UNIVERSIDADE FEDERAL DE SÃO CARLOS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA MOLECULAR

Renata Luiza Rosa de Moraes

Evolução cromossômica, diversidade genética e história demográfica em espécies do gênero *Pyrrhulina* (Characiformes, Lebiasinidae)

SÃO CARLOS -SP
2024

Renata Luiza Rosa de Moraes

Evolução cromossômica, diversidade genética e história demográfica em espécies do gênero *Pyrrhulina* (Characiformes, Lebiasinidae)

Tese de doutorado apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos – UFSCar, como parte dos requisitos necessários para obtenção do título de Doutora em Ciências (Ciências Biológicas), área de concentração: Genética e Evolução.

Orientador: Prof. Dr. Marcelo de Bello Cioffi

São Carlos-SP
2024

Dedico esse trabalho a todas as pessoas que mudaram minha vida: minha família, amigos, e colegas de trabalho que participaram e me apoiaram intensamente nessa jornada.

A realização deste estudo foi possível devido:

À Universidade Federal de São Carlos e ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular (UFSCar).

Aos laboratórios de Citogenética Evolutiva (UFSCar), Laboratory of Molekulare Zytogenetik da Universitätsklinikum (Uni Jena).

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo financiamento dos projetos: ″**Evolução cromossômica, diversidade genética e história demográfica em espécies do gênero *Pyrrhulina* (Characiformes, Lebiasinidae)**″ (Processo: FAPESP 2019/25045-3)
e
“**Análise do Satelitoma em espécies do gênero *Pyrrhulina* (Characiformes, Lebiasinidae): Insights para evolução cariotípica e dos cromossomos sexuais**” (BEPE – processo: FAPESP 2022/04964-3)

AGRADECIMENTO

Agradeço a todas as pessoas e instituições envolvidas que possibilitaram a realização deste projeto, em especial:

Ao meu orientador Dr. Marcelo de Bello Cioffi por todos os ensinamentos e oportunidades que compartilhou comigo ao longo de todos esses anos. Terei eterna gratidão pelo crescimento profissional que me ofereceu e pelas portas que me abriu. Muito obrigada!

Ao Dr. Orlando Moreira-Filho e Dr. Luiz Antônio Carlos Bertollo por toda sabedoria, respeito e admiração.

Aos funcionários do Departamento de Genética e Evolução, aos professores do Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular da UFSCar e a todos os colegas de pós-graduação.

A Manoela Maria Ferreira Marinho, por todo apoio prestado com relação a taxonomia dos indivíduos e pela identificação dos exemplares.

Aos técnicos de laboratório Luiz Henrique da Silva (Piau) e Antônio Aparecido Donizetti da Silva (Toninho) pelo suporte com as coletas.

Agradeço a todos os alunos e professores do Laboratory of Molekulare Zytogenetik da Universitätsklinikum na cidade de Jena (Alemanha), especialmente ao Dr. Thomas Liehr e Ahmed Al-Rikabi por toda a ajuda, paciência e ensinamentos oferecidos.

A todos que me auxiliaram e que estiveram comigo durante o período em que estive no exterior, especialmente à: Aida e Helena; Caio, Rodrigo, Francisco; Michelly, Luz e Fernando pela amizade, momentos de descontração e por toda parceria no laboratório.

Agradeço a cada um dos colegas que fiz ao longo de todos esses anos no Laboratório de

Citogenética Evolutiva: Ezequiel, Nathalia, Cássia, Geovana, Terumi, Manolo, Carla, Caio, Felipe, Pedro, Priscila, Gustavo, Mariannah, Guilherme, Alan, Nayori, Marina. E agradecimentos especiais aos alunos que admiro e que me apoiaram ao longo desse processo: Francisco, Fernando e Jhon agradeço pela parceria professional e pela amizade cultivada e conquistada ao longo desses anos. E a Geize e Príncia, agradeço por todo carinho, amizade e paciência nesse longo processo, vocês são admiráveis. Foi um prazer trabalhar com vocês, cada um foi fundamental no meu crescimento profissional e pessoal.

Agradeço a todos os meus amigos, particularmente a Ana Carolina pelos nossos momentos de desabafos e torcida pelo meu sucesso. E em especial minha amiga Giovana, que me acompanhou e me incentivou todos os dias, tanto profissionalmente como pessoalmente, e hoje me protege e me guia do céu!

Agradeço incondicionalmente aos meus familiares que me apoiaram todos os dias. Em especial meus pais (Manoela e Sérgio) e minha irmã (Laura) que me auxiliaram, apoiaram, e me guiaram em cada minuto dessa minha jornada! Amo todos vocês!

Quero agradecer a toda a família do meu noivo pela torcida e carinho que foram essenciais. Em particular a minha cunhada Mayara que teve uma contribuição particularmente especial.

Um agradecimento especial ao meu noivo (futuro marido) meu parceiro de vida. Obrigada por todo apoio, incentivo, paciência e dedicação. Amo você!

"...Depois de muito esperar, num dia como outro qualquer, decidi triunfar…
Decidi não esperar as oportunidades e sim, eu mesmo buscá-las.
Decidi ver cada problema como uma oportunidade de encontrar uma solução."

(Walt Disney)

DADOS CURRICULARES DA AUTORA

Renata Luiza Rosa de Moraes iniciou sua graduação em Ciências Biológicas na Universidade Federal de São Carlos, *campus* São Carlos em 2015 e obteve seu título em 2019. Durante a graduação foi bolsista FAPESP de Iniciação Científica no Laboratório de Citogenética Evolutiva (LEC) com supervisão do Prof. Dr. Marcelo de Bello Cioffi durante quatro anos. Nesse período aprendeu e desenvolveu habilidades na área da Citogenética clássica e molecular de peixes com enfoque principal em pequenos peixes da família Lebiasinidae. Em 2017 realizou um estágio de dois meses no Laboratory of Molekulare Zytogenetik da Universitätsklinikum na cidade de Jena (Alemanha) liderado pelo Dr. Thomas Liehr onde efetuou experimentos de microdissecção e pintura cromossômica, além de aprimorar seus conhecimentos citogenéticos e culturais. Ingressou no mestrado pelo Programa de pós-graduação em Genética Evolutiva e Biologia molecular como bolsista CNPq e após dois meses iniciou doutorado direto pelo mesmo programa como bolsista FAPESP sob orientação do Prof. Dr. Marcelo de Bello Cioffi. Em 2023, retornou à cidade de Jena e realizou um estágio de seis meses no Laboratory of Molekulare Zytogenetik da Universitätsklinikum onde fez novos experimentos de microdissecção e pintura cromossômica, além de aprimorar seus conhecimentos com o estudo dos DNAs satélites. Durante o doutorado também participou de diversos projetos de pesquisa além de participar de eventos nacionais e internacionais da área com apresentação de trabalho. Até o momento publicou quinze artigos, sendo cinco deles de primeira autoria.

Lista de artigos com participação da aluna durante o doutorado

Moraes, RLR, Sassi, FMC, Bertollo, LAC, Marinho, MMF, Viana, PF, Feldberg, E, Oliveira, VCS, Deon, GA, Al-Rikabi, ABH, Liehr, T, Cioffi, MB. (2021). Tracking the Evolutionary Trends Among Small-Size Fishes of the Genus *Pyrrhulina* (Characiforme, Lebiasinidae): New Insights From a Molecular Cytogenetic Perspective. **Frontiers in Genetics**, 12, 769984.

Moraes RLR, Sassi FMC, Marinho MMF, Ráb P, Porto JIR, Feldberg E, Cioffi MB (2023). Small Body, Large Chromosomes: Centric Fusions Shaped the Karyotype of the Amazonian Miniature Fish *Nannostomus anduzei* (Characiformes, Lebiasinidae). **Genes,** 14(1), 192.

Moraes, RLR, Sassi, FMC, Vidal, JAD, Goes, CAG, dos Santos, RZ, Stornioli, JHF, Porto-Foresti, F, Liehr, T, Utsunomia, R, Cioffi, MB (2023). Chromosomal Rearrangements and Satellite DNAs: Extensive Chromosome Reshuffling and the Evolution of Neo-Sex Chromosomes in the Genus *Pyrrhulina* (Teleostei; Characiformes). **International Journal Molecular Science**, 24, 13654.

Moraes, RLR, Sassi, FMC, de Souza FMC, Deon GA, Marinho MMF, Cioffi MB. The evolution of Lebiasinidae (Teleostei: Characiformes): Insights from cytogenetics. **Genetic and Molecular Biology** *(Submetido)*

Sassi FMC, Hatanaka T, **Moraes RLR**, Toma GA, Oliveira EA, Liehr T, Rab P, Bertollo LAC, Viana PF, Feldberg E, Nirchio MT, Marinho MMF, Souza JFS, Cioffi MB. (2020). An insight into the chromosomal evolution of Lebiasinidae (Teleostei, Characiformes). **Genes**, 11(4), 365.

Leite PPM, Sassi FMC, Marinho MMF, Nirchio M, **Moraes RLR**, Toma GA, Bertollo LAC, Cioffi MB. (2022). Tracking the evolutionary pathways among Brazilian *Lebiasina* species (Teleostei: Lebiasinidae): a chromosomal and genomic comparative investigation. **Netropical Ichthyology**, 20 (1), e210153.

Ferreira PHN, de Souza FHS, **Moraes RLR**, Perez M, Sassi FMC, Viana PF, Feldberg E, Ezaz T, Liehr T, Bertollo LAC, Cioffi MB. (2022). The influence of chromosomal rearrangements and sex chromosomes in the genetic differentiation of *Pyrrhulina* (Teleostei, Characiformes). **Frontiers in Genetics,** 13, 869073.

Vidal JAD; Sassi FMC; **Moraes RLR**; Artoni RF; Liehr T; Cioffi MB; Almeida MC (2023). Giant Sex Chromosomes in *Omophoita* species (Coleoptera, Chrysomelidae): Structural and Evolutionary Relationships revealed by Zoo-FISH and Comparative Genomic Hybridization (CGH). **Insects**, v. 14, p. 440.

Tura V; Kretschmer R; Sassi FMC; **Moraes RLR**; Barcellos SA; Rosso VO; Souza MS; Cioffi MB; Gunski RJ; Garnero AV (2023). Chromosomal evolution of Suboscines: Karyotype diversity and evolutionary trends in Ovenbirds (Passeriformes, Furnariidae). **Cytogenetic and Genome Research**, v. 1, p. 1.

Khensuwan S and Sassi FMC; **Moraes RLR**; Jantarat S; Seetapan K; Phintong K; Thongnetr W; Kaewsri S; Jumrusthanasan S; Supiwong W; Ráb P; Tanomtong A; Liehr T; Cioffi MB (2023). Chromosomes of Asian Cyprinid Fishes: Genomic Differences in Conserved Karyotypes of ´Poropuntiinae´ (Teleostei, Cyprinidae). **Animals**, v. 13, p. 1415.

Khensuwan S and Sassi FMC; **Moraes RLR**; Ráb P; Liehr T; Supiwong W; Seetapan K; Tanomtong A; Tantisuwichwong N; Arunsang S; Buasriyot P; Tongnunui S; Cioffi MB (2024). Chromosomes of Asian cyprinid fishes: Novel insight into the chromosomal evolution of Labeoninae (Teleostei, Cyprinidae). **PlOS One,** 19(2), e0292689

Molina, WF, Khensuwan, S, **Moraes, RLR**, Sassi, FMC, da Costa, GWWF, Miguel, DZ, Supiwong, W, Jantarat, S, Phintong, K, Seetapan, K, Ditcharoen, S, Tanomtong, A, Liehr, T, Cioffi, MB. (2024). Karyotypic stasis and its implications for extensive hybridization events in corallivores species of butterflyfishes (Chaetodontidae). **Heliyon**, 10(6).

RESUMO

A família Lebiasinidae (Teleostei, Characiformes) é composta por sete gêneros de pequenos peixes de água doce, que podem ser encontrados em riachos por toda América central e do Sul (exceto Chile). Os representantes desta família apresentam uma grande diversidade de formas e colorações, o que os torna muito atraentes no setor aquarista. Embora seja um grupo que se destaca dentro da ordem Characiformes por sua riqueza de espécies, tal família ainda é pouco explorada citogeneticamente principalmente por suas dificuldades de coleta, obtenção cromossômica e taxonomia confusa. Entretanto, diversos estudos citogenéticos e moleculares envolvendo o grupo vem sendo desenvolvidos ao longo dos anos, contribuindo significativamente no início do esclarecimento de diversas incertezas. Sendo assim, o presente estudo faz parte de grandes avanços citogenéticos e genômicos em andamento que buscam elucidar a história evolutiva dos peixes da família Lebiasinidae. Esta tese tem enfoque principal no gênero *Pyrrhulina*, onde foram combinadas análises citogenéticas, moleculares e de diversidade genética com o objetivo de i) analisar a evolução cariotípica do grupo, ii) compreender a influência do sistema sexual na diferenciação de espécies de *Pyrrhulina* e iii) organizar uma filogenia baseada em dados moleculares para a família Lebiasinidae. Foram analisadas cinco novas espécies de *Pyrrhulina* (*Pyrrhulina* aff. *marilynae*, *Pyrrhulina* cf. *laeta*, *Pyrrhulina marilynae*, *Pyrrhulina obermulleri* e *Pyrrhulina* sp.) por técnicas de citogenética clássica e molecular (Giemsa, bandamento C, mapeamento de DNA repetitivo, hibridização genômica comparativa (CGH) e pintura total cromossômica (WCP). Os resultados mostraram um 2n conservado para a maioria das espécies variando entre 40 e 42 cromossomos e um cariótipo majoritariamente acrocêntrico, sendo *P. marilynae* uma exceção, apresentando 2n=32 e a notável presença de quatro pares de cromossomos metacêntricos, cujo status plesiomórfico é discutido. A distribuição de microssatélites não difere significativamente entre as espécies, entretanto o número e posição dos sítios de DNAr sofreram mudanças significativas em cada uma delas. Os experimentos de CGH interespecíficos apresentaram divergências moderadas no conteúdo de DNA repetitivo entre os genomas das espécies e notavelmente, os experimentos de WCP reforçaram a hipótese sobre a origem do sistema sexual múltiplo X_1X_2Y de *Pyrrhulina semifasciata*. A fim de compreender melhor a formação de tal sistema sexual e a redução cariotípica observada em *P. marilynae* também foram feitas análises com sequências de DNA satélites dessas espécies. Os resultados mostraram a possível influencia dessas *sequência*s na redução do cariótipo de *P. marilynae* e reforçou a hipótese sobre a formação do sistema sexual de *P. semifasciata*. Além disso, para concluir os estudos feitos aqui, o primeiro cariótipo de *Copella* foi descrito juntamento com uma filogenia baseada em dados moleculares dos representantes de cada um dos gêneros de Lebiasinidae contribuindo para as discussões feitas sobre as relações filogenéticas da família. Resumidamente podemos concluir que os dados sugerem que as diferenciações cariotípicas aqui estudadas foram impulsionadas por rearranjos estruturais acompanhados por uma alta dinâmica de DNAs repetitivos, principalmente quando diz respeito ao gênero *Pyrrhulina.*

Palavras-Chave: Peixes, DNAs repetitivos, Rearranjos Cromossômicos, Evolução

ABSTRACT

The Lebiasinidae family includes seven genera of tiny freshwater fishes that live in streams throughout Central and South America (except Chile). The family's representatives have a wide range of shapes and colors, making them highly desirable in the aquarium industry. Despite being a group that stands out within the Characiform order for its diversity of species, this family has received little attention, owing to collection challenges, chromosome obtaining, and complex taxonomy. Although they constitute a substantial challenge, multiple cytogenetic and molecular research involving the group have been developed throughout the years, contributing considerably to the beginning of the resolution of various uncertainties. Thus, our study is part of a larger cytogenetic and genomic effort to understand the evolutionary history of the Lebiasinidae family. Continuing the studies, with a primary focus on the genus *Pyrrhulina*, cytogenetic, molecular, and genetic analyses were combined with the goal of analyzing the group's karyotypic evolution, understanding the influence of the sex chromosome system in *Pyrrhulina* species differentiation, and organizing a phylogeny based on molecular data from the Lebiasinidae family. Five new *Pyrrhulina* species (*Pyrrhulina* aff. *marilynae*, *Pyrrhulina* cf. *laeta*, *Pyrrhulina marilynae*, *Pyrrhulina obermulleri* and *Pyrrhulina* sp.) were studied using conventional and molecular cytogenetic techniques, including Giemsa, C-banding, repetitive DNA mapping, comparative genomic hybridization (CGH), and whole chromosome painting (WCP). The results showed a conserved 2n for most species, ranging from 40 to 42 chromosomes and a predominantly acrocentric karyotype, with *P. marilynae* being an exception, with 2n=32 and the unusual presence of four pairs of metacentric chromosomes, the plesiomorphic status of which is questioned. The distribution of microsatellites is similar among species, while the number and position of rDNA sites vary greatly. Interspecific CGH tests revealed moderate divergences in repetitive DNA content between the species' genomes. Surprisingly, the WCP studies supported the theory that *Pyrrhulina semifasciata*'s X_1X_2Y multiple sex chromosome system evolved. To better understand the formation of such a sex chromosome system and the karyotypic reduction observed in *P. marilynae*, analyses were also carried out using satellite DNA sequences from these species. The results showed the possible influence of these sequences on the karyotype reduction of *P. marilynae* and reinforced the hypothesis about the formation of the sex chromosome system of *P. semifasciata*. Furthermore, to conclude the studies conducted here, the first karyotype of *Copella* was described, along with a phylogeny based on molecular data of representatives from each of the genera of Lebiasinidae, which contributed to discussions about the family's internal phylogenetic relationships. In conclusion, the data point to the possibility that structural rearrangements combined with a high dynamic of repetitive DNAs drove the karyotypic differentiations under study, particularly in the case of the genus *Pyrrhulina*.

Key-Words: Fishes, Repetitive DNAs, Chromosomal rearrangements, Evolution

Lista de Ilustrações

INTRODUÇÃO

MATERIAL E MÉTODOS

CAPÍTULO 1

CAPÍTULO 2

CAPÍTULO 3

CAPÍTULO 4

Lista de Tabelas

LISTA DE ABREVIATURAS

2n	Número diploide
DNAr	DNA ribossômico
ITS	Sítio telomérico intersticial
mm	Milímetro
ml	Mililitro
mg/ml	Miligrama por mililitro
ng/ml	Nanograma por mililitro
µl	Microlitro
mM	Milimolar
µM	Micromolar
µm	Micrômetro
m	Cromossomo metacêntrico
sm	Cromossomo submetacêntrico
a	Cromossomo acrocêntrico
PCR	*Polymerase chain reaction*
dNTP	*Deoxynucleotide triphosphates*
dUTP	*Deoxyuridine Triphosphate*
DOP	*Degenerate oligonucleotide-primed*
FISH	*Fluorescence in situ hybridization*
WCP	*Whole chromosome painting*
CGH	*Comparative genomic hybridization*
DNAg	DNA genômico
DAPI	4',6-Diamidino-2-Phenylindole, Dihydrochloride
$MgCl_2$	Cloreto de Magnésio
rpm	Rotação por minuto
M	Molaridade
N	Normalidade
satDNA	DNA satélite

Sumário

1. Introdução

1.1. Peixes: diversidade e importância nos estudos evolutivos e genéticos

Os peixes são um grupo parafilético composto por uma riqueza taxonômica notável que explora uma diversidade de hábitos e habitats. Em geral, a maioria dos peixes se encontra nas regiões tropicais e subtropicais, se destacando principalmente na África tropical, no sudeste asiático e na Bacia do Rio Amazonas (Nelson *et al*., 2016). Embora também seja caracterizada por um clima tropical, a América Central é relativamente pobre em diversidade de peixes, possivelmente em decorrência da sua história geológica (Nelson *et al*., 2016; Tickner *et al*., 2020; Albert *et al*., 2021). Do total de espécies de peixes válidas, mais de 50% habitam ambientes de água doce, sendo assim, uma parcela significativa da diversidade ictiológica está desproporcionalmente distribuída em uma área relativamente pequena quando comparada com a superfície terrestre, já que apenas 1% da água do planeta é doce (Tagliacollo *et al*., 2021).

Notavelmente, a região Neotropical apresenta o maior número de peixes de água doce, com uma alta proporção de espécies endêmicas (Albert *et al*., 2020; Tonella *et al*., 2022). Agrupados em cinco ordens predominantes: Siluriformes (peixes-gato), Characiformes (tetras, piranhas e aliados), Cyprinodontiformes (killifishes, rivulídeos e aliados), Cichliformes (ciclídeos) e Gymnotiformes (peixes elétricos neotropicais) (Van Der Sleen e Albert, 2017; Malabarba e Malabarba, 2020; Dagosta e de Pinna, 2019; Van Der Sleen e Albert, 2021), os peixes dessa região se caracterizam principalmente por sua diversidade e pela abundância de espécies de tamanho diminuto (Castro e Polaz, 2020).

A ordem Characiformes mais especificamente, se destaca por apresentar peixes de baixa tolerância a água salgada e uma notável variação morfológica (Nelson *et al*., 2016; Betancur-R *et al*., 2019), o que os torna particularmente atrativos para estudos evolutivos. Por serem exclusivamente de água doce, os peixes dessa ordem podem apresentar relações evolutivas diretamente ligadas com a fragmentação dos continentes e o desenvolvimento de barreiras naturais durante a dispersão para habitats secundários (Betancur-R *et al*., 2019; Melo *et al*., 2022). Embora atualmente existam análises morfológicas e moleculares mais aprofundadas dentro da ordem Characiformes, ainda sim as relações filogenéticas são incertas em diversos grupos questionando sua monofilia (Arcila *et al*., 2017; Dai *et al*., 2018; Betancur-R *et al*., 2019; Faircloth *et al*. 2020; Melo *et al*., 2022).

Por apresentarem uma das primeiras linhagens a diversificar na filogenia dos vertebrados, os peixes são ótimos modelos para estudos genômicos e evolutivos (Shao *et al*.,

2019; Sember *et al*., 2021), colaborando no esclarecimento da filogenia dos vertebrados como um todo, já que sua grande diversidade possivelmente ocorra por conta de mudanças genômicas rápidas, os diferenciando dos outros grupos (Ravi e Venkatesh, 2018). Uma vez que os peixes são um grupo muito abundante, essas grandes diversidades podem dificultar não só os arranjos taxonômicos desses animais, mas também uma compreensão mais aprofundada de suas relações evolutivas, fazendo com que sejam em boa parte inexplorados (Dornburg e Near, 2021; Parey *et al*., 2023). Sendo assim, os estudos citogenéticos têm contribuído para o esclarecimento das relações filogenéticas dos peixes através de citogenética clássica e molecular, permitindo um avanço no aprofundamento dos estudos evolutivos nesse grupo tão diverso (Cioffi *et al*., 2018).

1.1.1. Família Lebiasinidae, com enfoque no gênero Pyrrhulina

A família Lebiasinidae apresenta 7 gêneros (*Lebiasina*, *Piabusina*, *Derhamia*, *Nannostomus*, *Pyrrhulina*, *Copella*, *Copeina*) e aproximadamente 75 espécies válidas, endêmicas da região Neotropical, amplamente distribuídas nas Américas do Sul e Central (Costa Rica e Panamá), com exceção do Chile (Netto-Ferreira e Marinho, 2013; Fricke *et al*., 2024; Froese e Pauly, 2024). Esse grupo é organizado em duas subfamílias: Lebiasininae e Pyrrhulininae (Netto-Ferreira e Marinho, 2013; Froese e Pauly, 2018; Fricke *et al*., 2024), das quais a segunda representa o clado mais diversificado composto pelos gêneros *Nannostomus, Copeína, Copella e Pyrrhulina* (Netto-Ferreira e Marinho, 2013). De modo geral, a família é representada em maior quantidade por peixes de porte pequeno que variam de 1,5 a 7,0 cm de comprimento e apresentam grandes variações de formatos corporais e coloração, o que os tornam particularmente atrativos para a aquariofilia (Weitzman e Vari, 1988; Weitzman e Weitzman, 2003).

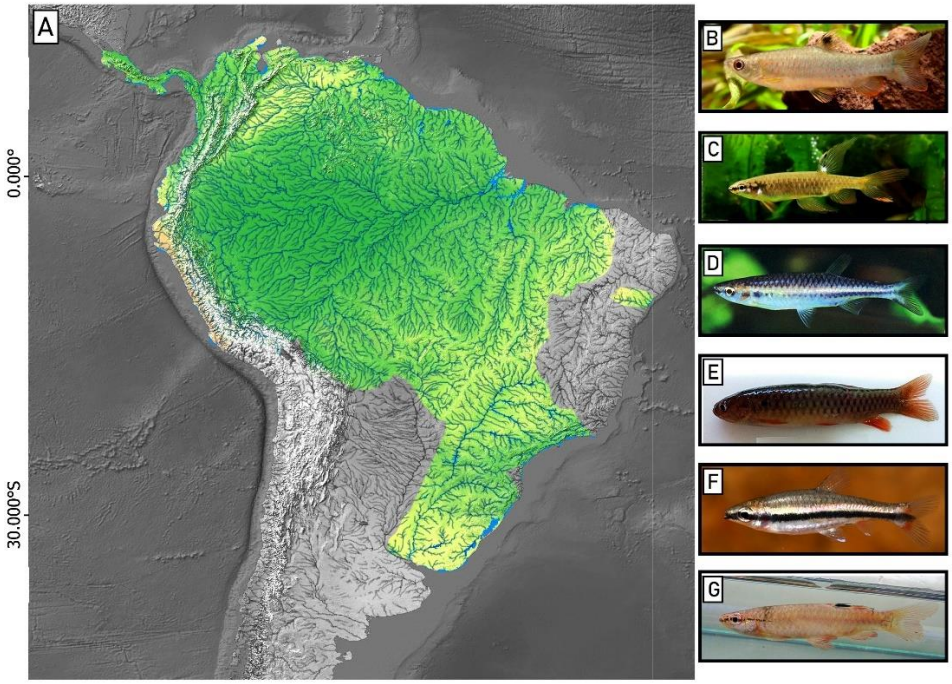


Figura 1: A) Mapa parcial da América do Sul, destacando em cores a distribuição de Lebiasinidae de acordo com GBIF (2024). B-G representantes dos gêneros *Copeina*, *Copella*, *Derhamia*, *Lebiasina*, *Nannostomus e Pyrrhulina*, respectivamente. *Copeina guttata* - https://www.seriouslyfish.com/species/copeina-guttata; *Copella arnoldi* - foto por Zikamoi - Inaturalist - https://www.inaturalist.org/photos/76418963; *Derhamia hoffmannorum* - https://lifecatalog.ru/cont/l/le/Lebiasinidae.html; Lebiasina melanoguttata - foto por Ezequiel A. de Oliveira; *Nannostomus beckfordi* - foto por Chen Hung-Jou - Inaturalist - https://www.inaturalist.org/photos/923588 e *Pyrrhulina obermulleri* - foto por Francisco C.M. Sassi.

Por muito tempo as relações filogenéticas da família Lebiasinidae eram incertas, sendo considerados grupo-irmão de diferentes grupos de peixes Neotropicais ao longo dos anos. Inicialmente, em decorrência de seus caracteres morfológicos a família Lebiasinidae era relacionada com Erythrinidae, Ctenoluciidae e Hepsetidae (Oyakawa, 1998; Buckup, 1998). Posteriormente, através de análises filogenéticas moleculares tal grupo foi considerado grupo irmão de Serrasalmidae, assim como Erythrinidae de Hepsetidae (Orti e Meyer, 1997). Anos mais tarde, diferentes estudos moleculares sugeriram uma relação entre Lebiasinidae e Ctenoluciidae (Calcagnotto *et al*., 2005; Oliveira *et al*., 2011), dados estes que foram corroborados por estudos moleculares posteriores (Arcila *et al*., 2017; Betancur-R *et al*., 2018; Melo *et al*., 2022).

Em particular, o gênero *Pyrrhulina* é o mais especioso da subfamília Pyrrhulininae (junto com *Nannostomus*) sendo composto por 19 espécies válidas (Fricke *et al*., 2024). Esse gênero está distribuído particularmente nas bacias Amazônica e Araguaia-Tocantins, bem como na rede hidrográfica Paraguai-Paraná-La Plata, e nos sistemas Laguna dos Patos e Rio Tramandaí no sul da América do Sul (Weitzman e Weitzman, 2003; Venere e Garutti, 2011; Bertaco *et al*., 2016; Dagosta e de Pinna, 2019). Considerado o mais derivado da família,

Pyrrhulina abriga a maioria das incertezas taxonômicas atuais dentro dos Lebisinidae (Netto-Ferreira e Marinho, 2013), incluindo uma série de inconsistências nas descrições e diagnósticos das espécies deste grupo (Gery, 1977), que são difíceis de distinguir com base em suas características morfológicas (Souza *et al*., 2023). Por conta da variedade de espécies e suas características individuais, o gênero ainda apresenta sua taxonomia confusa, mesmo depois de análises taxonômicas e moleculares recentes (Weitzman e Weitzman, 2003; Netto-Ferreira e Marinho, 2013; Souza *et al*., 2023).

1.1.2. Estudos citogenéticos em Pyrrhulina

Por conta do tamanho diminuto, a família Lebiasinidae por muito tempo foi pouco explorada citogeneticamente, possivelmente por conta das dificuldades de se obter boas preparações cromossômicas. Assim, os primeiros dados sobre a família se restringiam basicamente em referências ao número cromossômico haploide e diploide em algumas poucas espécies (Scheel, 1973; Arai, 2011), ou a descrição do cariótipo de uma única espécie da família, *Pyrrhulina* cf. *australis* (Oliveira *et al*., 1991). Anos mais tarde, foram analisadas por técnicas de citogenética mais avançadas duas espécies do gênero *Pyrrhulina* (*Pyrrhulina* aff. *australis* e *P. australis*) que deram um salto aos estudos citogenéticos da família. Os resultados de tais análises evidenciaram 2n=40 (4st+36a) sem a presença de cromossomos sexuais morfologicamente diferenciados em ambas as espécies (Moraes *et al*., 2017). Um acúmulo de heterocromatina C-positiva foi observado principalmente nas regiões centroméricas e teloméricas da maioria dos cromossomos de ambas as espécies, além disso foram evidenciadas bandas intersticiais presentes em *Pyrrhulina* aff. *australis* e ausentes em *P. australis* (Moraes *et al*., 2017). Os mapeamentos de DNAs repetitivos, hibridização genômica comparativa (CGH) e pintura total cromossômica (WCP) determinaram o número, a posição relativa dos DNAr 18S e 5S, e diversas particularidades e divergências entre as espécies (Moraes *et al*., 2017). Tais experimentos mostraram ainda que embora essas espécies apresentem diferenças, ambas compartilham muitas características, corroborando sua proximidade taxonômica (Moraes *et al*., 2017). Apesar das especificidades inerentes a cada grupo, algumas correlações quanto à estrutura cariotípica e distribuição de classes de DNAs repetitivos foram evidenciadas entre *Pyrrhulina* e *Erythrinus erythrinus* (Erythrinidae) sugerindo na época uma proximidade entre os grupos (Moraes *et al*., 2017).

Embora fosse um grupo desafiante em quesitos citogenéticos, os estudos dentro do gênero continuaram e contribuíram significativamente para a compreensão da evolução cariotípica do grupo. Mais duas espécies (*Pyrrhulina brevis* e *Pyrrhulina semifasciata*) foram

analisadas por diferentes técnicas citogenéticas e moleculares nos anos seguintes, trazendo como principais resultados um cenário que condiz com o anteriormente observado. *P. brevis* evidenciou um 2n=42 para ambos os sexos sem diferenciação sexual. Assim como observado em *Pyrrhulina* aff. *australis*, *P. brevis* apresentou acúmulo de heterocromatina C-positiva não apenas nas regiões pericentroméricas como também nas regiões intersticiais o que a diferencia das outras espécies até então estudadas (Moraes *et al.*, 2019). *P. semifasciata* por sua vez apresentou um cariótipo com 2n=42 para fêmeas e 2n=41 para machos com a presença de três cromossomos não pareados (1m+2a) sugerindo a formação de um sistema sexual múltiplo X1X2Y, originado por fusão cêntrica, que se destaca por ser o único sistema sexual heteromórfico evidente dentro da família (Moraes *et al*., 2019).

Para compreender melhor as relações entre as espécies estudadas e a possível origem do sistema sexual evidenciado em *P. semifasciata*, foram utilizadas ferramentas robustas como hibridização genômica comparativa (CGH) e pintura total cromossômicas (WCP). Assim, a presença do sistema sexual múltiplo observado em *P. semifasciata* foi confirmada por experimentos de WCP, os quais evidenciaram três cromossomos (1m+2a) não pareados e totalmente marcado nos machos da espécie e dois pares de cromossomos totalmente marcados tanto nas fêmeas da espécie como em outras espécies do gênero sugerindo que formação do cromossomo neo-Y desse sistema fosse a fusão de dois cromossomos acrocêntricos não homólogos (Moraes *et al*., 2019). Além disso, com os experimentos intraespecíficos de CGH foi possível observar sequências sexo especificas localizadas principalmente no cromossomo metacêntrico de *P. semifasciata* macho (cromossomo neo-Y) dado esse que corrobora com a presença de um sistema sexual múltiplo X_1X_2Y na espécie. Já os experimentos interespecíficos feitos entre *P. brevis* e *P. semifasciata* produziram uma gama de sinais espécie específicos não sobrepostos, sendo essa uma consequência de suas histórias evolutivas especificas (Moraes *et al*., 2019).

Portanto, em um contexto geral foi possível observar que embora as espécies de *Pyrrhulina* compartilhem uma gama de características cariotípicas e genômicas como: i) um 2n conservado variando entre 40-42 cromossomos; ii) acúmulo de heterocromatina C-positiva principalmente nas regiões pericentroméricas/ centroméricas e teloméricas; iii) múltiplos sítios de DNAr 18S e 5S; iiii) cariótipos marcados majoritariamente por cromossomos acrocêntricos (Moraes *et al*., 2017; Moraes *et al*., 2019), tais espécies também evidenciam características particulares que as diferenciam entre si, identificando diferentes processos evolutivos entre elas (Moraes *et al*., 2017; Moraes *et al*., 2019)

1.2. Marcadores citogenéticos

Os marcadores cromossômicos têm se mostrado como ferramentas cruciais na caracterização da biodiversidade, principalmente de peixes (Moreira-Filho e Bertollo, 1991; Bertollo *et al*., 2000; Cioffi *et al*., 2012). Os primeiros bandamentos cromossômicos permitiram a comparação de bandas para o pareamento correto dos cromossomos homólogos (Guerra, 1988), entretanto a citogenética molecular sofreu grandes aperfeiçoamentos ao longo dos anos permitindo, em conjunto com a citogenética clássica, análises evolutivas mais robusta. A seguir foram selecionadas algumas abordagens que vêm sendo utilizadas em análises evolutivas de peixes.

Descrito pela primeira vez por Sumner (1972) o bandamento C se originou com o objetivo de identificar regiões ricas em heterocromatina constitutiva. Tais regiões são comumente encontradas nas regiões pericentroméricas e teloméricas dos cromossomos e se caracterizam por serem compostas por diversas sequências de DNA repetitivo (Charlesworth, 1994; Plohl *et al*., 2008; López-Flores e Garrido-Ramos, 2012). Em espécies de peixes a distribuição de heterocromatina constitutiva pode auxiliar na diferenciação de espécies, na identificação de cromossomos acessórios ou até mesmo na evidência de sistemas cromossômicos sexuais (Galetti Jr. *et al*, 1991; Margarido e Galleti Jr., 1999; Margarido e Galetti Jr., 2000; Bertollo *et al*., 1997; Rocha-Reis *et al*., 2018).

Proposta por Howell e Black (1980) a detecção das regiões organizadoras de nucléolos (NORs) também é consideradas uma ferramenta primordial. Tal técnica consiste na impregnação de Nitrato de Prata nos cromossomos (Ag-86 NOR) com o objetivo de identificar famílias de DNA ribossomal (DNAr), dado que as NORs são compostas por DNAr 45S e proteínas acessórios. É digno de nota que os DNAr dão origem aos ribossomos e participam ativamente na síntese proteica. De modo geral são divididos em duas famílias multigênicas: DNAr 45S e DNAr 5S. No segmento DNAr 45S estão codificados os DNAs ribossomais 28S, 5,8S e 18S que se encontram separados por dois espaçadores internos transcritos (ITS1 e ITS2), dois espaçadores externos transcritos (ETS1 e ETS2) e apenas um espaçador não transcrito (NTS). Por sua vez, DNAr 5S é organizado por uma série de repetições, separadas por NTS ou também podem ser encontrados na forma de pseudogenes dispersos nos cromossomos (Lafontaine e Tollervey, 2001).

Entretanto, com o avanço da citogenética molecular foi possível investigar com mais propriedade a real posição dos genes ribossomais, assim como o seu número de sítios. A partir do surgimento das técnicas de hibridização fluorescente *in situ* (FISH) e do aperfeiçoamento das

sondas de DNAr, o mapeamento desses genes se revelou uma das principais opções na caracterização de biodiversidade e estudos de sistemática (Sochorová *et al*., 2018). É importante ressaltar que apesar da sua alta conservação, as sequências ribossomais tendem a mudar em número de sítios e sua posição nos cromossomos o que as torna *sequência*s extremamente atraentes em estudos evolutivos (Roy *et al*., 2005; McTaggart *et al*., 2007; Wang e Lemos, 2017; Sochorová *et al*., 2018; Glugoski *et al*., 2020; Sassi *et al*., 2021).

Sequências como microssatélites também tem se demostrado ótimas ferramentas em estudos evolutivos. Tais *sequência*s podem ser encontradas em todos os eucariotos (López- Flores e Garrido Ramos, 2012; Cioffi e Bertollo, 2012) e se caracterizam por serem pequenas repetições em tandem que podem apresentar até seis nucleotídeos. Interessantemente, essas unidades podem se agrupar formando longas sequências que podem se associar a diferentes regiões cromossômicas incluindo regiões heterocromáticas (Martins, 2007; Cioffi *et al*., 2011; Haerter *et al*., 2023). Especificamente em peixes, é muito comum observarmos essas *sequência*s principalmente em regiões centroméricas e teloméricas, e até mesmo em cromossomos sexuais (Cioffi e Bertollo, 2012; Schemberger *et al*., 2019; Balini *et al*., 2024).

Por fim, mas não menos importante temos as *sequência*s teloméricas, que se destacam principalmente por conter importantes informações sobre a evolução cromossômica (Merlo *et al*., 2007; Wang *et al.,* 2009). Ricas em guanina (TTAGGG)n, as *sequência*s teloméricas desempenham papeis fundamentais na sobrevivência dos organismos estabilizando os cromossomos e permitindo a replicação completa das suas regiões terminais (Blackburn, 1994). Embora essas sequências se encontrem majoritariamente nas regiões terminais dos cromossomos, em diferentes cenários elas podem ser detectadas em regiões intersticiais (sítios teloméricos intersticiais (ITS) tendo como função principal identificar possíveis processos de fusão cromossômica (Sola *et al*., 2003; Cioffi *et al*., 2010; Blanco *et al*., 2012; Ocalewicz, 2013; Moraes *et al*., 2022).

Dentre as ferramentas utilizadas na citogenética molecular temos a hibridização genômica comparativa (CGH) e a Pintura total cromossômica (WCP). A técnica de CGH consiste na utilização de sondas de genomas inteiros que possibilitam a comparação de regiões compartilhadas entre sexos ou espécies, identificação de ganho ou perda de sequências de DNA entre as espécies analisadas, bem como a variação do número de cópias de uma sequência genômica específica. Entre os peixes tal procedimento vem se destacando cada vez mais, conforme observado na quantidade de trabalhos publicados nos últimos anos utilizando a abordagem (Oliveira e Sember *et al*., 2017; Freitas *et al*., 2017; Moraes *et al*., 2019; Sassi *et al*., 2020),

certamente CGH mostrando-se uma ferramenta valiosa na identificação de homologias entre sistemas de cromossomos sexuais, bem como para estudos evolutivos como um todo.

Incrementando ainda mais os estudos citogenéticos os experimentos de pintura total cromossômica (WCP) surgiram com um enfoque predominantemente clínico (Blough *et al*., 1998), sendo posteriormente utilizados em estudos evolutivos envolvendo diferentes espécies animais (Yang *et al*., 2009). Aqui utilizamos sondas de cromossomos inteiros ou até mesmo de segmentos cromossômicos, que são obtidos através de microdissecção ou citometria de fluxo, seguido pela amplificação por primers degenerados (DOP-PCR) (Yang *et al*., 2009). Os estudos em peixes especificamente se iniciaram em 2010, permitindo estudos principalmente sobre a origem e evolução de sistemas sexuais e autossomos em diferentes espécies (Oliveira e Sember *et al*., 2017; Freitas *et al*., 2017; Barby *et al*., 2019; Sassi *et al*., 2020; Yano *et al*., 2021)

1.2.1 Sequências de DNAs satélite e o estudo do Satelitoma

DNAs satélites (satDNAs) são sequências repetidas em tandem não codificantes, com comprimento de unidade, número de cópias e organização cromossômica variáveis (Tautz, 1993; Richard *et al*., 2008; Garrido-Ramos, 2017). Os satDNAs são uma das *sequência*s mais abundantes da maioria dos genomas, constituindo longas matrizes localizadas principalmente em regiões heterocromáticas centroméricas e teloméricas, além de já terem sido relatadas mais de uma vez em regiões eucromáticas (Kuhn *et al*., 2012; Plohl *et al*., 2012; Brajković *et al*., 2012; Pavlek *et al*., 2015; Garrido-Ramos *et al*., 2017; Šatović-Vukšić e Plohl, 2023).

Com a integração da citogenética e os sequenciamentos de próxima geração (NGS), as coleções de diferentes famílias de satDNAs de diversas espécies foram caracterizadas, fornecendo insights sobre diversas questões evolutivas, como evolução cariotípica, diversidade do genoma e relações filogenéticas (Montiel *et al*., 2021; Ruiz-Ruano *et al*., 2016; Flynn *et al*., 2023; Ferretti *et al*., 2020; Bardella *et al*., 2020; Vozdova *et al*., 2021; Cabral-de-Mello *et al*., 2021; Cabral-de-Mello *et al*., 2023; Gržan *et al*., 2023; Peona *et al*., 2023). Em geral, grupos de espécies relacionadas compartilham diversas famílias de satDNAs que geralmente evoluem de forma independente dentro de cada linhagem, de acordo com a chamada hipótese de biblioteca (Fry e Salser, 1977). Embora por muito tempo tenha sido caracterizado como DNA "lixo", estudos recentes evidenciaram que os satDNAs apresentam importantes papéis na regulação de genes e genomas (revisado em Garrido-Ramos, 2017).

Acredita-se que sequências repetitivas são ótimas ferramentas para estudos da evolução dos cromossomos sexuais, onde é comum observar um acúmulo diferencial significativo dessas sequências entre machos e fêmeas, além de enfatizar a existência de vários satDNAs W/Y-

específicos (Flynn *et al*., 2023; Ferretti *et al*., 2020; Cabral-de-Mello *et al*., 2023). Supõem-se ainda que tais *sequência*s desempenham papéis essenciais na evolução e organização de cromossomos sexuais e na especiação cromossômica (Henikoff *et al*., 2001; O'Neill *et al*., 2004; Plohl *et al*., 2012; Weissensteiner e Suh, 2019; Utsunomia *et al*., 2019; Shatskikh *et al*., 2020; Ferretti *et al*., 2020; Flynn *et al*.,2023), além de induzirem rearranjos cromossômicos, impactando diretamente na evolução do cariótipo (Silva *et al*., 2017; Utsunomia *et al*., 2019; Crepaldi *et al*., 2021). Assim, a ocorrência de fusões e fissões cêntricas podem estar frequentemente associadas a satDNAs que se encontram principalmente em regiões centroméricas e pericentroméricas (Kopecna *et al*., 2014; Vozdova *et al*., 2019).

Embora os satDNAs tenham sido relacionados a diversas funções dentro do genoma, ainda sim seu papel na evolução genômica e cariotípica é incipiente, necessitando de mais investigações (Silva *et al*., 2017; Utsunomia *et al*., 2019; Crepaldi *et al*., 2021).

1.3. Citogenética Evolutiva em Lebiasinidae

A família Lebiasinidae é um grupo desafiador. Por conter representantes de tamanhos muito pequenos, o grupo apresenta grandes dificuldades de ser explorado tanto taxonomicamente como molecularmente. Entretanto, tal cenário vem sofrendo mudanças significativas ao longo dos anos através de análises citogenéticas clássicas e moleculares em diferentes gêneros da família (Moraes *et al*., 2017; Moraes *et al*., 2019; Sassi *et al*., 2019; Toma *et al*., 2019; Sassi *et al*., 2020; Sember *et al*., 2020; Leite *et al*., 2022; Moraes *et al*., 2023a). Até então, as análises mais específicas da família se resumiam com referências ao número cromossômico haploide e diploide em algumas poucas espécies (Scheel, 1973; Arai, 2011), ou com a descrição do cariótipo em uma única espécie da família, *Pyrrhulina* cf. *australis* (Oliveira *et al*., 1991).

