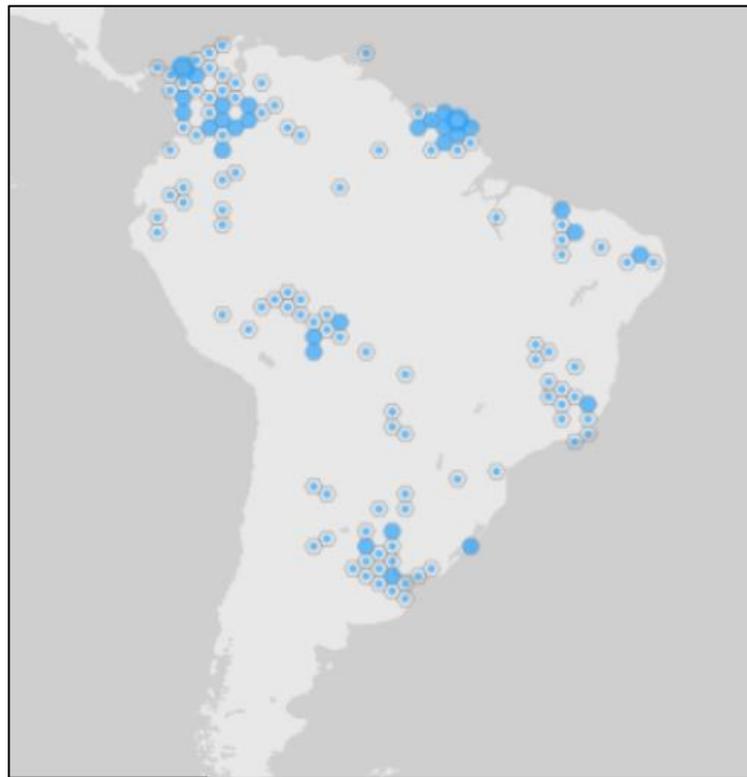
A maior parte dos estudos de caráter genético em *H. malabaricus* e *E. erythrinus* possuem um enfoque citogenético, vêm sendo realizados há mais de 20 anos, e ainda assim continuam levantando questionamentos a respeito dessas espécies. Além dos trabalhos responsáveis pela caracterização dos cariomorfos distintos de *H. malabaricus* (BERTOLLO et al. 2000) e *E. erythrinus* (BERTOLLO et al. 2004), pesquisas posteriores revelaram as possíveis relações evolutivas entre alguns cariomorfos, bem como os possíveis processos de surgimento de alguns cariomorfos e seus sistemas sexuais (BERTOLLO et al. 2004; CIOFFI et al. 2011; FREITAS et al, 2018; DE OLIVEIRA et al. 2018).

As investigações genéticas que implementaram dados de marcadores genéticos ou sequências são escassos em *H. malabaricus* e inexistentes em *E. erythrinus*. Dentre os trabalhos de caráter genético, Utsunomia e colaboradores (2014) identificaram um indivíduo triploide, um caso raro híbrido entre os cariomorfos A e D com 61 cromossomos, e também evidenciaram

alta diferenciação entre a sequência *cyt b* dos cariomorfos A e D (10,4%) de acordo com valores de distância genética calculados utilizando o modelo de substituição de Kimura K2P, diferença quase igual a encontrada entre o cariomorfo D de *H. malabaricus* e outra espécie, *H. microleptis* (UTSUNOMIA *et al.* 2014). Utilizando dados do fragmento COI de indivíduos dos cariomorfos A e C de *H. malabaricus*, Ferreira e colaboradores (2021) identificaram a presença de diferentes linhagens na Bacia do Amazonas, fornecendo suporte a hipótese de que *H. malabaricus* representa na verdade um complexo de espécies.

#### 1.4.1 *Hoplias malabaricus*

*H. malabaricus* é a espécie que possui a distribuição mais ampla na família Erythrinidae. Esta espécie está presente em muitas das bacias sul-americanas (OYAKAWA, 2003). A espécie é encontrada desde as Guianas até a Argentina, bem como na maioria das bacias brasileiras. Devido a sua ampla distribuição, a espécie é conhecida por diversos nomes comuns como traíra, dientudo, taraíra, maturaqué.



**Figura 1** - Potencial distribuição de *Hoplias malabaricus*. Modificado de <https://www.gbif.org/pt/species/2352230> utilizando apenas dados de coleta.

A espécie é caracterizada taxonomicamente por apresentar o corpo com forma cilíndrica e alongada, com o ponto mais alto do indivíduo localizado na parte anterior da nadadeira dorsal. (BIFI, 2013). O perfil dorsal destes indivíduos pode variar, podendo ser reto, levemente

côncavo ou mesmo levemente convexo da região do focinho até a margem posterior dos parietais. O perfil ventral varia e pode ser convexo ou levemente convexo da região do focinho até a origem da nadadeira anal e côncavo até a nadadeira caudal (BIFI, 2013). De forma geral, a coloração dos indivíduos desta espécie varia do preto ao marrom podendo apresentar regiões em coloração marrom-amarelada (BIFI, 2013). A quantidade de escamas da linha lateral, bem como outras características taxonômicas podem variar de acordo com o cariomorfo que o espécime apresenta (BIFI, 2013).



**Figura 2** - Espécime de *Hoplias malabaricus*. Barra = 2cm

*H. malabaricus* tem preferência por ambientes com baixa velocidade de correnteza e habita lagoas, rios e regiões de pastagens alagadas (COSTA NOVAES e CARVALHO, 2011). A alimentação da espécie é composta principalmente por peixes de pequeno porte (COSTA NOVAES e CARVALHO, 2011), porém, *H. malabaricus* é considerado um predador oportunista, ou seja, não possui um tipo de presa específico e aproveita-se das oportunidades que surgem para capturar as presas disponíveis (OYAKAWA, 2003). A espécie apresenta hábitos sedentários (OYAKAWA, 2003) e sua reprodução ocorre ao longo do ano, apresentando maior intensidade durante o verão e a primavera (COSTA NOVAES e CARVALHO, 2011). As características predatórias desta espécie fazem com que a mesma desempenhe importante papel na mediação de populações de outras espécies de peixes, principalmente em pequenas lagoas durante o período de seca (PETRY et al. 2010). De acordo com a Lista Vermelha de Espécies Ameaçadas da IUCN (LYONS, 2020), *H. malabaricus* é considerada uma espécie “Least Concern”, ou seja, não ameaçada. Entretanto, a espécie é bastante visada em tanto na pesca de subsistência quanto na pesca artesanal (COSTA NOVAES

e CARVALHO, 2011).

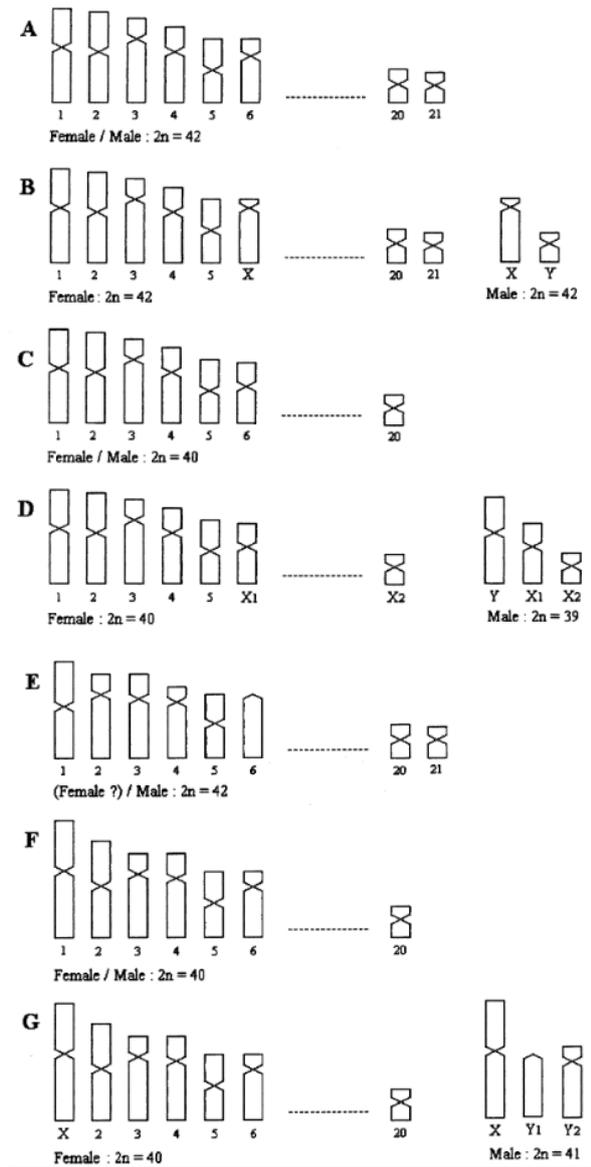
A distribuição bastante ampla de *H. malabaricus* favorece a presença de divergências populacionais. Do ponto de vista citogenético, a espécie possui um nível surpreendente de diversidade cariotípica, e até o momento sete cariomorfos distintos foram encontrados. Os cariomorfos identificados foram nomeados de A até G e cada um deles apresenta características cariotípicas específicas (BERTOLLO, 2007; CIOFFI et al. 2012). Algumas das características que se diferenciam entre os cariomorfos são o número de cromossomos, a presença de diferentes sistemas de cromossomos sexuais bem como a morfologia dos cromossomos (CIOFFI et al. 2012).

O cariomorfo A apresenta 42 cromossomos tanto no macho quanto na fêmea e possui uma distribuição ampla podendo ser encontrado na Argentina, Uruguai e na maior parte do Brasil. Seus cromossomos são metacêntricos e submetacêntricos, sem indícios da presença de um sistema de cromossomos sexuais (CIOFFI et al. 2009). O cariomorfo B apresenta uma distribuição mais específica, habitando principalmente a região da bacia do Rio Doce e a região do Rio Iguaçu. O cariomorfo B é caracterizado pela presença de 42 cromossomos em ambos os sexos, e é semelhante ao cariomorfo A com exceção da presença de um sistema de cromossomos sexuais simples classificado como XX/XY. O cariomorfo C possui 40 cromossomos tanto em machos quanto em fêmeas e não apresenta sistema de cromossomos sexuais. Este cariomorfo apresenta uma distribuição que vai do norte do Brasil até a Argentina.

O cariomorfo D é um dos mais emblemáticos de *H. malabaricus*, as fêmeas apresentam 40 cromossomos enquanto os machos possuem 39. O cariomorfo D apresenta um sistema de cromossomos neo-sexuais múltiplos do tipo  $X_1X_1X_2X_2/X_1X_2Y$ , o que o torna um dos cariomorfos mais distintos. Os indivíduos que apresentam este cariomorfo possuem distribuição restrita à bacia do Alto Paraná (BERTOLLO et al. 2000). De acordo com estudos citogenéticos de bandeamento realizados previamente, uma hipótese para a formação do sistema neo-sexual foi proposta. No estudo realizado por Bertollo e colaboradores (1997) foi evidenciado que o cromossomo X deste cariomorfo surgiu graças a um rearranjo cromossômico, mais precisamente a fusão entre um cromossomo e o braço curto de outro cromossomo que provavelmente assemelha-se ao cromossomo  $X_1$  atual (BERTOLLO et al. 1997). O cariomorfo E é composto por 42 cromossomos, mas pode ser diferenciado por apresentar o par número seis na forma acrocêntrica, bem como o par um consideravelmente maior do que nos cariomorfos A e B. Este cariomorfo apresenta distribuição bastante restrita, na região do Rio Trombetas. O cariomorfo F apresenta 40 cromossomos em ambos os sexos

assim como o cariomorfo C, e até o momento não foi identificado nenhum sistema de cromossomos sexuais. Os indivíduos deste cariomorfo estão presentes principalmente na região leste do Brasil. O cariomorfo F pode ser diferenciado do cariomorfo C por apresentar um grande par de cromossomos metacêntricos. As fêmeas do cariomorfo G apresentam 40 cromossomos enquanto os machos possuem 41. Este cariomorfo, apresenta um sistema de cromossomos sexuais múltiplos do tipo  $XX/XY_1Y_2$ . Os indivíduos deste cariomorfo tem uma distribuição mais restrita à Bacia Amazônica.

A grande diversidade dos cariomorfos de *H. malabaricus* é a sua característica mais interessante (**Figura 3**). Estes cariomorfos podem ser classificados em dois grupos principais de acordo com suas relações evolutivas. Baseado nos estudos de Bertollo e colaboradores (2000) o cariomorfo B surgiu a partir do cariomorfo A. O cariomorfo D surgiu do cariomorfo C (CIOFFI et al. 2011). Os cariomorfos F e G, por sua vez derivam provavelmente do cariomorfo E. No cariomorfo F o mecanismo proposto indica que fusões em tandem acarretaram na formação do sistema sexual simples do tipo  $XX/XY$  (FREITAS et al. 2017). O cariomorfo G também surgiu provavelmente a partir de fusões em tandem, gerando o sistema sexual múltiplo do tipo  $XX/XY_1Y_2$  (DE OLIVEIRA et al. 2017). Os cariomorfos A, B, C e D compõem o Grupo I e os cariomorfos E, F e G formam o Grupo II.



**Figura 3** – Idiograma parcial dos cariomorfos de *Hoplias malabaricus*. Fonte: Bertollo et al. 2000.

A presença de cariomorfos com número de cromossomos distintos e sistemas sexuais distintos sugere que *H. malabaricus* é na verdade um complexo de espécies, uma hipótese que vem sendo amplamente defendida recentemente (BERTOLLO, 2007; CIOFFI et al. 2012; DA ROSA et al. 2012; MARTINS et al. 2013). Do ponto de vista taxonômico, Bifi (2013) identifica diversas características distintas que variam entre os diferentes cariomorfos de *H. malabaricus*, evidenciando desta forma que existe também certa diversidade morfológica entre estes indivíduos. Adicionalmente, alguns outros fatores reforçam esta hipótese como a presença de cariomorfos distintos em simpatria e até mesmo em sintopia sem que haja presença de híbridos, com apenas uma exceção. O único caso de híbrido registrado até o momento foi o de um indivíduo triploide coletado em uma região de sintopia de cariomorfos (UTSUNOMIA et

al. 2014). Até mesmo os cariomorfos que seriam mais instáveis do ponto de vista meiótico, como o cariomorfo D geram gametas balanceados (BERTOLLO e MESTRINER, 1998). Com base em todos os aspectos apresentados até o momento, há indícios bastante consistentes de que *H. malabaricus* é um complexo de espécies que é composto por diferentes unidades evolutivas (BERTOLLO et al. 2000).

#### 1.4.2 *Erythrinus erythrinus*

O gênero *Erythrinus* apresenta apenas duas espécies conhecidas, *E. erythrinus* e *E. kessleri* (OYAKAWA, 2003). Este é o gênero menos estudado da família Erythrinidae, havendo pouca literatura acerca de características citogenéticas e genéticas. A espécie *E. kessleri* apresenta distribuição mais específica na região do estado da Bahia, em contraponto *E. erythrinus* apresenta uma distribuição mais ampla pela América do Sul (OYAKAWA, 2003).



**Figura 4** - Potencial distribuição de *Erythrinus erythrinus*. Modificado de <https://www.gbif.org/pt/species/2352207> utilizando apenas dados de coleta.

Popularmente conhecido como Jejú, ou Morobá, *E. erythrinus* geralmente é encontrado em ambientes pantanosos, graças à modificação em sua bexiga natatória a espécie apresenta capacidade de sobreviver em ambientes menos oxigenados (BOUJARD, 1997). Os indivíduos

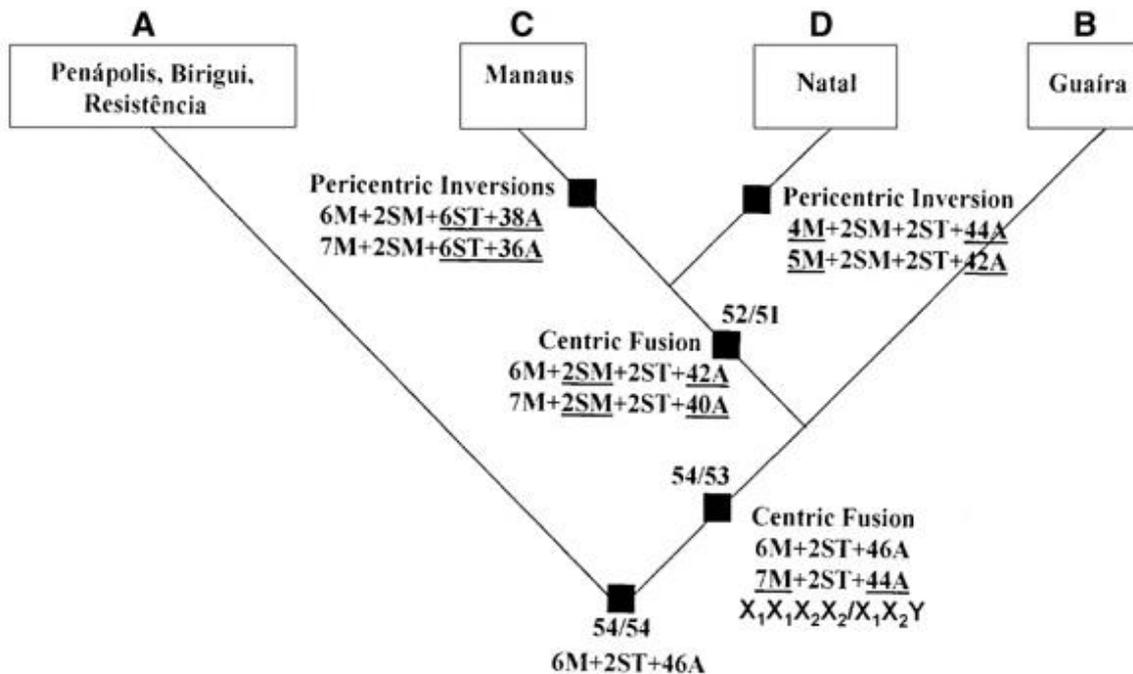
alimentam-se principalmente de peixes, bem como de pequenos insetos. Em um estudo realizado por Bertollo e colaboradores (2004) foram identificados cariomorfos com características distintas, bem como aquelas presentes em *H. malabaricus*. Ao todo, quatro cariomorfos distintos foram descritos e nomeados de A a D.



**Figura 5** - Espécime de *Erythrinus erythrinus*. Fonte: <https://www.gbif.org/pt/occurrence/624180106>.

O cariomorfo A apresenta um cariótipo composto por  $2n=54$  com 6 cromossomos metacêntricos, 2 cromossomos subtelocêntricos e 46 cromossomos acrocêntricos (6M 2ST e 46A), não possui sistema de cromossomos sexuais identificado (BERTOLLO et al. 2004). O cariomorfo B possui um sistema sexual múltiplo do tipo  $X_1X_1X_2X_2/X_1X_2Y$ , os machos apresentam  $2n=53$  enquanto as fêmeas apresentam  $2n=54$  (BERTOLLO et al. 2004). O cariomorfo C também apresenta sistema sexual múltiplo do tipo  $X_1X_1X_2X_2/X_1X_2Y$  tendo os machos  $2n=51$  (7M 2SM 6ST e 36A) e as fêmeas  $2n=52$  (6M 2SM 6ST e 38A) (BERTOLLO et al. 2004). O cariomorfo D apresenta sistema sexual múltiplo do tipo  $X_1X_1X_2X_2/X_1X_2Y$  assim como os cariomorfos B e C, e possui machos com  $2n=51$  e fêmeas com  $2n=52$  assim como o cariomorfo C. A principal diferença entre os cariomorfos C e D são a estrutura cariotípica onde o cariomorfo D apresenta nos machos 5 cromossomos metacêntricos, 2 submetacêntricos, 2 subtelocêntricos e 41 acrocêntricos. Nas fêmeas o cariomorfo D é caracterizado por 4 cromossomos metacêntricos, 2 submetacêntricos, 2 subtelocêntricos e 44 acrocêntricos (BERTOLLO et al. 2004). Através da análise citogenética dos quatro cariomorfos distintos de *E. erythrinus* um modelo para as alterações cromossômicas que

resultaram nos cariomorfos atuais foi proposta (BERTOLLO *et al.* 2004).



**Figura 6** - Modelo evolutivo dos cariomorfos A-D em *E. erythrinus* proposto por Bertollo et al. 2004. Fonte: Bertollo et al. 2004.

De acordo com o modelo proposto, o cariomorfo A é a forma cariotípica mais basal em *Erythrinus erythrinus*. Após um evento de fusão centrípeta que deu origem ao sistema sexual múltiplo, o cariomorfo B surgiu. Posteriormente, um evento adicional de fusão centrípeta ocorreu, seguido por eventos individuais de inversão pericêntrica, estes eventos deram origem respectivamente aos cariomorfos C e D.

## 1.5 OBSERVAÇÕES ACERCA DA METODOLOGIA

### 1.5.1 Métodos de redução de complexidade do genoma

Os SNPs (*Single nucleotide polymorphisms*) são a forma de variação genética que pode ser encontrada em maior quantidade nos organismos, portanto métodos que permitam explorar estes SNPs são de extrema relevância (JORDAN et al. 2002). Os métodos de sequenciamento existentes antes do desenvolvimento de métodos de redução de complexidade modernos tornavam muito inviáveis os estudos de genética de populações devido ao alto custo (DAVEY e BLAXTER, 2010). As metodologias de sequenciamento de nova geração permitiram que genomas completos pudessem ser sequenciados, entretanto essa não é a melhor abordagem para a genética de populações. Nos estudos de genética de populações é necessário obter marcadores

genéticos de diversos indivíduos amostrados para que seja possível realizar as análises propriamente (DAVEY e BLAXTER, 2010). O uso de marcadores como microssatélites foi muito comum antes do advento das técnicas de redução de complexidade, e continua sendo utilizado ainda hoje, entretanto uma de suas desvantagens é a baixa quantidade de microssatélites estudados. Portanto, era necessário desenvolver técnicas que possibilitassem a recuperação de um grande número de marcadores genéticos de diversos indivíduos sem que houvesse custos excessivos.

Uma das primeiras e mais conhecidas técnicas de sequenciamento por redução de complexidade é a técnica de RADSeq. Esta técnica inicialmente utiliza uma enzima de restrição para fragmentar o genoma, e posteriormente liga os adaptadores de sequenciamento e a sequência de *barcode* nas regiões a ser sequenciadas, após amplificação por PCR o material é sequenciado e analisado (DAVEY e BLAXTER, 2010). As vantagens das técnicas de sequenciamento por redução de complexidade incluem a possibilidade de analisar organismos que não possuam um genoma de referência, e a obtenção de um alto número de marcadores moleculares que podem ser utilizados em estudos de genética de populações, filogenética e filogeografia.

Diversas variações do RADSeq surgiram ao longo do tempo, uma delas é a ddRAD que utiliza duas enzimas de restrição ao invés de uma, e permite um maior controle do tamanho dos fragmentos obtidos além de reduzir a quantidade de *reads* necessários devido à sua maior capacidade de cobrir o genoma de forma pouco repetitiva (BRANT et al. 2012). Após o surgimento do ddRAD, novas técnicas que utilizam duas enzimas de restrição surgiram para atender propósitos específicos. Uma dessas técnicas é a de sequenciamento DArTseq que utiliza as enzimas de restrição PstI e SbfI. A SbfI é uma enzima de corte frequente, enquanto que a PstI é sensível a regiões CG metiladas e não cliva as regiões altamente repetitivas do genoma, dessa maneira ocorre o enriquecimento de sequências informativas (KILIAN et al. 2012).

Entre as vantagens dos métodos de sequenciamento por redução de complexidade que utilizam duas enzimas de restrição estão o custo reduzido que permite o sequenciamento de mais indivíduos, e a possibilidade de ajustar a cobertura do genoma ou mesmo o tamanho dos fragmentos obtidos (ILLUMINA, 2022). Entre as desvantagens pode-se citar a possibilidade de haver *gaps* na cobertura do genoma analisado, bem como a necessidade de utilizar um DNA de alta qualidade (ILLUMINA, 2022).

### **1.5.2 O uso de SNPs na genética de populações**

A estimativa de parâmetros de diversidade genética e o agrupamento em populações é

um importante passo na obtenção de informações de populações naturais (DESALLE e AMATO, 2004). Antes da popularização das técnicas de sequenciamento de nova geração os marcadores do tipo microssatélite foram amplamente utilizados em estudos populacionais devido a sua distribuição ao longo do genoma e ao alto grau de polimorfismo (BAUMUNG, SIMIANER e HOFFMANN, 2004; SCHLÖTTERER, 2004). Entretanto os microssatélites são obtidos em baixa quantidade, e conforme a obtenção de SNPs foi democratizada o uso dos microssatélites passou a diminuir. Com a evolução das metodologias de sequenciamento é possível amostrar uma grande quantidade de marcadores com alto custo-benefício (SCHLÖTTERER, 2004). Os SNPs apresentam uma distribuição mais uniforme ao longo do genoma e são mais abundantes (MORIN, LUIKART e WAYNE, 2004). Adicionalmente, a capacidade dos mecanismos de sequenciamento por redução de complexidade de gerar SNPs para espécies que não possuem genoma de referência (ILLUMINA, 2022) tornou ainda mais atrativo o uso destes marcadores. Entretanto, por possuir apenas quatro estados possíveis (A, T, C, G) os SNPs apresentam níveis de diversidade alélica menores do que os microssatélites (XING et al. 2005).

