

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

Pedro Henrique Narciso Ferreira

Os cromossomos sexuais como fonte geradora de diferenciação genética. O gênero *Pyrrhulina* (Characiformes, Lebiasinidae) como modelo.

São Carlos

2021

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

Pedro Henrique Narciso Ferreira

Os cromossomos sexuais como fonte geradora de diferenciação genética. O gênero *Pyrrhulina* (Characiformes, Lebiasinidae) como modelo.

Trabalho de conclusão de curso submetido como requisito para obtenção do grau de Bacharel em Biotecnologia

Orientador: Prof. Dr. Marcelo de Bello Cioffi

São Carlos
2021

Dedicado aos meus pais Pedro e Creuza (*in memoriam*), irmãos e sobrinhos, por todo suporte e amor.

Agradecimentos

Agradeço aos meus pais, os quais devo tudo que sou e realizo e que com muita dedicação construíram uma família exemplar pela qual muito me orgulho.

Aos meus irmãos, em especial a minha irmã Karina por sua presença essencial em todas as etapas da minha vida e pelo amor dedicado a todos nós.

A meu orientador Prof. Dr. Marcelo Bello Cioffi pela oportunidade, pelo apoio constante, dedicação e compreensão, pelo exemplo de competência científica e ensino, sendo a mim sempre fonte de inspiração e motivação.

Ao Dr. Manolo e Fernando, parceiros de Covil e referências como pesquisadores e pessoas, pela paciência, didática, sugestões, discussões e disponibilidade. Sendo também responsáveis por esse trabalho e grande parte da minha formação profissional e pessoal.

Aos meus amigos, pelo convívio e por fazerem a minha jornada na graduação uma experiência incrível, infinitamente mais leve e feliz, em especial meus amigos Júnior, Nico, La, Fernando, Estêvão, Iva, Guerra, Flor, Breno, Rafa, Fê e Amanda.

A minha tia Maria e primos por terem me proporcionado um lar em São Carlos, oferecendo-me muito mais que um teto, sendo conforto e carinho em meus dias.

À equipe do Laboratório de Citogenética de Peixes, pessoas fantásticas e muito solícitas que me ensinaram muito e participaram direta ou indiretamente desse projeto. Em especial a Renata e Francisco, por toda colaboração.

À todas essas pessoas, expresso meu muito obrigado. Dividam comigo o mérito desse trabalho!

Resumo

Os processos alopátricos são geralmente considerados uma das principais causas da notável biodiversidade da região Neotropical. No entanto, o papel dos rearranjos cromossômicos e dos cromossomos neo-sexuais na diversidade genética e especiação ainda é pouco investigado. Aqui, nosso objetivo foi inferir a contribuição relativa da alopatria, rearranjos cromossômicos e cromossomos sexuais para a divergência genética em cinco espécies de *Pyrrhulina* usando análises genômicas populacionais juntamente com dados citogenéticos. Nossos resultados destacaram que a distribuição da diversidade genética é consistente com seus dados cromossômicos e geográficos e foi influenciada principalmente pela alopatria de longa distância. Os resultados da PCoA indicaram um claro agrupamento de espécies localizadas na mesma região geográfica, exceto para *P. semifasciata*, que está em simpatria com *P. brevis*. No resultado do fastSTRUCTURE $K = 4$, apenas *P. australis* e *P. marilynae* foram agrupadas, e com $K = 3$ *P. obermulleri* e *P. brevis* foram posteriormente combinados. Neste caso, mostrou que espécies próximas geograficamente são geneticamente semelhantes, indicando mecanismos de especiação que são amplamente dependentes da alopatria. O efeito da alopatria também foi amplificado quando acoplado à ocorrência de rearranjos cromossômicos, como observado para *P. marilynae*, a espécie com cariótipo altamente rearranjado ($2n = 32$). Além disso, a estrutura genética indicou grande diferenciação em *P. semifasciata*, evidenciando que a presença do sistema cromossômico sexual X_1X_2Y , presente exclusivamente em indivíduos desta espécie, causou alta diferenciação genética e mostrando que tal efeito foi significativo mesmo na simpatria. Em suma, a distribuição da diversidade genética entre as espécies analisadas é influenciada principalmente pela alopatria de longa distância, mas a ocorrência de rearranjos cromossômicos, principalmente os relacionados ao sexo, contribuíram para a fixação de diferenças genéticas que podem levar ao isolamento reprodutivo e consequente processo de especiação dentro do gênero.

Palavras-chave: Peixes, cromossomos sexuais, diversidade genética, SNPs, especiação, alopatria.

Abstract

Allopatric processes are generally considered among the main causes of the striking Neotropical biodiversity. However, the role of chromosomal rearrangements and neo-sex chromosomes in genetic diversity and speciation remains still little investigated. Here, we aimed to infer the relative contribution of allopatry, chromosomal rearrangements, and sex chromosomes to genetic divergence in five *Pyrrhulina* species using population genomic analyses coupled with cytogenetic data. Our results highlighted that the distribution of the genetic diversity is consistent with their chromosomal and geographical data and was mainly influenced by long-distance allopatry. The results of PCoA indicated a clear clustering of species located in the same geographical region, except for *P. semifasciata*, that is in sympatry with *P. brevis*. In fastStructure K=4 result, only *P. australis* and *P. marilynae* were grouped, and with K=3 *P. obermulleri* and *P. brevis* were further combined. In this case, geographically close species showed genetically similarity, indicating speciation mechanisms that are largely dependent of allopatry. The effect of allopatry also were amplified when coupled with the occurrence of chromosomal rearrangements, as observed for *P. marilynae*, the species with highly rearranged karyotype (2n=32). Besides, the genetic structure analyzes indicated great differentiation in *P. semifasciata*, suggesting that the presence of the X1X2Y sex chromosome system, exclusively present in individuals of this species, caused high genetic differentiation and showing that such effect was significant even in sympatry. In summary, the distribution of genetic diversity among the analyzed species is mainly influenced long-distance allopatry, but the occurrence of chromosomal rearrangements, especially the sex-related ones contributed to the fixation of genetic differences which may lead to reproductive isolation and a consequent speciation process within the genus.

Key-words: Fish, sex chromosomes, genetic diversity, SNPs, speciation, allopatry

Lista de Figuras

Figura 1. Mapa da América do Sul indicando as espécies do gênero *Pyrrhulina* analisadas. As cores indicam diferentes bacias hidrográficas listadas na legenda e os círculos coloridos indicam os pontos de coletas das espécies: *P. australis* (azul), *P. brevis* (verde), *P. obermulleri* (roxo); *P. marilynae* (vermelho); e *P. semifasciata* (laranja). A elipse vermelha indica ponto de coleta onde as espécies estão em simpatria. Os ideogramas na direita representam parte do cariótipo de cada uma das espécies analisadas e os cromossomos sexuais presentes apenas em *P. semifasciata*.

15

Figure 2. Distribuição da diversidade genética nas espécies de *Pyrrhulina* de acordo com à PCoA. Indivíduos de diferentes espécies estão representados por círculos de diferentes cores. As abreviações representam os pontos de coleta: AM = Amazonas, MT = Mato Grosso.

27

Figure 3. Resultados do fastStructure para $K = 3$ e $K = 4$. Result for fastStructure with $K = 3$ and $K = 4$. As barras coloridas indicam os grupos em que o fastStructure classificou as espécies.

27

Lista de Tabelas

- Tabela 1.** Espécies analisadas, evidenciando seus número diplóide de cromossomos ($2n$), sistema de cromossomo sexual, número de indivíduos analisados citogeneticamente (Cito_N) de cada ponto de coleta e número de indivíduos sequenciados (DArT_N) 16
- Table 2.** Níveis de diversidade genética das espécies de *Pyrrhulina*. Abreviações: A_R - Riqueza alélica; P_A - Alelos privados; H_O - Heterozigosidade observada; H_E - Heterozigosidade esperada e F_{is} - Coeficiente de endogamia 28
- Table 3.** Resultado AMOVA, porcentagem de variação dentro e entre cada um dos grupos testados: (1) por espécies, (2) pelo melhor K resultante no fastStructure, (3) pelos grupos gerados na PCoA, (4) por região geográfica e (5) pela presença e ausência do sistema múltiplo de cromossomos sexuais. 26

Sumário

1. INTRODUÇÃO	10
1.1. A Região Neotropical	10
1.2. Alopatria, rearranjos cromossômicos e o processo de especiação	11
1.3. A família Lebiasinidae e gênero <i>Pyrrhulina</i>	12
2. OBJETIVOS	14
2.1. Diversidade genética e presença de estrutura genética	14
2.2. Papel dos sistemas sexuais múltiplos na diversidade genética	14
3. MATERIAIS E MÉTODOS	15
3.1. Coleta dos indivíduos	15
3.2. Sequenciamento DArT-Seq	16
3.3. Tratamento dos dados brutos de sequenciamento	17
3.4. Detecção de marcadores sobre seleção	17
3.5. Estrutura Genética	18
3.6. Diversidade genética e Análise de variância molecular	18
4. RESULTADOS E DISCUSSÃO	19
Abstract	20
Materials and Methods	22
Sampling	22
DArTseq Sequencing Procedure and Data Filtering	24
Assessment of Markers under Selection	25
Genetic Diversity and Analysis of Molecular Variance (AMOVA).	25
Genetic Structure	25
Results	26
Sequencing, filtering, and detection of markers under selection	26
Genetic diversity and structure	26
Discussion	28
The impact of sex chromosomes in genetic differentiation and speciation	30
5. CONSIDERAÇÕES FINAIS	31
6. REFERÊNCIAS	32

