# Phylogeography and population genetic structure of the cardinal tetra (*Paracheirodon axelrodi*) in the Orinoco basin and Negro River (Amazon basin): evaluating connectivity and historical patterns of diversification.

**Diana Sanchez-Bernal** [Equal first author, 1], **José Gregorio Martinez** [Equal first author, 2], **Izeni Pires Farias** [3], **Tomas Hrbek** [3], **Susana Caballero** [Corresp. 1]

[1] Biological Sciences Department, Universidad de Los Andes, Bogota, Colombia

[2] Grupo de Investigación Biociencias, Facultad de Ciencias de la Salud, Institución Universitaria Colegio Mayor de Antioquia, Medellín, Antioquia, Colombia

[3] Departamento de Genética, Universidade do Amazonas, Manaos, Brazil

Corresponding Author: Susana Caballero
Email address: sj.caballero26@uniandes.edu.co

The Neotropics contain one of the most diverse assemblages of freshwater fishes worldwide. Part of this diversity is shared between the Orinoco and Amazon basins. These basins have been separated for a long time due to the Vaupes Arch, rising between 10 - 11 Ma. Today, there is only one permanent connection between the Orinoco and Negro (Amazon) basins, known as the Casiquiare Canal. However, alternative corridors allowing fish dispersion between both basins have been proposed. The cardinal tetra (*Paracheirodon axelrodi),* the most important fish in the ornamental world market, is distributed in both basins. Here we investigated *P. axelrodi* phylogeography, population structure, and potential routes of migration and connectivity between the two basins. A total of 468 bp of the mitochondrial gene (COI), 555 bp of the nuclear gene fragment (MYH6), and 8 microsatellite loci were analyzed. As a result, we found two major genetic clusters as the most likely scenario (K=2), but they were not discreetly distributed between basins. A gradient of genetic admixture was observed in Cucui and São Gabriel da Cachoeira, between the upper Negro River and the upper Orinoco. Samples from the middle-lower Negro River were highly structured. Cucui (Negro basin) was more similar to the Orinoco than to the rest of the Negro basin populations. However, substructure was also observed by the discriminant analysis, fixation indices and other hierarchichal structure analyses (K=3-6), showing three major geographic clusters: Orinoco, Cucui, and the remaining of the Negro basin. Unidirectional migration patterns were detected between basins: via Cucui toward Orinoco and via the remaining of the Negro basin toward Orinoco. Results from the Relaxed Random Walk analysis support a very recent origin of this species in the headwater Orinoco basin (Western Guiana Shield, at late Pleistocene)

with a later rapid colonization of the remaining Orinoco basin and almost simultaneously the Negro River via Cucui, between 0.115 until about 0.001 Ma. Historical biogeography and population genetic patterns observed here for Cardinal tetra, seem to be better explained by river capture, physical, or ecological barriers than due to the geographic distance.

# **Phylogeography and population genetic structure of the cardinal tetra (*Paracheirodon axelrodi*) in the Orinoco basin and Negro River (Amazon basin): evaluating connectivity and historical patterns of diversification.**

Diana Sanchez-Bernal[1]*, José Gregorio Martínez[2]*, Izeni Pires Farias[3], Tomas Hrbek[3], Susana Caballero[1]

[1]Laboratorio de Ecología Molecular de Vertebrados Acuáticos LEMVA, Universidad de los Andes, Bogotá, Colombia.

[2]Grupo de Investigación Biociencias, Facultad de Ciencias de la Salud, Institución Universitaria Colegio Mayor de Antioquia, Medellín, Colombia.

[3]Laboratório de Evolução e Genética Animal LEGAL, Departamento de Genética, Universidade Federal do Amazonas UFAM, Manaus, Brazil.

**\***Share equal authorship

Corresponding author:

Susana Caballero-Gaitán.

Carrera 1 #18a-12, Bogotá, 111711, Colombia.

Email address: sj.caballero26@uniandes.edu.co

## **Abstract**

The Neotropics contain one of the most diverse assemblages of freshwater fishes worldwide. Part of this diversity is shared between the Orinoco and Amazon basins. These basins have been separated for a long time due to the Vaupes Arch, rising between 10 - 11 Ma. Today, there is only one permanent connection between the Orinoco and Negro (Amazon) basins, known as the Casiquiare Canal. However, alternative corridors allowing fish dispersion between both basins have been proposed. The cardinal tetra (*Paracheirodon axelrodi),* the most important fish in the ornamental world market, is distributed in both basins. Here we investigated *P. axelrodi* phylogeography, population structure, and potential routes of migration and connectivity between the two basins. A total of 468 bp of the mitochondrial gene (COI), 555 bp of the nuclear gene fragment (MYH6), and 8 microsatellite loci were analyzed. As a result, we found two major genetic clusters as the most likely scenario (K=2), but they were not discreetly distributed

33  between basins. A gradient of genetic admixture was observed in Cucui and São Gabriel da
34  Cachoeira, between the upper Negro River and the upper Orinoco. Samples from the middle-
35  lower Negro River were highly structured. Cucui (Negro basin) was more similar to the Orinoco
36  than to the rest of the Negro basin populations. However, substructure was also observed by the
37  discriminant analysis, fixation indices and other hierarchichal structure analyses (K=3-6),
38  showing three major geographic clusters: Orinoco, Cucui, and the remaining of the Negro basin.
39  Unidirectional migration patterns were detected between basins: via Cucui toward Orinoco and
40  via the remaining of the Negro basin toward Orinoco. Results from the Relaxed Random Walk
41  analysis support a very recent origin of this species in the headwater Orinoco basin (Western
42  Guiana Shield, at late Pleistocene) with a later rapid colonization of the remaining Orinoco basin
43  and almost simultaneously the Negro River via Cucui, between 0.115 until about 0.001 Ma.
44  Historical biogeography and population genetic patterns observed here for Cardinal tetra, seem
45  to be better explained by river capture, physical, or ecological barriers than due to the geographic
46  distance.
47
48  **Keywords:** Microsatellite, COI, Ornamental fish, Casiquiare, Gene flow.
49

## Introduction

51
52  The Neotropics contain one of the most diverse assemblages of freshwater fishes in the world
53  (Reis, Kullander & Ferraris, 2003; Turner et al., 2004; Albert, Tagliacollo & Dagosta, 2020).
54  Within the Neotropics, the Orinoco and Amazon basins are characterized by their high
55  ichthyofauna diversity, exceeding 9,000 total species, of which ~ 5,160 are freshwater species
56  (Reis et al., 2016; Albert, Tagliacollo & Dagosta, 2020).
57      The origins of this ichthyofauna date back to the middle Miocene where most of the area
58  of what is now the western Amazon comprised the "Paleo Amazonas-Orinoco" drainage. It is
59  believed that this system drained northward into the Magdalena River, and, after the rise of the
60  Cordillera Oriental and Sierra de Perijá, into Lake Maracaibo (Hoorn, 1994; Hoorn et al., 1995;
61  Díaz de Gamero, 1996). This ancient river drained areas that are now part of the Magdalena,
62  Orinoco, and Amazonas River basins, implying an extensive, biogeographically interconnected
63  region (Albert, Lovejoy & Crampton, 2006; Winemiller & Willis, 2011).
64      The rise of the Vaupes Arch in the late Miocene (between 10 - 11 Ma) severed this
65  connectivity, resulting in the formation of two basins: the Orinoco, running north and northeast,
66  and the Amazon, flowing towards the east, until it reached its current delta in the Atlantic Ocean,
67  once the Purús Arch ruptured (Hoorn et al., 1995; Díaz de Gamero, 1996; Lundberg et al., 1998;
68  Hubert & Renno, 2006; Winemiller & Willis, 2011; Mora et al., 2011). At present, there is only
69  one permanent connection between the two basins: the Casiquiare Canal. The Casiquiare Canal
70  originated by a partial headwater capture of the upper Orinoco River, with the southern branch
71  emptying into the upper Negro River, the largest tributary of the Amazon River, and the northern
72  branch continuing as the Orinoco River (Winemiller et al., 2008).
73      Along its course, the Casiquiare Canal has a gradient of water types, ranging from clear
74  water (waters low in organic matter and sediment) near its source in the upper Orinoco, to black
75  waters (waters rich in organic matter but low in sediments) near its confluence with the Negro
76  River (Sioli, 1984; Crampton, 2011). This environmental gradient appears to act as an ecological