As primeiras análises citogenéticas mais aprofundadas em Lebiasinidae se deram durante a Iniciação cientifica da presente aluna, na qual foram analisadas por citogenética clássica e molecular quatro espécies do gênero *Pyrrhulina* (*Pyrrhulina* aff. *australis*, *Pyrrhulina australis*, *Pyrrhulina brevis* e *Pyrrhulina semifasciata*) (Moraes *et al*., 2017; Moraes *et al*., 2019). Posterior a esse cenário, outros trabalhos envolvendo os outros gêneros do grupo sofreram destaque (Sassi *et al*., 2019; Toma *et al*., 2019; Sassi *et al*., 2020; Sember *et al*., 2020; Leite *et al*., 2022), contribuindo impreterivelmente para uma melhor compreensão da evolução cariotípica e filogenética da família.

Mesmo antes com a escassez de dados, já era possível observar uma variação do número diploide entre os Lebiasinideos, variando de 2n=22 em *Nannostomus unifasciatus* a 2n=46 em *Nannostomus trifasciatus* (revisado em Oliveira *et al*., 2007 e Arai, 2011). Com os avanços da citogenética clássica juntamente com a molecular foi possível dar uma nova perspectiva aos desafios enfrentados ao estudar o grupo, sendo possível contribuir com novos dados citogenéticos e moleculares. No total, 13 espécies de Lebiasinidae foram analisadas por estudos citogenéticos mais aprofundados, sendo três espécies de *Lebiasina* (*L. melanoguttata*, *L. bimaculata* e *L. minuta*), uma de *Copeina* (*C. gutatta*), cinco de *Nannostomus* (*N. breckfordi*, *N. eques*, *N. marginatus*, *N. unifasciatus* e *N. anduzei*) e quatro de *Pyrrhulina* (*Pyrrhulina* aff. *australis*, *P. australis*, *P. brevis* e *P. semifasciata*). Tais estudos, além de corroborar com dados já existentes contribuíram com dados inéditos sobre a família, possibilitando uma visão mais robusta sobre as relações filogenéticas do grupo (Moraes *et al*., 2017; Moraes *et al*., 2019; Sassi *et al*., 2019; Toma *et al*., 2019; Sember *et al*., 2020; Leite *et al*., 2022; Moraes *et al*., 2023a).

Dentre os gêneros que compõem a família, *Lebiasina* é considerado o gênero basal, sendo representado por um cariótipo conservado com 2n=36 composto por cromossomos de dois braços (Sassi *et al*., 2019; Leite *et al*., 2022). Por sua vez os outros representantes da família apresentam grandes variações na estrutura cariotípica e no 2n, variando de 22 a 46 cromossomos. *Copeina* por exemplo evidenciou um 2n=42 (2m + 4sm + 36 st-a) com um cariótipo majoritariamente acrocêntrico, reforçando a tendência evolutiva divergente entre as espécies de Lebiasinidae (Toma *et al*., 2019). Já *Nannostomus* e *Pyrrhulina* são os dois grupos mais especiosos da família, apresentando diferentes variações de 2n e estrutura cariotípica. Assim como as outras espécies do grupo, *Nannostomus* e *Pyrrhulina* apresentam em sua maioria cariótipos majoritariamente acrocêntricos ou até mesmo exclusivamente acrocêntricos como é o caso de *N. beckfordi* e *N. eques* (Moraes *et al*., 2017; Moraes *et al*., 2019; Sember *et al*., 2020; Moraes *et al*., 2023a).

Embora de modo geral as espécies apresentem cariótipos relativamente conservados, algumas espécies se destacam por se diferenciarem do grupo. *N. anduzei* e *N. unifasciatus*, por exemplo, são as espécies que apresentam os menores números diploides já documentados para peixes teleósteos (2n=22) (Arai, 2011), com cariótipo composto exclusivamente por grandes cromossomos metacêntricos originados por possíveis fusões cêntricas (Sember *et al*., 2020; Moraes *et al*., 2023a), cenário esse diferente ao observado para as outras espécies do gênero. *P. semifasciata* por sua vez também se destaca entre as espécies do gênero apresentando o único sistema sexual heteromórfico documentado na família até o presente momento (Moraes *et al*., 2019).

Embora os estudos em Lebiasinidae tenham apresentado grandes avanços, ainda sim as relações filogenéticas dentro do grupo se apresentam incertas. Ao longo dos anos duas tendências evolutivas foram sugeridas, sendo i) um cariótipo conservado com 2n=36 composto exclusivamente por cromossomos meta/submetacêntricos representando a condição basal do grupo, como visto em *Lebiasina*; e ii) um cariótipo com 2n mais elevado, rearranjos cromossômicos estruturais e um cariótipo majoritariamente acrocêntrico (Sassi *et al*., 2020). Sendo assim, teríamos a subfamília Lebiasininae composta por *Lebiasina* e *Dehamia*, e Pyrrhulilinae composta por *Copeina*, *Copella*, *Nannostomus* e *Pyrrhulina*. Embora tais tendências tenham sido reafirmadas mais de uma vez em diferentes estudos, e os dados citogenéticos e taxonômicos até então relatados corroborem com essas tendências, estudos moleculares recentes propuseram algumas mudanças nesse cenário, tais como a inclusão de *Nannostomus* dentro de Lebiasininae (Casimiro *et al*., 2023). Mesmo com o conjunto de todos esses dados ainda sim fica evidente a necessidade de maiores estudos envolvendo principalmente as relações internas da família, bem como a inclusão de estudo citogenéticos envolvendo *Dehamia* e *Copella* dados esses ausentes até o presente momento.

Em suma, mesmo com todas as dificuldades e desafios encontrados ao trabalhar com a família Lebiasinidae, recentes estudos vem conseguindo superar boa parte das dificuldades e contribuir significativamente com dados citogenéticos e moleculares que auxiliaram no esclarecimento de diversas dúvidas sobre as relações internas do grupo. Adicionalmente, é evidente que a combinação de mais estudos moleculares, morfológicos e cromossômicos são muito bem-vindos para superar as incertezas ainda existentes no grupo, bem como novas filogenias com maiores números amostrais objetivando o esclarecimento das relações filogenéticas ainda incertas.

2. Objetivos

2.1. Objetivo geral

Avançar com os conhecimentos sobre a família Lebiasinidae, dando enfoque no gênero *Pyrrhulina*, tendo em vista analisar as relações cromossômicas intragenérica e intrafamiliar, possibilitando assim abordagens comparativas importantes do ponto de vista evolutivo.

2.2. Objetivos específicos

1) Caracterizar o cariótipo de machos e fêmeas de cinco espécies de *Pyrrhulina* (*Pyrrhulina* aff. *marilynae*, *Pyrrhulina* cf. *laeta*, *Pyrrhulina marilynae*, *Pyrrhulina obermulleri* e *Pyrrhulina* sp.) quanto ao número e morfologia cromossômica, assim como a distribuição de heterocromatina C-positiva.
2) Investigar a distribuição e número de sítios de sequências de DNA repetitivo através do mapeamento dos DNAs ribossomais 5S e 18S e sequências de microssatélites.
3) Avaliar o compartilhamento de frações do genoma entre todas as espécies de *Pyrrhulina* analisadas citogeneticamente até o momento através da hibridização genômica comparativa (CGH).
4) Identificar por experimentos de pintura total cromossômica (WCP) utilizando sondas do cromossomo neo-Y (cromossomo Y de *Pyrrhulina semifasciata*) os possíveis cromossomos sexuais putativos nas demais espécies de *Pyrrhulina* não portadoras de cromossomos sexuais heteromórficos, visando auxiliar na compreensão da possível origem do sistema sexual múltiplo X_1X_2Y de *P. semifasciata*.
5) Analisar o papel dos rearranjos cromossômicos e a presença do sistema sexual na diferenciação das espécies de *Pyrrhulina*.
6) Compreender a evolução e as tendências evolutivas em Lebiasinidae, através de análises filogenéticas da família baseada em dados moleculares.

3. Materiais e Métodos

3.1. Material

Para a realização desse trabalho foram amostradas nove espécies de peixes do gênero *Pyrrhulina* localizadas principalmente na Bacia Amazônica. Os locais de coleta, número e sexo dos espécimes investigados são apresentados na **Figura 1** e **Tabela 1**. Os animais foram coletados com autorização do órgão ambiental brasileiro ICMBIO/SISBIO (licença n°48628-14) e SISGEN (A96FF09). Todas as espécies foram devidamente identificadas por critérios morfológicos pela Dra. Manoela Maria Ferreira Marinho e os exemplares depositados na coleção de peixes do Museu de Zoologia da Universidade de São Paulo (MZUSP) sob os números de comprovante (119077, 119079, 123073,123080) e da Universidade Federal da Paraíba (UFPR) sob o número de voucher (12079, 12080, 12082, 12083). Todos os experimentos seguiram condutas éticas e anestésicas, e foram aprovados pelo comitê de Ética em Experimentação animal da Universidade Federal de São Carlos (processo número CEUA 1853260315)

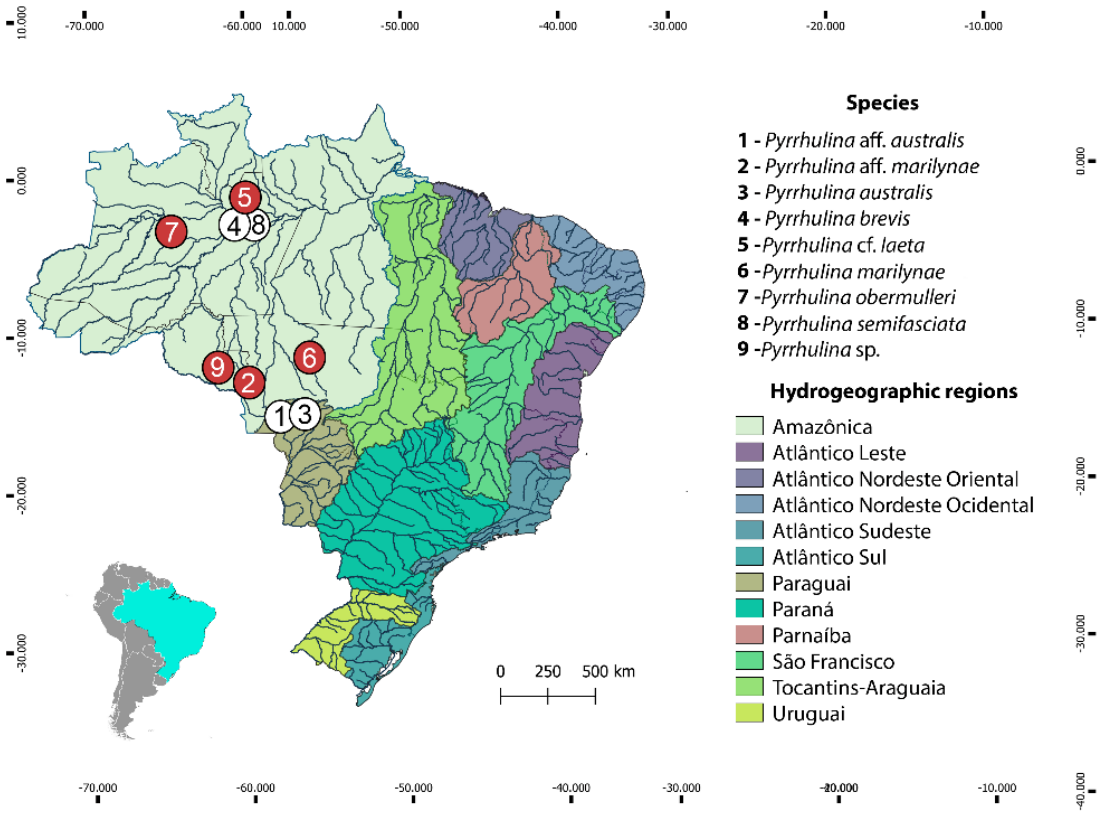


Figura 2: Locais de coleta no Brasil das espécies de *Pyrrhulina* investigadas citogeneticamente em Moraes *et al*. (2021) (círculos vermelhos) e as anteriormente analisadas citogeneticamente (círculos brancos: dados de (Moraes *et al*., 2017; Moraes *et al*., 2019)

Tabela 1 - Coordenadas geográficas e número amostral das espécies de *Pyrrhulina* (Characiformes, Lebiasinidae) coletadas no Brasil

Espécies	Localização	Número Amostral
Pyrrhulina aff. *australis*	Rio Sepotuba, Lambari D'Oeste – MT (15°11'28.0"S 57°41'30.7"W)	16♂ 22♀
Pyrrhulina aff. *marilynae*	Igarapé 12 de Outubro, Comodoro – MT (12°58'41.0"S 60°00'34.0"W)	14♂ 10♀
P. australis	Barra do Bugres – MT (15°04'27.5"S 57°11'05.4"W)	18♂ 30♀
P. brevis	Reserva Florestal Adolpho Ducke, Manaus – AM (2°58'20.7"S 59°55'53.0"W)	17♂ 13♀
Pyrrhulina cf. *laeta*	Presidente Figueiredo – AM (1°59'10.8"S 60°03'40.8"W)	07♂ 05♀
P. marilynae	Ipiranga do Norte – MT (11°36'02.0"S 55°56'27.0"W)	14♂ 08♀
P. obermulleri	Tefé – AM (3°25'50.7"S 64°44'54.8"W)	21♂ 12♀
P. semifasciata	Careiro – AM (3°51'00.0"S 60°04'00.0"W)	12♂ 09♀
Pyrrhulina sp.	Represa, Alto Alegre dos Parecis – RO (12°11'58.0"S 61°46'47.7"W)	19♂ 29♀

3.2. Métodos

3.2.1. Obtenção de cromossomos mitóticos

Os cromossomos mitóticos foram obtidos de acordo com Bertollo *et al*. (2015). Todos os animais aqui coletados foram previamente submetidos a um tratamento com Colchicina (1 ml de solução a 0,005%/100 gramas de peso), e mantidos em aquário aerado por um período de 45 minutos. Após procedimentos anestésicos, foram extraídos fragmentos do rim anterior, posterior e das guelras, que foram transferidos para uma cubeta contendo 10ml de solução hipotônica de cloreto de Potássio (KCl) 0,075M. Com o auxílio de uma seringa desprovida de agulha os tecidos foram dissociados com leves movimentos até se obter uma solução homogênea. Tal suspensão foi mantida em estufa a 37 °C durante 30min. Em seguida, toda a suspensão foi transferida para um tubo de centrifuga com auxílio de uma pipeta Pasteur de vidro, onde foi feito o processo de pré-fixação pingando dez gotas de fixador (proporção 3:1 de álcool metílico e ácido acético) e deixando a temperatura ambiente por 15min. Posteriormente a solução foi colocada em uma centrifuga a 1000 rpm por 10min, foi feito o descarte do sobrenadante e o conteúdo celular foi ressuspendido em 10ml de fixador e centrifugado novamente nas mesmas condições por mais duas vezes. Na última centrifugação, o sobrenadante foi removido e descartado e cerca de 1,5 ml de fixador foi adicionado, na qual

a solução final foi transferida para um microtubo e armazenada a -20°C. Todos os procedimentos seguiram as condutas éticas e de anestesia aprovadas pelo Comitê de Ética em Experimentação e Uso Animal da Universidade Federal de São Carlos (Processo CEUA 7994170423)

3.2.1.1. Preparação das lâminas

Foram pingadas duas gostas de 10ul da suspensão celular sobre lâminas limpas e aquecidas a 55°C. As lâminas foram secas ao ar e em seguida os cromossomos foram corados com Giemsa 5% (pH 6.8) por 8 minutos. O excesso de corante foi removido e as lâminas lavadas em água corrente. Após a análise em microscópio de campo claro e a marcação em lâmina branca das melhores metáfases, o excesso de óleo de imersão foi retirado com álcool 100% e as lâminas descoradas em solução de fixador (proporção 3:1 de álcool metílico e ácido acético) por 15 minutos em temperatura ambiente. Em seguida as lâminas foram lavadas em água corrente e deixadas para secar em temperatura ambiente.

3.2.2. Detecção de heterocromatina constitutiva

Para a detecção de heterocromatina C-positiva foi utilizado o protocolo de Sammer *et al*. (1972). Após os experimentos de FISH, as melhores metáfases foram selecionadas para os experimentos sequência de Bandamento C. Sendo assim, as lâminas contendo preparações cromossômicas foram tratadas em solução de ácido clorídrico (0,2N) a 45°C por 2 minutos e lavadas em água destilada. Ainda úmidas as laminas foram incubadas em solução recém preparada e filtrada de Hidróxido de Bário $Ba(OH)_2$ a 45°C de 50 segundos a 1 minuto dependendo da espécie, e em seguida foram lavadas em solução de HCl 0,2N e em água corrente. Por fim, as lâminas foram incubadas em solução salina (2xSSC) durante 20 minutos em estufa à 60°C. Passado esse período as lâminas foram lavadas com água destilada e deixadas para secar ao ar. Os cromossomos foram corados com 20ul da solução antifading e iodeto de Propídio (50 mg/ml) de acordo com Lui *et al*. (2012) e analisadas em microscópio de fluorescência.

3.2.3. Hibridização fluorescente in situ (FISH) - mapeamento de DNAs repetitivo

Foram utilizadas sondas dos DNAs ribossomais 5S e 18S, isoladas do genoma de *Hoplias malabaricus* (Characiformes, Erythrinidae), clonadas em vetores plasmidiais e propagadas em células competentes de *Escherichia coli* DH5α (Invitrogen, San Diego, CA, USA). A sonda de rDNA 5S incluiu 120 pares de bases (pb) da região codificadora do gene rDNA 5S e

200 pb de espaçador não transcrito (NTS) (Pendás *et al*., 1994). A sonda de rDNA 18S foi composta por um segmento de 1400 pb da região codificadora de rDNA 18S (Cioffi *et al*., 2009). Ambas as sondas foram marcadas diretamente utilizando o *Nick-Translation Mix Kit* (Jena Bioscience, Jena, Alemanha) – rDNA 18S com ATTO488-dUTP e rDNA 5S com ATTO550-dUTP, de acordo com as instruções do fabricante. Além das famílias multigênicas descritas, também foram utilizadas sondas de sequências repetitivas de pequeno tamanho, correspondendo aos microssatélites $(CA)_{15}$, $(GA)_{15}$ e $(CGG)_{10}$. Tais microssatélites foram marcados diretamente com Cy3 durante a síntese, segundo Kubat *et al*. (2008). Além disso, a sequência telomérica (TTAGGG)n também foi utilizada. Tal sonda foi gerada por PCR na ausência de molde de acordo com Ijdo *et al*. (1991) e posteriormente marcada por Nick-Translation Mix Kit (Jena Bioscience, Jena, Alemanha) com ATTO550-dUTP.

Os experimentos de FISH seguiram a metodologia descrita em Yano *et al*. (2017) e Sassi *et al*. (2022). Os cromossomos foram tratados com 100ul de solução de RNAse (1ul de RNAse A + 100ul de 2xSSC) por 1,5 horas incubados em câmara úmida em estufa a 37°C. Após esse tempo, as lâminas foram lavadas sob agitação em PBS 1x por 5 minutos, seguido por um tratamento de solução de Pepsina (1ml de H2O miliq autoclavada + 10µl HCl 1M + 3 µl Pepsina 2%) em câmara úmida por 10 minutos. Em seguida uma nova lavagem de PBS 1x sob agitação por 5 minutos seguida por uma rápida lavagem em serie alcóolica 70%, 85% e 100% por 2 minutos cada. Para a desnaturação dos cromossomos, as lâminas foram submetidas ao tratamento de formamida 70% a 72°C por 3 minutos, sucedida de uma lavagem em série alcóolica como no passo anterior com a alteração para álcool 70% gelado. Por fim o mix de hibridização conteve 2ul de cada sonda desejada completando para 20ul com DS (sulfato de dextrano – 2,5ng/ul, formamida 50% deionizada, 10% de sulfato de dextrano) sendo desnaturado no termociclador a 86°C por 8 minutos + 4°C a 2 minutos. Após esse tempo, foi aplicado 20ul em cada lâmina e tudo foi incubado em câmara úmida a 37°C por no mínimo 16 horas. Passado esse tempo, as lâminas foram lavadas sob agitação em SSC 1x a 65°C por 5 minutos, seguido de uma lavagem de 4xSSC/Tween sob agitação por 5 minutos, uma rápida lavagem sob agitação em PBS 1x, concluindo com uma lavagem em série alcóolica 70%, 85%, 100% por 2 minutos cada. Após estarem completamente secas, os cromossomos foram corados com 20ul de DAPI + Antifading (Vactashield da Vector Laboratories, Burlingame, CA).

3.2.4. Hibridização genômica comparativa (CGH)

Os DNAs genômicos (gDNAs) de machos e fêmeas de *Pyrrhulina* aff. *australis*, *Pyrrhulina* aff. *marilynae*, *Pyrrhulina australis*, *Pyrrhulina brevis*, *Pyrrhulina* cf. *laeta*, *Pyrrhulina marilynae*, *Pyrrhulina obermulleri*, *Pyrrhulina semifasciata* e *Pyrrhulina* sp. foram extraídos a partir do musculo e do fígado, seguindo o método de fenol-clorofórmio-álcool iso-amílico (Sambrook e Russell, 2001). As comparações interespecíficas, os gDNAs derivados de machos e fêmeas de *P. australis* (Paus), *Pyrrhulina* aff. *australis* (Pafa), *P. semifasciata* (Psem), *P. brevis* (Pbre), *P. marilynae* (Pmar), *Pyrrhulina* aff. *marilynae* (Pafm), *Pyrrhulina* sp (Psp), *P. obermulleri* (Pobe) e *Pyrrhulina* cf. *laeta* (Pcfl) foram hibridizados contra cromossomos metafásicos de *P. marilynae*. Tal espécie foi selecionada por apresentar o menor registro de 2n (2n=32) até então identificado no gênero juntamente com uma notável diferenciação cariotípica. Para esse propósito o gDNA derivado de machos de *P. marilynae* foi marcado para todos os ensaios com ATTO550-dUTP, enquanto os gDNAs das outras espécies foram marcados com ATTO488-dUTP (*P. australis*, *Pyrrhulina* aff. *marilynae*, *P. brevis* e *P. obermulleri*) ou ATTO425-dUTP (*Pyrrhulina* aff. *australis*, *Pyrrhulina* sp., *P. semifasciata* e *Pyrrhulina* cf. *laeta*) todos marcados por Nick-Translation (Jena Bioscience, Jena, Alemanha).

As comparações foram divididas em um conjunto de quatro lâminas: 1) Na primeira, a mistura final de sondas foi composta por 500ng de gDNA + 10ug de Cot-1 DNA de macho das seguintes espécies: *P. marilynae*, *P. australis* e *Pyrrhulina* aff. *australis*. 2) Na segunda a mistura final de sondas foi composta por 500ng de gDNA + 10ug de Cot-1 DNA derivado de macho das seguintes espécies: *P. marilynae*, *Pyrrhulina* aff. *marilynae* e *Pyrrhulina* sp. 3) Na terceira a mistura final de sondas foi composta por 500ng de gDNA + 10ug de Cot-1 DNA derivado de macho das seguintes espécies: *P. marilynae*, *P. brevis* e *P. semifasciata*. 4) Finalmente, na quarta lâmina, a mistura final de sondas foi composta por 500ng de gDNA + 10ug de Cot-1 DNA derivado de macho das seguintes espécies: *P. marilynae*, *P. obermulleri* e *Pyrrhulina* cf. *laeta*. Para todos os ensaios a proporção de sonda versus a quantidade de Cot-1 DNA foi baseada em experimentos anteriormente realizados em nossos estudos com peixes (Moraes *et al*., 2019; Toma *et al*., 2019 e Sassi *et al*., 2020) e os experimentos de CGH seguiram a metodologia descrita por Sember *et al*. (2018).

3.2.5. Pintura total cromossômica (WCP)

Como *Pyrrhulina semifasciata* representa a única espécie de *Pyrrhulina* com a presença de um sistema sexual múltiplo X_1X_2Y, uma sonda do cromossomo Y (PSEMI-Y) foi previamente preparada por microdissecção como descrito em Moraes *et al*. (2019). Aproximadamente 20 cópias do cromossomo Y de *P. semifasciata* foram microdissectadas manualmente usando agulhas de vidro, sob um microscópio invertido (Zeiss Axiovert 135). Os cromossomos foram amplificados por PCR, seguindo o protocolo descrito em Yang *et al*. (2009). Em seguida, 1ul do produto de amplificação primária foi usado como DNA modelo para uma marcação secundária de DOP-PCR com Spectrum Orange-dUTP (Vysis, Downers Grove, EUA) em 30 ciclos, seguindo Yang e Graphodatsky (2009).

Metáfases de macho e fêmea de *Pyrrhulina* aff. *marilynae*, *Pyrrhulina* cf. *laeta, P. marilynae*, *Pyrrhulina* sp., *P. obermulleri* foram utilizadas para experimentos Zoo-FISH com a sonda PSEMI-Y, de acordo com os procedimentos descritos em Yano *et al*. (2017) e Sassi *et al*. (2022). A mistura final de sonda foi composta por 500ng da sonda PSEMI-Y e 25ug de Cot-1 DNA isolado do genoma de machos de *P. semifasciata*. As hibridizações foram realizadas em câmara úmida a 37°C por um período de 72h, seguido por duas lavagens sob agitação de 1xSSC por 5 minutos a 65°C, mais duas lavagens sob agitação de 4xSSC/Tween por 5 minutos, uma lavagem rápida em 1xPBS, finalizando com a serie alcóolica 70%, 85% e 100% por dois minutos cada. Os cromossomos foram corados com 20ul de DAPI + Antifading (Vactashield da Vector Laboratories, Burlingame, CA).

3.2.6. Análises de bioinformática

3.2.6.1. Procedimento de sequenciamento DArTseq e filtragem de dados

O tecido hepático de todos os indivíduos foi usado para realizar procedimentos de sequenciamento DArTseq na Diversity Arrays Technology Pty Ltd. Esse método de sequencimento usa uma combinação de enzimas de restrição (Sbfl e Pstl) que enriquecem regiões hipometiladas (Kilian et al., 2012). O sequenciamento das bibliotecas obtidas foi realizado na plataforma Ilumina HiSeq2500.

Os dados brutos foram processados de duas maneiras gerando os seguintes conjuntos (I) uma matriz de SNPs usando Ipyrad v. 0.9.84 (Eaton e Overcast, 2020) e (II) um arquivo de sequência em fases obtido com pyRAD v3.0.66 (Eaton, 2014). Em ambos, os adaptadores de sequenciamento foram aparados, e todas as sequências menores que 35pb ou aparentando mais de cinco bases de baixa qualidade (Q<20) não foram consideradas. No conjunto de dados I, um único SNP por locus foi selecionado, para reduzir a inclusão de SNPs vinculados, e os SNPs

foram codificados como 0 para homozigotos de estado de referência, 1 para heterozigotos e 2 para homozigotos de estado alternativo.

Os dados brutos estão disponíveis no banco de dados do National Center for Biotechnology Information, com o código de acesso PRJNA804560 (link de acesso: https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA804560&o=acc_s%3Aa).

3.2.6.2. Avaliadores de marcação em seleção

Uma análise BayeScan foi realizada para procurar marcadores potenciais sob seleção em ambos os conjuntos de dados (Foll e Gaggiotti, 2008). Realizamos um total de 5.000 execuções MCMC (Markov Chain Monte Carlo) e desbaste de 10, com um valor de probabilidades anterior de 100. Valores de False Discovery Rate (FDR) foram usados para classificar outliers, e apenas loci com valores menores que 0,01 foram considerados outliers e removidos da análise subsequente.

3.2.6.3. Diversidade genética

Utilizando o conjunto de dados II, foi realizada uma análise no software DnaSP v. 6.12.03 (Rozas et al., 2019), para obter a diversidade de haplótipos (Hd), D de Tajima (D) e duas medidas de diversidade de nucleotídeos por sítio, π e teta de Watterson (θw). A diferenciação também foi estimada a partir da divergência de nucleotídeos entre amostras usando o número médio de substituições de nucleotídeos por sítio entre pares de amostras de espécies diferentes (Dxy) e a divergência líquida, corrigida para variação dentro das amostras analisadas (Da).

3.2.6.4. Estrutura genética e análise de variância molecular (AMOVA)

Uma análise de coordenadas principais (PCoA) no pacote R dartR (Gruber et al., 2018) foi usada para investigar a estrutura genética entre as espécies. A estrutura genética também foi investigada com fastSTRUCTURE (Raj et al., 2014), usando o pipeline "Lizards-are-awesome" (Melville et al., 2017). Usamos o comando chooseK.py para selecionar o número de clusters que maximiza a probabilidade e é mais informativo para a estrutura do nosso conjunto de dados. Os resultados do fastStructure foram visualizados com Clumpak (Kopelman et al., 2015). Além disso, uma análise de variância molecular (AMOVA) também foi realizada usando o conjunto de dados I, com amostras agrupadas da seguinte forma: 1) por espécie, 2) pelos clusters do melhor valor K no fastStructure, 3) pelo padrão de agrupamento gerado no PCoA e 4) pela presença ou ausência do sistema de cromossomos sexuais múltiplos.

3.2.6.5. Árvore de espécies

Usamos o pacote SNAPP no BEAST 2.6.4 (Bouckaert et al., 2014) para inferir a topologia da árvore de espécies. Incluímos apenas locais polimórficos, usamos o pior padrão para taxa de coalescência e calculamos taxas de substituição para trás e para frente (parâmetros u e v) com base nos dados empíricos. Aplicamos um comprimento de cadeia de dois milhões de gerações, amostramos a cada 5000 interações. A convergência foi avaliada no Tracer 1.7.1 (Rambaut et al., 2018). Obtivemos a árvore MCC com base em alturas ancestrais comuns no TreeAnnotador, com as primeiras 25% das árvores geradas descartadas com burn-in. Exportamos a árvore de consenso no FigTree 1.4.4.

3.2.6.6. Análise de introgressão

Avaliamos a mistura entre linhagens calculando as estatísticas D de Patterson (teste ABBA-BABA) no Ipyrad v. 0.9.84 (Eaton e Overcast, 2020) usando frequências de SNP agrupadas de indivíduos de cada espécie (Durand et al., 2011). Realizamos testes de 4 táxons de acordo com a topologia SNAPP recuperada, usando *P. marilynae* ou *P. australis* como grupo externo, *P. brevis* como P_3 e *P. semifasciata* e *P. obermulleri* como P_1 ou P_2. A significância foi medida com 1.000 réplicas bootstrap por reamostragem de loci com substituição (seguindo Eaton e Ree, 2013). Os resultados foram considerados significativos de acordo com seus escores Z (valores significativos > 3), que quantificam o número de desvios-padrão do bootstrap nos quais os valores da estatística D se desviam de seu valor esperado de zero.

3.2.7. Sequências DNA satélites

3.2.7.1. Sequenciamento

Para o sequenciamento de DNA, selecionamos um macho *P. marilynae* e um macho e uma fêmea de *P. semifasciata*. O DNA total de tais espécies foi extraído dos tecidos musculares utilizando um protocolo baseado em spin-column (Cellco Biotech, São Carlos – SP, Brasil). Os DNAs purificados foram *sequência*dos pela plataforma BGISEQ-500 (2 x 150 pb; BGI Shenzhen Corporation, Shenzhen, China). O sequenciamento de Short-read rendeu entre 2,14 GB (*P.semifasciata*) e 2,44 GB (*P. marilynae*). Todos os dados brutos foram depositados em um arquivo de leitura de sequências (SRA-NVBI) e estão disponíveis sob os números de acesso (SRR25467276, SRR25476502 e SRR25476501).

3.2.7.2 Análises Bioinformáticas e Biblioteca de satDNA

Inicialmente, foram cortadas as leituras brutas com Trimmomatic (Bolger *et al*., 2014) para selecionar as leituras finais do par com Q> 20 para todos os nucleotídeos. Em seguida, os

catálogos de satDNA de *P. semifasciata* e *P. marilynae* foram caracterizados independentemente no TAREAN (Novák *et al*., 2020) com o pipeline satMiner (Ruiz-Ruano *et al*., 2016). As *sequência*s de consenso de satDNA produzidas no TAREAN foram filtradas das bibliotecas genômicas com DeconSeq (Schmieder e Eswards, 2011) e interações subsequentes foram realizadas no TAREAN até que nenhum satDNA fosse encontrado. Em seguida, foi realizada uma busca de homologia com RepeatMasker (Smith *et al*., 2020) para agrupar as *sequência*s em: Variantes, famílias e superfamílias conforme sugeridas por Ruiz-Ruano (2016). Também foram calculados os valores de abundância e divergência das famílias de satDNA selecionando 10.000.000 leituras (2 x 5.000.000) de cada biblioteca genômica e mascarando seu próprio catálogo de satDNAs com RepeatMasker (Smith *et al*., 2020). Depois disso, foram nomeadas as famílias de satDNA de acordo com sua abundância em cada espécie.

Como *P. semifasciata* apresenta um sistema de cromossomos sexuais múltiplos, foi calculada a razão de abundância machos/fêmeas (M/F) de cada família de satDNA, em seguida foram selecionados aqueles que apresentavam uma relação M/F> 1,2 como satDNAs sexo específicos, acumulados supostamente em machos. Finalmente todos os satélites foram pesquisados pelo BLAST (Altschul *et al*., 1990) contra a coleção de nucleotídeos do NCBI para verificar a presença de satDNAs conservados. Além disso, foi construída uma árvore geradora mínima (MST) para PseSat55 usando PHYLOIZ (Nascimento *et al*., 2017) para descrever as proporções dos haplótipos de machos e fêmeas.

3.2.7.3. Design de Primer e Amplificação de DNA via Reação em Cadeia da Polimerase (PCR)

Para a confecção dos primers, foram projetados primers para 21 dos 70 PseSatDNAs e para 10 dos 71 PmaSatDNAs que foram caracterizados. Como critério de seleção para produção dos primers, foram selecionados os dez mais abundantes para *P. marilynae* (**Tabela 2**), os cinco mais abundantes e aqueles com alguma diferença de abundância entre os sexos para *P. semifasciata* (**Tabela 3**). Para os procedimentos de PCR foram usadas temperaturas e concentrações de DNA ideais para cada DNA satélite, de acordo com Ruiz-Ruano (2016).

Tabela 2 – Dez satDNAs mais abundantes em *P. marilynae.* Destaques em azul são os PmaSatDNAs selecionados para FISH.

Família satDNA	RUL	Abundância	Divergência	A+T (%)
PmaSat01-1627	1627	0.0395879278883	1.33	58.6
PmaSat02-68	68	0.0126290012853	2.63	66.2
PmaSat03-45	45	0.0083896090511	2.92	44.4
PmaSat04-50	50	0.00772308401902	3.93	60
PmaSat05-226	226	0.00728803566553	9.69	64.2
PmaSat06-198	198	0.00647761332963	9.25	67.2
PmaSat07-45	45	0.00639111613516	5.97	62.2
PmaSat08-33	33	0.00268802096213	2.10	57.6
PmaSat09-335	335	0.00220671410467	9.24	70.1
PmaSat10-4663	4663	0.00209498220866	5.93	43.6

Tabela 3 – 21 satDNAs selecionados para confecção de primer de *P.semifasciata*. Destaques em azul são os PseSatDNAs selecionados para FISH e os asteriscos (*) são todos aqueles que hibridizaram em algum dos cromossomos sexuais.

Família SatDNA	RUL	Abundância (M)	Abundância (F)	Abundância (M/F)	Divergência (M)	Divergência (F)	A+T (%)
PseSat01- 304*	304	0,005606905	0,006386826	0,877885927	9,78	9,65	68
PseSat02- 45	45	0,005289071	0,00496321	1,065655359	4,97	4,98	62,2
PseSat03- 68	68	0,005288433	0,005331731	0,991879186	2,52	2,54	66,1
PseSat04-226*	226	0,004975278	0,005546355	0,897035567	9,76	9,62	64,6
PseSat05 - 45	45	0,004215905	0,004087883	1,03131743	4,4	4,43	44,4
PseSat06- 198	198	0,002799141	0,003195114	0,876069106	8,88	8,81	66,6
PseSat27- 422	422	0,000581057	0,000722512	0,804218246	6,35	6,6	0,575829
PseSat32- 186	186	0,000390964	0,000658971	0,593294997	8,44	8,02	0,709677
PseSat34-165	165	0,000282004	0,000328703	0,857928629	6,61	6,58	0,642424
PseSat38- 300*	300	0,000230661	0,000059944	3,847935851	4,96	6,47	0,576667
PseSat39- 80	80	0,000225943	0,000317697	0,711188427	15,55	15,44	0,6625
PseSat40- 39	39	0,000215363	0,000275735	0,781050911	10,74	9,56	0,641026
PseSat50- 1125	1125	0,000160486	0,000123408	1,300450538	9,97	12,7	0,656
PseSat54- 159	159	0,00013754	0,000194867	0,705815943	5,1	4,7	0,672956
PseSat55- 43*	43	0,000133589	0,000159353	0,838325057	8,62	7,99	0,488372
PseSat56- 87	87	0,000126526	0,000159655	0,792494666	6,25	5,81	0,678161
PseSat57- 162	162	0,000124983	8,86007E-05	1,410636489	4,99	5,06	0,67284
PseSat61- 213	213	0,000109063	0,000137659	0,792264963	5,14	5,29	0,615023
PseSat63- 463	463	0,000100169	0,000113188	0,884981918	3,94	3,69	0,591793
PseSat64- 182	182	0,000098656	5,97373E-05	1,651496552	10,95	13,61	0,681319
PseSat67- 198	198	8,07013E-05	0,000103397	0,780502273	4,81	4,26	0,60101

Para cada uma das *sequência*s foi usado o seguinte ciclo: desnaturação inicial a 95°C por 5 min; 29-35 ciclos com desnaturação a 95°C por 15s; anelamento de 50°C a 62°C (**Tabela 4**); extensão a 72°C por 10s e extensão final a 72°C por 10 min. Para validar a amplificação e garantir a integridade dos satDNAs, os produtos de PCR foram analisados por eletroforese em géis de agarose a 1% e 2%. Por fim todos os produtos foram quantificados no

634 espectrofotômetro NanoDrop (ThermoFisher Scientific, Branchburg, NJ, EUA).

635

636 **Tabela 4 -** Condições de PCR (primer, temperatura e concentração de DNA template) para a
637 amplificação ideal de DNAs satélites de *P. marilynae* e *P. semifasciata*
638

Satélite	Primer F	Primer R	Faixa de temperatura de anelamento	[DNA] ng/µl
PmaSat01-1627	5'CACACCTTTGGCATTCTAGC	5'TCTGTTAAGCATGGTGGAGG	56°C – 57,4°C	100
PmaSat04-50	5'GGGTGTGGTTATCTCTGTAC	5'CCCCTCTAAACAGAGTATAAAC	50°C – 51,7°C	100; 10; 1; 0,1
PmaSat05-226	5'GCAAGCTGAATACATTCATG	5'GACTGTCTGAGAGCACAAAC	53°C - 54,5°C	100; 10; 1; 0,1
PmaSat06-198	5'AAGACAGCTTCTGCATCCATG	5'TTGCAGATTTGCCCAAAAAC	55°C - 55,6°C	100; 10; 1; 0,1
PmaSat07-45	5'CCTCTGTAACACATTAAACTGT	5'TAGGAGAGTGTAGTGTTAGTCT	51,7°C – 58°C	100; 10
PmaSat09-335	5'CTGCAACCACTTCCAGTGAT	5'ACACTTAGTTGTGTCTGAAA	53°C - 53,5°C	100; 10; 1; 0,1
PmaSat10-4663	5'AGAACGGAGGTCTCTTGCGT	5'CTCCATTCATTTCCATCACGC	56°C – 62,5°C	100
PseSat01-304	5'CTGCAACCACTTCCAGTGAT	5'ACACTTAGTTGTGTCTGAAA	53°C - 53,5°C	100; 10; 1; 0,1
PseSat04-226	5'GCAAGCTGAATACATTCATG	5'GACTGTCTGAGAGCACAAAC	53°C - 54,5°C	100; 10; 1; 0,1
PseSat06-198	5'AAGACAGCTTCTGCATCCATG	5'TTGCAGATTTGCCCAAAAAC	55°C - 55,6°C	100; 10; 1; 0,1
PseSat32-186	5'TGAGAGAAGACTTTACAAGC	5'TGACACATTTAAGGCATTTC	50°C - 54,9°C	100
PseSat34-165	5'CATGGCTGATAGGGTAAAAG	5'ATGTGACCACACTGTATCCC	56,5°C - 63°C	1; 0,1
PseSat38-300	5'ATAGACGGATGGATTAACGG	5'AAAGACTGACAGACCATTAT	53°C - 61°C	100; 10; 1; 0,1
PseSat39-80	5'CTTACTGTCTCATATCTGTG	5'AGTTAAGACCAGTATCTCTA	50,5°C - 57,4°C	100
PseSat50-1125	5'CTGTACAGTGGATTATGGAG	5'TTCGTGTTACTCAGAATGTC	53 - 53,5	100; 10
PseSat54-159	5'AATCTCTGCATATAAAATGGC	5'AGATGGACCAAAAGGTGTAT	50°C - 58°C	100; 10; 1; 0,1
PseSat55-43	5'CTGTGGTGCACTAACCAGA	5'CCCTATCATTTACAGTACA	50°C - 58°C	100; 10; 1; 0,1
PseSat56-87	5'CACCCAGCCACTTTTTTA	5'TGGAGGTAGTTAGTTAGATG	50°C - 51,5°C	100
PseSat57-162	5'AGGGTCAGTACTCTCACT	5'CACTTCATAACAGTCATTTTAA	52°C - 60°C	100; 10; 1; 0,1
PseSat61-213	5'ATTCCAGGCAATAATCTGCC	5'CAGGCGAGAATTCTACCACT	53°C - 57,9°C	100; 10; 0,1
PseSat63-463	5'ATTGGTCAGATATTGTGAAG	5'ATTGCGCCGTTCATATTCAC	52°C - 60°C	100; 10
PseSat64-182	5'TTTACTGCTGTTAAGATTTC	5'CCTTTGGTAGGACATGTAGC	50°C - 50,7°C	100; 10; 1; 0,1
PseSat67-198	5'CTAACTTCATTCGGCGTTCT	5'CCACCATGGCACACCTGATA	55°C - 55,6°C	10; 1; 0,1

3.2.8. Análises

Aproximadamente 20 metáfases por indivíduo foram analisadas para confirmar o número diploide, as estruturas cariotípicas de cada espécie e os resultados de FISH. As imagens foram capturadas usando microscópio de epifluorescência Olympus BX50 (Olympus Corporation, Ishikawa, Japão) e Axioplan II (Carl Zeiss Jena GmbH, Jena, Alemanha) acoplado a uma câmera CoolSNAP e as imagens processadas no ISIS (MetaSystems Hard e Software GmbH, Altlussheim, Alemanha). Os cromossomos foram classificados em metacêntricos (m), submetacêntricos (sm), subtelocêntricos (st) ou acrocêntricos (a), de acordo com suas proporções de braços (Levan *et al*., 1964). Além disso, a elaboração dos mapas foi realizada no QGIS Desktop 3.6.3, Inkscape 0.92 e Adobe Photoshop CC 2020, sendo esse último também utilizado para as montagens de cariótipos, pranchas, ideogramas e tratamentos de imagem.

4. Resultados e Discussão

Os resultados obtidos no presente estudo e suas respectivas discussões foram compilados e organizados na forma de capítulos, os quais correspondem aos artigos científicos listados abaixo que se encontram publicados ou submetidos para publicação:

CAPÍTULO 1:

Tracking the Evolutionary Trends Among Small-Size Fishes of the Genus *Pyrrhulina* (Characiforme, Lebiasinidae): New Insights From a Molecular Cytogenetic Perspective

CAPÍTULO 2:

The Genetic Differentiation of *Pyrrhulina* (Teleostei, Characiformes) Species is Likely Influenced by Both Geographical Distribution and Chromosomal Rearrangements

CAPÍTULO 3:

Chromosomal Rearrangements and Satellite DNAs: Extensive Chromosome Reshuffling and the Evolution of Neo-Sex Chromosomes in the Genus *Pyrrhulina* (Teleostei; Characiformes)

CAPÍTULO 4:

The evolution of Lebiasinidae (Teleostei: Characiformes): Insights from cytogenetics

Capítulo 1

Tracking the Evolutionary Trends Among Small-Size Fishes of the Genus *Pyrrhulina* (Characiforme, Lebiasinidae): New Insights From a Molecular Cytogenetic Perspective

Renata Luiza Rosa de Moraes, Francisco de Menezes Cavalcante Sassi, Luiz Antônio Carlos Bertollo, Manoela Maria Ferreira Marinho, Patrik Ferreira Viana, Eliana Feldberg4, Vanessa Cristina Sales Oliveira, Geize Aparecida Deon, Ahmed B. H. Al-Rikabi, Thomas Liehr and Marcelo de Bello Cioffi

Frontiers in Genetics (2021) – DOI: 10.3389/fgene.2021.769984

Abstract

Miniature fishes have always been a challenge for cytogenetic studies due to the difficulty in obtaining chromosomal preparations, making them virtually unexplored. An example of this scenario relies on members of the family Lebiasinidae which include miniature to medium-sized, poorly known species, until very recently. The present study is part of undergoing major cytogenetic advances seeking to elucidate the evolutionary history of lebiasinids. Aiming to examine the karyotype diversification more deeply in *Pyrrhulina*, here we combined classical and molecular cytogenetic analyses, including Giemsa staining, C-banding, repetitive DNA mapping, comparative genomic hybridization (CGH), and whole chromosome painting (WCP) to perform the first analyses in five *Pyrrhulina* species (*Pyrrhulina* aff. *marilynae*, *Pyrrhulina* sp., *P. obermulleri*, *P. marilynae* and *Pyrrhulina* cf. *laeta*). The diploid number (2n) ranged from 40 to 42 chromosomes among all analyzed species, but *P. marilynae* is strikingly differentiated by having 2n = 32 chromosomes and a karyotype composed of large meta/submetacentric chromosomes, whose plesiomorphic status is discussed. The distribution of microsatellites does not markedly differ among species, but the number and position of the rDNA sites underwent significant changes among them. Interspecific comparative genome hybridization (CGH) found a moderate divergence in the repetitive DNA content among the species' genomes. Noteworthy, the WCP reinforced our previous hypothesis on the origin of the X1X2Y multiple sex chromosome system in *P. semifasciata*. In summary, our data suggest that the karyotype differentiation in *Pyrrhulina* has been driven by major structural rearrangements, accompanied by high dynamics of repetitive DNAs.