1. INTRODUÇÃO

1.1. A Região Neotropical

Estima-se que os peixes de água doce representam a um quarto do número total de vertebrados e metade dos peixes em geral, que somam mais de 60.000 espécies, distribuídas em 85 ordens e 536 famílias, principalmente na região Neotropical do planeta (Cavin, 2017; Nelson et al., 2016). Ainda que grande parte dessa fauna ainda esteja desconhecida, sabe-se que cerca de 6000 espécies válidas de peixe habitam essa região, tal diversidade é concordante com um padrão global de diversificação, conhecido como gradiente latitudinal de biodiversidade, reconhecido pelo decréscimo da biodiversidade de muitos grupos de organismos dos trópicos para os polos (Reis et al, 2003; Meseguer e Condamine, 2019).

Existem inúmeros modelos para explicar a origem de tal diversidade, embora ainda não haja um consenso na comunidade científica (Hillebrand, 2004). Dentre esses modelos, podemos citar a Hipótese de Conservação Tropical. De acordo com ele, novas espécies tendem a manter as mesmas características ecológicas que seus ancestrais, portanto considerando que a maioria das espécies se originaram em regiões tropicais, uma vez que essas regiões ocuparam uma área mais extensa do globo por um período de tempo muito maior, o conservadorismo de nicho juntamente com a complexa história geológica e climática da região Neotropical, podem explicar a maior biodiversidade dessa porção do planeta (Wiens e Graham, 2005 ; Wiens e Donoghue, 2004; Hinojosa, 2015; Hoorn et al., 2010; van Soelen et al., 2017). Além desta, umas das hipóteses mais aceitas denominada “Out of the Tropics”, tem como premissa a baixa taxa de extinção somada à alta taxa de especiação da região Neotropical (Jablonski et al., 2006). As causas associadas a esse processo também estão relacionadas à complexa história geológica e climática da região (Hoorn et al., 2010; van Soelen et al., 2017). Dentre esses eventos, podemos citar o resfriamento global após períodos de altas temperaturas durante o Eoceno, que afetou negativamente as espécies presentes em regiões distantes dos trópicos, por estas serem mais adaptadas ao clima quente, além de dificultar sua dispersão (Meseguer e Condamine, 2019)

Dessa maneira, fica evidente que o clima e geologia característica da região tiveram papel fundamental para a impressionante composição taxonômica da ictiofauna Neotropical (Albert & Reis, 2011). Apesar disso, os processos evolutivos responsáveis por essa grande

biodiversidade permanecem pouco explorados, justificando esforços para uma melhor compreensão e avaliação dos mecanismos genéticos que moldam a enorme biodiversidade entre os peixes nas regiões de baixa latitude.

1.2. Alopatria, rearranjos cromossômicos e o processo de especiação

Dentre as diversas definições existentes, a especiação pode ser descrita como um processo pelo qual uma população geneticamente independente é dividida em outras populações, como resultado de eventos que levam à diferenciação genética e ao isolamento reprodutivo (Mayr 1954). Nesse sentido, a alopatria é frequentemente considerada a forma mais comum de especiação, causando divergência entre populações por meio de reconfigurações da paisagem e, portanto, pode ter contribuído para a biodiversidade da região Neotropical (Seehausen & Wagner, 2014).

A evolução de barreiras reprodutivas é crucial para que o processo de especiação prossiga e há evidências de que os rearranjos cromossômicos, principalmente envolvendo cromossomos sexuais, desempenham um papel fundamental na evolução das barreiras reprodutivas entre as espécies, uma vez que há redução do fluxo gênico (Laurie 1997; Masly & Presgraves 2007; Presgraves & Orr 1998; Turner et al. 2014). Entre os animais, a maioria dos casos de isolamento pós-zigótico são causados por incompatibilidades genéticas, entre as quais ressaltamos os rearranjos cromossômicos e desenvolvimento de cromossomos sexuais (Bateson, 1909; Dobzhansky, 1933; Muller, 1942; Coyne & Orr, 2004). No entanto, pouco se sabe sobre as contribuições dos cromossomos sexuais múltiplos para o isolamento reprodutivo e sua relevância na origem da biodiversidade na região Neotropical.

Do ponto de vista genético, avanços consideráveis permitem a compreensão dos tipos mais comuns de mecanismos de isolamento reprodutivo, sua evolução e causas subjacentes, particularmente no que diz respeito ao isolamento pós-zigótico. Dessa forma, o aperfeiçoamento recente das técnicas de sequenciamento de nova geração forneceram novas perspectivas para a avaliação da diversidade genética e estrutura populacional em muitas espécies animais e vegetais, em especial entre os peixes (Yang et al., 2013; Alkimin et al., 2018; Alfaro et al. 2009, Rabosky et al. 2013).

1.3. A família Lebiasinidae e gênero *Pyrrhulina*

A família Lebiasinidae representa um grupo de peixes de água doce Neotropicais, com 75 espécies (Fricke et al., 2021), pertencente à ordem Characiforme, está amplamente distribuídos por toda a América Central e do Sul, com exceção do Chile (Froese e Pauly, 2014) Dividida em duas subfamílias Lebiasininae e Pyrrhulininae e sete gêneros *Lebiasinina*, *Piabucina*, *Derhamia*, *Nannostomus*, *Copeina*, *Copella* e *Pyrrhulina*, seus exemplares são geralmente caracterizados pelo seu tamanho diminuto, variando entre 1,6 e 20 cm de comprimento, e também pelo corpo cilíndrico e alongado com alguns dimorfismos sexuais (Weitzman e Weitzman, 2003). Pyrrhulininae é o claro mais diversificado da família, composto pelos gêneros *Derhamia*, *Nannostomus*, *Copella*, *Copeina* e *Pyrrhulina* (Fricke et al., 2021). Recentemente trabalhos envolvendo análises citogenéticas e moleculares de algumas espécies dos gêneros têm contribuído consideravelmente para a compreensão das relações filogenéticas internas do grupo (Moraes et al., 2017, 2019, 2021; Sassi et al., 2019; Toma et al., 2019; Sember et al., 2020).

Estudos prévios indicam que o número cromossômico no gênero *Pyrrhulina* varia de 32 a 42 cromossomos, onde *Pyrrhulina aff. australis*, *Pyrrhulina australis*, *Pyrrhulina aff. marilynae* e *Pyrrhulina sp* apresentam $2n = 40$ cromossomos, sem diferenciação cariotípica entre os sexos (Moraes et al., 2017; 2021). Por sua vez, *P. brevis*, *P. obermulleri* e *Pyrrhulina cf. laeta* tem $2n = 42$ cromossomos, tanto em machos quanto em fêmeas. Ao contrário das demais espécies do gênero, *P. semifasciata* apresenta $2n = 42$ nas fêmeas e $2n = 41$ cromossomos nos machos, com presença de três cromossomos não pareados, dois pequenos acrocêntricos e um metacêntrico exclusivo, evidenciando a presença de um sistema sexual múltiplo do tipo $X_1X_1X_2X_2 / X_1X_2Y$ sendo a única espécie na família com cromossomos sexuais heteromórficos até o momento (Moraes et al., 2017; Moraes et al., 2019). Além disso, trabalhos atuais também mostraram que *P. marilynae* apresenta $2n = 32$, diferindo dos demais representantes do grupo, provavelmente devido a uma série de fusões cromossômicas (Moraes et al., 2021).

Adicionalmente, experimentos de hibridização genômica comparativa (CGH) e análises de pintura cromossômica total (WCP) foram empregados, onde o gDNA e o cromossomo Y de *P. semifasciata* foram mapeados nas outras espécies de *Pyrrhulina*

evidenciando as relações entre as espécies e confirmaram a ocorrência do sistema de cromossomos sexuais múltiplos do tipo $X_1X_1X_2X_2/X_1X_2Y$ exclusivamente em *P. semifasciata* (Moraes et al., 2019).