77    filter for several fish species, while the canal itself allows for connectivity between the
78    ichthyofaunas of the Orinoco and the Amazon basins (Winemiller et al., 2008).
79        In addition to the Casiquiare Canal, alternative corridors allowing movement of fishes
80    between the Orinoco and Amazon basins have been proposed (Winemiller et al., 2008; Escobar
81    et al., 2015). One of these corridors between the two basins is via the Essequibo River and the
82    Rupununi savanna during the high-water season. This alternative corridor has been used to
83    explain the dispersion of the ancestors of marine species that colonized the freshwater
84    environments of the two basins (Lovejoy & Araujo, 2000), and to explain the biogeographical
85    processes behind the diversification of species between the two basins (Escobar et al., 2015).
86        Although the Orinoco and Amazon basins have been separated for a long time, these two
87    basins share a large number of species (Reis, Kullander & Ferraris, 2003; Escobar et al., 2015).
88    One such species is *Paracheirodon axelrodi,* commonly known as the cardinal tetra. This fish is
89    distributed in the upper Orinoco basin and the upper and middle Negro basin in the Amazon
90    Basin. This species uses a flooded forest environment for feeding and reproduction. During the
91    flooding cycle they move from small streams to the flooded forest, returning to small streams
92    during low water season. These dynamics of flood cycles provide conditions for migration and
93    genetic exchange among populations (Cooke, Chao & Beheregaray, 2009).
94        The cardinal tetra is the most important ornamental fish in the world market (Chao, 2001;
95    Terencio, Schneider & Porto, 2012) comprising more than 80% of the total ornamental fish
96    exported from the Negro River each year (Chao, 2001; Beheregaray et al., 2004; De Oliveira et
97    al., 2008). Moreover, it is one of the most commercialized species in the Orinoco basin, ranking
98    in the first place between the years 1998 and 2004 (Tovar et al., 2009). Despite its extensive
99    geographic distribution spanning two major South American river basins, and extensive
100   commercial exploitation, very little is known about its population dynamics, population
101   structuring or evolutionary history of this species using molecular markers.
102       The first molecular study reported for *P. axelrodi* was that of Harris & Petry (2001). The
103   authors analyzed the mitochondrial control region for specimens from two geographically
104   proximate localities from the vicinity of Barcelos (Brazil) and specimens of unknown geographic
105   origin and found 3.9% and 4.3% sequence divergence between these specimens. In a much more
106   comprehensive study focusing on the Negro River, Cooke & Beheregaray (2007) found
107   extremely high levels of genetic diversity in the S72 intron in specimens sampled from the
108   tributaries of the Negro River. Subsequently, Cooke et al. (2009) evaluated population structure,
109   history of colonization, and genealogical relationships, concluding that headwater populations
110   showed higher genetic diversity than downstream populations. Bittencourt et al. (2017) analyzed
111   temporal genetic variation within a single population using five microsatellite loci and concluded
112   that changes in allelic frequencies over time are associated with droughts instigated by El Niño,
113   although these events do not significantly reduce genetic diversity. Despite these studies, none
114   has so far revealed the evolutionary behavior of the Orinoco population and its relationship to the
115   Amazon populations, combining information from mitochondrial and nuclear markers.
116       In this study we provide the first phylogeographic hypothesis and analysis of population
117   structure for *P. axelrodi* in its entire geographic distribution in the Orinoco and Amazon basins
118   using different molecular markers, including mitochondrial (COI), nuclear sequences (MYH6),
119   and eight microsatellite loci. Additionally, we test the hypothesis of basins connectivity and
120   provide support regarding potential routes of gene flow and colonization between the two basins.

121
## Materials & Methods
123
### Sample collection and DNA extraction
125

126  A total of 163 muscle tissue samples (one per individual) were collected along the Orinoco and
127  Amazon basins on the distribution area for *P. axelrodi*. The samples were collected by
128  ornamental fish fishermen from four localities on the Orinoco basin (Colombia) including San
129  José del Guaviare (*n*=20), Puerto Carreño (*n*=20), Puerto Gaitan (*n*=21), and Puerto Inirida
130  (*n*=21) (Fig. 1), and four localities in Negro River in the Amazon basin (Brazil) including Cucui
131  (*n*=21), Santa Isabel (*n*=20), Barcelos (*n*=20), and São Gabriel da Cachoeira (*n*=20) (Fig. 1).
132  Samples from São Gabriel da Cachoeira (Negro River – Amazon basin) were obtained from the
133  Laboratório de Genética Animal – LGA at Instituto Nacional de Pesquisas da Amazônia - INPA.
134  However, these samples could not be used for mitochondrial and nuclear DNA analysis due to
135  the low integrity of DNA, but they were suitable for microsatellite analysis. Samples from all
136  other Amazon localities were also obtained from local artisanal fishermen working in these areas
137  selling ornamental fish. Tissue samples and DNAs were deposited on the Laboratorio de
138  Ecología Molecular de Vertebrados Acuáticos – LEMVA at Universidad de los Andes,
139  Colombia. Details on sample information and depositaries can be consulted in the Table S1.
140      Once the fish were received, they were euthanized by rapid chilling (hypothermic shock)
141  on ice, following the American Veterinary Medical Association (AVMA) guidelines for the
142  euthanasia of animals: 2020 Edition ([https://www.avma.org/resources-tools/avma-policies/avma-](https://www.avma.org/resources-tools/avma-policies/avma-guidelines-euthanasia-animals)
143  [guidelines-euthanasia-animals](https://www.avma.org/resources-tools/avma-policies/avma-guidelines-euthanasia-animals)). This procedure is recommended for small-bodied (3.8-cm-long
144  or smaller) tropical and subtropical stenothermic fish like cardinal tetra. Then, all individuals and
145  tissue samples were preserved in ethanol at 95% and stored at 4°C. Genomic DNA was extracted
146  using MoBio UltraClean® Tissue and Cells DNA Isolation Kit or following a phenol/chloroform
147  protocol (Sambrook, Fritsch & Maniatis, 1989).
148
### PCR amplification and Sequencing
150
*Mitochondrial DNA*
152

153  In this study, 96 mitochondrial fragments of the cytochrome oxidase I (COI) gene corresponding
154  to 468 base pairs (bp) length each one (7-21 individuals per locality), was amplified using the
155  primers designed by Ivanova et al. (2007). PCR reactions were carried out in a final volume of
156  15 μL and contained 8.3 μL of ddH$_2$O, 1.2 μL of MgCl$_2$ (25 mM), 1.5 μL of 10x buffer
157  ((NH$_4$)$_2$SO$_4$), 1.2 μL de dNTPs (10 mM), 1.5 μL of both primers (2 μM), 0.3 μL of Taq DNA
158  Polymerase (1 U/μL) and 1 μL of DNA (~ 50 ng/μL). PCR conditions were as follows: 72°C for
159  1 min, 35 cycles of 94°C for 10 sec, 50°C for 35 sec and 72°C for 90 sec, with a final extension
160  period of 72°C for 5 min.
161
*Nuclear DNA*
163

164  Nuclear genome sequences corresponding to 555 bp of exonic region myosin heavy polypeptide
165  6 cardiac muscle alpha (MYH6), were amplified only for samples of *P. axelrodi* representing
166  unique haplotypes for COI gene within each locality, previously identified as described in the

167 "**Phylogeographic analyses**" section. We used the primers designed by Li et al. (2007). For
168 nuclear gene (MYH6) amplification, nested PCR was performed. For this, two independent
169 working solution of 2 µM were set up by combining equimolarly the forward and reverse
170 primers for each gene: MYH6_F459 (5' CATMTTYTCCATCTCAGATAATGC 3') with
171 MYH6_R1325 (5' ATTCTCACCACCATCCAGTTGAA 3'). Then, the 1st PCR reaction was
172 carried out in a final volume of 15 µL and contained 5.1 µL of ddH$_2$O, 1.5 µL of BSA (10
173 mg/mL), 1.2 µL of MgCl$_2$ (25 mM), 1.5 µL of 10x buffer ((NH$_4$)$_2$SO$_4$), 1.2 µL of dNTPs (10
174 mM), 1.5 µL of each primer working solution, 0.5 µL of Taq DNA Polymerase (1 U/µL) and 1
175 µL of DNA (~ 50 ng/µL). The 2$^{nd}$-PCR reaction was carried out in a final volume of 15 µL and
176 contained 8.3 µL of ddH$_2$O, 1.2 µL of MgCl$_2$ (25 mM), 1.5 µL of 10x buffer ((NH$_4$)$_2$SO$_4$), 1.2 µL
177 de dNTPs (10 mM), 1.5 µL of both primers (2 µM), 0.3 µL of Taq DNA Polymerase (1 U/µL)
178 and 1 µL of 1$^{st}$ PCR product. PCR reagents were obtained from Fermentas (Vilnius, Lithuania).
179 PCR conditions for nuclear genes were as follows: 1st PCR - 68°C for 1 minute, 30
180 cycles of 93°C for 10 sec, 48°C for 35 sec and 68°C for 90 sec, with a final extension of 68°C
181 for 7 min. 2nd–PCR: 68°C for 1 min, 35 cycles of 93°C for 10 sec, 55°C for 35 sec and 68°C for
182 90 sec., with a final extension period of 68°C for 7 min. PCR products were purified using EXO-
183 SAP, (Exonuclease—Shrimp Alcaline Phosphatase) following the manufacturer's suggested
184 protocol (Werle et al., 1994). All PCR products (mtDNA and nuDNA) were sequenced on an
185 ABI 3500 automated sequencer (Applied Biosystems) using the BigDye Terminator Kit Cycle
186 Sequencing Kit (Applied Biosystems) according to the manufacturer's instructions. The
187 assembly, alignment, and quality check of the nucleotide sequences were performed using
188 Geneious (Kearse et al., 2012).
189 Eight polymorphic microsatellite loci (*Pa7*, *Pa13*, *Pa19*, *Pa26*, *Pa27*, *Pa32*, *Pa33*, *Pa37*)
190 developed for *P. axelrodi* by Beheregaray et al. (2004) were used. PCR reactions were carried
191 out in a final volume of 10 µL and contained 3.7 µL of ddH$_2$O, 1.0 µL of MgCl$_2$ (25 mM), 1.0
192 µL of 10x buffer((NH$_4$)$_2$SO$_4$) ,1.0 µL of dNTPs (10 mM), 1.0 µL of reverse primer ( 2 µM), 0.5
193 µL of forward primer (2 µM), 0.5 µL of FAM – labelled M13 primer (2 µM), 0.3 µL of Taq
194 DNA Polymerase (1 U/µL) and 1 µL of DNA (~ 25 ng/µL). PCR reagents were obtained from
195 Fermentas (Vilnius, Lithuania). Amplification conditions were as follows: 94°C for 1 min, 25
196 cycles of 94°C for 30 s, 60°C for 30 s and 68°C for 40 s, 30 cycles of 94°C for 20 s, 52°C for 30
197 s and 72°C for 1 min, with a final extension of 72°C for 30 min. Then, 1 µL of PCR product was
198 combined with 1 mL of ROX size standard (DeWoody et al., 2004) and 8 mL of Hi-Di
199 formamide. All microsatellite loci were amplified separately. They were then combined and
200 genotyped on an Applied Biosystems ABI 3100 Genetic Analyzer (Universidade Federal do
201 Amazonas – UFAM). Allele sizes were inferred using the pUC19 ROX-labelled size standard
202 (DeWoody et al., 2004). Subsequently, a matrix of genotypes for each individual was generated.
203
204 **Phylogeographic analyses**
205
206 All sequences were edited manually and aligned using the software Geneious v4.7 (Kearse et al.,
207 2012). After the alignment by the ClustalW algorithm (Thompson, Higgins & Gibson, 1994),
208 unique COI haplotypes within each population was inferred using the online fasta sequence
209 toolbox DNAcollapser from FaBox v1.61 (https://birc.au.dk/~palle/php/fabox/; Villesen, 2007).
210 Then, a haplotype network was constructed for each gene to determine ancestry relationships. In
211 the case of the MYH6, sequences were first phased into individual alleles using the PHASE
212 algorithm (Stephens, Smith & Donnelly, 2001) incorporated into DnaSP v5.1 software (Librado