Um estudo realizado por Zimmerman, Aldridge e McCance (2020) comparou estimativas de diversidade genética e diferenciação genética utilizando dados de SNPs e microssatélites. O estudo evidenciou que o uso de SNPs quando comparado ao uso de microssatélites gera intervalos de confiança menores nas estimativas de diversidade. Adicionalmente, quanto maior o número de marcadores utilizados, maior a precisão dos resultados (menor o intervalo de confiança), entretanto isso não ocorreu para todas as métricas analisadas. As técnicas de *clusterização* e avaliação de estrutura populacional também apresentaram melhor capacidade de separar os indivíduos em grupos diferentes utilizando SNPs do que utilizando microssatélites (ZIMMERMAN, ALDRIDGE e MCCANCE, 2020). Desta forma, a não ser em casos onde o DNA amostrado apresenta baixa qualidade, o uso de SNPs pode ser mais vantajoso do que o uso de microssatélites em estudos de genética de populações.

## 1.6 PANORAMA GERAL

O processo de especiação é complexo, e pode ser estimulado por diversos fatores. A influência da alopatria na especiação é conhecida, entretanto, a influência de eventos como o surgimento de sistemas de cromossomos sexuais múltiplos e de rearranjos cromossômicos no processo de especiação e diferenciação genética, apesar de ser discutida há muito tempo, ainda não apresenta um consenso a respeito. Desta forma, torna-se necessário explorar de forma mais

aprofundada essa problemática, avaliando populações naturais de espécies com características apropriadas para compreensão do efeito destes processos.

De maneira geral, a família Erythrinidae apresenta-se como um modelo importante para abordar essas questões, pois reúne espécies e populações naturais que apresentam diversas variações citogenéticas, variando em número de cromossomos, na ocorrência de rearranjos cromossômicos e na presença de sistemas sexuais simples, múltiplos e ausentes. Adicionalmente, devido a ampla distribuição dessas espécies, é possível comparar o efeito destas variações cromossômicas ao efeito da alopatria. O uso de marcadores moleculares amplamente distribuídos pelo genoma obtidos por sequenciamento de nova geração, em conjunto com informações citogenéticas abundantes já disponíveis para o grupo, e outras que serão obtidas no presente trabalho permitirão uma investigação ampla a respeito da influência do desenvolvimento dos sistemas de cromossomos sexuais múltiplos na diversidade genética, e no processo de especiação.

## **2. OBJETIVOS**

### **2.1 OBJETIVOS GERAIS**

O objetivo principal desta investigação é analisar o impacto da ocorrência de rearranjos cromossômicos, especialmente aqueles relacionados ao sexo (levando, portanto, ao surgimento de sistemas de cromossomos sexuais múltiplos) na diversidade genética de populações de duas espécies, *H. malabaricus* e *E. erythrinus*, em busca de melhor compreender a correlação destes eventos cromossômicos com o processo de especiação.

## 2.2 OBJETIVOS ESPECÍFICOS

- (i) A diversidade genética para os indivíduos amostrados será quantificada e analisada para melhor compreensão das relações entre as populações e da influência dos sistemas sexuais na diversidade encontrada.
- (ii) A presença de estrutura populacional será investigada para os grupos de indivíduos amostrados, e os agrupamentos genéticos serão comparados com os padrões cariotípicos presentes em cada população, por meio de abordagens distintas, que consideram ou não as distâncias geográficas para inferir populações estruturadas.
- (iii) As correlações entre SNPs e as condições ambientais serão investigadas por meio de uma análise de Associação Genótipo-Ambiente (GEA) utilizando diversas variáveis ambientais.
- (iv) Para *Erythrinus erythrinus*, uma nova localidade será avaliada com técnicas citogenéticas clássicas e moleculares para verificar possíveis diferenças entre os cariomorfos.

### 3. MATERIAIS E MÉTODOS

A metodologia referente a coleta de espécimes, preparação de cromossomos mitóticos e hibridização *in situ* fluorescente (FISH), Hibridização Genômica Comparativa (CGH), Microscopia e processamento de imagens, sequenciamento e filtragem dos dados, detecção de marcadores potencialmente sob seleção e Análise de Associação Genótipo-Ambiente encontra-se descrita nos Capítulos 1 e 2.

#### 3.1 ANÁLISES DE ESTRUTURA POPULACIONAL

A estrutura populacional para ambas as espécies foi avaliada através de diferentes métodos. Inicialmente, uma Análise de Coordenadas Principais (PCoA) foi realizada para obtenção de um panorama inicial da distribuição da diversidade genética e da possível estrutura populacional apresentada pelos cariomorfos amostrados. Esta análise inicial foi realizada utilizando a função “gl.pcoa” do pacote “dartR” no software Rstudio v.1.3.959 (RSTUDIO TEAM, 2020). Posteriormente, o software fastStructure foi utilizado para identificar a estrutura populacional de forma mais específica pois este software incorpora premissas da genética de populações e foi desenvolvido especialmente para este fim (RAJ; STEPHENS; PRITCHARD, 2014). O fastStructure é um software que utiliza uma abordagem Bayesiana na inferência da estrutura populacional. Este método baseia-se no software STRUCTURE porém é otimizado para grandes conjuntos de dados como os empregados neste trabalho (PRITCHARD; STEPHENS; DONNELLY, 2000). Para realizar a análise, o *pipeline* intitulado “Lizards are awesome” foi utilizado para converter os arquivos de *input* para o formato *ped*, um dos formatos aceitos pelo fastStructure. As análises foram realizadas com todos os parâmetros utilizando valores padrão, os valores de K (número de grupos formados) testados foram de onze para *H. malabaricus*, e de quatro para *Erythrinus erythrinus*. Após o fim da análise, todos os resultados foram compactados em um arquivo no formato *zip* e enviados ao Clumpak (KOPELMAN et al. 2015) para geração da representação gráfica dos resultados. Adicionalmente, os peixes representam um grupo suscetível ao isolamento por distância (PEREZ et al. 2018). Partindo desta premissa, foi realizada uma análise espacial de estrutura populacional, ou seja, uma análise que considere os pontos geográficos de coleta dos indivíduos durante o processo de classificação das amostras em populações. Para realizar esta análise optou-se pelo método ConStruct (BRADBURD; COOP; RALPH, 2018). A análise foi realizada no software Rstudio e o parâmetro de análise espacial foi utilizado. As distâncias par-a-par foram obtidas através do Genealex (PEAKALL e SMOUSE, 2006; PEAKALL e SMOUSE, 2012) utilizando como base

as coordenadas de coleta. Para *H. malabaricus* a análise foi realizada com valores de K variando de 1 até 10 (dez foi o máximo de grupamentos permitidos pela análise) e para cada valor de K foram realizadas de 500.000 a 1.000.000 de iterações. Para *E. erythrinus* a análise foi realizada com valores de K variando de 1 até 4, e para cada valor de K um total de 1.000.000 de iterações foi executado. Após todas as corridas, a contribuição de cada valor de K para a estrutura populacional foi analisada através da metodologia de comparação da contribuição de cada camada de K, como indicado em BRADBURD (2019). Os resultados do ConStruct para *H. malabaricus* não atingiram o nível de convergência necessária para serem confiáveis, portanto não serão apresentados nos resultados e discussão, porém serão apresentados nos apêndices.

#### **4. RESULTADOS E DISCUSSÃO**

Os resultados obtidos através da investigação apresentada nesta dissertação, bem como a discussão dos mesmos encontram-se organizados na forma de dois artigos científicos, um deles está em processo de revisão, o outro encontra-se publicado. A tradução dos artigos será apresentada nos dois capítulos a seguir.

## CAPÍTULO 1

### **Elucidando as complexas relações das populações de *Hoplias malabaricus* (Teleostei; Characiformes). Uma abordagem genética, cromossômica e espacial.**

Fernando H. S. de Souza<sup>1</sup>, Manolo F. Perez<sup>1</sup>, Deborah Charlesworth<sup>2</sup>; Pedro H. N. Ferreira<sup>1</sup>, Luiz A. C. Bertollo<sup>1</sup>, Tariq Ezaz<sup>3</sup> e Marcelo de B. Cioffi<sup>1</sup>

<sup>1</sup>Laboratório de Citogenética de Peixes, Departamento de Genética e Evolução, Universidade Federal de São Carlos, 13565-905 São Carlos, SP, Brazil;

<sup>2</sup>Institute for Evolutionary Biology, Ashworth Laboratories, King's Buildings, University of Edinburgh, West Mains Road, EH9 3JT, Edinburgh, UK;

<sup>3</sup> Institute for Applied Ecology, University of Canberra, Canberra, Australia;

\* Correspondence: mbcioffi@ufscar.br; Tel.: +5533066885

Tipo de artigo: *Letter*

Status: Submetido ao periódico “*Evolution Letters*”

## RESUMO

A notável diversidade de peixes neotropicais permanece pouco compreendida. Embora a alopatria seja um dos principais fatores que favorecem a especiação, os rearranjos cromossômicos e a evolução dos cromossomos sexuais também podem desempenhar papéis importantes. Rearranjos cromossômicos envolvendo cromossomos sexuais podem produzir híbridos que apresentam desequilíbrios em genes envolvidos na determinação do sexo, levando ao isolamento entre populações. Neste trabalho, nosso objetivo foi inferir a contribuição relativa da alopatria, rearranjos cromossômicos e cromossomos sexuais para a divergência genética populacional no peixe lobo *Hoplias malabaricus*, usando análises populacionais com dados genômicos e citogenéticos. Analisamos 10 localidades simpátricas e alopátricas que diferem por suas i) características cromossômicas (denominadas cariomorfos), ii) distribuições simpátricas versus alopátricas e iii) presença e ausência de cromossomos sexuais heteromórficos. Os resultados indicam que, embora os cariomorfos representem populações geneticamente divergentes, sua história evolutiva pode ser melhor compreendida quando considerada juntamente com suas distribuições geográficas, condições ambientais e distâncias genéticas. Em alguns casos, diferentes locais de amostragem com o mesmo cariomorfo apresentaram diferenciação genética acentuada, sugerindo evolução em isolamento. Os cariomorfos A e D são particularmente interessantes, pois suas populações abrangem todas as características analisadas: condições alopátricas/simpátricas, diferenciação cromossômica e presença/ausência de cromossomos sexuais. O cariomorfo D tem cromossomos neo-X<sub>1</sub>X<sub>2</sub>Y, enquanto o cariomorfo A não apresenta um sistema de cromossomos sexuais. O sistema cromossômico neo-X<sub>1</sub>X<sub>2</sub>Y do cariomorfos D está associado a uma maior diferenciação genética interpopulacional em comparação com o efeito da alopatria dentro do cariomorfo A, que é encontrado em vários locais geograficamente distantes. Portanto, o sistema de cromossomos neo-sexuais pode estar provavelmente associado à fixação de diferenças genéticas e a promoção do isolamento reprodutivo. Mais evidências são necessárias para compreender se ele atuou como fator de divergência ou amplificou a diferenciação alopátrica já existente nesses dois cariomorfos sintópicos (A e D).

**Palavras-chave:** Diversidade genética, evolução, especiação, cromossomos sexuais, cariótipo, estrutura populacional

## INTRODUÇÃO

A região Neotropical apresenta uma biodiversidade ictiológica vasta, com mais 5.200 espécies descritas (FRICKE; ESCHMEYER; VAN DER LAAN, 2021). Esta alta biodiversidade está alinhada com a tendência bem estabelecida dos trópicos apresentarem níveis de biodiversidade mais elevados quando comparados à outras latitudes, por ambientes distintos (Gradiente latitudinal de biodiversidade) (MESEGUER e CONDAMINE, 2019). A hipótese “Out of the tropics” (JABLONSK; ROY; VALENTINE, 2006) é uma das diversas propostas que tentam explicar este padrão, sugerindo que regiões próximas aos trópicos tem altas taxas de especiação aliadas a baixas taxas de extinção. Adicionalmente, aspectos específicos da região Neotropical, incluindo sua história geológica e climática complexas podem afetar os níveis de especiação (ALBERT, 2011; RULL, 2018).

A especiação alopátrica, com o isolamento reprodutivo evoluindo como um subproduto da divergência genética de populações isoladas, é geralmente considerada a principal forma de processo evolutivo, sendo que o isolamento das populações pode ser devido à distância, barreiras geográficas, ou características da paisagem que isolem as populações (DE QUEIROZ 2007). Os Neotrópicos com certeza possuem uma história complexa, com diversos eventos tectônicos e climáticos que ocorreram principalmente durante o Neogeno e o Pleistoceno (HOORN, 2010; GARZÓN-ORDUÑA; BENETTI-LONGHINI; BROWER, 2014) e podem ter desempenhado importante papel na geração desta alta diversidade.

Os Characiformes são uma das maiores ordens de peixes de água doce, com aproximadamente 2.300 espécies (NELSON; GRANDE; WILSON, 2016), incluindo a família Erythrinidae que compreende apenas três gêneros *Hoplias* Gill, 1903; *Hoplerythrinus* Gill, 1895 e *Erythrinus* Scopoli, 1977, todos endêmicos à América do Sul e América Central. Dentro do gênero *Hoplias*, a “traíra comum”, *Hoplias malabaricus*, está amplamente distribuída em várias bacias hidrográficas da América do Sul (OYAKAWA, 2003) e populações geograficamente isoladas podem desenvolver diferenças, devido ao número cromossômico e/ou morfologia, e pela ocorrência de diferentes sistemas de cromossomos sexuais, com sete cariomorfos distintos (A a G,) atualmente identificados (BERTOLLO, 2007; CIOFFI et al. 2012b; **Tabela 1; Figura 7**).

Enquanto alguns cariomorfos são amplamente distribuídos, outros são endêmicos em áreas pequenas, até mesmo apresentando distribuição simpátrica, de forma que dois cariomorfos distintos podem ser encontrados no mesmo habitat ao mesmo tempo (sintópicos). Com exceção de um único caso envolvendo uma triploidia natural (UTSUNOMIA et al. 2014),

não há evidências de que cariomorfos distintos reproduzam-se e produzam híbridos. A espécie nominal *H. malabaricus* é então provavelmente um complexo de espécies, com unidades evolutivas independentes (BERTOLLO et al. 2000), ao invés de uma única espécie (BERTOLLO, 2007; CIOFFI et al. 2012b)

Os cromossomos sexuais dos cariomorfos A e E não são morfologicamente diferenciados, enquanto o cariomorfo B apresenta um sistema sexual heteromórfico do tipo XX/XY (BORN e BERTOLLO, 2000), que aparenta ser derivado do cariomorfo A (CIOFFI e BERTOLLO 2010). Os cariomorfos C e F apresentam também sistemas do tipo XY, porém em um estágio de diferenciação mais incipiente que o cariomorfo B (CIOFFI e BERTOLLO 2010; FREITAS et al. 2018). Por fim, os cariomorfos D e G apresentam sistemas de cromossomos sexuais múltiplos do tipo  $X_1X_1X_2X_2/X_1X_2Y$  e  $XX/XY_1Y_2$  respectivamente (BERTOLLO et al. 2000). O sistema  $X_1X_2Y$  no cariomorfo D aparenta ser derivado do sistema XY do cariomorfo C devido a uma fusão do Y com um cromossomo autossomo (CIOFFI e BERTOLLO 2010). O cariomorfo homomórfico E possivelmente deu origem aos dois cariomorfos heteromórficos F e G, devido a fusão entre o cromossomo X e um autossomo (DE OLIVEIRA et al. 2015). A **Figura 7** apresenta os cariomorfos e indica a potencial relação evolutiva entre eles.

O presente estudo utiliza abordagens genômicas para investigar como a evolução dos sistemas sexuais múltiplos pode estar relacionada com a diferenciação genética, fluxo gênico e especiação. Nós analisamos 10 localidades alopátricas e simpátricas de *H. malabaricus*, com indivíduos que variam tanto em número de cromossomos diplóides quanto na presença ou ausência de sistemas de cromossomos sexuais e morfologia dos cromossomos. Nossa hipótese é de que presença de sistemas sexuais múltiplos pode ter contribuído para a diferenciação genética das populações, o que pode potencialmente ter acelerado o processo de especiação.

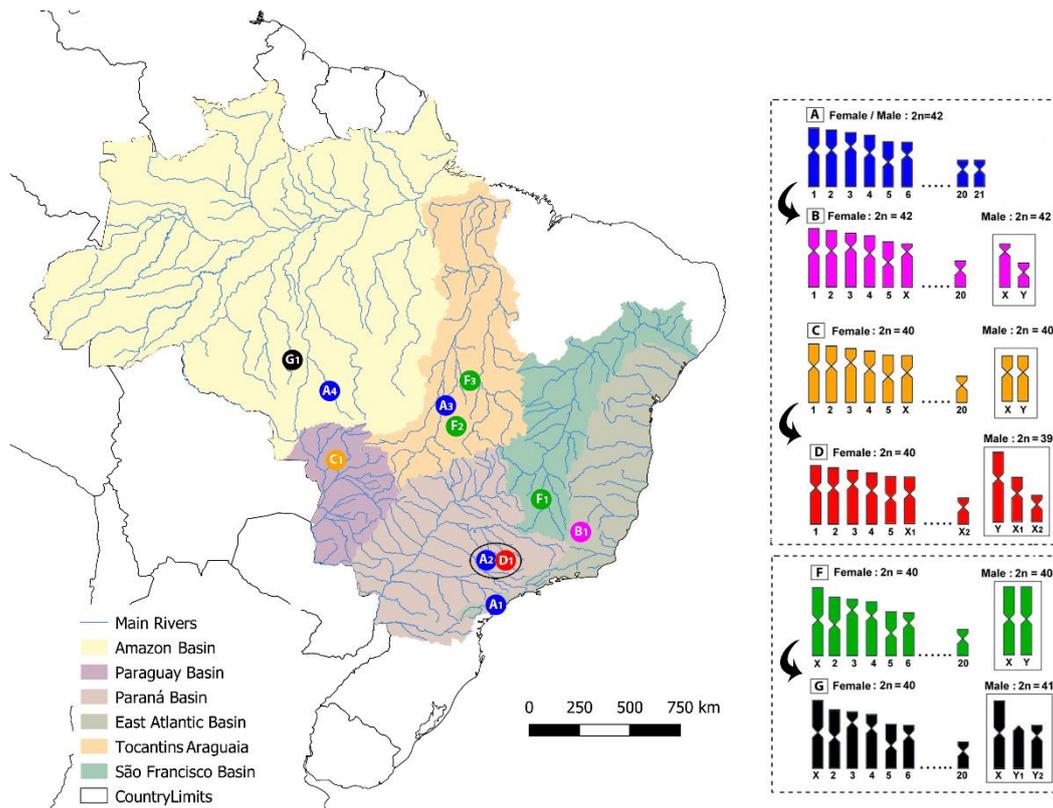
## MÉTODOS

### AMOSTRAGEM DE ESPÉCIMES

Os pontos de coleta, quantidade e sexo dos espécimes investigados estão dispostos na **Tabela 1** e **Figura 7**. Algumas das amostras são de locais previamente analisados. Todos os animais foram coletados com autorização da agência ambiental brasileira ICMBIO/SISBIO (licença nº. 48628-14) e SISGEN (A96FF09). Todos os experimentos seguiram protocolos de anestesia e procedimentos éticos previamente aprovados pelo Comitê de Ética em Experimentação Animal da Universidade Federal de São Carlos (processo número CEUA 1853260315).

**Tabela 1** - Indivíduos analisados, seus cariomorfos, número de cromossomos diploides (2n), Sistema de cromossomos sexuais, Código de identificação dos grupos (indicando o local de coleta e o cariomorfo), número de indivíduos analisados citogeneticamente e número de indivíduos sequenciados.

Cariomorfo	2n	Cromossomos sexuais	Código	Local de coleta	Latitude/ Longitude	Análise citogenética	Análise por sequenciamento	Referência
A	♀♂ 42	Homomórfico	A1	Rio Ribeira de Iguapé (SP)	-24.489722 -47.836111	09♂ 09♀	6	SANTOS et al. 2009
A	♀♂ 42	Homomórfico	A2	Córrego do Monjolinho (SP)	-21.985556 -47.881944	09♂ 09♀	6	CIOFFI et al. 2010
A	♀♂ 42	Homomórfico	A3	Araguaia River (GO)	-13.179504 -50.583301	09♂ 11♀	8	BLANCO et al. 2010
A	♀♂ 42	Homomórfico	A4	Bacia do Rio Xingu (MT)	-12.404056 -56.960861	14♂ 11♀	6	BLANCO et al. 2010
B	♀♂ 42	XY	B1	Rio Doce (MG)	-20.258019 -42.901313	10♂09♀	1	CIOFFI et al. 2009
C	♀♂ 40	XY - Diferenciação inicial	C1	Rio Poconé (MT)	-16.252905 -56.574296	19♂ 08♀	6	CIOFFI e BERTOLLO, 2010
D	♀ 40 ♂ 39	X1X1X2X2 / X1X2Y	D1	Córrego do Monjolinho (SP)	-21.985556 -47.881944	09♂ 14♀	6	CIOFFI e BERTOLLO, 2010
F	♀♂ 40	XY - Diferenciação inicial	F1	Rio Três Marias (MG)	-18.524139 -45.234917	14♂ 10♀	6	FREITAS et al. 2018
F	♀♂ 40	XY - Diferenciação inicial	F2	Rio Peixe (GO)	-14.358889 -49.825861	13♂ 06♀	6	Trabalho atual
F	♀♂ 40	XY - Diferenciação inicial	F3	Rio Gurupi (TO)	-11.691366 -48.970221	09♂ 12♀	3	Trabalho atual
G	♀ 40 ♂ 41	XX / XY1Y2	G1	Rio Aripuanã (MT)	-10.753389 -59.259667	12♂ 09♀	6	OLIVEIRA et al. 2018



**Figura 7** - Mapa da América do Sul indicando as localidades com amostras de *Hoplias malabaricus* analisadas. As cores indicam as diferentes bacias hidrográficas no Brasil, listadas na legenda, e os círculos coloridos indicam os pontos de coleta, com os cariomorfos. O idiograma à direita indica os cariótipos parciais de cada cariomorfo e o seu sistema de cromossomos sexuais, quando presente.

## SEQUENCIAMENTO E FILTRAGEM

Foram utilizados tecidos do fígado para realização do procedimento de sequenciamento DArTseq na Diversity Arrays Technology Pty Ltd (Austrália). Este método de sequenciamento é um método de redução de complexidade que utiliza as enzimas PstI e SbfI para digerir DNA antes do sequenciamento, uma abordagem que enriquece as sequências de regiões pouco metiladas (KILIAN et al. 2012). O sequenciamento foi realizado na plataforma Illumina HiSeq 2500, e os dados brutos foram processados utilizando o software pyRAD v3.0.66 (EATON, 2014). Os adaptadores de sequenciamento foram removidos, as sequências com mais de cinco bases de baixa qualidade ( $Q < 20$ ) também foram removidas. Os *reads* foram alinhados e as sequências consenso foram obtidas para cada uma das amostras, e usadas para definir a base de referência em cada posição do alinhamento. Nós obtivemos *loci* que estavam presentes em todos os indivíduos, e geramos dois diferentes conjuntos de dados para as análises subsequentes: (1) um arquivo contendo sequências para cada *loci* com a fase determinada, que foram utilizados nas análises de diversidade genética e diferenciação genética e (2) uma matriz de polimorfismos de nucleotídeo único (SNPs), selecionando apenas um único SNP por *locus*,

a fim de evitar a presença de SNPs ligados. Quando presentes, *indels* não foram considerados nas análises. Os SNPs foram codificados da seguinte forma: 0 para homozigoto para a base de referência, 1 para heterozigotos, e 2 para os homozigotos para a base alternativa. Este segundo conjunto de dados foi utilizado em todas as demais análises.

## DETECÇÃO E EXCLUSÃO DOS MARCADORES POTENCIALMENTE SOB SELEÇÃO

Com o objetivo de testar a presença de marcadores com níveis incomuns de diferenciação interpopulacional (tanto altos quanto baixos), que podem refletir pressões seletivas em algumas populações, indicando alelos localmente adaptados nos *loci* sequenciados, uma análise no software BayeScan foi implementada para buscar por *outliers* (FOLL e GAGGIOTTI, 2008). Nós realizamos as corridas utilizando um valor de *prior odds* de 100, um valor de *thinning* de 10 e 50000 iterações de cadeia de Markov Monte Carlo (MCMC). Loci que apresentaram valores de *False Discovery Rate* (FDR) menores que 0,01 foram classificados como possivelmente sob seleção. Os *loci* identificados foram posteriormente investigados através de uma análise com a ferramenta de alinhamento NCBI BlastX utilizando o banco de dados Uniprot/Swiss-Prot com o objetivo de identificar possíveis funções. Os marcadores sob seleção foram removidos das análises subsequentes.