Portanto, esses trabalhos ajudaram a levantar hipóteses sobre as relações evolutivas e origem e evolução do sistema de cromossomos sexuais múltiplos no gênero. Contudo, apenas os dados citogenéticos geralmente não permitem testar hipóteses explícitas de isolamento reprodutivo ou diferenciação genética como os dados genômicos, por serem frequentemente associados com padrões geográficos (climáticos e/ou geológicos). O uso de pares de espécies e/ou populações em que há variações relacionadas a efeitos de alopatria, rearranjos cromossômicos e ocorrência de sistemas de cromossomos sexuais constitui uma abordagem frutífera para a compreensão do papel desses eventos na origem da biodiversidade. Em vista disso, neste estudo nós analisamos cinco espécies de *Pyrrhulina* simpátricas e alopátricas para testar a hipótese de que a diferenciação dos cromossomos sexuais representa um fator chave na diferenciação genética de espécies portadoras de tais sistemas, promovendo o isolamento reprodutivo em maior proporção quando comparado à alopatria.

2. OBJETIVOS

Levando em consideração o embasamento acima, o presente trabalho objetiva a continuidade e avanço nos estudos até então realizados na família Lebiasinidae, analisando espécies do gênero *Pyrrhulina* pela abordagem da genômica comparativa, correlacionando os resultados obtidos com parâmetros cromossômicos e genômicos. Tais procedimentos possibilitarão uma melhor compreensão dos mecanismos de evolução genômica e cromossômica destas espécies, em consonância com as seguintes propostas investigativas:

2.1. Diversidade genética e presença de estrutura genética

Serão utilizados marcadores moleculares do tipo “Single Nucleotide Polymorphism” (SNP) obtidos por sequenciamento de nova geração, de modo a compreender melhor como a diversidade se distribui entre as diferentes espécies, bem como investigar a presença de estrutura genética entre elas.

2.2. Papel dos sistemas sexuais múltiplos na diversidade genética

Uma alta variabilidade cariotípica e espécies com e sem sistemas sexuais múltiplos podem ser encontradas no gênero *Pyrrhulina*. No entanto, as relações das influências da alopatria sobre a ocorrência de rearranjos e a origem de cromossomos sexuais ainda é muito discutida. Sendo assim, não há um consenso relacionado à real influência dos rearranjos como redutores de fluxo gênico, tornando os estudos das populações naturais que apresentam sistemas de cromossomos sexuais múltiplos muito interessantes à comunidade científica. Nesse sentido, a presença de um sistema sexual múltiplo em *Pyrrhulina*, nos leva aos seguintes questionamentos: A presença de um sistema sexual múltiplo tem implicações na diversidade genética de uma espécie? O surgimento de cromossomos sexuais múltiplos pode representar uma força evolutiva mais efetiva do que somente a ação da alopatria na geração da diversidade genética e na redução do fluxo gênico?

3. MATERIAIS E MÉTODOS

3.1. Coleta dos indivíduos

Utilizamos 28 espécimes de cinco espécies do gênero *Pyrrhulina*, os pontos de coletas, número amostral, dados citogenéticos, sexo, número de indivíduos sequenciados e espécie investigada são apresentados na **Figura 1** e **Tabela 1**. Estudos prévios realizados com as mesmas amostras providenciaram os dados citogenéticos (Moraes et al., 2017, 2019, 2021). Todos as amostras foram coletadas com autorização do Instituto Chico Mendes de Conservação da Biodiversidade e do Sistema de Autorização e Informação em Biodiversidade (ICMBio/ SISBIO - licença n°. 48628-14), bem como do Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado (SISGEN, A96FF09).

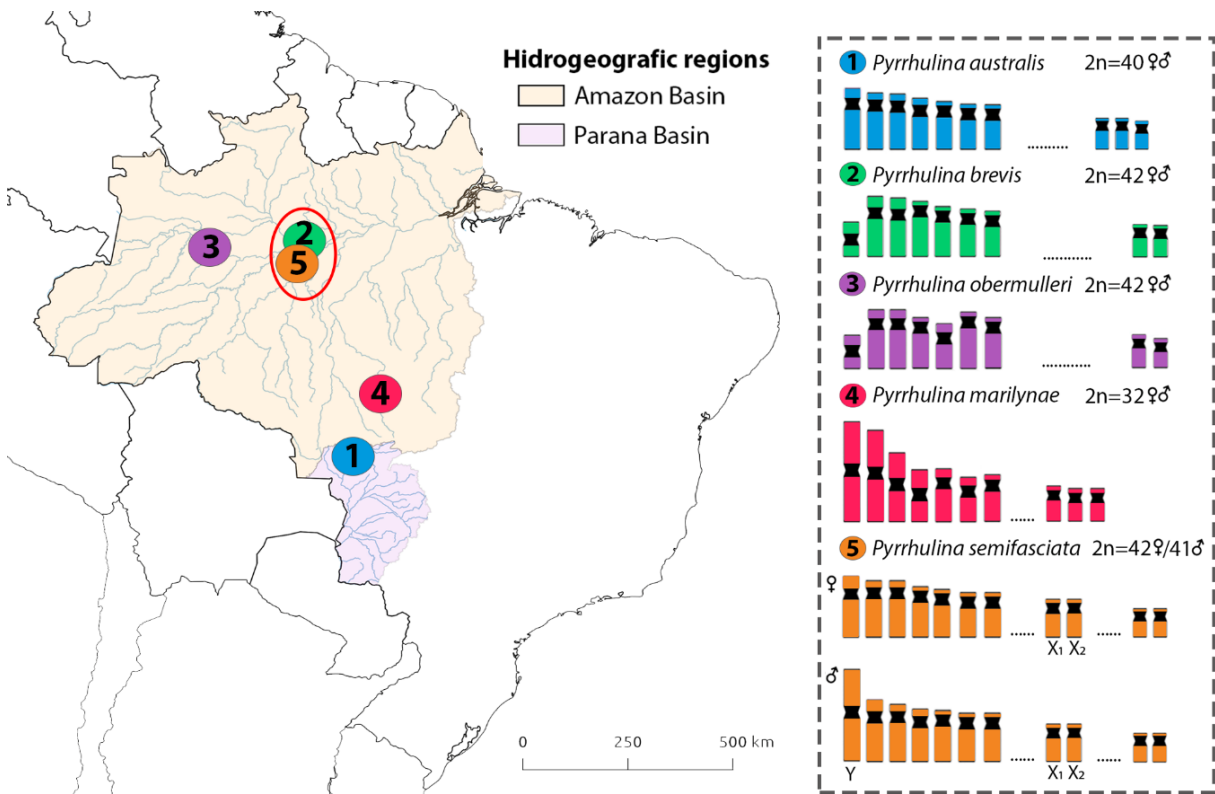


Figura 1. Mapa da América do Sul indicando as espécies do gênero *Pyrrhulina* analisadas. As cores indicam diferentes bacias hidrográficas listadas na legenda e os círculos coloridos indicam os pontos de coletas das espécies: *P. australis* (azul), *P. brevis* (verde), *P. obermulleri* (roxo); *P. marilynae* (vermelho); e *P. semifasciata* (laranja). A elipse vermelha indica ponto de coleta onde as espécies estão em simpatria. Os ideogramas na direita representam parte do cariótipo de cada uma das espécies analisadas e os cromossomos sexuais presentes apenas em *P. semifasciata*.

Tabela 1. Espécies analisadas, evidenciando seus número diplóide de cromossomos (2n), sistema de cromossomo sexual, número de indivíduos analisados citogeneticamente (Cito_N) de cada ponto de coleta e número de indivíduos sequenciados (DArT_N).

			Ponto de coleta		Número de indivíduos		
Espécies	2n	Cromossomos sexuais	Nome	Latitude/Longitude	Cito_N	DArT_N	Referência
<i>P. australis</i>	♀♂40	Homomórfico	Barra do Bugres – MT	-15.04275, -57.11054	18♂ 30♀	6	Moraes et al., 2017
<i>P. brevis</i>	♀♂42	Homomórfico	Adolpho Ducke Reserve – AM	-2.58207, -59.55530	17♂ 13♀	6	Moraes et al., 2019
<i>P. semifasciata</i>	♀42 ♂41	X ₁ X ₁ X ₂ X ₂ /X ₁ X ₂ Y	Careiro – AM	-3.51000, -60.04000	12♂ 07♀	6	Moraes et al., 2019
<i>P. obermulleri</i>	♀♂42	Homomórfico	Tefé – AM	-3.25507, -64.44548	21♂ 12♀	6	Moraes et al., 2021
<i>P. marilynae</i>	♀♂32	Homomórfico	Ipiranga do Norte – MT	-11.36020, -55.56270	14♂08 ♀	6	Moraes et al., 2021

3.2. Sequenciamento DArT-Seq

Tecido do fígado de todos os indivíduos (**Tabela 1**) foram enviadas à empresa Diversity Array Technology Pty Ltd localizada em Canberra, Austrália, para realização do sequenciamento DarTseq, descrito em Kilian et al. (2012), na plataforma Illumina HiSeq 2500. Essa técnica de sequenciamento consiste no uso de duas diferentes enzimas de restrição combinadas para redução da complexidade do genoma. A utilização da enzima *PstI* de corte raro e sensível a determinados tipos de metilação e a *SbfI* enzima de corte frequente, permite que regiões metiladas e repetitivas possam ser excluídas antes da realização das etapas subsequentes (Akbari et al., 2006; Wenzl et al., 2006). Sendo assim o processo tende a

sequenciar as regiões hipometiladas do DNA, potencialmente expressas, e por isso apesar da redução da complexidade o método apresenta uma boa cobertura (Akbari *et al.*, 2006).