213 & Rozas, 2009). Then, we used a maximum likelihood approach in RAxML v8.2.X for
214 phylogenetic analyses (Stamatakis, 2014) in CIPRES (Miller, Pfeiffer & Schwartz, 2010)
215 because of its ability to efficient and fast maximum likelihood tree search algorithm that returns
216 trees with good likelihood scores. We used a GTR+GAMMA model of sequence evolution, the
217 default and unique available for RAxML analysis, according the authors of the program in the
218 version 8.2.X manual, for single full ML tree searches. In addition, we generate 1,000 replicates
219 of RAxML's rapid bootstrap algorithm to account for uncertainty in the estimation of the
220 topology (Stamatakis, 2014). A consensus tree was constructed from the bootstrap output file
221 without ruling out any of the 1,000 replicates, for the allele network visualization into
222 Haploviewer software (Salzburger, Ewing & Von Haeseler, 2011), a haplotype viewer program
223 specifically developed for the purpose of reconstructing haplotypes networks using traditional
224 phylogenetic algorithms, neighbour-joining, maximum parsimony and **maximum likelihood**
225 genealogies from closely related, and hence, highly similar haplotype sequence data.
226      As a final analysis, we used non-concatenated COI and MYH6 sequences in a unique
227 phylogeographic analysis with a Bayesian statistical approach based on a relaxed random walk
228 model in continuous space and time in BEAST v1.8.2 software (Drummond et al., 2012), on the
229 CIPRES platform (Miller, Pfeiffer & Schwartz, 2010). This analysis was used for the
230 reconstruction of the origin and dispersal routes of *P. axelrodi* between Amazon and Orinoco
231 basins. This method infers evolutionary history in a continuum landscape combining DNA
232 sequences with geographic coordinates, generating a genealogy, and estimating the ancestral
233 localities of the internal nodes, taking into consideration uncertainties in topology (Lemey et al.,
234 2010).
235      The models of substitution selected for the Bayesian analysis were HKY+GAMMA for
236 COI and, GTR+GAMMA for MYH6. Both substitution models were previously estimated in
237 JModelTest (Posada, 2008). As tree prior, a relaxed non-correlated model with a relaxed
238 molecular clock was used, with the tree coalescence prior of the Bayesian Skyline. For the time
239 calibration of the resulting tree, the geometric median derived from a substitution rate of
240 $0.68 \times 10^{-8}$ mutations per site per year for COI (from the widely accepted mtDNA substitution rate
241 for poikilotherm vertebrates in Martin & Palumbi (1993), and $1.0 \times 10^{-9}$ for MYH6 (Freeland,
242 2005). Two independent runs were performed independents with a chain length of a 100 million,
243 with a log every 10,000 steps. Run convergence was evaluated in Tracer v1.6.0 considering the
244 stationary behavior of the chains and the Effective Sample Size (ESS) > 300. Summary of the
245 topologies to generate a clade credibility tree, was done on TreeAnnotator v1.8.2. Then, the two
246 independent runs were combined. To generate space and time projections of the genealogies,
247 keyhole markup language (kml) was used in the software SpreaD3 v0.9.6 (Bielejec et al., 2016),
248 and the results were accessed and visualized using Google Maps.
249
250 **Gene flow and population genetic structure analyses**
251
252 To establish the genetic structure and connectivity between the Orinoco and Negro River
253 populations, a matrix microsatellites allele was initially constructed in GeneMapper v4.1
254 (Applied Biosystems). Null allele frequencies were estimated in FreeNA Software (Chapuis &
255 Estoup, 2007).
256      Arlequin v3.5 (Excoffier & Lischer, 2010) was used to run an Analysis of molecular
257 variance AMOVA and Pairwise *Fst* among localities. Additionally, the number of alleles per
258 locus ($N_A$), the observed ($Ho$), and the expected heterozygosity ($H_E$) for every locus were

259   estimated. Deviations from Hardy-Weinberg equilibrium (HWE) were calculated for each locus
260   for each sampling locality. Similar analyzes were done for COI, including pairwise *Φst* index,
261   haplotype (h) and nucleotide (π) diversity. In all cases, significance (P-values) were calculated
262   using 10,000 non-parametric permutations. P-values were adjusted using Bonferroni correction
263   for multiple comparisons (Rice, 1989).
264        Additionally, isolation by distance (IBD) was analyzed for COI and microsatellites, using
265   the Mantel test from a matrix of genetic (*Fst/Φst;* previously calculated) and geographic (km)
266   distances among the sampled localities, using the Arlequin v3.5 software (Excoffier & Lischer,
267   2010) and following all steps included in its manual
268   (http://cmpg.unibe.ch/software/arlequin35/man/Arlequin35.pdf). The geographical distances
269   were obtained following the course of the rivers via the coordinates of each location using
270   GoogleEarth Pro (Google Inc). The analyses were performed to evaluate if the spatial processes
271   are driving population structure, or, in other words, if the IBD is the best explanation to the
272   observed genetic structure patterns for the sedentary *P. axelrodi.*
273        To estimate the number of genetic clusters and to determinate the degree of admixture
274   and gene flow among populations, we used the Bayesian clustering approach as implemented in
275   the software STRUCTURE v2.3.3 (Pritchard, Stephens & Donnelly, 2000). We used the
276   admixture and correlated frequencies priors and performed 10 replicates for K = 1 to 8 (testing
277   each locality as a possible cluster). Assignment analyses were executed with 1 million steps
278   Markov-Chain-Monte-Carlo (MCMC) with a burn-in of 10%. The convergence of the MCMC
279   was inferred from the plot of α value of each independent run. The runs were analyzed in the
280   program Structure Harvester v0.6.92 (Earl & VonHoldt, 2011), and independent replicates for
281   each K were summarized in the program CLUMPP (Jakobsson & Rosenberg, 2007) and
282   visualized using the software Distruct (Rosenberg, 2003). The most likely number of biological
283   groups (K) was inferred using the method of Evanno, Regnaut & Goudet (2005). However, since
284   DeltaK (Evanno) leads to wrong inferences on hierarchical structure and downward-biased
285   estimates of the true number of sub-populations, we used alternatively the Puechmaille method
286   (Puechmaille, 2016) to infer the number of biological groups (K), using the web-based software
287   StructureSelector (Li & Liu, 2018).
288        Patterns of genetic structure were further explored, using the diploid genotypes of 8 loci
289   (16 variables) in the complete genotype matrix of the 163 individuals; through a discriminant
290   analysis of principal components (DAPC) using R- package Adegenet (Jombart, 2008), and
291   following the instructions found in the manual developed by Jombart & Collins in 2015
292   (https://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf). DAPC construct linear
293   combinations of the original variables (alleles) which have the largest between-group variance
294   and the smallest within-group variance. First, the genotype matrix (STRUCTURE format) was
295   converted to a *genind* object using the function "read.structure". Then, it was implemented the
296   function "dapc", which first transforms the data using PCA, and then performs a Discriminant
297   Analysis on the retained principal components.
298        According to the observed genetic structure patterns and biological groups formation
299   through STRUCTURE and DAPC analyses, we defined biogeographical populations, and then,
300   quantified the gene flow in terms of the effective number of migrants per generation (*Nm*)
301   between theme, using the program Migrate v3.6.11 (Beerli, 2006), following the spirit of the
302   methodology previously implemented by Lagostina et al. (2018). For this purpose, we tested
303   different migrations models between the paired predefined populations [1:full (xxxx);
304   2:unidirectional (x0xx); 3:unidirectional (xx0x); 4:panmictic (x)], and then, we selected the best

305    one by comparing theirs Marginal Likelihood based on the Bezier approximation score, through
306    the Bayes Factor method. For the Theta and M parameters calculations, we implemented
307    uniforms priors. The search strategy used static heating scheme with four chains at different
308    temperatures (1.0, 1.5, 3.0, 1000000.0) and one replicate of one long chain with 1 million
309    generations, discarding a burn-in of 100,000, and sampling every 500 generations, as previously
310    described in Muniz et al. (2017). We diced to evaluate (1) the effective sample size (ESS>500),
311    (2) autocorrelation values (<0.2), and, (3) the correct distribution of the probability density plots
312    for estimates of the parameters; to verify the convergence of parameter estimates, as previously
313    described in Muniz et al. (2017) and in Lagostina et al. (2018). To convert the scaled parameter
314    "M" to biological data in terms of number of effective migrants (Nm), we assumed a
315    conservative generation time of 1 year
316    (https://animaldiversity.org/accounts/Paracheirodon_axelrodi/); and a conservative substitution
317    rate for fish microsatellites of $1.0 \times 10^{-5}$ mutations per locus per generation [reviewed in
318    Chistiakov, Hellemans & Volckaert (2006)].
319