## DIVERSIDADE GENÉTICA

A diversidade genética foi estimada através do software DnaSP v. 6.12.03 (ROZAS et al. 2017) para obtenção das médias de diversidade haplotípica ( $H_d$ ),  $D$  de Tajima ( $D$ ) e médias de duas outras medidas de diversidade nucleotídica por sítio ( $\pi$ ) e Theta de Watterson ( $\theta_w$ ). A divergência nucleotídica entre amostras de diferentes localidades ou cariomorfos foi estimada utilizando o número médio de substituições nucleotídicas por sítio entre pares de amostras de diferentes populações ou cariomorfos ( $D_{XY}$ ), e a divergência total, corrigida pela variação dentro das amostras analisadas ( $D_A$ ), que melhor reflete os tempos relativos em que começaram a evoluir de forma independente. Estas estimativas nos permitem ter uma boa ideia dos tempos relativos à ancestralidade comum das amostras, usando a divergência estimada entre elas juntamente com a diversidade dentro delas.

## ANÁLISES DE ESTRUTURA POPULACIONAL

Nós investigamos a estrutura populacional por meio de uma Análise de Coordenadas Principais (PCoA) para visualizar a distribuição da diversidade genética e a estrutura

populacional em cada um dos cariomorfos. A PCoA foi realizada utilizando a função “gl.pcoa” do pacote *dartR*. Posteriormente realizamos uma análise com o software *fastStructure* (RAJ; STEPHENS; PRITCHARD, 2014), que aplica uma abordagem Bayesiana para avaliar a estrutura populacional. Este método é similar ao do software *STRUCTURE* (PRITCHARD; STEPHENS; DONNELLY, 2000), porém otimizado para conjuntos de dados maiores. Para realizar as análises, o pipeline “Lizards are awesome” foi utilizado para converter os arquivos de entrada para o formato *ped* e posteriormente realizar a análise no *fastStructure*. Nós utilizamos as configurações padrão, com valor de K máximo configurado em 11. Os resultados foram comprimidos em formato *zip* e enviados ao *Clumpak* (KOPELMAN et al. 2015) para representação gráfica dos resultados.

Uma análise de Associação Genótipo-Ambiente (GEA) foi realizada para identificar correlações entre SNPs e as condições ambientais, usando as variáveis ambientais de água doce compiladas em Domisch, Amatulli e Jetz (2015). Incluímos todas as onze variáveis de temperatura hidroclimática (temperatura média anual, intervalo diurno médio, isothermalidade, sazonalidade da temperatura, temperatura máxima do mês mais quente, temperatura mínima do mês mais frio, intervalo de temperatura anual, temperatura média do trimestre mais úmido, temperatura média do trimestre mais seco, temperatura média do trimestre mais quente e temperatura média do trimestre mais frio) junto com oito variáveis de precipitação (precipitação anual, precipitação do mês mais chuvoso, precipitação do mês mais seco, sazonalidade da precipitação, precipitação do trimestre mais chuvoso, precipitação do trimestre mais seco, precipitação do trimestre mais quente e precipitação do trimestre mais frio). Além disso, também incluímos quatro variáveis para elevações e declives (LEHNER; VERDIN; JARVIS, 2008), junto com quatro variáveis para carbono orgânico e pH do solo (HENGL et al. 2014). Para cada variável ambiental, realizamos uma PCA e mantivemos apenas os dois primeiros valores de Componente Principal (PC) obtidos. As variáveis transformadas resultantes foram testadas para correlação e as variáveis com maiores valores de fator de inflação de variância (VIF) foram removidas. Para levar em consideração a possível confusão com o componente espacial (MEIRMANS, 2015), usamos as coordenadas geográficas de nossos locais de amostragem (decompostos como um polinômio de terceiro grau) para testar as correlações. Finalmente, realizamos RDA com as variáveis ambientais transformadas, testando correlações com os genótipos enquanto controlamos as variáveis espaciais, seguindo uma abordagem semelhante à sugerida por Brauer et al. (2018).

## RESULTADOS

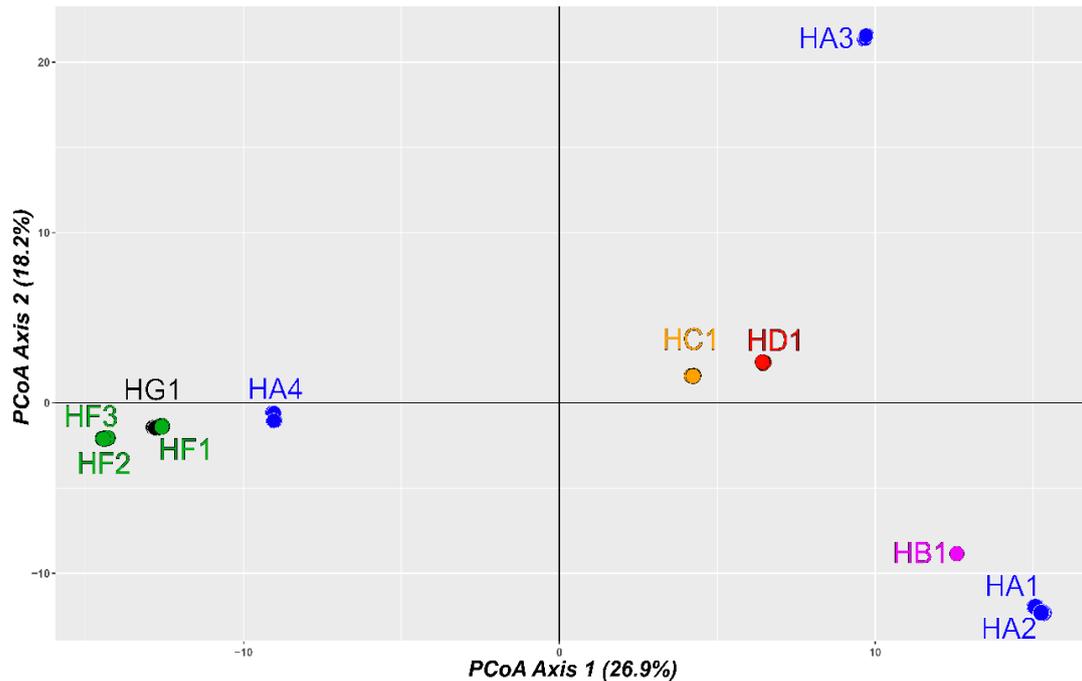
### DADOS CITOGENÉTICOS, SEQUENCIAMENTO, FILTRAGEM E DETECÇÃO DE MARCADORES SOB SELEÇÃO

Oito das dez populações estudadas já haviam sido categorizadas em diferentes cariomorfos. Duas localidades adicionais foram analisadas pela primeira vez, ambas atribuídas ao cariomorfo F, com  $2n=40$  e um sistema de cromossomos sexuais XY críptico, de acordo com Bertollo et al. (2000) e Freitas et al. (2018) (**Tabela 1**).

Nosso sequenciamento resultou em aproximadamente dois milhões de *reads* por amostra. Após a filtragem e remoção dos adaptadores de sequenciamento, um total de 6.848 *loci* foram obtidos para todos os indivíduos sequenciados, incluindo 14.105 SNPs variáveis. Nós extraímos 6.200 SNPs polimórficos depois de descartar do conjunto de dados os SNPs que estavam presentes na mesma sequência e manter apenas um único SNP por *locus*. A análise no Bayescan indicou que 73 *loci* estavam possivelmente sob seleção e podem ter diferentes alelos favorecidos em diferentes ambientes, em vez de evoluir como esperado se as diferenças fossem neutras. Apenas três delas corresponderam às sequências disponíveis no banco de dados Swissprot. Um deles correspondeu, com 97% de *query cover* e *e-value* de  $8e-06$ , a um receptor acoplado à proteína G que é conservado entre várias espécies, de *Homo sapiens* a *Danio rerio*. Outra sequência correspondeu à subunidade M1 da Ribonucleosídeo-difosfato redutase com 97% de *query cover* e *e-value* de  $5e-09$ . Esta proteína é altamente conservada e está envolvida em várias funções metabólicas importantes, como o metabolismo da glutatona, a síntese e conversão de difosfatos e trifosfatos de nucleotídeos e o metabolismo das purinas. A terceira correspondência foi observada com a sequência Homeobox da proteína Zinc Finger 4 de *Mus musculus* com alto *query cover* (97%), mas *e-value* de apenas 0,010. Como esses três marcadores podem estar evoluindo de forma não neutra, eles foram removidos das análises posteriores junto com os demais marcadores potencialmente sob seleção.

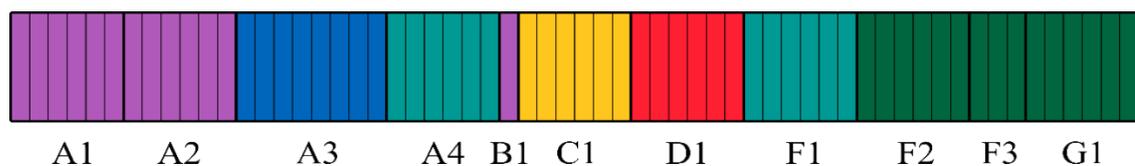
### ESTRUTURA POPULACIONAL

Os resultados da PCoA evidenciariam uma forte estruturação populacional (**Figura 8**). Inesperadamente, o cariomorfo A formou três grupos distintos com alta diferenciação de sequência entre amostras de diferentes regiões geográficas, embora as duas localidades da bacia do Paraná se agrupem (HA1 e HA2).



**Figura 8** - Análise de Coordenadas Principais. Indivíduos de diferentes cariomorfos estão representados por cores distintas de acordo com a Figura 7.

Os resultados do fastStructure indicaram que o valor de  $K=6$  maximizou a probabilidade posterior (**Figura 9**), e teve conclusões similares as da PCoA, com algumas pequenas diferenças no grupo que inclui o cariomorfo A4. A PCoA agrupa F1 com G1, enquanto o fastStructure agrupa F1 com A4, e também G1 com F2 e F3. Adicionalmente, os cariomorfos C e D foram separadas pelo fastStructure.



**Figura 9** - Resultado do FastStructure para  $K=6$ . Cada barra vertical representa um indivíduo, os cariomorfos e locais de coleta estão representados abaixo; a cor das barras representa a população na qual a análise classificou cada indivíduo.

## DIVERSIDADE GENÉTICA E DIFERENCIAÇÃO GENÉTICA

Todas as localidades amostradas apresentaram baixos valores de diversidade genética (**Tabela 5**). As duas medidas de diversidade para as diferentes localidades,  $\pi$  e  $\theta_w$ , foram concordantes e são todas baixas; mesmo excluindo a localidade com diversidade excepcionalmente alta, A4, o  $R^2$  para a correlação é de 84%. No geral, a diversidade nucleotídica ( $\pi$ ) e o  $\theta_w$  foram especialmente altas nos cariomorfos A e F, para os quais vários locais de amostragem foram analisados, com populações alopátricas de cariomorfo A

mostrando valores de divergência tão altos quanto os valores entre diferentes cariomorfos (**Figura 10**). As amostras da maioria das localidades têm valores de  $\pi$  em torno de 0,1% (**Tabela 3**), com exceção de um valor 4 vezes maior para a localidade A4 (que teve o maior valor de diversidade para todas as medidas de diversidade).

**Tabela 2** - Valores médios de diversidade genética estimada para *H. malabaricus* por cariomorfo. Os tamanhos amostrais são indicados, bem como a diversidade haplotípica e o desvio padrão ( $H_d$ ), Diversidade nucleotídica ( $\pi$ ), Theta de Watterson por sitio ( $\theta_w$ ) e D de Tajima ( $D$ ).

Cariomorfo	Tamanho amostral	$H_d \pm DP$	$\pi$	$\theta_w$	$D$
A	26	0,2983 (0,0459)	0,0069	0,0054	0,4962
B	1	0,0397 (0,9963)	0,0007	0,0007	n.a.
C	6	0,0299 (0,0405)	0,0006	0,0006	-0,0828
D	6	0,0658 (0,0540)	0,0013	0,0012	0,1671
F	15	0,1520 (0,0433)	0,0031	0,0026	0,3978
G	6	0,0457 (0,0483)	0,0009	0,0009	-0,0013

**Tabela 3** – Valores médios de diversidade genética estimada para *H. malabaricus* por ponto de coleta. Os tamanhos amostrais são indicados, bem como a diversidade haplotípica ( $H_d$ ) e o desvio padrão, Diversidade nucleotídica ( $\pi$ ), Theta de Watterson por sitio ( $\theta_w$ ) e D de Tajima ( $D$ ).

Código	Tamanho amostral	$H_d \pm DP$	$\pi$	$\theta_w$	$D$
A1	6	0,0356 (0,0434)	0,0007	0,0007	0,0075
A2	6	0,0272 (0,0372)	0,0006	0,0006	0,0210
A3	8	0,0662 (0,0548)	0,0012	0,0015	-0,3525
A4	6	0,1860 (0,0510)	0,0040	0,0027	1,2311
B1	1	0,0397 (0,0996)	0,0007	0,0007	n.a.
C1	6	0,0299 (0,0405)	0,0006	0,0006	-0,0828
D1	6	0,0658 (0,0540)	0,0013	0,0012	0,1671
F1	6	0,0462 (0,0489)	0,0009	0,0009	-0,0706
F2	6	0,0492 (0,0509)	0,0009	0,0010	-0,0777
F3	3	0,0598 (0,0635)	0,0011	0,0010	0,3328
G1	6	0,0457 (0,0483)	0,0009	0,0009	-0,0013

Nossos resultados mostram que as estimativas de divergência entre indivíduos amostrados de diferentes populações são geralmente baixas (**Figura 10**). Especificamente, as estimativas de divergência líquida entre indivíduos amostrados de diferentes populações (valores de  $D_A$ ) estão principalmente entre 0,5 e 1%, de 5 a 10 vezes mais altas do que entre pares de indivíduos dentro dessas populações (estimado por  $\pi$ ). As únicas exceções, com estimativas de divergência mais baixas, foram dois pares de amostras, A1 e A2 (as amostras

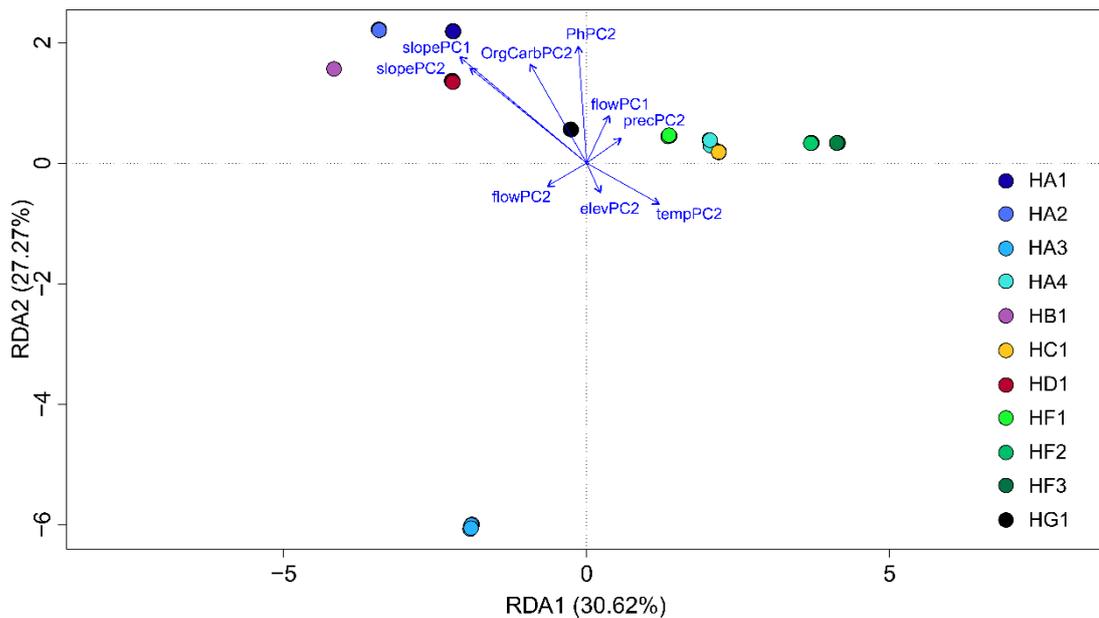
geograficamente mais próximas de cariomorfo A, ambas da bacia do Paraná), e F2 e F3 (as duas amostras de cariomorfo F com as localizações geográficas mais próximas, ambos da mesma drenagem). Por outro lado, os maiores valores de divergência foram observados para comparações envolvendo amostras entre A3 e D1 (**Figura 10**). Vários grupos do cariomorfo A quando comparados a outros grupos do cariomorfo A apresentam valores de divergência  $D_A$  muito maiores do que os valores de diversidade nucleotídica ( $\pi$ ) apresentado pelas amostra do cariomorfo A (**Tabela 3**). O mesmo ocorre para as amostras do cariomorfo F (**Tabela 3**).

	A1	A2	A3	A4	B1	C1	D1	F1	F2	F3	G1
A1		0.0014	0.0099	0.0095	0.0035	0.0084	0.0097	0.0094	0.0096	0.0096	0.0095
A2	0.0008		0.0099	0.0095	0.0035	0.0085	0.0097	0.0094	0.0096	0.0096	0.0095
A3	0.0090	0.0091		0.0109	0.0102	0.0100	0.0116	0.0107	0.0108	0.0109	0.0109
A4	0.0073	0.0073	0.0084		0.0099	0.0092	0.0110	0.0065	0.0065	0.0066	0.0064
B1	0.0028	0.0029	0.0093	0.0076		0.0089	0.0102	0.0099	0.0100	0.0100	0.0099
C1	0.0078	0.0079	0.0092	0.0070	0.0082		0.0089	0.0090	0.0091	0.0092	0.0091
D1	0.0088	0.0089	0.0105	0.0084	0.0092	0.0092		0.0107	0.0109	0.0110	0.0108
F1	0.0087	0.0088	0.0096	0.0041	0.0091	0.0083	0.0097		0.0052	0.0053	0.0055
F2	0.0088	0.0089	0.0098	0.0041	0.0092	0.0084	0.0099	0.0043		0.0014	0.0058
F3	0.0088	0.0088	0.0097	0.0041	0.0092	0.0084	0.0099	0.0043	0.0004		0.0059
G1	0.0088	0.0089	0.0099	0.0040	0.0092	0.0084	0.0098	0.0046	0.0049	0.0049	

**Figura 10** - Valores de  $D_{XY}$  par-a-par por ponto de coleta são representados na diagonal superior, e os valores de  $D_A$  par-a-par na diagonal inferior. Valores mais altos são verdes, e conforme os valores diminuem a cor muda para amarelo, laranja e então vermelho para os menores valores.

## ASSOCIAÇÃO GENÓTIPO-AMBIENTE

Após o teste de correlação entre as variáveis ambientais, o segundo PC foi mantido para as variáveis carbono orgânico, pH, precipitação, elevação e temperatura (Org Carb PC2, pH PC2, Prec PC2, Elev PC2 e TempPC2), e o primeiro e segundo PCs para as variáveis fluxo e inclinação (Flow PC1, Flow PC2, Slope PC1 e Slope PC2, respectivamente) para realizar a RDA subsequente. O modelo da RDA foi significativo ( $p < 0,001$ ) com os eixos RDA1 e RDA2 explicando 30,62% e 27,27% da variação, respectivamente (ambos significativos;  $p < 0,001$ ). As variáveis mais importantes para a variação restrita foram PhPC2 (18,2%), flowPC1 (13,34%) elevPC2 (12,5%) e slopePC1 (9,60%). Ao plotar os indivíduos ao longo dos dois eixos RDA, foi possível observar que as variáveis ambientais também estão afetando fortemente a diversidade genética de *H. malabaricus*, apresentando uma estrutura clara (**Figura 11**). A RDA1 separou as localidades B1, A1, A2 e D1 de A4, C1, F1, F2 e F3, enquanto G1 apresentou valores intermediários. A RDA2 separou claramente A3 das demais localidades.



**Figura 11** - Gráfico apresentando dois eixos principais da RDA. Os círculos representam os indivíduos, com cores indicando seu local de amostragem e cariomorfo, conforme a legenda à direita. Os fatores ambientais considerados significativos são representados por linhas azuis, com setas indicando a localização na RDA em que os dados estão concentrados.

## DISCUSSÃO

Neste estudo, analisamos as características cromossômicas juntamente com a diversidade genética, distribuição espacial e condições ambientais de diferentes populações, para melhor compreender a evolução da espécie *H. malabaricus*. Nossos resultados mostraram que agrupar diferentes populações de *H. malabaricus* por seus cariomorfos, ou por suas localidades geográficas, não pode por si só explicar completamente sua diferenciação. Às vezes, diferentes cariomorfos mostram diferenças genéticas que os separam de indivíduos em outros cariomorfos, mesmo que não sejam alopátricos. No entanto, a separação geográfica também parece ser importante e pode estar associada a diferenças ambientais. Alguma diferenciação entre cariomorfos provavelmente reflete pequenas diferenças que evoluíram em populações pequenas e talvez isoladas, e essas condições podem ter permitido que diferenças cromossômicas fossem fixadas, mesmo aquelas que são desfavorecidas em heterozigotos entre diferentes tipos cromossômicos. No entanto, não podemos explicar toda a diferenciação de cariomorfos dessa maneira, pois até mesmo nossa pequena amostragem detectou diversidade em todos os cariomorfos amostrados.

## DIVERSIDADE GENÉTICA E ESTRUTURA POPULACIONAL

A metodologia aqui aplicada resulta em dados enriquecidos para sequências potencialmente codificantes, o que pode ter contribuído para os baixos valores de diversidade apresentados (consulte Métodos). Também esperávamos níveis mais baixos de diversidade genética porque muitos sítios estarão sob restrições de seleção. Portanto, as estimativas de diversidade sinônima podem ser maiores, porém a anotação do genoma dessa espécie não nos permite gerar tais estimativas. Observando os dados de diferenciação genética (**Figura 10**), concluímos que as populações foram isoladas por tempo suficiente para ter divergido geneticamente. Os valores de diversidade nucleotídica em uma população que esteja em equilíbrio de Hardy-Weinberg não devem depender do tamanho da amostra. A observação de que os dois cariomorfos com as maiores amostras, A e F, apresentaram as estimativas mais altas é consistente com o isolamento geográfico causando diferenciação e, portanto, maiores estimativas de diversidade das populações agrupadas, já que ambos os cariomorfos A e F foram amostrados em vários locais, separados por grandes distâncias (**Tabela 2**).

Para a população A3, os valores de divergência de ambas as análises de diferenciação sugerem isolamento geográfico por tempo considerável, pois suas sequências estão entre as mais diferentes das sequências das outras amostras de cariomorfo A (quase tão diferentes quanto sequências de outros cariomorfos); isso é consistente com ambas as análises de diferenciação. O valor de  $D_A$  também é alto entre A2 e a amostra simpátrica D1, sugerindo que os dois cariomorfos são tão isolados quanto a maioria dos pares de populações de locais distantes. A diferenciação clara dos cariomorfos A1 e A2 nas análises de diferenciação genética pode refletir amplamente sua baixa diversidade dentro da amostra. No outro extremo, apesar da localização separada de A4, a alta diversidade dentro da amostra e os valores de divergência líquidos mais baixos quando comparada a outras amostras exceto A3 (e também valores mais baixos do que A1 e A2) podem indicar a possibilidade de que população A4 recebeu contribuições genéticas através de introgressão em algum momento.

A alta diversidade desta amostra pode refletir a introgressão de outras populações. A população A4 também diferencia-se das outras amostras por ter um valor de D de Tajima altamente positivo, sugerindo que esta população pode possuir contribuições substanciais de sequências ancestralmente divergentes. Uma possibilidade alternativa, mas menos plausível, é que estes indivíduos sejam um resquício de uma população ancestral com um tamanho efetivo muito grande (e, portanto,  $\pi$  alto), e que um evento de gargalo muito recente causou a perda de variantes raras, reduzindo seu valor de  $\theta_w$  e levando a um D de Tajima elevado. A divergência

líquida é particularmente baixa entre A4 e todas as três amostras do cariomorfo F, com valores muito semelhantes à diversidade apenas na amostra A4 (0,4%), enquanto a diversidade dentro das amostras do cariomorfo F é menor (com  $\pi$  em torno de 0,1% em todas as três). Isso sugere que as mudanças que produziram o cariomorfo F podem ter ocorrido em uma população pequena e isolada, e o quadro de forte diferenciação nas análises de diferenciação pode refletir amplamente a baixa diversidade dentro das amostras do cariomorfo F. Apesar do complicado quadro geral, os resultados sugerem que tanto o isolamento geográfico quanto as diferenças cromossômicas contribuem para a divergência.