3.3. Tratamento dos dados brutos de sequenciamento

Após o sequenciamento, os dados brutos já demultiplexados preparados pela empresa foram enviados para o Laboratório de Citogenética de Peixes para as etapas de triagem e obtenção dos marcadores, na forma de 28 arquivos comprimidos do tipo FASTQ. O processamento foi realizado através do uso do software ipyrad v. 0.7.28 (Eaton e Overcast, 2017). Para realização da filtragem dos dados um arquivo de parâmetros foi criado, além um arquivo contendo o barcode (disponibilizado pela empresa) de todas as amostras para a remoção deles. Dessa maneira, os adaptadores foram removidos, além de todas as sequências menores que 35 pares de bases ou com mais de cinco bases de baixa qualidade (considerando $Q < 20$). Posteriormente, sequências consenso foram obtidas por meio do alinhamento e clusterização dos *reads* separadamente e pré-*locus* foram obtidos. Os pré-*locus* foram então agrupados com base nos indivíduos e apenas aqueles presentes em todos os indivíduos foram mantidos, gerando *loci* definitivos. Na intenção de evitar a obtenção de regiões genômicas com potencial paralogia, um SNP aleatório foi selecionado para cada *loci* definitivo. Os marcadores obtidos foram compilados em uma matriz de SNPs, com *loci* codificados como 0 para homocigotos de referência, 1 para heterocigotos e 2 para homocigotos de estado alternativo.

3.4. Detecção de marcadores sobre seleção

Para a busca de marcadores potencialmente sob seleção, foi realizada uma análise no software BayeScan (Foll e Gaggiotti, 2008). As execuções foram realizadas com valores de probabilidade a priori de 100, um total de 5000 MCMC (Markov Chain Monte Carlo) foram realizadas e o valor de *thinning* foi estabelecido como dez. Os valores de *False Discovery Rate* (FDR) foram usados para classificar *outliers*, e apenas *loci* com valores menores que 0,01 foram considerados *outliers* e potencialmente sob seleção, sendo removidos das análises seguintes.

3.5. Estrutura Genética

Para visualizar a distribuição da diversidade genética entre as espécies, uma análise de coordenadas principais (PCoA) foi realizada utilizando o pacote `dartR` do R Studio (Gruber et al., 2018). Posteriormente, a estrutura genética foi investigada com o método não espacial `fastSTRUCTURE` (Raj et al., 2014), baseado no método Bayesiano `Structure` (Pritchard et al. 2000), adaptado para conjuntos de dados genômicos. A análise foi performada utilizando o pipeline “*Lizards-are awesome*” (Melville et al., 2017), com o intervalo de K estabelecido entre um e cinco, também utilizamos o comando `choosek.py` para selecionar o valor de K que se mostrou mais informativo para nossos dados. Os arquivos de saída do `fastSTRUCTURE` foram compactados e visualizados graficamente no `Clumpak` (Kopelman et al., 2015).

3.6. Diversidade genética e Análise de variância molecular

O pacote `dartR` do R Studio também foi usado para estimar a diversidade genética das espécies. Para isso, foram obtidos os seguintes índices de diversidade: Riqueza Alélica (A_R), Alelos Privados (P_A), Heterozigosidade Observada (H_O), Heterozigosidade Esperada (H_E), e o Coeficiente de Endogamia (F_{IS}). Uma análise de variância molecular (AMOVA) também foi realizada. As análises de AMOVA foram executadas com as amostras agrupadas das seguintes formas: (1) por espécie, (2) pelo melhor K resultante no `fastSTRUCTURE`, (3) pelo agrupamento gerado no PCoA, (4) pela região geográfica da coleta e (5) pela presença ou ausência do sistema de cromossomos sexuais múltiplos. As estatísticas foram calculadas de acordo com Sherwin et al. (2017).

4. RESULTADOS E DISCUSSÃO

Os resultados obtidos no presente estudo e a sua discussão encontram-se organizados no formato de capítulo que constitui um artigo científico a ser submetido para publicação em periódico especializado da área.

Allopatry is only part of the story: Impact of chromosomal rearrangements and sex chromosomes in genetic differentiation and speciation in the genus *Pyrrhulina* (Teleostei, Characiformes).

Pedro H. N. Ferreira¹, Renata L. R. de Moraes¹, Manolo F. Perez¹, Fernando H. S. de Souza¹, Francisco de M. C. Sassi¹, Patrik F. Viana², Eliana Feldberg², Manoela Maria Ferreira Marinho^{3,4}, Luiz A. C. Bertollo¹ and Marcelo de B. Cioffi¹

1 Laboratório de Citogenética de Peixes, Departamento de Genética e Evolução, Universidade Federal de São Carlos, 13565-905 São Carlos, SP, Brazil.

2 Instituto Nacional de Pesquisas da Amazônia, Coordenação de Biodiversidade, Laboratório de Genética Animal, Av. André Araújo 2936, Petrópolis, CEP 69067-375, Brazil.

3 Museu de Zoologia da Universidade de São Paulo (MZUSP), São Paulo, Brazil,

4 Laboratório de Sistemática e Morfologia de Peixes, Departamento de Sistemática e Ecologia (DSE), Universidade Federal da Paraíba (UFPB), João Pessoa, Brazil

Abstract

Allopatric processes are generally considered among the main causes of the striking Neotropical biodiversity. However, the role of chromosomal rearrangements and neo-sex chromosomes in genetic diversity and speciation remains still little investigated. Here, we aimed to infer the relative contribution of allopatry, chromosomal rearrangements, and sex chromosomes to genetic divergence in five *Pyrrhulina* species using population genomic analyses coupled with cytogenetic data. Our results highlighted that the distribution of the genetic diversity is consistent with their chromosomal and geographical data and was mainly influenced by long-distance allopatry. The results of PCoA indicated a clear clustering of species located in the same geographical region, except for *P. semifasciata*, that is in sympatry with *P. brevis*. In fastStructure K=4 result, only *P. australis* and *P. marilynae* were grouped, and with K=3 *P. obermulleri* and *P. brevis* were further combined. In this case, geographically close species showed genetically similarity, indicating speciation mechanisms that are largely dependent of allopatry. The effect of allopatry also were amplified when coupled with the occurrence of chromosomal rearrangements, as observed for *P. marilynae*, the species with highly rearranged karyotype (2n=32). Besides, the genetic structure analyzes indicated great differentiation in *P. semifasciata*, suggesting that the presence of the X₁X₂Y sex chromosome system, exclusively present in individuals of this species, caused high genetic differentiation and showing that such effect was significant even in sympatry. In summary, the distribution of genetic diversity among the analyzed species is mainly influenced long-distance allopatry, but the occurrence of chromosomal rearrangements, specially the sex-related ones contributed to the fixation of genetic differences which may lead to reproductive isolation and a consequent speciation process within the genus.

Introduction

The Neotropical region harbors most of the fish diversity, with more than 5200 valid fish species (Fricke et al., 2021). However, despite representing such an important group, a large portion of this diversity remains undescribed. With countless questions yet to be answered, some studies indicate that only less than half of Neotropical freshwater fish species are known (Almeida-Toledo, 1998; Pereira et al., 2013). There are numerous possible explanations for this spectacular diversity, and in this regard, allopatry is often considered the most common form of speciation, promoting divergence among populations through landscape reconfigurations, thus contributing to the impressive biodiversity of the Neotropical region (Seehausen & Wagner, 2014). In fact, the evolution of reproductive barriers is crucial for the speciation process to proceed since it reduces gene flow and plays a key role in the evolution of reproductive incompatibilities between species. Among animals, most cases of post-zygotic isolation are caused by genetic incompatibilities, among which chromosomal rearrangements and development of sex chromosomes are of fundamental importance (Bateson, 1909; Dobzhansky, 1933; Muller, 1942; Coyne & Orr, 2004). However, little is known about the contribution of chromosomal rearrangements and multiple sex chromosomes to reproductive isolation and their relevance in the origin of biodiversity in the Neotropical region.

The Neotropical Lebiasinidae is a fish family distributed throughout Central and South America, with 75 species divided into two subfamilies: Lebiasininae and Pyrrhulinae (Fricke et al., 2021; Froese & Pauly, 2014; Weitzman & Weitzman, 2003). Pyrrhulinae is the most diverse subfamily, composed of the genera *Derhamia*, *Nannostomus*, *Copella*, *Copeina*, and *Pyrrhulina* (Fricke et al., 2021). Recent studies involving cytogenetic and molecular analyses of some of the species from these genera have contributed considerably to the comprehension of the phylogenetic relationships in the family (Moraes et al., 2017, 2019, 2021; Sassi et al., 2019; Sember et al., 2020; Toma et al., 2019). The 2n in the genus *Pyrrhulina* ranges from 32 to 42 chromosomes, with *Pyrrhulina* aff. *australis*, *P. australis*, *Pyrrhulina* aff. *marilynae* and *Pyrrhulina* sp present 2n = 40 chromosomes, without karyotypic differentiation between sexes (Moraes et al., 2017; 2021). In turn, *P. brevis*, *P. obermulleri* and *Pyrrhulina* cf. *laeta* have 2n = 42 chromosomes, both in males and females.