## Results

321

### Phylogeographic patterns

323

324    We identified 2-4 haplotypes per locality, reaching 20 in total (see Table S1 and GenBank
325    accessions MG384556 - MG384575). The COI haplotype network did not show shared
326    haplotypes between the Orinoco and Negro basins. However, no reciprocal monophyly was
327    found between basins (Fig. 2; Fig. S1). In addition, for the 20 nuclear MYH6 gene sequences
328    (see Table S1 and GenBank accessions MG384576 – MG384595), they were inferred 40 alleles,
329    and the resultant network showed 25 alleles shared among populations between basins (Fig. 3;
330    Fig. S2), mainly for individuals from the Cucui and Santa Isabel with individuals from the
331    Orinoco basin, breaking the reciprocal monophyly between basins (Fig. S2).
332            The analysis of phylogenetic reconstruction Relaxed Random Walk (Fig. 4; Fig. S3)
333    estimated that *P. axelrodi* ancestral population was located in the Orinoco basin in the Late
334    Pleistocene (*ca.* 0.255 Ma), between Inírida and Cucui, on the Vaupés Arch near the border of
335    the Guiana Shield. The signal of both markers suggested that from there (approximately 0.115
336    Ma), that ancestral population of *P. axelrodi* dispersed faster along the Orinoco basin reaching
337    all localities studied simultaneously. Later, around 0.019 Ma, it colonized the Negro River basin
338    via Cucui, arriving in its middle basin through Santa Isabel in the Holocene, approximately
339    between 0.003 Ma and 0.001 Ma. Finally, *P. axelrodi* reached the lower Negro River through
340    Barcelos, the last location, to be colonized in the last 1,000 years.
341

### Gene flow and population genetic structure

343

344    FreeNA results showed that no null alleles were present in any of the loci analyzed. Genetic
345    diversity values such as expected (*HE*) and observed Heterozygosity (*Ho*), and number of alleles
346    per population (n) were obtained from eight loci screened in all individual of *P. axelrodi* along
347    with deviations from HW equilibrium. Microsatellite loci had between 2 and 22 alleles per locus.
348    Heterozygosity values were very similar for Puerto Carreño and Puerto Gaitan while *HE* was the
349    highest in Cucui and lowest in Barcelos (Table S2), after Bonferroni correction (*P-value* =

350  0.0007812) deviation from HWE was observed at the majority of loci in Orinoco basin
351  populations except *(Pa7)*. Among the Negro basin populations, deviation from HWE was
352  observed most loci except *(Pa7, Pa13, and Pa26)* (Table S2).
353      The $F_{ST}$ pairwise population differentiation analysis showed that within Orinoco basin
354  there was no differentiation between Puerto Gaitán and Puerto Inírida ($F_{ST}: \sim 0.010$; P>0.05),
355  showing panmixia between them. There was significant differentiation among the rest of
356  populations with values ranging from ~ 0.10 to 0.24; P<0.05), showing structuring among them
357  (Table S3). Additionally, inside Negro basin all populations were found to be differentiated ($F_{ST}$
358  values from ~ 0.03 to 0.17; P<0.05). According to the AMOVA, most of genetic variation in *P.*
359  *axelrodi* was within populations (86.53%; P>0.05). However, a small but significant
360  portion (13.47%; P<0.0001) was occurring among populations.
361      The pairwise $\Phi_{st}$ index for COI showed a high degree of significant structuring, except
362  among some Orinoco populations ($\Phi_{st}= \sim 0.05$-0.115; P>0.05). For instance, in the Negro basin,
363  each location was highly structured, showing an $\Phi_{st}$ index between ~ 0.68 and 0.95 (P <0.05).
364  Similarly, pairwise differentiation values between basins, they ranged from 0.58 up to 0.98,
365  indicating a highly significant structure (P<0.05) (Table S4; Supplemental Information).
366      According to Mantel Test, the correlation between pairwise $F_{st}$ and geographic distances
367  for the eight *P. axelrodi* populations was moderate but statistically significant (r=0.591787;
368  P=0.008). However, no robust evidence was found to infer Isolation by Distance, since only 35%
369  of genetic differentiation observed among *P. axelrodi* populations can be explained by the
370  geographic distance ($R^2$=0.350211).
371      The assignment test analysis conducted in STRUCTURE showed two biological clusters
372  (K=2) as the highest posterior probability of population structure [Mean Ln(K)= -5063.080],
373  according the Evanno method (Fig. 5; Fig. S4). The two populations units detected by
374  STRUCTURE were: (1) Orinoco River populations (Guaviare, Puerto Carreño, Puerto Gaitan,
375  and Puerto Inirida) + Cucui (*n*=103); and (2) the remaining of the Amazon River populations
376  (Barcelos, Santa Isabel and São Gabriel da Cachoeira) (*n* = 60) (Fig. 5). However, the
377  Puechmaille method estimators suggested the existence of a genetic substructure, with K=6 as
378  the highest probability of biological clusters (Fig. S4). As an observable pattern from K=2 to
379  K=6 (Fig. 5), the Rio Negro populations (except Cucui) are highly structured compared to those
380  of the Orinoco. Cucui is shown as an independent population or more associated with the
381  Orinoco River basin, but less associated with the Negro River basin. In all of them, Cucui and
382  San Gabriel are observed as a transition zone between the Orinoco and Negro basins. The DAPC
383  is concordant with the observed pattern K=2 – K=6 described above (Fig. 6). Both,
384  STRUCTURE and DAPC showed Puerto Carreño (Orinoco basin) as the locality with the
385  highest signals of co-ancestry or relatedness with the populations from the Negro River basin
386  (except Cucui).
387      According to the genetic structure pattern observed, we tested in Migrate the gene flow
388  among three biogeographical groups: Orinoco basin, Cucui and rest of Negro basin. As a result,
389  we found that the unidirectional pattern was the best model explaining migration among these
390  groups: Cucui toward Orinoco basin [*Nm*=4.8 migrants per generation, Supplementary material
391  1]; Cucui toward rest of Negro basin [*Nm*=6.58 migrants per generation, Supplementary material
392  2] and rest of Negro basin toward Orinoco basin [*Nm*=4.5 migrants per generation,
393  Supplementary material 3]. Supplementary material 1, 2, and 3 can be accessed through
394  DRYAD, https://doi.org/10.5061/dryad.7kd80.

## Discussion

### Genetic structure and phylogeographic patterns

Despite its sedentary condition and evidence of structuring between basins, cardinal tetra populations appear to have had a historical connection between basins. The nuclear alleles network and the genetic structure results, shown signals of this fact, and also, that Cucui (Negro basin) is more associated with the Orinoco basin, while Santa Isabel and Barcelos are shown as another independent biological group, being São Gabriel da Cachoeira the transition area between them. This observed pattern (Orinoco+Cucui / rest of the Negro basin) is probably explained by an ancient geological event such as the formation of waterfalls and rapids located in São Gabriel da Cachoeira. Our hypothesis suggests that the waterfalls may be the first split event dividing the Orinoco basin + Cucui populations from the rest of the Negro basin populations (São Gabriel de Cachoeira, Santa Isabel and Barcelos). It is known that the waterfalls system and rapids near São Gabriel da Cachoeira act as a physical barrier structuring the populations of small species between the upper section and the remaining area of the Negro River (Albert & Reis, 2011; Terencio, Schneider & Porto, 2012).

However, when the results are explored hierarchically in STRUCTURE (K3 - K6) and DAPC, it is found a substructure that shows Cucui isolated from the Orinoco, currently showing itself as an independent population from the Negro River and Orinoco River. This suggests that after the São Gabriel waterfalls formation, a second more recent split event occurred between the headwaters of the Negro River (where Cucui is located) and the Orinoco, until then connected probably through the Inírida River. This separation event was most likely influenced by the final Vaupes Arch rising. It is known that this rising happened from west to east, towards the Guiana Shield, and studies evidenced the influence of this process on fish fauna divergence between the resulting basins (Mora et al., 2011; Caputo & Soares, 2016). These split hypotheses modeling part of the structure of *P. axelrodi* between basins are also supported by the Mantel test, where the distance does not explain by itself the current patterns of genetic structuring.

Another of the remarkable genetic structure patterns was the found for Santa Isabel and Barcelos. STRUCTURE (K2- K6), DAPC and fixation indices, coincide in showing Santa Isabel and Barcelos as a genetically different stocks from São Gabriel da Cachoeira and Cucui. Since the Isolation by Distance hypothesis was rejected, the explanation for this pattern of genetic structure may be associated with the influence of the whitewater river in the section that comprises the distribution of these two populations or other of ecological order. This pattern of high genetic structure can also be observed in ornamental fish such as *Nannostomus eques* (Terencio, Schneider & Porto, 2012) and in *Carnegiella strigata* (Schneider et al., 2012). In these species, physicochemical differences in types of water produced by at least three whitewater tributaries such as: Demini, Padauari, and Branco Rivers and higher pH levels, sediment load, and turbidity, were showed as the more likely factors causing the observed structure in these populations. It is known that *P. axelrodi* is a blackwater fish (Chao, 2001; Terencio, Schneider & Porto, 2012), which allows it, as well as other native teleost to the Negro River, have a special adaptation that allows them ionoregulatory specializations for exceptional tolerance of ion-poor, acidic waters of this blackwater river (Gonzalez et al., 1998; Gonzalez & Wilson, 2001; Matsuo & Val, 2007; Wood et al., 2014).