1. Introduction

Characiformes comprise a very diverse and abundant freshwater order (Nelson *et al*., 2016), in which the family Lebiasinidae is represented by 75 valid species (Fricke *et al*., 2021) widely distributed across South and Central America (Weitzman and Weitzman, 2003). The phylogenetic relationships of the Lebiasinidae remained in doubt for a long time, but more recent phylogenetic analysis indicate their proximity to the Ctenoluciidae (Calcagnotto *et al*., 2005; Oliveira *et al*., 2011), which was also reinforced by the different studies (Arcila *et al*., 2017; Betancur-R *et al*., 2019; Melo *et al*., 2021). Most Lebiasinidae species reach about 60 mm of Standard Length (SL), but miniature species, not surpassing a maximum of 26 mm SL, is found within the Pyrrhulininae, whereas medium-sized species up to 150 mm SL can be found within Lebiasininae (Weitzman and Weitzman, 2003).

Because of their small sizes and difficulties in obtaining good chromosomal preparations, species of Lebiasinidae were, for a long time, little analyzed in terms of cytogenetics, with scarce references mainly on the chromosomal number of few species (Scheel, 1973; Oliveira *et al*., 1991; Arai, 2011). However, this scenario has recently undergone significant changes with the methodological advance of cytogenetics and its applicability among small to miniature fishes of *Pyrrhulina*, Lebiasina, Copeina, and Nannostomus genus (Moraes *et al*., 2017, Moraes *et al*., 2019; Sassi *et al*., 2019; Toma *et al*., 2019; Sassi *et al*., 2020; Sember *et al*., 2020).

Pyrrhulina is one of the most speciose genera of the subfamily Pyrrhulininae, with 19 valid small species (Fricke *et al*., 2021), ranging from 30.4 to 85 mm SL (Weitzman and Weitzman, 2003; Netto-Ferreira and Marinho, 2013). The genus is among the most problematic, with many poorly known species, species complexes, and old taxonomic problems (Netto-Ferreira and Marinho, 2013). The first *Pyrrhulina* species to have some chromosomal data evidenced was *Pyrrhulina* cf. *australis*, with 2n = 40 chromosomes, mainly acrocentric ones (Oliveira *et al*., 1991). Taxonomic boundaries of *P. australis* are still poorly defined, demonstrated in subsequent studies (Moraes *et al*., 2017; Moraes *et al*.,2019) of two morphotypes. Both *P. australis* and *Pyrrhulina* aff. *australis* showed similar data 2n = 40 (4st + 36a), distinct from *P. brevis*, 2n = 42 (2sm + 4st + 36a), with no evidence of heteromorphic sex chromosomes in the three species (Moraes *et al*., 2017; Moraes *et al*., 2019). Another species, *P. semifasciata*, was analyzed, presenting 2n = 42 (4st + 38a) in females, and 2n = 41 (1m + 4st + 36a) in males, the latter with three unpaired chromosomes because of a multiple X1X1X2X2/X1X2Y sex

chromosome system (Moraes *et al*., 2019). This occurrence was also confirmed by comparative genomic hybridizations (CGH) and whole-chromosome painting (WCP), with some indications that the Y chromosome originated by centric fusions of non-homologous acrocentric chromosomes (Moraes *et al*., 2019).

To improve the knowledge of the evolutionary processes within the genus *Pyrrhulina*, we combined classical and molecular cytogenetic analyses, including Giemsa staining, C-banding, repetitive DNA mapping, comparative genomic hybridization (CGH), and whole chromosome painting (WCP to perform the first analyses in five *Pyrrhulina* species (*Pyrrhulina* aff. *marilynae, Pyrrhulina* sp., *P. obermulleri, P. marilynae* and *Pyrrhulina* cf. *laeta*). The results highlighted relationships and particular evolutionary paths at the chromosomal and genomic levels among the species. In addition, the hypothesis on the origin of the multiple sex chromosome system in *P. semifasciata* is validated.

2. Materials and Methods

2.1. Animals

The collection sites, number, and sex of the specimens investigated are presented in **Figure 1**, **Table 1**. Part of the sampling (**Figure 1**, white circles) resembles the one previously analyzed by Moraes *et al*. (2017), Moraes *et al*. (2019) with different cytogenetic and molecular methods. Animals were collected with the authorization of the Brazilian environmental agency ICMBIO/SISBIO (license no. 48628-14) and SISGEN (A96FF09). All species were properly identified by morphological criteria, and the specimens were deposited in the fish collection of the Museu de Zoologia da Universidade de São Paulo (MZUSP) under the voucher numbers (119077, 119079, 123073, 123080) and the Universidade Federal da Paraíba (UFPB) museum under the voucher number (12079, 12080, 12082 and 12083). Experiments followed ethical and anesthesia conducts and were approved by the Ethics Committee on Animal Experimentation of the Universidade Federal de São Carlos (process number CEUA 1853260315).

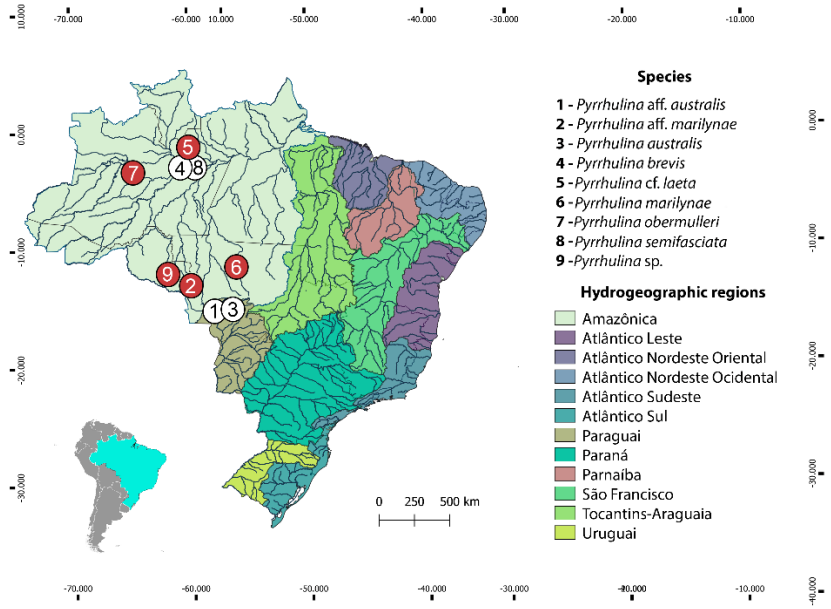


Figure 1: Brazilian collection sites of the *Pyrrhulina* species cytogenetically investigated in the present study (red circles) and the ones previously cytogenetically analyzed (white circles: data from (Moraes *et al*., 2017; Moraes *et al*., 2019).

Table 1. Geographical coordinates and sample size of *Pyrrhulina* (Characiformes, Lebiasinidae) species collected in Brazil.

Species	Locality	Sample Size
Pyrrhulina aff. *australis*	Rio Sepotuba, Lambari D'Oeste – MT (15°11'28.0"S 57°41'30.7"W)	16♂ 22♀
Pyrrhulina aff. *marilynae*	Igarapé 12 de Outubro, Comodoro – MT (12°58'41.0"S 60°00'34.0"W)	14♂ 10♀
P. australis	Barra do Bugres – MT (15°04'27.5"S 57°11'05.4"W)	18♂ 30♀
P. brevis	Reserva Florestal Adolpho Ducke, Manaus – AM (2°58'20.7"S 59°55'53.0"W)	17♂ 13♀
Pyrrhulina cf. *laeta*	Presidente Figueiredo – AM (1°59'10.8"S 60°03'40.8"W)	07♂ 05♀
P. marilynae	Ipiranga do Norte – MT (11°36'02.0"S 55°56'27.0"W)	14♂ 08♀
P. obermulleri	Tefé – AM (3°25'50.7"S 64°44'54.8"W)	21♂ 12♀
P. semifasciata	Careiro – AM (3°51'00.0"S 60°04'00.0"W)	12♂ 09♀
Pyrrhulina sp.	Represa, Alto Alegre dos Parecis – RO (12°11'58.0"S 61°46'47.7"W)	19♂ 29♀

2.2. Chromosomal Preparations and Analysis of the Constitutive Heterochromatin

Mitotic chromosomes were obtained from kidney cells by the protocol described in Bertollo *et al*. (2015). The distribution of constitutive heterochromatin was observed by the C-banding method according to (Sumner, 1972).

2.3. Repetitive DNA Mapping with Fluorescence in situ Hybridization (FISH)

The 5S rDNA probe included 120 base pairs (bp) of the 5S rDNA gene coding region and 200 bp of non-transcribed spacer (NTS) (Pendás *et al*., 1994). The 18S rDNA probe was composed of a 1,400-bp-long segment of the 18S rDNA coding region (Cioffi *et al*., 2009). Both probes were directly labeled with the Nick-Translation Mix Kit (Jena Bioscience, Jena, Germany)—18S rDNA with ATTO488-dUTP and 5S rDNA with ATTO550-dUTP, according to the manufacturer's instructions. The (CA)15, (GA)15, (CGG)10 microsatellite probes were directly labeled with Cy3 during the synthesis, according to Kubat *et al*. (2008). In addition, since it contains the lowest 2n, telomeric (TTAGGG)n sequence was also used as probe in *P. marylinae*. This probe was generated by PCR in the absence of a template according to Ijdo *et al*. (1991) and later labeled with ATTO550-dUTP with the Nick-Translation Mix Kit (Jena Bioscience, Jena, Germany). FISH experiments followed the methodology described in Yano *et al*. (2017). Metaphase chromosomes were treated with RNAse A (40 µg/ml) for 1.5 h at 37°C and the DNA denatured in 70% formamide/2× SSC at 72°C for 3.15 min. A hybridization mixture (2.5 ng/µL probes, 50% deionized formamide, 10% dextran sulfate) was then dropped on the slides, and the hybridization process was performed overnight at 37°C in a moist chamber. The first post-hybridization wash was performed with 1× SSC for 5 min at 65°C, followed by the second one performed with 4xSSC/Tween for 5 min, at room temperature. Chromosomes were then counterstained with DAPI, and the slides were mounted with an antifade solution (Vectashield from Vector Laboratories, Burlingame, CA).

2.4. FISH for Whole Chromosome Painting

As *P. semifasciata* represents the only *Pyrrhulina* species that harbors an X1X2Y multiple sex system, a Y-chromosome probe, named PSEMI-Y, was previously prepared by microdissection, as described in (Moraes *et al*., 2019) Male and female metaphases of *P. marilynae*, *Pyrrhulina* aff. *marilynae*, *Pyrrhulina* sp., *P. obermulleri*, *Pyrrhulina* cf. *laeta* were used for Zoo-FISH experiments with the PSEMI-Y probe, according to procedures described in Yano *et al*. (2017). The hybridization was performed for 72 h at 37°C in a moist chamber, with post-hybridization washes with 1xSSC for 5 min at 65°C, and in 4xSSC/Tween (RT). 10 µg of male-derived C0t-1 DNA from *P. semifasciata* was used as suppressor in each experiment. Chromosomes were stained with DAPI (1.2 µg/ml) and the slides were mounted with an antifade solution, as described above.

2.5. Probes for Comparative Genomic Hybridization

The genomic DNAs (gDNAs) from male and female specimens of *P. marilynae*, *Pyrrhulina* aff. *marilynae*, *Pyrrhulina* sp., *P. obermulleri*, *Pyrrhulina* cf. *laeta*, *P. australis*, *Pyrrhulina* aff. *australis*, *P. brevis*, and *P. semifasciata* were extracted from liver tissue by the standard phenol-chloroform-isoamyl alcohol method (Sambrook and Russell, 2001). For intraspecific comparisons, the male-derived gDNAs of all species were labeled with ATTO550-dUTP and the female gDNAs with ATTO 488-dUTP, by nick translation (Jena Bioscience, Jena, Germany). The repetitive sequences were blocked using unlabeled C0t-1 DNA in all experiments, according to (Zwick *et al*., 1997). The final hybridization mixture for each slide (20 µL) was composed of male- and female-derived gDNAs (500 ng each), plus 25 µg of female-derived C0t-1 DNA from the respective species. The probe was ethanol-precipitated, and the dry pellets were mixed in a hybridization mixture containing 50% formamide, 2× SSC, 10% SDS, 10% dextran sulfate, and Denhardt´s buffer, pH 7.0.

For interspecific comparisons, the gDNA of male specimens of *P. australis* (Paus), *Pyrrhulina* aff. *australis* (Pafa), *P. semifasciata* (Psem), *P. brevis* (Pbre), *P. marilynae* (Pmar), *Pyrrhulina* aff. *marilynae* (Pafm), *Pyrrhulina* sp. (Psp), *P.obermulleri* (Pobe) and *Pyrrhulina* cf. *laeta* (Pcfl) were hybridized against metaphase chromosomes of *P. marilynae*. This species was selected since it harbors the lowest 2n = 32 until now register for the genus, coupled with a remarkable karyotype differentiation. For this purpose, male-derived gDNA of *P. marilynae* was labeled with ATTO 550-dUTP, while the gDNAs of the other species were labeled with ATTO 488-dUTP (P. *australis*, *Pyrrhulina* aff. *marilynae*, *P. brevis* and *P. obermulleri*) or ATTO 425-dUTP (*Pyrrhulina* aff. *australis*, *Pyrrhulina* sp., *P. semifasciata* and *Pyrrhulina* cf. *laeta*), both through nick translation (Jena Bioscience, Jena, Germany).

The interspecific comparisons were divided into a set of four slides. In the first slide, the final probe mixture was composed of 500 ng of male-derived gDNA plus 10 µg of male-derived C0t-1 DNA of each of the following species: *P. marilynae*, *P. australis*, and *Pyrrhulina* aff. *australis*. In the second slide, the final probe mixture was composed of 500 ng of male-derived gDNA plus 10 µg of male-derived C0t-1 DNA of each one of the following species: *P. marilynae*, *Pyrrhulina* aff. *marilynae* and *Pyrrhulina* sp. In the third slide, the final probe mixture was composed of 500 ng of male-derived gDNA plus 10 µg of male-derived C0t-1 DNA of each one of the following species: *P. marilynae*, *P. brevis*, and *P. semifasciata*. Finally, in the fourth slide, the final probe mixture was composed of 500 ng of male-derived gDNA plus and 10 µg of male-derived C0t-1 DNA of each one of the following species: *P. marilynae*, *P. obermulleri*, and *Pyrrhulina* cf. *laeta*. The chosen ratio of probe vs. C0t-1 DNA amount was

based on previous experiments performed in our fish studies (Moraes *et al.*, 2019; Toma *et al*., 2019; Sassi *et al*., 2020). The CGH experiments followed the methodology described in Sember *et al*. (2018).

2.6. Microscopy and Images Processing

To confirm the diploid number, karyotype structure and FISH results inat least 30 metaphase spreads were analyzed per individual. The microscopy images were captured using an Olympus BX50 epifluorescence microscope (Olympus Corporation, Ishikawa, Japan) coupled with a CoolSNAP camera, and the images were processed using Image-Pro Plus 4.1 Software (Media Cybernetics, Silver Spring, MD, United States). Final images were optimized and arranged using Adobe Photoshop, version CC 2020. Chromosomes were classified as metacentric (m), submetacentric (sm), subtelocentric (st), or acrocentric (a), according to their arm ratios (Levan, 1964). As the males and females results showed no differences, only male metaphases were represented in the figures.

3. Results

3.1. Karyotypes and Heterochromatin Distribution

The diploid number ranged from 2n = 40 to 42 among the following four species: *Pyrrhulina* sp. (2n = 40; 2st+38a), *Pyrrhulina* aff. *marilynae* (2 = 40; 40a), *P. obermulleri* (2n = 42; 2m/sm+8st+32a) and *Pyrrhulina* cf. *laeta* (2n = 42; 2m/sm+4st+36a), the two latter also sharing a characteristic small metacentric/submetacentric pair. On the other hand, *P. marilynae* differed by presenting a very distinct karyotype composition (2n = 32; 8m/sm+4st+20a). These results represent the first cytogenetic data for the abovementioned species. The constitutive heterochromatin was distributed at the pericentromeric region of several chromosome pairs in *P. marilynae* and *Pyrrhulina* aff. *marilynae*. In its turn, *Pyrrhulina* sp., *P. obermulleri*, and *Pyrrhulina* cf. *laeta* presented a remarkable series of interstitial and pericentromeric C-bands, in addition to telomeric ones (**Figure 2**). In our sampling, we did not observe any karyotype differences between males and females.

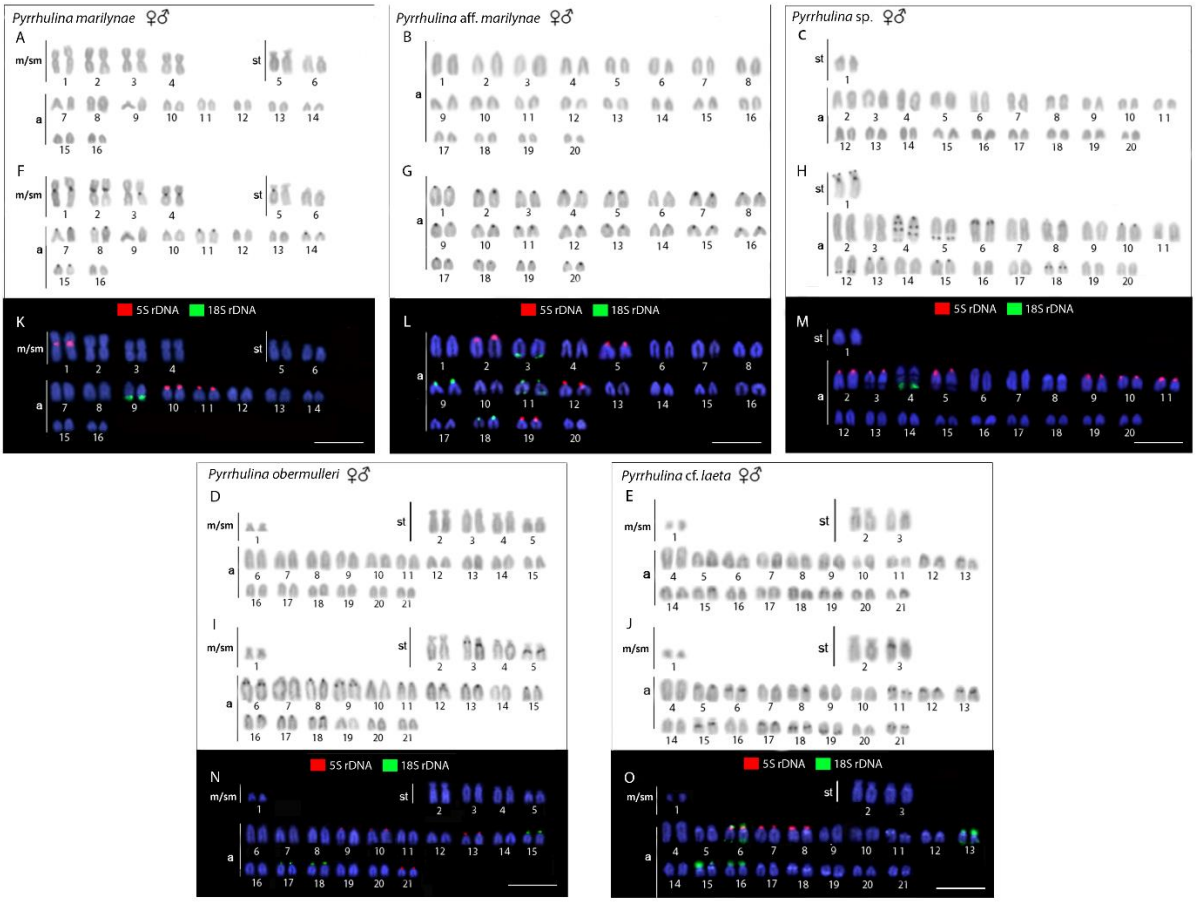


Figure 2. Male and female karyotypes of *Pyrrhulina marilynae* **(A, F, and K)**, *Pyrrhulina* aff. *marilynae* **(B, G and L)**, *Pyrrhulina* sp. **(C, H and M)**, *P. obermulleri* **(D, I and N)** and *Pyrrhulina* cf. *laeta* **(E, J and O)** arranged after Giemsa staining **(A-E)**, C-banding **(F-J)**, and dual-color *in situ* hybridization (FISH) with 18S (green) and 5S (red) ribosomal DNA probes **(K-O)**. Chromosomes were counterstained with 4′,6-diamidino-2-phenylindole (DAPI). Scale bar = 5 μm.

3.2. Chromosomal Mapping of Repetitive DNA Sequences

All the five species differ by the distribution of the multigene rDNA families. *Pyrrhulina* sp. and *P. marilynae* were the only species with only one chromosome pair bearing 18S rDNA sites, found at the telomeric region of acrocentric pairs 4 and 9, respectively. Six to twelve centromeric or telomeric sites occur in the other three species, including bitelomeric sites in *Pyrrhulina* aff. *marilynae* (pair 11) and *Pyrrhulina* cf. *laeta* (pairs 6 and 13). As for the 5S rDNA, from six to twelve centromeric sites occured among species, including a syntenic condition for the 5S and 18S rDNA repeats in the chromosome pair 6 of *Pyrrhulina* cf. *laeta*, the same pair that displays bitelomeric 18S rDNA signals in this species (**Figure 2**). The distribution of the microsatellites (CA)15, (GA)15, and (CGG)10 does not differ significantly among species, having a preferential location in the centromeric and telomeric regions of the chromosomes, in addition to some interstitial sites. However, (CA)15 differs quantitatively, with a greater number of conspicuous sites compared to the other microsatellites, especially in *Pyrrhulina* aff.

marilynae and *Pyrrhulina* cf. *laeta*. In the same way, (CGG)10 occurs in smaller amounts in the five species (**Figure 3**). The (TTAGGG)n repeats showed the expected hybridization signals on telomeres of *P. marylinae* (**Figure 4F**). Whole chromosome painting–WCP.

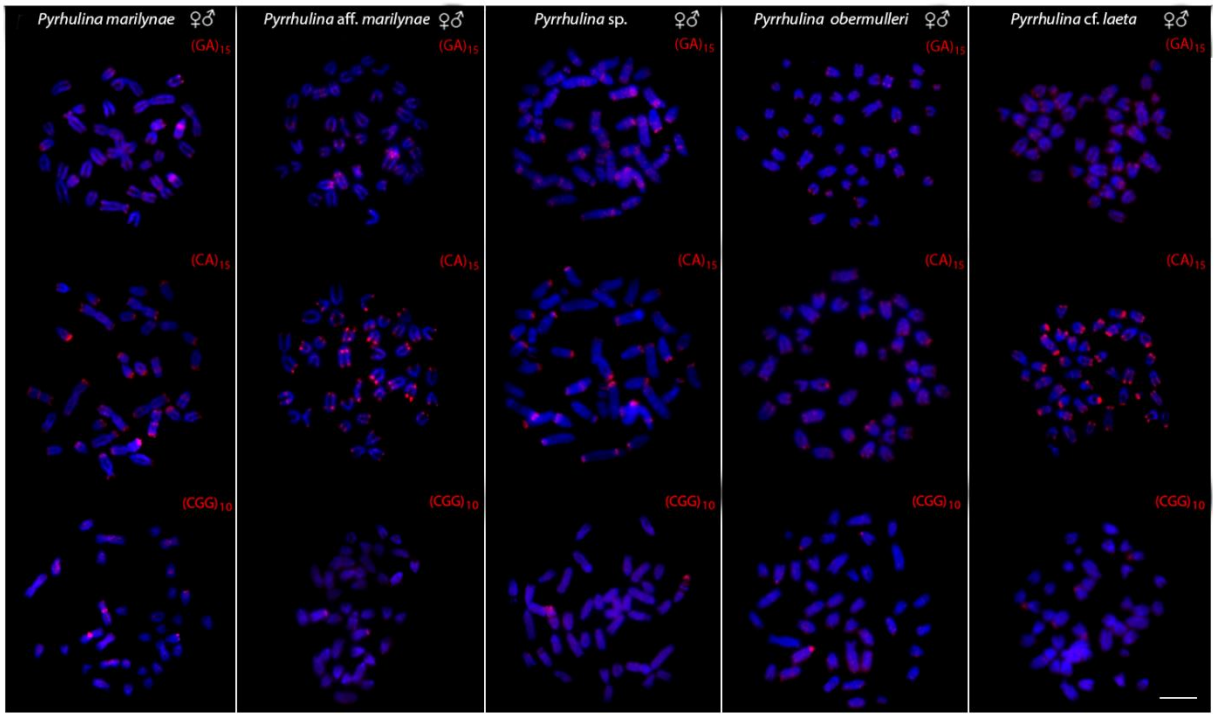


Figure 3. Male and female metaphase plates of *Pyrrhulina marilynae*; *Pyrrhulina* aff. *marilynae*; *Pyrrhulina* sp.; *P. obermulleri* and *Pyrrhulina* cf. *laeta* shows the general distribution of the microsatellites $(GA)_{15}$, $(CA)_{15}$ and $(CGG)_{10}$ on chromosomes. Bar = 5 μm.

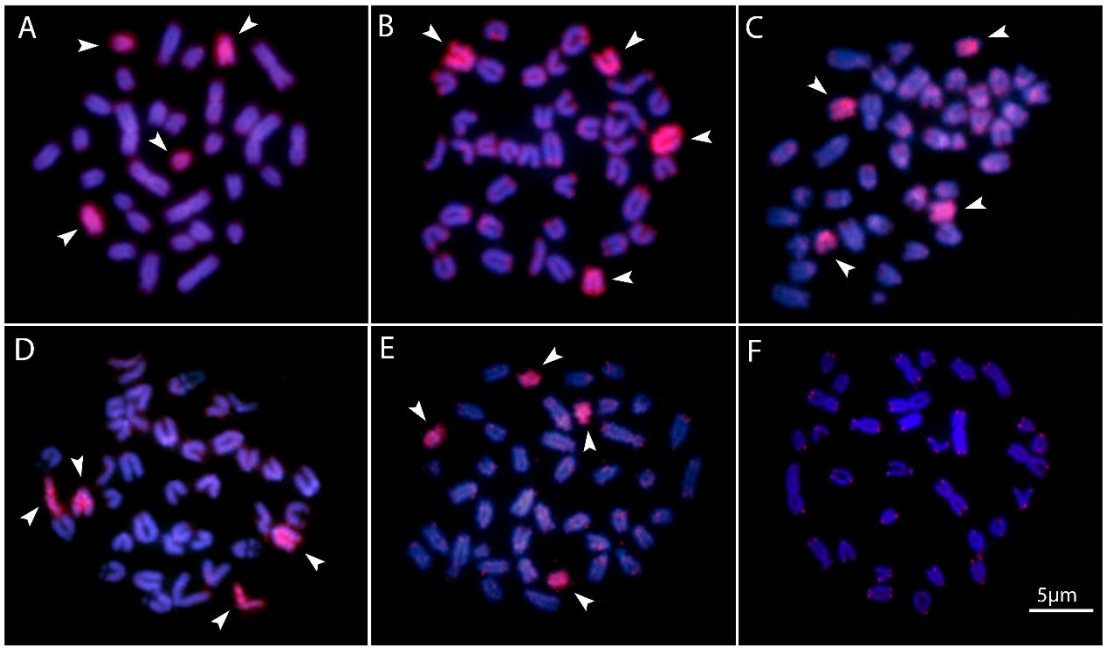


Figure 4. Zoo-FISH with the PSEMI-Y probeon male metaphase plates of *P. marilynae***(A)**, *Pyrrhulina* aff. *marilynae* **(B)**, *Pyrrhulina* cf. *laeta* **(C)**, *Pyrrhulina* sp. **(D)**, and *P. obermulleri* **(F)** shows the distribution of the telomeric (TTAGGG)n repeats in *P. marilynae*. Bar=5 μm.

Two acrocentric chromosome pairs were entirely painted with the PSEMI-Y probe in *Pyrrhulina marilynae*, *P. obermulleri*, *Pyrrhulina* sp., *Pyrrhulina* aff. *marilynae* and *Pyrrhulina* cf. *laeta* (**Figures 4A–E**).

3.3. Comparative Genomic Hybridization–CGH

The interespecific genomic comparison among *Pyrrhulina marilynae* and other *Pyrrhulina* species (*P. semifasciata*, *P. australis*, *P. brevis*, *P. obermulleri*, *Pyrrhulina* aff. *australis*, *Pyrrhulina* sp., *Pyrrhulina* aff. *marilynae*, *Pyrrhulina* cf. *laeta*) revealed a high level of DNA compartmentalization, within all species presenting a distinct composition of repetitive DNA sequences and specific signals. However, *P. marilynae* shows more evident species-specific arrangements when compared to the other species. (**Figure 5**). Intraspecific genomic hybridization between males and females did not show any clustering for sex-specific sequences in all species (data not shown).

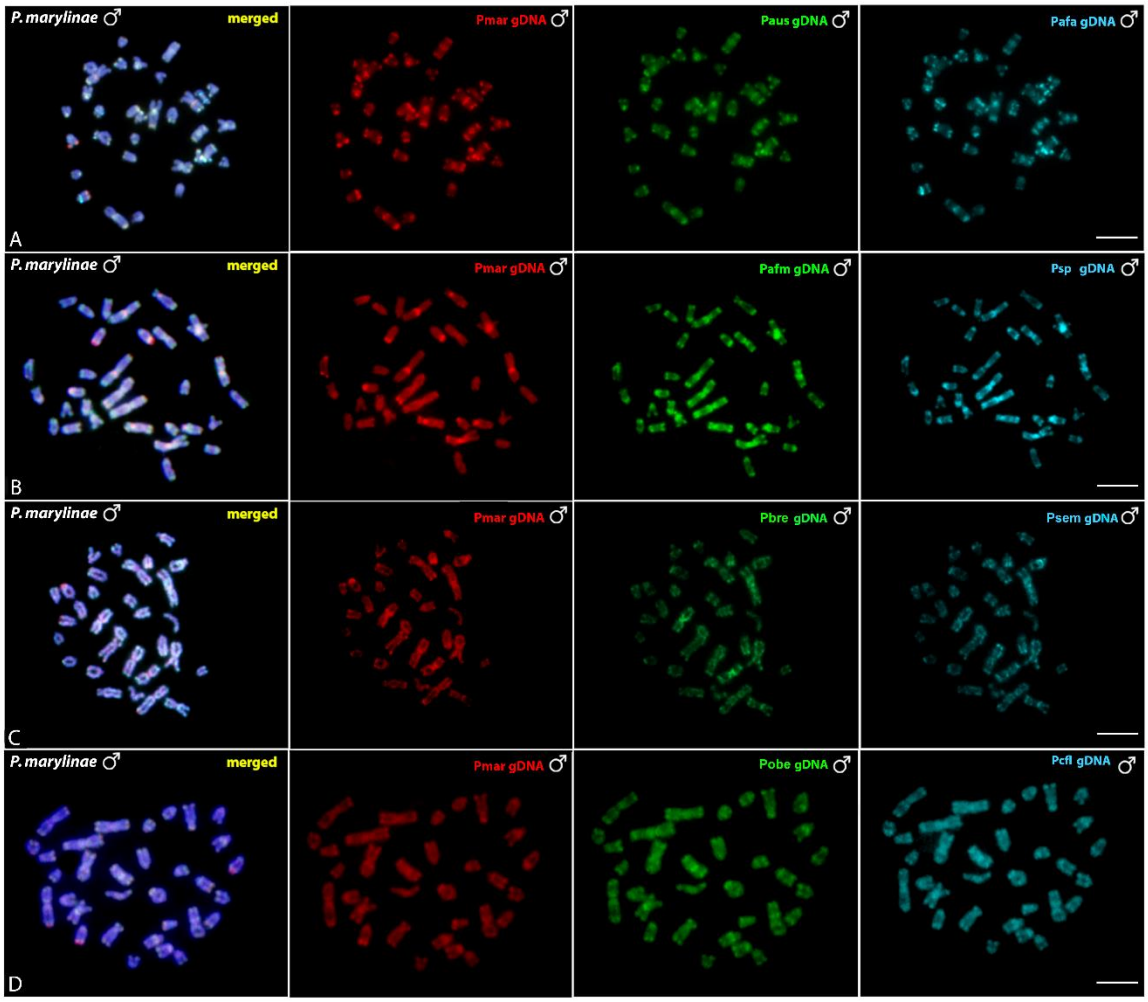


Figure 5. Comparative genomic hybridization (CGH) using male-derived genomic probes

from *Pyrrhulina* species hybridized onto male chromosomes of *P. marilynae*. The common genomic regions are depicted in the 1st column in each line representing the experiments A-D. Hybridization between the gDNA of *P. marilynae* (Pmar), *P. australis* (Paus) and *Pyrrhulina* aff. *australis* (Pafa) **(A)**; *P. marilynae* (Pmar), *Pyrrhulina* aff. *marilynae* (Pafm) and *Pyrrhulina* sp. (Psp) **(B)**; *P. marilynae* (Pmar), *P. brevis* (Pbre) and *P. semifasciata* (Psem) **(C)**; *P. marilynae* (Pmar), *P. obermulleri* (Pobe) and *Pyrrhulina* cf. *laeta* (Pcfl) **(D)**. Bar = 5 μm.

4. Discussion

Overall, two main evolutionary trends are proposed for the karyotypic evolution of the Lebiasinidae: 1) the conservation of a plesiomorphic karyotype in the subfamily Lebiasininae, with 2n = 36 bi-armed chromosomes and, 2) high variations in diploid numbers and karyotypic structures in the subfamily Pyrrhulininae, with the predominance of acrocentric chromosomes (Sassi *et al*., 2020). It is noteworthy that the karyotypic structure of Lebiasininae, 2n = 36 biarmed chromosomes, is similar to that found in the sister family Ctenoluciidae (de Souza e Sousa *et al*., 2017; Sassi *et al*., 2019; de Souza e Sousa *et al*., 2021). Therefore, in this scenario, the majority of the acrocentric chromosomes found in the species of the Pyrrhulininae are probably derived from rearrangements such as centric fissions (Sassi *et al*., 2020). However, unlike other *Pyrrhulina* species, *P. marilynae* has the smallest 2n identified in the genus so far, 2n = 32, including four typical meta/submetacentric pairs. Some exceptions within the subfamily showed secondary fusion events of acrocentric chromosomes giving rise to metacentric chromosomes, reducing the diploid number as observed in *Nannostomus unifasciatus* (Sember *et al*., 2020). Biarmed chromosomes could also represent remnants of the ancestral karyotype condition that were maintained during the evolutionary processes. However, no ITS was found in any chromosome of *P. marilynae*, but such a scenario does not exclude the hypothesis of fusion, given that telomeric regions can be lost after the rearrangement occurs (Bolzán, 2017). Thus, to corroborate such hypotheses and to determine whether the evolutionary trajectory of karyotype change in *Pyrrhulina* is directed mainly towards centric fusions or fissions, cytogenetic data should be discussed in a larger phylogenetic framework of interspecific and intergeneric relationships of Lebiasinidae.

CGH procedures have greatly assisted cytogenetic studies (Symonová *et al*., 2013; Cioffi *et al*., 2017; Cioffi *et al*., 2019), as among all *Pyrrhulina* studied so far. In fact, despite showing close genomic similarities, the species also show considerable divergences, in addition to species-specific repetitive DNA and C-band patterns, thus helping to understand their differential evolutionary paths, considering the taxonomic problems still pending in this fish group. In addition, multiple and syntenic ribosomal sites are not frequently observed among

fishes, but these chromosomal features are very informative cytotaxonomic markers regarding Pyrrhulininae species. Comparatively, they occur more frequently among *Pyrrhulina* than in other species of this subfamily (Moraes *et al*., 2017; Moraes *et al*., 2019; Sassi *et al*., 2019; Sassi *et al*.,2020; Toma *et al*., 2019; Sember *et al*., 2020). Like *Pyrrhulina* aff. *australis* (Moraes *et al*., 2017), *Pyrrhulina* sp., and *P. marilynae* present multiple 5S rDNA sites and only one 18S rDNA site, thus differentiating them from *Pyrrhulina* aff. *marilynae*, *P. obermulleri*, and *Pyrrhulina* cf. *laeta*, as well as from some other *Pyrrhulina* species (Moraes *et al*., 2017; Moraes *et al*., 2019), which have multiple 5S and 18S rDNA sites. Furthermore, the syntenic condition for the 18S/5S rDNAs in *Pyrrhulina* cf. *laeta* is shared with *P. brevis* and *P. australis*, indicating a high rDNA diversity. (**Figure 6**). In its turn, the 18S rDNA clusters are distributed on distal chromosome positions for all investigated *Pyrrhulina* species (Moraes *et al*., 2017; Moraes *et al*., 2019; this study), as also occur among *Copeina* (Toma *et al*., 2019), *Lebiasina* (Sassi *et al*., 2019), and *Nannostomus* (Sember *et al*., 2020), so as in the species of the sister family, Ctenoluciidae (de Souza e Sousa *et al*., 2017; de Souza e Sousa *et al*., 2021).

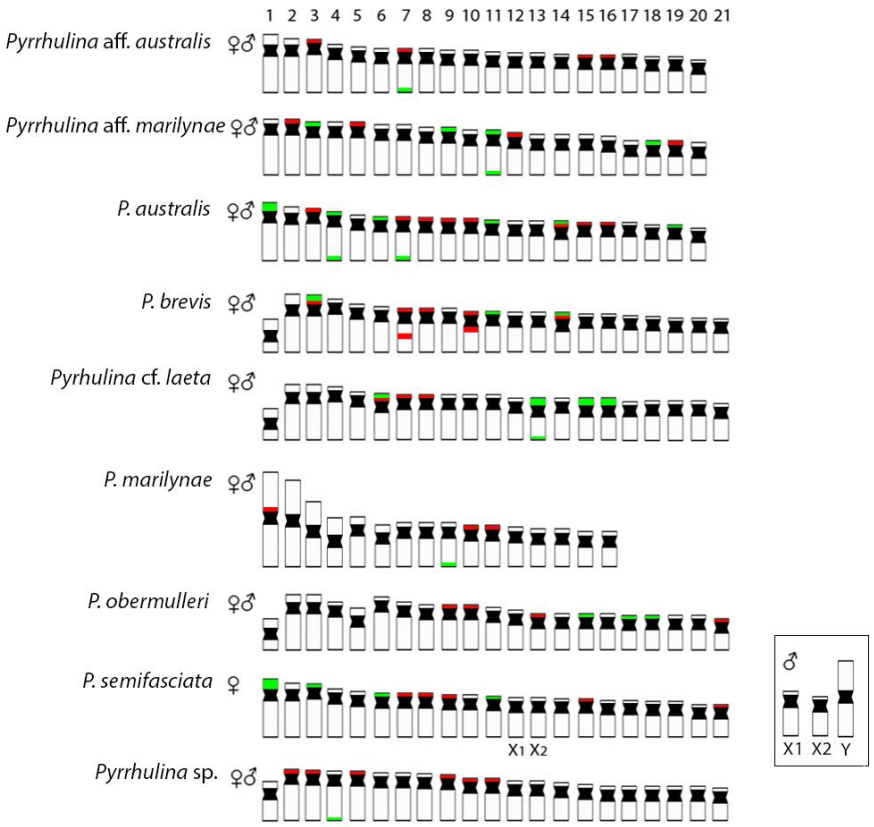


Figure 6. Representative idiograms of *Pyrrhulina* species showing the distribution of the 18S (green) and 5S rDNA (red) sites on chromosomes, based on the present study and some other previous data (Moraes *et al*., 2017; Moraes *et al*., 2019). Bar = 5 µm.

Microsatellite distribution patterns have significantly contributed to evolutionary studies in fish species, especially regarding sex chromosome differentiation (Kubat *et al*., 2008; Cioffi *et al*., 2012; Terencio *et al*., 2012; Kejnovský *et al*., 2013; Poltronieri *et al*., 2014; Yano *et al*., 2014; de Freitas *et al*., 2018). Among the five *Pyrrhulina* species now investigated, as well as in other previous analyzed ones (Moraes *et al*., 2017; Moraes *et al*., 2019), the distribution of the microsatellites did not significantly differ among them, although the (CA)15 repeats present a greater number of more conspicuous sites than the other microsatellites, especially in *Pyrrhulina* aff. *marilynae* and *Pyrrhulina* cf. *laeta*. It is noteworthy that microsatellites have a preferential location in the telomeric and centromeric regions of fish chromosomes (Cioffi and Bertollo, 2012), as occur with the (CA)15 and (GA)15 motifs in *Pyrrhulina*, despite some interstitial and pericentromeric signs in *Pyrrhulina* cf. *laeta*, *P. marilynae*, *Pyrrhulina* aff. *marilynae* and *Pyrrhulina* sp., thus differentiating these species from others previously studied (Moraes *et al*., 2017; Moraes *et al*., 2019). Furthermore, it is also frequent that microsatellites and other repetitive sequences occur in the association among fish (Cioffi and Bertollo, 2012), such as in Hepsetus odoe (Carvalho *et al*., 2017), Lebiasina bimaculata (Sassi *et al*., 2019), and Silurichthys phaiosoma (Ditcharoen *et al*., 2020), for example. This is the scenario that also occurs in *Pyrrhulina* sp., in which the (CGG)10 microsatellite located in the telomeric region of pair 4 shares the same chromosomal region with 18S rDNA repeats.

Fish, besides presenting high diversity in morphological and genetic characteristics, also have a variety of sex chromosome systems (Sember *et al*., 2021). About nine differentiated systems, involving the XX/XY and ZZ/ZW sex chromosomes and their variations, have been identified among species, including several Neotropical ones (Sember *et al*., 2021). It is noteworthy that among the multiple systems, the ♀X1X1X2X2/♂X1X2Y is the most prevalent one, and commonly originated by centric or tandem fusions of the ancestral Y with an autosomal member of the karyotype, giving rise to neo-Y chromosomes, as identified in a variety of fish species (Sember *et al*., 2021). That includes *P. semifasciata*, the only Lebiasinidae representative highlighting heteromorphic sex chromosomes so far (Moraes *et al*., 2019), in addition to a putative ZZ/ZW sex system present in *Lebiasina bimaculata* (Sassi *et al*., 2019). Although our intraspecific CGH results in the current analyzed species did not reveal any sex-specific differentiated region, our WCP experiment with the Y-derived probe of *P. semifasciata* entirely painted two acrocentric pairs, suggesting that putative proto-XY chromosomes may occur in these species. Thus, it supports our previous hypothesis on the origin of the *P. semifasciata* sex chromosome system through centric fusion between the non-homologous acrocentric, giving rise

to the large metacentric Y chromosome. That can be considered as an apomorphy of this species when compared to others of the genus. Furthermore, the experiments also showed that although the karyotype of *P. marilynae* has large metacentric chromosomes, these do not correspond to the heteromorphic sex chromosome of *P. semifasciata* (**Figure 4**).

5. Conclusion

Our data advances the understanding of evolutionary trends of the Lebiasinidae, particularly concerning *Pyrrhulina*. Karyotypes with 2n = 40–42, with the predominance of mono-armed chromosomes, are more frequent among its species, except for *P. marilynae*, which has a smaller diploid number (2n = 32), and several atypical biarmed chromosomes, a characteristic that differentiates this species from the others analyzed in the genus. However, we cannot rule out the hypothesis that this karyotypic reduction (2n = 32) may have been generated by secondary fusions that allowed the formation of the four meta/submetacentric pairs identified in *P. marilynae*. The present data also highlighted the putative proto-XY chromosomes that may occur in these species and support the occurrence, through centric fusion, of the multiple sex chromosome system of *P. semifasciata* as an independent evolutionary event of this Lebiasinidae species. Our results highlight the importance of chromosomal data as valuable markers for understanding the evolutionary relationships among Lebiasinidae species.

References

All references are compiled in the end of this thesis.