A análise da estrutura populacional sugeriu que o cariomorfo A é composto por três grupos em diferentes localizações alopátricas, sendo apenas A1 e A2, de uma mesma bacia hidrográfica, agrupados juntos (**Figuras 7 e 9**). Da mesma forma, o cariomorfo F também compreende dois grupos distintos, um com indivíduos das localidades F2 e F3, e o outro com os indivíduos de F1. Esses grupos são explicados por sua distribuição geográfica, uma vez que F2 e F3 são ambos da bacia hidrográfica do Tocantins-Araguaia, enquanto F1 está localizado na bacia do Rio São Francisco. Além disso, a PCoA sugeriu que B1 está intimamente relacionado com A1, D1 com C1 e F1 com G1. Esses resultados são consistentes com as relações propostas para esses cariomorfos na **Figura 7**, com o cariomorfo B sendo derivado de A, o cariomorfo D de C (BERTOLLO et al. 2000) e os cariomorfos F e G derivados de E (BERTOLLO et al. 2000; OLIVEIRA e SEMBER et al. 2018). Os resultados do fastStructure são parcialmente congruentes com o resultado da PCoA. Nesta análise os indivíduos com os cariomorfos C e D formaram, cada um, uma população. Os grupos F1 e A4 formaram uma população, enquanto F2, F3 e G1 formaram outra, embora F1 esteja mais próximo de G1 na PCoA. Dada a diversidade globalmente baixa dentro das amostras, é importante considerar se a diferenciação detectada nas análises de PCoA e no fastStructure é devido à divergência genética ou reflete principalmente a baixa diversidade dentro das amostras (o que pode gerar valores altos quando a diferenciação é estimada), uma vez que a maioria das diferenças deve ser entre as amostras.

Além das diferenças geográficas e cariotípicas, a variação genética em *H. malabaricus* também é afetada pelas variações ambientais. A RDA foi corrigida para o efeito dos dados geográficos, de modo que os padrões observados na **Figura 11** devem refletir a correlação entre o ambiente e as diferenças genéticas apenas. As variáveis ambientais que melhor contribuíram para o modelo foram Slope PC1 e Slope PC2 (declive médio), Org Carb PC2 (carbono orgânico do solo), pH PC2 (pH do solo em H<sub>2</sub>O) e Temp PC1 (temperatura média). Consequentemente,

essas características bioclimáticas separando os locais amostrados também afetam a diversidade genética entre os cariomorfos. No geral, esses resultados reforçam a visão de que considerar apenas o cariomorfo ou a localização geográfica não descreve completamente a estrutura populacional atual de *H. malabaricus*.

## CROMOSSOMOS SEXUAIS MÚLTIPLOS E O PROCESSO DE ESPECIAÇÃO

Mudanças estruturais, como fusões e translocações entre um autossomo e um cromossomo sexual (X, Y, Z, W) podem acarretar no surgimento de um sistema de cromossomos sexuais múltiplo. Diferente do processo de supressão de recombinação que ocorre em um sistema simples do tipo XY, ZW ou XO, os sistemas de cromossomos sexuais evoluem, algumas vezes em diversos passos chamados estratos evolutivos (evolutionary strata) (LAHN e PAGE, 1999). O surgimento de um sistema de cromossomos sexuais múltiplo pode imediatamente suprimir a recombinação em regiões próximas aos pontos de quebra (TOUPS et al. 2019; SEMBER et al. 2021). Adicionalmente, além de formar novos grupos de ligação entre genes que estava anteriormente em cromossomos distintos, esses rearranjos aumentam o número de genes ligados ao sexo, e podem acelerar o acúmulo de incompatibilidades genéticas entre diferentes populações (VIEIRA; COELHO; VIEIRA, 2003). Genes que desempenham papel importante no isolamento reprodutivo entre populações podem acumular-se em cromossomos neo-sexuais, levando conseqüentemente à especiação (KITANO e PEICHEL, 2012). No entanto, os aspectos possivelmente mais importantes dos rearranjos envolvendo cromossomos sexuais são suas conseqüências meióticas. Ao contrário dos sistemas cromossômicos sexuais simples, nos quais a segregação meiótica envolve apenas um par de cromossomos, um trivalente meiótico característico é formado nos heterozigotos em espécies com sistemas múltiplos, que podem não segregar regularmente, podendo levar a uma barreira pós-zigótica eficaz por meio da produção de gametas desbalanceados (revisado em ZHANG et al. 2021).

Cromossomos sexuais são importantes no processo de especiação, de diversas outras formas (como revisado em COYNE e ORR 1989). Genes ligados ao X são super-representados entre os fatores que reduzem a aptidão dos híbridos, fenômeno denominado "efeito do X grande" (COYNE e ORR 1989; COYNE, 1992 PRESGRAVES, 2008). A inviabilidade e a esterilidade do híbrido afetam principalmente o sexo heterogamético, um efeito conhecido como regra de Haldane (HALDANE 1922; COYNE 1992; PRESGRAVES 2008). Além disso, os genes nos cromossomos X e Z tendem a apresentar maior divergência de sequência do que os genes dos autossomos (PRESGRAVES, 2018).

Conforme revisado por Charlesworth, Campos e Jackson (2018), quando mutações recessivas aparecem nos cromossomos X e Z, elas são expostas à seleção no sexo hemizigótico, o que pode resultar em fixação mais rápida de mutações adaptativas. Por sua vez, o isolamento reprodutivo é mais fraco em espécies com cromossomos sexuais homomórficos em comparação com aquelas com cromossomos heteromórficos (PRESGRAVES e ORR 1998; LIMA, 2014). Análises genômicas recentes têm frequentemente identificado cromossomos sexuais como *outliers* de diferenciação, nos quais o isolamento reprodutivo evolui mais rápido entre as espécies com cromossomos sexuais heteromórficos, provavelmente devido à divergência de sequência mais rápida (PAYSEUR; PRESGRAVES; FILATOV, 2018).

Por fim, o surgimento de sistema de cromossomos sexuais múltiplos resultantes de rearranjos entre um autossomo e um cromossomo sexual pode criar novos grupos de ligação entre genes de cromossomos distintos, aumentando o número de genes ligados ao sexo nos quais incompatibilidades genéticas entre populações podem se acumular (VIEIRA; COELHO; VIEIRA, 2003). Acredita-se que a formação de um sistema neo-sexual contendo *loci* que afetam a esterilidade de híbridos masculinos e o isolamento comportamental tenha ajudado na especiação em peixes esgana-gatas (KITANO et al. 2009).

Um processo de especiação similar pode estar ocorrendo em *H. malabaricus*. Mesmo que A2 e D1 tenham sido coletados juntos na mesma localidade, a diferenciação genética entre eles foi alta tanto na PCoA quanto no fastStructure, e isso não pode ser completamente explicado pela baixa diversidade presente nessas sequências (**Tabela 3**), já que D1 apresentou o segundo maior valor de  $\pi$  (0.0013), e, apesar de A2 apresentar o menor valor de  $\pi$  (0.0061), este valor é apenas metade do valor de D1. A diferenciação é muito similar àquela entre A2 e A3, e maior do que a observada entre A2 e A4 (**Tabela 3**), que são populações muito distantes e isoladas. As diferenças no número diploide e estrutura dos cromossomos, bem como os sistemas sexuais distintos claramente anulam a possibilidade de cruzamento entre os cariomorfos A e D. Ademais, juntamente com os resultados apresentados, as sequências de *cytb* dos cariomorfos A e D são altamente distintas, e tão diferenciadas quanto as sequências de *H. malabaricus* e *H. microlepis*, reforçando indícios de isolamento reprodutivo (UTSUNOMIA et al. 2014).

## CONSIDERAÇÕES FINAIS

*H. malabaricus* é, há muito tempo, considerado um provável complexo de espécies (BERTOLLO; MOREIRA-FILHO; GALETTI, 1986; DERGAM e BERTOLLO 1990; BERTOLLO et al. 1997; LOPES et al. 1998; BERTOLLO et al. 2000; GRAÇA e PAVANELLI, 2007). Cariomorfos distintos foram observados em simpatria sem que a presença de híbridos fosse detectada (SCAVONE; BERTOLLO; CAVALLINI, 1994; BERTOLLO et al. 1997; LOPES et al. 1998), sugerindo que alguns destes cariomorfos pudessem estar isolados reprodutivamente. O presente estudo mostra isso de forma particularmente clara para os cariomorfos A e D, que, além das diferenças no número de cromossomos, estrutura cariotípica e cromossomos sexuais, também diferem em sequências, mesmo quando em sintopia. O sistema neo-sexual pode ter acelerado a diferenciação genética do cariomorfo D, acelerando seu processo de especiação. Alternativamente, como a população D é restrita a uma área geográfica específica, a fixação de um rearranjo cromossômico relacionado ao sexo pode ter ocorrido por acaso; se os cariomorfos A e D evoluíram durante um período de alopatria, sua simpatria poderia refletir uma migração recente (como sugerido para explicar a alta diversidade dentro da população A4, onde a ausência de diferenças cromossômicas parece ter permitido a mistura de populações divergentes).

Em relação aos outros cariomorfos, especialmente A e F, a manutenção da mesma estrutura cariotípica entre suas diferentes populações não reflete parentesco muito próximo. Diferenças de sequência já evoluíram entre essas populações, indicando considerável evolução independente. No entanto, a divergência de sequência, mesmo entre cariomorfos, não é grande, e o número de alterações cromossômicas, incluindo dois eventos separados de fusões entre cromossomos autossomos e sexuais são marcantes para os tempos de divergência correspondentes a essas pequenas diferenças de sequência, novamente talvez apontando para populações fortemente isoladas no recente passado evolutivo.

## REFERÊNCIAS BIBLIOGRÁFICAS

As referências deste capítulo encontram-se reunidas ao final da dissertação.

## CAPÍTULO 2

# **Integrando citogenética e genômica populacional: a alopatria e os cromossomos neo-sexuais podem ter moldado a divergência genética no complexo de espécies *Erythrinus erythrinus* (Teleostei, Characiformes)**

**Fernando H. S. de Souza<sup>1</sup>, Francisco de M. C. Sassi<sup>1</sup>, Pedro H. N. Ferreira<sup>1</sup>, Luiz A. C. Bertollo<sup>1</sup>, Tariq Ezaz<sup>2</sup>, Thomas Liehr<sup>3\*</sup>, Manolo F. Perez<sup>1</sup> & Marcelo de B. Cioffi<sup>1</sup>**

<sup>1</sup> Laboratório de Citogenética de Peixes, Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos – SP 13565-905, Brazil;

<sup>2</sup> Institute for Applied Ecology, University of Canberra, Canberra, Australia;

<sup>3</sup> University Hospital Jena, Friedrich Schiller University, Institute of Human Genetics, 07747 Jena, Germany.

Publicado no periódico “Biology”.

DOI: [10.3390/biology11020315](https://doi.org/10.3390/biology11020315)

## RESUMO

A diversidade encontrada em peixes neotropicais de água doce é notável. Pode até dificultar uma delimitação adequada de muitas espécies, sendo o peixe-lobo *Erythrinus erythrinus* (Teleostei, Characiformes) um exemplo notável. Esta espécie nominal apresenta notável variação intra-específica, com extensa diversidade cariotípica encontrada entre as populações em termos de diferentes números de cromossomos diplóides ( $2n$ ), composições cariotípicas e sistemas de cromossomos sexuais. Aqui, analisamos três populações distintas (uma delas investigada citogeneticamente pela primeira vez) que diferiram em suas características cromossômicas (denominadas cariomorfos) e pela presença ou ausência de cromossomos sexuais heteromórficos. Combinamos análises citogenética com abordagens genômicas para investigar como a evolução de múltiplos cromossomos sexuais juntamente com a alopatria está ligada à diversidade genética e especiação. Os resultados indicaram a presença de alta diferenciação genética entre as populações tanto do ponto de vista citogenético quanto genômico, sendo a alopatria de longa distância potencialmente o principal agente de divergência genética. Uma população apresentou um sistema cromossômico sexual neo- $X_1X_2Y$  e nós hipotetizamos que esse sistema está associado a uma diferenciação genética interpopulacional aprimorada que poderia ter potencialmente acelerado a especiação em comparação com o efeito da alopatria sozinha.

## INTRODUÇÃO

A biodiversidade dos peixes Neotropicais de água doce é notável, pois compreendem pelo menos 5.200 espécies (FRICKE; ESCHMEYER; VAN DER LAAN, 2021). A família Erythrinidae é amplamente distribuída nesta região, com conhecida preferência por ambientes lênticos como pequenos e grandes rios e lagoas (OYAKAWA, 2003). A família contém atualmente 18 espécies descritas, dentro de três gêneros: *Hoplias* (13), *Hoplerythrinus* (3) e *Erythrinus* (2) (OYAKAWA, 2003). No entanto, alguns autores sugeriram a necessidade de uma revisão taxonômica (BERTOLLO et al. 2004; CIOFFI et al. 2011), pois todos os gêneros apresentam vários morfotipos não descritos, que atualmente estão incluídos na mesma espécie nominal (OYAKAWA e MATTOX, 2009).

Tal diversidade críptica é pelo menos parcialmente explicada pela alta diversidade cariotípica encontrada em algumas espécies, como *E. erythrinus* e *H. malabaricus*, nas quais muitos estudos mostraram a presença de variantes intraespecíficas, com extensas diferenças observadas no número de cromossomos diplóides (2n), composições cariotípicas e diferentes sistemas de cromossomos sexuais entre populações revisadas em Cioffi e colaboradores (2011). Quatro cariomorfos diferentes são encontrados em *E. erythrinus* (classificados como A, B, C e D) com prevalência de cromossomos acrocêntricos, que diferenciam os cariótipos da espécie daqueles de outros eritrinídeos (BERTOLLO et al. 2004; GIULIANO-CAETANO, 1986). Enquanto o cariomorfo A apresenta 2n=54 cromossomos tanto em machos quanto em fêmeas, sem a presença de cromossomos sexuais diferenciados, os cariomorfos B, C e D apresentam um sistema múltiplo de cromossomos sexuais ♀X<sub>1</sub>X<sub>1</sub>X<sub>2</sub>X<sub>2</sub>/♂X<sub>1</sub>X<sub>2</sub>Y, em que machos e fêmeas possuem 2n=51 e 2n=52, respectivamente (BERTOLLO et al. 2004; CIOFFI; MARTINS; BERTOLLO, 2010).

Embora a alopatria seja geralmente considerada o principal fator para o surgimento do isolamento reprodutivo, por meio da redução do fluxo gênico que culmina no processo de especiação (MAYR, 1964), o impacto dos rearranjos cromossômicos na especiação também é amplamente discutido desde o trabalho de Haldane (1922). Em animais, a maioria dos casos de isolamento pós-zigóticos são causados por incompatibilidades genéticas, dentre as quais rearranjos cromossômicos e o surgimento de cromossomos sexuais podem desempenhar um papel fundamental (BATESON, 1909; DOBZHANSKY, 1933; MULLER, 1942; MULLER, 1942; COYNE e ORR, 2004). Os rearranjos cromossômicos têm o potencial de limitar a introgressão, facilitando assim a origem e a manutenção do isolamento reprodutivo por meio da supressão da recombinação (FARIA e NAVARRO, 2010; SICHOVA et al. 2015). Além

disso, rearranjos cromossômicos envolvendo cromossomos sexuais podem levar à formação de novos grupos de ligação entre genes originalmente encontrados em cromossomos diferentes, o que pode acelerar o acúmulo de incompatibilidades genéticas entre grupos de indivíduos (VIEIRA;COELHO;VIEIRA, 2003). Esses rearranjos aumentam a quantidade de genes ligados ao sexo e aceleram o acúmulo de incompatibilidades genéticas entre populações (VIEIRA;COELHO;VIEIRA, 2003). Quando comparadas, espécies que apresentam cromossomos sexuais heteromórficos apresentam isolamento reprodutivo mais forte do que espécies nas quais eles estão ausentes (PRESGRAVES e ORR, 1998; LIMA, 2014). Particularmente, ao contrário do processo gradual de supressão de recombinação que ocorre em sistemas de cromossomos sexuais simples, sistemas de cromossomos sexuais múltiplos podem suprimir imediatamente a recombinação em regiões próximas aos pontos de quebra (TOUPS et al. 2019; SEMBER et al. 2021). Esses sistemas geralmente surgem após eventos de translocação ou fusão entre um autossomo e um cromossomo sexual existente. Na segregação meiótica, forma-se um multivalente meiótico que pode causar problemas na segregação, gerando gametas desbalanceados, acarretando em uma barreira pós-zigótica (revisado em ZHANG et al. 2021). Portanto, a formação de um sistema de cromossomos sexuais múltiplos pode promover uma rápida supressão da recombinação, atuar como uma barreira pós-zigótica e contribuir para o isolamento reprodutivo.

Recentemente, as tecnologias de sequenciamento de alto rendimento iniciaram uma nova era na genética de populações, pois permitem a avaliação de cenários demográficos complexos, como a investigação de como o fluxo gênico e a introgressão afetam a diversidade genética e a especiação (ELLEGREN, 2014). Nesse contexto, nossa abordagem busca combinar ferramentas citogenéticas e genômicas de forma integrada. Para isso, analisamos três populações distintas de *E. erythrinus* (uma delas citogeneticamente investigada pela primeira vez) e utilizamos abordagens genômicas para investigar como a evolução de cromossomos sexuais múltiplos juntamente com a alopatria estão ligadas à diversidade genética e especiação.

Nossa hipótese é que a presença de cromossomos sexuais múltiplos tenha contribuído diferencialmente para a diferenciação genética das populações, o que poderia potencialmente ter acelerado o processo de especiação.

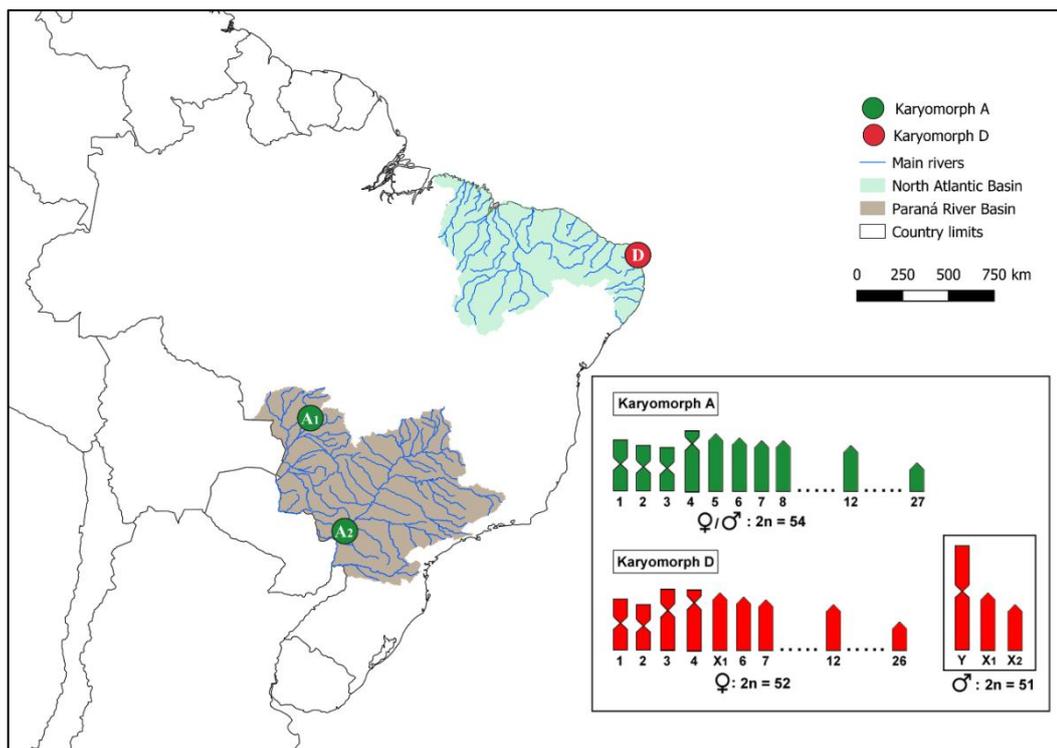
## MATERIAIS E MÉTODOS

### AMOSTRAGEM DOS ESPÉCIMES

Os locais de coleta, quantidade, e sexo dos espécimes investigados estão disponíveis na **Tabela 4** e **Figura 12**. Parte da amostragem (Grupos A1 e D) foi previamente analisada através de diferentes técnicas citogenéticas e moleculares (CIOFFI; MARTINS; BERTOLLO, 2010; MARTINS et al. 2013). Todos os animais foram coletados com autorização apropriada da Agencia Ambiental Brasileira ICMBIO/ SISBIO (Licença 48628-14) e SISGEN (A96FF09). Os experimentos seguiram os protocolos de ética aprovados pelo Comitê de Ética em Experimentação Animal da Universidade Federal de São Carlos (Processo número CEUA 1853260315)

**Tabela 4** - Lista de indivíduos de *E. erythrinus* analisados, evidenciando cariomorfo, número de cromossomos (2n), sistema de cromossomos sexual, código do grupo (indicando o ponto de coleta e o cariomorfo), número de indivíduos analisados citogeneticamente, número de indivíduos sequenciados e referências.

Cariomorfo	2n	Sistema sexual	Código	Ponto de coleta	Latitude/ Longitude	Análises citogenéticas	DArT	Referência
A	♀♂ 54	Homomórfico	A1	Rio Cuiabá (MT)	-16.1713/ -55.9573	14♂11♀	6	MARTINS et al. 2013
A	♀♂ 54	Homomórfico	A2	Rio Paraná (PR)	-23.3774/ -53.7805	09♂07♀	5	<b>Trabalho atual</b>
D	♀52/♂51	♀X <sub>1</sub> X <sub>1</sub> X <sub>2</sub> X <sub>2</sub> ♂X <sub>1</sub> X <sub>2</sub> Y	D	Córrego do Sangue (RN)	-5.7770/ -35.2092	10♂09♀	7	CIOFFI; MARTINS; BERTOLLO 2010



**Figura 12** - Mapa da América do Sul indicando os pontos de coleta para *E. erythrinus*. As cores indicam as diferentes bacias hidrográficas brasileiras, coloridas de acordo com a legenda. Os círculos coloridos indicam os pontos de coleta, com os respectivos cariomorfos coloridos da seguinte forma: A em verde e D em vermelho. O idiograma à direita indica os cariótipos parciais de cada cariomorfo e o seu sistema de cromossomos sexuais, quando presente.

## PREPARAÇÃO CROMOSSÔMICA E HIBRIDIZAÇÃO IN SITU FLUORESCENTE (FISH) PARA MAPEAMENTO DO DNAR

Neste trabalho apresentamos os primeiros dados citogenéticos para os indivíduos da localidade A2. Os cromossomos mitóticos foram obtidos através do protocolo descrito por Bertollo e colaboradores (2017). As sondas de DNAr 5S incluíram a região codificante do DNAr 5S, com 120 pares de base (bp), e o espaçador longo não transcrito de 200 bp (NTS) (PENDAS et al. 1994). As sondas de DNAr 18S foram obtidas por PCR utilizando os primers descritos em Cioffi et al. (2019). As sondas de DNAr 18S e 5S foram marcadas diretamente com o kit de marcação Nick-translation labeling kit (Jena Bioscience, Jena, Alemanha), onde o DNAr 18S foi marcado com o Atto488 (fluorescência verde) e o DNAr 5S com Atto550 (fluorescência vermelha), de acordo com o manual do fabricante. O FISH foi realizado sob condições de alta estrigência descritas em Yano, Bertollo e Cioffi (2017).

## HIBRIDIZAÇÃO GENÔMICA COMPARATIVA (CGH)

Os experimentos de hibridização genômica comparativa (CGH) foram realizados de acordo com o protocolo descrito em Sember et al. (2018) com adaptações na proporção sonda/DNA

Cot-1 baseada em estudos anteriores (SEMBER et al. 2018; TOMA et al. 2019; DE MORAES et al. 2021). As sondas genômicas de macho para todas as populações foram co-hibridizadas aos cromossomos dos machos das localidades A1 e D separadamente. Para este esquema experimental, DNAs de macho derivados do genoma-total (gDNAs) de todas as populações analisadas foram extraídos do tecido de fígado pelo método padrão fenol-clorofórmio-isoamil-álcool (SAMBROOK e RUSSELL, 2001) e marcados com ATTO550-dUTP e AF488-dUTP utilizando o kit Nick Translation mix kit (Jena Bioscience, Jena, Alemanha). Para bloquear o excesso de sequências repetitivas compartilhadas, um DNA Cot-1 não marcado, obtido de cada população de acordo com Zwick et al. (1997) foi incluso na mistura de sondas final. Para cada lâmina, sondas genômicas de macho (500 ng cada) e 15 µg de DNA Cot-1 derivado de fêmea das respectivas populações foram precipitados em 96% de etanol, e os pellets secos foram ressuspensos em 20 µl de solução tampão de hibridização contendo o 50% formamida, 2×SSC, 10% sulfato de dextrana, e tampão de Denhardt (pH 7.0). O FISH foi realizado sob condições de alta estrigência descritas em Yano, Bertollo e Cioffi (2017).

