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

Pedro H. N. Ferreira[1], Fernando H. S. Souza[1], Renata L. de Moraes[1], Manolo F. Perez[1], Francisco de M. C. Sassi[1], Patrik F. Viana[2], Eliana Feldberg[2], Tariq Ezaz[3], Thomas Liehr[4]*, Luiz A. C. Bertollo[1] and Marcelo de B. Cioffi[1]

[1]Laboratório de Citogenética de Peixes, Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, Brazil, [2]Laboratório de Genética Animal, Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil, [3]Institute for Applied Ecology, University of Canberra, Canberra, NSW, Australia, [4]Institute of Human Genetics, Friedrich Schiller University, University Hospital Jena, Jena, Germany

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.

Keywords: fishes, neo-sex chromosomes, chromosomal rearrangements, cytogenetic, genetic diversity

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, 2019, 2021; Sassi et al., 2019; Toma et al., 2019; Sember et al., 2020). 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, Meyers and Myers, 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 ♀$X_1X_1X_2X_2$/♂X_1X_2Y 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.

MATERIALS AND METHODS

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 e Depinna 2021).

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=PRJNA804560&o=acc_s %3Aa). Additionally, datasets one and two are available at https://github.com/PHnarciso/Dataset-1-and-2---Manuscript-*Pyrrhulina*.

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.

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

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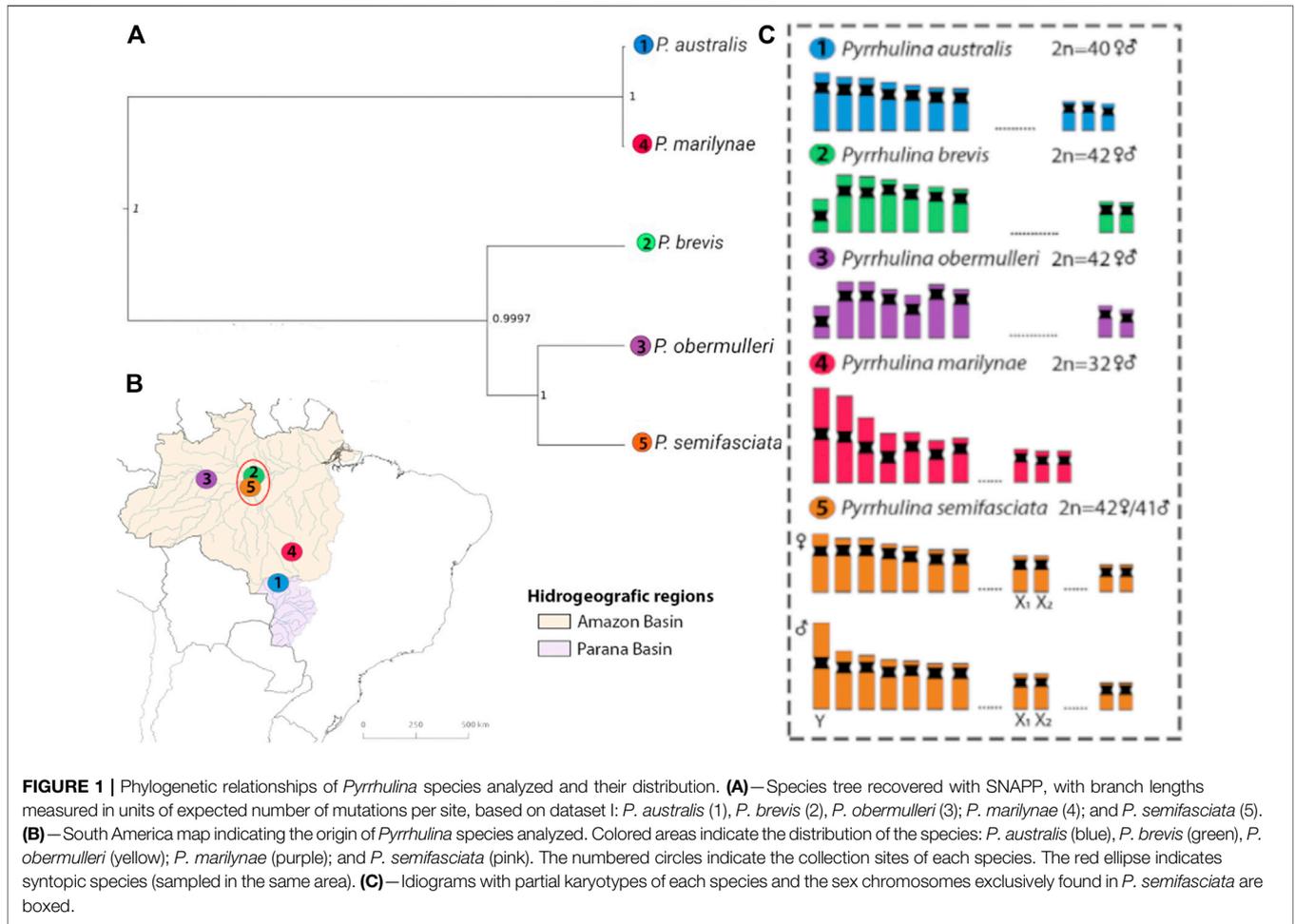
TABLE 1 | Species, diploid numbers (2n), sex chromosome systems, numbers of individuals cytogenetically analyzed, and the number of individuals sequenced.