440     On the other hand, we observed that unlike the Negro River, there is not a very strong
441 structure among the populations of the Orinoco basin. This can be explained by several
442 characteristics of the area in which the Orinoco is distributed. For example, the topography,
443 being the Llanos region a plain area [savannahs (alluvial plain of the Orinoco River (0–50
444 m.a.s.l.)] (Stallard, 1985), which does not have relevant geological formations that may have
445 caused a marked genetic vicariant or allopatric pattern of genetic structure in *P. axelrodi*. So, as
446 there were no such strong barriers, this could have caused a more rapid, recent dispersal and
447 establishment of populations, where there was not enough time to accumulate differences, as
448 shown by Random Walk and pairwise $F_{ST}$ indices. However, according to STRUCTURE (K4-
449 K6) and DAPC, we found genetic structure among Guaviare and the rest of the Orinoco
450 populations. This and other structuring process ongoing for the Cardinal tetra, may be explained
451 by three factors. First, since the *P. axelrodi* is a blackwater fish (Chao, 2001; Terencio,
452 Schneider & Porto, 2012), it is likely that Guaviare River, which is whitewater, can be a barrier
453 preventing the current dispersion and connectivity of *P. axelrodi* with the rest of the Orinoco
454 populations. Second, it is possible that the highest elevation of the Vaupes Arch in its western
455 area (Winemiller & Willis, 2011; Mora et al., 2011), at the Andean Piedmont, where the
456 Guaviare population is located (>200 m.a.s.l.), may be causing a gradual isolation of such
457 individuals from the rest of Orinoco basin (mean altitude=98 m.a.s.l.). Third one, the life history
458 and the small body size of this species can help prevent dispersion, generating structure
459 (Godinho, Lamas & Godinho, 2010). This is expected for sedentary fishes such as *P. axelrodi*,
460 especially between populations separated by large distances (Lovejoy & Araujo, 2000; Cooke,
461 Chao & Beheregaray, 2009).
462
463 **Connectivity and dispersion routes**
464
465 It is clear that the Vaupés Arch fragmented the transcontinental Paleo Amazonas-Orinoco River
466 about ~ 5-10 Ma (Hoorn et al., 1995), generating until today 61.2% of exclusive taxa to Amazon
467 and 16.6% to Orinoco (Winemiller & Willis, 2011). However, 22.2% of the taxa are still shared
468 between basins until the present, probably by the persistent connections between the Orinoco and
469 Amazon basins via the Casiquiare Canal and other blackwater tributaries (Winemiller et al.,
470 2008). Until today, it was unclear to which of these three groups *P. axelrodi* belonged. However,
471 as expected, in this study it was demonstrated that Cardinal tetra is a species still shared between
472 Orinoco and Negro basins, establishing current (or at least until a very recent past) gene flow in a
473 biological scale of ~ 4 migrants per generation between them. These results suggest a connection
474 via Cucui toward Inírida route (*e.g.* across western Casiquiare canal tributaries), or through a
475 Cucui independent route, via rest of the Negro basin toward Orinoco (*e.g.* across eastern
476 Casiquiare canal tributaries, as suggested by the marked co-ancestry between Puerto Carreño and
477 rest of the Negro basin). Patterns of connection between the Orinoco and Negro basins, has also
478 been suggested previously in Cardinal tetra by Cooke et al. (2009). Likewise, other previous
479 genetic studies demonstrated that non-migratory fishes such as the cychlids *Cichla temensis*, *C.
480 monoculus*, and *C. orinocensis* (Willis et al., 2010) or the black arowana *Osteoglossum ferreirai*
481 (Olivares et al., 2013), experienced some degree of connectivity and gene flow between both
482 basins, using the Casiquiare as their connection route. Contrary, aquatic species from white water
483 environments such as the riverine turtle *Podocnemis unifilis* (Escalona et al., 2009) or the fish
484 *Piaractus brachypomus* (Escobar et al., 2015), are unable to currently establish connection/gene
485 flow between the Orinoco and Amazon. It suggests that since these alternative connections are

486  between blackwater drainages, dispersal between the Orinoco and Negro basins could be
487  facilitated for small species that are adapted to blackwater environments (Winemiller & Willis,
488  2011), being *Paracheirodon axelrodi* one of such species.
489      There are currently at least five alternative connections among the tributaries of the upper
490  Orinoco and Negro to the west of the Casiquiare canal (Dungel, 2009; Winemiller & Willis,
491  2011). One such connection is via the isthmus of Pimichin. The upper reaches of the Temi
492  branch of the Atabapo River (Orinoco drainage) are connected with a 16 km marshy and
493  swampy plain with the Pimichin creek, an affluent of the upper Guainia (Negro drainage) River.
494  Another potential connection is via a Llanos floodplain, less than 10 km wide, that connects the
495  left-bank affluents of the upper Guainia River and the Guacamayo creek that drain into the
496  Inírida River (Orinoco basin). Yet another alternate connection is via the right bank affluent of
497  the Conochirite River and tributaries of the Atacavi and Temi Rivers (Orinoco basin) separated
498  by less than 10 km of flat, and at least seasonally marshy terrain. The Conochirite River is a left
499  bank affluent of the Guainia River, and during the high-water season is also connected to the
500  upper Casiquiare Canal. A fourth connection is between the floodplains of the Baria River
501  (affluent of the Pasimony River, Casiquiare canal drainage) and the Maturaca River (affluent of
502  the Cauaburi River, Negro River drainage). The connection via the Pimichin isthmus was already
503  mentioned by Alexander von Humboldt and confirmed by (Rice, 1921). Based on the
504  examination of topographic maps, Winemiller & Willis (2011) suggest that this connection may
505  have existed until very recently, or perhaps is current but ephemeral. The Conochirite River
506  connection was also used by Humboldt instead of the lower Casiquiare Canal to arrive at the
507  Guainia River near Meme da Maroa. The last potential connection is much further to the west
508  across a less than 10 km flat floodplain between the upper Içana River, a tributary of the Negro
509  River, and the Papunáua River, a tributary of the Inírida River.
510      Although the mitochondrial DNA did not show shared haplotypes between basins, this
511  result is not conclusive to determine the absence of gene flow among Orinoco and Negro basins
512  populations. This could mean that females are philopatric, a behavioral observation of males
513  leaving the natal troop before sexual maturity and females staying with the troop for life
514  (Caparroz, Miyaki & Baker, 2009). So, when a species with sex-biased dispersal, the
515  simultaneous use of markers with different modes of inheritance (i.e., uni- and biparental) can
516  reveal the contrasting patterns of spatial distribution of the species' genetic variation. As a result,
517  stronger population structure can be detected with the mtDNA than with nuclear DNA
518  (Caparroz, Miyaki & Baker, 2009), such as the case of *P. axelrodi*.
519      However, there are other potential reasons for this result. Firstly, there were technical
520  limitations in COI amplification for all 163 individuals in the study [uneven sampling (7-15
521  samples per locality)]. In addition, localities such as São Gabriel da Cachoeira [with strong
522  signals of connection with Orinoco (according microsatellites)] were not sampled for the COI
523  marker due to low DNA quality, as explained above. Additionally, microsatellites are
524  codominant markers that more robustly examine the demographic behavior of both genders in
525  the population and not only the behavior of females (mtDNA) (Freeland, 2005).
526
527  **Historical biogeography and diversification patterns**
528
529  In this study, according to mitochondrial and nuclear diversity indices, we found higher values in
530  Orinoco River populations such as Puerto Inírida, Guaviare and Puerto Carreño. This may
531  suggest that the ancestral origin of *P. axelrodi* is in the Orinoco basin, as reported by Cooke et al.

532 (2009), founding that Negro River headwater populations had the highest gene diversity values
533 compared to the middle Negro River populations.
534      Random Walk analysis confirmed that ancient populations belong to the Orinoco basin,
535 on the Western Guiana Shield, near Inírida. From here, it rapidly colonized the rest of Orinoco
536 and almost simultaneously the upper Negro River (via Cucui), followed by the median and lower
537 part of this river. These results are in accordance with the chronological Negro River formation
538 proposed by Latrubesse & Franzinelli (2005). These authors suggested that Negro River
539 headwaters belong to an ancient formation while middle and lower Negro River floodplains
540 during the final of Quaternary, belongs to a more recent formation. This observed headwater
541 toward lower Negro colonization pattern, was also confirmed for *P. axelrodi* by Cooke et al.
542 (2009). These authors showed that populations in the headwater and upper sections of Negro
543 River represent ancient lineages, while downstream populations represent more recent lineages.
544      Although the ancestral population of *P. axelrodi* was originated around 0.255 Ma, the
545 colonization of the middle and lower Negro basin occurred only about <3,000 years ago.
546 According to Ruokolainen et al. (2018), the middle and lower sections of the Negro basin were
547 not totally blackwaters until about 1,000 years ago, because of the direct influence of the
548 whitewater Japurá River, that until then, drained into the Negro River, probably causing
549 extensive changes in the physical-chemical characteristics of the water downstream middle
550 Negro River, preventing the colonization of *P. axelrodi,* a restricted blackwater fish.
551      The colonization of *P. axelrodi* along the blackwater systems between the Orinoco and
552 Negro basin, occurred rapidly and very recently, from the Late Pleistocene until the Holocene
553 (~255,000 – 1,000 years ago; see Fig. S3), having its origin at the Western Guiana Shield.
554 According to Albert, Val & Hoorn (2018), during this period, it was evidenced a generalized
555 pattern of river capture for the coastal-draining basins of the Western Guiana Shield. It was
556 observed in the many signature elbows, U-turns, rapids and waterfalls of the uppermost Negro
557 and Branco basin headwaters (*e.g.* Casiquiare, Siapa, Mucajaí, Uraricoera and Taku-tu). These
558 River captures resulted in geodispersal of coastal-basin headwaters to the Amazon (and probably
559 to Orinoco), thereby enriching the Amazonian biota, and vicariances isolating sister species on
560 either side of the Western Guiana Shield-Amazon divide [reviewed in Albert, Val & Hoorn
561 (2018)]. For instances, it is known that until about 1 million years ago, the upper Branco River
562 drained into the Essequibo River (Schaefer & Júnior, 1997), and only in the mid-late Pleistocene,
563 about 0.25-0.1 Ma (Cremon et al., 2016), it was captured by the Negro River to become part of
564 the Amazon basin.
565
566 ## Conclusions
567
568 This study presents the first comprehensive analyses of population structure, connectivity, and
569 phylogeographic patterns for cardinal tetra (*Paracheirodon axelrodi*) from the Orinoco and
570 Amazon basins using coalescent- and frequency-based analyses of mitochondrial and nuclear
571 DNA markers. Our results suggest that *P. axelrodi* is a young species originated at Western
572 Guiana Shield drainages (near Inírida) at late Pleistocene (~0.255 Ma), with a very rapid and
573 recent colonization of both basin from 0.115 Ma until about 1,000 years ago. Currently, the
574 species is distributed between basins in a genetic structuring pattern of three geographical
575 populations, even in the face of a restricted gene flow among them: (1) Orinoco [Guaviare,
576 Puerto Gaitán, Puerto Inírida and Puerto Carreño], (2) Cucui [headwater Negro basin] and (3)
577 rest of the Negro basin [São Gabriel da Cachoeira, Santa Isabel and Barcelos]. Vaupés Arch is