## MICROSCOPIA E PROCESSAMENTO DE IMAGEM

Para confirmar o número diploide, a estrutura cariotípica e os resultados de hibridização in situ fluorescente, pelo menos 30 metáfases foram analisadas por indivíduo. As imagens de microscopia foram capturadas utilizando um microscópio de epifluorescência Olympus BX50 (Olympus Corporation, Ishikawa, Japão) acoplado a uma câmera CoolSNAP, e as imagens foram processadas utilizando o software Image-Pro Plus 4.1 Software (Media Cybernetics, Silver Spring, MD, Estados Unidos da América). As imagens finais foram otimizadas e organizadas utilizando o software Adobe Photoshop, versão CC 2020. Os cromossomos foram classificados como metacêntrico (m), submetacêntrico (sm), subtelocêntrico (st), ou acrocêntrico (a), de acordo com a proporção do tamanho dos braços (LEVAN; FREDGA; SANDBERG, 1964). Tendo em vista que os resultados de machos e fêmeas não apresentaram diferenças, apenas as metáfases de macho foram apresentadas. Os idiogramas foram construídos no Adobe Photoshop, versão CC 2020, com base no estudo anterior e atual.

## SEQUENCIAMENTO E FILTRAGEM DOS DADOS

O procedimento de sequenciamento DArTseq foi realizado usando tecidos hepáticos, que foram enviados para a Diversity Arrays Technology Pty Ltd. (Austrália). Este método envolve uma etapa de redução de complexidade em que o DNA é tratado com as enzimas PstI e SbfI antes do sequenciamento, o que pode enriquecer o material para sequências em regiões

pouco metiladas (KILIAN et al. 2018). O sequenciamento foi realizado em uma plataforma Illumina HiSeq 2500. Esta técnica de sequenciamento foi implementada em vários artigos recentes investigando a genômica de populações de peixes (CIOFFI et al. 2019; SOUZA et al. 2019; SASSI et al. 2021). O processamento dos dados brutos e todas as etapas subsequentes para a obtenção do conjunto de dados foi executado utilizando o software pyRAD v3.0.66 (EATON, 2014). O parâmetro de *depth* de cobertura mínimo para a formação de um *cluster* foi definido como 6. O limite de *clustering* implementado foi de 0,88. Na última etapa de alinhamento, a cobertura mínima de amostras para um *locus* foi definido como 18 (número de indivíduos sequenciados) portanto apenas os dados presentes em todos os indivíduos foram mantidos. Os adaptadores de sequenciamento foram removidos e as sequências com mais de cinco bases de baixa qualidade foram retiradas ( $Q < 20$ ). Posteriormente, os *reads* foram alinhados e foram obtidas as sequências consenso para cada amostra, que foram utilizadas para definir a base de referência em cada posição em cada alinhamento. Obtivemos *loci* que estavam presentes em todos os indivíduos amostrados e geramos dois conjuntos de dados diferentes para análises posteriores: (1) um arquivo de sequência com a fase definida para cada *loci* baseado no formato “.alleles”, que foi implementado apenas nas análises de diversidade e diferenciação genética e (2) uma matriz de polimorfismo de nucleotídeo único (SNPs), com um único SNP por *locus* para evitar a presença de SNPs intimamente ligados. A matriz de SNPs foi codificada da seguinte forma: 0 para homozigotos para a base de referência, 1 para heterozigotos e 2 para os homozigotos de base alternada. *Indels* não foram considerados, e este segundo *input* foi utilizado nas demais análises.

## DETECÇÃO DE MARCADORES POTENCIALMENTE SOB SELEÇÃO

Uma análise no software BayeScan foi realizada para testar marcadores com níveis incomuns de diferenciação interpopulacional (baixo ou alto), o que pode indicar alelos adaptativos nos *loci* sequenciados (FOLL e GAGGIOTTI, 2008). Realizamos a análise com um valor de *prior odds* de 100, um *thinning* de 10 e 5000 cadeias MCMC. Todos os *loci* com valores de *False Discovery Rate* (FDR) menores que 0,01 foram classificados como possivelmente sob seleção.

## DIVERSIDADE GENÉTICA

Os índices de diversidade genética foram obtidos com o software DnaSP v. 6.12.03 (ROZAS et al. 2017). Calculamos as médias de diversidade haplotípica ( $H_d$ ), o D de Tajima ( $D$ ), e duas medidas de diversidade nucleotídica por sítio,  $\pi$  e Theta de Watterson ( $\theta_w$ ). Nós

também estimamos a divergência nucleotídica entre pares de amostras de diferentes locais utilizando o número médio de substituições nucleotídicas por sítio ( $D_{XY}$ ) e a divergência líquida, corrigida pela variação dentro das amostras analisadas ( $D_A$ ).

## ANÁLISES DE ESTRUTURA POPULACIONAL

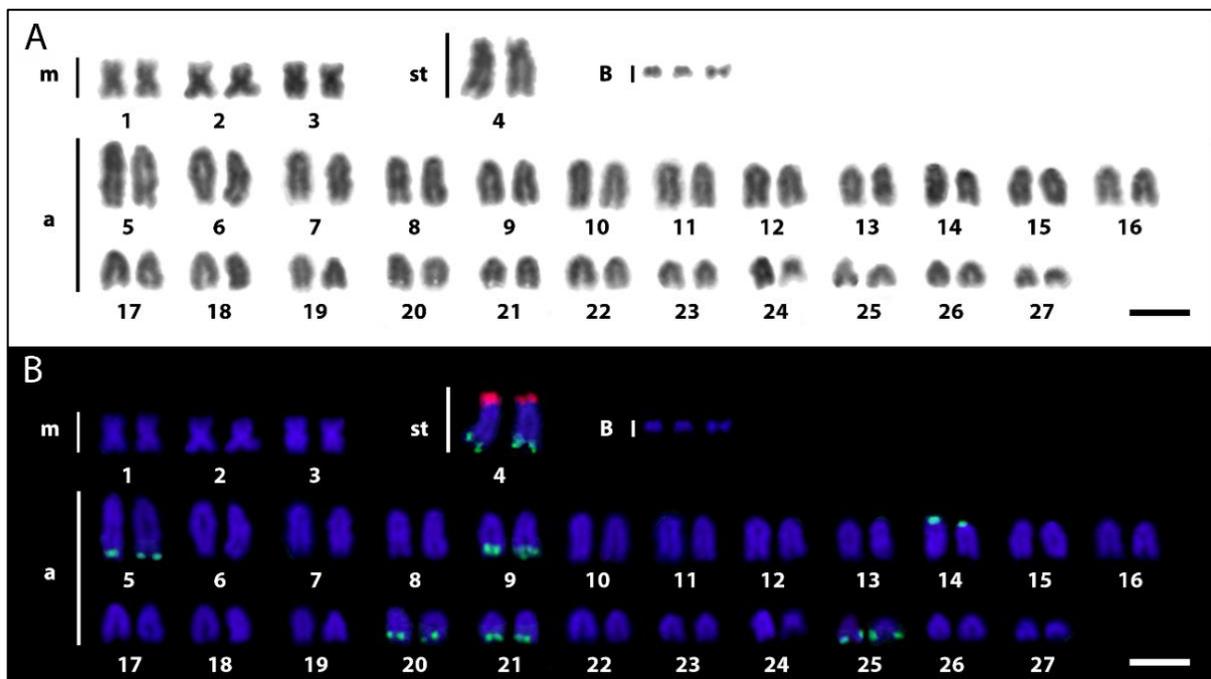
Avaliamos a estrutura populacional com uma Análise de Coordenadas Principais (PCoA) para visualizar a distribuição da diversidade genética em cada um dos cariomorfos. A PCoA foi realizada utilizando a função “gl.pcoa” do pacote dartR. Em seguida, implementamos a abordagem Bayesiana fastStructure (RAJ; STEPHENS; PRITCHARD, 2014). Este método é análogo ao utilizado pelo software amplamente utilizado STRUCTURE (PRITCHARD; STEPHENS; DONNELLY, 2000), mas otimizado para conjuntos de dados maiores. Realizamos a análise com as configurações padrão e o valor máximo de K definido como quatro. Todos os resultados foram compactados e enviados para Clumpak (KOPELMAN et al. 2015) para gerar uma representação gráfica dos resultados. Além disso, realizamos a abordagem espacial no pacote ConStruct (BRADBURD; COOP; RALPH, 2018) pois os peixes de água doce são um dos grupos mais suscetíveis aos efeitos de confusão do isolamento por distância na estrutura populacional (PEREZ et al. 2018). Executamos a análise no Rstudio com parâmetros espaciais ativados e distâncias par-a-par com base nas coordenadas dos locais de amostragem. Realizamos a análise com valores de K variando de 1 a 4, com um total de 1.000.000 de iterações para cada valor de K. Após a conclusão das corridas, analisamos a contribuição de cada camada de K para a estrutura populacional de *E. erythrinus*, medida pelo seu impacto na covariância do modelo. Portanto, quando um determinado valor de K não contribuiu mais para o resultado, o valor de K anterior foi selecionado (BRADBURD, 2019).

## RESULTADOS

### CARIÓTIPOS E MAPEAMENTO DE DNA RIBOSSÔMICO

O cariótipo do grupo recém analisado A2 mostrou características gerais apresentadas pelo cariomorfo A de *E. erythrinus* previamente analisados em outra localidade, ou seja, muitos cromossomos acrocêntricos e poucos cromossomos com dois braços. O cariótipo apresentou  $2n= 54$  ( $6m + 2st + 46a$ ), sem cromossomos sexuais diferenciados. Adicionalmente, foram observados até 3 cromossomos supranumerários B, que aparecem como microcromossomos e coram mais fracamente comparado aos demais (**Figura 13**). Das 30 metáfases analisadas, as metáfase de número 20, 08 e 02 apresentaram 3, 2 e 1 cromossomos B, respectivamente. Padrões semelhantes de cariótipo, de FISH e CGH foram observados para todas as metáfases analisadas em todos os cariomorfos.

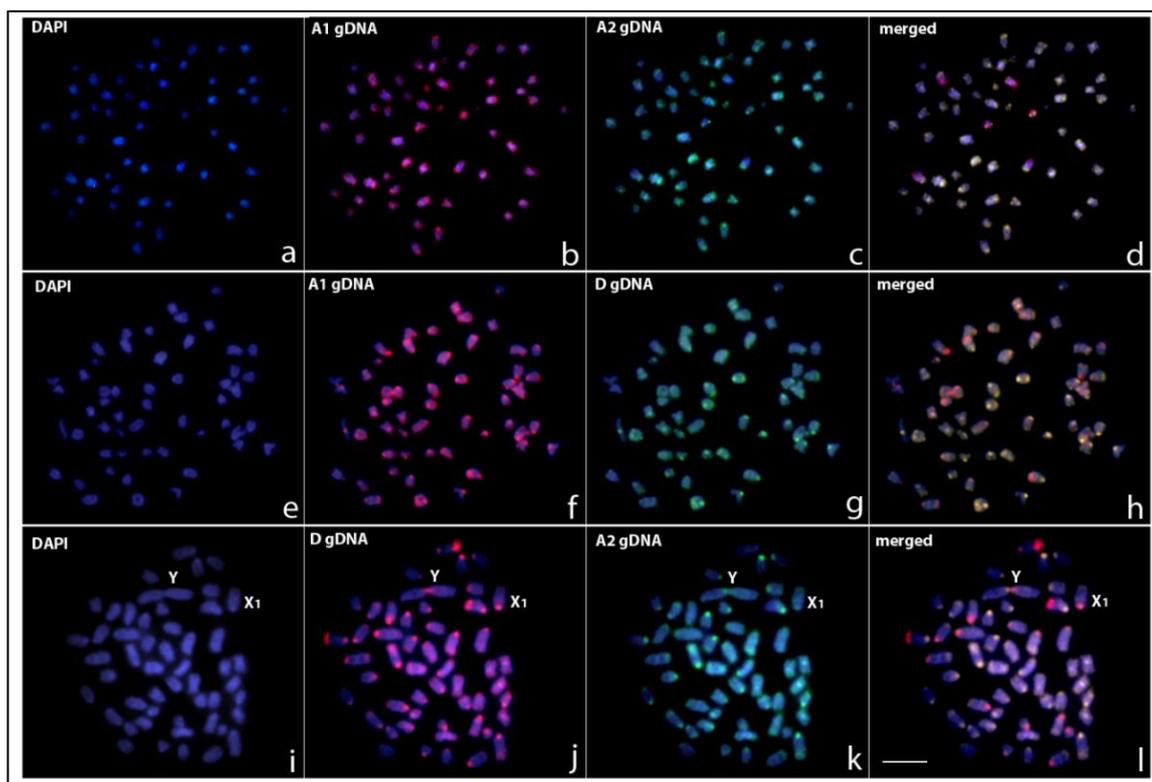
Sítios de DNAr 18S foram mapeados nas regiões terminais dos braços longos de seis pares acrocêntricos (pares 5, 9, 20, 21 e 25) e no par subtlocêntrico 4, além de um par de cromossomos com sítios centroméricos (par 14). Por sua vez, apenas um par acrocêntrico foi observado com sítios de DNAr 5S na região telomérica dos braços curtos. Esse mesmo par de cromossomos também possui sítios de DNAr 18S em seus braços longos, evidenciando uma condição sintênica incomum em relação a essas duas classes de DNAs ribossômicos (**Figura 13**).



**Figura 13** - Cariótipos de *E. erythrinus* do cariomorfo A da localidade A2 analisados pelo método de coloração com Giemsa (A) e FISH com sequências de DNAr 18S (verde) e 5S (vermelho) como sondas (B). Tamaho da barra: 5  $\mu$ m.

## HIBRIDIZAÇÃO GENÔMICA COMPARATIVA

Os experimentos de CGH tiveram o objetivo de comparar os genomas entre as diferentes populações estudadas. A hibridização genômica comparativa das sondas de gDNAs para os indivíduos A1 e A2 produziu muitos sinais sobrepostos, o que evidencia as suas similaridades genômicas (com uma ligação mais forte preferencialmente em regiões heterocromáticas terminais ou pericentroméricas) e sugerem homologia significativa. Alguns sinais população-específicos observados indicam alguns sinais super-representados, provavelmente ligados a um número maior de cópias de sequências repetitivas em algumas regiões cromossômicas (**Figura 14a-d**). Por outro lado, a hibridização comparativa intra-cariomorfo entre os grupos A1 e A2 com o cariomorfo D produziu vários sinais população-específicos com apenas um número limitado de sinais sobrepostos mostrando um nível mais alto de diferenciação genômica (**Figura 14e-l**).



**Figura 14** - Cromossomos mitóticos de machos de *E. erythrinus* após o CGH – comparação entre cariomorfos. (a-d) A sonda genômica derivada de machos do cariomorfo A1 (b) e A2 (c) mapeadas contra cromossomos de macho do cariomorfo A1. (e-h) Sondas genômicas derivadas de machos d (f) e D (g) mapeadas contra cromossomos de machos do cariomorfo A1. (i-l) Sondas genômicas de macho do cariomorfo de D (j) e A2 (k) mapeadas contra cromossomos de macho do cariomorfo D. Na primeira coluna (a,e,i): imagens de DAPI (azul); Segunda e terceira colunas (b,c,f,g,j,k): padrão de hibridização utilizando as sondas derivadas de macho (vermelho) de cada cariomorfo analisado; Quarta coluna (d,h,l): imagens sobrepostas de ambas as sondas genômicas e coloração com DAPI. As regiões genômicas comuns de ambos os cariomorfos comparados são representadas em amarelo. Tamanho da barra = 10  $\mu$ m.

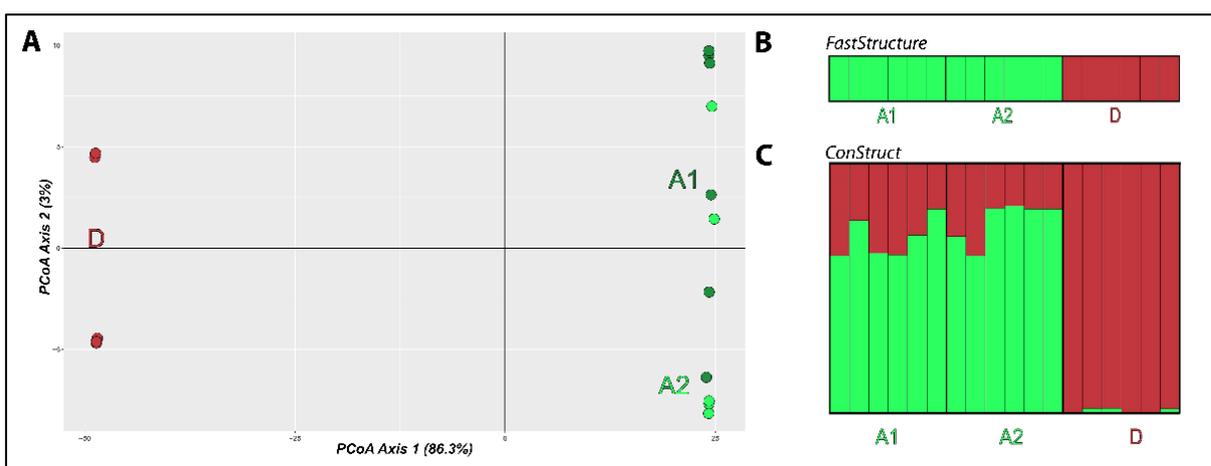
## SEQUENCIAMENTO, PREPARO DOS DADOS E BUSCA DE MARCADORES SOB SELEÇÃO

O sequenciamento gerou cerca de 2 milhões de *reads* para cada amostra. Após as etapas de *trimming* e filtragem, obtivemos 14.467 *loci* com a fase determinada presentes em ao menos seis indivíduos, que então denominamos Conjunto de Dados 1. Em seguida, criamos o Conjunto de Dados 2 mantendo apenas um SNP por *locus*, o que resultou em 8.597 SNPs polimórficos. A análise no Bayescan não considerou nenhum *locus* como potencialmente sob seleção, portanto, nenhum marcador foi removido das análises subsequentes.

### ESTRUTURA POPULACIONAL

A PCoA indica uma clara separação entre os cariomorfos A e D. O cariomorfo D formou dois grupos menores que estão intimamente relacionados, e os cariomorfos A1 e A2 formaram um grupo maior cada, com alguns indivíduos espalhados entre esses dois grupos (**Figura 15A**).

A abordagem bayesiana não geográfica fastStructure foi congruente com o resultado da PCoA e apontou K=2 como o valor que maximiza a probabilidade posterior e melhor explica os dados. A análise agrupou A1 e A2 como uma única população e o cariomorfo D como uma segunda população (**Figura 15B**). A abordagem espacial ConStruct apresentou um resultado que também é congruente com os resultados do PCoA e do fastStructure. Os indivíduos do cariomorfo A são muito semelhantes, enquanto o cariomorfo D formou um grupo distinto (**Figura 15C**).



**Figura 15-** (A) Análise de Coordenadas Principais. Indivíduos estão representados em cores diferentes: A1 em verde escuro; A2 em verde claro; D em vermelho. (B) Resultado do fastStructure para K=2. Cada barra vertical representa um indivíduo, cariomorfos e locais de coleta estão indicados abaixo das barras, a cor das barras indica a população na qual cada indivíduo foi classificado. (C) Resultado do ConStruct para K=2. Cada barra vertical representa um indivíduo, cariomorfos e locais de coleta estão indicados abaixo das barras.

## DIVERSIDADE GENÉTICA

De forma geral, todos os grupos analisados apresentaram baixos valores de diversidade genética. Os valores de diversidade nucleotídica ( $\pi$ ) e Theta de Watterson ( $\theta_w$ ) foram mais altos na localidade A1 enquanto o menor valor foi encontrado para o cariomorfo D. Todos os grupos amostrados apresentaram um valor negativo de D de Tajima, com o menor valor sendo apresentado por A1, seguido por A2 e então D. Os valores de diferenciação genética  $D_{XY}$  e  $D_A$  dentro do cariomorfo A foram aproximadamente dez vezes menores quando em comparação com os valores entre os cariomorfos A e D (**Tabelas 5 e 6**).

**Tabela 5** – Valores médios de diversidade genética estimada por ponto de coleta. Estão evidenciados os tamanhos amostrais, a diversidade haplotípica ( $H_d$ ), diversidade nucleotídica ( $\pi$ ), Theta de Watterson por sítio ( $\theta_w$ ), e D de Tajima ( $D$ ).

<b>Código</b>	<b>Tamanho amostral</b>	<b><math>H_d \pm DP</math></b>	<b>p</b>	<b><math>q_w</math></b>	<b>D</b>
A1	6	0,0529 (0,0584)	0,00102	0,00125	-0,4848
A2	6	0,0430 (0,0502)	0,00085	0,00093	-0,25292
D1	12	0,0211 (0,0353)	0,00044	0,00047	-0,19581

**Tabela 6** - Valores de  $D_{XY}$  par-a-par por ponto de coleta indicados na diagonal superior, e valores de  $D_A$  par-a-par na diagonal inferior.

	<b>A1</b>	<b>A2</b>	<b>D</b>
<b>A1</b>		0,00102	0,01084
<b>A2</b>	0,00012		0,01082
<b>D</b>	0,01014	0,01022	

## DISCUSSÃO

### DIVERSIDADE CROMOSSÔMICA ENTRE POPULAÇÕES DE *E. ERYTHRINUS*

Uma macroestrutura cromossômica geral conservadora dos cariomorfos de *E. erythrinus* não é mantida quando se comparam características intracromossômicas das subpopulações estudadas. Isso é evidenciado pelas distribuições diferenciais da fração de DNA repetitivo entre eles (**Figuras 14 e 15**). As novas características cromossômicas do grupo A2 de *E. erythrinus* concordam com as encontradas em outras populações pertencentes aos cariomorfos A, ou seja,  $2n=54$ , cariótipos dominados por cromossomos acrocêntricos e múltiplos sítios de DNAr (**Figura 13**). Enquanto a maioria dos peixes eritrínídeos tende a manter os cariótipos dominados por cromossomos com dois braços, *E. erythrinus* apresenta exceções a essa regra geral (BERTOLLO et al. 2004; CIOFFI; MARTINS; BERTOLLO, 2010; MARTINS et al. 2014, presente estudo).

O mapeamento de DNA ribossômico tem sido amplamente utilizado em muitas investigações citogenéticas modernas (revisado em SOCHOROVÁ et al. 2018; SOCHOROVÁ et al. 2021), constituindo um marcador valioso para citotaxonomia (CIOFFI et al. 2012; DEON et al. 2020; PAZZA; KAVALCO; BERTOLLO, 2006; FERREIRA et al. 2017). Em peixes, a variação do número de cópias de DNAr é frequentemente observada, uma vez que seus processos de regulação gênica parecem ser mais relaxados em comparação com vertebrados superiores (GORNUNG e DARWIN, 2013). Para *E. erythrinus*, o DNAr 5S provavelmente está associado ao transposon *Rex3*, o que pode explicar a alta dispersão desta sequência em seus cromossomos (CIOFFI; MARTINS; BERTOLLO, 2010). O cariomorfo A1 apresenta sinais nos braços curtos do 8º par, enquanto A2 nos braços p do grande cromossomo subtelocêntrico (par 4). Além disso, o cariomorfo D apresenta dez cromossomos com sequências 5S, principalmente nos braços curtos (**Figura 16**). Um número substancial de espécies de peixes apresenta sítios de DNA ribossômico mapeados em cromossomos sexuais heteromórficos (SOCHOROVÁ et al. 2021), o que pode ser explicado pela presença de *Evolutionary Breakpoint Regions* (EBRs), que são propensas a promover regiões de quebras de fita dupla sendo observadas em vários grupos de peixes (DEON et al. 2020; RUIZ-HERRERA; CASTRESANA; ROBINSON, 2006; CARBONE et al. 2009; FARRÉ et al. 2011). Essa característica é evidente no cariomorfo D, onde as sequências de DNAr 5S podem ser identificadas nos cromossomos  $X_1$  e  $X_2$ , assim como no cromossomo neo-Y (CIOFFI; MARTINS; BERTOLLO, 2010) (**Figura 16**). Por outro lado, o DNAr 18S apresenta um padrão de distribuição mais conservado nos cromossomos em *E. erythrinus*, uma vez que ambos os

cariomorfos A1 e D revelam os mesmos resultados de hibridização para esta sonda. No entanto, o cariomorfo A2 (aqui descrito) apresenta seis cromossomos mapeados com a sonda de DNAr 18S (pares 5, 9, 20, 21 e 25 na região terminal dos braços q e pares 14 p). Diferenças na distribuição do DNAr 18S entre populações já foram descritas para peixes da mesma família, como no caso de *H. malabaricus* (revisado em SEMBER et al. 2018). Uma investigação em três ordens distintas de peixes (Characiformes, Siluriformes e Perciformes) também revelou uma associação entre a qualidade do ambiente aquático e a dispersão de sequências de DNAr (DA SILVA et al. 2017).