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.

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*).

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.

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.

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 P_3 and *P. semifasciata* and *P. obermulleri* as either P_1 or P_2. 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.

RESULTS

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.

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.

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**).

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**).

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**).

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

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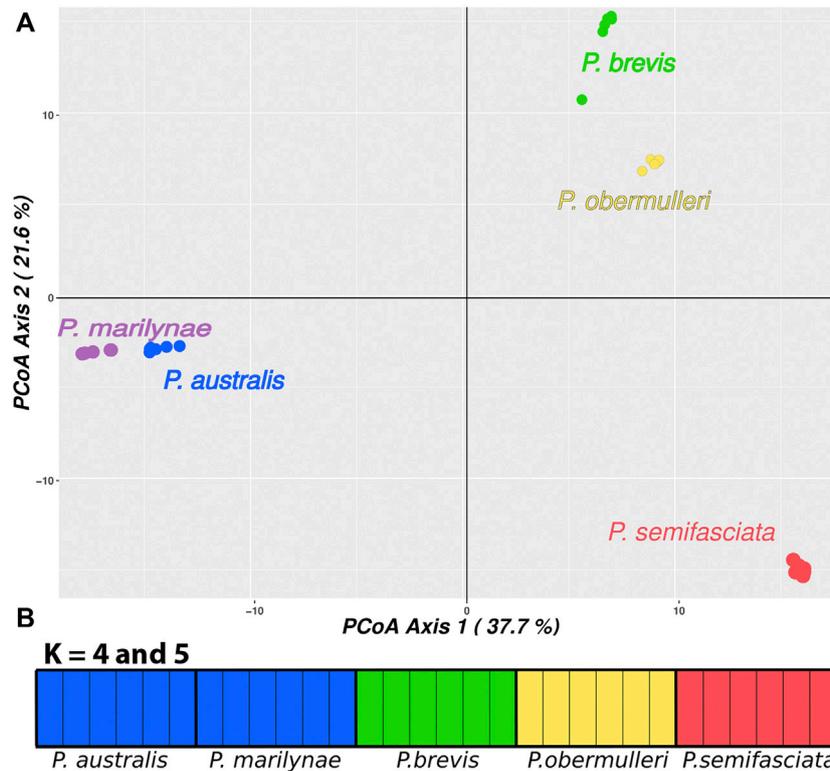
FIGURE 3 | (A) Genetic diversity in *Pyrrhulina* species according to a PCoA. The PCoA recovered 37.7% of the total variation in the first principal coordinate (PC1), and 21.6% in the second principal coordinate (PC2). **(B)**—Results for fastStructure with $K = 4$ and $K = 5$. Each vertical bar represents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

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

chromosome systems might have contributed to increase differentiation.

　　Overall, the genetic diversity indexes (π; θ_W) (**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

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.

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 & 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$_1$X$_2$Y-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.

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 X_1X_2Y 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.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: SRA, PRJNA804560.

ETHICS STATEMENT

The animal study was reviewed and approved by the Ethics Committee on Animal Experimentation of the Universidade Federal de São Carlos (Process number CEUA 1853260315).

AUTHOR CONTRIBUTIONS

PF, FS, RR, and MP carried out the analysis and drafted the manuscript. TE, FS, and PV helped in the analysis, drafted and revised the manuscript. TL, LB, EF, and MC drafted and revised the manuscript. All authors read and approved the final version of the manuscript.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

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

Alkimim, E. R., Caixeta, E. T., Sousa, T. V., da Silva, F. L., Sakiyama, N. S., and Zambolim, L. (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

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. (1909). *Heredity and Variation in Modern Lights. Darwin and Modern Science*. Editor A. C. Seward (Cambridge: Cambridge University Press), 85–101.