578 not a barrier to gene flow between basins for *P. axelrodi*. Instead, permanent connection between
579 Orinoco and Negro basins, as the Casiquiare Canal or alternative black-water corridors drainages
580 are allowing connection between both basins (currently or at least until a very recent past). This
581 study provides key information to scientists, managers, and governmental agencies regarding the
582 management of this fish as an important resource for ornamental industries in northern South
583 American countries.
584

## Acknowledgments

586
587 The authors thanks Valéria Nogueira Machado, Mario Nunes and Sandra Hernández for their
588 support in field and laboratory work.
589

## References

591

592 Albert JS, Lovejoy NR, Crampton WGR. 2006. Miocene tectonism and the separation of cis- and
593      trans-Andean river basins: Evidence from Neotropical fishes. *Journal of South American*
594      *Earth Sciences* 21:14–27. DOI: 10.1016/J.JSAMES.2005.07.010.
595 Albert JS, Reis RE. 2011. *Historical biogeography of Neotropical freshwater fishes*. Berkely:
596      University of California Press.
597 Albert JS, Tagliacollo VA, Dagosta F. 2020. Diversification of Neotropical Freshwater Fishes.
598      *Annual Review of Ecology, Evolution, and Systematics* 51:27–53. DOI: 10.1146/annurev-
599      ecolsys-011620-031032.
600 Albert JS, Val P, Hoorn C. 2018. The changing course of the Amazon River in the Neogene :
601      center stage for Neotropical diversification. *Neotropical Ichthyology* 16:1–23. DOI:
602      10.1590/1982-0224-20180033.
603 Beerli P. 2006. Comparison of Bayesian and maximum-likelihood inference of population
604      genetic parameters. *Bioinformatics* 22:341–345. DOI: 10.1093/bioinformatics/bti803.
605 Beheregaray LB, Möller LM, Schwartz TS, Chao NL, Caccone A. 2004. Microsatellite markers
606      for the cardinal tetra Paracheirodon axelrodi, a commercially important fish from central
607      Amazonia. *Molecular Ecology Notes* 4:330–332. DOI: 10.1111/j.1471-8286.2004.00686.x.
608 Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A, Lemey P. 2016. SpreaD3:
609      Interactive Visualization of Spatiotemporal History and Trait Evolutionary Processes.
610      *Molecular Biology and Evolution* 33:2167–2169. DOI: 10.1093/molbev/msw082.
611 Bittencourt PS, Marshall BG, Hrbek T, Farias IP. 2017. Life after the drought: Temporal genetic
612      structure of Paracheirodon axelrodi Schultz, 1956 (Characiformes: Characidae) in the
613      middle Negro River. *Pan-American Journal of Aquatic Sciences* 12:184–193.
614 Caparroz R, Miyaki CY, Baker AJ. 2009. Contrasting phylogeographic patterns in mitochondrial
615      DNA and microsatellites: Evidence of female philopatry and male-biased gene flow among
616      regional populations of the blue-and-yellow Macaw (Psittaciformes: Ara ararauna) in
617      Brazil. *Auk* 126:359–370. DOI: 10.1525/auk.2009.07183.
618 Caputo MV, Soares EAA. 2016. Eustatic and tectonic change effects in the reversion of the
619      transcontinental Amazon River drainage system. *Brazilian Journal of Geology* 46:301–328.
620      DOI: 10.1590/2317-4889201620160066.

621 Chao NL. 2001. The fishery, diversity, and conservation of ornamental fishes in the rio Negro
622     Basin, Brazil - A review of Project Piaba (1989-99). In: Chao NL, Prang G, Sonneschein L,
623     Tlusty M eds. *Conservation and management of ornamental fish resources of the rio Negro*
624     *Basin, Amazonia, Brazil - Project Piaba*. Manaus: Universidade Federal do Amazonas,
625     161–201.
626 Chapuis MP, Estoup A. 2007. Microsatellite null alleles and estimation of population
627     differentiation. *Molecular Biology and Evolution* 24:621–631. DOI:
628     10.1093/molbev/msl191.
629 Chistiakov DA, Hellemans B, Volckaert FAM. 2006. Microsatellites and their genomic
630     distribution, evolution, function and applications: A review with special reference to fish
631     genetics. *Aquaculture* 255:1–29. DOI: 10.1016/j.aquaculture.2005.11.031.
632 Cooke GM, Beheregaray LB. 2007. Extremely high variability in the S72 intron of the
633     Amazonian cardinal tetra (Paracheirodon axelrodi). *Journal of Fish Biology* 71:132–140.
634     DOI: 10.1111/j.1095-8649.2007.01478.x.
635 Cooke GM, Chao NL, Beheregaray LB. 2009. Phylogeography of a flooded forest specialist fish
636     from central Amazonia based on intron DNA: the cardinal tetra Paracheirodon axelrodi.
637     *Freshwater Biology* 54:1216–1232. DOI: 10.1111/j.1365-2427.2009.02172.x.
638 Crampton WGR. 2011. An ecological perspective on diversity and distributions. In: Albert JS,
639     Reis RE eds. *Historical Biogeography of Neotropical Freshwater Fishes*. University of
640     California Press, 165–192. DOI: 10.1525/california/9780520268685.003.0010.
641 Cremon ÉH, Rossetti D de F, Sawakuchi A de O, Cohen MCL. 2016. The role of tectonics and
642     climate in the late Quaternary evolution of a northern Amazonian River. *Geomorphology*
643     271:22–39. DOI: 10.1016/J.GEOMORPH.2016.07.030.
644 DeWoody JA, Schupp J, Kenefic L, Busch J, Murfitt L, Keim P. 2004. Universal method for
645     producing ROX-labeled size standards suitable for automated genotyping. *BioTechniques*
646     37:348–352.
647 Díaz de Gamero ML. 1996. The changing course of the Orinoco River during the Neogene: a
648     review. *Palaeogeography, Palaeoclimatology, Palaeoecology* 123:385–402. DOI:
649     10.1016/0031-0182(96)00115-0.
650 Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti
651     and the BEAST 1.7. *Molecular Biology and Evolution* 29:1969–1973. DOI:
652     10.1093/molbev/mss075.
653 Dungel J. 2009. *Deep in the Jungle*. Praga: Euromedia Praga.
654 Earl DA, VonHoldt BM. 2011. STRUCTURE HARVESTER: a website and program for
655     visualizing STRUCTURE output and implementing the Evanno method. *Conservation*
656     *Genetics Resources* 4:359–361. DOI: 10.1007/s12686-011-9548-7.
657 Escalona T, Engstrom TN, Hernandez OE, Bock BC, Vogt RC, Valenzuela N. 2009. Population
658     genetics of the endangered South American freshwater turtle, Podocnemis unifilis, inferred
659     from microsatellite DNA data. *Conservation Genetics* 10:1683–1696. DOI:
660     10.1007/s10592-008-9746-3.
661 Escobar MD, Andrade-López J, Farias IP, Hrbek T. 2015. Delimiting Evolutionarily Significant
662     Units of the Fish, Piaractus brachypomus (Characiformes: Serrasalmidae), from the Orinoco
663     and Amazon River Basins with Insight on Routes of Historical Connectivity. *The Journal of*
664     *heredity* 106 Suppl:428–38. DOI: 10.1093/jhered/esv047.
665 Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the
666     software STRUCTURE: a simulation study. *Molecular ecology* 14:2611–20. DOI:

667         10.1111/j.1365-294X.2005.02553.x.

668    Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform
669         population genetics analyses under Linux and Windows. *Molecular ecology resources*
670         10:564–7. DOI: 10.1111/j.1755-0998.2010.02847.x.

671    Freeland JR. 2005. *Molecular Ecology*. Chichester, England: John Wiley & Sons, Ltd. DOI:
672         10.1002/9780470015902.a0003268.pub2.

673    Godinho AL, Lamas IR, Godinho HP. 2010. Reproductive ecology of Brazilian freshwater
674         fishes. *Environmental Biology of Fishes* 87:143–162. DOI: 10.1007/s10641-009-9574-4.

675    Gonzalez RJ, Wilson RW. 2001. Patterns of ion regulation in acidophilic fish native to the ion-
676         poor, acidic Rio Negro. *Journal of Fish Biology* 58:1680–1690. DOI:
677         10.1006/jfbi.2001.1577.

678    Gonzalez RJ, Wood CM, Wilson RW, Patrick ML, Bergman HL, Narahara A, Val AL. 1998.
679         Effects of water pH and calcium concentration on ion balance in fish of the Rio Negro,
680         Amazon. *Physiological Zoology* 71:15–22. DOI: 10.1086/515893.

681    Harris P, Petry P. 2001. Preliminary report on the genetic population structure and
682         phylogeography of cardinal tetra (Paracheirodon axelrodi) in the rio Negro Basin. In: Chao
683         N, Petry P, Prang G, Sonneschien L, Tlusty M eds. *Conservation and Management of*
684         *Ornamental Fish Resources of the Rio Negro Basin, Amazonia, Brazil - Project Piaba*.
685         Manaus, Brazil: University of Amazonas Press, 205–225.

686    Hoorn C. 1994. Fluvial palaeoenvironments in the intracratonic Amazonas Basin (early
687         Miocene-early middle Miocene, Colombia). *Palaeogeography, Palaeoclimatology,*
688         *Palaeoecology* 109:1–54.

689    Hoorn C, Guerrero J, Sarmiento GA, Lorente MA. 1995. Andean tectonics as a cause for
690         changing drainage patterns in Miocene northern South America. *Geology* 23:237–240.

691    Hubert N, Renno J-F. 2006. Historical biogeography of South American freshwater fishes.
692         *Journal of Biogeography* 33:1414–1436. DOI: 10.1111/j.1365-2699.2006.01518.x.