Unlike the other species of the genus, *P. semifasciata* show a $2n = 42$ in females and $2n = 41$ chromosomes in males, being the only analyzed species of the family with heteromorphic sex chromosomes. It presents three unpaired chromosomes, two small acrocentrics and one metacentric, evidencing the presence of a multiple $X_1X_1X_2X_2 / X_1X_2Y$ sex system (Moraes et al., 2017; Moraes et al., 2019). Besides, recent studies have shown a $2n=32$ in *P. marilynae*, differing from the rest of the group by showing a series of chromosomal fusions (Moraes et al., 2021).

The development and improvement of large-scale genotyping-by-sequencing techniques allowed high-resolution analysis using Single Nucleotide Polymorphisms (SNP) for genetic diversity studies in non-model organism investigations (Steane et al., 2011; Yang et al., 2013; Sánchez-Sevilla et al., 2015; Alkimin et al., 2018). Combining such genomic data with cytogenetic information is a powerful approach to test explicit hypotheses of reproductive isolation and genetic differentiation. The use of pairs of species/ populations in which there are variations related to allopatry, chromosome rearrangements and sex chromosome systems constitutes a fruitful approach to understanding the role of these events in the origin of biodiversity. Therefore, in this study we analyzed five *Pyrrhulina* species with distinct sex chromosome systems and geographical distribution (sympatric and allopatric) to measure the impact of allopatry, chromosomal rearrangements and sex chromosome differentiation in the genetic differentiation of these species.

Materials and Methods

Sampling

The collection sites, number, and sex of the specimens are given in **Table 1** and **Figure 1**. Previous studies already provided cytogenetic information for these same samples (Moraes et al., 2017, 2019, 2021). Sampling authorizations were obtained from the Brazilian Environmental Agency, Instituto Chico Mendes de Conservação da Biodiversidade/ Sistema de Autorização e Informação em Biodiversidade (ICMBio/ SISBIO - license nº. 48628-14) and Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado (SISGEN - A96FF09).

Table 1. Species analyzed, their diploid chromosome numbers (2n), sex chromosome system, numbers of individuals cytogenetically analyzed from each site and number of individuals sequenced. Abbreviations: AM = Amazonas, MT = Mato Grosso Brazilian States.

Species	2n	Sex chromosomes	Sampling site		Number of individuals		Reference
			Name	Latitude/Longitude	Cytological analysis	Sequence analysis	
<i>P. australis</i>	♀♂40	Homomorphic	Barra do Bugres – MT	-15.04275, -57.11054	18♂ 30♀	6	Moraes et al., 2017
<i>P. brevis</i>	♀♂42	Homomorphic	Adolpho Ducke Reserve – AM	-2.58207, -59.55530	17♂ 13♀	6	Moraes et al., 2019
<i>P. semifasciata</i>	♀42 ♂41	X ₁ X ₁ X ₂ X ₂ /X ₁ X ₂ Y	Careiro – AM	-3.51000, -60.04000	12♂ 07♀	6	Moraes et al., 2019
<i>P. obermulleri</i>	♀♂42	Homomorphic	Tefé – AM	-3.25507, -64.44548	21♂ 12♀	6	Moraes et al., 2021
<i>P. marilynae</i>	♀♂32	Homomorphic	Ipiranga do Norte – MT	-11.36020, -55.56270	14♂08♀	6	Moraes et al., 2021

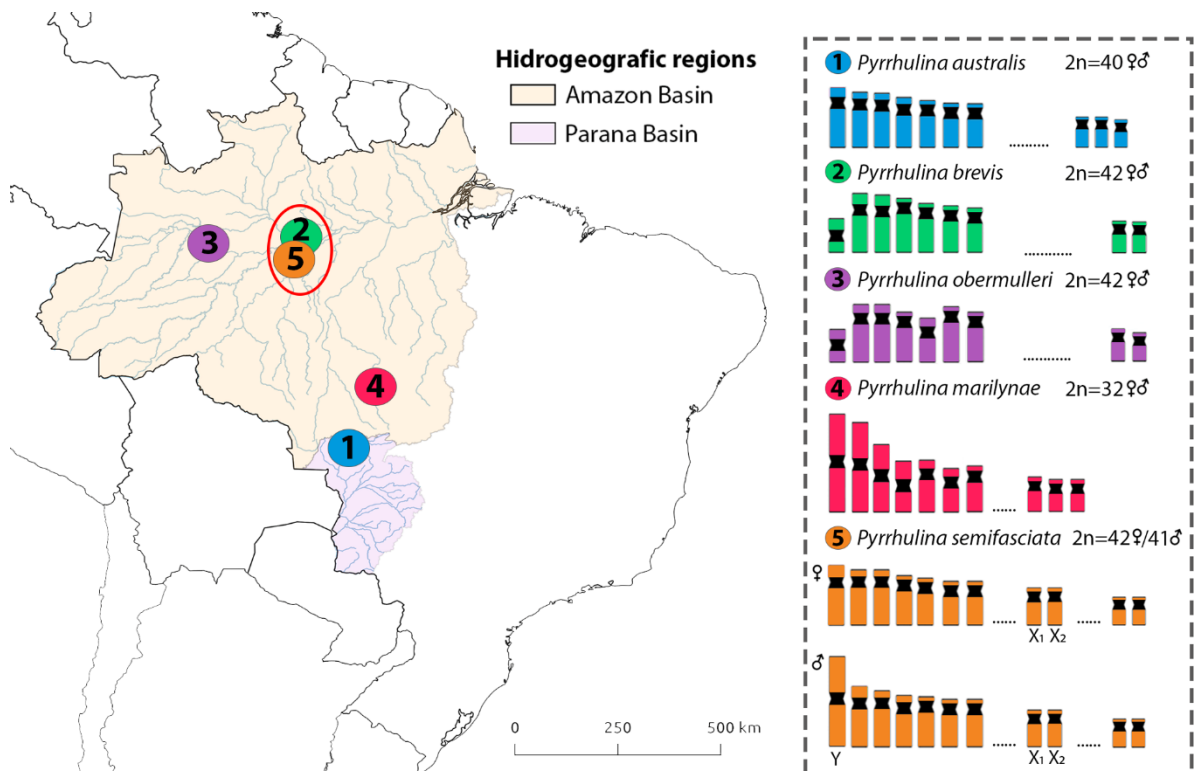


Figure 1. Map of South America indicating the *Pyrrhulina* species analyzed. Colored contours indicate the different hydrographic basins listed in the legend and the colored circles show the collection sites of the species: *P. australis* (blue), *P. brevis* (green), *P. obermulleri* (purple); *P. marilynae* (red); and *P. semifasciata* (orange). The ellipse indicates a site with sympatric species. The box shows Ideograms with partial karyotypes of each species and the sex chromosomes present exclusively in *P. semifasciata*.

DARtseq Sequencing Procedure and Data Filtering

Liver tissue of all individuals (Table 1) were used to perform DARtseq sequencing procedures at Diversity Arrays Technology Pty Ltd. This sequencing technique uses a combination of restriction enzymes (SbfI and PstI) that enrich for hypomethylated regions (Kilian et al., 2012). Sequencing of the obtained libraries was carried out on the Illumina HiSeq 2500 platform. Raw sequencing data was treated with Ipyrad (Eaton & Overcast, 2017). Sequencing adapters were trimmed, and all sequences shorter than 35 base pairs or presenting more than 5 low-quality bases ($Q < 20$) were discarded. Afterwards, alignment of reads was carried out, and consensus sequences for each sample were obtained. A polymorphic SNP matrix was created with only one SNP per loci, to ensure the absence of

linked SNPs. Genotypes were coded as 0 for reference homozygotes, 1 for heterozygotes, and 2 for alternate state homozygotes.

Assessment of Markers under Selection

To search for markers that are potentially under selection, we applied the analysis from the BayeScan software (Foll & Gaggiotti, 2008). We performed a total of 5000 MCMC (Markov Chain Monte Carlo) runs and thinning of 10, with a prior odds value of 100. Values of False Discovery Rate (FDR) were used to classify outliers, and only loci with values lower than 0.01 were considered outliers and removed from subsequent analysis.

Genetic Diversity and Analysis of Molecular Variance (AMOVA).