Beaudry, F. E. G., Barrett, S. C. H., and 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

Bertaco, V. A., Ferrer, J., Carvalho, F. R., and Malabarba, L. R. (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

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., et al. (2014). BEAST 2: a Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput. Biol.* 10 (4), e1003537. doi:10.1371/journal.pcbi.1003537

Coyne, J. A., and Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.

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

de M. C. Sassi, F., Perez, M. F., Oliveira, V. C. S., Deon, G. A., de Souza, F. H. S., and Ferreira, P. H. N. (2021). High Genetic Diversity Despite Conserved Karyotype Organization in the Giant Trahiras from Genus *Hoplias* (Characiformes, Erythrinidae). *Genes* 12 (2), 252. doi:10.3390/genes12020252

de Moraes, R. L. R., Bertollo, L. A. C., Marinho, M. M. F., Yano, C. F., Hatanaka, T., and Barby, F. F. (2017). Evolutionary Relationships and Cytotaxonomy Considerations in the Genus *Pyrrhulina* (Characiformes, Lebiasinidae). *Zebrafish* 14 (6), 536–546. doi:10.1089/zeb.2017.1465

de Moraes, R. L. R., Sassi, F. D. M. C., Bertollo, L. A. C., Marinho, M. M. F., Viana, P. F., and Feldberg, E. (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

de Moraes, R. L. R., Sember, A., Bertollo, L. A. C., De Oliveira, E. A., Ráb, P., and Hatanaka, T. (2019). Comparative Cytogenetics and Neo-Y Formation in Small-Sized Fish Species of the Genus *Pyrrhulina* (Characiformes, Lebiasinidae). *Front. Genet.* 10, 678. doi:10.3389/fgene.2019.00678

Dobzhansky, T. (1933). Geographical Variation in Lady-Beetles. *Am. Nat.* 67 (709), 97

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.

Faria, R., and Navarro, A. (2010). Chromosomal Speciation Revisited: Rearranging Theory with Pieces of Evidence. *Trends Ecol. Evol.* 25 (11), 660–669. doi:10.1016/j.tree.2010.07.008

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

Fricke, R., Eschmeyer, W. N., and van der Laan, R. (2022). Eschmeyer's Catalog of Fishes: Genera, Species, References. Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (Accessed July 13, 2021).

Froese, R. (2009). FishBase. World Wide Web Electronic Publication. Available at: http://www.fishbase.org (Accessed April 4, 2021).

Gruber, B., Unmack, P. J., Berry, O. F., and Georges, A. (2018). Dartr : An R Package to Facilitate Analysis of SNP Data Generated from Reduced Representation Genome Sequencing. *Mol. Ecol. Resour.* 18 (3), 691–699. doi:10.1111/1755-0998.12745

Haldane, J. B. S. (1922). Sex Ratio and Unisexual Sterility in Hybrid Animals. *J. Gen.* 12, 101–109. doi:10.1007/bf02983075

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., et al. (2012). "Diversity Arrays Technology: a Generic Genome Profiling Technology on Open Platforms," in *Data Production and Analysis in Population Genomics* (Totowa, NJ: Humana Press), 67–89. doi:10.1007/978-1-61779-870-2_5

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., Kume, M., Jones, F. C., Chan, Y. F., et al. (2009). A Role for a Neo-Sex Chromosome in Stickleback Speciation. *Nature* 461 (7267), 1079–1083. doi:10.1038/nature08441

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., and Mayrose, I. (2015). Clumpak : a Program for Identifying Clustering Modes and Packaging Population Structure Inferences across K. *Mol. Ecol. Resour.* 15 (5), 1179–1191. doi:10.1111/1755-0998.12387

Lindholm, A., and Breden, F. (2002). Sex Chromosomes and Sexual Selection in Poeciliid Fishes. *Am. Nat.* 160 (S6), S214–S224. doi:10.1086/342898

Machado, C. A., Haselkorn, T. S., and 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

Marinho, M. M. F. (2014). *Relações filogenéticas e revisão taxonômica das espécies do gênero* Copella *Myers. Characiformes: Lebiasinidae)*. [dissertattion thesis]. São Paulo, Brazil: Unesp.