.

Capítulo 2

The Genetic Differentiation of *Pyrrhulina* (Teleostei, Characiformes) Species is Likely Influen-ced by Both Geographical Distribution and Chromosomal Rearrangements

Pedro Henrique Narciso Ferreira, Fernando Henrique Santos de Souza, Renata Luiza Rosa de Moraes, Manolo Fernandez Perez, Francisco de Menezes Cavalcante Sassi, Patrik Ferreira Viana, Eliana Feldberg, Tariq Ezaz, Thomas Liehr, Luiz Antônio Carlos Bertollo and Marcelo de Bello Cioffi

Frontiers in Genetic (2022) – DOI: 10.3389/fgene.2022.869073

Abstract

Allopatry is generally considered to be one of the main contributors to the remarkable Neotropical biodiversity. However, the role of chromosomal rearrangements including neo-sex chromosomes for genetic diversity is still poorly investigated and understood. Here, we assess the genetic divergence in five *Pyrrhulina* species using population genomics and combined the results with previously obtained cytogenetic data, highlighting that molecular genetic diversity is consistent with their chromosomal features. The results of a principal coordinate analysis (PCoA) indicated a clear difference among all species while showing a closer relationship of the ones located in the same geographical region. This was also observed in genetic structure analyses that only grouped *P. australis* and *P. marilynae*, which were also recovered as sister species in a species tree analysis. We observed a contradictory result for the relationships among the three species from the Amazon basin, as the phylogenetic tree suggested *P. obermulleri* and *P. semifasciata* as sister species, while the PCoA showed a high genetic difference between *P. semifasciata* and all other species. These results suggest a potential role of sex-related chromosomal rearrangements as reproductive barriers between these species.

1. Introduction

The Neotropical region harbors most of the known fish diversity, with more than 5,200 valid species, however, there is additional evidence that many of them are yet to be described (Fricke *et al*., 2021). Several reasons are suggested to be related to this huge diversity, and among them, the evolution of reproductive barriers reducing gene flow plays a key role in creating reproductive incompatibilities between species (Mayr, 1963). The evolution of strong reproductive barriers is essential for coexistence of species, especially for sympatric ones (those whose geographical distribution overlaps), even if the isolation in sympatry has preceded the overlap of species that currently coexist in the same region, in a secondary contact (Mayr, 1999; Coyne and Orr, 2004).

Among animals, most cases of post-zygotic isolation are caused by genetic incompatibilities, among which chromosomal rearrangements including specific sex chromosome-systems are of fundamental importance (Bateson, 1909; Dobzhansky, 1933; Muller, 1942; Coyne and Orr, 2004). Such chromosomal rearrangements can prevent introgression and reduce gene flow by suppressing recombination (Machado *et al*., 2007; Yannic *et al*., 2009; McGaugh and Noor, 2012; Ostberg, *et al*., 2013; Presgraves, 2018; Potter, *et al*., 2021). However, the contribution of these processes for biodiversity in Neotropical fish species is still poorly understood.

The Lebiasinidae family is distributed throughout Central and South America, with 75 species divided into two subfamilies: Lebiasininae and Pyrrhulininae (Weitzman and Weitzman, 2003; Froese and Pauly, 2014; Fricke *et al*., 2021). The subfamily Pyrrhulininae is the most diverse, including the genera *Derhamia, Nannostomus, Copella, Copeina*, and *Pyrrhulina* (Fricke *et al*., 2021). Recent cytogenetic and molecular data have added significant contributions to understanding evolutionary relationships of these genera (Moraes *et al*., 2017, Moraes *et al*., 2019, Sassi *et al*., 2019; Toma *et al*., 2019; Sember *et al*., 2020; Moraes *et al*., 2021). Particularly, concerning *Pyrrhulina*, its diploid number (2n) ranges from 32 to 42 chromosomes, with *Pyrrhulina* aff. *australis, P. australis* (Eigenmann and Kennedy, 1903), *Pyrrhulina* aff. *marilynae* and *Pyrrhulina* sp. present 2n = 40 chromosomes, whereas *P. brevis* (Steindachner, 1876), *P. obermulleri* (Allen *et al*., 1926), and *Pyrrhulina* cf. *laeta* have 2n = 42 chromosomes, both in males and females. Differently from the others, *P. semifasciata* (Steindachner, 1876) has divergent 2n among sexes (2n = ♀42 and ♂41) due to its exclusive multiple ♀X1X1X2X2/♂X1X2Y sex chromosome system (Moraes *et al*., 2017; Moraes *et al*., 2019). In addition, *P. marilynae* (Netto-Ferreira and Marinho, 2013) also differs from its

congeneric species having 2n = 32 chromosomes, due to a series of chromosomal fusions (Moraes *et al*., 2021).

The development and improvement of large-scale genotyping-by-sequencing (GBS) procedures allowed high-resolution analysis using Single Nucleotide Polymorphisms (SNPs) for genetic diversity studies in non-model organisms (Steane *et al*., 2011; Yang *et al*., 2013; Sánchez-Sevilla *et al*., 2015; Alkimin *et al*., 2018; Souza *et al*., 2019; Oliveira *et al*., 2020; Sassi *et al*., 2021). Combining such genomic and cytogenetic data for pairs of species/populations in which there are variations related to geographic distribution, analysis of chromosomal rearrangements, and sex chromosome systems provide a critical tool to understanding the role of these events to generate biodiversity. The present study applies cytogenetic and genomic approaches in five *Pyrrhulina* species with distinct geographic distributions. Here, we aimed to assess the genetic diversity in five *Pyrrhulina* species by combining population genomic with previously obtained cytogenetic data. Our results highlighted that genomic diversity is consistent with the chromosomal features and provided hypothesis on the potential role of chromosomal rearrangements, especially those involving the sex chromosomes, in fostering genetic differentiation.

2. Materials and Methods

2.1. Sampling

Collection sites, number, and sex of the specimens are given in **Figure 1**; **Table 1**. 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). The distribution data was obtained from the analysis of scientific collection materials and previous literature (Bertaco *et al*., 2016; Dagosta and Depinna, 2021).

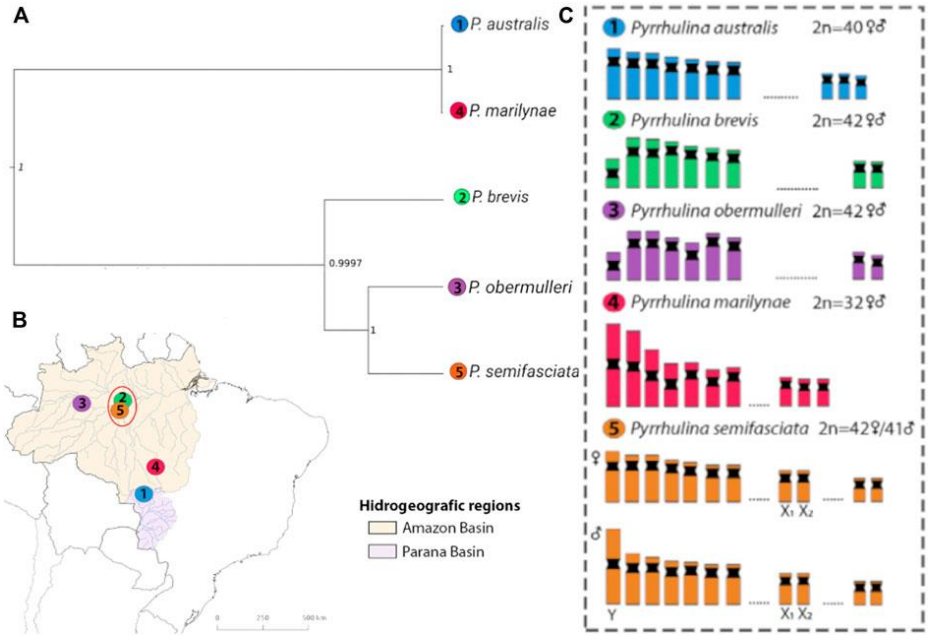


Figure 1: Phylogenetic relationships of *Pyrrhulina* species analyzed and their distribution. (A)—Species tree recovered with SNAPP, with branch lengths measured in units of expected number of mutations per site, based on dataset I: *P. australis* (1), *P. brevis* (2), *P. obermulleri* (3); *P. marilynae* (4); and *P. semifasciata* (5). (B)—South America map indicating the origin of *Pyrrhulina* species analyzed. Colored areas indicate the distribution of the species: *P. australis* (blue), *P. brevis* (green), *P. obermulleri* (yellow); *P. marilynae* (purple); and *P. semifasciata* (pink). The numbered circles indicate the collection sites of each species. The red ellipse indicates syntopic species (sampled in the same area). (C)—Idiograms with partial karyotypes of each species and the sex chromosomes exclusively found in *P. semifasciata* are boxed.

Table 1. Species, diploid numbers (2n), sex chromosome systems, numbers of individuals cytogenetically analyzed, and the number of individuals sequenced

Species	2n	Sex chromosomes	Sampling location	Latitude/ Longitude	Cytological analysis	Sequence analysis	References
P. australis	♀♂40	Homomorphic	Barra do Bugres—MT	-15.04275,-57.11054	18♂ 30♀	6	Moraes et al. (2017)
P. brevis	♀♂42	Homomorphic	Adolpho Ducke Reserve—AM	-2.58207,-59.55530	17♂ 13♀	6	Moraes et al. (2019)
P. semifasciata	♀42 ♂41	$X_1X_1X_2X_2/X_1X_2Y$	Careiro—AM	-3.51000,-60.04000	12♂ 07♀	6	Moraes et al. (2019)
P. obermulleri	♀♂42	Homomorphic	Tefé—AM	-3.25507,-64.44548	21♂ 12♀	6	Moraes et al. (2021)
P. marilynae	♀♂32	Homomorphic	Ipiranga do Norte—MT	-11.36020,-55.56270	14♂ 08♀	6	Moraes et al. (2021)

Abbreviations: AM, Amazonas; MT, Mato Grosso Brazilian states. Chromosomal data refer to previous studies, according to references.

2.2. DArTseq Sequencing Procedure and Data Filtering

Liver tissue from all individuals was used to perform DArTseq sequencing procedures at Diversity Arrays Technology Pty Ltd. This sequencing method uses a combination of restriction

enzymes (SbfI and PstI) that enrich hypomethylated regions (Kilian *et al*., 2012). Sequencing of the obtained libraries was carried out on the Illumina HiSeq 2500 platform.

The raw data were processed in two different ways to generate the following datasets (I) a matrix of SNPs using Ipyrad v. 0.9.84 (Eaton and Overcast, 2020) and (II) a phased sequence file obtained with pyRAD v3.0.66 (Eaton, 2014). In both, the sequencing adapters were trimmed, and all sequences shorter than 35 base pairs or presenting more than five low-quality bases (Q < 20) were not considered. In dataset I, a single SNP per locus was selected, to reduce the inclusion of linked SNPs, and the SNPs were coded as 0 for reference state homozygotes, 1 for heterozygotes, and 2 for alternative state homozygotes.

The raw data is available in the database of the National Center for Biotechnology Information, whose project access code is PRJNA804560 (access link: https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA804560eo=acc_s %3Aa). Additionally, datasets one and two are available at https://github.com/PHnarciso/Dataset-1-and-2---Manuscript-*Pyrrhulina*.

2.3. Assessment of Markers Under Selection

A BayeScan analysis was carried to search for potential markers under selection in both datasets (Foll and Gaggiotti, 2008). We performed a total of 5,000 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.

2.4. Genetic Diversity

Using dataset II, an analysis was performed in the DnaSP software v. 6.12.03 (Rozas *et al*., 2019), to obtain the haplotype diversity (Hd), Tajima's D (D), and two measures of nucleotide diversity per site, π and Watterson's theta (θw). Differentiation was also estimated from the nucleotide divergence between samples using the average number of nucleotide substitutions per site between pairs of samples from different species (Dxy) and the net divergence, corrected for variation within the samples analyzed (Da).

2.5. Genetic Structure and Analysis of Molecular Variance (AMOVA)

A principal coordinate analysis (PCoA) in the R package dartR (Gruber *et al*., 2018) was used to investigate the genetic structure among species. The genetic structure was also investigated with fastSTRUCTURE (Raj *et al*., 2014), using the "Lizards-are-awesome" pipeline

(Melville *et al*., 2017). 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 from fastStructure were visualized with Clumpak (Kopelman *et al*., 2015). Additionally, an analysis of molecular variance (AMOVA) was also performed using dataset I, with samples grouped as follows: 1) by species, 2) by the clusters from the best K value on fastStructure, 3) by the clustering pattern generated on the PCoA, and 4) by the presence or absence of the multiple sex chromosome system.

2.6. Species Tree

We used the package SNAPP in BEAST 2.6.4 (Bouckaert *et al*., 2014) to infer the species tree topology. We included only polymorphic sites, used the default prior for coalescence rate, and calculated backwards and forward substitution rates (parameters u and v) based on the empirical data. We applied a chain length of two million generations, sampling every 5000 iterations. Convergence was evaluated in Tracer 1.7.1 (Rambaut *et al*., 2018). We obtained the MCC tree based with common ancestor heights in TreeAnnotator, with the first 25% generated trees discarded as burn-in. We exported the consensus tree in FigTree 1.4.4.

2.7. Analysis of Introgression

We assessed admixture between lineages by calculating Patterson's D-statistics (ABBA-BABA test) in Ipyrad v. 0.9.84 (Eaton and Overcast, 2020) using pooled SNP frequencies of individuals from each species (Durand *et al*., 2011). We performed 4-taxon tests according to the recovered SNAPP topology (**Figure 1**), using either *P. marilynae* or *P. australis*as the outgroup, *P. brevis* as P3 and *P. semifasciata* and *P. obermulleri* as either P1 or P2. Significance was measured with 1,000 bootstrap replicates by resampling loci with replacement (following Eaton and Ree, 2013). Results were considered significant according to their Z-scores (significant values >3), which quantifies the number of bootstrap standard deviations in which the D-statistic values deviates from its expected value of zero.

3. Results

3.1. Sequencing, Filtering, and Detection of Markers Under Selection

Sequencing of DArTSeq resulted in two million reads per sample, on average, a total of 5,149 SNP-markers were obtained after performing all filtering steps in dataset I, and a total of 5,604 loci were obtained in dataset II. According to the BayeScan result, none of the loci were

potentially under selection, as all FDR values were higher than 0.65 and 0.97 for dataset I and dataset II respectively. Thus, all markers were maintained in subsequent analyses.

3.2. Genetic Diversity and Structure

In general, all species showed low diversity values. Values of π and θw were closely related across all sampled groups, with slightly higher estimates in *P. obermulleri* and *P. semifasciata*, while *P. marilynae* showed the lowest values. (**Table 2**).

Table 2. Estimated genome-wide genetic diversity of genus *Pyrrhulina* by species.

Species	Sample size	*Hd*	*π*	*θw*	*D*
P. australis	6	0.1230	0.0024	0.0026	–0.2004
P. brevis	6	0.1048	0.0020	0.0024	–0.3512
P. obermulleri	6	0.2033	0.0042	0.0044	–0.1656
P. marilynae	6	0.0338	0.0007	0.0006	0.1380
P. semifasciata	6	0.1712	0.0033	0.0039	–0.4412

Abbreviations: Hd - *haplotype diversity;* π - *nucleotide diversity;* θw - *Watterson theta per site and* D - *Tajima's D.*

The nucleotide divergence (Dxy values) and net divergence (Da values) estimates between individuals sampled of different species ranged from 0.01770 to 0.00660 as shown in **Figure 2**. The lowest divergence values estimates were presented for *P. australis* and *P. marilynae*, and the highest values were observed for comparisons of *P. semifasciata* with these species.

	P. australis	*P. brevis*	*P. obermulleri*	*P. marilynae*	*P. semifasciata*
P. australis		0.01320	0.01550	0.00800	0.01710
P. brevis	0.01110		0.01210	0.01380	0.01380
P. obermulleri	0.01230	0.00900		0.01600	0.01460
P. marilynae	0.00660	0.01250	0.01360		0.01770
P. semifasciata	0.01440	0.01120	0.01090	0.01580	

Figure 2. Pairwise Dxy per Species is represented in the upper diagonal, and Da in the lower diagonal. Higher values are shown in green, and as the values decrease the color changes to yellow, orange, and then red for lower values.

The PCoA results separate all species and indicate a higher similarity for species in the same regions, as *P. brevis* and *P. obermulleri* are more similar to each other, as are *P. australis* and *P. marilynae*. In turn, *P. semifasciata* appeared as the most different species of the group (**Figure 3**).

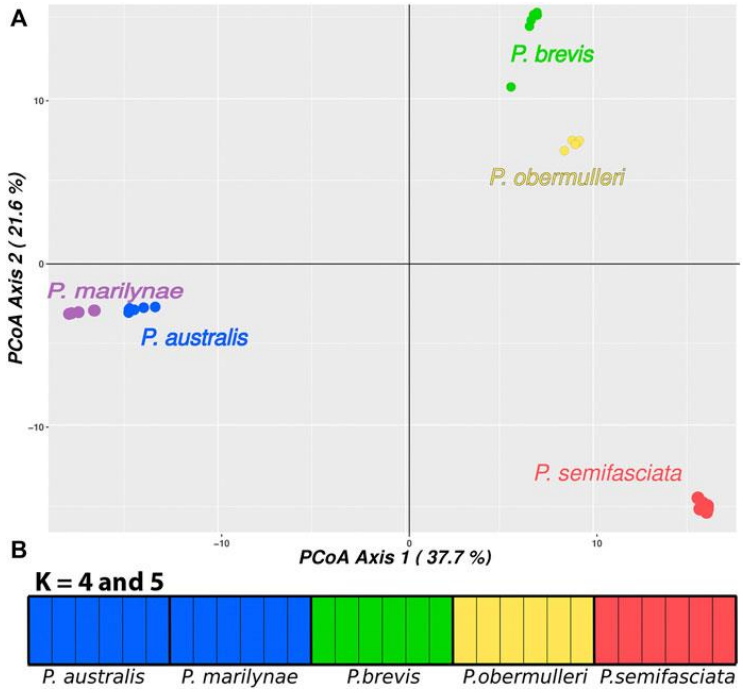


Figure 3. (A) Genetic diversity in *Pyrrhulina* species according to a PCoA. The PCoA recovered 37.7% of the total variation in the first principal coordinate (PC1), and 21.6% in the second principal coordinate (PC2). (B)—Results for fastStructure with K = 4 and K = 5. Each vertical bar represents an individual; the bar colors represent the group in which the fastStructure classified the individual; the legend below indicates their species.

Results from fastStructure suggested K = 4 as the number of clusters that maximizes both likelihood and is more informative for genetic structure, being the one that best explains the complexity of the data. It is worth mentioning that the clusterization presented in K = 5 is identical to that presented by K = 4, with both results combining *P. australis*and *P. marilynae*. AMOVA results suggested that the clustering strategy presenting the highest variation between groups and the lowest variation within groups was achieved when each separate species was considered (83.66% variation between groups), followed by the fastStructure K = 4 result and by the PCoA (**Table 3**).

Table 3. AMOVA percentage of variation 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, and 4) by the presence or absence of sex chromosomes.

Clustering	Species	FastStructure K = 4	PCoA	Presence/absence of sex chromosomes
Variation between groups	83.664	73.916	61.593	46.534
Variation within group	2.016	12.465	25.087	41.805

3.3. Species Tree and Analysis of Introgression

The species tree recovered in SNAPP showed a relationship between species similar to the structure shown in PCoA (**Figure 3**), showing *P. australis* and *P. marilynae* as sister species. A sister species relationship was also recovered for *P. semifasciata* and *P. obermulleri*, with *P. brevis* closer to this last two (**Figure 1**). Results of the admixture tests using D-statistics were not able to detect significant introgression in *P. semifasciata* or *P. obermulleri*, regardless of whether *P. marilynae* or *P. australis* were considered as the outgroup (Table 4). This result agrees with the lack of admixture suggested by fastSTRUCTURE (**Figure 3**).

Table 4. Results of the ABBA-BABA test for admixture assessment.

Outgroup	P_3	P_2	P_1	D	Z	ABBA	BABA
P. marilynae	*P. brevis*	*P. obermulleri*	*P. semifasciata*	−0.0214164	0.44500958	180.753565	188.665185
P. australis	*P. brevis*	*P. obermulleri*	*P. semifasciata*	0.00523111	0.1052923	182.665167	180.764028
P. marilynae	*P. brevis*	*P. semifasciata*	*P. obermulleri*	−0.0052311	0.1035013	180.764028	182.665167
P. australis	*P. brevis*	*P. semifasciata*	*P. obermulleri*	−0.0052311	0.10302304	180.764028	182.665167

Abbreviations: D, overrepresentation of ABBA against BABA patterns as measured by the D-statistics; Z, Z-score test to assess whether the D-statistics is significantly (Z > 3) different from zero; ABBA, frequency of ABBA patterns; BABA, Frequency of BABA patterns.

4. Discussion

Here the first genomic evaluation of five *Pyrrhulina* species was performed. Our results highlight that genetic diversity among the analyzed *Pyrrhulina* species is consistent with their chromosomal features. In addition, evidence was found that the occurrence of chromosomal rearrangements including variation in sex chromosome systems might have contributed to increase differentiation.

Overall, the genetic diversity indexes (π; θW) (**Table 2**) were low, being the highest ones observed in *P. obermulleri* and *P. semifasciata* and the lowest in *P. marilynae* followed by *P. brevis*. Low diversity values are somewhat expected due to the nature of our sequencing technique (see methods) and similar results were found in populations from *Hoplias malabaricus* under the same sequencing technique (Souza *et al*. submitted). According to Dxy and Da values (**Figure 2**), *P. semifasciata* and *P. marilynae* are very distinct from each other, presenting the highest observed values (0.01770 and 0.0158). They are also distinct morphologically (Netto-Ferreira and Marinho, 2013) and present non-overlapping distributions in the Amazon basin. *Pyrrhulina* semifasciata occur in the Purus, Juruá, Negro, Branco rivers, Amazonas main channel (Dagosta and De Pinna, 2021), and Madeira River while *P. marilynae* is restricted to the upper portions of Tapajós and Xingu (Netto-Ferreira e Marinho, 2013) (**Figure 1**). Similarly, values of Dxy and Da between *P. semifasciata* and *P. australis* are also very high (0.0171 and

0.0144), indicating their genetic distinctiveness. Furthermore, their distribution does not overlap (**Figure 1**) and they are also quite distinct morphologically and probably distantly related species, according to a morphology-based phylogenetic analysis (Marinho, 2014). This results are also in agreement with the species tree, as *P. marilynae* and *P. australis* were recovered as sister species in a distinct clade (**Figure 1**). However, it is important to emphasize that a more comprehensive species-level sampling is still necessary before claiming sister species relationships, especially in taxa that are apparently divergent such as *P. brevis* and *P. obermulleri*.

Pyrrhulina semifasciata, the only species in the family up to now analyzed with heteromorphic sex chromosomes, presents high genetic differentiation values when compared to all other species. This could suggest that the presence of such multiple sex system promoted genetic differentiation in the species. The association of sex chromosomes and a high genetic differentiation was also evidenced in previous studies for some populations of other species such as the wolf fishes Hoplias malabaricus (Souza *et al*., submitted) and *Erythrinus erythrinus* (Souza *et al*., 2022). In addition, other studies suggest that an event of sex chromosome turnover in plants could have promoted the reduction of gene flow between groups of Rumex hastatulus, which could have contributed to further accumulation of genetic differences (Beaudry *et al*., 2019). Furthermore, the emergence of a neo-sex chromosome system is suggested as a factor that may have driven the speciation process in sticklebacks (Kitano *et al*., 2009). Therefore, the emergence of sex neo-chromosomes could result in a rapid evolutionary mechanism fixing genetic variation in different species (Beaundry *et al*., 2019). Indeed, our results suggest that *P. semifasciata*, which was collected in syntopy with *P. brevis* and presents a different chromosomal structure with a multiple sex chromosome system, is also genetically highly divergent from that latter species.

Therefore, we present three main hypotheses for the observed scenario (I) *P. brevis* and *P. semifasciata* evolved in allopatry. Later, their distribution overlapped and the current genetic differences are the result of this previous allopatric period and not related to the sex chromosomes emergence (II) these species evolved in allopatry and the sex chromosomes present in *P. semifasciata* may have acted as an additional barrier to introgression, fostering their genetic differentiation after they became sympatric, and (III) the emergence of a multiple sex chromosome system of the X_1X_2Y-type in *P. semifasciata* have acted as key factor to reduction of the gene flow and contributed to the speciation process and the significant genetic differentiation observed for this species.

The PCoA result also suggested that *P. semifasciata* stayed isolated from all other species, while *P. brevis* clustered with *P. obermulleri* (**Figure 3**). Many works highlight the important role of neo-sex chromosomes in preventing introgression, which is agreement with the lack of introgression between sympatric species detected here (**Table 4**), due to pairing problems associated with meiotic segregation of heterozygous rearrangements (e.g., McKee *et al*., 1998; Kitano and Peichel, 2012; Beaudry *et al*., 2019). Similarly, sex chromosomes might promote speciation, as some genes that play a role in reproductive isolation between populations can be accumulated in neo-sex chromosomes (Reinhold, 1998; Lindholm and Breden, 2002; Kitano and Peichel, 2012). Haldane's rule postulates that hybrids of the heterogametic sex from closely related species are often sterile, thus hybrid incompatibility also may act as a reproductive isolating barrier and contribute to prevent introgression (Haldane, 1922; Coyne and Orr, 2004). In both scenarios, the neo-sex chromosomes system could have contributed to genetic differentiation or speciation. Hence, it may be possible that the chromosomal rearrangements present in *P. semifasciata* may have had an impact on the diversity of *Pyrrhulina* species.

In contrast, some species showed low differentiation, indicating that they are more similar genetically. The Dxy and Da values between *P. marilynae* and *P. australis* were the lowest ones (0.008 and 0.0066 respectively) and these two species were recovered as sister species in our phylogenetic analysis (**Figure 1**). Also, both species share with *P. vittata* a series of unique morphological characters which indicate that they form a monophyletic clade (Netto-Ferreira and Marinho, 2013). *Pyrrhulina* brevis and *P. obermulleri* presented low Dxy and Da values as well (0.0121 and 0.0090) in addition to be close to each other in the PCoA analysis (**Figure 3**) and be recovered in the same clade in the species tree, although not as sister species (**Figure 1**). All Tajima's D values were negative except for *P. marilynae*. Negative results may suggest a recent event of population expansion, or that purifying selection is acting mainly in *P. semifasciata* and *P. brevis*. The small positive value presented by *P. marilynae* could indicate a bottleneck event, but this seems unlikely due to the low value.

The results of PCoA indicated a clear clustering of species sampled in the same geographical region, except for *P. semifasciata*, which was collected in sympatry with *P. brevis* and did not cluster with any other species (**Figure 3**). In fastStructure K = 4 result, only *P. australis* and *P. marilynae* were grouped. In both fastStructure and PCoA results, *P. australis* and *P. marilynae* were closely related, and such a connection is also presented in the speciestree (**Figure 1**) and supported by the species tree presented by Marinho (2014).

In addition, AMOVA results suggest that grouping individuals by species maximize the variation between groups, although the groups generated by fastStructure and PCoA have also shown informative values (**Table 3**). Therefore, AMOVA results show that the clustering of individuals by species best explains the complexity of the data and therefore suggests that the species are well separated. The relationships between river reorganizations and biodiversity are a complex issue, as previously isolated habitats and populations can be merged, increasing dispersal and local diversity, and reducing speciation and extinction rates. They can also separate previously connected environments, leading to increased genetic isolation and speciation, and extinction rates (Tagliacollo *et al*., 2015; Albert *et al*., 2018). Therefore, the complex biogeographic history of the region complicates the understanding of *Pyrrhulina* genetic diversity considering that the current geographic distribution of species may not be the same as in the past and a comprehensive species-level phylogeny is still necessary to infer past geographic distributions.

It is widely known that chromosomal rearrangements might have a role in speciation (Rieseberg, 2001; Navarro and Barton, 2003; Basset *et al*., 2008). Chromosomal rearrangements have the potential to limit introgression, thus facilitating the origin and maintenance of reproductive isolation through recombination suppression (Faria and Navarro, 2010; Sichová *et al*., 2015). Here, *P. australis* and *P. marilynae* were recovered as sister species in our phylogenetic analysis. This suggests a low genetic difference despite their different chromosomal formulae. Conversely, the species *P. semifasciata*, the only with a multiple sex chromosome system, showed high differentiation to all other species. This furnishes additional evidence for the role of chromosomal rearrangements, especially those associated with sex chromosomes, as potential important reproductive barriers between these species.

5. Conclusion

Our results highlighted the distribution of the genetic diversity among the analyzed *Pyrrhulina* species and contributed to our understanding of their evolutionary history. However, the scarcity of phylogenetic data involving all *Pyrrhulina* species does not permit us to do a cause-and-effect correlation. Besides, the occurrence of chromosomal rearrangements leading to the presence of the X1X2Y sex chromosome system in *P. semifasciata* could have led to an intensification of the genetic differentiation of this species by preventing introgression with the other species. Future work with a better population and species-level sampling is still needed to bring new insights to 1) species delimitation, 2) infer demographic trends, 2) infer past

distributional ranges, and 4) study the tempo and mode of speciation, including chromosomal rearrangements as a mechanism for isolation.

References

All references are compiled in the end of this thesis.

Capítulo 3

Chromosomal Rearrangements and Satellite DNAs: Extesive Chromosome Reshuffling and the Evolution of Neo-Sex Chromosomes in the Genus *Pyrrhulina* (Teleostei; Characiformes)

Renata Luiza Rosa de Moraes, Francisco de Menezes Cavalcante Sassi, Jhon Alex Dziechciarz Vidal, Caio Augusto Gomes Goes, Rodrigo Zeni dos Santos, José Henrique Forte Stornioli, Fábio Porto-Foresti, Thomas Liehr, Ricardo Utsunomia and Marcelo de Bello Cioffi

International Journal of Molecular Sciences (2023) – DOI: 10.3390/ijms241713654

Abstract

Chromosomal rearrangements play a significant role in the evolution of fish genomes, being important forces in the rise of multiple sex chromosomes and in speciation events. Repetitive DNAs constitute a major component of the genome and are frequently found in heterochromatic regions, where satellite DNA sequences (satDNAs) usually represent their main components. In this work, we investigated the association of satDNAs with chromosome-shuffling events, as well as their potential relevance in both sex and karyotype evolution, using the well-known *Pyrrhulina* fish model. *Pyrrhulina* species have a conserved karyotype dominated by acrocentric chromosomes present in all examined species up to date. However, two species, namely *P. marilynae* and *P. semifasciata*, stand out for exhibiting unique traits that distinguish them from others in this group. The first shows a reduced diploid number (with 2n = 32), while the latter has a well-differentiated multiple X_1X_2Y sex chromosome system. In addition to isolating and characterizing the full collection of satDNAs (satellitomes) of both species, we also in situ mapped these sequences in the chromosomes of both species. Moreover, the satDNAs that displayed signals on the sex chromosomes of *P. semifasciata* were also mapped in some phylogenetically related species to estimate their potential accumulation on proto-sex chromosomes. Thus, a large collection of satDNAs for both species, with several classes being shared between them, was characterized for the first time. In addition, the possible involvement of these satellites in the karyotype evolution of *P. marilynae* and *P. semifasciata*, especially sex-chromosome formation and karyotype reduction in *P. marilynae*, could be shown.

1. Introduction

Fishes are an incredibly diverse group with many chromosomal variations, including polyploidy, supernumerary chromosomes, distinct sex chromosome systems, and polymorphisms (Oliveira *et al*., 2009). In fact, most cases of postzygotic isolation are caused by genetic incompatibilities, among which chromosomal rearrangements play a fundamental role (Bateson, 1909; Dobzhansky, 1933; Muller, 1942). Chromosomal changes have the potential to limit introgression, thus facilitating the origin and maintenance of reproductive isolation through recombination suppression (Faria and Navarro, 2010; Šíchová *et al*., 2015). However, one of the most interesting evolutionary events refers to the emergence of neo-sex systems, when multiple sex chromosomes arise because of rearrangements between an autosome and a sex chromosome. This evolutionary step, also known as sex chromosome turnover, has the potential to suppress recombination next to breakpoints, creating new linkage groups between genes from distinct chromosomes, increasing the number of sex-linked genes, and accelerating the accumulation of genetic incompatibilities between populations (Vieira *et al*., 2003).

The impact of chromosomal rearrangements in fish karyotype evolution has been studied primarily from a cytogenetic point of view, with a particular emphasis on the chromosomal mapping of repetitive DNA sequences. The latter has proven to be a valuable source of information on the role of such sequences in genome organization and evolution (Biémont and Vieira, 2006; Khosraviani *et al*., 2019). Satellite DNAs (satDNAs) are one of the most common repeated sequences, forming extensive arrays of largely similar repeating units (monomers) that make up a significant percentage of genomes (reviewed in Šatović-Vukšić and Plohl, 2023). Recently, given the integration of cytogenetics with high-throughput sequencing data from next-generation sequencing methods (NGS), the whole collection of different satDNA families (satellitome) of several species has been characterized, providing insights into several evolutionary issues, such as karyotype evolution, genome diversity, and phylogenetic relationships (Ruiz-Ruano *et al*., 2016; Ferretti *et al*., 2020; Bardella *et al*., 2020; Cabral-de-Mello *et al*., 2021; Montiel *et al*., 2021; Vozdova *et al*., 2021; Cabral-de-Mello *et al*., 2023; Flynn *et al*., 2023; Gržan *et al*., 2023; Peona *et al*., 2023). These satellites (satDNAs) are also thought to play a role in the evolution and structure of sex chromosomes, as well as chromosome-based speciation (Henikoff *et al*., 2001; O'Neill *et al*., 2004; Plohl *et al*., 2012; Weissensteiner and Suh, 2019; Ferretti *et al*., 2020; Shatskikh *et al*., 2020; Flynn *et al*., 2023).

Pyrrhulina Valenciennes 1846 (Characiformes, Lebiasinidae) is the most diverse genus of the fish subfamily Pyrrhulininae, with 19 valid species (Frecke *et al*., 2023). Many species remain unexplored due to their small sizes and, thus, difficult sampling; consequently, the genus

presents unsolved taxonomic issues (Netto-Ferreira and Marinho, 2013). Recent research, however, has led to a better understanding of the evolution of Lebiasinidae species, including several *Pyrrhulina* ones, particularly from a cytogenetic and molecular genetics standpoint (Benzaquem *et al*., 2015; Moraes *et al*., 2017; Moraes *et al*., 2019; Sassi *et al*., 2019; Toma *et al*., 2019; Sember *et al*., 2020; Liu *et al*., 2020; Moraes *et al*., 2021; dos Santos *et al*., 2022). In general, about half (i.e., nine out of the 19) *Pyrrhulina* species have been cytogenetically documented, demonstrating a quite conserved diploid chromosome number, ranging from 40 to 42 chromosomes with karyotypes predominantly formed by acrocentric chromosomes (Moraes *et al*., 2017; Moraes *et al*., 2019; Moraes *et al*., 2021; Froese and Pauly, 2023). Aside from distinctive karyotypes and diploid numbers, genomic content comparisons among all analyzed species reveal a significant degree of similarity between their genomes, with most of the variations related to their repetitive content (Moraes *et al*., 2021). Most species have multiple 5S rDNA and 18S rDNA sites, with some species having a syntenic arrangement of these rDNAs (Moraes *et al*., 2017; Moraes *et al*., 2019; Moraes *et al*., 2021). Among all *Pyrrhulina* species, *P. marilynae* and *P. semifasciata* stand out for exhibiting characteristics that distinguish them from other species in the genus. In the first case, *P. marilynae* shows a significant karyotypic reduction, presenting 2n = 32 chromosomes with four metacentric pairs not observed in other species, presumably due to secondary fusions (Moraes *et al*., 2021). *P. semifasciata*, on the other hand, contains the sole morphologically differentiated sex chromosome system found in the genus, i.e., the multiple $X_1X_1X_2X_2/X_1X_2Y$ system (Moraes *et al*., 2019).

In this study, we selected those two *Pyrrhulina* species that underwent substantial, cytogenetically visible chromosomal rearrangements to examine the involvement of satDNAs in these chromosome-shuffling events and their putative role in both sex and karyotype evolution. Apart from performing a comprehensive analysis of their satellitomes, the satDNAs located on *P. semifasciata*'s sex chromosomes were also mapped in two phylogenetically related species (*P. brevis* e *P. obermulleri*) to check their possible accumulation on proto-sex chromosomes.

2. Materials and Methods

2.1. Material, Mitotic Chromosomes and DNA Sequencing

Samples of four *Pyrrhulina* species were collected in Brazilian rivers, according to **Table 1**. To obtain mitotic metaphase chromosomes, the animals were euthanized with eugenol, as approved by the Ethics Committee on Animal Experimentation of the Universidade Federal de São Carlos, Brazil (Process number CEUA 7994170423), and cell suspensions were obtained

from kidney cells (Bertollo, 1978). For DNA sequencing, we selected *P. marilynae* (one male) and *P. semifasciata* (one male and one female) individuals. Then, total DNA was extracted from muscle tissues using a spin-column-based protocol (Cellco Biotech, São Carlos–SP, Brazil). The purified DNAs were then sequenced in the BGISEQ-500 platform (2 × 150 bp; BGI Shenzhen Corporation, Shenzhen, China). Short-read sequencing yielded between 2.14 Gb (in *P. semifasciata* female) and 2.44 Gb (in *P. marilynae*). Raw reads were deposited in the sequence read archive (SRA-NVBI) and are available under the accession numbers (SRR25467276, SRR25476502, and SRR25476501).

Table 1. Species, locality, and number of individuals (N) used in the present study.

Species	Locality	N	Voucher
P. brevis	Adolfo Ducke Reserve- Igarapé Barro Branco, Manaus - AM (2°56'04.6"S 59°58'10.6"W)	04♂; 07♀	MZUSP 123077
P. marilynae	Ipiranga do Norte - MT (11°36'02"S 55°56'27"W)	13♂; 04♀	UFPB 12080
P. obermulleri	Tefé- AM (3°25'50.7"S 64°44'54.8"W)	06♂; 04♀	UFPB 12079
P. semifasciata	Adolfo Ducke Reserve- Igarapé Barro Branco, Manaus - AM (2°56'04.6"S 59°58'10.6"W)	07 ♀; 12 ♂	MZUSP 123080

2.2. Bioinformatic Analyses and satDNA Library

Firstly, we trimmed the raw reads with Trimmomatic (Bolger *et al*., 2014) to select the pair-end reads with Q > 20 for all nucleotides. Then, the satDNA catalogs from *P. semifasciata* and *P. marilynae* were independently characterized on TAREAN (Novák *et al*., 2020) with the satMiner pipeline (Ruiz-Ruano, 2016). Specifically, the consensus satDNA sequences outputted in TAREAN were filtered from the genomic libraries with DeconSeq (Schmieder and Eswards, 2011), and subsequent iterations were performed on TAREAN until no satDNA was found. Then, a homology search with RepeatMasker (Smith *et al*., 2020) was performed to group the sequences into variants, families, and superfamilies, as suggested by (Ruiz-Ruano, 2016). We also calculated the abundance and divergence values of the satDNA families by selecting 10,000,000 reads (2 × 5,000,000) from each genomic library and masking against their own catalog of satDNAs with RepeatMasker (Smith *et al*., 2020). After that, we named the satDNA families according to their abundance in each species.

Since *P. semifasciata* exhibits a multiple-sex chromosome system, we calculated the male: female (M/F) abundance ratio of each satDNA family in this species (quotient between the

abundance of the same satellite DNA in male and female). Then, we selected those with an M/F ratio greater than 1.2 as putatively sex-specific accumulating satDNAs. Finally, both satellitomes were BLAST-searched (Altschul *et al*., 1990) against the NCBI nucleotide collection to check for the presence of conserved satDNAs. We constructed a minimum spanning tree (MST) for PseSat55 using PHYLOVIZ (Nascimento *et al*., 2017) to describe the proportions of the male and female haplotypes.

2.3. Primer Design and DNA Amplification via Polymerase Chain Reaction (PCR)

We designed primers for 21 out of the 70 PseSatDNAs and for 10 of the 71 PmaSatDNAs that were characterized. As a criterion for primer selection, we selected the ten most abundant ones (in *P. marylinae*, **Table S1**), the five most abundant, and those with some difference in abundance between sexes (in *P. semifasciata*, **Table S2**). PCR procedures used the optimal amplification temperatures and DNA template concentrations for each satDNA, according to (Ruiz-Ruano, 2016). For each sequence, the following cycles were used: initial denaturation at 95 °C for 5 min; 29–35 cycles with denaturation at 95 °C for 15 s; annealing at 50 °C to 62 °C for 30 s (**Table S3**); extension at 72 °C for 10 s, and final extension at 72 °C for 10 min. To validate the amplification and ensure the integrity of the satDNAs, the PCR products were checked by electrophoresis on 2% and 1% agarose gels. Finally, they were quantified using the NanoDrop spectrophotometer (ThermoFisher Scientific, Branchburg, NJ, USA).

2.4. Fluorescence In Situ Hybridization (FISH)

The probes derived from the satDNA's PCRs were labeled with Atto550-dUTP or Atto488-dUTP by Nick-Translation (Jena Biosciences, Jena, Germany) and used for FISH experiments. The hybridization mixes were composed of 100 ng of each labeled satellite DNA plus 50% formamide, 2 × SSC, 10% SDS, 10% dextran sulfate, and Denhardt's buffer at pH 7.0 in a total volume of 20 µL, following high-stringency conditions for FISH (Pinkel *et al*., 1986). We hybridized the above-mentioned selected sequences in both species (i.e., *P. marilynae* and *P. semifasciata*), then selected those satDNAs that displayed positive signals on the sex chromosomes of *P. semifasciata* to hybridize in *P. brevis* and *P. obermulleri* chromosomes. For the satDNA FISH experiments, glass slides containing metaphase chromosomes were aged for 1 h at 60 °C, following a treatment at 37 °C for 5 min with 0.005% pepsin solution (99 µL H2O, 10 µL HCl, and 2.5 µL pepsin (20 mg/mL). Chromosomes were denatured in 70% formamide/2 × SSC at 72 °C for 3 min, while probes were at 85 °C for 10 min, then cooled at 4 °C for 2 min before application. Hybridization occurred overnight in a moist chamber at 37 °C. Next, the

slides were washed for 5 min in 1 × SSC at 65 °C and 4 × SSC/Tween at room temperature, following a quick wash in 1 × PBS for 1 min. The slides were dehydrated in an ethanol row (70%, 85%, and 100%) before the counterstaining of the chromosomes with DAPI mounted in Vectashield (Vector Laboratories, Burlingame, USA). We also used whole-chromosome painting (WCP) using a Y-specific probe (Psemi-Y) previously obtained (Moraes *et al*., 2019) to detect the sex chromosomes of *P. semifasciata* and the proto-sex pairs in *P. brevis* and *P. obermulleri*. For this, sequential FISH/WCP was performed following (Sassi *et al*., 2022). In total, 10 PseSatDNAs showed visible FISH signals in *P. semifasciata*, and 07 PmaSatDNA showed a visible FISH signal in *P. marilynae*.

2.5. Images and Confirmation of Results

To confirm the FISH results, we analyzed a minimum of 30 metaphase spreads per individual. Images were captured with CoolSNAP on an Axioplan II microscope (Carl Zeiss Jena GmbH, Jena, Germany) and processed with ISIS (MetaSystems Hard e Software GmbH, Altlussheim, Germany).

3. Results

3.1. SatDNA Content of P. marilynae and P. semifasciata

We applied the satMiner pipeline using short-read libraries of *P. semifasciata* (female) and *P. marilynae* (for more detailed information, see the material and methods section). After three iterations of each of the satMiner protocols, we found 70 and 71 satDNA families for *P. marilynae* (Pma) and *P. semifasciata* (Pse), respectively. The repeat unit lengths ranged from 23 to 4663 bp, with a median of 443.5 bp for *P. marilynae*, and from 6 to 2510 bp, with a median of 39 bp in *P. semifasciata*. In *P. marilynae*, the A + T content of satDNAs ranged from 39.2 to 71.8% with a mean of 60%, whereas in *P. semifasciata*, it ranged from 39.2 to 78.5% with a mean of 60%. In total, 64 and 77 satDNAs in *P. marilynae* and *P. semifasciata*, respectively, had an A + T content of more than 50%. Long satDNAs (>100 bp sensu (Ruiz-Ruano *et al*., 2012) were predominant in both satellitomes, with 39 and 44 satDNA families in *P. marilynae* and *P. semifasciata*, respectively. The complete results for each satellitome are described in Tables S1 and S2. Sequences are available on the NCBI-Genbank, under the accession numbers OR094701-OR094771 (*P. semifasciata*) and OR094772-OR094841 (*P. marilynae*).