The R package *dartR* was also used to estimate genetic diversity. We calculated allelic richness (A_R), private alleles (P_A), observed heterozygosity (H_O), expected heterozygosity (H_E) and inbreeding coefficient (F_{IS}). We performed an analysis of molecular variance (AMOVA) with samples grouped as follows: (1) by species, (2) by the clusters from best K on *fastStructure*, (3) by the clustering pattern generated on the PCoA, (4) by geographical region and (5) by the presence or absence of the multiple sex chromosome system.

Genetic Structure

To visualize the distribution of genetic diversity among species, we performed a principal coordinate analysis (PCoA) in the R package *dartR* (Gruber et al., 2018). The genetic structure was also investigated with *fastSTRUCTURE* (Raj et al., 2014), using the “Lizards-are-awesome” pipeline. We used the *choosek.py* command to select the number of clusters that maximizes likelihood and is more informative for the structure from our dataset. The results of *fastStructure* were visualized with *Clumpak* (Kopelman et al., 2015).

Results

Sequencing, filtering, and detection of markers under selection

Sequencing of DArTSeq resulted in an average of 2 million reads per sample. After all filtering steps, a total of 5,149 single nucleotide polymorphism (SNP) markers were obtained. According to the BayeScan result, none of the loci were detected as potentially under selection, as all FDR values were higher than 0.65. Thus, all markers were maintained in subsequent analysis.

Genetic diversity and structure

The obtained genetic diversity indexes are presented in **Table 3**. We observed slightly higher indices in *P. obermulleri* and *P. semifasciata*, both collected in AM. Contrarily, the species *P. marilynae*, collected in MT, showed the lowest values for all summary statistics, except for the number of private alleles (P_A). The PCoA recovered 37.7% of the total variation in the first principal coordinate (PC1), and 21.6% in the second principal coordinate (PC2). The results indicate that *P. brevis* and *P. obermulleri* are more closely related to each other, as are *P. australis* and *P. marilynae*. In turn, *P. semifasciata* proved to be the most distantly related species of the group (**Figure 2**).

Table 3. Genetic diversity levels of *Pyrrhulina* species. Abbreviations: N - Number of individuals; A_R - Allelic richness; P_A - Private alleles; H_O - Observed heterozygosity; H_E - Expected heterozygosity and *Fis* - Inbreeding coefficient.

Genetic diversity						
Species	N	A_R	P_A	H_O	H_E	<i>Fis</i>
<i>P. australis</i>	6	1.132	4459	0.043	0.040	0.010
<i>P. brevis</i>	6	1.107	4583	0.016	0.028	0.422
<i>P. obermulleri</i>	5	1.193	4098	0.061	0.058	0.033
<i>P. marilynae</i>	5	1.02	5017	0.008	0.007	-0.008
<i>P. semifasciata</i>	6	1.197	4115	0.052	0.050	0.0301

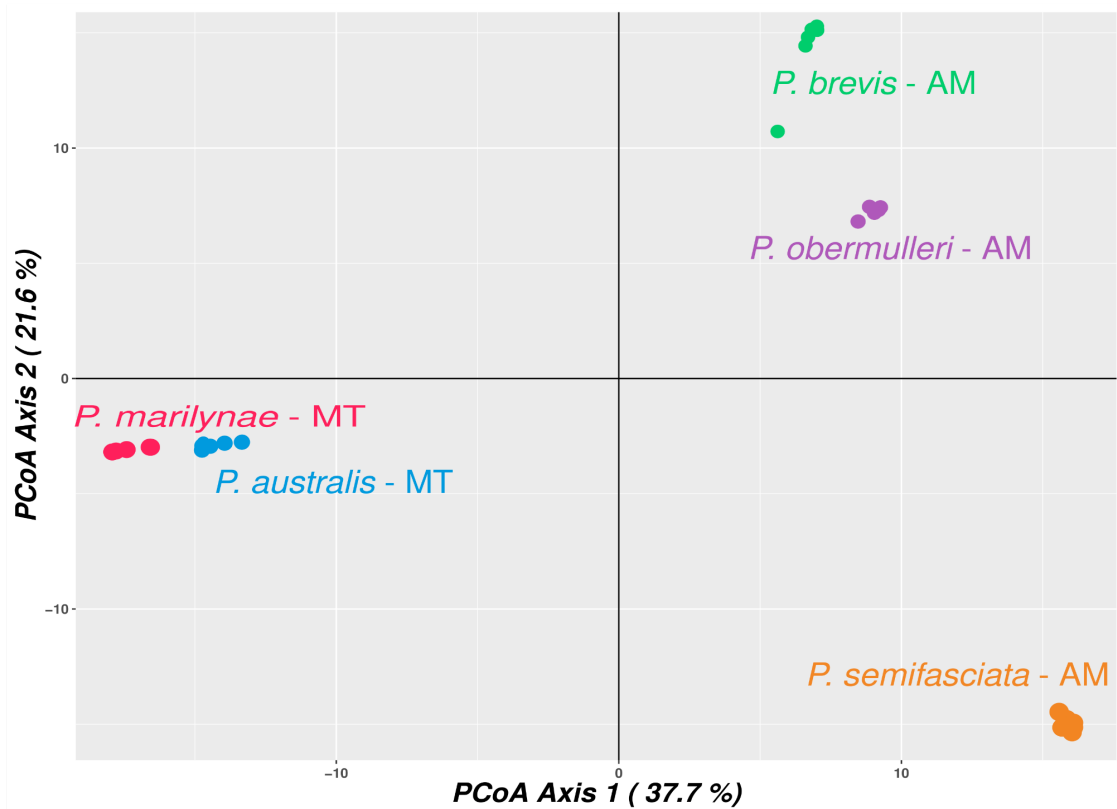


Figure 2. Distribution of the genetic diversity in *Pyrrhulina* species according to a PCoA. Individuals of different species are represented by circles of different colors. Abbreviations represented the sampling sites: AM= Amazonas, MT = Mato Grosso Brazilian States

Results from fastStructure suggested a $K = 4$ as the most likely value. Additionally, we decided to present the clustering pattern generated with $K = 3$, which was congruent with the aforementioned PCoA result (**Figure 3**).

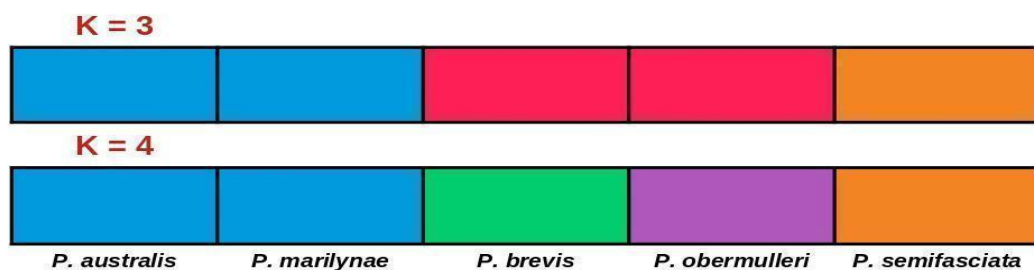


Figure 3. Result for fastStructure with $K = 3$ and $K = 4$. Bar colors represent the group in which fastStructure classified the species.

The AMOVA suggested that the clustering strategy presenting the highest variation between and the lowest variation value within groups was achieved with groups formed by each separate species (83.66% variation between groups), followed by the fastStructure K = 4 result and by the PCoA, as presented in **Table 2**.

Table 2. AMOVA percentage of variation results within and between each of the tested groups: (1) by species, (2) by the best K in the fast structure, (3) by the groups generated in PCoA, (4) by geographical region, and (5) by the presence or absence of the multiple system of sex chromosomes.

Clustering	Species	fastStructure K=4	PCoA	Geographical Region	Presence/absence of multiple sex system
Variation between groups	83.664	73.916	61.593	48.534	46.534
Variation between samples within group	2.016	12.465	25.087	38.960	41.805

Discussion

Our results highlighted that the distribution of the genetic diversity among the analyzed *Pyrrhulina* species is consistent with their geographical distribution and chromosomal diversity and contribute to our understanding of their evolutionary history. We demonstrated that allopatry, particularly the long-distance isolation, represents the most relevant driving force of speciation in most cases. However, the occurrence of chromosomal rearrangements and especially those related to sex (leading to the presence of multiple sex chromosome systems) amplifies this differentiation and is a driving force in the formation of reproductive barriers between species.

The results of PCoA indicated a clear clustering of species located in the same geographical region, except for *P. semifasciata*, that is in sympatry with *P. brevis* (**Figure 2**). In fastStructure K=4 result, only *P. australis* and *P. marilynae* were grouped, and with K=3 *P.*

obermulleri and *P. brevis* were further combined (**Figure 3**). In this case, geographically close species showed genetically similarity, indicating speciation mechanisms that are largely dependent of allopatry. AMOVA results suggest that grouping individuals by species maximizes the variation between groups, although the groups generated by fastStructure and PCoA have also shown informative values (**Table 2**).