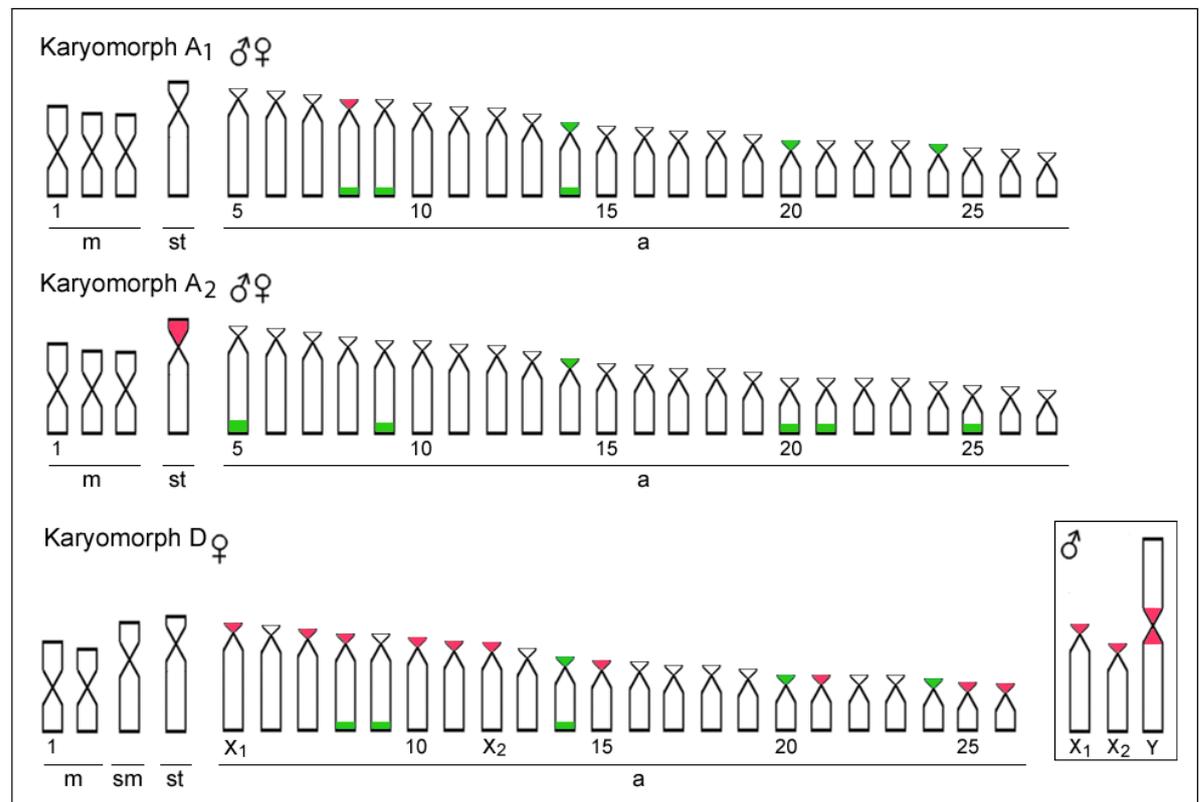


Figura 16 - Idiogramas representativos dos cariomorfos A1, A2, e D de *E. erythrinus* evidenciando a distribuição do DNAr 18S (verde) e DNAr 5S (vermelho). Os cromossomos sexuais encontram-se representados no retângulo.

O conjunto de nossos experimentos de CGH teve como objetivo comparar os genomas entre as diferentes populações estudadas. Tais comparações foram consideradas muito eficazes na identificação da biodiversidade oculta em peixes (revisado em SEMBER et al. 2018) e, neste estudo, elas abriram novas visões e ampliaram nossa compreensão dos processos de diferenciação de genoma entre cariomorfos que estão em andamento. O grau de divergência genômica concorda com suas características cromossômicas, uma vez que os cariomorfos exibiram um estágio avançado de divergência de sequência (**Figura 14**). Embora populações distintas do cariomorfo A possuam uma macroestrutura cariotípica conservada, também foram

observadas divergências genômicas intracariomorfo. Esses resultados são indicativos de processos evolutivos em andamento que podem ser responsáveis por conduzir a divergência e possivelmente também à especiação dentro dos cariomorfos de *E. erythrinus*.

Entre os peixes, espécies com macroestruturas cariotípicas conservadas também são relatadas como tendo um padrão variável de organização genômica, como observado para *Lebiasina* (LEITE et al. submetido à publicação), para as trairas gigantes (SASSI et al. 2021) e para algumas espécies de Osteoglossiformes (BARBY et al. 2018). Cromossomos sexuais diferenciados também parecem desempenhar um papel importante na divergência de regiões genômicas repetitivas, uma vez que espécies que apresentam tais características também apresentam padrões distintos de hibridização em CGH, como observado em bagres (SASSI et al. 2020), traíras (DE OLIVEIRA et al. 2017), peixes-lápis (MORAES et al. 2019) e caraciformes com um sistema estável de cromossomos sexuais ZZ/ZW (YANO et al. 2021). De fato, o papel dos cromossomos sexuais na divergência genômica entre espécies relacionadas não está bem estabelecido. Nossos resultados citogenéticos e genômicos apontam para uma maior divergência genômica na população de *E. erythrinus* que apresenta um sistema de cromossomos sexuais diferenciado (cariomorfo D). Isso está de acordo com descobertas anteriores que evidenciam que o surgimento de um novo *locus* ligado ao sexo ou a translocação de um já existente é um dos passos fundamentais na diferenciação dos cromossomos sexuais dos peixes (BEUKEBOOM e PERRIN, 2014; SEMBER et al. 2021). Tal *turnover* pode ser responsável pelo padrão maior de divergência genômica observado no CGH (**Figura 14**).

## DIVERSIDADE GENÉTICA E ESTRUTURA POPULACIONAL

Este é o primeiro estudo que investigou a diversidade genética em *E. erythrinus*, portanto, só é possível comparar nossos resultados com os de outras espécies. O valor médio da diversidade nucleotídica ( $\pi$ ) para *E. erythrinus* foi de 0,00077, enquanto estudos em outras espécies de peixes encontram valores mais elevados, como em *Nassau grouper* (0,3566 com dados de RAD-Seq) (SHERMAN et al. 2020). Por outro lado, um valor médio semelhante de  $\pi$  foi obtido para *H. malabaricus* com dados DArTSeq (0,001172) (SOUZA et al. submetido à publicação). Essa estimativa de diversidade menor quando comparada ao RAD-seq, mas semelhante ao DArTSeq era esperada devido às características do método de sequenciamento implementado em nosso estudo, uma vez que o procedimento de DArTSeq pode enriquecer regiões que estão potencialmente sendo expressas e, portanto, podem estar sob restrições

seletivas (KILIAN et al. 2018). É importante notar, no entanto, que a análise Bayescan não encontrou nenhum marcador potencialmente sob seleção.

Além disso, nossos resultados indicam que os cariomorfos A e D já são altamente diferenciados tanto em características cromossômicas quanto genéticas. Ao considerar as duas métricas de diferenciação aqui implementadas (valores de  $D_{XY}$  e  $D_A$ ) fica evidente que os grupos A1 e A2 são muito semelhantes, enquanto o cariomorfo D é muito mais diferente desses grupos, com valores pelo menos dez vezes maiores. Todos os grupos apresentaram valores negativos de  $D$  de Tajima, que mede a diferença entre o número médio de diferenças par-a-par e o número de sítios segregantes. Existem três hipóteses principais para esse resultado: (i) a seleção purificadora está agindo removendo a variação desses grupos e mantendo sua diversidade genética estável (o que também pode ser consequência da metodologia DArTSeq); (ii) a deriva genética pode estar influenciando a diversidade genética desses grupos, principalmente por serem sedentários e, portanto, podem favorecer a ocorrência de endogamia; (iii) o tamanho da população era baixo e então uma expansão de tamanho da população ocorreu recentemente.

Os resultados atuais e os padrões de divergência observados são provavelmente explicados por dois fatores principais, o primeiro é que se espera que amostras do mesmo cariomorfo sejam mais semelhantes, embora isso não seja uma regra geral. O segundo fator é a distância entre esses grupos e, conseqüentemente, o tempo provável em que ocorreu o evento de isolamento. Todos os grupos do cariomorfos A estão na Bacia do Rio Paraná, embora estejam em regiões diferentes. No entanto, a Bacia do Paraná é muito fragmentada no sentido de que possui cerca de 450 barragens (PAIVA, 1982). Adicionalmente, *E. erythrinus* apresenta comportamento sedentário e não é uma espécie migratória. Levando em conta todos esses fatos, embora as duas localidades de amostragem estejam na mesma bacia geográfica, elas são consideravelmente isoladas. Ao comparar os cariomorfos A e D temos um cenário diferente, pois além de estarem geograficamente distantes entre si, também são diferentes quanto à presença de cromossomos sexuais, o que poderia acelerar a diferenciação dessas populações e potencialmente levar à especiação.

As análises de estrutura populacional (**Figura 15**) indicaram que existem dois grupos principais de indivíduos separados de acordo com seu cariomorfo. Como pode ser visto na PCoA, os cariomorfos A e D estão muito distantes no eixo horizontal que compreende 86,3% da variação. Olhando atentamente para cada cariomorfo no eixo vertical, que compreende 3% da variação, existem dois grupos distintos de indivíduos dentro do cariomorfo D. Além disso, dentro do cariomorfo A há uma pequena diferença entre os grupos A1 e A2, mas muitos

indivíduos são distribuídos ao longo do eixo Y. Isso indica que há uma maior diversidade dentro e uma baixa diferença entre as localidades do cariomorfo A. Os resultados da abordagem da estrutura populacional não espacial Bayesiana fastStructure (**Figura 15B**), da PCoA e da abordagem geográfica ConStruct (**Figura 15C**) são todos congruentes. De acordo com essas análises, existem duas populações principais onde o cariomorfo A compreende uma população e o cariomorfo D compreende outra. Além disso, é possível observar na **Figura 15C** que os indivíduos dos grupos A1 e A2 apresentam maior diversidade, com composição genética mais diversificada. Esses resultados também estão de acordo com os resultados citogenéticos, que também destacaram essa variabilidade cromossômica interna. Em contraste, os indivíduos cariomorfos D apresentam uma composição genética quase uniforme.

## SISTEMA DE CROMOSSOMOS SEXUAIS MÚLTIPLOS E O PROCESSO DE ESPECIAÇÃO

Os resultados indicam que a presença do sistema de cromossomos sexuais múltiplos no cariomorfo D pode ter contribuído para a fixação ou aumento de diferenças genéticas, o que pode ter estimulado isolamento reprodutivo e consequente especiação no processo evolutivo dentro de *Erythrinus*.

Entre os eritrinídeos, foi recentemente destacado em outro estudo que a presença de um sistema de cromossomos sexuais múltiplos pode estar associada ao processo de especiação (SOUZA et al. submetido à publicação). Em *H. malabaricus*, uma espécie que também é considerada “um complexo de espécies” e é dividida em sete cariomorfos distintos que também variam em sua estrutura cariotípica  $2n$ , e a presença de sistemas de cromossomos sexuais simples e múltiplos (BERTOLLO et al. 2000). Análises genômicas e citogenéticas envolvendo várias populações de diferentes cariomorfos reforçaram que *H. malabaricus* é provavelmente um complexo de espécies e não uma única espécie e destacaram o potencial papel do sistema de cromossomos neo-sexuais para a fixação de diferenças genéticas entre populações, favorecendo assim o processo de especiação (SOUZA et al. submetido à publicação). Em *H. malabaricus*, indivíduos dos cariomorfos A (cromossomos sexuais homomórficos) e D (sistema múltiplo ♀ $X_1X_1X_2X_2$ /♂ $X_1X_2Y$ ) foram coletados em sintopia (mesma localização) e apresentaram alta diferenciação genética. De acordo com os valores de  $D_{XY}$ , a diferenciação genética entre indivíduos sintópicos dos cariomorfos A e D foi equivalente à diferenciação genética de populações alopátricas de longa distância do cariomorfo A. Esses dados nos permitiram levantar a hipótese de que a presença de um sistema de cromossomos neo-sexuais pode ter promovido um aumento na diferenciação genética do cariomorfo D, potencialmente

acelerando seu processo de especiação (SOUZA et al. submetido à publicação). Além disso, outros estudos indicaram que o surgimento de cromossomos neo-sexuais pode ter contribuído para o processo de especiação, como demonstrado em sticklebacks (KITANO et al. 2009) e na planta *Rumex hastatulus* (BEAUDRY; BARRETT; WRIGHT, 2020). Em todos os casos mencionados acima, eventos de *turnover* dos cromossomos sexuais que levaram ao surgimento de sistemas de cromossomos neo-sexuais com potencial de fixar a variação genética principalmente devido à rápida evolução desses cromossomos (BEAUDRY; BARRETT; WRIGHT, 2020).

## CONCLUSÕES

Nossos resultados indicaram a presença de uma alta diferenciação genética nas populações do complexo de espécies *E. erythrinus*, tanto do ponto de vista citogenético quanto genético. Embora os índices de diversidade genética sejam globalmente baixos, os cariomorfos A e D são muito diferenciados geneticamente. Além disso, três hipóteses principais para a diferenciação encontrada podem ser propostas: (I) a alopatria de longa distância pode ter sido o principal agente de diferenciação genética entre esses cariomorfos; (II) o surgimento de cromossomos neo-sexuais pode ter sido o principal fator que promoveu essa diferenciação genética; (III) tanto a alopatria de longa distância quanto o surgimento dos cromossomos neo-sexuais no cariomorfo D podem ter contribuído para a fixação e acúmulo de divergências. A hipótese III é nossa hipótese principal. Em resumo, nossos dados forneceram uma camada adicional de evidência sobre o status do táxon *E. erythrinus* e corroboraram estudos anteriores na conclusão de que este é provavelmente um complexo de espécies com taxonomia não resolvida. Estudos futuros usando sequenciamento de alto rendimento com foco na determinação do sexo neste complexo ajudarão a destacar a possível ligação dos cromossomos sexuais com o processo de especiação neste grupo.

## REFERÊNCIAS BIBLIOGRÁFICAS

As referências deste capítulo encontram-se reunidas ao final da dissertação.

## 5. CONSIDERAÇÕES FINAIS

De forma geral, por meio do presente trabalho foi possível avançar na compreensão acerca das possíveis relações evolutivas e da hipótese de possível contribuição de sistemas de cromossomos sexuais múltiplos na diferenciação de peixes da família Erythrinidae.

Os dados aqui apresentados evidenciaram que não é possível classificar os grupos de indivíduos de *H. malabaricus* apenas com base na informação de seus cariomorfos, pois a localização geográfica dos indivíduos e as variáveis ambientais também afetam as características genéticas destes indivíduos e precisam ser integradas, tornando a problemática ainda mais complexa. Adicionalmente, os resultados obtidos reforçaram a ideia de que *H. malabaricus* é na verdade um complexo de espécies. A presença dos cariomorfos A e D em sintopia, apresentando altos valores de diferenciação genética contribuem para esta ideia, e também indicam que a presença do sistema de cromossomos neo-sexuais do tipo  $X_1X_1X_2X_2/X_1X_2Y$  no cariomorfo D pode ter contribuído para o isolamento reprodutivo, a fixação de diferenças genéticas, e conseqüentemente ao processo de diferenciação destas espécies. Entretanto, é possível que os cariomorfos A e D tenham evoluído em alopatria, e sua atual condição simpátrica seja fruto apenas de um evento de migração recente. De toda forma, a contribuição da alopatria para a diversidade genética de diferentes grupos de indivíduos pode ser claramente evidenciada ao analisar os indivíduos do cariomorfo A e sua grande diversidade e diferenciação genética.

Com relação a *Erythrinus erythrinus*, assim como *H. malabaricus* este grupo compõe um provável complexo de espécies, devido a natureza altamente distinta de seus cariomorfos e a ausência de híbridos conhecidos. Devido a amostragem de apenas dois dos quatro cariomorfos existentes e a localidade dos mesmos, torna-se mais desafiador inferir as relações destes cariomorfos e avaliar a influência do sistema neo-sexual na diferenciação genética. Entretanto, a análise citogenética de uma nova localidade do cariomorfo A realizada pela primeira vez no presente trabalho ajuda a compreender melhor os aspectos citogenéticos das diferentes populações de *E. erythrinus*. Embora os índices de diversidade genética estimados sejam baixos, as análises evidenciaram alta diferenciação genética entre os cariomorfos A e D. Devido as diferenças cariotípica e à longa distância que separa estas populações as principais hipóteses apresentadas foram que: (I) a alopatria pode ter sido o principal agente de diferenciação genética entre esses cariomorfos; (II) o surgimento de cromossomos neo-sexuais pode ter sido o principal fator que promoveu essa diferenciação genética; (III) tanto a alopatria de longa distância quanto o surgimento dos cromossomos neo-sexuais no cariomorfo D podem ter

contribuído para a fixação e acúmulo de divergências. A hipótese III é considerada a mais provável.

Desta forma, o presente trabalho traz avanços na compreensão das relações entre diferentes populações de *H. malabaricus* e *Erythrinus erythrinus*, explorando dados citogenéticos e genéticos e também fornece dados acerca de populações naturais com sistemas de cromossomos neo-sexuais que podem contribuir para a compreensão do efeito destes sistemas no surgimento e na manutenção da diversidade genética. Estudos futuros implementando sequenciamento especificamente em cromossomos sexuais, permitirão a análise e comparação de características genéticas destes cromossomos com cromossomos autossomos fornecendo grandes contribuições para a problemática aqui apresentada. A coleta e sequenciamento de cariomorfos não abordados neste trabalho também permitirá que as relações entre cariomorfos de ambas as espécies seja melhor compreendida.

## REFERÊNCIAS

- ALBERT, J. **Historical Biogeography of Neotropical Freshwater Fishes**. University of California Press, 2011. DOI: 10.1525/california/9780520268685.001.0001.
- ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of molecular biology**, v. 215, n. 3, p. 403–410, 1990. DOI: 10.1016/S0022-2836(05)80360-2.
- ANTONELLI, A. et al. Amazonia is the primary source of Neotropical biodiversity. **Proceedings of the National Academy of Sciences of the United States of America**, v. 115, n. 23, p. 6034–6039, 2018. DOI: 10.1073/PNAS.1713819115/-/DCSUPPLEMENTAL.
- AXELROD, D. I. Mesozoic Paleogeography and Early Angiosperm History. **Botanical Review**, v. 36, n. 3, p. 277–319, 1970.
- BARBY, F. F. et al. Emerging patterns of genome organization in Notopteridae species (Teleostei, Osteoglossiformes) as revealed by Zoo-FISH and Comparative Genomic Hybridization (CGH). **Scientific Reports**, v. 9, n. 1, p. 1112, 2019. DOI: 10.1038/s41598-019-38617-4.
- BATESON, W. Heredity and variation in modern lights. In: Seward, A.C. ED., **Darwin and Modern Science**, Cambridge University Press, Cambridge, p. 85-101, 1909. DOI: 10.1017/CBO9780511693953.006.
- BAUMUNG, R.; SIMIANER, H.; HOFFMANN, I. Genetic diversity studies in farm animals – a survey. **Journal of Animal Breeding and Genetics**, v. 121, n. 6, p. 361–373, 2004. DOI: 10.1111/J.1439-0388.2004.00479.X.
- BERTOLLO, L. A. C. Chromosome Evolution in the Neotropical Erythrinidae Fish Family: An Overview. In: PISANO, E. et al. (org.). **Fish Cytogenetics**. 1. ed. Boca Raton, Florida: CRC Press, 2007. p. 195–212. DOI: 10.1201/B10746-7.
- BERTOLLO, L. A. C. et al. A biodiversity approach in the neotropical erythrinidae fish, *Hoplias malabaricus*. Karyotypic survey, geographic distribution of cytotypes and cytotaxonomic considerations. **Chromosome research**, v. 8, n. 7, p. 603–13, 2000. DOI: 10.1023/a:1009233907558.
- BERTOLLO, L. A. C. et al. Chromosome evolution in the erythrinid fish, *Erythrinus erythrinus* (Teleostei: Characiformes). **Heredity**, v. 93, n. 2, p. 228–233, 2004. DOI: 10.1038/sj.hdy.6800511.
- BERTOLLO, L. A. C. et al. Contributions to the cytogenetics of the Neotropical fish fauna. **Comparative cytogenetics**, v. 11, n. 4, p. 665–690, 2017. DOI: 10.3897/COMP CYTOGEN.V11I4.14713.

- BERTOLLO, L. A. C. et al. The X1X2Y sex chromosome system in the fish *Hoplias malabaricus*. I. G-, C- and chromosome replication banding. **Chromosome research**, v. 5, n. 7, p. 493–499, 1997. DOI: 10.1023/A:1018477232354.
- BERTOLLO, L. A. C.; MESTRINER, C. A. The X1X2Y sex chromosome system in the fish *Hoplias malabaricus* (Pisces, Erythrinidae). II. Meiotic analyses. **Chromosome Research**, v. 6, n. 2, p. 141–147, 1998. DOI: 10.1023/a:1009243114124.
- BERTOLLO, L. A. C.; MOREIRA-FILHO, O.; GALETTI, P. M. Cytogenetics and taxonomy: considerations based on chromosome studies of freshwater fish. **Journal of Fish Biology**, v. 28, n. 2, p. 153–159, 1986. DOI: 10.1111/j.1095-8649.1986.tb05153.x.
- BIFI, A. G. **Revisão taxonômica das espécies do grupo *Hoplias malabaricus* (Bloch, 1794) (Characiformes: Erythrinidae) da bacia do rio da Prata**. 2013. Tese (Doutorado em Ciências Ambientais) - Universidade Estadual de Maringá, Maringá, 2013.
- BORN, G. G.; BERTOLLO, L. A. C. An XX/XY Sex Chromosome System in a Fish Species, *Hoplias malabaricus*, with a Polymorphic NOR-Bearing X Chromosome. **Chromosome Research**, v. 8, n. 2, p. 111–118, 2000. DOI: 10.1023/A:1009238402051.
- BOUJARD, T. et al. **Poissons de Guyane**: Guide écologique de l'Approuague et de la réserve des Nouragues. Institut National de la Recherche Agronomique, Editions Inra, Paris, 219 p. 1997.
- BRADBURD, G. How to run a conStruct analysis. 2019. Disponível em < <https://cran.r-project.org/web/packages/conStruct/vignettes/run-conStruct.html>>. Acesso em Jan. 2022.
- BRADBURD, G. S.; COOP, G. M.; RALPH, P. L. Inferring continuous and discrete population genetic structure across space. **Genetics**, v. 210, n. 1, p. 33–52, 2018. DOI: 10.1534/genetics.118.301333.
- BRAUER, C. J. et al. On the roles of landscape heterogeneity and environmental variation in determining population genomic structure in a dendritic system. **Molecular Ecology**, v. 27, n. 17, p. 3484–3497, 2018. DOI: 10.1111/MEC.14808.
- CARBONE, L. et al. Evolutionary breakpoints in the gibbon suggest association between cytosine methylation and karyotype evolution. **PLoS genetics**, v. 5, n. 6, 2009. DOI: 10.1371/JOURNAL.PGEN.1000538.
- CHARLESWORTH, B.; CAMPOS, J. L.; JACKSON, B. C. Faster-X evolution: Theory and evidence from *Drosophila*. **Molecular Ecology**, v. 27, n. 19, p. 3753–3771, 2018. DOI: 10.1111/MEC.14534.
- CHRISTOFFELS, A. et al. Fugu Genome Analysis Provides Evidence for a Whole-Genome Duplication Early During the Evolution of Ray-Finned Fishes. **Molecular Biology and Evolution**, v. 21, n. 6, p. 1146–1151, 2004. DOI: 10.1093/MOLBEV/MSH114.

CIOFFI, M. B. et al. Cross-species chromosome painting tracks the independent origin of multiple sex chromosomes in two cofamiliar Erythrinidae fishes. **BMC Evolutionary Biology**, v. 11, n. 1, p. 1–7, 2011. DOI: 10.1186/1471-2148-11-186/.

CIOFFI, M. B. et al. Deciphering the Evolutionary History of Arowana Fishes (Teleostei, Osteoglossiformes, Osteoglossidae): Insight from Comparative Cytogenomics. **International Journal of Molecular Sciences**, v. 20, n. 17, p. 4296, 2019. DOI: 10.3390/IJMS20174296.

CIOFFI, M. B. et al. The key role of repeated DNAs in sex chromosome evolution in two fish species with ZW sex chromosome system. **Molecular Cytogenetics**, v. 5, n. 1, p. 1–7, 2012. DOI: 10.1186/1755-8166-5-28/FIGURES/4.