3.2. *Chromosomal Distribution of PmaSatDNA in P. marilynae*

To examine the chromosomal location and distribution of the PmaSatDNAs, we used both female and male mitotic metaphase plates of *P. marylinae* in our two-color fluorescence in situ hybridization (FISH) experiments. Within 10 successfully amplified satDNAs families, six of them produced visible FISH signals, yielding the same result in both sexes (**Figure 1**). PmaSat04, PmaSat07, and PmaSat10 were mostly found in the telomeric and centromeric regions of most chromosomes, with the presence of bitelomeric signals for PmaSat04 and PmaSat10 (**Figure 1**). In addition, PmaSat06, PmaSat07, PmaSat09, and PmaSat10 were observed in the pericentromeric regions of some chromosomes (**Figure 1**). The satDNAs PmaSat02, PmaSat03, and PmaSat08 did not produce any FISH signal.

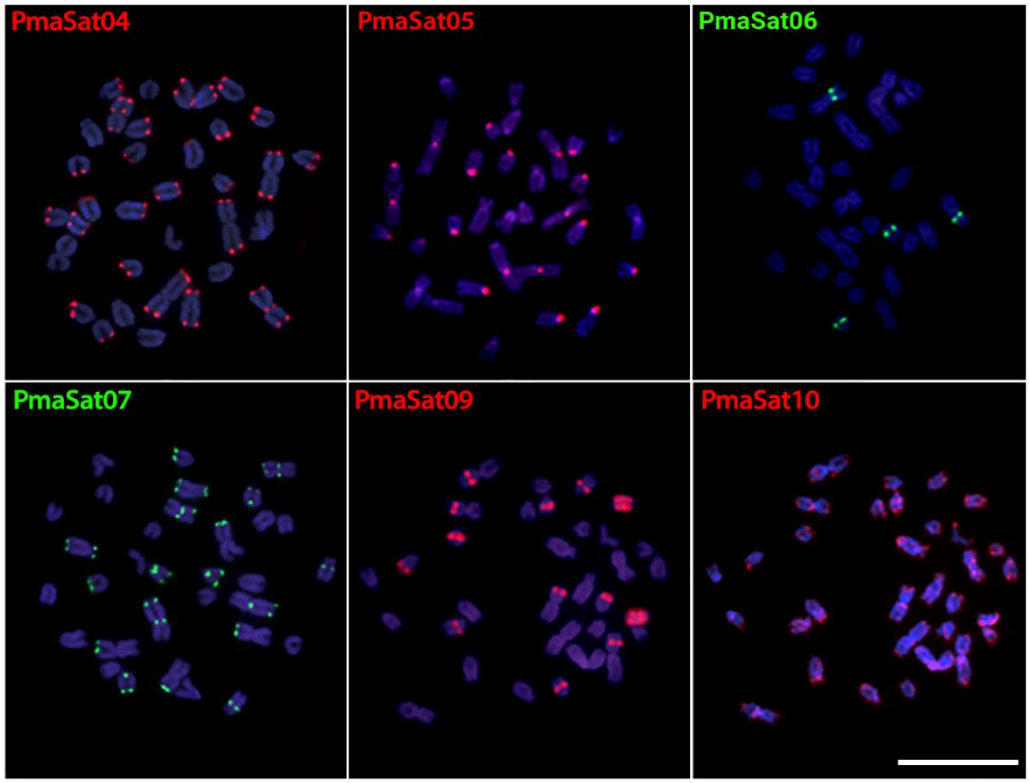


Figure 1. Metaphase plates of *Pyrrhulina marilynae* highlighting the chromosomal location PmaSatDNAs. The satDNA family names are indicated on the left top, in red (Atto550-labeled) or green (Atto488-labeled). Scale bar = 10 µm.

3.3. *Chromosomal Distribution of PseSatDNA in P. semifasciata*

To examine the chromosomal location and distribution of PseSatDNAs, we used both female and male mitotic metaphase plates of *P. semifasciata* in the same two-color FISH sets as before for *P. marilynae*. Within the 16 successfully amplified satDNA families, ten produced visible FISH signals, yielding the same result in both sexes, except the ones located on the sex chromosomes. Most of the analyzed PseSatDNAs were found in the centromeric and

pericentromeric regions of the autosomal chromosomes (**Figure 2**). PseSat01, PseSat04, PseSat38, and PseSat55 hybridized in the autosomes and sex chromosomes of *P. semifasciata* (**Figure 2**). PseSat01, the most abundant satellite DNA, was located on two autosomes, the X2 and the Y chromosome (**Figure 2** and **Figure S1**). The sequences PseSat06, PseSat32, PseSat39, PseSat57, PseSat61, and PseSat67 did not produce any visible FISH signals.

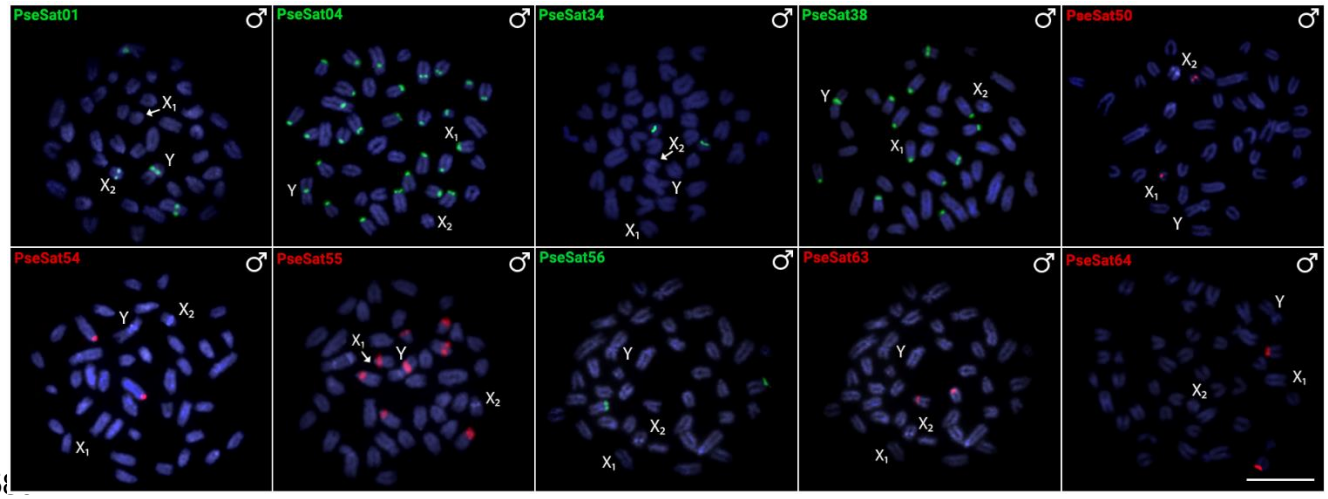


Figure 2. Male metaphase chromosomes of *Pyrrhulina* semifasciata after FISH with 10 PseSatDNAs. The satDNA family names are indicated in the top left corner in red (Atto550-labeled) or green (Atto488-labeled). The sex chromosomes X1, X2, and Y are indicated. Scale bar = 10 µm.

In all experiments, a second FISH experiment using PseSat01 and/or the Y-specific PSEMI-Y probe was carried out to accurately identify the X_1, X_2, and Y sex chromosomes (**Figure 3**).

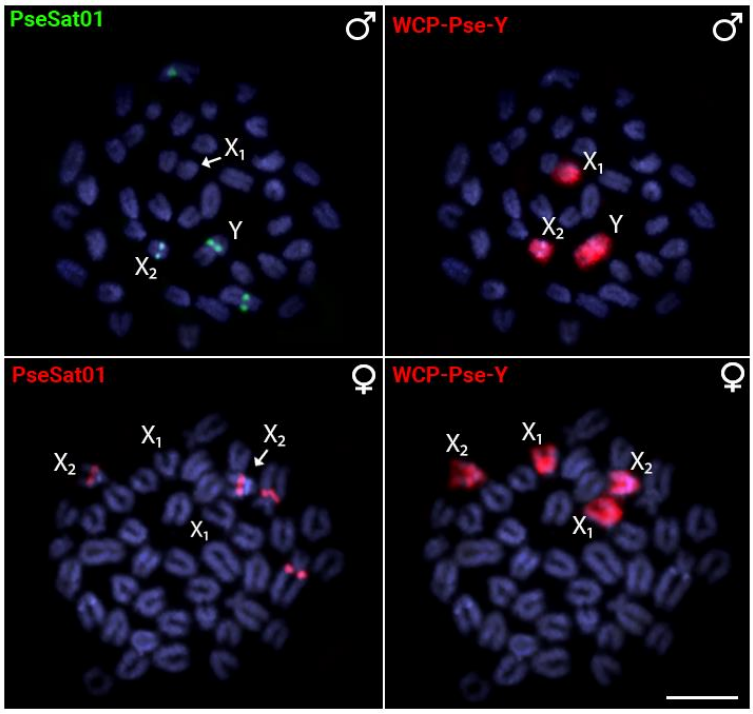


Figure 3. Male and female metaphase plates of *Pyrrhulina* semifasciata showing that the hybridization

pattern of PseSat01 (first column) is coincident with the X2 and Y sex chromosomes, as indicated by the whole-chromosome painting with the PSEMI-Y probe (second column), which is derived from the microdissection of the Y chromosome. Scale bar = 10 µm.

3.4. Chromosomal Distribution of PseSatDNA in Other Pyrrhulina Species

All PseSatDNAs located on the sex chromosomes of *P. semifasciata* (i.e., PseSat01, PseSat04, PseSat38, and PseSat55) were also hybridized against *P. obermulleri*, *P. brevis*, and *P. marilyanae* metaphase chromosomes. All these satDNAs delivered evaluable results in all species (**Figure 4**), except for PseSat38, which was not visible in the chromosomes of *P. obermulleri* (**Figure 4c**). PseSat04 hybridized in the centromeric region of almost all chromosomes in *P. obermulleri* and *P. brevis* (**Figure 4b,f**), whereas in *P. marilynae*, visible signals were seen in the centromeric, pericentromeric, or telomeric regions on most chromosomes (**Figure 4j**). PseSat01, on the other hand, was present in only two chromosome pairs in *P. obermulleri* and *P. marilynae* (**Figure 4a,i**) and in nearly all chromosomes in *P. brevis* (**Figure 4e**). PseSat55 was found in the centromeric and telocentromeric regions of most *P. obermulleri* and *P. brevis* chromosomes (**Figure 4d,h**). In *P. marilynae*, on the other hand, it was mapped exclusively in the centromeric region of most chromosomes (**Figure 4l**). Except for PseSat38, which was not present in *P. obermulleri*, and PseSat01, which did not exhibit visible signals in *P. marilynae* proto-sex pairs, all three species had at least one pair of putative proto-sex pairs that exhibited positive signals for each of the selected PseSatDNAs (**Figure 5**). Again, in all slides, a second FISH experiment with the Y-specific PSEMI-Y probe was carried out to accurately identify the proto-sex chromosomes.

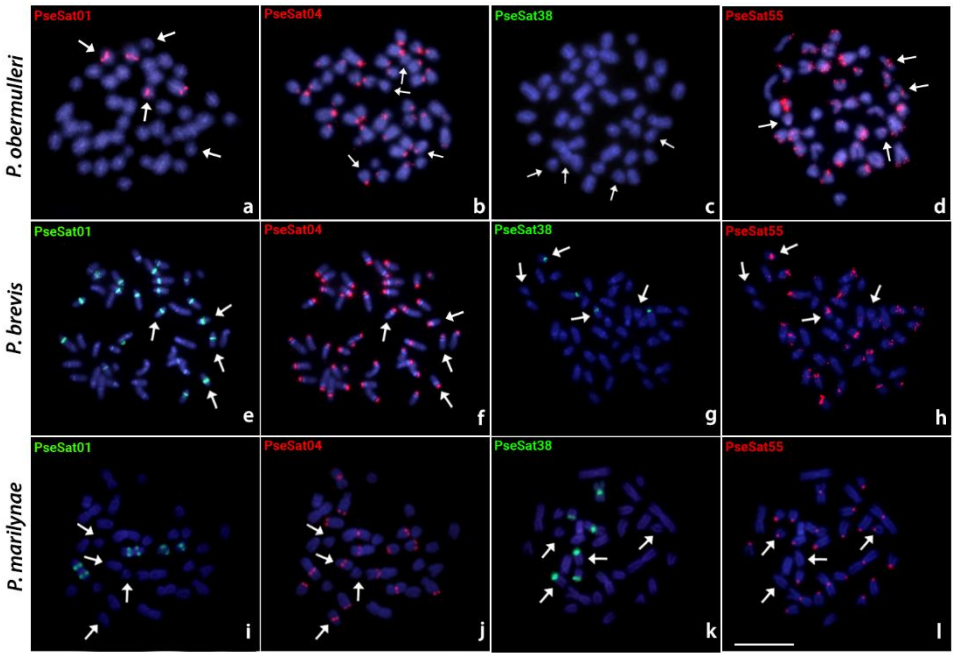


Figure 4. Metaphase plates of *Pyrrhulina* obermulleri (a–d), *P. brevis* (e–h), and *P. marilynae* (i–l)

highlighting the chromosomal location of PseSatDNAs that were mapped in the sex chromosomes of *P. semifasciata*. The satDNA family names are indicated in the upper left in red (if labeled with Atto550-dUTP) or green (if labeled with Atto488-dUTP). The arrows indicate the proto-sex chromosomes. Scale bar = 10 µm.

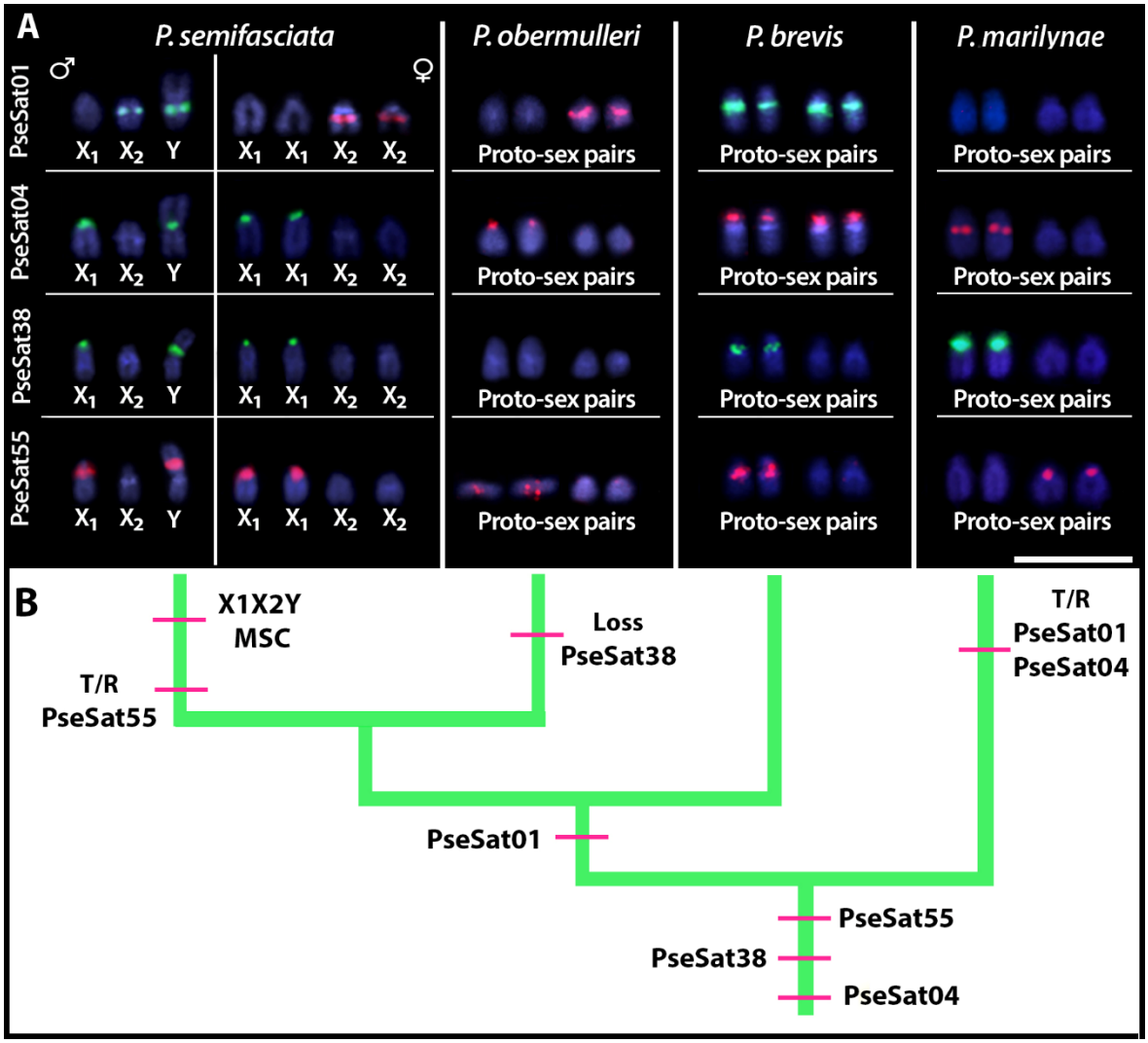


Figure 5. (A) *Pyrrhulina* semifasciata male (first column) and female (second column) sex chromosomes highlighting the hybridization pattern of the four PseSatDNAs that produced visible signals, with each line matching to a satellite sequence indicated on the left. The third, fourth, and fifth columns show the hybridization patterns of those satellites in the *P. obermulleri*, *P. brevis*, and *P. marilynae* chromosomes, respectively. Scale bar = 10 µm. (B) Phylogenetic relationships of *Pyrrhulina* species (green lines) based on Ferreira *et al*. (2022) plotted with main events of origin, loss, and T/R (transposition or recombination) of PseSatDNAs, and origin of the multiple sex chromosome system (MSC) (pink lines).

3.5. Minimum Spanning Trees: MSTs

We have chosen PseSat55 to generate minimum spanning trees (MST). The other satDNAs clustered in *P. semifasciata*'s sex chromosomes contain more than 150 bp, making it impossible to create an MST. In addition to the identical locations of the X1 and Y chromosomes, PseSat55 is found in the pericentromeric regions of six autosomes in both males and females (**Figure 6**).

Even among the less common haplotypes, this sharing is observed, with only a few sequences being exclusive, mostly in females. The MST results show that PseSat55 is homogeneous in both sexes, with just a few unique haplotypes, suggesting a certain degree of recombination between the X and Y chromosomes.

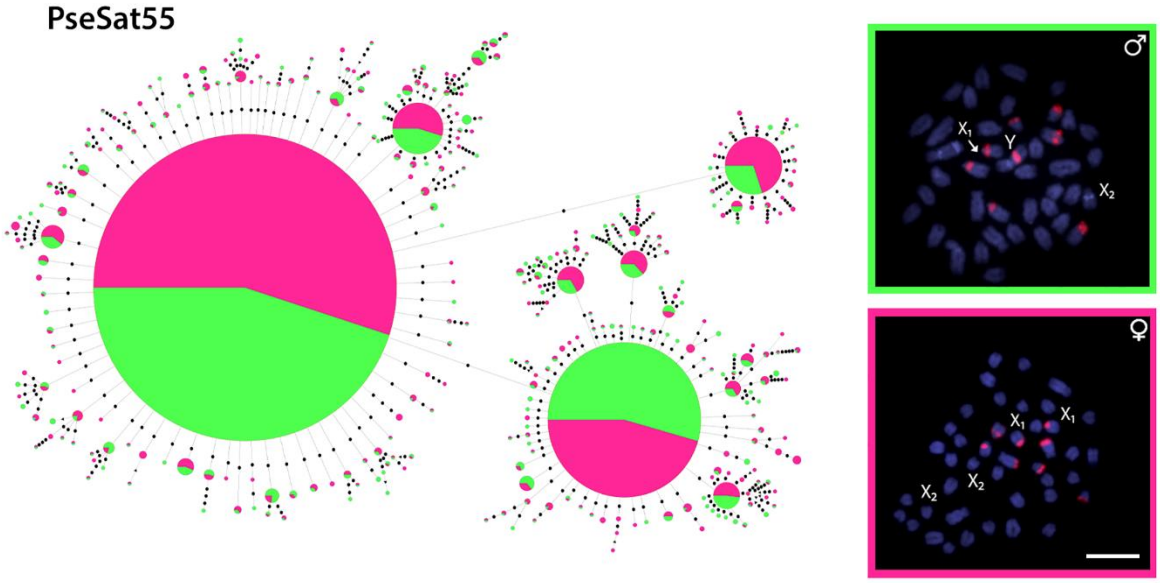


Figure 6. Linear MSTs of PseSat55 obtained from reads of females (pink) and males (green). The diameter of the circle is proportional to the abundance of the haplotype. Black circles represent 01 base of divergence between haplotypes, and the black triangles represent 05 bases of divergence. The in situ mapping in male and female metaphase plates is highlighted in boxes. Bar = 5 µm.

4. Discussion

4.1. General Features P. marilynae and P. semifasciata Satellitomes

Here, we show that a significant proportion of satDNA families (38 satDNAs) are shared between the satellitomes of *P. marilynae* (70 satDNAs) and *P. semifasciata* (71 satDNAs) (**Table S4**). Such a scenario (i.e., the conservation and sharing of satDNA families) is not an exclusive feature of *Pyrrhulina* but is also observed in other fish species (Silva *et al*., 2017; dos Santos *et al*., 2021; Goes *et al*., 2023), as well as several vertebrates such as snakes (Lisachov *et al*., 2023), primates (Ahmad *et al*., 2020), and true toads (Guzmán-Markevich *et al*., 2022). According to the library hypothesis (Fry and Salser, 1977), satellite sequences are preserved over long evolutionary timescales because closely related species share a common library of these sequences, with quantitative changes brought on by differential amplification. Investigations suggested that the satDNA libraries can disappear or be formed de novo following cladogenetic events (Camacho *et al*., 2022). According to this theory, the acquisition of a biological function that will ultimately be preserved by natural selection is necessary for conservation over a lengthy evolutionary time. Although both species analyzed retained about half of their satDNA families,

evidencing a common library between them, the observed differences in the abundance of these shared families are also remarkable, a fact predicted by the library hypothesis (Fry and Salser, 1977). Among the 38 shared satDNAs, six of them were selected for FISH experiments (**Table S4**) and revealed the same accumulation pattern in both species.

The heterochromatic regions differ substantially between *Pyrrhulina* species, with large accumulations in centromeric and telomeric regions, as well as pericentromeric and interstitial regions (Moraes *et al*., 2017; Moraes *et al*., 2019; Moraes *et al*., 2021). The karyotype of *P. semifasciata* has considerable accumulations of heterochromatin in centromeres and telomeres (Moraes *et al*., 2019), whereas *P. marilynae* has these accumulations mainly in the centromeric region of most chromosomes (Moraes *et al*., 2021). In the same way, the *in situ* mapping revealed a predominance of PseSatDNAs in the centromeric regions of *P. semifasciata* (**Figure 2**). This type of association can be suggestive of a probable relationship with centromere development, as well as an essential role in genome integrity by preserving the higher-level nucleus structure (Podgornaya, 2022). Although satellite DNA sequences are often abundant in heterochromatic regions (Thakur *et al*., 2021; Šatović-Vukšić and Plohl, 2023), as indicated in *P. semifasciata*, their occurrence in euchromatic regions has been emphasized in many species, including bivalves (Tunjić-Cvitanić *et al*., 2021), insects (Cabral-de-Mello *et al*., 2021; Pereira *et al*., 2021; Flynn *et al*., 2023), mammals (Valeri *et al*., 2021), and fishes (Silva *et al*., 2017; dos Santos *et al*., 2021; Crepaldi *et al*., 2021), as well as *P. marilynae* (**Figure 1**). PmaSat01 constitutes approximately 4% of the *P. marilynae* genome (**Table S1**), while the most abundant satellite of *P. semifasciata* (PseSat01) only represents about 0.5% of its genome (**Table S2**). However, the satellitomes of both species are very similar in the number of clusters recovered by TAREAN (i.e., 70 PmaSatDNAs and 71 PseSatDNAs). Among fishes, other Characiformes species, such as representatives of *Astyanax*, *Characidium*, and *Psalidodon*, also carry similar numbers of satDNA families (Serrano-Freiras *et al*., 2020; Goes *et al*., 2023). However, species with more satDNAs have previously been identified, such as in *Megaleporinus*, demonstrating a great dynamic of such sequences in fish genomes. Owing to the various processes involved in their dynamics throughout cladogenetic genomic development, the variety of satellitomes among different species of animals and plants is rather significant, both in terms of the number of satDNA families and the proportion of the genome they occupy (Garrido-Ramos, 2017; Šatović-Vukšić and Plohl, 2023).

4.2. Satellite DNA Contribution in the Significant 2n Reduction Observed in P. marilynae

P. marilynae stands out for presenting the most rearranged karyotype among all *Pyrrhulina* species, with 2n = 32 and formed by large and unusual metacentric pairs (Moraes *et al*., 2021). Many studies have found that satellite DNAs can induce chromosomal rearrangements and thus have a direct impact on karyotype evolution due to their dynamic and fast-evolving nature (Paço *et al*., 2013; Vieira-da-Silva *et al*., 2015; Gatto *et al*., 2018; Escudeiro *et al*., 2019; de Lima and Ruiz-Ruano, 2022). Thus, the occurrence of centric fusions and fissions is frequently associated with the fast dynamics of satDNAs, which are located mostly in the centromeric and pericentromeric regions of chromosomes (Slamovits *et al*., 2001; Ugarković and Plohl, 2014; Kopecna *et al*., 2014; Vozdova *et al*., 2019). Interestingly, our experiments demonstrate that PmaSatDNAs hybridized at telomeres or centromeres of *P. marilynae* chromosomes, including bitelomeric signals from PmaSat04 and PmaSat10, and pericentromeric signals on both arms with PseSat01 (**Figure 1**). These results imply that chromosomal reduction in *P. marilynae* is directly linked to centric fusion processes that result in the large metacentric chromosomes found in this species' karyotype. Notably, the satellite DNA sequences did not experience any accumulation or loss after the fusion process.

PmaSatDNAs 04, 07, 09, and 10 accumulate in telomeres and (peri)centromeres, especially in the large metacentric pairs. These large chromosomes probably originate from chromosome fusions, restricting recombination to the distal ends (Franchini *et al*., 2017; Franchini *et al*., 2020; Vara *et al*., 2021). Repetitive sequences, such as satellites and rDNAs, often make up the subtelomere region, also known as the buffer zone between the internal chromosome and telomere. Given that the latter is made up of telomeric motifs interspersed with other repeated sequences (Naish *et al*., 2021), such subtelomere sequences can also be found in interstitial telomeric sites (ITSs). As sub and telomeric arrays help to stabilize new ends (Yang *et al*., 2005), it is likely that the FISH signals observed directly represent the fusion process that *P. marilynae* went through throughout its karyotypic reduction (**Figure 1**). The expansion and contraction of those telomere-associated satellite motifs that remained even after the inactivation of the telomeric region near the newly formed centromere may be fostered by repair mechanisms (Kipling *et al*., 1991). This scenario is not exclusive to *P. marilynae*. Other fish species also carry a series of repetitive DNA in their chromosome fusion points, as observed in *Rineloricaria* (Glugoski *et al*., 2022; Glugoski *et al*., 2023), for example.

SatDNAs are a significant and prominent component of the so-called "dark matter of genomes" (Sedlazeck *et al*., 2018). In fact, multiple pieces of evidence show that satellite DNAs

are sequences that can participate in centromere and telomere formation besides presenting fundamental roles and specific functions in the genome (Garrido-Ramos, 2017; Louzada *et al*., 2020; dos Santos *et al*., 2021; Montiel *et al*., 2022). Although no functional experiments have been performed here, the majority of the PmaSatDNAs are found in centromeric and telomeric regions and may be directly linked to their formation, in addition to possibly playing a role in *P. marilynae*'s reduction and karyotypic evolution.

4.2. SatDNAs and the Evolution of Multiple Sex Chromosomes

Repetitive sequences are great tools for the study of sex chromosomes, where species with heteromorphic sex chromosomes show a difference in the accumulation of some sequences between males and females, emphasizing the existence of several W/Y-specific satDNAs (Ferretti *et al*., 2020; Sember *et al*., 2021; Cabral-de-Mello *et al*., 2023; Flynn *et al*., 2023). Despite the presence of a well-differentiated X1X2Y sex chromosome system, minor differences in haplotype accumulation between males and females were observed (**Figure 6**). Although four PseSatDNAs were mapped in this study to either X1, X2, or neo-Y chromosomes, no neo-Y-specific satDNA was identified (**Figure 5**), contrasting with *Eneoptera surinamensis* (Palacios-Gimenez *et al*., 2017) and *Ronderosia bergii* (Ferretti *et al*., 2020), which show a large accumulation of satDNAs in neo-sex chromosomes. These species, however, lack synapses between sex chromosomes, demonstrating significant differentiation between them. Multiple-sex chromosome systems, in contrast to simple ones, are known to have a more recent origin and still exhibit considerable recombination rates, resulting in few or no sex-specific sequences (Sember *et al*., 2021). A similar scenario has already been reported in other animal species, such as the frog *Proceratophrys boiei*. Despite having a simple and heteromorphic ZZ/ZW sex chromosome system, it is distinguished by the absence of sex-specific satDNAs and low sex divergence, indicating their early stage of sex chromosome differentiation (da Silva *et al*., 2023).

There are no dated phylogenetic reconstructions for *Pyrrhulina* species. As a result, determining when the multiple-sex chromosome system emerged is not feasible, regardless of other cytogenetic traits that suggest its recent origin. The X1, X2, and Y sex chromosomes lack strong differences in size and accumulation of heterochromatic regions (Moraes *et al*., 2019), as in the satellite distribution. Accordingly, the MST generated for PseSat55 also did not point to any difference in the abundance of the haplotypes obtained from reads of females and males (**Figure 6**). Such low heterochromatin is also observed in other species with recent diversification of multiple sex chromosomes, such as Hoplias malabaricus and *Erythrinus erythrinus* (Bertollo *et al*., 1997; Bertollo *et al*., 2004). Considering the formation of a trivalent

during the meiosis process, multiple-sex chromosome systems (such as the X1X2Y of *P. semifasciata*) could not accumulate large heterochromatic blocks; otherwise, they could impair its correct segregation (Cioffi *et al*., 2012). In this sense, the presence of PseSatDNAs in the homologous regions of the multiple sex chromosomes of *P. semifasciata* can be indicative of their role in the modulation of gene expression. Under stress circumstances, euchromatic copies of pericentromeric satDNAs in *Tribolium castaneum* are functionally relevant in modifying chromatin and the expression of adjacent genes (Feliciello *et al*., 2015). In Drosophila melanogaster, more than a thousand euchromatic copies of satDNAs are mainly found near genes and are thought to have a function in the modulation of gene expression (Kuhn *et al*., 2012). Similarly, those euchromatin-dominant satellite DNAs share characteristics regarding their structure, organization, and evolution (reviewed in Šatović-Vukšić and Plohl (2023) and Plohl *et al*. (2008).

Despite the absence of neo-Y-specific satDNA sequences in *P. semifasciata*, we demonstrate that several motifs are shared between this species' sex chromosome and the proto-sex pairs in other *Pyrrhulina* representatives (**Figure 5**). Considering the absence of specific sequences such as PseSat38 in *P. obermulleri*, it is also noteworthy to emphasize the high dynamism of these sequences in the cladogenetic history. Both a seemingly complete absence of these sequences and their low-copy numbers can be explained by the absence of detectable FISH signals. Unequal crossing over between sister chromatids or intra and interchromosomal recombination may significantly decrease the copy number of satellite DNAs (Pralhongcheep *et al*., 2017). On the other hand, sequences such as PseSat01, PseSat04, and PseSat55 suffered expansions and retractions in the number of loci, including changes between both proto-sex pairs. In some instances, site-directed recombination between homologous motifs in satellite repeats and in the target genomic sequence, most likely mediated by extrachromosomal circular DNA, can explain the dispersion of satellites (Feliciello *et al*., 2020). Alternatively, such translocation can be due to transposable elements (TEs) that are known to participate in the origin and the dispersal of satDNA repeats (Zattera and Bruschi, 2022). Through amplifying tandem repeats within them, TEs may also capture satDNA motifs and generate new satDNAs, which can then be disseminated over other regions of the genome and preserved through nonreciprocal transfer mechanisms such as unequal crossover (Scalvenzi and Pollet, 2014). Indeed, TEs such as Rex3 were already mapped in *Pyrrhulina* chromosomes, showing a high dispersal tendency (Moraes *et al*., 2017). However, the extremely similar size and shape of those chromosomes, as well as the frequent translocations between them, might also be the result of putative pseudo-homologous regions (PHRs), which are thought to be due to recombination

between non-homologous sequences (Guarracino *et al*., 2023) or due to the presence of pseudoautosomal regions (PARs) with high recombination rates (Yazdi *et al*., 2023) in these proto-sex chromosomes.

It is noteworthy that some conserved PseSatDNAs across distinct *Pyrrhulina* species (**Figure 4**) are present in the pericentromeric region of the sex chromosomes in *P. semifasciata* and the proto-sex pairs of *P. brevis*, *P. obermulleri*, and *P. marilynae*. From the canonical model of sex chromosome evolution, when the recombination between the sex-related chromosomes is ceased, the sex-specific chromosome is invaded by several repetitive sequences, including satDNAs (reviewed in Kratochvíl *et al*. (2023). However, the presence of such satDNAs in the proto-sex chromosomes leads us to two main questions: i) Can the accumulation of the same PseSatDNAs (i.e., PseSat01, PseSat04, PseSat38, and PseSat55) in proto-sex and multiple-sex chromosomes indicate the presence of a plesiomorphic and homomorphic XX/XY in *P. brevis*, *P. obermulleri*, and *P. marilynae*? Or ii) Were those sequences simply conserved in these acrocentric pairs, given their putative role in the centromeric structure and genome integrity? To fully solve these questions, however, a high-quality reference genome is mandatory, and the next steps involve generating a high-quality genome assembly with PacBio having access to an end-to-end solution.

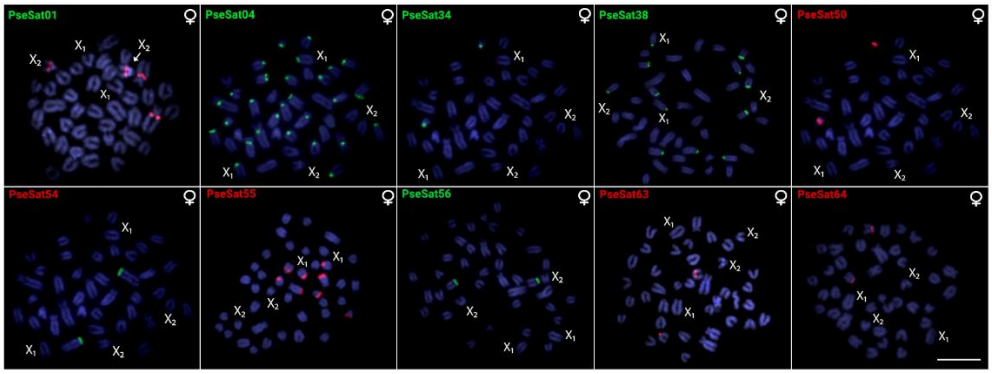
References

All references are compiled in the end of this thesis.

Supplementary Materials

The following supporting information can be downloaded at https://www.mdpi.com/1422-0067/24/17/13654/s1.

Figure S1: Females metaphase plates of *Pyrrhulina semifasciata* highlighting the chromosomal location of 10 satDNA. The satDNA family names are indicated on the left top, in red (ATTO550 labeled) or green (ATTO488 labeled). The sex chromosomes (X_1X_2Y) are shown. Scale bar = 10 μm.



Supplementary Table 1. Main characteristics of 71 satDNAs found in *Pyrrhulina marilynae*. The sequences highlighted in blue correspond to those selected for the FISH experiments.

Família de satDNAs	RUL	Abundância	Divergência	A+T (%)
PmaSat01-1627	1627	0.0395879278883	1.33	58.6
PmaSat02-68	68	0.0126290012853	2.63	66.2
PmaSat03-45	45	0.0083896090511	2.92	44.4
PmaSat04-50	50	0.00772308401902	3.93	60
PmaSat05-226	226	0.00728803566553	9.69	64.2
PmaSat06-198	198	0.00647761332963	9.25	67.2
PmaSat07-45	45	0.00639111613516	5.97	62.2
PmaSat08-33	33	0.00268802096213	2.10	57.6
PmaSat09-335	335	0.00220671410467	9.24	70.1
PmaSat10-4663	4663	0.00209498220866	5.93	43.6
PmaSat11-2486	2486	0.00207120592574	10.40	62.4
PmaSat12-842	842	0.00204339498168	9.60	63.5
PmaSat13-4322	4322	0.00196875721573	13.26	63.7
PmaSat14-54	54	0.00185443026527	3.79	57.4
PmaSat15-1437	1437	0.0015994228488	8.96	58.1
PmaSat16-670	670	0.000982230583972	7.52	59.4
PmaSat17-192	192	0.000746618253954	10.82	63
PmaSat18-42	42	0.000666838894057	4.21	64.3
PmaSat19-84	84	0.000628643771793	13.99	53.6
PmaSat20-142	142	0.000590982313911	3.14	66.2
PmaSat21-712	712	0.000582114970286	1.67	62.2
PmaSat22-32	32	0.000555307988527	3.71	53.1
PmaSat23-23	23	0.000553234717251	7.65	52.2
PmaSat24-1283	1283	0.000550808970056	13.84	56.5
PmaSat25-36	36	0.000541725784436	5.34	44.4
PmaSat26-21	21	0.000528704967553	17.33	47.6
PmaSat27-39	39	0.000519611880924	8.11	64.1
PmaSat28-165	165	0.000493380147786	7.48	66.7
PmaSat29-51	51	0.000480608836328	7.36	39.2
PmaSat30-1062	1062	0.00047705239391	8.96	64.8
PmaSat31-47	47	0.000450014718642	5.81	59.6
PmaSat32-33	33	0.000425481998641	3.84	60.6
PmaSat33-176	176	0.00040900473953	4.45	64.8
PmaSat34-1411	1411	0.000394930455293	0.99	58
PmaSat35-845	845	0.000394410652323	3.59	62.7
PmaSat36-445	445	0.00032788478302	4.71	57.1
PmaSat37-28	28	0.000311260988976	9.21	60.7
PmaSat38-162	162	0.000284214402801	4.71	66
PmaSat39-88	88	0.000276674784478	8.15	68.2

PmaSat40-211	211	0.000276362902695	6.43	59.2
PmaSat41-915	915	0.000237961839444	1.67	67.4
PmaSat42-314	314	0.000236262826307	5.27	58.3
PmaSat43-165	165	0.00023088459824	4.09	64.8
PmaSat44-43	43	0.000227744988298	11.66	48.8
PmaSat45-568	568	0.000214584567189	5.78	59
PmaSat46-470	470	0.000206393462476	5.08	61.3
PmaSat47-672	672	0.000203039000641	3.93	58.5
PmaSat48-31	31	0.000193220170055	5.51	67.7
PmaSat49-80	80	0.00017571122581	17.60	67.5
PmaSat50-428	428	0.000172995379053	7.23	57.2
PmaSat51-414	414	0.000172949834411	7.17	61.6
PmaSat52-387	387	0.000167174575885	4.40	58.1
PmaSat53-36	36	0.000162056744354	7.38	61.1
PmaSat54-44	44	0.000156811189807	4.92	61.4
PmaSat55-231	231	0.000144396314672	6.74	55
PmaSat56-544	544	0.000139795315808	4.75	58.8
PmaSat57-98	98	0.00013769828211	7.53	61.2
PmaSat58-165	165	0.000130510149605	6.24	63.6
PmaSat59-167	167	0.000125421031	6.72	70.7
PmaSat60-31	31	0.000124081424488	7.00	51.6
PmaSat61-142	142	0.000117378441422	8.02	71.8
PmaSat62-47	47	0.000115912101995	4.91	59.6
PmaSat63-33	33	0.000114132890685	2.66	60.6
PmaSat64-39	39	0.0001140388311	5.89	66.7
PmaSat65-64	64	0.000106678421039	2.74	54.7
PmaSat66-188	188	0.000101115044104	8.31	66.5
PmaSat67-39	39	0.00000989358	4.82	53.8
PmaSat68-173	173	0.00000899655	8.21	62.4
PmaSat69-53	53	0.0000078606	4.77	56.6
PmaSat70-33	33	0.00000701149	4.78	51.5

Supplementary Table S2. Main characteristics of 70 satDNAs found in *Pyrrhulina semifasciata*, highlighting, in blue, those chosen for FISH mapping. The asterisk (*) indicates the satDNAs mapped in the sex chromosomes.

Família de satDNA	RUL	Abundância (M)	Abundância (F)	Abundância (M/F)	Divergência (M)	Divergência (F)	A+T (%)
PseSat01- 304*	304	0,005606905	0,006386826	0,877885927	9,78	9,65	68
PseSat02- 45	45	0,005289071	0,00496321	1,065655359	4,97	4,98	62,2
PseSat03- 68	68	0,005288433	0,005331731	0,991879186	2,52	2,54	66,1
PseSat04-226*	226	0,004975278	0,005546355	0,897035567	9,76	9,62	64,6
PseSat05 - 45	45	0,004215905	0,004087883	1,03131743	4,4	4,43	44,4
PseSat06- 198	198	0,002799141	0,003195114	0,876069106	8,88	8,81	66,6
PseSat07- 174	174	0,002764281	0,002730314	1,012440816	4,55	4,66	62,6
PseSat08- 2005	2005	0,002341543	0,00191567	1,222310384	10,37	10,38	62,3
PseSat09- 50	50	0,002080652	0,001881225	1,10600892	8,37	9,01	60
PseSat10- 610	610	0,001957573	0,001948731	1,004537311	16,83	16,92	66
PseSat11-2510	2510	0,001864105	0,001601655	1,163861305	12,58	11,71	62
PseSat12- 2219	2219	0,001725149	0,001426195	1,209616448	3,71	3,84	63,8
PseSat13- 35	35	0,001527915	0,001394119	1,095972222	1,86	1,8	57,1
PseSat14- 1235	1235	0,001207345	0,001038919	1,162116533	11,49	11,91	57,4
PseSat15- 38	38	0,001082321	0,001210707	0,893957307	13,9	14,32	63,1
PseSat16- 42	42	0,00096017	0,000568407	1,689228734	7,36	7,61	59,5
PseSat17- 381	381	0,000850538	0,000685979	1,23988987	8,74	8,25	61,4
PseSat18- 615	615	0,000729584	0,000612211	1,191720497	6,62	6,4	58,2

PseSat19- 461	461	0,000715265	0,000645695	1,10774328	11,51	9,8	63,5
PseSat20- 1284	1284	0,000705413	0,00074095	0,952039049	20,91	20,17	56,3
PseSat21- 42	42	0,000696777	0,000711353	0,979510398	4,22	4,15	64,2
PseSat22- 54	54	0,000664187	0,000593948	1,118258389	6,05	6,44	57,4
PseSat23- 155	155	0,000658736	0,000516891	1,274418737	6,28	6,5	63,8
PseSat24- 84	84	0,000643957	0,000633686	1,016208869	14,6	14,94	0,535714
PseSat25- 52	52	0,000643279	0,000613842	1,047954794	13,95	14,18	0,557692
PseSat26- 828	828	0,000613564	0,000594593	1,03190528	7,57	7,25	0,509662
PseSat27- 422	422	0,000581057	0,000722512	0,804218246	6,35	6,6	0,575829
PseSat28- 14	14	0,000557749	0,000571732	0,975542154	6,92	6,86	0,785714
PseSat29- 6	6	0,000552197	0,000573221	0,963321905	26,22	26,24	0,5
PseSat30- 192	192	0,000529303	0,0005363	0,986953819	15,32	15,9	0,640625
PseSat31- 51	51	0,000441051	0,000393715	1,120227305	9,09	9,1	0,392157
PseSat32- 186	186	0,000390964	0,000658971	0,593294997	8,44	8,02	0,709677
PseSat33- 880	880	0,000336683	0,000305781	1,101059365	3	3,46	0,669318
PseSat34-165	165	0,000282004	0,000328703	0,857928629	6,61	6,58	0,642424
PseSat35- 23	23	0,000282019	0,000250551	1,125598016	11,53	12,52	0,521739
PseSat36- 195	195	0,000237211	0,000183166	1,295058399	3,36	3,42	0,707692
PseSat37- 1041	1041	0,000231577	0,000229568	1,008752672	10,1	9,49	0,623439
PseSat38- 300*	300	0,000230661	0,000059944	3,847935851	4,96	6,47	0,576667
PseSat39- 80	80	0,000225943	0,000317697	0,711188427	15,55	15,44	0,6625
PseSat40- 39	39	0,000215363	0,000275735	0,781050911	10,74	9,56	0,641026
PseSat41- 21	21	0,000203179	0,000205974	0,986431944	17,77	16,97	0,47619
PseSat42- 1093	1093	0,000198728	0,000179851	1,104957057	1,5	1,56	0,600183
PseSat43- 124	124	0,000194009	0,000155268	1,249512671	7,66	8,11	0,564516
PseSat44- 28	28	0,00019343	0,000196448	0,984637156	9,47	9,48	0,571429
PseSat45- 445	445	0,00018569	0,000175259	1,059515613	3,95	4,07	0,573034
PseSat46- 837	837	0,000176346	0,000178717	0,986731375	7,57	7,21	0,624851
PseSat47- 21	21	0,000174861	0,000236993	0,737829475	11,25	11,75	0,666667
PseSat48- 32	32	0,000171967	0,000124967	1,376098287	5,94	7,75	0,53125
PseSat49- 33	33	0,000170543	0,000131041	1,301441784	6,45	7,5	0,606061
PseSat50- 1125	1125	0,000160486	0,000123408	1,300450538	9,97	12,7	0,656
PseSat51- 713	713	0,000146761	0,00014616	1,004109652	5,34	5,72	0,622721
PseSat52- 673	673	0,000141336	0,000132213	1,069004987	7,81	7,72	0,576523
PseSat53- 23	23	0,000138389	0,000075714	1,827790545	8,88	10,24	0,565217
PseSat54- 159	159	0,00013754	0,000194867	0,705815943	5,1	4,7	0,672956
PseSat55- 43*	43	0,000133589	0,000159353	0,838325057	8,62	7,99	0,488372
PseSat56- 87	87	0,000126526	0,000159655	0,792494666	6,25	5,81	0,678161
PseSat57- 162	162	0,000124983	8,86007E-05	1,410636489	4,99	5,06	0,67284
PseSat58- 372	372	0,000122859	0,00012864	0,955063226	2,88	2,95	0,567204
PseSat59- 46	46	0,000116657	0,000119628	0,975167464	12,34	12,69	0,608696
PseSat60- 944	944	0,000116336	0,000101468	1,146528955	7,8	7,73	0,610169
PseSat61- 213	213	0,000109063	0,000137659	0,792264963	5,14	5,29	0,615023
PseSat62- 469	469	0,000103981	0,000108045	0,962379989	4,37	3,96	0,614072
PseSat63- 463	463	0,000100169	0,000113188	0,884981918	3,94	3,69	0,591793
PseSat64- 182	182	0,000098656	5,97373E-05	1,651496552	10,95	13,61	0,681319
PseSat65- 187	187	9,83113E-05	8,62127E-05	1,14033514	5,38	5,61	0,620321
PseSat66- 44	44	8,66047E-05	0,000077618	1,115780704	4,41	4,93	0,613636
PseSat67- 198	198	8,07013E-05	0,000103397	0,780502273	4,81	4,26	0,60101
PseSat68- 179	179	7,91173E-05	7,52593E-05	1,051262745	6,5	6,33	0,614525
PseSat69- 592	592	7,38767E-05	7,23727E-05	1,020781326	4,78	4,41	0,586149
PseSat70- 759	759	7,23553E-05	6,96727E-05	1,038503861	4,52	4,48	0,586298
PseSat71- 182	182	6,96327E-05	6,92727E-05	1,005196855	6,83	7,08	0,538462

Supplementary Table S3. PCR conditions (primer, temperature, and concentration of template DNA) for the optimal amplification of satellite DNAs from *P. marilynae* and *P. semifasciata*.