In summary, these results outline that the effect of allopatry is more significant across long geographical distances. In fact, allopatric speciation is usually considered a major speciation process and may have played a significant role in the Neotropical fish biodiversity (Hoorn et al., 2010; Garzon-Orduña et al., 2014). South America has the greatest biodiversity of freshwater fishes on Earth (Tisseuil et al., 2013; Reis et al., 2016) and much of this richness is located in the western Amazon Basin (Boschman et al., 2021). In the paleobiogeographic history of South America, the gradual uplift of the Andes and the repeated flooding of large areas of the Amazon Basin represented the main factors that influenced the richness pattern observed today (Oberdorff et al., 2019). Species of the genus *Pyrrhulina* are very abundant in small forested streams (Zuanon et al., 2015), thus flooding caused by rainfall can directly impact the diversity of these species. Indeed, the relationships between river reorganizations and biodiversity is a complex issue, as previously isolated habitats and populations can be merged, increasing dispersal and local diversity, and reducing speciation and extinction rates, or they can separate previously connected environments, leading to increased genetic isolation and speciation and extinction rates (Tagliacollo et al., 2015; Albert et al., 2018).

The effect of allopatry can potentially be amplified when coupled with the occurrence of chromosomal rearrangements, as observed in *P. marilynae*, the species with the highly rearranged karyotype ($2n=32$) which showed the lowest diversity values among the analyzed species, together with a high number of private alleles (**Figure 2; Table 3**). It's widely known that chromosomal rearrangements might have a role in speciation (Rieseberg, 2001; Navarro & Barton, 2003; Basset et al., 2008). In fact, chromosomal rearrangements have the potential to limit introgression, thus facilitating the origin and maintenance of reproductive isolation through recombination suppression (Faria & Navarro, 2010; Sichová et al., 2015).

The impact of sex chromosomes in genetic differentiation and speciation

Sex chromosomes have long been thought to be associated with population genetic differentiation and several authors have highlighted their connection with speciation (Coyne & Orr 2004; Payseur et al., 2007; O'Neill & O'Neill, 2017; Beaundry et al., 2019). Here, our results pointed that the emergence of a multiple sex chromosome system of the X_1X_2Y -type in the species *P. semifasciata* largely contributed to the fixation of genetic differences in this species. Even though *P. semifasciata* was collected in sympatry with *P. brevis*, the former showed much higher diversity values (**Table 2**). Additionally, *P. brevis* showed a high *Fis* value (0.422) suggesting inbreeding within this species. *P. semifasciata* also presented the highest allelic richness value among all species (**Table 2**). The PCoA result also evidenced that *P. semifasciata* stayed isolated from all other species, while *P. brevis* clustered with *P. obermulleri*. Therefore, even being sympatric with *P. brevis*, none of the analyses grouped these species, and *P. semifasciata* seems to be more differentiated.

Although the impact of sex chromosomes in genetic differentiation and speciation has been largely studied over the recent years, most of such studies involves simple (XY or ZW) sex chromosome systems and little is known about the role of multiple sex chromosomes in fostering the reproductive isolation and increasing genetic diversity. Chromosomal rearrangements involving sex chromosome and autosome are by far the most prevalent mechanism giving rise to multiple sex chromosomes (Pennell et al., 2018; Sember et al., 2021). Among fishes, the most prevalent multiple sex chromosome system is $\text{♀}X_1X_1X_2X_2/\text{♂}X_1X_2Y$. Such system usually arises through Y-autosome fusions (most common) or by fission of the X chromosome in species with a male heteromorphic (Kitano and Peichel; 2012). There is a lot of evidence for the important role of neo-sex chromosomes in speciation, due to problems associated with meiotic segregation of heterozygous rearrangements for example (McKee et al., 1998; Kitano & Peichel, 2012; Beaundry et al., 2019). Previous works suggest that the of sex chromosomes might promote speciation, as some genes that play a role in reproductive isolation between populations can be accumulated in neo-sexual chromosomes (Kitano & Peichel, 2012; Reinhold, 1998; Lindholm & Breden, 2002). Similarly, Haldane's rule postulates that hybrids of closely related species of heterogametic sex are often sterile, thus hybrid incompatibility also may act as a reproductive

isolating barrier (Haldane, 1922; Coyne & Orr, 2004). In both scenarios, the neo-sex chromosomes system may lead to differentiation genetics and speciation.

Therefore, within the chromosomal rearrangement events involving sex chromosomes associated with reproductive isolation, the emergence of sex neo-chromosomes is one of the most expressive, due to its rapid evolution and fundamental action in fixing genetic variation in different species (Beaundry et al., 2019). In the present work, we can observe their impact on the genetic diversity of *Pyrrhulina* species. The genetic structure analyzes indicated great differentiation in *P. semifasciata*, suggesting that the presence of the X_1X_2Y sex chromosome system, exclusively present in individuals of this species, caused high genetic differentiation and showing that such effect was significant to increase genetic differentiation sympatry. In summary, the results suggests that the distribution of genetic diversity among the analyzed species is mainly influenced long-distance allopatry and the presence of a neo-sex chromosome system in *P. semifasciata* largely contributed to the fixation of genetic differences which may lead to reproductive isolation and a consequent speciation process within the genus.

5. CONSIDERAÇÕES FINAIS

Os resultados obtidos através da investigação aqui apresentada destacam que a distribuição da diversidade genética entre as espécies de *Pyrrhulina* analisadas é consistente com sua distribuição geográfica e diversidade cromossômica, contribuindo consideravelmente para os estudos acerca da história evolutiva do gênero. Oferecendo evidências que estabelecem a alopatria, particularmente o isolamento de longa distância, como a força motriz mais relevante da especiação na maioria dos casos. Apesar disso, a ocorrência dos cromossomos neo X_1X_2Y , exclusivamente presente em *P. semifasciata*, reforçou uma elevada diferenciação genética, mesmo em simpatria. Dessa forma, a presença do sistema cromossômico sexual desempenhou um papel importante na fixação da diferenciação genética em *Pyrrhulina*, sendo um fator de especiação particular por promover o isolamento reprodutivo.

6. REFERÊNCIAS

- AKBARI, Mona et al. Diversity arrays technology (DART) for high-throughput profiling of the hexaploid wheat genome. **Theoretical and Applied Genetics**, v. 113, n. 8, p. 1409–1420, 3 Nov 2006.
- ALBERT, James S. et al. Upland and lowland fishes: a test of the river capture hypothesis. **Mountains, climate and biodiversity**, p. 273-294, 2018.
- ALBERT, James S.; REIS, Roberto (Ed.). Historical biogeography of Neotropical freshwater fishes. **Univ of California Press**, 2011.
- ALFARO, Michael E. et al. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. **Proceedings of the National Academy of Sciences**, v. 106, n. 32, p. 13410-13414, 2009.
- ALKIMIM, Emily Ruas et al. High-throughput targeted genotyping using next-generation sequencing applied in Coffea canephora breeding. **Euphytica**, v. 214, n. 3, p. 1-18, 2018.
- Basset P, Yannic G, Hausser J (2008) Chromosomal rearrangements and genetic structure at different evolutionary levels of the Sorex araneus group. **Journal of Evolutionary Biology**, 21,842–852.
- BATESON W. Heredity and variation in modern lights. In: SEWARD, A.C. (ed.), Darwin and Modern Science. **Cambridge University Press**, pp. 85-101, 1909.
- BEAUDRY, Felix EG; BARRETT, Spencer CH; WRIGHT, Stephen I. Ancestral and neo-sex chromosomes contribute to population divergence in a dioecious plant. **Evolution**, v. 74, n. 2, p. 256-269, 2019.
- CAVIN, Lionel. Freshwater Fishes: 250 million years of evolutionary history. **ISTE Press-Elsevier**, 2017.
- COYNE, Jerry A. et al. Speciation. **Sunderland, MA: Sinauer Associates**, 2004.
- DE MENEZES CAVALCANTE SASSI, Francisco et al. Chromosomal evolution and evolutionary relationships of Lebiasina species (Characiformes, Lebiasinidae). **International journal of molecular sciences**, v. 20, n. 12, p. 2944, 2019.
- 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. **Front. Genet. 12: 769984. doi: 10.3389/fgene. 2021.769984. Frontiers in Genetics| www.frontiersin.org**, v. 12, 2021.
- DE MORAES, Renata Luiza Rosa et al. Comparative cytogenetics and neo-Y formation in small-sized fish species of the genus Pyrrhulina (Characiformes, Lebiasinidae). **Frontiers in genetics**, v. 10, p. 678, 2019.
- DE MORAES, Renata Luiza Rosa et al. Evolutionary relationships and cytotaxonomy considerations in the genus Pyrrhulina (Characiformes, Lebiasinidae). **Zebrafish**, v. 14, n. 6, p. 536-546, 2017.

DOBZHANSKY T. Geographical variation in lady-beetles. **Am Nat**, 67 (709), p. 97-126, 1933.

Eaton, D.A.R.; Overcast, I. ipyrad v. 0.7.28 2017. **Disponível online: <https://github.com/dereneaton/ipyrad>**.