CIOFFI, M. B.; BERTOLLO, L. A. C. Initial steps in XY chromosome differentiation in *Hoplias malabaricus* and the origin of an X1X2Y sex chromosome system in this fish group. **Heredity**, v. 105, n. 6, p. 554–561, 2010. DOI: 10.1038/hdy.2010.18.

CIOFFI, M. B.; MARTINS, C.; BERTOLLO, L. A. C. Chromosome spreading of associated transposable elements and ribosomal DNA in the fish *Erythrinus erythrinus*. Implications for genome change and karyoevolution in fish. **BMC Evolutionary Biology**, v. 10, n. 1, p. 1–9, 2010. DOI: 10.1186/1471-2148-10-271/FIGURES/4.

CIOFFI, M. B.; MARTINS, C.; BERTOLLO, L. A. C. Comparative chromosome mapping of repetitive sequences. Implications for genomic evolution in the fish, *Hoplias malabaricus*. **BMC Genetics**, v. 10, n. 1, p. 1–11, 2009. DOI: 10.1186/1471-2156-10-34/FIGURES/6.

COSTA NOVAES, J. L.; CARVALHO, D. Population structure and stock assessment of *Hoplias malabaricus* (Characiformes: Erythrinidae) caught by artisanal fishermen in river-reservoir transition area in Brazil. **Revista de Biología Tropical**, v. 59, n. 1, p. 71–83, 2011.

COX, C. B. The biogeographic regions reconsidered. **Journal of Biogeography**, v. 28, n. 4, p. 511–523, 2001. DOI: 10.1046/J.1365-2699.2001.00566.X.

COYNE, J. A. Genetics and speciation. **Nature**, v. 355, n. 6360, p. 511–515, 1992. DOI: 10.1038/355511a0.

COYNE, J. A., & ORR, H.A. Two rules of speciation In: Otte, D.; Endler, J. (Org.) **Speciation and Its Consequences**, Sinauer Associates, Sunderland, MA, pp. 180–207. 1989.

COYNE, J. A.; ORR, H. A. **Speciation**. Sinauer Associates, Sunderland, MA, 2004.

DA ROSA, R. et al. Evolutionary trends in *Hoplerethrinus unitaeniatus* (Agassiz 1829) (Characiformes, Erythrinidae). **Reviews in Fish Biology and Fisheries**, v. 22, n. 2, p. 467–475, 2012. DOI: 10.1007/s11160-011-9237-3.

DA SILVA, F. A. Effects of environmental pollution on the DNAomics of Amazonian fish. **Environmental Pollution**, v. 252, p. 180–187, 2019. DOI: 10.1016/J.ENVPOL.2019.05.112.

DAVEY, J. L.; BLAXTER, M. W. RADSeq: next-generation population genetics. **Briefings in Functional Genomics**, v. 9, n. 5–6, p. 416, 2010. DOI: 10.1093/BFGP/ELQ031.

DAWKINS, R. **The selfish gene**. Oxford University Press: Oxford, Inglaterra, 1989.

ddRADSeq. Disponível em: <https://www.illumina.com/science/sequencing-method-explorer/kits-and-arrays/ddradseq.html>. Acesso em: 3 mar. 2022.

DE MORAES, R. L. et al. Comparative cytogenetics and neo-Y formation in small-sized fish species of the genus *Pyrrhulina* (Characiformes, Lebiasinidae). **Frontiers in Genetics**, v. 10, n. 6, p. 678, 2019. DOI: 10.3389/FGENE.2019.00678/BIBTEX.

DE MORAES, R. L. R. et al. Tracking the Evolutionary Trends Among Small-Size Fishes of the Genus *Pyrrhulina* (Characiforme, Lebiasinidae): New Insights From a Molecular Cytogenetic Perspective. **Frontiers in Genetics**, v. 12, p. 1953, 2021. DOI: 10.3389/FGENE.2021.769984/BIBTEX.

DE OLIVEIRA, E. A. et al. Comparative cytogenetics in the genus *Hoplias* (Characiformes, Erythrinidae) highlights contrasting karyotype evolution among congeneric species. **Molecular Cytogenetics**, v. 8, n. 1, p. 1–10, 2015. DOI: 10.1186/S13039-015-0161-4.

DE OLIVEIRA, E. A. et al. Tracking the evolutionary pathway of sex chromosomes among fishes: characterizing the unique XX/XY1Y2 system in *Hoplias malabaricus* (Teleostei, Characiformes). **Chromosoma**, v. 127, n. 1, p. 115–128, 2018. DOI: 10.1007/s00412-017-0648-3.

DE QUEIROZ, K. Species Concepts and Species Delimitation. **Systematic Biology**, v. 56, n. 6, p. 879–886, 2007. DOI: 10.1080/10635150701701083.

DEON, G. A. et al. Highly Rearranged Karyotypes and Multiple Sex Chromosome Systems in Armored Catfishes from the Genus *Harttia* (Teleostei, Siluriformes). **Genes**, v. 11, n. 11, p. 1366, 2020. DOI: 10.3390/GENES11111366.

DERGAM, J. A.; BERTOLLO, L. A. C. Karyotypic Diversification In *Hoplias Malabaricus* (Osteichthyes, Erythrinidae) of The São Francisco And Alto Paraná Basins, Brazil). **Rev. Brasil. Genet**, p. 755–766, 1990.

DOBZHANSKY, T. Evolution in the Tropics, by Theodosius Dobzhansky. **American Scientist**, v. 38, p. 209–221, 1950.

DOBZHANSKY, T. On the Sterility of the Interracial Hybrids in *Drosophila Pseudoobscura*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 19, n. 4, p. 397–403, 1933. DOI: 10.1073/PNAS.19.4.397.

DOMISCH, S.; AMATULLI, G.; JETZ, W. Near-global freshwater-specific environmental variables for biodiversity analyses in 1 km resolution. **Scientific Data**, v. 2, n. 1, p. 1–13, 2015. DOI: 10.1038/sdata.2015.73.

EATON, D. A. R. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. **Bioinformatics**, v. 30, n. 13, p. 1844–1849, 2014. DOI: 10.1093/BIOINFORMATICS/BTU121.

ELLEGREN, H. Genome sequencing and population genomics in non-model organisms. **Trends in ecology & evolution**, v. 29, n. 1, p. 51–63, 2014. DOI: 10.1016/J.TREE.2013.09.008.

FARIA, R.; NAVARRO, A. Chromosomal speciation revisited: rearranging theory with pieces of evidence. **Trends in Ecology & Evolution**, v. 25, n. 11, p. 660–669, 2010. DOI: 10.1016/j.tree.2010.07.008.

FARRÉ, M. et al. Assessing the role of tandem repeats in shaping the genomic architecture of great apes. **PLoS ONE**, v. 6, n. 11, 2011. DOI: 10.1371/JOURNAL.PONE.0027239.

FERREIRA, A. M. V.; RIBEIRO, L. B.; FELDBERG, E. Molecular analysis reveals high diversity in the *Hoplias malabaricus* (Characiformes, Erythrinidae) species complex from different Amazonian localities. **Biodiversity and Conservation, Acta Amazonica**, v. 51, n. 2, p. 139–144, 2021. DOI: 10.1590/1809-4392202001652.

FERREIRA, M. et al. The *Bunocephalus coracoideus* species complex (Siluriformes, Aspredinidae). Signs of a speciation process through chromosomal, genetic and ecological diversity. **Frontiers in Genetics**, v. 8, p. 120, 2017. DOI: 10.3389/FGENE.2017.00120/BIBTEX.

FOLL, M.; GAGGIOTTI, O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. **Genetics**, v. 180, n. 2, p. 977–993, 2008. DOI: 10.1534/genetics.108.092221.

FREITAS, N. L. et al. Early Stages of XY Sex Chromosomes Differentiation in the Fish *Hoplias malabaricus* (Characiformes, Erythrinidae) Revealed by DNA Repeats Accumulation. **Current Genomics**, v. 19, n. 3, p. 216–226, 2018. DOI: 10.2174/1389202918666170711160528.

FRICKE, R.; ESCHMEYER, W.; VAN DER LAAN, R. **Eschmeyer's Catalog of Fishes**. v. 3882. 2021. DOI: 10.11646/zootaxa.3882.1.1.

GARZÓN-ORDUÑA, I. J.; BENETTI-LONGHINI, J. E.; BROWER, A. V. Z. Timing the diversification of the Amazonian biota: butterfly divergences are consistent with Pleistocene refugia. **Journal of Biogeography**, v. 41, n. 9, p. 1631–1638, 2014. DOI: 10.1111/jbi.12330.

GBIF – Global Biodiversity Information Facility. **Erythrinus erythrinus (Bloch & Schneider, 1801)**. Disponível em <<https://www.gbif.org/species/2352207>>. Acesso em Jan. 2022.

GBIF – Global Biodiversity Information Facility. **Hoplias malabaricus (Bloch, 1794)**. Disponível em <<https://www.gbif.org/species/2352230>>. Acesso em Jan. 2022.

GIULIANO-CAETANO, L. Estudo citogenético em *Hoplerethrinus unitaeniatus* (PISCES, ERYTHRINIDAE) de diferentes bacias hidrográficas brasileiras. 1986. Dissertação de Mestrado, Universidade Federal de São Carlos, São Carlos, 1986.

GLASAUER, S. M. K.; NEUHAUSS, S. C. F. Whole-genome duplication in teleost fishes and its evolutionary consequences. **Molecular genetics and genomics**, v. 289, n. 6, p. 1045–1060, 2014. DOI: 10.1007/S00438-014-0889-2.

GORNUNG, E.; DARWIN, C. Twenty Years of Physical Mapping of Major Ribosomal RNA Genes across the Teleosts: A Review of Research. **Cytogenetic and Genome Research**, v. 141, n. 2–3, p. 90–102, 2013. DOI: 10.1159/000354832.

GRAÇA, W. J.; PAVANELLI, C. S. **Peixes da Planície de Inundação do Alto Rio Paraná e Áreas Adjacentes**. Maringá: EDUEM, 2007.

HALDANE, J. B. S. Sex ratio and unisexual sterility in hybrid animals. **Journal of Genetics**, v. 12, n. 2, p. 101-109, 1922.

HENGL, T. et al. Soil Grids 1km — Global Soil Information Based on Automated Mapping. **PLOS ONE**, v. 9, n. 8, 2014. DOI: 10.1371/JOURNAL.PONE.0105992.

HILLEBRAND, Helmut. On the Generality of the Latitudinal Diversity Gradient. **The American Naturalist**, v. 163, n. 2, p. 192–211, 2004. DOI: 10.1086/381004.

HOORN, C. et al. Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. **Science**, v. 330, n. 6006, p. 927–931, 2010. DOI: 10.1126/science.1194585.

HUBERT, N.; RENNO, J. F. Historical biogeography of South American freshwater fishes. **Journal of Biogeography**, v. 33, n. 8, p. 1414–1436, 2006. DOI: 10.1111/j.1365-2699.2006.01518.x.

Illumina. Disponível em: <https://www.illumina.com/science/sequencing-method-explorer/kits-and-arrays/ddradseq.html> . Acesso em: 3 mar. 2022.

INOUE, J. et al. Rapid genome reshaping by multiple-gene loss after whole-genome duplication in teleost fish suggested by mathematical modeling. **Proceedings of the National Academy of Sciences of the United States of America**, v. 112, n. 48, p. 14918–14923, 2015. DOI: 10.1073/PNAS.1507669112/-/DCSUPPLEMENTAL.

JABLONSKI, D.; ROY, K.; VALENTINE, J. W. Out of the Tropics: Evolutionary Dynamics of the Latitudinal Diversity Gradient. **Science**, v. 314, n. 5796, p. 102–106, 2006. DOI: 10.1126/science.1130880.

JATLLON, O. et al. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. **Nature**, v. 431, n. 7011, p. 946–957, 2004. DOI: 10.1038/nature03025.

JORDAN, B. et al. Genome complexity reduction for SNP genotyping analysis. **Proceedings of the National Academy of Sciences of the United States of America**, v. 99, n. 5, p. 2942–2947, 2002. DOI: 10.1073/PNAS.261710699/.

KILIAN, A. et al. Diversity Arrays Technology: A Generic Genome Profiling Technology on Open Platforms. *In: Methods in molecular biology (Clifton, N.J.)*, v. 888p. 67–89, 2018. DOI: 10.1007/978-1-61779-870-2\_5.

KITANO, J. et al. A role for a neo-sex chromosome in stickleback speciation. **Nature**, v. 461, n. 7267, p. 1079–1083, 2009. DOI: 10.1038/NATURE08441.

KITANO, J.; PEICHEL, C. L. Turnover of sex chromosomes and speciation in fishes. **Environmental Biology of Fishes**, v. 94, n. 3, p. 549–558, 2012. DOI: 10.1007/s10641-011-9853-8.

KOPELMAN, N. M. et al. Clumpak: a program for identifying clustering modes and packaging population structure inferences across *K*. **Molecular Ecology Resources**, v. 15, n. 5, p. 1179–1191, 2015. DOI: 10.1111/1755-0998.12387.

LAHN, B. T.; PAGE, D. C. Four Evolutionary Strata on the Human X Chromosome. **Science**, v. 286, n. 5441, p. 964–967, 1999. DOI: 10.1126/SCIENCE.286.5441.964.

LEHNER, B.; VERDIN, K.; JARVIS, A. New Global Hydrography Derived From Spaceborne Elevation Data. **Eos, Transactions American Geophysical Union**, v. 89, n. 10, p. 93–94, 2008. DOI: 10.1029/2008EO100001.

LEITE, P. P. M. et al. Tracking the evolutionary pathways among Brazilian *Lebiasina* species (Teleostei: Lebiasinidae): a chromosomal and genomic comparative investigation. **Neotropical Ichthyology**, 2021. No prelo.

LEVAN, A.; FREDGA, K.; SANDBERG, A. A. Nomenclature for Centromeric Position on Chromosomes. **Hereditas**, v. 52, n. 2, p. 201–220, 1964. DOI: 10.1111/J.1601-5223.1964.TB01953.X.

LIMA, T. G. Higher levels of sex chromosome heteromorphism are associated with markedly stronger reproductive isolation. **Nature Communications**, v. 5, n. 1, p. 1–8, 2014. DOI: 10.1038/ncomms5743.

LOPES, P. A. et al. Cytotaxonomy of *Hoplias malabaricus* (Osteichthyes, Erythrinidae) in the Aguapey River (Province of Corrientes, Argentina). **Source: Copeia**, n. 2, p. 485–487, 1998.

LYNCH, M.; CONERY, J. S. The evolutionary fate and consequences of duplicate genes. **Science**, v. 290, n. 5494, p. 1151–1155, 2000. DOI: 10.1126/SCIENCE.290.5494.1151.

LYONS, T. J. *Hoplias malabaricus* (Trahira). 2020. Disponível em: <https://www.iucnredlist.org/species/186393/1812387>. Acesso em: Jan. 2022.

MARTINS, N. F. et al. Differentiation and evolutionary relationships in *Erythrinus erythrinus* (Characiformes, Erythrinidae): Occurrence and distribution of B chromosomes. **Genetics and Molecular Research**, v. 13, n. 3, p. 7094–7101, 2014. DOI: 10.4238/2014.FEBRUARY.14.11.

MARTINS, N. F. et al. Differentiation and evolutionary relationships in *Erythrinus erythrinus* (Characiformes, Erythrinidae): comparative chromosome mapping of repetitive sequences. **Reviews in Fish Biology and Fisheries**, v. 23, n. 2, p. 261–269, 2013. DOI: 10.1007/s11160-012-9292-4.

MATOS-MARAVÍ, P. et al. Mesoamerica is a cradle and the Atlantic Forest is a museum of Neotropical butterfly diversity: insights from the evolution and biogeography of Brassolini (Lepidoptera: Nymphalidae). **Biological Journal of the Linnean Society**, v. 133, n. 3, p. 704–724, 2021. DOI: 10.1093/biolinnean/blab034.

MATTOX, G. M. T.; TOLEDO-PIZA, M.; OYAKAWA, O. T. **Taxonomic Study of *Hoplias Aimara* (Valenciennes, 1846) and *Hoplias macrophthalmus* (Pellegrin, 1907) (Ostariophysi, Characiformes, Erythrinidae)**. v. 2006, n. 3, p. 516–528, 2006. DOI: 10.1643/0045-8511(2006).

MAYR, E. The evolution of living systems. **Proceedings of the National Academy of Sciences**, v. 51, n. 5, p. 934–941, 1964. DOI: 10.1073/pnas.51.5.934

MCGRATH, C. L. et al. Differential retention and divergent resolution of duplicate genes following whole-genome duplication. **Genome Research**, v. 24, n. 10, p. 1665, 2014. DOI: 10.1101/GR.173740.114.

MCKENNA, D. D.; FARRELL, B. D. Tropical forests are both evolutionary cradles and museums of leaf beetle diversity. **Proceedings of the National Academy of Sciences**, v. 103, n. 29, p. 10947–10951, 2006. DOI: 10.1073/PNAS.0602712103.

MEIRMANS, P. G. Seven common mistakes in population genetics and how to avoid them. **Molecular Ecology**, v. 24, n. 13, p. 3223–3231, 2015. DOI: 10.1111/mec.13243.

MESEGUER, A. S.; CONDAMINE, F. L. Ancient tropical extinctions at high latitudes contributed to the latitudinal diversity gradient\*. **Evolution**, v. 74, n. 9, p. 1966–1987, 2020. DOI: 10.1111/EVO.13967.

MESEGUER, A. S.; CONDAMINE, F. L. Ancient tropical extinctions contributed to the latitudinal diversity gradient. **bioRxiv**, p. 236646, 2019. DOI: 10.1101/236646.

MOREAU, C. S.; BELL, C. D. Testing The Museum Versus Cradle Tropical Biological Diversity Hypothesis: Phylogeny, Diversification, And Ancestral Biogeographic Range Evolution Of The Ants. **Evolution**, v. 67, n. 8, p. 2240–2257, 2013. DOI: 10.1111/EVO.12105.

MORIN, P. A.; LUIKART, G.; WAYNE, R. K. SNPs in ecology, evolution and conservation. **Trends in Ecology & Evolution**, v. 19, n. 4, p. 208–216, 2004. DOI: 10.1016/J.TREE.2004.01.009.

MULLER, H. Isolating Mechanisms, Evolution And Temperature. **Biological Symposia**, v. 6, p. 71–125, 1942.

NELSON, J. S.; GRANDE, T. C.; WILSON, M. V. H. **Fishes of the World: Fifth Edition**. 5. ed. Hoboken, New Jersey: Wiley Blackwell, 2016. DOI: 10.1002/9781119174844.

OYAKAWA, O. T. Erythrinidae. In: ROBERTO E. REIS; SVEN O. KULLANDER; CARL J. FERRARIS, Jr. (Org.). **Check List of the Freshwater Fishes of South and Central America**. Porto Alegre: EDIPUCRS, p. 238–241, 2003

OYAKAWA, O. T.; MATTOX, G. M. T. Revision of the Neotropical trahiras of the *Hoplias lacerdae* species-group (Ostariophysi: Characiformes: Erythrinidae) with descriptions of two new species. **Neotropical Ichthyology**, v. 7, n. 2, p. 117–140, 2009. DOI: 10.1590/S1679-62252009000200001.

PAIVA, M. P. **Grandes Represas do Brasil**. Editerra: Brasília, Brazil, 1982.

PAYSEUR, B. A.; PRESGRAVES, D. C.; FILATOV, D. A. Introduction: Sex chromosomes and speciation. **Molecular Ecology**, v. 27, n. 19, p. 3745–3748, 2018. DOI: 10.1111/MEC.14828.

PAZZA, R.; KAVALCO, K. F.; BERTOLLO, L. A. C. Chromosome polymorphism in *Astyanax fasciatus* (Teleostei, Characidae). 1. Karyotype analysis, Ag-NORs and mapping of the 18S and 5S ribosomal genes in sympatric karyotypes and their possible hybrid forms. **Cytogenetic and genome research**, v. 112, n. 3–4, p. 313–319, 2006. DOI: 10.1159/000089886.

PEAKALL, R.; SMOUSE, P. E. GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. **Bioinformatics**, v. 28, n. 19, p. 2537–2539, 2012. DOI: 10.1093/BIOINFORMATICS/BTS460.

PEAKALL, R.; SMOUSE, P. E. genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. **Molecular Ecology Notes**, v. 6, n. 1, p. 288–295, 2006. DOI: 10.1111/J.1471-8286.2005.01155.X.

PENDAS, A. M. et al. Chromosomal mapping and nucleotide sequence of two tandem repeats of Atlantic salmon 5S DNAr. **Cytogenetic and Genome Research**, v. 67, n. 1, p. 31–36, 1994. DOI: 10.1159/000133792.

PEREZ, M. F. et al. Assessing population structure in the face of isolation by distance: Are we neglecting the problem? **Diversity and Distributions**, v. 24, n. 12, p. 1883–1889, 2018. DOI: 10.1111/ddi.12816.

PERRY, G. H. et al. Diet and the evolution of human amylase gene copy number variation. **Nature Genetics**, v. 39, n. 10, p. 1256–1260, 2007. DOI: 10.1038/ng2123.

PETERSON, B. K. et al. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. **PLoS ONE**, v. 7, n. 5, 2012. DOI: 10.1371/JOURNAL.PONE.0037135.

- PETRY, A. C. The role of the predatory trahira (Pisces: Erythrinidae) in structuring fish assemblages in lakes of a Neotropical floodplain. **Hidrobiologia**, p. 115–126, 2010. DOI: 10.1007/s10750-010-0281-0.
- PRESGRAVES, D. C. Evaluating genomic signatures of “the large X-effect” during complex speciation. **Molecular Ecology**, v. 27, n. 19, p. 3822–3830, 2018. DOI: 10.1111/MEC.14777.
- PRESGRAVES, D. C. Sex chromosomes and speciation in Drosophila. **Trends in genetics**, v. 24, n. 7, p. 336–343, 2008. DOI: 10.1016/J.TIG.2008.04.007.
- PRESGRAVES, D. C.; ORR, H. A. Haldane’s rule in taxa lacking a hemizygous X. **Science**, v. 282, n. 5390, p. 952–954, 1998. DOI: 10.1126/SCIENCE.282.5390.952.
- PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945–959, 2000. DOI: 10.1093/GENETICS/155.2.945.
- RAJ, A.; STEPHENS, M.; PRITCHARD, J. K. fastSTRUCTURE: variational inference of population structure in large SNP data sets. **Genetics**, v. 197, n. 2, p. 573–89, 2014. DOI: 10.1534/genetics.114.164350.
- RIBEIRO, A. C. Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: an example of faunal evolution associated with a divergent continental margin. **Neotropical Ichthyology**, v. 4, n. 2, p. 225–246, 2006. DOI: 10.1590/S1679-62252006000200009.
- RIBEIRO, A. C. Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: an example of faunal evolution associated with a divergent continental margin. **Neotropical Ichthyology**, v. 4, n. 2, p. 225–246, 2006. DOI: 10.1590/S1679-62252006000200009.
- ROZAS, J. et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. **Molecular biology and evolution**, v. 34, n. 12, p. 3299–3302, 2017. DOI: 10.1093/MOLBEV/MSX248.
- RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.
- RUIZ-HERRERA, A.; CASTRESANA, J.; ROBINSON, T. J. Is mammalian chromosomal evolution driven by regions of genome fragility? **Genome Biology**, v. 7, n. 12, p. 1–16, 2006. DOI: 10.1186/GB-2006-7-12-R115/FIGURES/2.
- RULL, V. Neotropical diversification: historical overview and conceptual insights. **PeerJ Preprints**, v. 6, 2018. DOI: 10.7287/peerj.preprints.27294v1.
- RULL, V.; CARNAVAL, A. C. **Correction to: Neotropical Diversification: Patterns and Processes**. DOI: 10.1007/978-3-030-31167-4\_31.

SAMBROOK, J.; RUSSELL, D. W. **Molecular Cloning: A laboratory manual**. Vol. 1, 2, 3; Cold Springs Harb. Lab. Press: Cold Spring Harbor, NY, USA, 2001.

SASSI, F. M. C. et al. High genetic diversity despite conserved karyotype organization in the giant trahiras from genus *hoplias* (Characiformes, erythrinidae). **Genes**, v. 12, n. 2, p. 1–13, 2021. DOI: 10.3390/genes12020252.

SASSI, F. M. C. et al. Multiple Sex Chromosomes and Evolutionary Relationships in Amazonian Catfishes: The Outstanding Model of the Genus *Harttia* (Siluriformes: Loricariidae). **Genes**, v. 11, n. 10, p. 1179, 2020. DOI: 10.3390/GENES11101179.

SCAVONE, M. D. P.; BERTOLLO, L. A. C., CAVALLINI, M. M. Sympatric occurrence of two karyotypic forms of *Hoplias malabaricus* (Pisces, Erythrinidae). **Cytobios**, v. 80, p. 223–227, 1995.