Satellite	Primer F	Primer R	Anneraling temperature range	[DNA] ng/µl
PmaSat01-1627	5'CACACCTTTGGCATTCTAGC	5'TCTGTTAAGCATGGTGGAGG	56°C – 57,4°C	100
PmaSat04-50	5'GGGTGTGGTTATCTCTGTAC	5'CCCCTCTAAACAGAGTATAAAC	50°C – 51,7°C	100; 10; 1; 0,1
PmaSat05-226	5'GCAAGCTGAATACATTCATG	5'GACTGTCTGAGAGCACAAAC	53°C - 54,5°C	100; 10; 1; 0,1
PmaSat06-198	5'AAGACAGCTTCTGCATCCATG	5'TTGCAGATTTGCCCAAAAAC	55°C - 55,6°C	100; 10; 1; 0,1
PmaSat07-45	5'CCTCTGTAACACATTAAACTGT	5'TAGGAGAGTGTAGTGTTAGTCT	51,7°C – 58°C	100; 10
PmaSat09-335	5'CTGCAACCACTTCCAGTGAT	5'ACACTTAGTTGTGTCTGAAA	53°C - 53,5°C	100; 10; 1; 0,1
PmaSat10-4663	5'AGAACGGAGGTCTCTTGCGT	5'CTCCATTCATTTCCATCACGC	56°C – 62,5°C	100
PseSat01-304	5'CTGCAACCACTTCCAGTGAT	5'ACACTTAGTTGTGTCTGAAA	53°C - 53,5°C	100; 10; 1; 0,1
PseSat04-226	5'GCAAGCTGAATACATTCATG	5'GACTGTCTGAGAGCACAAAC	53°C - 54,5°C	100; 10; 1; 0,1
PseSat06-198	5'AAGACAGCTTCTGCATCCATG	5'TTGCAGATTTGCCCAAAAAC	55°C - 55,6°C	100; 10; 1; 0,1
PseSat32-186	5'TGAGAGAAGACTTTACAAGC	5'TGACACATTTAAGGCATTTC	50°C - 54,9°C	100
PseSat34-165	5'CATGGCTGATAGGGTAAAAG	5'ATGTGACCACACTGTATCCC	56,5°C - 63°C	1; 0,1
PseSat38-300	5'ATAGACGGATGGATTAACGG	5'AAAGACTGACAGACCATTAT	53°C - 61°C	100; 10; 1; 0,1
PseSat39-80	5'CTTACTGTCTCATATCTGTG	5'AGTTAAGACCAGTATCTCTA	50,5°C - 57,4°C	100
PseSat50-1125	5'CTGTACAGTGGATTATGGAG	5'TTCGTGTTACTCAGAATGTC	53 - 53,5	100; 10
PseSat54-159	5'AATCTCTGCATATAAAATGGC	5'AGATGGACCAAAAGGTGTAT	50°C - 58°C	100; 10; 1; 0,1
PseSat55-43	5'CTGTGGTGCACTAACCAGA	5'CCCTATCATTTACAGTACA	50°C - 58°C	100; 10; 1; 0,1
PseSat56-87	5'CACCCAGCCACTTTTTTA	5'TGGAGGTAGTTAGTTAGATG	50°C - 51,5°C	100
PseSat57-162	5'AGGGTCAGTACTCTCACT	5'CACTTCATAACAGTCATTTTAA	52°C - 60°C	100; 10; 1; 0,1
PseSat61-213	5'ATTCCAGGCAATAATCTGCC	5'CAGGCGAGAATTCTACCACT	53°C - 57,9°C	100; 10; 0,1
PseSat63-463	5'ATTGGTCAGATATTGTGAAG	5'ATTGCGCCGTTCATATTCAC	52°C - 60°C	100; 10
PseSat64-182	5'TTTACTGCTGTTAAGATTTC	5'CCTTTGGTAGGACATGTAGC	50°C - 50,7°C	100; 10; 1; 0,1
PseSat67-198	5'CTAACTTCATTCGGC	5'CCACCATGGCACACCT	55°C - 55,6°C	10; 1; 0,1

Supplementary Table S4. Shared SatDNA between *P. marilynae* and *P. semifasciata*. The sequences highlighted in blue correspond to those selected for the FISH experiments

Pyrrhulina marilynae	*Pyrrhulina semifasciata*	Classificação
PmaSat02-68	PseSat03-68	SV
PmaSat03-45	PseSat05-45	SV
PmaSat04-50	PseSat09-50	SV
PmaSat05-226	PseSat04-226	SV
PmaSat07-45	PseSat02-45	SV
PmaSat08-33	PseSat13-35	V
PmaSat11-2486	PseSat11-2510	V
PmaSat14-54	PseSat22-54	SV
PmaSat17-192	PseSat30-192	SV
PmaSat18-42	PseSat21-42	SV
PmaSat19-84	PseSat24-84	SV
PmaSat21-712	PseSat51-713	SV
PmaSat22-32	PseSat48-32	SV
PmaSat23-23	PseSat35-23	SV
PmaSat24-1283	PseSat28-1284	V
PmaSat26-21	PseSat47-21	SV
PmaSat27-39	PseSat40-39	SV
PmaSat28-165	PseSat54-159	V
PmaSat29-51	PseSat31-51	SV
PmaSat32-33	PseSat49-33	SV
PmaSat35-845	PseSat46-837	V
PmaSat36-445	PseSat45-445	V
PmaSat37-28	PseSat44-28	SV
PmaSat38-162	PseSat57-162	SV
PmaSat39-88	PseSat56-87	SV
PmaSat40-211	PseSat61-213	V
PmaSat41-915	PseSat33-880	V
PmaSat44-43	PseSat55-43	SV
PmaSat45-568	PseSat69-592	V
PmaSat46-470	PseSat62-469	V
PmaSat47-672	PseSat52-673	V
PmaSat49-80	PseSat39-80	SV
PmaSat52-387	PseSat58-372	V
PmaSat54-44	PseSat66-44	SV
PmaSat12-842	PseSat15-38	SF
PmaSat15-1437	PseSat14-1235	SF
PmaSat66-188	PseSat06-198	SF
PmaSat55-231	PseSat71-182	SF

Author Contributions

Formal analysis, J.A.D.V., R.U. and M.d.B.C.; Funding acquisition, T.L. and M.d.B.C.; Investigation, T.L., R.L.R.d.M., F.d.M.C.S., R.Z.d.S., J.H.F.S., F.P.-F. and M.d.B.C.; Methodology, R.L.R.d.M., F.d.M.C.S., J.A.D.V., C.A.G.G., R.Z.d.S., J.H.F.S., F.P.-F., R.U. and M.d.B.C.; Project administration, M.d.B.C.; Software, R.L.R.d.M., J.A.D.V., C.A.G.G., J.H.F.S. and R.U.; Supervision, M.d.B.C.; Validation, R.L.R.d.M., F.d.M.C.S., J.A.D.V., C.A.G.G., R.Z.d.S., F.P.-F., R.U. and M.d.B.C.; Visualization, T.L. and F.d.M.C.S.; Writing—original draft, R.L.R.d.M., F.d.M.C.S., J.A.D.V. and C.A.G.G.; Writing—review and editing, T.L., J.A.D.V., C.A.G.G., R.Z.d.S., J.H.F.S., F.P.-F., R.U. and M.d.B.C. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by São Paulo Research Foundation (FAPESP) grants 2022/04964-3 and 2019/25045-3 (RLRDM), 2020/11772-8 (MBC). We also acknowledge support for TL by the German Research Foundation Projekt-Nr. 512648189 and the Open Access Publication Fund of the Thueringer Universitaets- und Landesbibliothek Jena. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Institutional Review Board Statement

Sample was approved by the Brazilian Environmental Agency ICMBIO/SISBIO (License 48628-14) and SISGEN (A96FF09). All experiments followed the guidelines and were approved by the Ethics Committee on Animal Experimentation of the Universidade Federal de São Carlos (Process number CEUA 7994170423).

Informed Consent Statement

Not applicable.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request. The datasets generated and analyzed during the

current study are available in the GenBank repository under the accession numbers OR094701-OR094771 and OR094772-OR094841.

Acknowledgments

We appreciate the contributions of Elixir CZ, CESNET, and CERIT-SC in the maintenance of the public Galaxy server where RepeatExplorer2 analysis was performed. The authors are grateful to Claudio Oliveira for the space on the server and to Francisco J. Ruiz-Ruano for providing Python scripts to analyze satellite sequences.

Conflicts of Interest

The authors declare no conflict of interest.

References

All references are compiled in the end of this thesis.

Capítulo 4

The evolution of Lebiasinidae (Teleostei: Characiformes): Insights from cytogenetics

Renata Luiza Rosa de Moraes, Francisco de Menezes Cavalcante Sassi, Fernando Henrique Santos de Souza, Geize Aparecida Deon, Manoela Maria Ferreira Marinho and Marcelo de Bello Cioffi

Artigo em Submissão

Abstract

Cytogenetics and molecular data play critical roles in understanding the evolution of non-model species. When coupled, they offer an integrated approach that enables a more robust understanding of genetic diversity, population structure, and phylogenetic relationships within groups. Cytogenetic and molecular studies on Lebiasinidae fishes are challenging, mainly because of their small sizes. However, with advancements in cytogenetic and molecular methods, the number of studies on the family has increased significantly. In this work, we first filled an important gap by describing, for the first time, the karyotype of a *Copella* species (*C. callolepis*), showing a 2n = 36 with a predominance of acrocentric chromosomes and multiple 5S and 18S rDNA sites. Furthermore, by coupling previous phylogenetic studies, a nuclear marker-based phylogenetic tree, and a NeighborNet network, we were able to achieve a comprehensive understanding of the genetic relationships among species. In light of these new data and previous morphology-based phylogenetic analyses, we evaluated the chromosomal evolution patterns of Lebiasinidae and put forth their most plausible evolutionary pathways. There was a significant karyotypic remodeling pattern in Lebiasinidae fishes, involving different sets of chromosomal rearrangements. This research is the continuation of a series of previous investigations into the Lebiasinidae family and settles one more step in the elucidation of its evolutionary history.

1. Introduction

Research on non-model species generally presents several challenges when compared to model ones. Genomic resources, such as scaffold whole genome data or well-annotated reference genomes, are usually scarce, or even unavailable (Ekblom and Galindo, 2011; Russel *et al*., 2017). Cytogenetic techniques are a very powerful tool in the study of non-model organisms since they characterize the chromosomal morphologies, ploidy, and position of specific genomic markers without the need for prior information. As a result, it can aid in clarifying taxonomic issues and improving comprehension of the relationships between species or populations. For example, cytogenetics has also contributed to finding cryptic diversity and even possible species complexes in fishes (Castro *et al*., 2015; Rocha-Reis *et al*., 2018). Other approaches, such as the use of mitochondrial markers like 16S rRNA, cytochrome oxidase I (COI), and the Control Region (CR) are widely applied to comprehend how different groups of organisms evolved and study population genetics (Imoto *et al*., 2013; Bronstein, Kroh and Haring, 2018). However, mitochondrial DNA analysis can generate distortions in the results when reticulate events are present since it represents a maternally inherited marker (Toews and Brelsford, 2012). Sampling nuclear markers is required to overcome such inheritance biases and provide a more reliable perspective on the evolutionary history of the species under study (Tollesfrud, 2009). Furthermore, incomplete lineage sorting and other frequent phylogenetic issues can be resolved with the use of several genetic markers (Mirarab *et al.*, 2016; Simmons and Gatesy, 2021). Therefore, using complexity reduction techniques that take advantage of the high throughput of NGS techniques is one of the primary methods used to access the diversity and genetic structure of these species at the nuclear level. This approach allows for the cost-effective acquisition of a large number of independent genetic markers across the genome (Davey *et al*., 2011). Among complexity reduction methods available, DArTseq sequencing (Diversity Arrays Technology, Canberra, Australia) allows for the enrichment of hypomethylated regions (Jaccoud *et al*., 2001; Kilian *et al*., 2012). Since molecular cytogenetics approaches primarily focus on repetitive and hypermethylated regions, DArTseq is a useful complementary strategy for cytogenetics investigations (Cioffi *et al*., 2019; de Oliveira *et al*., 2019; Panthum *et al*., 2021; Schimek *et al*., 2022; Toma *et al*., 2023).

The Lebiasinidae family includes seven genera (*Lebiasina*, *Piabucina*, *Derhamia*, *Nannostomus*, *Pyrrhulina*, *Copella*, and *Copeina*) and 74 valid species that are native to the Neotropical region and widely spread in South and Central America (Costa Rica and Panama),

except for Chile (Fricke *et al*., 2024). Two subfamilies are recognized based on morphological characters: Lebiasininae, which includes *Derhamia*, *Lebiasina,* and *Piabucina*, and Pyrrhulininae, the most varied clade, encompassing the genera *Nannostomus*, *Copeina*, *Copella*, and *Pyrrhulina* (Géry and Zarske, 2002; Weitzman and Cobb, 1975). Unpublished phylogenetic analysis, however, recovers *Derhamia* as a basal taxa within the Pyrrhulininae (Marinho, 2014; Netto-Ferreira, 2010). In general, Lebiasinidae is composed of small-sized fishes that range in length from 1.5 to 7.0 cm and exhibit a wide diversity of body forms and colors, making them very appealing to aquarium hobbyists (Weitzman and Weitzman, 2003; Weitzman and Vari, 1988). Due to this richness and to old species description, taxonomic identification of species is sometimes difficult, demanding the use of additional techniques such as cytogenetics or barcoding (Xu *et al*., 2024). In the same way, issues with phylogenetic placement also occurred. For a long time, the phylogenetic relationships of Lebiasinidae were uncertain, being related to different groups of Neotropical fishes. Phylogenetic analysis based on morphological characters recovered Lebiasinidae as closely related to Erythrinidae, Ctenoluciidae, and Hepsetidae (Buckup, 1998) and Tarumanidae (de Pinna *et al*., 2018) whereas those based on molecular data recovered Lebiasinidae as close to Serrasalmidae, Erythrinidae, and Hepsetidae (Ortí and Meyer, 1997) or Ctenoluciidae. Combined data have hypothesized Lebiasinidae as a sister to Ctenoluciidae (Calcagnotto *et al*., 2005; Oliveira *et al*., 2011; Arcila *et al*., 2017; Betancur-R *et al*., 2019; Mirande, 2019; Cassemiro *et al*., 2023; Near and Thacker, 2024;).

Because of their small size, representatives of the Lebiasinidae were little explored cytogenetically. One big problem with studying the cytogenetics of such small fish was getting accurate information on the number and quantity of chromosome spreads. As a result, the studies were limited to describing haploid and/or diploid chromosome numbers in some species (Oliveira *et al*., 1991; Scheel, 1973; Arai, 2011). Thanks to improvements in methods, solid studies using both classical and molecular cytogenetics have made a considerable difference in our understanding of how the karyotypes of this group of fish have changed over time (Moraes *et al*., 2017; Moraes *et al*., 2019; Sassi *et al*., 2019; Toma *et al*., 2019; Sassi *et al*, 2020; Sember *et al*., 2020; Moraes *et al*., 2021; Ferreira *et al*., 2022; Leite *et al*., 2022; Moraes *et al*., 2023a). Such investigations included the first cytogenetic data for *Lebiasina, Copeina, Nannostomus,* and *Pyrrhulina*. In general, two evolutionary trends have been proposed for the karyotypic evolution of Lebiasinidae: 1) the conservation of a plesiomorphic karyotype in the subfamily Lebiasininae with 2n = 36 biarmed chromosomes, and 2) the Pyrrhulininae subfamily exhibits significant variances in diploid numbers and karyotypes, with acrocentric chromosomes predominating in most species (Sassi *et al*., 2020). In this scenario, the majority of Pyrrhulininae's

acrocentric chromosomes are the result of rearrangements such as centric fissions (Sassi *et al*., 2020). However, some exceptions within the subfamily have revealed secondary fusion events, resulting in metacentric chromosomes in some species and a decrease in 2n, as seen in *Pyrrhulina marilynae* (2n=34), *Nannostomus anduzei* (2n= 22), and *Nannostomus unifasciatus* (2n=22), the last two representing one of the lowest diploid chromosome numbers found among teleost fishes (Sember *et al*., 2020; Moraes *et al*., 2021; Moraes *et al*., 2023a).

Here, to gain insights into the evolution of Lebiasinidae fishes, we describe, for the first time, the karyotype of a *Copella* species (named *C. callolepis*), filling an important gap in the cytogenetics knowledge of the family. Next, we reconstruct the evolutionary relationships within the family using genome-wide markers. With the assistance of this new data, we assessed the chromosomal evolution patterns of Lebiasinidae and proposed their likely evolutionary pathway.

2. Material and Methods

2.1. Sampling

Thirteen individuals of *Copella callolepis* were collected in Tefé - AM (3°25'50.7"S, 54°44'54.8"W). To collect tissues of the liver, kidney, spleen, and gills, a stereomicroscope was used. Chromosome isolation from collected tissues followed the conventional air-drying technique (Bertollo, 1978). Following the dissection, the complete fish bodies were preserved at 100% ethanol and deposited in the tissue bank of the Laboratório de Citogenética de Peixes of Universidade Federal de São Carlos (São Carlos, São Paulo, Brazil). All individuals were collected under the appropriate authorization of the Brazilian environmental agency ICMBIO/SISBIO (License number 48628-14) and SISGEN (A96FF09). The experiments followed ethical standards, and anesthesia followed the Ethics Committee on Animal Experimentation of the Universidade Federal de São Carlos (Process number CEUA 7994170423). A list of all individuals used in Cytogenetics and sequencing analysis is presented in **Table 1**. Sampling sites for all species are presented in **Figure 1**.

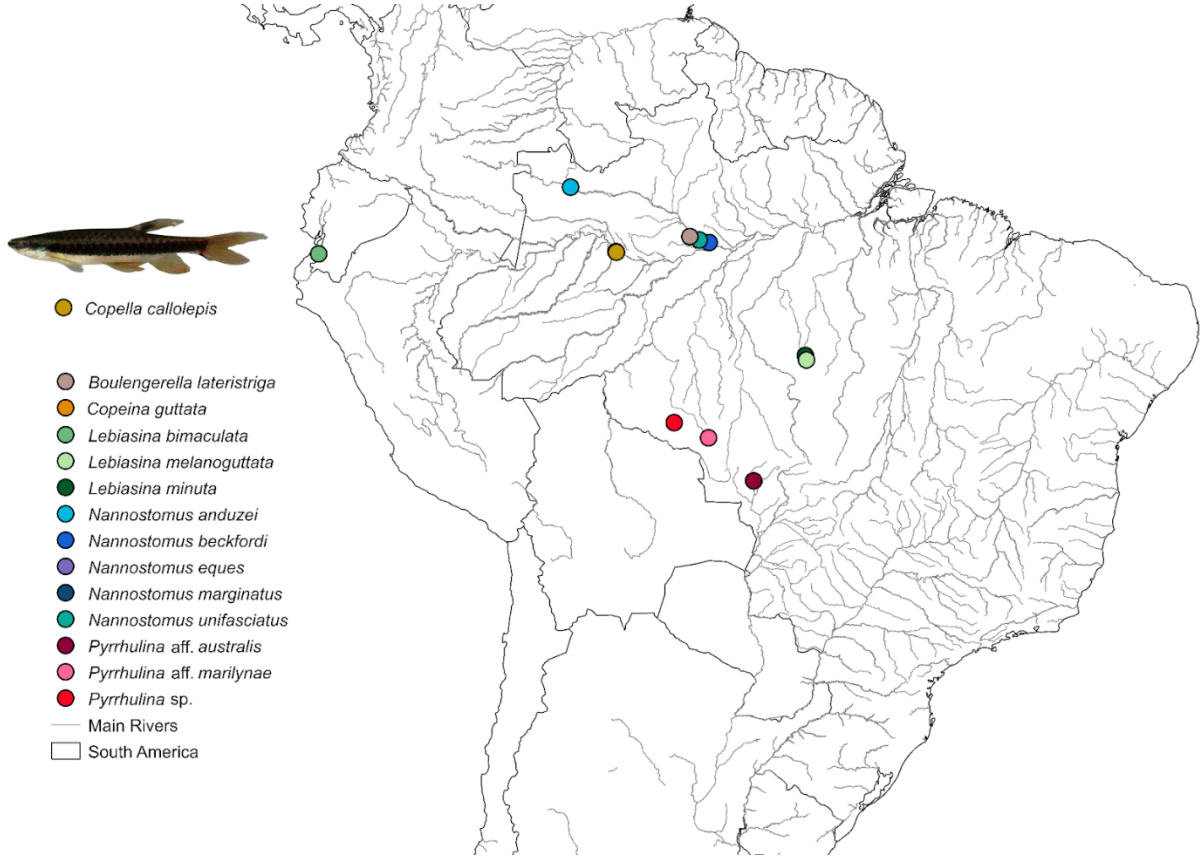


Figure 1. Map of South America with sampling sites indicated as dots. Each color represents a distinct species, with colors defined on the legend in the left corner. A *Copella callolepis* individual is presented above the legend (photo by José Luis Olivan Birindelli). The map was produced using the software QGis 3.4.4 (https://qgis.org), Inkscape 0.92 (https://inkscape.org), and Adobe Photoshop CC 2020 (San Jose, CA, USA).

2.2. Conventional and molecular cytogenetics

Constitutive heterochromatin was revealed by the application of a standard C-banding procedure (Sumner, 1972). The 5S and 18S rDNA probes of ribosomal 5S and 18S rDNA were isolated from the *Hoplias malabaricus* genome (Cioffi *et al*., 2009; Pendás *et al*., 1994). The 5S rDNA probe included 120 base pairs (bp) of the 5S rDNA gene coding region and 200 bp of non-transcribed spacer (NTS) (Pendás *et al*., 1994). The 18S rDNA probe was composed of a 1400-bp-long segment of the 18S rDNA coding region (Cioffi *et al*., 2009). Both probes were directly labeled with the Nick Translation Mix Kit (Jena Bioscience, Jena, Germany) according to the manufacturer's instructions, the 18S rDNA probe with Atto488-dUTP and the 5S rDNA with Atto550-dUTP. Both rDNA sequences were mapped by fluorescence in situ hybridization (FISH) with high stringency conditions (Sassi *et al*., 2022). Briefly, metaphase chromosomes were treated with RNAse A (40 g/mL) for 1.5 h at 37°C and the metaphase chromosomes were denatured in 70% formamide/2xSSC at 72°C for 3 min. A total of 20 µL of the hybridization mixture (2.5 ng/L probes, 50% deionized formamide, and 10% dextran sulfate) was then

applied, and the hybridization process was performed overnight at 37°C in a moist chamber. The first post-hybridization wash was performed with 1xSSC for 5 min at 65°C in a shaker, followed by 4xSSC/Tween for 5 min at room temperature. Chromosomes were counterstained with DAPI in Vectashield (Vector Laboratories, Burlingame, CA, USA).

2.3. Sequencing and obtention of SNPs

We sequenced a total of 41 individuals from 14 different Lebiasinidae species (**Table 1**). From each individual, we extracted a small fragment of liver or muscle and stored it in 100% ethanol. We performed the DNA extraction according to the protocol described in Sambrook and Russell (Sambrook and Russell, 2001). We assessed the DNA quality with NanoDrop (ThermoFisher Scientific, Branchburg, NJ, EUA) and then sent the samples to Diversity Arrays Technology Pty Ltd. (Canberra, Australia), for the DArTseq procedure.

DArTseq is based on the use of two distinct restriction enzymes: a frequent cutter and a nuclease with affinity for hypomethylated regions, generally PstI and SphI (Kilian *et al*., 2012). The sequencing was performed on the Illumina HiSeq 2500 platform, using 3 individuals from each species. After sequencing, all generated adapters were trimmed, and sequences went through a quality filtering process. The quality of the reads was checked and the resulting reads were processed with proprietary DArT software, resulting in high-quality sequences. We extracted and concatenated the sequences of up to 69 bp with the gl2fasta function of the package DArTR v.2.9.7 (Gruber *et al*., 2018) in a dataset for phylogenetic analysis. A final dataset comprising a set of Single Nucleotide Polymorphisms (SNP) was generated, with SNPs coded as 1 for heterozygotes, 0 for reference homozygotes, and 2 for alternate homozygotes.

Table 1: Samples used in this study.

Species	2n	Sampling Site	Geographical coordinates	References
Boulengerella lateristriga	36	Novo Airão - AM, Brazil	2°37′28.5″S, 60°58′16.8″W	Souza e Sousa *et al*. 2017
Copeina guttata	42	Tefé River - AM, Brazil	03°39'49.5"S, 64°59'40"W	Toma *et al*. 2019
Copella callolepis	36	Tefé River - AM, Brazil	3°25'50.7"S, 54°44'54.8"W	**Present study**
Lebiasina bimaculata	36	Arenillas River - El Oro, Ecuador	03°30'57.2"S, 80°3'44.2"W	Sassi *et al*. 2019
Lebiasina melanoguttata	36	Cachoeira da Serra - PA, Brazil	08°46' 59.4"S, 54°58'26.9"W	Sassi *et al*. 2019
Lebiasina minuta	36	Cachoeira da Serra - PA, Brazil	8°44'39.0"S 55°02'03.0"W	Leite *et al*. 2022
Nannostomus anduzei	22	Zamula Stream, Barcelos - AM, Brazil	0°04'57.5"S 67°06'23.8"W	Moraes *et al*. 2023a
Nannostomus beckfordi	44	Agenor Stream - AM, Brazil	2°55'53.9"S 59°58'30.7"W	Sember *et al*. 2020
Nannostomus eques	36	Cuieiras River - AM, Brazil	2°47'58.1"S 60°29'19.8"W	Sember *et al*. 2020
Nannostomus marginatus	42	Adolpho Ducke Reserve - AM, Brazil	2°55'53.9"S 59°58'30.7"W	Sember *et al*. 2020
Nannostomus unifasciatus	22	Cuieiras River - AM, Brazil	2°47'58.1"S 60°29'19.8"W	Sember *et al*. 2020
Pyrrhulina aff. *australis*	40	Branco River - MT, Brazil	15°11'28.0"S 57°41'30.7"W	Moraes *et al*. 2017

Pyrrhulina aff. *marilynae*	40	12 de Outubro Stream - MT, Brazil	12°58'41.0"S 60°00'34.0"W	Moraes *et al*. 2021
Pyrrhulina sp.	40	Alto Alegre do Parecis - RO, Brazil	12°11'58.0"S 61°46'47.4"W	Moraes *et al*. 2021

2.4. Principal Component Analysis

To access an overview of the distribution of genetic diversity among samples, we carried out a Pearson Principal Component Analysis (PCA). We performed the analysis with two distinct datasets, the first comprising data from all species, and the second only with *Nannostomus* data. We imported the SNP matrix in RSTUDIO v.4.3.2 as a genlight object, and the PCA was performed using the Package DArTR v.2.9.7 (Gruber *et al*., 2018).

2.5. Species tree

To better understand the genetic relationships between species, we estimated a maximum likelihood species tree using RAxML (Stamatakis, 2014). Before the analysis, we first concatenated all sequences from each sample in a PHYLIP format file using the DArTR package. Initially, we conducted a model test to define which substitution model better adjusted to our data. We estimated the substitution model with the maximum likelihood approach using ModelTest-NG (Darriba *et al*., 2020) with a template parameter adjusted to RAxML, datatype set as DNA, and seed number defined as 12345. We set all remaining parameters at default. We considered the model with the smallest Bayesian Information Criterion (BIC) as the best model. We performed the phylogenetic analysis using the GTR+Gamma+I model with four rate categories, setting parameters to generate 100 distinct starting trees, and performed a rapid bootstrap analysis. We used a random seed number of 12345 for both the parsimony starting trees and the bootstraps. We estimated a consensus tree in GENEIOUS Prime v2023.2. with the Majority Rule consensus method, with a support threshold of 70%. The consensus tree is based on the 100 trees generated, we exported it and plotted it using FigTree v 1.4.4.

To test for any reticulate structure that deviates from the bifurcating tree-like evolution, we implemented a NeighborNet network analysis (Bryant and Moulton, 2004) in SplitsTree software.

3. Results

3.1. Conventional and molecular cytogenetics

The diploid chromosome number in all analyzed *Copella callolepis* specimens was 2n = 36 (**Figure 2**). The karyotype consisted predominantly of acrocentric chromosomes except for two subtelocentric chromosomes (**Figure 2A**), with C-positive heterochromatin blocks found in the centromeric region of almost all chromosomes (**Figure 2B**). Multiple 5S rDNA and 18S

rDNA sites were found in the centromeric regions of pairs 5, 9, and 14 (5S rDNA) and 6 and 12 (18S rDNA) (**Figure 2C**).

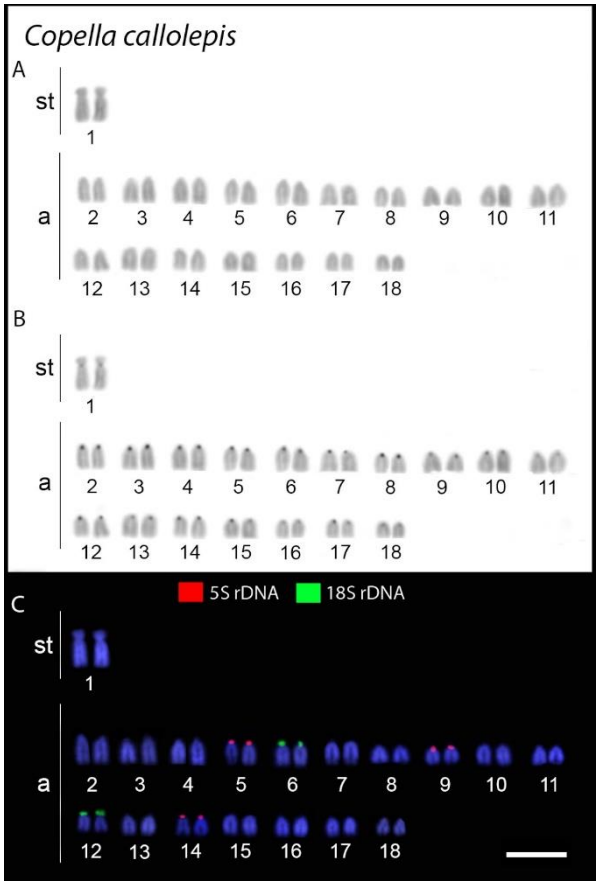


Figure 2. Karyotypes of *Copella callolepis* arranged after different cytogenetic protocols. Giemsa staining (A), C-banding (B), and dual-color FISH with 18S (green) and 5S (red) rDNA probes (C). Chromosomes are counterstained with DAPI (blue). Scale bar = 10μm

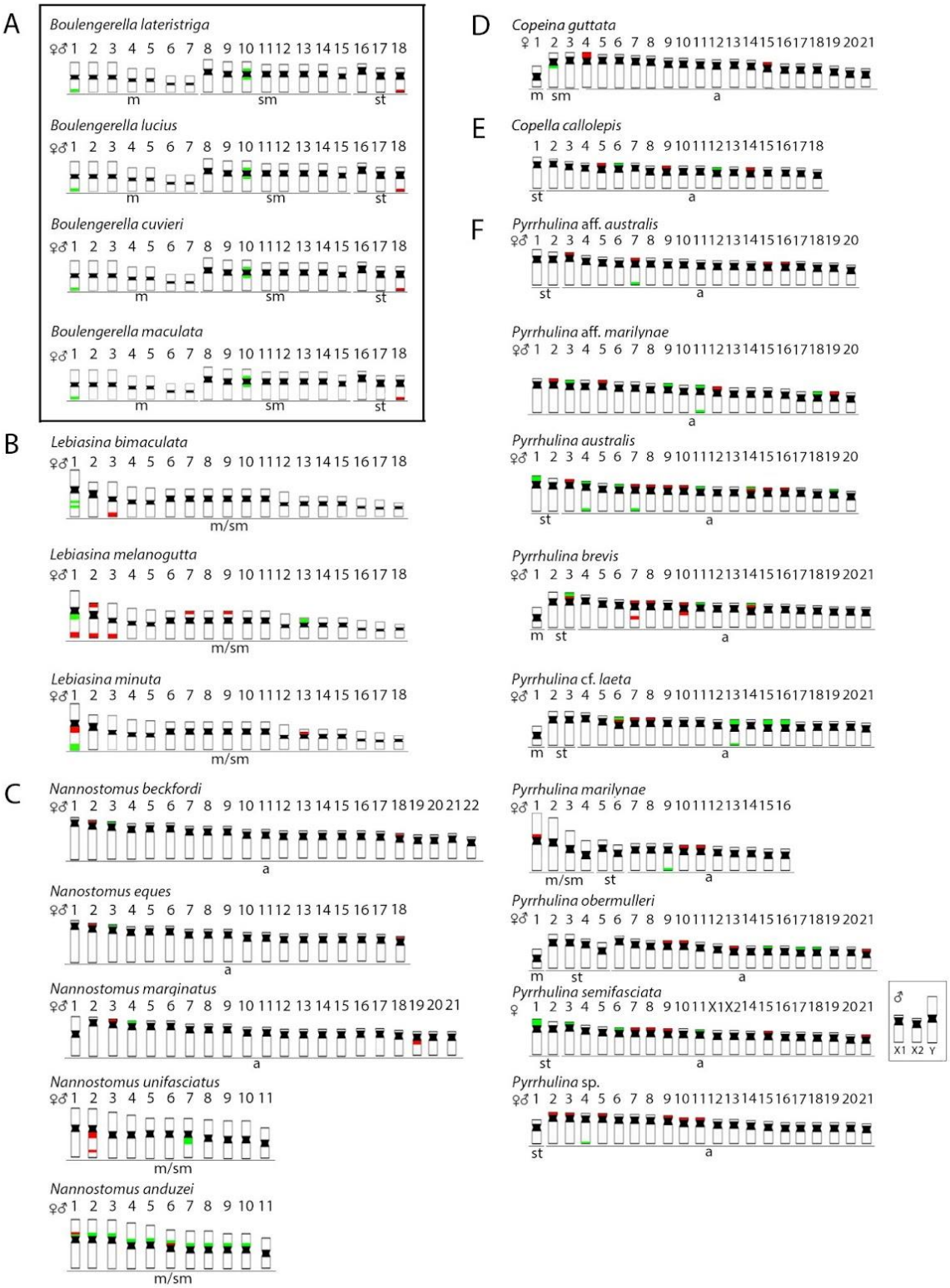


Figure 3. Schematic representation of chromosomes of Lebiasinidae and Ctenoluciidae species, highlighting the position of 18S rDNA (green) and 5S rDNA (red). The small box highlights a sex chromosome system in *Pyrrhulina semifasciata*, while the bigger box highlights the Ctenoluciidae members. FISH data were taken from Souza e Sousa *et al*. 2017; Moraes *et al*, 2017; Sassi *et al*. 2019; Toma *et al*. 2019; Sember *et al*. 2020; Leite *et al*. 2022; Moraes *et al*., 2019; Moraes *et al*., 2021; Moraes *et al*. 2023a. Letters correspond to the genera: (A) *Boulengerella*, (B) *Lebiasina*, (C) *Nannostomus*, (D) *Copeina*, (E) *Copella*, and (F) *Pyrrhulina*.

3.2. *Principal component analysis*

The principal component analysis shows a clear clustering of samples mainly at the genera level, evidencing the genetic differentiation between the species examined (**Figure 4A**), the PCA comprising only *Nannostumus* samples (**Figure 4B**) helps to unravel the differentiation between *Nannostumus beckfordi*, *Nannostomus eques*, and *Nannostomus unifasciatus* which appeared clustered on the main PCA.

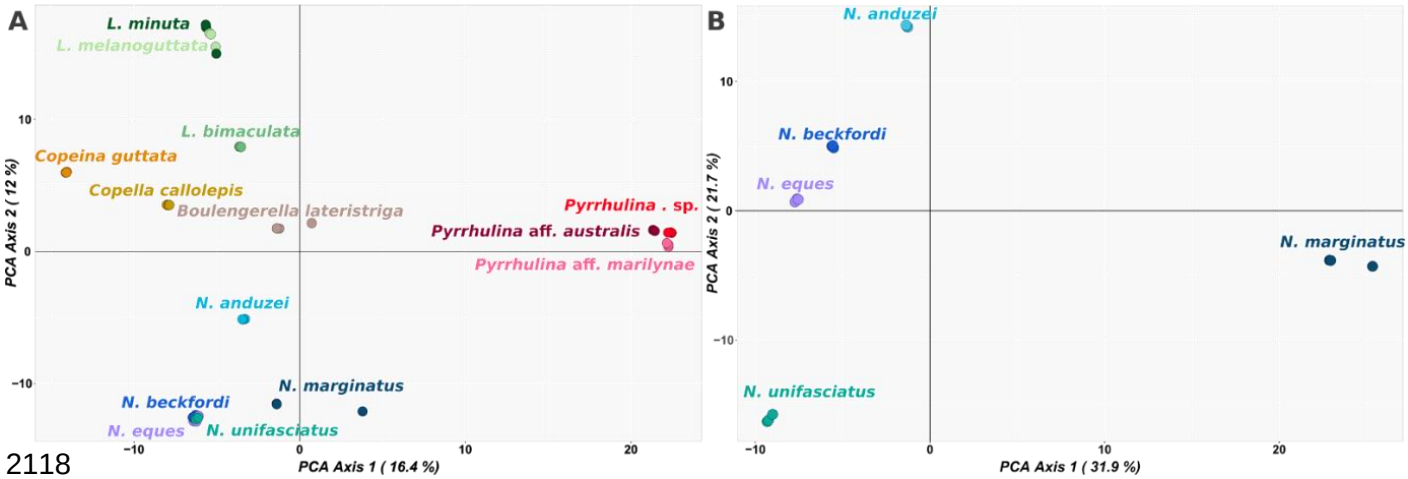


Figure 4. Principal component analysis results, each circle represents an individual, species are colored differently according to the names, percentage values indicate the explanatory capability of each axis, (A) displays the results for the PCA analysis with the full dataset; (B) displays the results for the dataset comprising only *Nannostomus* samples.

3.3. *Species tree and Neighbor-net*

The estimated species tree is presented in **Figure 5A**. Individuals of the same species are collapsed together to simplify visualization, and the full tree is presented in the supplementary material (**Figure S1**). Overall, the topology is well supported, with all nodes having 100 bootstrap support. The tree depicts the relationships among sampled Lebiasinidae species, placing *Copella callolepis* in a broader phylogenetic context with nuclear molecular data for the first time.

The NeighborNet network (**Figure 5B**) presents a structure congruent to the species tree, the main difference being a minor clustering of a *Pyrrhulina* aff. *marilynae* individual with *Pyrrhulina* sp. samples.

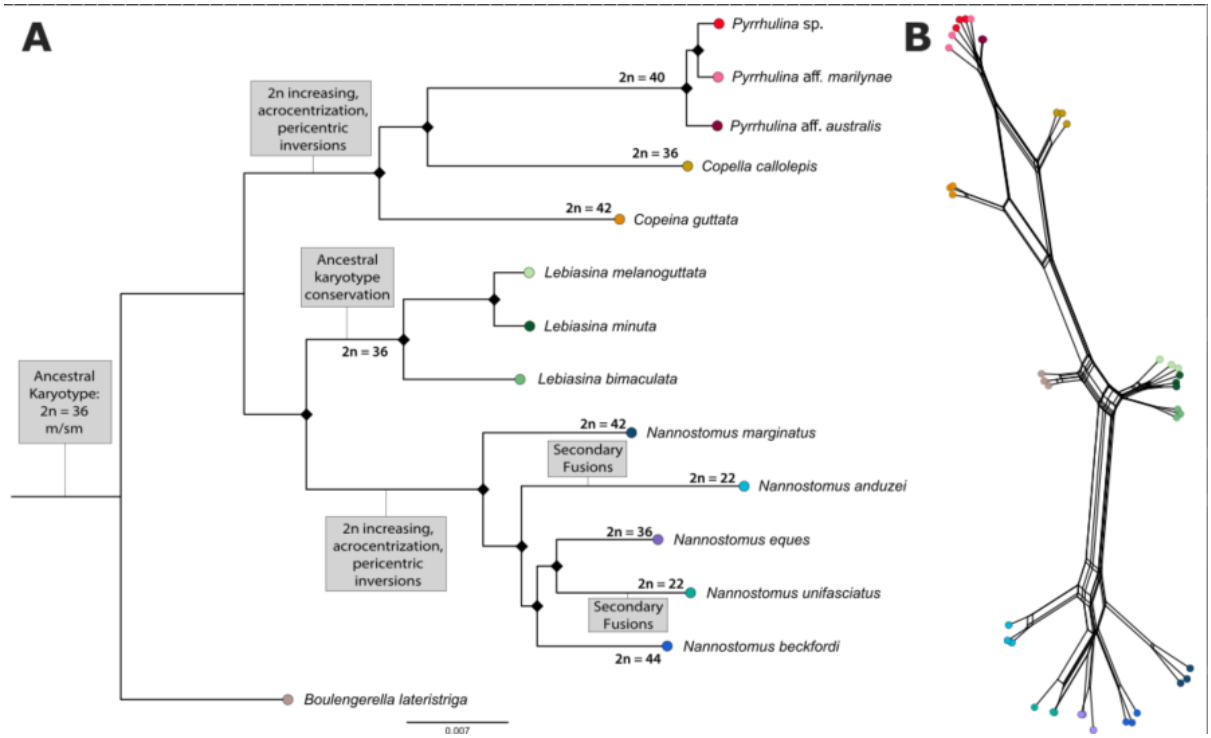


Figure 5. Maximum likelihood species tree (A), diamond symbols indicate nodes with bootstrap values of 100, gray boxes indicate relevant the main cytogenetic features of each clade. NeighborNet network circles are colored in the same scheme as the phylogenetic tree, each circle represents a sample that was collapsed into one in the phylogeny (B).

4. Discussion

This study presents a comprehensive analysis of the cytogenetics of Lebiasinidae, including the karyotype of a *Copella* species (*C. callolepis*), which has now been defined for the first time. In addition, by using molecular phylogenetic research and a NeighborNet network, we successfully reconstructed a species tree, providing a thorough comprehension of the genetic links between Lebiasinidae species. Based on this new information, we analyzed the chromosomal evolution patterns of Lebiasinidae and proposed their most likely evolutionary pathways.