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

Mayr, E. (1999). *Systematics and the Origin of Species, from the Viewpoint of a Zoologist*. Cambridge, MA: 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., and Ren, X.-j. (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

Melville, J., Haines, M. L., Boysen, K., Hodkinson, L., Kilian, A., Smith Date, K. L., 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

Muller, H. (1942). Isolating Mechanisms, Evolution, and Temperature. *Biol. Symp.* 6, 71

Netto-ferreira, A. L., Marinho, M. M. F., Netto-Ferreira, A. L., and Marinho, M. M. (2003). New Species of Pyrrhulina (Ostariophysi: Characiformes: Lebiasinidae) from the Brazilian Shield, with Comments on a Putative Monophyletic Group of Species in the genus. *Science Zootaxa* 3003664 (56173), 321369. doi:10.11646/zootaxa.3664.3.7

Oliveira, E. A., Perez, M. F., Bertollo, L. A. C., Gestich, C. C., Ráb, P., Ezaz, T., 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

Ostberg, C. O., Hauser, L., Pritchard, V. L., Garza, J. C., and Naish, K. A. (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

Potter, S., Bragg, J. G., Turakulov, R., Eldridge, M. D. B., Deakin, J., Kirkpatrick, M., et al. (2022). Limited Introgression between Rock-Wallabies with Extensive

Chromosomal Rearrangements. *Mol. Biol. Evol.* 39 (1), 333. doi:10.1093/molbev/msab333

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., and 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., Baele, G., and Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* 67 (5), 901–904. doi:10.1093/sysbio/syy032

Reinhold, K. (1998). Sex Linkage Among Genes Controlling Sexually Selected Traits. *Behav. Ecol. Sociobiol.* 44 (1), 1–7. doi:10.1007/s002650050508

Rieseberg, L. H. (2001). Chromosomal Rearrangements and Speciation. *Trends Ecol. Evol.* 16 (7), 351–358. doi:10.1016/s0169-5347(01)02187-5

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S., et al. (2019). *DnaSP Version 6 for 32-bit and 64-bit Environments.*

Sánchez-Sevilla, J. F., Horvath, A., Botella, M. A., Gaston, A., Folta, K., Kilian, 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., Oliveira, E. A., Bertollo, L. A. C., Nirchio, M., Hatanaka, T., Marinho, M. M. F., et al. (2019). Chromosomal Evolution and Evolutionary Relationships of *Lebiasina* Species (Characiformes, Lebiasinidae). *Int. J. Mol. Sci.* 20 (12), 2944. doi:10.3390/ijms20122944

Sember, A., de Oliveira, E. A., Ráb, P., Bertollo, L. A. C., de Freitas, N. L., Viana, P. F., et al. (2020). Centric Fusions behind the Karyotype Evolution of Neotropical Nannostomus Pencilfishes (Characiforme, Lebiasinidae): First Insights from a Molecular Cytogenetic Perspective. *Genes* 11 (1), 91. doi:10.3390/genes11010091

Šíchová, J., Voleníková, A., Dincă, V., Nguyen, P., Vila, R., Sahara, K., et al. (2015). Dynamic Karyotype Evolution and Unique Sex Determination Systems in *Leptidea* Wood White Butterflies. *BMC Evol. Biol.* 15 (1), 1–16. doi:10.1186/s12862-015-0375-4

Souza, F. H. S., Perez, M. F., Bertollo, L. A. C., Oliveira, E. A. D., Lavoué, S., Gestich, C. C., et al. (2019). Interspecific Genetic Differences and Historical Demography in South American Arowanas (Osteoglossiformes, Osteoglossidae, *Osteoglossum*). *Genes* 10 (9), 693. doi:10.3390/genes10090693

Souza, F. H. S., Perez, M. F., Charlesworth, D., Ferreira, P. H. N., Bertollo, L. A. C., Ezaz, T., et al. (2021). Elucidating the Complex Relationship of *Hoplias malabaricus'* (Teleostei; Characiformes) Population. A Genetic, Chromosomal and Spatial Approach. *Evol. Lett.*

Steane, D. A., Nicolle, D., Sansaloni, C. P., Petroli, C. D., Carling, J., Kilian, A., 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. (1876). Beiträge zuer Kenntniss der Characinen des Amazonenstromes. *Sitzungsber. Kaiserl. Akad. Wiss., Wien* 72, 6

Tagliacollo, V. A., Roxo, F. F., Duke-Sylvester, S. M., Oliveira, C., and Albert, J. S. (2015). Biogeographical Signature of River Capture Events in Amazonian Lowlands. *J. Biogeogr.* 42 (12), 2349–2362. doi:10.1111/jbi.12594

Toma, G. A., de Moraes, R. L. R., Sassi, F. D. M. C., Bertolo, L. A. C., de Oliveira, E. A., Rab, P., 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 (12), e0226746. doi:10.1371/journal.pone.0226746

Weitzman, M., and Weitzman, S. H. (2003). Family Lebiasinidae. Check List of the Freshwater Fishes of South and Central America, Porto Alegre. *Edipucrs* 729p, 241

Yang, W., Kang, X., Yang, Q., Lin, Y., and Fang, M. (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