FOLL, Matthieu e GAGGIOTTI, Oscar. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. **Genetics**, v. 180, n. 2, p. 977–93, 1 Out 2008.

FRICKE, R e ESCHMEYER, W e VAN DER LAAN, R. **Eschmeyer's Catalog of Fishes**. 2021. v. 3882.

FROESE R, PAULY D. FishBase: World Wide Web electronic publication. **www.fishbase.org**, 2014.

GARZÓN-ORDUÑA, Ivonne J.; BENETTI-LONGHINI, Jennifer E.; BROWER, Andrew VZ. **Timing the diversification of the Amazonian biota: butterfly divergences are consistent with Pleistocene refugia**. 2014.

GRUBER, B. GEORGES, A. BERRY, O. UNMACK, P. dartR: Importing and Analysing SNP and Silicodart Data Generated by Genome-Wide Restriction Fragment Analysis. R package version 1.0.5.; **The R Foundation for Statistical Computing: Vienna, Austria**, 2008.

HALDANE, John BS. Sex ratio and unisexual sterility in hybrid animals. **Journal of genetics**, v. 12, n. 2, p. 101-109, 1922.

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

HINOJOSA, Luis Felipe et al. Non-congruent fossil and phylogenetic evidence on the evolution of climatic niche in the Gondwana genus *Nothofagus*. **Journal of Biogeography**, v. 43, n. 3, p. 555-567, 2016.

HOORN, C. e colab. Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity. **Science**, v. 330, n. 6006, p. 927–931, 12 Nov 2010.

JABLONSKI, D. e ROY, K. e VALENTINE, J. W. Out of the Tropics: Evolutionary Dynamics of the Latitudinal Diversity Gradient. **Science**, v. 314, n. 5796, p. 102–106, 6 Out 2006.

KILIAN, Andrzej et al. Diversity Arrays Technology: A Generic Genome Profiling Technology on Open Platforms. **Humana Press, Totowa, NJ**, 2012. p. 67–89.

KITANO, Jun; PEICHEL, Catherine L. Turnover of sex chromosomes and speciation in fishes. **Environmental biology of fishes**, v. 94, n. 3, p. 549-558, 2012.

KOPELMAN, Naama 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, 1 Set 2015.

LAURIE, Cathy C. The weaker sex is heterogametic: 75 years of Haldane's rule. **Genetics**, v. 147, n. 3, p. 937, 1997.

- LINDHOLM, Anna; BREDEN, Felix. Sex chromosomes and sexual selection in poeciliid fishes. **The american naturalist**, v. 160, n. S6, p. S214-S224, 2002.
- MASLY, John P.; PRESGRAVES, Daven C. High-resolution genome-wide dissection of the two rules of speciation in *Drosophila*. **PLoS biology**, v. 5, n. 9, p. e243, 2007.
- MAYR, Ernst. Geographic speciation in tropical echinoids. **Evolution**, p. 1-18, 1954.
- MCKEE, Bruce D. et al. Male sterility and meiotic drive associated with sex chromosome rearrangements in *Drosophila*: role of XY pairing. **Genetics**, v. 149, n. 1, p. 143-155, 1998.
- MELVILLE, Jane et al. Identifying hybridization and admixture using SNPs: application of the DArTseq platform in phylogeographic research on vertebrates. **Royal Society Open Science**, v. 4, n. 7, p. 161061, 19 Jul 2017.
- MESEGUER, Andrea S. e CONDAMINE, Fabien L. Ancient tropical extinctions contributed to the latitudinal diversity gradient. **bioRxiv**, p. 236646, 9 Abr 2019.
- MULLER H. J. Isolating mechanisms, evolution, and temperature. **Biology Symposium**, 6, p. 71-125, 1942.
- NAVARRO, Arcadi; BARTON, Nick H. Chromosomal speciation and molecular divergence--accelerated evolution in rearranged chromosomes. **Science**, v. 300, n. 5617, p. 321-324, 2003.
- NELSON, Joseph S e GRANDE, Terry e WILSON, Mark V H. **Fishes of the world**. 2016.
- O'NEILL, Michael J.; O'NEILL, Rachel J. Sex chromosome repeats tip the balance towards speciation. **Molecular ecology**, v. 27, n. 19, p. 3783-3798, 2018.
- OBERDORFF, Thierry et al. Unexpected fish diversity gradients in the Amazon basin. **Science advances**, v. 5, n. 9, p. eaav8681, 2019.
- PAYSEUR, Bret A.; PRESGRAVES, Daven C. ; FILATOV, Dmitry A. Introduction: Sex chromosomes and speciation. **Molecular Ecology**, v. 27, n. 19, p. 3745-3748, 2018.
- PENNELL, Matthew W.; MANK, Judith E.; PEICHEL, Catherine L. Transitions in sex determination and sex chromosomes across vertebrate species. **Molecular Ecology**, v. 27, n. 19, p. 3950-3963, 2018.
- PRESGRAVES, Daven C.; ORR, H. Allen. Haldane's rule in taxa lacking a hemizygous X. **Science**, v. 282, n. 5390, p. 952-954, 1998.
- PRITCHARD, Jonathan K.; STEPHENS, Matthew; DONNELLY, Peter. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945-959, 2000.
- RABOSKY, Daniel L. et al. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. **Nature communications**, v. 4, n. 1, p. 1-8, 2013.
- RAJ, Anil e STEPHENS, Matthew e PRITCHARD, Jonathan K. fastSTRUCTURE: variational inference of population structure in large SNP data sets. **Genetics**, v. 197, n. 2, p. 573-89, Jun 2014.
- REINHOLD, Klaus. Sex linkage among genes controlling sexually selected traits. **Behavioral Ecology and Sociobiology**, v. 44, n. 1, p. 1-7, 1998.

- REIS, R. E. et al. Fish biodiversity and conservation in South America. **Journal of fish biology**, v. 89, n. 1, p. 12-47, 2016.
- REIS, Roberto E. e colab. Check list of the freshwater fishes of South and Central America. **EDIPUCRS**, 2003
- RIESEBERG, Loren H. Chromosomal rearrangements and speciation. **Trends in ecology & evolution**, v. 16, n. 7, p. 351-358, 2001.
- SEEHAUSEN, Ole; WAGNER, Catherine E. Speciation in freshwater fishes. **Annual review of ecology, evolution, and systematics**, v. 45, p. 621-651, 2014.
- SEMBER, Alexandr et al. Centric fusions behind the karyotype evolution of neotropical nannostomus pencilfishes (characiforme, lebiasinidae): First insights from a molecular cytogenetic perspective. **Genes**, v. 11, n. 1, p. 91, 2020.
- SEMBER, Alexandr 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, p. 20200098, 2021.
- SHERWIN, William B. et al. Information theory broadens the spectrum of molecular ecology and evolution. **Trends in Ecology & Evolution**, v. 32, n. 12, p. 948-963, 2017.
- ŠÍCHOVÁ, Jindra et al. Dynamic karyotype evolution and unique sex determination systems in Leptidea wood white butterflies. **BMC evolutionary biology**, v. 15, n. 1, p. 1-16, 2015.
- TAGLIACOLLO, Victor A. et al. Biogeographical signature of river capture events in Amazonian lowlands. **Journal of Biogeography**, v. 42, n. 12, p. 2349-2362, 2015.
- TISSEUIL, Clement et al. Global diversity patterns and cross-taxa convergence in freshwater systems. **Journal of Animal Ecology**, v. 82, n. 2, p. 365-376, 2013.
- TOMA, Gustavo Akira 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.
- TURNER, Leslie M.; HARR, Bettina. Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions. **Elife**, v. 3, p. e02504, 2014.
- VAN SOELEN, E. E., KIM, J.-H., SANTOS, R. V., DANTAS, E. L., VASCONCELOS DE ALMEIDA, F., PIRES, J. P., RODDAZ, M., SINNINGHE DAMSTÉ, J. S. A 30 Ma history of the Amazon River inferred from terrigenous sediments and organic matter on the Ceará Rise. **Earth Planet Sc Lett**, 474, p. 40–48, 2017.
- WEITZMAN, M.; WEITZMAN, S. H. Family Lebiasinidae. Check List of the freshwater fishes of South and Central America, Porto Alegre, **Edipucrs**, 729p, p. 241-250, 2003.
- WENZL, Peter 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, 12 Ago 2006.
- WIENS, John J.; DONOGHUE, Michael J. Historical biogeography, ecology and species richness. **Trends in ecology & evolution**, v. 19, n. 12, p. 639-644, 2004.
- WIENS, John J.; GRAHAM, Catherine H. Niche conservatism: integrating evolution, ecology, and conservation biology. **Annu. Rev. Ecol. Evol. Syst.**, v. 36, p. 519-539, 2005.

YANG, Wanjie et al. Review on the development of genotyping methods for assessing farm animal diversity. **Journal of animal science and biotechnology**, v. 4, n. 1, p. 1-6, 2013.

ZUANON, Jansen et al. **Guia de peixes da Reserva Adolpho Ducke**. 2015.