4.1. Chromosomal features of Copella callolepis

Based on our findings, *C. callolepis* preserved the proposed ancestral 2n for the family (2n = 36), but not its karyotype structure. In contrast to the ancestral karyotype, which was mostly made up of metacentric chromosomes, *C. callolepis* karyotype exhibited a large number of acrocentrics, which reinforces the important role of pericentric inversions in the karyotype evolution of the group. Similar scenarios are observed in other fishes, such as the giant trahiras *Hoplias* (Sassi *et al*., 2021), *Nothobranchius* killifishes (Krysanov *et al*., 2023), *Rineloricaria* armored catfishes (Marajó *et al*., 2023), and *Ictalurus* catfishes (Waldbieser *et al*., 2023). Besides pericentric inversions, centromere repositioning (Schubert, 2018) might also explain the

high divergence of acrocentrics between karyotypes of related species. Indeed, a molecular study in the medaka fish *Oryzias javanicus* suggests the repositioning of centromere-associated repeats as the main mechanism for the increase of acrocentrics in karyotype, rather than pericentric inversions (Ansai *et al*., 2023). In addition to acrocentrics, the karyotype of *C. callolepis* has a single pair of large subtelocentric chromosomes, which is very similar in shape and size to pair 5 of *P. marylinae*, pair 1 of *Pyrrhulina* sp., and pair 2 of *P. obermulleri* and *Pyrrhulina* cf. *laeta* (Moraes *et al*., 2021), suggesting that it may have originated before the split of *Pyrrhulina* and *Copella* (**Figure 5**). The presence of metacentric and submetacentric chromosomes suggests that the karyotype of *Copeina guttata* retained other features of the basal karyotype of the family (Toma *et al*., 2019). However, given their similar sizes, it's also possible that the subtelocentric pair observed in *Pyrrhulina* and *Copella* is also present in the karyotype of *Copeina guttata*, either as pair 2, 3, or 4. To shed light on this scenario, additional investigations employing whole-chromosome painting seem necessary.

While the 2n of other *Copella* species ranges from 24 to 44 chromosomes (Arai, 2011; Scheel, 1973), these studies were merely descriptive and omitted important details like the karyotype formula or the number of individuals karyotyped, which impairs their reproducibility. However, such 2n variation indicates that a series of rearrangements have taken place during the chromosomal evolution of the genus, which might have included chromosome fissions, fusions, translocations, and inversions, as also observed in other lebiasinids (Sember *et al*., 2020; Moraes *et al*., 2021; Moraes *et al*., 2023). Moreover, this series of rearrangements is also highlighted by the presence of multiple sites for both 5S and 18S rDNA (**Figures 2** and **3**), an unusual characteristic for fishes (Gornung, 2013; Sochorová *et al*., 2018). Ribosomal DNAs are known to spread through the genome either by their association with transposable elements (TEs) (Garcia *et al*., 2024) or by the presence of pseudo-homologous regions across heterologous acrocentric chromosomes, allowing them to recombine and exchange ribosomal repeats (Guarracino *et al*., 2023). Given the presence of C-positive heterochromatin in the centromeres of *C. callolepis* (**Figure 2B**) associated with ribosomal repeats, and hence decreased or repressed recombination (Ellermeier *et al*., 2010; Roberts, 1965), the association with TEs is more likely to be the mechanism behind the occurrence of multiple rDNA sites in its karyotype. Albeit the association of rDNAs and sex chromosomes in several fish species (Cioffi and Bertollo, 2012), the multiple X_1X_2Y that arose by a Robertsonian translocation in *P. semifasciata* (Moraes *et al*., 2019) do not accumulate such sequences. Nevertheless, repeated DNA differences between the sexes supported the likely ZZ/ZW sex system of *L. bimaculata* (Sassi *et al*., 2019). Regretfully, although *C. callolepis* exhibits parental care and sexual dimorphism

(Marinho and Menezes, 2017), our study was unable to access the female karyotype to detect any sex-related variations using cytogenetic analysis.

4.2. Trends in karyotype evolution of Lebiasinidae in light of new genomic data

According to TimeTree 5 (Kumar *et al*., 2022), the split between *Lebiasina* (Lebiasinidae) and *Boulengerella* (Ctenoluciidae) occurred around 70 million years ago (mya), while the one between *Lebiasina* and *Nannostomus* occurred 65 mya. By that time, with the uplifting of the Andes Mountain range, several river capture events occurred, raising diversification rates and creating an extraordinary richness in South American fishes, mostly through allopatric speciation events (Boschman *et al*., 2023; Cassemiro *et al*., 2023). In contrast, the split between *Pyrrhulina* and *Copella* took place more recently, around 55 mya, during the Paleogene. This is associated with the Paleocene–Eocene thermal maximum, which raised temperatures and significantly increased carbon input into the ocean and atmosphere. Given the unique community structures identified in fossilized ray-finned fish, this split indicates the age of modern fishes (Haynes and Hönisch, 2020). As most Lebiasinidae populations are small and allopatric, both hypotheses support the large cytogenetic diversity observed within the family. The conservation of the ancestral karyotype in Lebiasininae (which, according to morphological data, includes *Derhamia, Lebiasina*, and *Piabucina*) and extensive rearrangements leading to the formation of acrocentrics in the karyotypes of Pyrrhulininae (including *Copeina*, *Copella, Nannostomus*, and *Pyrrhulina*) were the two main evolutionary pathways that occurred in Lebiasinidae, as proposed by Sassi *et al*. (2020). In the light of molecular data, however, a distinct scenario is proposed. In light of recent molecular studies that included Lebiasinidae representatives, the position of *Nannostomus* in Pyrrhulininae is contested (Cassemiro *et al*., 2023) and herein reinforced. Our molecular phylogeny based on DArTseq SNPs suggests that *Nannostomus* is more likely to be a sister genus to *Lebiasina* and these two genera to be sisters of the clade composed of *Copeina*, *Copella*, and *Pyrrhulina* (**Figure 5**). By that, it is possible to infer that Lebiasininae is likely to be composed of *Lebiasina* and *Nannostomus*, while Pyrrhulininae comprises *Copeina, Copella*, and *Pyrrhulina* (**Figure 5**). Although nearly all phylogenetic research on Lebiasinidae is not congruent in terms of species samples, up to date, no species tree includes all recognized species. On the other hand, these investigations provide us with predictions of relationships between genera and occasionally between species. Even though most of their samples consisted of species not included in our analysis, our results line up with those of Casemiro *et al*. (2023), who identified *Pyrrhulina* as a sister genus to *Copella* and *Lebiasina* as a sister genus to *Nannostomus*. Four different *Nannostomus* species—*N. marginatus*, *N.*

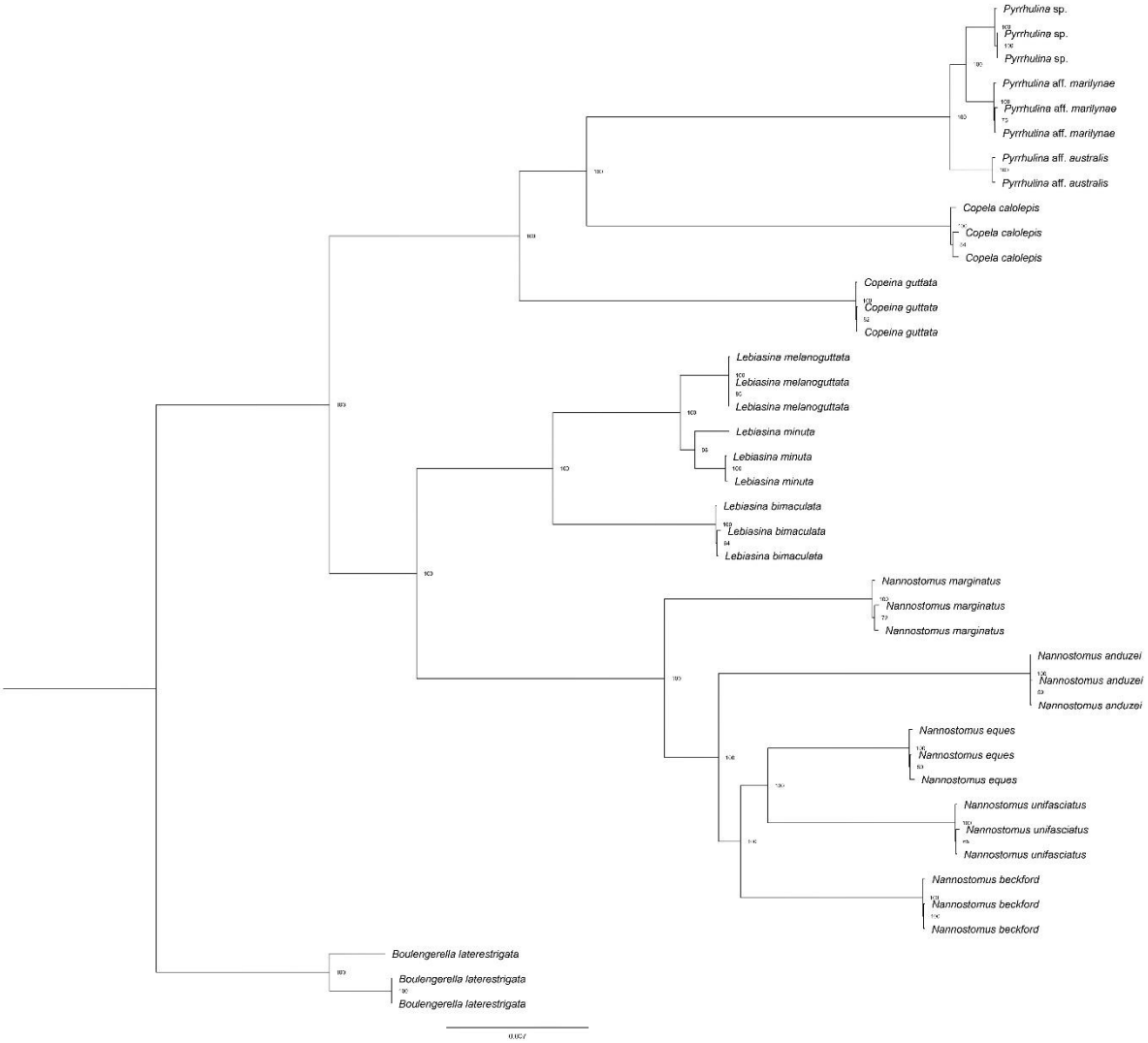
beckfordi, and *N. unifasciatus*—had their mitogenomes constructed in a recent work (Xu *et al*., 2024). The estimated phylogeny tree retrieved *N. beckfordi* and *N. unifasciatus* as closely related, with *N. marginatus* positioned more distant from both. This result is consistent with both the PCA and our phylogenetic analysis (**Figures 4** and **5**).

We can see from these phylogenetic studies that the ancestral 2n of the family (36), which is the same in both Lebiasinidae subfamilies, is shared by both *C. callolepis* and *Lebiasina* species. This indicates that Lebiasinidae represents another Neotropical Characiform family with extensive interspecific karyotype diversity, along with Erythrinidae (Bertollo, 2007; Cioffi *et al*., 2012) and Characidae (Arefjev, 1990; Pazza and Kavalco, 2007; Soto *et al*., 2018). Although the majority of *Nannostomus* and *Pyrrhulina* have higher 2n and a karyotype dominated by acrocentric chromosomes, some species exhibit metacentric chromosomes that result from secondary fusions (Sember *et al*., 2020; Moraes *et al*., 2023). This feature and the molecular phylogeny indicate that the two evolutionary pathways proposed by Sassi *et al*. (2020) may only partially reflect the evolutionary history of Lebiasinidae. With that in mind, we can now focus on a primary pattern of high karyotypic reorganization that results in acrocentrics dominating Pyrrhulininae karyotypes. As for Lebiasininae, while the ancestral karyotype is fully conserved in *Lebiasina* (Sassi *et al*., 2019; Leite *et al*., 2022), it has experienced substantial modification in *Nannostomus*, leading to karyotypes made up of acrocentrics, which in certain species may undergo full secondary fusion into metacentrics. As was already indicated, a small number of studies on *Copella* species have previously reported the 2n with reliable results **(Table 1)** nonetheless, these studies, when combined with our findings, suggest that this genus also experiences processes similar to the evolution of *Nannostomus* karyotypes.

Recent molecular research suggests that *Nannostomus* is classified under the Lebiasininae subfamily (Cassemiro *et al*., 2023). However, the cytogenetic and taxonomy data available so far align with the patterns proposed by Sassi *et al*. (2020). Future investigations on *Derhamia* and other unstudied species might enhance our comprehension of the evolutionary connections within this family. Moreover, the integration of additional morphological, molecular, and cytogenetic studies with a larger sample size might establish a conclusive Lebiasinidae phylogeny, enhancing comprehension of their evolutionary pathways.

Supplementary Materials

Figure S1: Maximum likelihood species tree with all sampled individuals in this study. Values close to each node indicate bootstrap values.



References

All references are compiled in the end of this thesis.

5. Considerações Finais

A família Lebiasinidae é um grupo desafiante por conta das dificuldades de coleta, identificação, obtenção cromossômica e manuseio de animais tão pequenos. Entretanto, ao longo dos anos com os avanços metodológicos foi possível analisar citogenética e molecularmente uma diversidade de espécies da família que contribuíram significativamente para os avanços sobre o grupo. Sendo assim, nesse trabalho foram caracterizadas citogeneticamente cinco espécies de *Pyrrhulina* que resumidamente apresentaram i) um 2n conservado variando entre 40-42 cromossomos; ii) acúmulo de heterocromatina C-positiva principalmente nas regiões pericentroméricas/centroméricas e teloméricas; iii) múltiplos sítios de DNAr 18S e 5S; iiii) cariótipos marcados majoritariamente por cromossomos acrocêntricos. Adicionalmente foi observada a presença de um cariótipo diferenciado entre as espécies estudadas, das quais *P. marilynae* apresentou uma redução cariotípica (2n=32) com a presença de quatro cromossomos metacêntricos grandes e incomuns oriundos de possíveis fusões secundárias.

Os experimentos de hibridização genômica comparativa, pintura total cromossômica e análises de DNA satélite auxiliaram na compreensão das relações entre as espécies de *Pyrrhulina*, a redução cariotípica ocorrida em *P. marilynae* e a origem do único sistema heteromórfico X_1X_2Y observado na família. Assim, a presença do sistema sexual múltiplo observado em *P. semifasciata* foi confirmada por experimentos de WCP, os quais evidenciaram três cromossomos (1m+2a) não pareados e totalmente marcado nos machos da espécie e dois pares de cromossomos totalmente marcados tanto nas fêmeas da espécie como em outras espécies do gênero sugerindo que formação do cromossomo neo-Y desse sistema fosse a fusão de dois cromossomos acrocêntricos não homólogos. Ainda nesse cenário, sequências de DNA satélite foram utilizadas para auxiliar na corroboração da hipótese da origem desse sistema (X_1X_2Y), assim, a presença de poucos sinais sexo específicos e a ausência de sequências exclusivas dos cromossomos sexuais, levaram a crer que o sistema sexual presente em *P. semifasciata* seria evolutivamente recente. Além disso, todos os satDNAs presentes em cromossomos sexuais de *P. semifasciata* hibridizaram em suas espécies próximas, corroborando assim com a possível formação do sistema por fusão cêntrica. Sendo assim, a presença marcante dos cromossomos sexuais heteromórficos em *P. semifasciata* podem ter contribuído significativamente para o aumento da diferenciação genética na espécie.

Com relação a redução cariotípica de *P. marilynae* não foi possível concluir efetivamente as causas desse processo, entretanto as hibridizações genômicas comparativas mostraram que

tal espécie apresenta uma gama de regiões espécie-especificas indicando uma diferente tendência evolutiva quando comparada com outras espécies do gênero. Além disso, foi observado também um acúmulo de satDNAs na região centromérica dos cromossomos da espécie sugerindo que tais sequências podem estar relacionadas as possíveis fusões que geraram a redução do cariótipo de *P. marilynae*. Em um contexto geral foi possível observar que embora as nove espécies de *Pyrrhulina* compartilhem uma gama de características cariotípicas e genômicas, foram evidenciadas uma série de particularidades em cada uma das espécies, identificando diferentes processos evolutivos entre elas.

Por fim foi caracterizado pela primeira vez o cariótipo de *Copella* que preservou o 2n ancestral proposto para família (2n=36), mas não sua estrutura cariotípica. Em contraste com o cariótipo ancestral que era composto principalmente de cromossomos metacêntricos, *C. callolepis* exibiu um grande número de acrocêntricos, reforçando o importante papel das inversões pericêntricas na evolução do cariótipo do grupo. Sabe-se que duas tendências evolutivas foram sugeridas, para as subfamílias de Lebiasinidae: i) um cariótipo conservado com 2n=36 composto exclusivamente por cromossomos meta/submetacêntricos representando a condição basal do grupo, como visto em *Lebiasina*; e ii) um cariótipo com 2n mais elevado, rearranjos cromossômicos estruturais e um cariótipo majoritariamente acrocêntrico. Sendo assim, teríamos a subfamília Lebiasininae composta por *Lebiasina* e *Dehami*, e Pyrrhulilinae composta por *Copeina*, *Copella*, *Nannostomus* e *Pyrrhulina*. Embora essas tendências tenham sido reafirmadas mais de uma vez em diferentes estudos, e os dados citogenéticos e taxonômicos até então relatados corroborem com essas tendências, nossos estudos moleculares propuseram algumas mudanças nesse cenário, tais como a inclusão de *Nannostomus* dentro de Lebiasininae. Mesmo com o conjunto de todos esses dados ainda sim fica evidente a necessidade de maiores estudos envolvendo principalmente as relações internas da família, bem como a inclusão de estudo citogenéticos envolvendo *Dehamia*, dados esses ausentes até o presente momento.

6. Referências bibliográficas

Ahmad, S.F., Singchat, W., Jehangir, M., *et al*. (2020). Dark Matter of Primate Genomes: Satellite DNA Repeats and Their Evolutionary Dynamics. *Cells* 9, 2714. doi:10.3390/cells9122714

Albert, J.S., Craig, J.M., Tagliacollo, V.A., *et al*. (2018). Upland and Lowland Fishes: A Test of the River Capture Hypothesis. *Mt. Clim. Biodivers* 273.

Albert, J.S., Tagliacollo, V.A., Dagosta, F. (2020). Diversification of Neotropical Freshwater Fishes. *Annual Review of Ecology, Evolution, and systematics* 51(1), 27-53.

Albert, J.S., Bernt, M.J., Fronk, A.H., *et al*. (2021). Late Neogene megariver captures and the Great Amazonian biotic interchange. *Global and Planetary Change* 205, 103554. doi: 10.1016/j.gloplacha.2021.103554

Alkimim, E.R., Caixeta, E.T., Sousa, T.V., *et al*. (2018). High-throughput Targeted Genotyping Using Next-Generation Sequencing Applied in Coffea Canephora Breeding. *Euphytica* 214(3), 1–18. doi:10.1007/s10681-018-2126-2

Allen, W.R., Ternetz, C., Meyers, G.S., and Myers, G.S. (1926). Descriptions of a New Characin Fish and a New Pygidiid Catfish from the Amazon Basin. *Copeia* 156, 150–152. doi:10.2307/1436719

Altschul, S.F., Gish, W., Miller, W., *et al*. (1990). Basic local alignment search tool. *J. Mol. Biol*. 215, 403-410. doi: 10.1016/S0022-2836(05)80360-2

Ansai, S., Toyoda, A., Yoshida, K., *et al*. (2023). Repositioning of centromere-associated repeats during karyotype evolution in *Oryzias* fishes. *Mol. Ecol.* doi: 10.1111/mec.17222

Arai, R., 2011. Fish karyotypes: a check list. *Springer Science and Business Media*

Arcila, D., Ortí, G., Vari, R., *et al*. (2017). Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. *Nat. Ecol. Evol*. 1, 0020. doi: 10.1038/s41559-016-0020

Arefjev, V. (1990). Problems of karyotypic variability in the family Characidae (Pisces, Characiformes) with the description of somatic karyotypes for six species of tetras. *Caryologia* 43, 305–319. doi: 10.1080/00087114.1990.10797009

Balini, L. C., Fernandes, C. A., Portela-Castro, A. L. D. B., *et al*. (2024). Initial Steps of XY Sex Chromosome Differentiation in the Armored Catfish *Hypostomus albopunctatus* (Siluriformes: Loricariidae) Revealed by Heterochromatin Accumulation. *Zebrafish* 21(3), 265-273. doi: 10.1089/zeb.2023.0100

Barby, F.F., Bertollo, L.A.C., de Oliveira, E.A., *et al*. (2019). Emerging patterns of genome organization in Notopteridae species (Teleostei, Osteoglossiformes) as revealed by Zoo-FISH and Comparative Genomic Hybridization (CGH). *Sci. Rep.* 9, 1112.

Bardella, V.B., Milani, D., Cabral-de-Mello, D.C. (2020). Analysis of *Holhymenia histrio* genome provides insight into the satDNA evolution in an insect with holocentric chromosomes. *Chromosome Res*. 28, 369-380. doi: 10.1007/s10577-020-09642-1

Basset, P., Yannic, G., and Hausser, J. (2008). Chromosomal Rearrangements and Genetic Structure at Different Evolutionary Levels of the Sorex araneus Group. *J. Evol. Biol.* 21, 842–852. doi:10.1111/j.1420-9101.2008.01506.x

Bateson, W. "Heredity and variation in modern lights". In Seward A.C (ed.) **Darwin and modern Science**. Cambridge: Cambridge University Press, pp 85-101, 1909.

Beaudry, F.E.G., Barrett, S.C.H., Wright, S.I. (2019). Ancestral and Neo-sex Chromosomes Contribute to Population Divergence in a Dioecious Plant. *Evolution* 74(2), 256–269. doi:10.1111/evo.13892

Benzaquem, D.C., Oliveira, C., da Silva Batista, J., *et al*. (2015). DNA Barcoding in Pencilfishes (Lebiasinidae: *Nannostomus*) Reveals Cryptic Diversity across the Brazilian Amazon. *Plos One* 10, e0112217. doi: 10.1371/journal.pone.0112217

Bertaco, V.A., Ferrer, J., Carvalho, F.R., *et al*. (2016). Inventory of the Freshwater Fishes from

a Densely Collected Area in South America-a Case Study of the Current Knowledge of Neotropical Fish Diversity. *Zootaxa* 4138(3), 401–440. doi:10.11646/ZOOTAXA.4138.3.1

Bertollo, L.A.C., Cioffi, M.B., Moreira-Filho, O. "Direct chromosome preparation from freshwater teleost fishes". In: Ozouf-Costaz, C., Pisano, E., Foresti, F.; Almeida-Toledo, L.F. (eds.) **Fish Cytogenetic Techniques (Chondrichthyans and Teleosts).** Florida: CRC Press, pp. 21–26, 2015.

Bertollo, L.A.C., Fontes, M.S., Fenocchio, A.S., *et al*. (1997). The X 1 X 2 Y sex chromosome system in the fish *Hoplias malabaricus*. I. G-, C-and chromosome replication banding. *Chromosome Res.* 5, 493-499. doi: 10.1023/A:1018477232354

Bertollo, L.A.C., Born, G.G., Dergam, J.A., *et al*. (2000). A biodiversity approach in the neotropical Erythrinidae fish, *Hoplias malabaricus*. Karyotypic survey, geographic distribution of cytotypes and cytotaxonomic considerations. *Chromosom. Res.* 8, 603–61.

Bertollo, L.A.C., Oliveira, C., Molina, W.F., *et al*. (2004). Chromosome evolution in the erythrinid fish, *Erythrinus erythrinus* (Teleostei: Characiformes). *Heredity* 93, 228-233. doi: 10.1038/sj.hdy.6800511

Bertollo, L. (1978). Cytotaxonomic considerations on *Hoplias lacerdae* (Pisces, Erythrinidae). *Braz. J. Genet.* 1, 103–120.

Bertollo, L.A.C. (2007). Chromosome evolution in the Neotropical Erythrinidae fish family: an overview. *Fish Cytogenet*. 195, 211.

Betancur-R, R., Arcila, D., Vari, R.P., *et al.* (2019). Phylogenomic incongruence, hypothesis testing, and taxonomic sampling: The monophyly of characiform fishes. *Evolution* 73, 329–345. doi: 10.1111/evo.13649

Biémont, C. and Vieira, C. (2006). Junk DNA as an evolutionary force. *Nature* 443, 521-524. doi: 10.1038/443521a

Blackburn, E.H. (1994). Telomeres: no end in sight. *Cell* 77(5), 621-623.

Blanco, D.R., Vicari, M.R., Artoni, R.F., *et al*. (2012). Chromosomal Characterization of Armored Catfish *Harttia longipinna* (Siluriformes, Loricariidae): First Report of B Chromosomes in the Genus. *Zoolog. Sci*. 29(9), 604–609. doi:10.2108/zsj.29.604

Blough, R.I., Heerema, N.A., Ulbright, T.M., *et al*. (1998). Interphase chromosome painting of paraffin-embedded tissue in the differential diagnosis of possible germ cell tumors. *Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc* 11(7), 634-641.

Bolger, A.M., Lohse, M., Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114-2120. doi: 10.1093/bioinformatics/btu170

Bolzán, A.D. (2017). Interstitial telomeric sequences in vertebrate chromosomes: Origin, function, instability and evolution. *Mutation research/Reviews in mutation research* 773, 51-65. doi: 10.1016/j.mrrev.2017.04.002

Boschman, L.M., Carraro, L., Cassemiro, F.A.S., *et al.* (2023). Freshwater fish diversity in the western Amazon basin shaped by Andean uplift since the Late Cretaceous. *Nat. Ecol. Evol.* 7, 2037–2044. doi: 10.1038/s41559-023-02220-8

Bouckaert, R., Heled, J., Kühnert, D., *et al*. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *Plos computational biology* 10(4), e1003537. doi: 10.1371/journal.pcbi.1003537

Brajković, J., Feliciello, I., Bruvo-Mađarić, B., *et al*. (2012). Satellite DNA-like elements associated with genes within euchromatin of the beetle Tribolium castaneum. *G3: Genes Genomes Genetics*, 2(8), 931-941. doi: 10.1534/g3.112.003467

Bronstein, O., Kroh, A., Haring, E. (2018). Mind the Gap! The Mitochondrial Control Region and Its Power as a Phylogenetic Marker in Echinoids. *BMC Evolutionary Biology* 18(1), 80. doi:10.1186/s12862-018-1198-x.

Bryant, D. and Moulton, V. (2004). Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol*. 21, 255–265. doi:10.1093/molbev/msh018

Buckup, P.A., Malabarba, L.R., Reis, R.E., *et al*. (1998). Relationships of the Characidiinae and phylogeny of characiform fishes (Teleostei: Ostariophysi). *Phylogeny and classification of neotropical fishes*, 123-144.

Cabral-de-Mello, D.C., Zrzavá, M., Kubíčková, S., *et al*. (2021). The role of satellite DNAs in genome architecture and sex chromosome evolution in *Crambidae* moths. *Front in Genet*. 12, 661417. doi: 10.3389/fgene.2020.661417

Cabral-de-Melo, D.C., Mora, P., Rico-Porras, J.M., *et al*. (2023). The spread of satellite DNAs in euchromatin and insights into the multiple sex chromosome evolution in Hemiptera revealed by repeatome analysis of the bug *Oxycarenus hyalinipennis*. *Insect Mol Biol.*12868. doi: 10.1111/imb.12868

Calcagnotto, D., Schaefer, S.A., DeSalle, R. (2005). Relationships among characiform fishes inferred from analysis of nuclear and mitochondrial gene sequences. *Mol. Phylogenet. Evol*. 36, 135–153. doi: 10.1016/j.ympev.2005.01.004

Camacho, J.P.M., Cabrero, J., López-León, M.D., *et al*. (2022). Satellitome comparison of two oedipodine grasshoppers highlights the contingent nature of satellite DNA evolution. *BMC Biol*. 20, 1-24. doi: 10.1186/s12915-021-01216-9

Carvalho, P.C., de Oliveira, E.A., Bertollo, L.A.C., *et al*. (2017). First chromosomal analysis in Hepsetidae (Actinopterygii, Characiformes): Insights into relationship between African and Neotropical fish groups. *Front. Genet.* 8, 203. doi:10.3389/fgene.2017.00203.

Cassemiro, F.A., Albert, J.S., Antonelli, A., *et al*. (2023). Landscape dynamics and diversification of the megadiverse South American freshwater fish fauna. *Proc. Natl. Acad. Sci.* 120, e2211974120. doi: 10.1073/pnas.2211974120

Castro, R., and Polaz, C.N. (2020). Small-sized fish: the largest and most threatened portion of the megadiverse neotropical freshwater fish fauna. *Biota Neotropica* 20, e20180683. doi: 10.1590/1676-0611-BN-2018-0683

Castro, J.P., Moura, M.O., Moreira-Filho, O., *et al*. (2015). Diversity of the *Astyanax scabripinnis* species complex (Teleostei: Characidae) in the Atlantic Forest, Brazil: species limits and evolutionary inferences. *Rev. Fish Biol. Fish*. 25, 231–244. doi: 10.1007/s11160-014-9377-3

Charlesworth, B. *Evolution in age-structured populations*. Cambridge: Cambridge University Press, 1994.

Cioffi, M.B., Kejnovský, E., Marquioni, V., *et al*. (2012). The key role of repeated DNAs in sex chromosome evolution in two fish species with ZW sex chromosome system. *Mol. Cytogenet*. 5, 1-7. doi: 10.1186/1755-8166-5-28

Cioffi, M.B., Martins, C., Centofante, L., *et al*. (2009). Chromosomal Variability among Allopatric Populations of Erythrinidae Fish *Hoplias malabaricus*: Mapping of Three Classes of Repetitive DNAs. *Cytogenet. Genome Res.* 125, 132–141. doi: 10.1159/000227838

Cioffi M.B., Martins C., Bertollo L.A.C. (2010). Chromosome spreading of associated transposable elements and ribosomal DNA in the fish *Erythrinus erythrinus*. Implications for genome change and karyoevolution in fish. *BMC Evol Biol*. 10, 271. doi: 10.1186/1471-2148-10-271.

Cioffi, M.B., Sánchez, A., Marchal, J.A., *et al*. (2011). Whole chromosome painting reveals independent origin of sex chromosomes in closely related forms of a fish species. *Genetica*, 139(8), 1065-1072.

Cioffi, M.B. and Bertollo, L. (2012). Chromosomal distribution and evolution of repetitive DNAs in fish. *Repetitive DNA* 7, 197–221.

Cioffi, M.B., Moreira-Filho, O., Ráb, P., *et al*. (2018). Conventional cytogenetic approaches—Useful and indispensable tools in discovering fish biodiversity. *Curr. Genet. Med. Rep.* 6, 176–186.

Cioffi, M.B., Ráb, P., Ezaz, T., *et al*. (2019). Deciphering the evolutionary history of arowana fishes (Teleostei, Osteoglossiformes, Osteoglossidae): Insight from comparative cytogenomics. *Int. J. Mol. Sci*. 20. doi: 10.3390/ijms20174296

Coyne, J.A. and Orr, H.A. (2004). Speciation: a catalogue and critique of species concepts. *Philosophy of biology: an anthology*, 272-92.

Crepaldi, C., Martí, E., Gonçalves, E.M., *et al*. (2021). Genomic differences between the sexes in a fish species seen through satellite DNAs. *Front. Genet*. 12, 728670. doi: 10.3389/fgene.2021.728670

da Silva, M.J., Gazoni, T., Haddad, C.F.B., *et al*. (2023). Analysis in *Proceratophrys boiei* genome illuminates the satellite DNA content in a frog from the Brazilian atlantic forest. *Front. Genet*. 14, 1101397. doi: 10.3389/fgene.2023.1101397

Dagosta, F.C. and De Pinna, M. (2019). The fishes of the Amazon: distribution and biogeographical patterns, with a comprehensive list of species. *Bulletin of the American Museum of Natural History*, 2019(431), 1-163.

Dagosta, F.C.P. and De Pinna, M.C.C. (2021). Two New Catfish Species of Typically Amazonian Lineages in the Upper Rio Paraguay (Aspredinidae: Hoplomyzontinae and Trichomycteridae: Vandelliinae), with a Biogeographic Discussion. *Pap. Avulsos Zool*. 61, e20216147. doi:10.11606/1807-0205/2021.61.47.

Dai, W., Zou, M., Yang, L. *et al.* (2018). Phylogenomic Perspective on the Relationships and Evolutionary History of the Major Otocephalan Lineages. *Sci Rep* 8, 205. doi: 10.1038/s41598-017-18432-5

Darriba, D., Posada, D., Kozlov, A.M., *et al*. (2020). ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Mol. Biol. Evol.* 37, 291–294. doi: 10.1093/molbev/msz189

Davey, J.W., Hohenlohe, P.A., Etter, P.D., *et al*. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12, 499–510. doi: 10.1038/nrg3012

de Lima, L.G. and Ruiz-Ruano, F.J. (2022). In-depth satellitome analyses of *Drosophila* species illuminate repetitive DNA evolution in the *Drosophila* genus. *Genome Biol. Evol.* 14, evac064. doi: 10.1093/gbe/evac064

de Oliveira, E.A., Bertollo, L.A.C., Rab, P., *et al*. (2019). Cytogenetics, genomics and biodiversity of the South American and African Arapaimidae fish family (Teleostei, Osteoglossiformes). Plos One 14, e0214225. doi: 10.1371/journal.pone.0214225

de Pinna, M., Zuanon, J., Rapp Py-Daniel, L., *et al*. (2018). A new family of neotropical freshwater fishes from deep fossorial Amazonian habitat, with a reappraisal of morphological characiform phylogeny (Teleostei: Ostariophysi). *Zool. J. Linn. Soc.* 182, 76–106.

de Souza e Sousa, J.F., Guimarães, E., Pinheiro-Figliuolo, V., *et al*. (2021). Chromosomal Analysis of Ctenolucius hujeta Valenciennes, 1850 (Characiformes): A New Piece in the Chromosomal Evolution of the Ctenoluciidae. *Cytogenet. Genome Res.*, 1–8. doi:10.1159/000515456.

de Souza e Sousa, J.F., Viana, P.F., Bertollo, L.A.C., *et al*. (2017). Evolutionary relationships among *Boulengerella* species (Ctenoluciidae, Characiformes): genomic organization of repetitive DNAs and highly conserved karyotypes. *Cytogenet. Genome Res.* 152, 194–203. doi:10.1159/000480141.

Ditcharoen, S., Sassi, F.M.C., Bertollo, L.A.C., *et al*. (2020). Comparative chromosomal mapping of microsatellite repeats reveals divergent patterns of accumulation in Siluridae

(Teleostei: Siluriformes) species. *Genet. Mol. Biol.* 43. doi:10.1590/1678-4685-GMB-2020-0091.
Dobzhansky, T. (1933). On the sterility of the interracial hybrids in *Drosophila pseudoobscura*. *Proc Natl Acad Sci USA* 19, 397-403. doi:10.1073/pnas.19.4.397
Dornburg, A. and Near, T.J. (2021). The emerging phylogenetic perspective on the evolution of actinopterygian fishes. *Annual Review of Ecology, Evolution, and Systematics*, 52(1), 427-452. doi: 10.1146/annurev-ecolsys-122120-122554
dos Santos, L.L., Benone, N.L., Brasil, L.S., *et al*. (2022). The use of taxonomic families as biological surrogates of the diversity of the Amazonian stream fish. *Ecol Ind* 141, 109094. doi: 10.1016/j.ecolind.2022.109094
dos Santos, R.Z., Calegari, R.M., Silva, D.M.Z.D.A., *et al*. (2021). A long-term conserved satellite DNA that remains unexpanded in several genomes of Characiformes fish is actively transcribed. *Genome Biol Evol*. 13, evab002. doi: 10.1093/gbe/evab002
Eaton, D.A.R. and Overcast, I. (2017). Ipyrad V. 0.7. Available at: https://github.com/dereneaton/ipyrad
Eaton, D.A.R. (2014). PyRAD: Assembly of De Novo RADseq Loci for Phylogenetic Analyses. *Bioinformatics* 30(13), 1844–1849. doi:10.1093/bioinformatics/btu121
Eigenmann, C.H. and Kennedy, C.H. (1903). On a Collection of Fishes from Paraguay, with a Synopsis of the American Genera of Cichlids. *Proc. Acad. Nat. Sci. Phila.* 55, 497–537.
Ekblom, R. and Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107, 1–15. doi: 10.1038/hdy.2010.152
Ellermeier, C., Higuchi, E.C., Phadnis, N., *et al*. (2010). RNAi and heterochromatin repress centromeric meiotic recombination. *Proc. Natl. Acad. Sci*. 107, 8701–8705. doi: 10.1073/pnas.0914160107
Escudeiro, A., Ferreira, D., Mendes-da-Silva, A., *et al*. (2019). Bovine satellite DNAs–a history of the evolution of complexity and its impact in the Bovidae family. *The European Zool. Journal* 86, 20-37. doi: 10.1080/24750263.2018.1558294
Faircloth, B.C., Alda, F., Hoekzema, K., *et al*. (2020). A target enrichment bait set for studying relationships among ostariophysan fishes. *Copeia*, 108(1), 47-60.
Faria, R. and Navarro, A. (2010). Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends in Ecology and Evolution* 25, 660-669. doi: 10.1016/j.tree.2010.07.008
Feliciello, I., Akrap, I., Ugarković, Đ. (2015). Satellite DNA modulates gene expression in the beetle *Tribolium castaneum* after heat stress. *Plos Genetics* 11, e100566. doi: 10.1371/journal.pgen.1005547
Feliciello, I., Pezer, Ž., Kordiš, D., *et al*. (2020). Evolutionary history of alpha satellite DNA repeats dispersed within human genome euchromatin. *Genome biology and evolution*, 12(11), 2125-2138. doi: 10.1093/gbe/evaa224
Ferreira, P.H.N., Souza, F.H.S, Moraes, R.L.R.M., *et al*. (2022). The genetic differentiation of *Pyrrhulina* (Teleostei, Characiformes) species is likely influenced by both geographical distribution and chromosomal rearrangements. *Front Genet*. 13, 869073. doi:10.3389/fgene.2022.869073
Ferretti, A.B.S.M., Milani, D., Palacios-Gimenez, O.M., *et al*. (2020). High dynamism for neo-sex chromosomes: satellite DNAs reveal complex evolution in a grasshopper. *Heredity* 125, 124-137. doi: 10.1038/s41437-020-0327-7
Flynn, J.M., Hu, K.B., Clark, A.G. (2023). Three recent sex chromosome-to-autosome fusions in a *Drosophila virilis* strain with high satellite DNA content. *Genetics* 224(2), iyad062. doi: 10.1093/genetics/iyad062

Foll, M. and Gaggiotti, O. (2008). A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: a Bayesian Perspective. *Genetics* 180(2), 977–993. doi:10.1534/genetics.108.092221

Foresti, F., Almeida-Toledo, L.F., Toledo-Filho, S.A. (1981). Polymorphic nature of nucleolus organizer regions in fishes. *Cytogenet. Genome Res*. 31, 137–144. doi: 10.1159/000131639

Franchini, P., Irisarri, I., Fudickar, A., *et al*. (2017). Animal tracking meets migration genomics: transcriptomic analysis of a partially migratory bird species. *Mol. Ecol.* 26, 3204-3216. doi: 10.1111/mec.14108

Franchini, P., Kautt, A.F., Nater, A., *et al*. (2020). Reconstructing the evolutionary history of chromosomal races on islands: a genome-wide analysis of natural house mouse populations. *Molecular Biology and Evolution* 37, 2825-2837. doi: 10.1093/molbev/msaa118

Freitas, N.L., Al-Rikabi, A.B.H., Bertollo, L.A.C., *et al*. (2018). Early stages of XY sex chromosomes differentiation in the fish *Hoplias malabaricus* (Characiformes, Erythrinidae) revealed by DNA repeats accumulation. *Curr. Genomics* 19, 216–226. doi:10.2174/1389202918666170711160528.

Fricke, R., Eschmeyer, W. N., van der Laan, R. Eschmeyer's Catalog of Fishes: Genera, Species, References. 2021 Available online: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.as (accessed on 2 September 2021).

Fricke, R., Eschmeyer, W.N., van der Laan, R. Eschmeyer's Catalog of Fishes: Genera, Species, References. 2023. Available online: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp. (accessed on 12 July 2023).

Fricke, R., Eschmeyer, W.N., van der Laan, R. Eschmeyer's catalog of fishes: Genera, Species, References. 2024. Available online: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (accessed on 24 may 2024).

Froese, R. and Pauly, D. FishBase: World Wide Web electronic publication. 2014. Available online: http://www.fishbase.org (accessed on 22 april 2014).

Froese, R. and Pauly, D. FishBase: World Wide Web electronic publication. 2018. Available online: http://www.fishbase.org (accessed on 14 april 2018).

Froese, R. and Pauly, D. Fish Base: World Wide Web electronic publication. 2023. Available online: http://www.fishbase.org. (accessed on 22 July 2023).

Froese, R. and Pauly, D. Fish Base: World Wide Web electronic publication. 2024. Available online: http://www.fishbase.org. (accessed on 26 July 2024).

Fry, K. and Salser, W. (1977). Nucleotide sequences of HS-α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* 12, 1069-1084. doi: 10.1016/0092-867(77)90170-2

Galetti Jr, P.M., Mestriner, C.A., Venere, P.C., *et al*. (1991). Heterochromatin and karyotype reorganization in fish of the family Anostomidae (Characiformes). *Cytogenetic and Genome Research* 56(2), 116-121.