SCHLÖTTERER, C. The evolution of molecular markers — just a matter of fashion? **Nature Reviews Genetics**, v. 5, n. 1, p. 63–69, 2004. DOI: 10.1038/nrg1249.

SCLATER, P. L. On the general geographic distribution of the members of the class Aves. **Proceedings of the Linnean Society of London, Zoology**, v. 2, p. 130–145. 1858.

SEMBER, A. et al. Multiple sex chromosomes in teleost fishes from a cytogenetic perspective: state of the art and future challenges. **Philosophical Transactions of the Royal Society B**, v. 376, n. 1833, 2021. DOI: 10.1098/RSTB.2020.0098.

SEMBER, A. et al. Sex Chromosome Evolution and Genomic Divergence in the Fish *Hoplias malabaricus* (Characiformes, Erythrinidae). **Frontiers in Genetics**, v. 9, p. 71, 2018. DOI: 10.3389/FGENE.2018.00071.

SHERMAN, K. D. et al. RAD-Seq Analysis and in situ Monitoring of Nassau Grouper Reveal Fine-Scale Population Structure and Origins of Aggregating Fish. **Frontiers in Marine Science**, v. 7, p. 157, 2020. DOI: 10.3389/FMARS.2020.00157/BIBTEX.

ŠÍCHOVÁ, J. et al. Dynamic karyotype evolution and unique sex determination systems in Leptidea wood white butterflies. **BMC Evolutionary Biology**, v. 15, n. 1, p. 89, 2015. DOI: 10.1186/s12862-015-0375-4.

SOCHOROVÁ, J. et al. Analyses of the Updated “Animal DNAr Loci Database” with an Emphasis on Its New Features. **International Journal of Molecular Sciences**, v. 22, n. 21, p. 11403, 2021. DOI: 10.3390/IJMS222111403.

SOCHOROVÁ, J. et al. Evolutionary trends in animal ribosomal DNA loci: introduction to a new online database. **Chromosoma**, v. 127, n. 1, p. 141–150, 2018. DOI: 10.1007/S00412-017-0651-8/FIGURES/6.

SOUZA, F. H. S. et al. Elucidating the complex relationship of *Hoplias malabaricus*’ (Teleostei; Characiformes) population. A genetic, chromosomal and spatial approach. **Evolution Letters**. 2021. No prelo.

SOUZA, F. H. S. et al. Interspecific Genetic Differences and Historical Demography in South American Arowanas (Osteoglossiformes, Osteoglossidae, Osteoglossum). **Genes**, v. 10, n. 9, p. 693, 2019. DOI: 10.3390/genes10090693.

STEBBINS, G. L. 1974. **Flowering plants**: Evolution above the species level. The Belknap Press of Harvard University Press, Cambridge, MA, 399 p. 1974.

TOMA, G. A. et al. Cytogenetics of the small-sized fish, *Copeina guttata* (Characiformes, Lebiasinidae): Novel insights into the karyotype differentiation of the family. **PLoS ONE**, v. 14, n. 12, p. e0226746, 2019. DOI: 10.1371/JOURNAL.PONE.0226746.

TOUPS, M. et al. A reciprocal translocation radically reshapes sex-linked inheritance in the common frog. **Molecular ecology**, v. 28, n. 8, p. 1877–1889, 2019. DOI: 10.1111/MEC.14990.

UTSUNOMIA, R. et al. Genetic differentiation among distinct karyomorphs of the wolf fish *Hoplias malabaricus* species complex (Characiformes, Erythrinidae) and report of unusual hybridization with natural triploidy. **Journal of Fish Biology**, 2014. DOI: 10.1111/jfb.12526.

VANDEPOELE, K. et al. Major events in the genome evolution of vertebrates: Paraneome age and size differ considerably between ray-finned fishes and land vertebrates. **Proceedings of the National Academy of Sciences**, v. 101, n. 6, p. 1638–1643, 2004. DOI: 10.1073/PNAS.0307968100.

VASCONCELOS, T.; O'MEARA, B. C.; BEAULIEU, J. M. Retiring “cradles” and “museums” of biodiversity. **EcoEvoRxiv**, 2021. DOI: 10.32942/OSF.IO/SXAH8.

VIEIRA, C. P.; COELHO, P. A.; VIEIRA, J. Inferences on the evolutionary history of the *Drosophila americana* polymorphic X/4 fusion from patterns of polymorphism at the X-linked paralytic and elav genes. **Genetics**, v. 164, n. 4, p. 1459–1469, 2003.

WALLACE, A. **The geographical distribution of animals**; with a study of the relations of living and extinct faunas as elucidating the past changes of the Earth's surface. Harper and Brothers, New York, Vol. I & II. 650 p. 1876.

WALSH, J. B. How often do duplicated genes evolve new functions? **Genetics**, v. 139, n. 1, p. 421–428, 1995. DOI: 10.1093/GENETICS/139.1.421.

WENZL, P. et al. A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. **BMC Genomics**, v. 7, n. 1, p. 206, 2006. DOI: 10.1186/1471-2164-7-206.

XING, C. et al. Comparison of microsatellites, single-nucleotide polymorphisms (SNPs) and composite markers derived from SNPs in linkage analysis. **BMC Genetics**, v. 6, n. SUPPL.1, p. 1–5, 2005. DOI: 10.1186/1471-2156-6-S1-S29.

YANO, C. F. et al. Against the mainstream: exceptional evolutionary stability of ZW sex chromosomes across the fish families Triportheidae and Gasteropelecidae (Teleostei: Characiformes). **Chromosome Research**, p. 1–26, 2021. DOI: 10.1007/S10577-021-09674-1.

YANO, C. F.; BERTOLLO, L. A. C.; CIOFFI, M. B. Fish-FISH: Molecular Cytogenetics in Fish Species. **Fluorescence In Situ Hybridization (FISH)**, p. 429–443, 2017. DOI: 10.1007/978-3-662-52959-1\_44.

YOSHIDO, A. et al. Step-by-step evolution of neo-sex chromosomes in geographical populations of wild silkmths, *Samia cynthia* ssp. **Heredity**, v. 106, n. 4, p. 614–624, 2011. DOI: 10.1038/hdy.2010.94.

ZHANG, S. et al. Meiotic Heterogeneity of Trivalent Structure and Interchromosomal Effect in Blastocysts With Robertsonian Translocations. **Frontiers in Genetics**, v. 12, p. 161, 2021. DOI: 10.3389/FGENE.2021.609563.

ZIMMERMAN, S. J.; ALDRIDGE, C. L.; OYLER-MCCANCE, S. J. An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. **BMC Genomics**, v. 21, n. 1, p. 1–16, 2020. DOI: 10.1186/S12864-020-06783-9/.

ZWICK, M. S. et al. A rapid procedure for the isolation of C0t-1 DNA from plants. **Genome**, v. 40, n. 1, p. 138–142, 1997. DOI: 10.1139/g97-020.

## APÊNDICE 1 – Resultados da análise ConStruct para *H. malabaricus*

De acordo com os *outputs* as cadeias MCMC não convergiram, portanto os resultados não apresentam grande probabilidade posterior, dessa forma não foram apresentados no artigo, porém estão dispostos aqui como material complementar. Os dois melhores resultados de acordo com o indicado pelo arquivo *layer.contributions* foram os valores de K=3 e K=5.

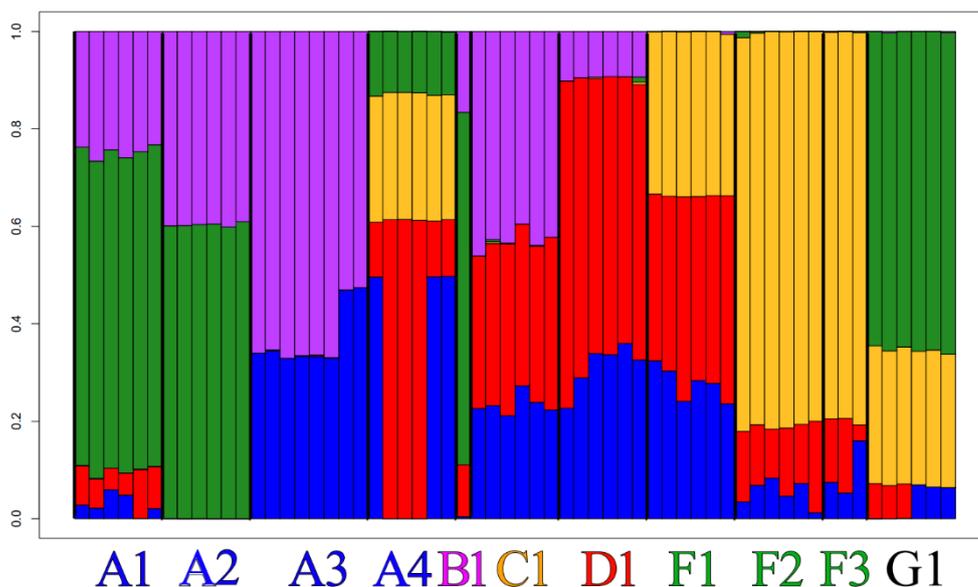


Figura 14 - Resultado da análise ConStruct para o valor de K=3. Fonte: Autor (2022).

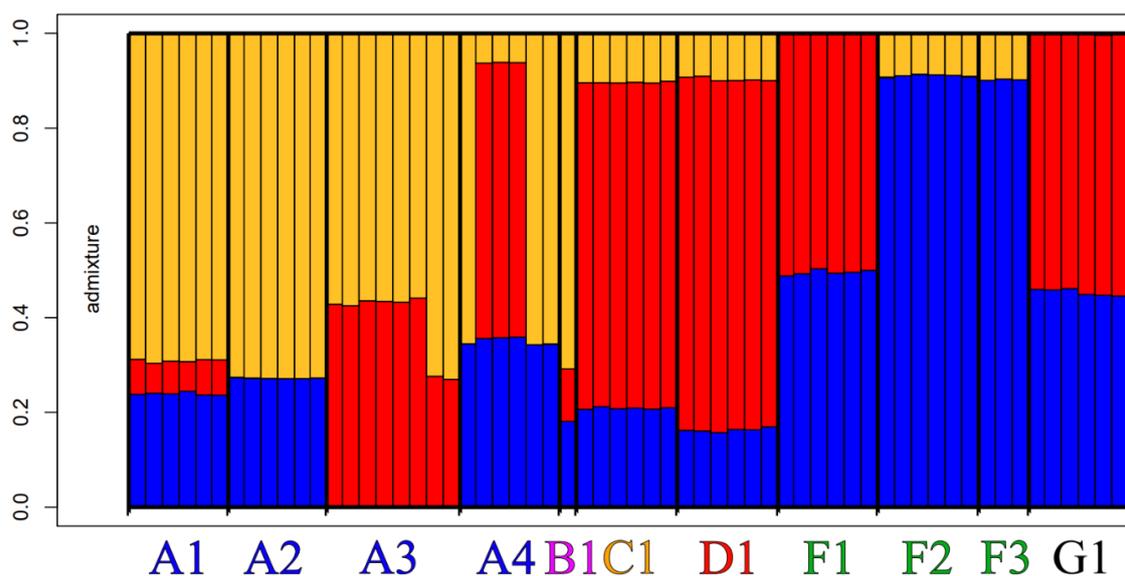


Figura 15 - Resultado da análise ConStruct para o valor de K=5. Fonte: Autor (2022).

**APÊNDICE 2 – Valores de  $D_{XY}$  e  $D_A$  par-a-par por cariomorfo para *H. malabaricus***

**Quadro 1 - Valores de  $D_{XY}$  par-a-par por cariomorfo**

<b><math>D_{XY}</math> par-a-par por cariomorfo</b>					
	HB	HC	HD	HF	HG
HA	0,006931574	0,00899648	0,010478493	0,009039902	0,009096628
HB		0,008851377	0,01019403	0,009875461	0,00994540
HC			0,008898986	0,009008059	0,009084258
HD				0,010814491	0,010841071
HF					0,005592311

Fonte: Autor (2022).

**Quadro 2 - Valores de  $D_A$  par-a-par por cariomorfo**

<b><math>D_A</math> par-a-par por cariomorfo</b>					
	HB	HC	HD	HF	HG
HA	0,003132886	0,005324302	0,006488482	0,004130634	0,005300797
HB		0,008225378	0,009231448	0,008001419	0,009186632
HC			0,00801214	0,007244743	0,008401305
HD				0,008713588	0,009837234
HF					0,003695139

Fonte: Autor (2022).

## APÊNDICE 3 – Lista de marcadores potencialmente sob seleção para *H. malabaricus*

Quadro 3 - Lista de marcadores potencialmente sob seleção

>20	GTGTGCAGGACGCTGTCTACAAGCCGAGGGGTTTACATGCA-TGAGA-TCGGAAGA-GCGGTTT-AGCAGGA
>112	TGGTGTGGTGGCTCCTGTGGATGTGGTGTAGGTGGAGCCAATTTTTGTGTTGGTACCAGGCTGA---
>127	CTTCGCCAAAACACAAGCCGCGGGATGGAGGAGCGTAAAGCACCGCATGA-GATC-GGAAGA-GCGGT---
>343	ACCTGTTTGTCTTCTAAAGCCCTATTCAGATTGGATACATTTTTCAAGAGGGTGTCTGTAATAAAAC-
>408	TTTGGAGTAAGATTATTGTCAAAAATAAACGTGTGCATGAG-AT-CGGAA-GAGC-GGTTACGC-AG-GAATGC-
>522	ATCATTGTTAGTCTCTTTTCTTTAGAGTCTCTTCACAAAATAAGAAATTGTGTGCATGTGTGTTGTG
>682	TCACACTGAGCTTGTGTCCACAAATTTTTGCATGAGA-TCGG-AAGAGCGG-TTCAGC-A-GGAATG-CCGAGA-
>829	CAGAGCGAGCACGTGCAGAGGAGAAAAGAGAAAGTGGGGAGATAGGTGTAATTAAGGCGTCTACATTC
>878	GTCTACCTGAATGACCGTGAGGCAGAGCTTTCTTCTGGACAGCATGA-GATC-GGAAGAGC-GGTT-CAGCA-
>979	ACACAGTGTCTGATCAGCGCTAAAATTTGCATCCCTAAGTAAATAAGTAAATAAATCAAGTGCATGAG
>1004	GACACAGTGAGGGAAGGAGGAGTGTCTGCTGGCCCACTGGGAGGACAGAGACATTGAGCTAATGACCCC
>1077	GATCGGAGACCTGTACACTGACCTGAGGGATGGACGCATGAGA-TCGG-AAGAG-CGGNTTCAGCAGGAATG
>1106	TAAGTGAGGAAAGCAGAGCTAACGTGACAACCTCCATAAGCTGTCTGTCTGTTACTACAGAGCAGC---
>1127	ACACAACTTTCCACATACGACACGCCATAAAGTACTCAAATAACAGTCATCAAGGCTTCGGCTTA
>1133	GAAGGAGTTTTGTGAGCTGTGCCTGGTGAAGAGCGGCAGACGGACAAGGTCTATGCATGAGATC---
>1141	ATAGAGGAGCATTGATTGACCAGAGCCAGTCCCTAATATACACATTGCTGAGCCTAACTATGGAAAG
>1175	ACACAAGCCACCTATACAGCCCCAATAAGTCCACAAACCAGAATCACACACATGCACACAACCACAT
>1255	TACATGCTCATTGTGTGAGCAGGCAAGGCTTCTATAAAACACTTTTTTAACACTCATTGGGTTGTT--
>1265	TACATGCTCATTGTGTGAGCAGGCAAGGCTTCTATAAAACACTTTTTTAACACTCATTGGGTTGTT--
>1270	CTCTGGGTTCAATGCAATTTAGGCTGATTTGAACAAAAGCCCCAAGAAACAGCCACATAACACAAT
>1325	GACCCCGCTCCACCTGCTTACCTCAGACAACGCCATCAACTGCCTGGAGGCCACCAGCGCACACA-
>1331	CACTGTGAAAACACCACAGGCTATGTAAGAGAGAGCAAGAGAGGCTTAACCGAAACCCCCAGAGGCAC
>1531	TCAAAAACAACACACTAGCTTTCTCCATTGAGCAGAGAGGGGCATCCACATTGAGAACCACAA--
>1594	AAGCCAAGTGGGGTCATGATACCGTACAAACAAGTGCTAGTTTGAAATTCTCCTCTTTAGGGAAGA--
>1738	GGCAGGAATTTCTATCAACTCTGCAATTACTACCATCCAGAAAAGCGGTTTGCTGTCAAAGCGTAC
>1836	TAGGGTTATTTAGACCATGCCAGATTTGCTTCAGAGCTTTTCAGCTTTCTTTTTACTATAAACCTCTT-
>1842	ATTCAAAGAGAAACTCCGAGGTGTCAATTCTAATTCCTCCACAATCGCCAAGACGTCTGAGTTTGAGG
>1973	TTTTGAGTTTTTGTGATGTGGCCTGTCAGCCAGAGTTTGAAGGGTGGCCACAGCTGACTGCTAACTT
>2037	TGTCGTGAGGGCTGCTCGTTCTGTCAGGACGATTCTCCGTGTTGACCCAGGAGAATGGTGTCTGCG
>2047	TGTCAGAAATGGGCTGCCTGATTTTTATTTGTTTTTCATCGGGGGTGGACATGTACTTTGTCTTG----
>2128	TGGGGACTCGGATACAGGCATTTGATAAAGGCGAGAGCTGCCACGGTCCCCTACCGATGACTGAT---
>2437	CAATGACCTCTTCAGCACACATAATGATAGCTCCTCTACTACACTGATTTCTTTTTGCAGTGTCCCT
>2448	CACTGTGAGCGTCAAAGACAGAACTGAAGCATGAG-AT-CGG-AAGAG-CGGTTCANGC-AGGAA-T-GCCGA----
>2459	-TGGAGCGCATTCCAGTTGTTGTGCCCCCTCTTTCCCAACCTCGAGCTAAACCCCTGCCATTTTGT-
>2494	TACAGTGAAGTGCTGAGATTAATAATGGATGCCCTTTGCCCTGCTGACACGCCATGTTAAGGTGG--
>2603	GCCTAAAGCTACAGGTGTGTTGCTGACAAGCCTGATGCTTGT-TGTTGCTAGCAGGCTTGAGGTTGTGA----
>2809	GACGGAGAAACAAACCTGTCTACTGGCAGCAAGCCTTTGGTTAATTGAC---ATCTACTGTATTGCTCAT

>2967	GAGTTTGCTTGTAGAAATGTATTATATTGACCAGTAGGTGGCATTGGAGACAGTGTGAATTAGAA---
>2975	GTAGTAAGGCCATTAATGTGTGCTGTATTTCTGCGTGACTGTCAGTTAGGTGTTTGTGGCCATGGTTG
>3005	GTTTATCAGTTGTGATATTGTGGCCATGCTGTGGCTGGGAGAGAGGGTTTGAAGTGTGAAGCAGGCT-
>3139	ACAATGCTGTCCGCTCTAAT-AACGGCTACGGCCCCGCTGAGAGCCTCATAACACCCGCCAGCAGCA---
>3254	TTGTGCACTCGCCAAAGCTGCCCCTTTTCCCCTCCCTATTGTCAGGAATTAACTTTTGATCACC---
>3438	GCCGCTCTCCCTGCTCAGATCTCGCCGAGTTCAGTGCCGCAGCGCGCGGAAAGCGGGGAGCTGCGC-
>3514	TCTACACTGCGTTTTAATCAAGATCCAGATCACGCTAGCAGCTCACTGTTTGTAGTTATAATTAATTA
>3540	CAGTGTACTATGATGTGGAGCTCATTACAGTTTTCTGTCTTTCAATAAAGTGCTGGGGACTGGAGCAGT
>3725	AAGGTTCAATACTTAAAATGGAGTCAGTTCCTTTCTGAGATTTTTGCTTTGTGGTGTGCGAGTGTTTTTG
>4008	GCTCCCGGAGGTGCTGTTACATTTATGTACGAG-TTTTTGTG-CATGA-GATC-GGAAGAG-CGGTT-CAGCAG
>4076	AATGTAATGGTACACAACGACATTTGGGTATCACAGCTGAAACTAAATGAGTAACCTCTGGACTTCGAA
>4122	GCGA-TTTGTCCCATAGAGATAGAGAGGG--GAAAAAAGGGTAACAGAACTACTATGGTATTACTGGGT
>4143	GGATGGCTATAATTGGCAGAATCATAAGAGAAATCATCTTTGGTGTGTGCGAGGTATGCGAGAATTCT
>4234	GGACCGAGGCGTCTCGCGCATTTTAATCTGGATCCCGTTTAATGTGGTGCTACATCAATTAAGAAGT-
>4350	TGCTGATTGTGTGAGGAGAGATATGTGTTTGTGTGCCAGCCAGAGAGATGAATGTGTCTGTAGTGGCT
>4371	ACGTCTCCAGGAGAGCAAATGCTGCACAGAAAGGTCAGCAGTTTTGCTTAATCCTTCTATAGAAAAGCAT
>4520	ACCCTGTTACATTCTCTTATACTAGGCATGAGAT-CGGAAG-AGCGG-TTCAG-CAGG-AATGNCCGAGACCG
>4755	GACAGCACCTCATCTCTGCAAATTGACTCCAGCTCGTTTAAAGTAGCAAGGCATGAGAT-CGGAAGAGC
>4835	TCCTCCAGGCTCCTCAGGGCTTTGTGGCTTTATCCATGATTATAAAGTGCCTGCTCCACCTCGTTT-
>4911	TCCTCCAGGCTCCTCAGGGCTTTGTGGCTTTATCCATGATTATAAAGTGCCTGCTCCACCTCGTTT-
>5146	GTGATCAGAACAAAAGTCTAGTGTAAATGGCTAGCTGCTAACGTAATAGTTTTACTGTTCCATAGT
>5225	CAGAGCTAGGCTCACACAAAGGCTGCATTGATGAGGAGGCAGACATAATAAAGGCTGCTCTGCCCCGC
>5294	ATATATAGCTACAGCACTGAAGAAAATCCAATGAGATAAGAAAATGTTTCAAATTGTTAGCTGG-
>5358	TGCACAGCAGCACATTTAACAGGTGCACCCTCCTAGGTCTAAACGCTGTCCGGCCAGCCGTTGCT-
>5451	GCTTTTATCCAAAGCTTTTATTTGACTTAATCACTCAAAAATACTTCAATCTCACTTCAAGCATGAGAT
>5489	ACGCTGCACATACGGTCAGTCTTACTATTCTCTTTCCATCAATCATGAACAAGAACTCGAGATACTT
>5621	GAACAGCAGCGTGACACATGAGAAGTACTTACTGCCCCCTGCTGGCTCTGAGACTTTTGGGA
>5631	AAGAGTCTTGGTTGGGGTACTGCTGAGAGACATGGCAGGAGAAAGGGTGGCTGGAGTAGGGAACCCAA
>5771	AGAAAGAACATAAACAGATCATAAGTCA-AAGTAATTGAGACTGAGAAAGAGACAAAAAGACAGCACTA
>5788	TCTCTTATAATAGAGCTTGCCCCAGAGGGACTGCCAGACGGTCAACGGAACAGGGAACAAGCCTTA
>5846	TTGTTTCTCCTAATGCTAAAATGTGATGACTTATTACAGTGCACATATTGTAAAGTAAAATCACCTTA
>5913	GGTTCAGTGGGCGATCGCGCTCTCGTACGCTGGCTGTCTGACAGGAAACTTGCCAGGTGTGCGCT
>5983	GGGAGCCCAGAAGCAATATACCTACTCGCTATTATGGAGCATG-AGATCGGNAAGAG-CGGTT--CA-GCAGGAA
>5994	TAGTCCACTCTGCTACTGTTACTGATGCCATTTCTGTCTTAAAAGCATGA-GATC-GGAAG-AGCGGTT
>5998	CGGCCAGCCGCTATTCAGTGCACCTGCCATTGTGTGGGCGCCCATCCAGGCCATTGTTGAGCGGAAA
>6146	GATTCCAGCAGGTTGCATTTGCAGATGCATGTCCTTGACCCGATGTCCGAAGCCTGGATTAAGGTGC

Fonte: Autor (2022).

**APÊNDICE 4 – Resultados de  $D_{XY}$  e  $D_a$  e  $F_{ST}$  entre cariomorfos de *E. erythrinus***

**Quadro 4** -  $D_{XY}$  entre cariomorfos

	<b>D</b>
<b>A</b>	0,0107855306805001

Fonte: Autor (2022).

**Quadro 5** -  $D_a$  entre cariomorfos

	<b>D</b>
<b>A</b>	0,0101377129514621

Fonte: Autor (2022).

**Quadro 6** -  $F_{ST}$  entre cariomorfos

	<b>D</b>
<b>A</b>	0,726498176008927

Fonte: Autor (2022).