Garcia, S., Kovarik, A., Maiwald, S., *et al*. (2024). The dynamic interplay between ribosomal DNA and transposable elements: a perspective from genomics and cytogenetics. *Mol. Biol. Evol.* 41, msae025. doi: 10.1093/molbev/msae025

Garrido-Ramos, M.A. (2017). Satellite DNA: an evolving topic. *Genes* 8, 230. doi: 10.3390/genes8090230

Gatto, K.P., Mattos, J.V., Seger. K.R., *et al*. (2018). Sex chromosome differentiation in the frog genus *Pseudis* involves satellite DNA and chromosome rearrangements. *Front. Genet.* 9, 301. doi: 10.3389/fgene.2018.00301

Géry, J. and Zarske, A. (2002). *Derhamia hoffmannorum* gen. et sp. n.-a new pencil fish

(Teleostei, Characiformes, Lebiasinidae), endemic from the Mazaruni River in Guyana. *Zool. Abh.-Staatl. Mus. Tierkd. Dresd*. 52, 35–47.
Géry, J. (1997). Characoids of the world.
Glugoski, L., Deon, G., Schott, S., *et al*. (2020). Comparative cytogenetic analyses in Ancistrus species (Siluriformes: Loricariidae). *Neotrop. Ichthyol.* 18, 1–16.
Glugoski, L., Nogaroto, V., Deon, G.A., *et al.* (2022). Enriched tandem repeats in chromosomal fusion points of *Rineloricaria latirostris* (Boulenger, 1900) (Siluriformes: Loricariidae). *Genome* 65, 479-489. doi:10.1139/gen-2022-0043
Glugoski, L., Deon, G.A., Nogaroto, V., *et al*. (2023). Robertsonian fusion site in *Rineloricaria pentamaculata (Siluriformes: Loricariidae)*: involvement of 5S rDNA and satellite sequences. *Cytogenet. Genome Res.* doi: 10.1159/000530636
Goes, C.A.G., Santos, R.Z., Aguiar, W.R.C., *et al. (2023).* Revealing the satellite DNA history in *Psalidodon* and *Astyanax* characid fish by comparative satellitomics. *Front Genet.* 13, 884072. doi: 10.3389/fgene.2022.884072
Gornung, E. (2013). Twenty years of physical mapping of major ribosomal RNA genes across the teleosts: A review of research. *Cytogenet. Genome Res.* 141, 90–102. doi:10.1159/000354832
Gruber, B., Unmack, P.J., Berry, O.F., *et al.* (2018). dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol. Ecol. Resour.* 18, 691–699. doi:10.1111/1755-0998.12745
Gržan, T., Dombi, M., Despot-Slade, E., *et al.* (2023). The Low-Copy-Number Satellite DNAs of the Model Beetle *Tribolium castaneum*. *Genes* 14, 999. doi: 10.3390/genes14050999
Guarracino, A., Buonaiuto, S., de Lima, L.G., *et al*. (2023). Recombination between heterologous human acrocentric chromosomes. *Nature* 617, 335-343. doi: 10.1038/s41586-023-05976-y
Guerra, M.S. (1988). *Introdução à citogenética geral*. Guanabara Koogan
Guzmán-Markevich, K., Roco, A.S., Ruiz-García, A., *et al*. (2022). Cytogenetic analysis in the toad species *Bufo spinosus, Bufotes viridis and Epidalea calamita (Anura, Bufonidae)* from the mediterranean area. *Genes* 13, 1475. doi: 10.3390/genes13081475
Haerter, C. A. G., Blanco, D. R., Traldi, J. B., *et al.* (2023). Are scattered microsatellites weak chromosomal markers? Guided mapping reveals new insights into Trachelyopterus (Siluriformes: Auchenipteridae) diversity. *Plos one* 18(6), e0285388. doi: 10.1371/journal.pone.0285388
Haldane, J. B. S. (1922). Sex Ratio and Unisexual Sterility in Hybrid Animals. *J. Gen.* 12, 101–109. doi:10.1007/bf02983075
Haynes, L.L. and Hönisch, B. (2020). The seawater carbon inventory at the Paleocene–Eocene Thermal Maximum. *Proc. Natl. Acad. Sci.* 117, 24088–24095. doi:10.1073/pnas.2003197117
Henikoff , S., Ahmad, K., Malik, H.S. (2001). The centromere paradox: Stable inheritance with rapidly evolving DNA. *Science* 293, 1098-1102. doi: 10.1126/science.1062939
Howell, W.M.T. and Black, D.A. (1980). Controlled silver-staining of nucleolus organizer regions with a protective colloidal developer: a 1-step method. *Experientia* 36.8: 1014-1015.
Ijdo, J. W., Wells, R. A., Baldini, A., *et al*. (1991). Improved telomere detection using a telomere repeat probe (TTAGGG) n generated by PCR. *Nucleic Acids Res.* 19, 4780
Imoto, J.M., Saitoh, K., Sasaki, T., *et al.* (2013). Phylogeny and Biogeography of Highly Diverged Freshwater Fish Species (Leuciscinae, Cyprinidae, Teleostei) Inferred from Mitochondrial Genome Analysis. *Gene* 514(2), 112–24. doi:10.1016/j.gene.2012.10.019.
Jaccoud, D., Peng, K., Feinstein, D., *et al.* (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res.* 29, e25–e25.

doi:10.1093/nar/29.4.e25
Kejnovský, E., Michalovova, M., Steflova, P., *et al*. (2013). Expansion of microsatellites on evolutionary young Y chromosome. *Plos One* 8, e45519. doi:10.1371/journal.pone.0045519.
Khosraviani, N., Ostrowski, L.A., Mekhail, K. (2019). Roles for non-coding RNAs in spatial genome organization. *Front. Cell Develop. Biol.* 7, 336. doi: 10.3389/fcell.2019.00336
Kilian, A., Wenzl, P., Huttner, E., *et al.* (2012). Diversity arrays technology: a generic genome profiling technology on open platforms. *Data Prod. Anal. Popul. Genomics Methods Protoc.* 67–89. doi:10.1007/978-1-61779-870-2_5
Kipling, D., Ackford, H.E., Taylor, B.A., *et al*. (1991). Mouse minor satellite DNA genetically maps to the centromere and is physically linked to the proximal telomere. *Genomics* 11, 235-241. doi:10.1016/0888-7543(91)90128-2
Kitano, J. and Peichel, C.L. (2012). Turnover of Sex Chromosomes and Speciation in Fishes. *Environ. Biol. Fish.* 94(3), 549–558. doi:10.1007/s10641-011-9853-8
Kitano, J., Ross, J. A., Mori, S., *et al*. (2009). A Role for a Neo-Sex Chromosome in Stickleback Speciation. *Nature* 461(7267), 1079–1083. doi:10.1038/nature08441
Kopecna, O., Kubickova, S., Cernohorska, H., *et al*. (2014). Tribe-specific satellite DNA in non-domestic Bovidae. *Chromosome Res.* 22, 277-291. doi: 10.1007/s10577-014-9401-4
Kopelman, N.M., Mayzel, J., Jakobsson, M., *et al*. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular ecology resources*, 15(5), 1179-1191. doi: 10.1111/1755-0998.12387
Kratochvíl, L., Stöck, M., Rovatsos, M., *et al*. (2021). Expanding the classical paradigm: what we have learnt from vertebrates about sex chromosome Evolution. *Phil. Trans. R. Soc. B.* 376, 20200097. doi:10.1098/rstb.2020.0097
Krysanov, E., Yu., Nagy, B., Watters, B.R., *et al*. (2023). Karyotype differentiation in the *Nothobranchius ugandensis* species group (Teleostei, Cyprinodontiformes), seasonal fishes from the east African inland plateau, in the context of phylogeny and biogeography. *Comp. Cytogenet.* 17, 13–29. doi:10.3897/compcytogen.v7.i1.97165
Kubat, Z., Hobza, R., Vyskot, B., *et al*. (2008). Microsatellite accumulation in the Y chromosome of Silene latifolia. *Genome* 51, 350-356. doi:10.1139/G08-024.
Kuhn, G.C., Küttler, H., Moreira-Filho, O., *et al*. (2012). The repetitive DNA of *Drosophila*: concerted evolution at different genomic scales and association with genes. *Mol. Biol. Evol.* 29, 7-11. doi:10.1093/molbev/msr173
Kumar, S., Suleski, M., Craig, J.M., *et al.* (2022). TimeTree 5: An Expanded Resource for Species Divergence Times. *Mol. Biol. Evol.* 39, msac174. doi: 10.1093/molbev/msac174
Lafontaine, D.L.J. and Tollervey, D. (2001). The function and synthesis of ribosomes. *Nature Reviews Molecular Cell Biology* 2(7), 514.
Leite, P.P.D.M., Sassi, F.M.C., Marinho, M.M.F., *et al.* (2022). Tracking the evolutionary pathways among Brazilian *Lebiasina* species (Teleostei: Lebiasinidae): a chromosomal and genomic comparative investigation. *Neotropical Ichthyol.* 20, e210153. doi:10.1590/1982-0224-2021-0153
Levan, A. (1964). Nomenclature for centromeric position on chromosomes. *Hereditas.* 52, 201–220. doi:10.1111/j.1601-5223.1964.tb01953.x
Lindholm, A. and Breden, F. (2002). Sex Chromosomes and Sexual Selection in Poeciliid Fishes. *Am. Nat.* 160(S6), S214–S224. doi:10.1086/342898
Lisachov, A., Rumyantsev, A., Prokopov, D., *et al*. (2023). Conservation of major satellite DNAs in snake heterochromatin. *Animals* 13, 334. doi: 10.3390/ani13030334
Liu, H., Sun, C., Zhu, Y., *et al*. (2020). Mitochondrial genomes of four American characins and phylogenetic relationships within the family Characidae (Teleostei: Characiformes). *Gene* 762, 145041. doi: 10.1016/j.gene.2020.145041

López-Flores, I., Garrido-Ramos, M.A. “The repetitive DNA content of eukaryotic genomes”. In Garrido-Ramos, M.A (ed.) **Repetitive DNA**. Switzerland: Karger Publishers, pp. 1–28, 2012.

Louzada, S., Lopes, M., Ferreira, D., *et a*. (2020). Decoding the role of satellite DNA in genome architecture and plasticity—An evolutionary and clinical affair. *Genes* 11, 72. doi:10.3390/genes11010072

Lui, R.L., Blanco, D.R., Moreira-Filho, O., *et al.* (2012). Propidium iodide for making heterochromatin more evident in the C-banding technique. *Biotech. Histochem.* 87, 433–43.

Machado, C.A., Haselkorn, T.S., Noor, M.A.F. (2007). Evaluation of the Genomic Extent of Effects of Fixed Inversion Differences on Intraspecific Variation and Interspecific Gene Flow in *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 175 (3), 1289–1306. doi:10.1534/genetics.106.064758

Malabarba, L.R. and Malabarba, M.C. “Phylogeny and classification of Neotropical fish”. In: Baldisseroto, B., Urbinati, E.C., Cyrino, J.E.P. (eds.) **Biology and physiology of freshwater neotropical fish**. Amsterdam: Elsevier, 2020.

Marajó, L., Viana, P.F., Ferreira, A.M.V., *et al*. (2023). Chromosomal rearrangements and the first indication of an ♀X1X1X2X2/♂X1X2Y sex chromosome system in *Rineloricaria* fishes (Teleostei: Siluriformes). *J. Fish Biol.* 102, 443–454. doi:10.1111/jfb.15275.

Margarido, V.P., Galetti Junior, P.M. (2000). Amplification of a GC-rich heterochromatin in the freshwater fish *Leporinus desmotes* (Characiformes, Anostomidae). *Genetics and Molecular Biology* 23(3), 569-573.

Margarido, V.P. and Galetti Junior, P.M. (1999). Heterochromatin patterns and karyotype relationships within and between the genera *Brycon* and *Salminus* (Pisces, Characidae). *Genetics and Molecular Biology* 22(3), 357-361.

Marinho, M.M.F. (2014). Relações filogenéticas e revisão taxonômica das espécies do gênero *Copella* Myers, 1956 (Characiformes: Lebiasinidae). Universidade Estadual Paulista (UNESP).

Marinho, M.M.F. and Menezes, N.A. (2017). Taxonomic review of *Copella* (Characiformes: Lebiasinidae) with an identification key for the species. *PloS One* 12. doi:10.1371/journal.pone.0183069

Martins, C. (2007). Chromosomes and repetitive DNAs: a contribution to the knowledge of fish genome. *Fish cytogenetics* 421, 452.

Mayr, E. (1963). Animal Species and Evolution. Cambridge: Harvard University Press.

Mayr, E. (1999). Systematics and the Origin of Species, from the Viewpoint of a Zoologist. Cambridge: Harvard University Press.

McGaugh, S.E. and Noor, M.A.F. (2012). Genomic Impacts of Chromosomal Inversions in parapatric *Drosophila* species. *Phil. Trans. R. Soc. B* 367(1587), 422–429. doi:10.1098/rstb.2011.025010.1186/1471-2164-14-570

McKee, B.D., Wilhelm, K., Merrill, C., *et al*. (1998). Male Sterility and Meiotic Drive Associated with Sex Chromosome Rearrangements in *Drosophila*: Role of X-Y Pairing. *Genetics* 149(1), 143–155. doi:10.1093/genetics/149.1.143

McTaggart, S.J., Dudycha, J.L., Omilian, A., *et al*. (2007). Rates of recombination in the ribosomal DNA of apomictically propagated *Daphnia obtusa* lines. *Genetics* 175, 311-320.

Melo, B.F., Sidlauskas, B.L., Near, T.J., *et al*. (2021). Accelerated Diversification Explains the Exceptional Species Richness of Tropical Characoid Fishes. *Syst. Biol.* doi:10.1093/sysbio/syab040.

Melo, B.F., Sidlauskas, B.L., Near, T.J., *et al*. (2022). Accelerated diversification explains the exceptional species richness of tropical characoid fishes. *Systematic Biology*, 71(1), 78-

92. doi: 10.1093/sysbio/syab040
Melville, J., Haines, M.L., Boysen, K., *et al*. (2017). Identifying Hybridization and Admixture Using SNPs: Application of the DArTseq Platform in Phylogeographic Research on Vertebrates. *R. Soc. Open Sci*. 4, 161061. doi:10.1098/rsos.161061
Merlo A., Cross I., Palazón JL., *et al*. (2007). Chromosomal mapping of the major and minor ribosomal genes, (GATA) n and (TTAGGG) n by one-color and doublecolor FISH in the toadfish *Halobatrachus didactylus* (Teleostei: Batrachoididae). *Genetica* 131, 195–200.
Mirande, J.M. (2019). Morphology, molecules and the phylogeny of Characidae (Teleostei, Characiformes). *Cladistics* 35, 282–300. doi:10.1111/cla.12345
Mirarab, S., Bayzid, M.S., Warnow, T. (2016). Evaluating Summary Methods for Multilocus Species Tree Estimation in the Presence of Incomplete Lineage Sorting. *Systematic Biology* 65(3), 366–80. doi:10.1093/sysbio/syu063.
Montiel, E.E., Panzera, F., Palomeque, T., *et al*. (2021). Satellitome analysis of *Rhodnius prolixus*, one of the main Chagas disease vector species. *Int J Mol Sci.* 22(11), 6052.
Montiel, E.E., Mora, P., Rico-Porras, J.M. *et al*. (2022). Satellitome of the red palm weevil, *Rhynchophorus ferrugineus* (Coleoptera: Curculionidae), the most diverse among insects. *Front. Ecol. Evol*. 10, 826808. doi: 10.3389/fevo.2022.826808
Moraes, J.N., Viana, P.F., Favarato, R.M., *et al*. (2022). Karyotype variability in six Amazonian species of the family Curimatidae (Characiformes) revealed by repetitive sequence mapping. *Genetics and molecular biology*, 45, e20210125. doi: 10.1590/1982-0224-2021-0153
Moraes, R.L.R., Bertollo, L.A.C., Marinho, M.M.F., *et al*. (2017). Evolutionary relationships and cytotaxonomy considerations in the genus *Pyrrhulina* (Characiformes, Lebiasinidae). *Zebrafish* 14, 536-546. doi: 10.1089/zeb.2017.1465
Moraes, R.L.R., Sember, A., Bertollo, L.A.C., *et al*. (2019). Comparative cytogenetics and neo-Y formation in sall-sized fish species of the genus *Pyrrhulina* (Characiformes, Lebiasinidae). *Front Genet*. 10, 678. doi: 10.3389/fgene.2019.00678
Moraes, R.L.R., Sassi, F.M.C., Bertollo, L.A.C., *et al*. (2021). 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
Moraes, R.L.R., Sassi, F.M.C., Marinho, M.M.F., *et al.* (2023). Small Body, Large Chromosomes: Centric Fusions Shaped the Karyotype of the Amazonian Miniature Fish *Nannostomus anduzei* (Characiformes, Lebiasinidae). *Genes* 14, 192. doi:10.3390/genes14010192
Moreira-filho, O. and Bertollo, L.A.C. (1991). *Astyanax scabripinnis* (Pisces, Characidae): A species complex. *Revista Brasileira de Genética* 14, 331–357.
Muller, H.J. Isolation mechanisms, evolution and temperature. *Biol Symp.* **1942**, *6*, 71. Available at: https://www.google.de/url?sa=terct=jeq=eesrc=sesource=webecd=eved=2ahUKEwiI1OuM3MyAAxXHhv0HHeVXDCAQFnoECBAQAQeurl=https%3A%2F%2Fwww.ucl.ac.uk%2Ftaxome%2Flit%2Fmuller%25201942%2520new%2520enh.pdfeusg=AOvVaw0Xf86o-hX4H9vA0QhXx295eopi=89978449
Naish, M., Alonge, M., Wlodzimierz, P., *et al*. (2021). The genetic and epigenetic landscape of the *Arabidopsis* centromeres. *Science*, 374, eabi7489. doi: 10.1126/science.abi7489
Nascimento, M., Sousa, A., Ramirez, M., *et al*. (2017). PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. *Bioinformatics* 33, 128-129. doi: 10.1093/bioinformatics/btw582
Navarro, A. and Barton, N.H. (2003). Accumulating postzygotic isolation genes in parapatry: A new twist on chromosomal speciation. *Evolution*, 57, 447–459.

Near, T.J. and Thacker, C.E. (2024). Phylogenetic Classification of Living and Fossil Ray-Finned Fishes (Actinopterygii). *Bull. Peabody Mus. Nat. Hist*. 65, 3–302. doi: 10.3374/014.065.0101

Nelson, J. S., Grande, T.C., and Wilson, M.V.H. (2016). Fishes of the World. *John Wiley e Sons* doi:10.2307/2412830.

Netto-Ferreira, A.L. and Marinho, M.M.F. (2013). New species of *Pyrrhulina* (Ostariophysi: Characiformes: Lebiasinidae) from the Brazilian shield, with comments on a putative monophyletic group of species in the genus. *Zootaxa* 3664, 369-76. doi: 10.11646/zootaxa.3664.3.7

Netto-Ferreira, A.L. (2010). Revisão taxonômica e relações interespecíficas de Lebiasinidae (Ostariophysi: Characiformes: Lebiasinidae). Universidade de São Paulo (USP).

Novák, P., Neumann, P., Macas, J. (2020). Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. *Nat Protoc*. 15, 3745-3776. doi: 10.1038/s41596-020-0400-y

O'Neill, R.J., Eldridge, M.D.B., Metcalfe, C.J. (2004). Centromere dynamics and chromosome evolution in marsupials. *J. Hered*. 95, 375-381. doi: 10.1093/jhered/esh063

Ocalewicz, K. (2013). Telomeres in fishes. *Cytogenetic and genome research*, 141(2-3), 114-125. doi: 10.1159/000354278

Oliveira, E.A., Sember, A., Bertollo, L.A.C., *et al.* (2017). Tracking the evolutionary pathway of sex chromosomes among fishes: characterizing the unique XX/XY_1Y_2 system in *Hoplias malabaricus* (Teleostei, Characiformes). *Chromosoma*, 127, 115–128.

Oliveira, C., Andreata, A.A., Almeida-Toledo, L.F., *et al*. (1991). Karyotype and nucleolus organizer regions of *Pyrrhulina* cf. *australis* (Pisces, Characiformes, Lebiasinidae). *Rev. Bras. Genet.* 14, 685–690.

Oliveira C., Almeida-Toledo L.F., Foresti F. "Karyotypic evolution in Neotropical fishes". In: Pisano E, Ozouf-Costaz C, Foresti F, Kapoor B.G. (eds.) **Fish Cytogenetics.** USA: Enfield, NH, pp.111-164, 2007.

Ortí, G. and Meyer, A. (1997). The radiation of characiform fishes and the limits of resolution of mitochondrial ribosomal DNA sequences. *Systematic Biology* 46, 75-100.

Oliveira, R.R., Feldberg, E., dos Anjos, M.B., *et al*. (2009). Occurrence of multiple sexual chromosomes (XX/XY_1Y_2 and $Z_1Z_1Z_2Z_2/Z_1Z_2W_1W_2$) in catfishes of the genus *Ancistrus* (Siluriformes: Loricariidae) from the Amazon basin. *Genetica* 134, 243-249. doi:10.1007/s10709-007-9231-9

Oliveira, C., Avelino, G.S., Abe, K.T., *et al.* (2011). Phylogenetic relationships within the speciose family Characidae (Teleostei: Ostariophysi: Characiformes) based on multilocus analysis and extensive ingroup sampling. *BMC Evol. Biol*. 11, 1–25. doi: 10.1186/1471-2148-11-275

Oliveira, E.A., Perez, M.F., Bertollo, L.A.C., *et al*. (2020). Historical Demography and Climate Driven Distributional Changes in a Widespread Neotropical Freshwater Species with High Economic Importance. *Ecography* 43(9), 1291–1304. doi:10.1111/ecog.04874

Ortí, G. and Meyer, A. (1997). The radiation of characiform fishes and the limits of resolution of mitochondrial ribosomal DNA sequences. *Syst. Biol.* 46, 75–100. doi: 10.1093/sysbio/46.1.75

Ostberg, C. O., Hauser, L., Pritchard, V. L., *et al.* (2013). Chromosome Rearrangements, Recombination Suppression, and Limited Segregation Distortion in Hybrids between Yellowstone Cutthroat Trout (Oncorhynchus Clarkii Bouvieri) and Rainbow Trout (O. mykiss). *BMC genomics* 14(1), 570–616. doi:10.1186/1471-2164-14-570

Oyakawa, O.T. Relações filogenéticas das famílias Pyrrhulinidae, Lebiasinidae e Erythrinidae (Osteichthyes: Characiformes). Ph.D. Thesis, Universidade de São Paulo, 1998.

Paço, A., Chaves, R., Vieira-da-Silva, A., *et al*. (2013). The involvement of repetitive sequences in the remodelling of karyotypes: The *Phodopus* genomes *(Rodentia, Cricetidae)*. *Micron* 46, 27-34. doi: 10.1016/j.micron.2012.11.010

Palacios-Gimenez, O.M., Dias, G.B., De Lima, L.G., *et al*. (2017). High-throughput analysis of the satellitome revealed enormous diversity of satellite DNAs in the neo-Y chromosome of the cricket *Eneoptera surinamensis*. *Sci. Rep*. 7, 6422. doi: 10.1038/s41598-017-06822-8

Panthum, T., Singchat, W., Laopichienpong, N., *et al*. (2021). Genome-wide SNP Analysis of Male and Female rice Field Frogs, *Hoplobatrachus rugulosus*, Supports a Non-genetic Sex Determination System. *Diversity* 13, 501. doi: 10.3390/d13100501

Parey, E., Louis, A., Montfort, J., *et al*. (2023). Genome structures resolve the early diversification of teleost fishes. *Science* 379(6632), 572-575. doi: 10.1126/science.abq4257

Pavlek, M., Gelfand, Y., Plohl, M., *et al*. (2015). Genome-wide analysis of tandem repeats in *Tribolium castaneum* genome reveals abundant and highly dynamic tandem repeat families with satellite DNA features in euchromatic chromosomal arms. *Dna research* 22(6), 387-401. doi: 10.1093/dnares/dsv021

Pazza, R. and Kavalco, K.F. (2007). Chromosomal evolution in the neotropical characin *Astyanax* (Teleostei, Characidae). *Nucleus-Calcutta*, *50*(3), 523.

Pendás, A.M., Móran, P., Freije, J.P., *et al*. (1994). Chromosomal location and nucleotide sequence of two tandem repeats of the Atlantic salmon 5S rDNA. *Cytogenet Cell Genet*. 67. doi: 10.1159/000133792

Peona, V., Kutschera, V.E., Blom, M.P., *et al*. (2023). Satellite DNA evolution in Corvoidea inferred from short and long reads. *Mol Ecol*. 32, 1288-1305. doi: 10. 1111/mec.16484

Pereira, J.A., Milani, D., Ferretti, A.B.S., *et al*. (2021). The extensive amplification of heterochromatin in *Melipona* bees revealed by high throughput genomic and chromosomal analysis. *Chromosoma* 130, 251-262. doi: 10.1007/s00412-021-00764-x

Pinkel, D., Straume, T., Gray, J.W. (1986). Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc. Natl. Acad. Sci. USA* 83, 2934–2938. doi: 10.1073/pnas.83.9.2934

Plohl, M., Luchetti, A., Meštrović, N., *et al*. (2008). Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero) chromatin. *Gene* 409, 72-78. doi:10.1016/j.gene.2007.11.013

Plohl, M., Meštrović, N., Mravinac, B. (2012). Satellite DNA evolution. *Repetitive DNA* 7, 126-152. doi: 10.1159/000337122

Podgornaya, O.I. (2022). Nuclear organization by satellite DNA, SAF-A/hnRNPU and matrix attachment regions. *Sem. Cell. Dev. Biol. 128*, 61-68. doi: 10.1016/j.semcdb.2022.04.018

Poltronieri, J., Marquioni, V., Bertollo, L.A.C., *et al*. (2014). Comparative chromosomal mapping of microsatellites in *Leporinus* species (Characiformes, Anostomidae): unequal accumulation on the W chromosomes. *Cytogenet. Genome Res.* 142, 40–45. doi:10.1159/000355908.

Potter, S., Bragg, J.G., Turakulov, R., *et al*. (2022). Limited Introgression between Rock-Wallabies with Extensive Chromosomal Rearrangements. *Mol. Biol. Evol.* 39(1), 333. doi:10.1093/molbev/msab333

Prakhongcheep, O., Thapana, W., Suntronpong, A., Singchat, W., Pattanatanang, K., Phatcharakullawarawat, R., Muangmai, N., Peyachoknagul K., Matsubara, K., Ezaz, T., *et al*. Lack of satellite DNA species-specific homogenization and relationship to chromosomal rearrangements in monitor lizards (Varanidae, Squamata). *BMC Evol. Biol*. **2017**, *17*, 1-14. doi: 10.1186/s12862-017-1044-6

Presgraves, D. C. (2018). Evaluating Genomic Signatures of "the Large X-Effect" during

Complex Speciation. *Mol. Ecol*. 27(19), 3822–3830. doi:10.1111/mec.14777
Raj, A., Stephens, M., Pritchard, J. K. (2014). Faststructure: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics* 197(2), 573–589. doi:10.1534/genetics.114.164350
Rambaut, A., Drummond, A. J., Xie, D., *et al.* (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol*. 67(5), 901–904. doi:10.1093/sysbio/syy032
Ravi, V. and Venkatesh, B. (2018). The divergent genomes of teleosts. *Annual review of animal biosciences*, 6(1), 47-68. doi: 10.1146/annurev-animal-030117-014821
Reinhold, K. (1998). Sex Linkage Among Genes Controlling Sexually Selected Traits. *Behav. Ecol. Sociobiol*. 44(1), 1–7. doi:10.1007/s002650050508
Richard, G. F., Kerrest, A., Dujon, B. (2008). Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiology and molecular biology reviews*, 72(4), 686-727. doi:10.1128/mmbr.00011-08
Rieseberg, L. H. (2001). Chromosomal Rearrangements and Speciation. *Trends Ecol. Evol.* 16 (7), 351–358. doi:10.1016/s0169-5347(01)02187-5
Roberts, P.A. (1965). Difference in the Behaviour of Eu- and Heterochromatin: Crossing-over. *Nature* 205, 725–726. doi: 10.1038/205725b0
Rocha-Reis, D.A., de Oliveira Brandão, K., de Almeida-Toledo, L.F. (2018). The persevering cytotaxonomy: discovery of a unique XX/XY sex chromosome system in catfishes suggests the existence of a new, endemic and rare species. *Cytogenet. Genome Res.* 156, 45–55. doi: 10.1159/000492959
Roy, V., Monti-Dedieu, L., Chaminade, N., *et al.* (2005). Evolution of the chromosomal location of rDNA genes in two *Drosophila* species subgroups: *ananassae* and *melanogaster*. *Heredity* 94, 388–395.
Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., *et al*. (2019). DnaSP Version 6 for 32-bit and 64-bit Environments.
Ruiz-Ruano, F.J., López-León, M.D., Cabrero, J., *et al*. (2016). High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Sci Rep*. 6, 28333. doi: 10.1038/srep28333
Sambrook, J.and Russell, D.W. Molecular Cloning, A Laboratory Manual. New York: Cold Spring, Harbor Laboratory Press, 2001
Sumner, A.T. (1972). A simple technique for demonstrating centromeric heterochromatin. *Exp. Cell Res.* 75, 304–306. doi:10.1016/0014-4827(72)90558-7
Sánchez-Sevilla, J.F., Horvath, A., Botella, M.A., *et al.* (2015). Diversity Arrays Technology (DArT) Marker Platforms for Diversity Analysis and Linkage Mapping in a Complex Crop, the Octoploid Cultivated Strawberry (Fragaria × Ananassa). *PLoS One* 10(12), e0144960. doi:10.1371/journal.pone.0144960
Sassi, F.M.C., Hatanaka, T., Moraes, R.L.R., *et al*. (2020). An Insight into the Chromosomal Evolution of Lebiasinidae (Teleostei, Characiformes). *Genes* 11, 365. doi: 10.3390/genes11040365
Sassi, F.M.C., Toma, G.A., Cioffi, M.B. "FISH-in fish chromosomes". In Liehr, T (ed.) **Cytogenetics and Molecular Cytogenetics**. USA: CRC Press, Boca Raton, pp. 281–297, 2022. doi: 10.1201/9781003223658
Sassi, F.M.C., Oliveira, E., Bertollo, L.A.C., *et al*. (2019). Chromosomal evolution and evolutionary relationships of *Lebiasina* species (Characiformes, Lebiasinidae). *Int. J. Mol. Sci*. 20, 2944. doi: 10.3390/ijms20122944
Sassi, F.M.C., Perez, M.F., Oliveira, V.C.S., *et al*. (2021). High Genetic Diversity despite Conserved Karyotype Organization in the Giant Trahiras from Genus *Hoplias* (Characiformes, Erythrinidae). *Genes* 12, 252. doi: 10.3390/genes12020252

Šatović-Vukšić, E. and Plohl, M. (2023). Satellite DNAs—From localized to highly dispersed genome components. *Genes* 14, 742. doi: 10.3390/genes14030742

Scalvenzi, T. and Pollet, N. (2014). Insights on genome size evolution from a miniature inverted repeat transposon driving a satellite DNA. *Mol. Phylogenet. Evol.* 81, 1-9. doi:10.1016/j.ympev.2014.08.014

Scheel, J. (1973). Fish chromosomes and their evolution. *Intern. Rep. Dan. Akvarvum* 22.

Schemberger, M.O., Nascimento, V.D., Coan, R., *et al*. (2019). DNA transposon invasion and microsatellite accumulation guide W chromosome differentiation in a Neotropical fish genome *Chromosoma*, 128, 547–560.

Schimek, C., Shams, F., Miura, I., *et al.* (2022). Sex-linked markers in an Australian frog *Platyplectrum ornatum* (Limnodynastidae) with a small genome and homomorphic sex chromosomes. *Sci. Rep.* 12, 20934. doi: 10.1038/s41598-022-25105-5

Schmieder, R. and Eswards, R. (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PloS One* 6, e17288. doi:10.1371/journal.pone.0017288

Schubert, I. (2018). What is behind "centromere repositioning"? *Chromosoma* 127, 229–234. doi: 10.1007/s00412-018-0672-y

Sedlazeck, F.J., Lee, H., Darby, C.A., *et al.* (2018). Piercing the dark matter: bioinformatics of long-range sequencing and mapping. *Nat. Rev. Genet*. 19, 329-346. doi: 10.1038/s41576-018-0003-4

Sember, A., Bertollo, L.A.C., Ráb, P. *et al*. (2018). Sex Chromosome Evolution and Genomic Divergence in the Fish *Hoplias malabaricus* (Characiformes, Erythrinidae). *Frontiers in Genetics* 5(9).

Sember, A., de Oliveira, E.A., Ráb, P., *et al*. (2020). Centric fusions behind the karyotype evolution of neotropical *Nannostomus* pencilfishes (Characiforme, Lebiasinidae): First insights from a molecular cytogenetic perspective. *Genes* 11, 91. doi: 10.3390/genes11010091

Sember, A., Nguyen, P., Perez, M.F., *et al.* (2021). Multiple sex chromosomes in teleost fishes from a cytogenetic perspective: State of the art and future challenges. *Phil. Trans. R. Soc. B.* 376, 20200098. doi: 10.1098/rstb.2020.0098

Serrano-Freitas, E.A., Silva, A.M.Z.A., Ruiz-Ruano, F.J., *et al*. (2020). Satellite DNA content of B chromosomes in the characid fish *Characidium gomesi* supports their origin from sex chromosomes. *Mol. Genet. Genomics* 295, 195-207. doi: 10.1007/s00438-019-01615-2

Shao, F., Han, M., Peng, Z. (2019). Evolution and diversity of transposable elements in fish genomes. *Sci. Rep.* 9, 15399. doi: 10.1038/s41598-019-51888-1

Shatskikh, A.S, Kotov, A.A., Adashev, V.E., *et al*. (2020). Functional significance of satellite DNAs: insights from *Drosophila*. *Front. Cell, Dev. Biol*. 8, 312. doi:10.3389/fcell.2020.00312

Šíchová, J., Voleníková, A., Dincă, V., *et al.* (2015). Dynamic karyotype evolution and unique sex determination systems in *Leptidea* wood white butterflies. *BMC Evol Biol.* 15, 1-16. doi: 10.1186/s12862-015-0375-4

Silva, D.M.D.A., Utsunomia, R., Ruiz-Ruano, F.J., *et al.* (2017). High-throughput analysis unveils a highly shared satellite DNA library among three species of fish genus *Astyanax. Sci Rep*. 7, 12726. doi: 10.1038/s41598-017-12939-7

Simmons, M.P. and Gatesy, J. (2021). Collapsing Dubiously Resolved Gene-Tree Branches in Phylogenomic Coalescent Analyses. *Molecular Phylogenetics and Evolution* 107092. doi: 10.1016/j.ympev.2021.107092.

Slamovits, C.H., Cook, J.A., Lessa, E.P., *et al*. (2001). Recurrent amplifications and deletions of satellite DNA accompanied chromosomal diversification in South American *tuco-tucos*

(genus *Ctenomys*, Rodentia: Octodontidae): a phylogenetic approach. *Mol. Biol. Evol. 18*, 1708-1709. doi: 10.1093/oxfordjournals.molbev.a003959

Smith, C.J., Castanon, O., Said, K., *et al*. (2020). Enabling large-scale genome editing at repetitive elements by reducing DNA nicking. *Nucleic Acids Res.* 48, 5183-5195. doi: 10.1093/nar/gkaa239

Sochorová, J., Garcia, S., Gálvez, F., *et al.* (2018). Evolutionary trends in animal ribosomal DNA loci: introduction to a new online database. *Chromosoma* 127, 141–150. doi: 10.1007/s00412-017-0651-8

Sola, L., Rossi, A.R., Annesi, F., *et al*. (2003). Cytogenetic studies in *Sparus auratus* (Pisces, Perciformes): molecular organization of 5S rDNA and chromosomal mapping of 5S and 45S ribosomal genes and of telomeric repeats. *Hereditas* 139, 232-236.

Soto, M.Á., Castro, J.P., Walker, L.I., *et al*. (2018). Evolution of trans-Andean endemic fishes of the genus *Cheirodon* (Teleostei: Characidae) are associated with chromosomal rearrangements. *Rev. Chil. Hist. Nat*. 91, 8. doi: 10.1186/s40693-018-0078-5

Souza e Sousa, J.F., Viana, P.F., Bertollo, L.A.C., *et al.* (2017). Evolutionary Relationships among *Boulengerella* Species (Ctenoluciidae, Characiformes): Genomic Organization of Repetitive DNAs and Highly Conserved Karyotypes. *Cytogenet. Genome Res.* 152. doi: 10.1159/000480141

Souza, J.S. Instituto Nacional de Pesquisas da Amazônia – INPA, Manaus, Brazil. Personal communication, 2019.

Souza, C.S., Melo, B.F., Mattox, G.M., *et al*. (2022). Phylogenomic analysis of the Neotropical fish subfamily Characinae using ultraconserved elements (Teleostei: Characidae). *Molecular Phylogenetics and Evolution*, *171*, 107462. doi: 10.1016/j.ympev.2022.107462

Souza, T.B.D., Ferreira, D.C., Silva, H.P.D., *et al*. (2023). DNA Barcoding of *Pyrrhulina australis* (Characiformes: Lebiasinidae) reveals unexpected cryptic diversity in the group. *Neotropical Ichthyology*, 21(4), e230037. doi: 10.1590/1982-0224-2023-0037

Souza, F.H.S., Sassi, F.M.C., Ferreira, P.H.N., *et al.* (2022). Integrating Cytogenetics and Population Genomics: Allopatry and Neo-Sex Chromosomes May Have Shaped the Genetic Divergence in the *Erythrinus erythrinus* Species Complex (Teleostei, Characiformes). *Biology 11*, 315. doi: 10.3390/biology11020315

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Steane, D.A., Nicolle, D., Sansaloni, C.P., *et al*. (2011). Population Genetic Analysis and Phylogeny Reconstruction in Eucalyptus (Myrtaceae) Using High-Throughput, Genome-wide Genotyping. *Mol. phylogenetics Evol*. 59(1), 206–224. doi: 10.1016/j.ympev.2011.02.003

Steindachner, F. Beiträge zuer Kenntniss der Characinen des Amazonenstromes. Wien: Sitzungsber. Kaiserl. Akad. Wiss., 72, 6, 1876

Symonová, R., Majtánová, Z., Sember, A., *et al*. (2013). Genome differentiation in a species pair of coregonine fishes: an extremely rapid speciation driven by stress-activated retrotransposons mediating extensive ribosomal DNA multiplications. *BMC Evol. Biol.* 13, 1–11. doi:10.1186/1471-2148-13-42.

Tagliacollo, V.A., Dagosta F.C.P., de Pinna M., *et al*. (2021). Assessing extinction risk from geographic distribution data in Neotropical freshwater fishes. *Neotrop Ichthyol*. 19(3), e210079. doi:10.1590/1982-0224-2021-0079

Tautz, D. "Notes on the defunction and nomenclature of tandemly repetitive DNA sequences". In: Pena, S.D.J., Chakraborty, R., Epplen, J.T., Jeffreys, A.J. (eds.) **DNA Fingerprinting: State of the Science**. Switzerland: Birkhauser, pp 21-28, 1993.

Terencio, M.L., Schneider, C.H., Gross, M.C., *et al*. (2012). Repetitive sequences associated

with differentiation of W chromosome in *Semaprochilodus taeniurus*. *Genetica* 140, 505–512. doi:10.1007/s10709-013-9699-4.
Thakur, J., Packiaraj, J., Henikoff, S. (2021). Sequence, chromatin and evolution of satellite DNA. *Int. J. Mol. Sci*. 22, 4309. doi: 10.3390/ijms22094309
Tickner, D., Opperman, J.J., Abell, R., *et al*. (2020). Bending the curve of global freshwater biodiversity loss: an emergency recovery plan. *BioScience* 70(4), 330-342. doi:10.1093/biosci/biaa002
Toews, D. P. and Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular ecology* 21(16), 3907-3930. doi: 10.1111/j.1365-294X.2012.05664.x
Tollesfrud, M.M., Sønstebø, J.H., Brochmann, C., *et al*. (2009). Combined Analysis of Nuclear and Mitochondrial Markers Provide New Insight into the Genetic Structure of North European *Picea Abies*. *Heredity* 102 (6), 549–62. doi: 10.1038/hdy.2009.16.
Toma, G.A., Moraes, R.L.R., Sassi, F.M.C., *et al*. (2019). Cytogenetics of the small-sized fish, *Copeina Guttata (Characiformes, Lebiasinidae)*: Novel insights into the karyotype differentiation of the family. *Plos One* 14, e0226746. doi: 10.1371/journal.pone.0226746
Toma, G.A., dos Santos, N., dos Santos, R., *et al*. (2023). Cytogenetics Meets Genomics: Cytotaxonomy and Genomic Relationships among Color Variants of the Asian Arowana *Scleropages formosus*. *Int. J. Mol. Sci.* 24, 9005. doi: 10.3390/ijms24109005
Tonella, L.H., Ruaro, R., Daga, V.S., *et al.* (2023). Neotropical freshwater fishes: A dataset of occurrence and abundance of freshwater fishes in the Neotropics. doi: 10.1002/ecy.3713
Tunjić-Cvitanić, M., Pasantes, J.J., García-Souto, D., *et al*. (2021). Satellitome analysis of the pacific oyster Crassostrea gigas reveals new pattern of satellite DNA organization, highly scattered across the genome. *Int. J. Mol. Sci.* 22, 6798. doi: 10.3390/ijms22136798
Ugarković, Đ. and Plohl, M. (2002). Variation in satellite DNA profiles—causes and effects. *The EMBO Jornal* 18, 1708-1719. doi: 10.1093/emboj/cdf612
Utsunomia, R., Silva, D.M.Z.D.A., Ruiz-Ruano, F.J., *et al.* (2019). Satellitome landscape analysis of *Megaleporinus macrocephalus (Teleostei, Anostomidae)* reveals intense accumulation of satellite sequences on the heteromorphic sex chromosome. *Sci Rep*. 9, 5856. doi: 10.1038/s41598-019-42383-8
Valeri, M.P., Dias, G.B., do Espírito Santo, A.A., *et al.* (2021). First description of a satellite DNA in manatees' centromeric regions. *Front Genet.* 12, 694866. doi: 10.3389/fgene.2021.694866
van der Sleen, P. and Albert, J.S. (eds.) *Field guide to the fishes of the Amazon, Orinoco, and Guianas*. USA: Princeton University Press, 15, 2017.
Van Der Sleen P. and Albert J.S. "Patterns of freshwater wetland fish species diversity". In: Tockner K, Mehner T, Wittmann F, Junk WJ (eds.) **Encyclopedia of Inland Waters**. Elsevier, pp 243–255, 2021.
Vara, C., Paytuví-Gallart, A., Cuartero, Y., *et al*. (2021). The impact of chromosomal fusions on 3D genome folding and recombination in the germ line. *Nat. Commun*. 12, 2981. doi: 10.1038/s41467-021-23270-1
Venere P.C. and Garutti V. Peixes do Cerrado: Parque Estadual da Serra Azul, rio Araguaia, MT. RiMa (ed.), São Carlos: Fapemat, 2011.
Vieira, C.P., Coelho, P.A. and Vieira, J. (2003). 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* 164, 1459-1469. doi: 10.1093/genetics/164.4.1459
Vieira-da-Silva, A., Louzada, S., Adega, F., *et al*. (2015). A high-resolution comparative chromosome map of *Cricetus cricetus* and *Peromyscus eremicus* reveals the involvement of

constitutive heterochromatin in breakpoint regions. *Cytogenetic Genome Res*. 145, 59-67. doi: 10.1159/000381840
Vozdova, M., Kubickova, S., Cernohorska, H., *et al.* (2019). Comparative study of the bush dog (*Speothos venaticus*) karyotype and analysis of satellite DNA sequences and their chromosome distribution in six species of Canidae. *Cytogenet. Genome Res*. 159, 88-96. doi: 10.1159/000503082
Vozdova, M., Kubickova, S., Martínková, N. (2021). Satellite DNA in neotropical deer species. *Genes* 12, 123. doi: 10.3390/genes12010123
Waldbieser, G.C., Liu, S., Yuan, Z., *et al.* (2023). Reference genomes of channel catfish and blue catfish reveal multiple pericentric chromosome inversions. *BMC Biol*. 21, 67. doi: 10.1186/s12915-023-01556-8
Wang, Y., Ghosh, G., Hendrickson, E.A. (2009). Ku86 represses lethal telomere deletion events in human somatic cells. *Proceedings of the National Academy of Sciences* 106(30), 12430-12435. doi: 10.1073/pnas.0903362106
Wang, M. and Lemos, B. (2017). Ribosomal DNA copy number amplification and loss in human cancers is linked to tumor genetic context, nucleolus activity, and proliferation. *PLoS genetics*, 13(9), e1006994. doi: 10.1371/journal.pgen.1006994
Weissensteiner, M.H. and Suh, A. "Repetitive DNA: The dark matter of avian genomics". In Kraus, R (ed.) **Avian genomics in ecology and evolution: From the lab into the wild.** Switzerland: Springer, pp 93-150, 2019. doi: 10.1007/978-3-030-16477-5_5
Weitzman, M. and Weitzman, S. Family Lebiasinidae. Check List Freshwater Fishes South Cent. Am. Porto Alegre: Edipucrs pp 241–250, 2003
Weitzman, S.H. and Cobb, J.S. (1975). A revision of the South American fishes of the genus *Nannostomus* Günther (Family Lebiasinidae).
Weitzman, S.H. and Vari, R.P. (1988). Miniaturization in South American freshwater fishes, an overview and discussion.
Xu, W., Tai, J., He, K. (2024). Complete Mitochondrial Genomes of *Nannostomus* Pencilfish: Genome Characterization and Phylogenetic Analysis. *Animals* 14, 1598. doi: 10.3390/ani14111598
Yang, F., Graphodatsky, A.S. "Animal Probes and ZOO-FISH". In: Liehr, T. (ed.) **Fluorescence In Situ Hybridization (FISH) - Application Guide**. Berlin: Springer Protocols Handbooks, p. 323-346, 2009.
Yang, F., Trifonove, V., Kosyakova, N., *et al*. "Generation of paint probes by flow-sorted and microdissected chromosomes". In: Liehr, T. (ed). **Fluorescence In Situ Hybridization (FISH) - Application Guide**. Berlin: Springer Protocols Handbooks, p. 35-52, 2009.
Yang, T.J., Yu, Y., Chang, S.B., *et al*. (2005). Toward closing rice telomere gaps: mapping and sequence characterization of rice subtelomere regions. *Theor. Appl. Genet*. 111, 467-478. doi:10.1007/s00122-005-2034-4
Yang, W., Kang, X., Yang, Q., *et al*. (2013). Review on the Development of Genotyping Methods for Assessing Farm Animal Diversity. *J. Anim. Sci. Biotechnol*. 4 (1), 1–6. doi:10.1186/2049-1891-4-2
Yannic, G., Basset, P., and Hausser, J. (2009). Chromosomal Rearrangements and Gene Flow over Time in an Inter-specific Hybrid Zone of the Sorex araneus Group. *Heredity* 102 (6), 616–625. doi:10.1038/hdy.2009.19
Yano, C.F., Bertollo, L.A.C., and de Cioffi, M.B.(2017). "Fish-FISH: molecular cytogenetics in fish species," in *Fluorescence in situ hybridization (FISH)* (Springer), 429–443. doi:10.1007/978-3-662-52959-1_44.
Yano, C.F., Bertollo, L.A.C., Molina, W.F., *et al*.(2014). Genomic organization of repetitive DNAs and its implications for male karyotype and the neo-Y chromosome differentiation in *Erythrinus erythrinus* (Characiformes, Erythrinidae). *Comp. Cytogenet.* 8, 139.

doi:10.3897/compcytogen.v8i2.7597.
Yano, C.F., Sember, A., Kretschmer, R. *et al.* (2021). Against the mainstream: exceptional evolutionary stability of ZW sex chromosomes across the fish families Triportheidae and Gasteropelecidae (Teleostei: Characiformes). *Chromosome Res* 29, 391–416. doi: 10.1007/s10577-021-09674-1
Yazdi, H.P., Olito, C., Kawakami, T., *et al*. (2023). The evolutionary maintenance of ancient recombining sex chromosomes in the ostrich. *PLoS Genetics* 19, e1010801. doi: 10.1371/journal.pgen.1010801
Zattera, M.L. and Bruschi, D.P. (2022). Transposable elements as a source of novel repetitive DNA in the eukaryote genome. *Cells* 11, 3373. doi:10.3390/cells11213373
Zwick, M.S., Hanson, R.E., Islam-Faridi, M.N., *et al*.(1997). A rapid procedure for the isolation of C 0 t-1 DNA from plants. *Genome* 40, 138–142. doi:10.1139/g97-020.