693    Ivanova N V., Zemlak TS, Hanner RH, Hebert PDN. 2007. Universal primer cocktails for fish
694         DNA barcoding. *Molecular Ecology Notes* 7:544–548. DOI: 10.1111/j.1471-
695         8286.2007.01748.x.

696    Jakobsson M, Rosenberg N a. 2007. CLUMPP: a cluster matching and permutation program for
697         dealing with label switching and multimodality in analysis of population structure.
698         *Bioinformatics (Oxford, England)* 23:1801–6. DOI: 10.1093/bioinformatics/btm233.

699    Jombart T. 2008. adegenet : a R package for the multivariate analysis of genetic markers.
700         *Bioinformatics* 24:1403–1405. DOI: 10.1093/bioinformatics/btn129.

701    Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A,
702         Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious
703         Basic: an integrated and extendable desktop software platform for the organization and
704         analysis of sequence data. *Bioinformatics (Oxford, England)* 28:1647–9. DOI:
705         10.1093/bioinformatics/bts199.

706    Lagostina E, Dal Grande F, Andreev M, Printzen C. 2018. The use of microsatellite markers for
707         species delimitation in Antarctic Usnea subgenus Neuropogon. *Mycologia* 110:1047–1057.
708         DOI: 10.1080/00275514.2018.1512304.

709    Latrubesse EM, Franzinelli E. 2005. The late Quaternary evolution of the Negro River, Amazon,
710         Brazil: Implications for island and floodplain formation in large anabranching tropical
711         systems. *Geomorphology* 70:372–397. DOI: 10.1016/j.geomorph.2005.02.014.

712    Lemey P, Rambaut A, Welch JJ, Suchard MA. 2010. Phylogeography Takes a Relaxed Random

713    Walk in Continuous Space and Time. *Molecular Biology and Evolution* 27:1877–1885.
714    DOI: 10.1093/molbev/msq067.
715    Li YL, Liu JX. 2018. StructureSelector: A web-based software to select and visualize the optimal
716    number of clusters using multiple methods. *Molecular Ecology Resources* 18:176–177.
717    DOI: 10.1111/1755-0998.12719.
718    Li C, Ortí G, Zhang G, Lu G. 2007. A practical approach to phylogenomics: the phylogeny of
719    ray-finned fish (Actinopterygii) as a case study. *BMC evolutionary biology* 7:44. DOI:
720    10.1186/1471-2148-7-44.
721    Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA
722    polymorphism data. *Bioinformatics (Oxford, England)* 25:1451–2. DOI:
723    10.1093/bioinformatics/btp187.
724    Lovejoy NR, Araujo LG. 2000. Molecular systematics, biogeography and population structure of
725    Neotropical freshwater needlefishes of the genus Potamorrhaphis. *Molecular Ecology*
726    9:259–268.
727    Lundberg JG, Marshall LG, Guerrero J, Horton B, Malabarba LMCS, Wesselingh. F. 1998. The
728    stage for Neotropical fish diversification: A history of tropical South American rivers. In:
729    Malabarba LR, Reis RE, Vari RP, Lucena ZMS, Lucena CAS eds. *Phylogeny and
730    Classification of Neotropical Fishes*. Porto Alegre: Porto Alegre Edipucrs, 13–48.
731    Martin AP, Palumbi SR. 1993. Body size, metabolic rate, generation time, and the molecular
732    clock. *Proceedings of the National Academy of Sciences* 90:4087–4091. DOI:
733    10.1073/pnas.90.9.4087.
734    Matsuo AYO, Val AL. 2007. Acclimation to humic substances prevents whole body sodium loss
735    and stimulates branchial calcium uptake capacity in cardinal tetras Paracheirodon axelrodi
736    (Schultz) subjected to extremely low pH. *Journal of Fish Biology* 70:989–1000. DOI:
737    10.1111/j.1095-8649.2007.01358.x.
738    Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference
739    of large phylogenetic trees. In: *2010 Gateway Computing Environments Workshop, GCE
740    2010*. 1–7. DOI: 10.1109/GCE.2010.5676129.
741    Mora A, Baby P, Roddaz M, Parra M, Brusset S, Hermoza W, Espurt N. 2011. Tectonic History
742    of the Andes and Sub-Andean Zones: Implications for the Development of the Amazon
743    Drainage Basin. In: Hoorn C, Wesselingh FP eds. *Amazonia: Landscape and Species
744    Evolution: A look into the past*. Oxford, UK: Wiley-Blackwell Publishing Ltd., 38–60. DOI:
745    10.1002/9781444306408.
746    Muniz FL, Campos Z, Hernández Rangel SM, Martínez JG, Souza BC, de Thoisy B, Botero-
747    Arias R, Hrbek T, Farias IP. 2017. Delimitation of evolutionary units in Cuvier's dwarf
748    caiman, Paleosuchus palpebrosus (Cuvier, 1807): insights from conservation of a broadly
749    distributed species. *Conservation Genetics*. DOI: 10.1007/s10592-017-1035-6.
750    Olivares AM, Hrbek T, Escobar MD, Caballero S. 2013. Population structure of the black
751    arowana (Osteoglossum ferreirai) in Brazil and Colombia: Implications for its management.
752    *Conservation Genetics* 14:695–703. DOI: 10.1007/s10592-013-0463-1.
753    De Oliveira SR, De Souza RTYB, Nunes ÉDSS, De Carvalho CSM, De Menezes GC, Marcon
754    JL, Roubach R, Akifumi Ono E, Affonso EG. 2008. Tolerance to temperature, pH,
755    ammonia and nitrite in cardinal tetra, Paracheirodon axelrodi, an amazonian ornamental
756    fish. *Acta Amazonica* 38:773–779. DOI: 10.1590/S0044-59672008000400023.
757    Posada D. 2008. jModelTest: phylogenetic model averaging. *Molecular biology and evolution*
758    25:1253–6. DOI: 10.1093/molbev/msn083.

759 Pritchard JK, Stephens M, Donnelly P. 2000. Inference of Population Structure Using Multilocus
760      Genotype Data. *Genetics* 155:945–959.
761 Puechmaille SJ. 2016. The program structure does not reliably recover the correct population
762      structure when sampling is uneven: Subsampling and new estimators alleviate the problem.
763      *Molecular Ecology Resources* 16:608–627. DOI: 10.1111/1755-0998.12512.
764 Reis RE, Albert JS, Di Dario F, Mincarone MM, Petry P, Rocha LA. 2016. Fish biodiversity and
765      conservation in South America. *Journal of fish biology* 89:12–47. DOI: 10.1111/jfb.13016.
766 Reis RE, Kullander SO, Ferraris CJ. 2003. *Check List of the Freshwater Fishes of South and*
767      *Central America*. Porto Alegre, Brasil: Edipucrs.
768 Rice AH. 1921. The Rio Negro, the Casiquiare Canal, and the Upper Orinoco, September 1919-
769      April 1920. *The Geographical Journal* 58:321–343. DOI: 10.2307/1780880.
770 Rice WR. 1989. Analyzing tables of statistical tests. *Evolution* 43:223–225.
771 Rosenberg N a. 2003. Distruct: a Program for the Graphical Display of Population Structure.
772      *Molecular Ecology Notes* 4:137–138. DOI: 10.1046/j.1471-8286.2003.00566.x.
773 Ruokolainen K, Moulatlet GM, Zuquim G, Hoorn C, Tuomisto H. 2018. River network
774      rearrangements in Amazonia shake Biogeography and civil security. *Preprints*
775      2018090168:1–16. DOI: 10.20944/preprints201809.0168.v1.
776 Salzburger W, Ewing GB, Von Haeseler A. 2011. The performance of phylogenetic algorithms
777      in estimating haplotype genealogies with migration. *Molecular ecology* 20:1952–63. DOI:
778      10.1111/j.1365-294X.2011.05066.x.
779 Sambrook J, Fritsch E, Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual*. New York,
780      NY: Cold Springs Harbor Laboratory Press, Cold Springs Harbor.
781 Schaefer C, Vale Júnior J. 1997. Mudanças climáticas e evolução da paisagem em Roraima: uma
782      resenha do Cretáceo ao recente. In: Barbosa R, Ferreira E, Castellón E eds. *Homem,*
783      *Ambiente e Ecologia no Estado de Roraima*.
784 Schneider CH, Gross MC, Terencio ML, Porto JIR. 2012. Cryptic diversity in the mtDNA of the
785      ornamental fish Carnegiella strigata. *Journal of fish biology* 81:1210–24. DOI:
786      10.1111/j.1095-8649.2012.03403.x.
787 Sioli H. 1984. The Amazon and its main affluents: Hidrography, morphology of the rivers
788      courses, and river type. In: Sioli H ed. *The Amazon: Limnology and Landscape Ecology of a*
789      *Mighty Tropical River and Its Basin*. Monographiae Biologicae. Dordrecht: Springer
790      Netherlands, 127–165. DOI: 10.1007/978-94-009-6542-3.
791 Stallard R. 1985. River Chemistry, Geology, Geomorphology, and Soils in the Amazon and
792      Orinoco Basins. In: Drever JI ed. *The Chemistry of Weathering*. New Jersey: Reidel
793      Publishing Company, 293–316.
794 Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of
795      large phylogenies. *Bioinformatics* 30:1312–1313. DOI: 10.1093/bioinformatics/btu033.
796 Stephens M, Smith NJ, Donnelly P. 2001. A New Statistical Method for Haplotype
797      Reconstruction from Population Data. *American Journal of Human Genetics* 68:978–989.
798 Terencio ML, Schneider CH, Porto JIR. 2012. Molecular signature of the D-loop in the brown
799      pencilfish Nannostomus eques (Characiformes, Lebiasinidae) reveals at least two
800      evolutionary units in the Rio Negro basin, Brazil. *Journal of fish biology* 81:110–24. DOI:
801      10.1111/j.1095-8649.2012.03320.x.
802 Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of
803      progressive multiple sequence alignment through sequence weighting, position-specific gap
804      penalties and weight matrix choice. *Nucleic Acids Research* 22:4673–4680. DOI:

805    10.1007/978-1-4020-6754-9_3188.
806 Tovar MO, Obando MJ, Gómez E, Caldas ML, Hurtado H. 2009. Histología y morfometría del
807    ojo del pez dulceacuícola paracheirodon axelrodi (characiformes: Characidae). *Revista de*
808    *Biologia Tropical* 57:1107–1118. DOI: 10.15517/rbt.v57i4.5450.
809 Turner TF, McPhee M V., Campbell P, Winemiller KO. 2004. Phylogeography and intraspecific
810    genetic variation of prochilodontid fishes endemic to rivers of northern South America.
811    *Journal of Fish Biology* 64:186–201.
812 Villesen P. 2007. FaBox: An online toolbox for FASTA sequences. *Molecular Ecology Notes*
813    7:965–968. DOI: 10.1111/j.1471-8286.2007.01821.x.
814 Werle E, Schneider C, Renner M, Volker M, Fiehn W. 1994. Convenient single-step, one tube
815    purification of PCR products for direct sequencing. *Nulceic Acid Research* 22:4354–4355.
816 Willis SC, Nunes M, Montaña CG, Farias IP, Ortí G, Lovejoy NR. 2010. The Casiquiare river
817    acts as a corridor between the Amazonas and Orinoco river basins: biogeographic analysis
818    of the genus Cichla. *Molecular ecology* 19:1014–30. DOI: 10.1111/j.1365-
819    294X.2010.04540.x.
820 Winemiller KO, López-Fernández H, Taphorn DC, Nico LG, Duque AB. 2008. Fish
821    assemblages of the Casiquiare River, a corridor and zoogeographical filter for dispersal
822    between the Orinoco and Amazon basins. *Journal of Biogeography* 35:1551–1563. DOI:
823    10.1111/j.1365-2699.2008.01917.x.
824 Winemiller KO, Willis SC. 2011. The Vaupes Arch and Casiquiare Canal: Barriers and
825    Passages. In: Albert JS, Reis RE eds. *Historical Biogeography of Neotropical Freshwater*
826    *fishes*. Los Angeles: University of California Press, 225–242.
827 Wood CM, Robertson LM, Johannsson OE, Val AL. 2014. Mechanisms of Na+ uptake,
828    ammonia excretion, and their potential linkage in native Rio Negro tetras (Paracheirodon
829    axelrodi, Hemigrammus rhodostomus, and Moenkhausia diktyota). *Journal of Comparative*
830    *Physiology B: Biochemical, Systemic, and Environmental Physiology* 184:877–890. DOI:
831    10.1007/s00360-014-0847-7.

832

833 **Field Study Permissions**

834

835 In Brazil, permission to collect tissue samples was granted by SISBIO/IBAMA (11325-
836 1/662933 - 1/10/2007), and in Colombia by ANLA (to Universidad de los Andes).

837

838 **Data accessibility**

839

840 DNA sequences: Genbank accessions numbers: (COI) MG384556 - MG384575; (MYH6)
841 MG384576 – MG384595. Microsatellite genotypes used in this study and Supplementary
842 material 1-3, have been uploaded to Dryad: DOI: https://doi.org/10.5061/dryad.7kd80

843

844 **Funding**

845

846 This study was supported by the Departamento Administrativo de Ciencia, Tecnología e
847 Innovación (COLCIENCIAS) in a cooperation with the Conselho Nacional de Desenvolvimento
848 Científico e Tecnologico (CNPq) of Brazil (code: 490682/2010-3); Laboratorio de Evolução
849 e Genética Animal (LEGAL) from Universidade Federal da Amazonas UFAM and, by a
850 Colombia Biodiversa grant from Fundación Alejandro Ángel Escobar and by the Proyecto
851 semilla from Faculty of Science from the Universidad de los Andes of Colombia.

852
853 **Author contributions**
854

855 All authors collaborated to design the study. DSB and JGM collected field data and performed
856 laboratory work, DSB, JGM and TH performed genetic analysis. IPF, TH and SC secured
857 funding and obtained permits. TH and SC oversaw laboratory and genetic data analysis and
858 edited the manuscript. DSB wrote the manuscript with the contribution of JGM, IPF, TH and SC.

859

860 **Figure captions**
861

862 **Figure 1:** Map showing the hydro-geological scenario of study and sampling localities of
863 *Paracheirodon axelrodi.* Negro basin (Amazon) locations are in purple. Orinoco basin Locations
864 are in green. Locations are (a) Puerto Carreño (*n*=20), (b) Puerto Gaitán (*n*=21), (c) San José del
865 Guaviare (*n*=20), (d) Puerto Inirida (*n*=21), (e) Cucui (*n*=21), (f) São Gabriel da Cachoeira
866 (n=20),(g) Santa Isabel (*n*=20), (h) Barcelos (*n*=20). Points A, B and C represent the three more
867 famous hypotheticals connectivity routes between Orinoco and Amazon basins for the
868 ichthyofauna (Lovejoy & Araujo, 2000; Winemiller & Willis, 2011; Escobar et al., 2015).
869 Alternative routes suggested in this study for the Cardinal tetra are represented in blue.
870

871 **Figure 2:** Mitochondrial *Cytochrome oxidase I* (COI) haplotype network obtained from
872 RAxML. The size circles reflect frequency of a particular haplotype in Orinoco and Negro
873 basins. Inside every circle is shown the number of individuals sharing one haplotype. The
874 parallel bars indicate a number of mutational steps.
875

876 **Figure 3:** Nuclear *myosin heavy polypeptide 6 cardiac muscle alpha* (MYH6) gene alleles
877 network obtained from RaxML. The size circles reflect frequency of a particular haplotype/allele
878 in Orinoco and Negro basins. Inside every circle is shown the number of individuals sharing one
879 haplotype. The parallel bars indicate a number of mutational steps.
880

881 **Figure 4:** Phylogeographical reconstruction analysis using non-concatenated dataset of the
882 mitochondrial COI and the nuclear MYH6 gene, based on a relaxed random walk in continuous
883 space and time method for *Paracheirodon axelrodi* from Orinoco and Negro basin, inferred in
884 BEAST v.1.8.2 and SPREAD3 and visualized in GoogleMaps. Graph is showing the
885 hypothetical spatial patterns and times of colonization for *P. axelrodi* in each basin, locating the
886 ancestral population from all of them was originated. All the observed colonization times are
887 median heights into 95% credibility intervals, which can be consulted in the supporting
888 information (Fig. S1).
889

890 **Figure 5:** Bayesian assignment for each individual of *Paracheirodon axelrodi* inferred by the
891 software Structure from K=2 to K=6, using eight microsatellite loci. Vertical lines indicate each
892 of the individuals and the colors represent the probability membership coefficient of that
893 individual to each genetic cluster. GV: San José del Guaviare; CA: Puerto Carreño; PG: Puerto
894 Gaitán; IN: Inírida; CUC: Cucuí; SGC: São Gabriel da Cachoeira; SI: Santa Isabel; BAR:
895 Barcelos.
896

897 **Figure 6:** Patterns of genetic structure by discriminant analysis of principal components (DAPC)
898 *(Adegenet package in R),* using eight microsatellite loci for samples of *Paracheirodon axelrodi*

899    from (a): Puerto Carreño; (b): Puerto Gaitán; (c): San José del Guaviare; (d): Inírida; (e): Cucuí;
900    (f): São Gabriel da Cachoeira; (g): Santa Isabel; (h): Barcelos.
901
902

## Supplemental Information

904

905    **Table S1**: Sample information of *Paracheirodon axelrodi* analyzed in the study.

906

907    **Table S2**: Genetic diversity for eight microsatellite loci in all *Paracheirodon axelrodi*
908    geographic locations analyzed.  N= Cardinal sample size; for each locus: n= total number of
909    alleles, Ho= observed heterozygosity, HE= expected heterozygosity. Loci out of equilibrium
910    after Bonferroni correction (0.00078125) are shown in bold. CAR: Puerto Carreño; GV: San José
911    del Guaviare; IN: Puerto Inírida; PG: Puerto Gaitán; CUC: Cucuí; SI: Santa Isabel; BAR:
912    Barcelos.

913

914    **Table S3**: Pairwise FST values using eight microsatellites loci among Orinoco and Negro basin
915    populations for *Paracheirodon axelrodi*. Significantly different values (P < 0.05) are shown in
916    bold. CAR: Puerto Carreño; GV: San José del Guaviare; IN: Puerto Inírida; PG: Puerto Gaitan;
917    CUC: Cucui; SGC: São Gabriel da Cachoeira; SI: Santa Isabel; BAR: Barcelos.

918

919    **Table S4**: Pairwise $\Phi st$ values for the mitochondrial COI marker, among Orinoco and Negro
920    basin populations for *Paracheirodon axelrodi*. Significantly different values (P<0.05) are shown
921    in bold. Haplotype (h) and nucleotide ($\pi$) diversity ± standard deviation (SD) are shown on the
922    diagonal for each geographic location. CAR: Puerto Carreño; GV: San José del Guaviare; IN:
923    Puerto Inírida; PG: Puerto Gaitán; CUC: Cucuí; SI: Santa Isabel; BAR: Barcelos.

924

925    **Fig. S1**: Maximum likelihood tree of Mitochondrial Cytochrome oxidase I (COI) partial region
926    (468 bp) of *Paracheirodon axelrodi* from Orinoco and Negro River (Amazon basin), using 1,000
927    replicates of RAxML's rapid bootstrap algorithm to account for uncertainty in the estimation of
928    the topology.

929

930    **Fig. S2**: Maximum likelihood tree of *myosin heavy polypeptide 6 cardiac muscle alpha* (MYH6)
931    partial region (555 bp) of *Paracheirodon axelrodi* from Orinoco and Negro River (Amazon
932    basin), using 1,000 replicates of RAxML's rapid bootstrap algorithm to account for uncertainty
933    in the estimation of the topology.

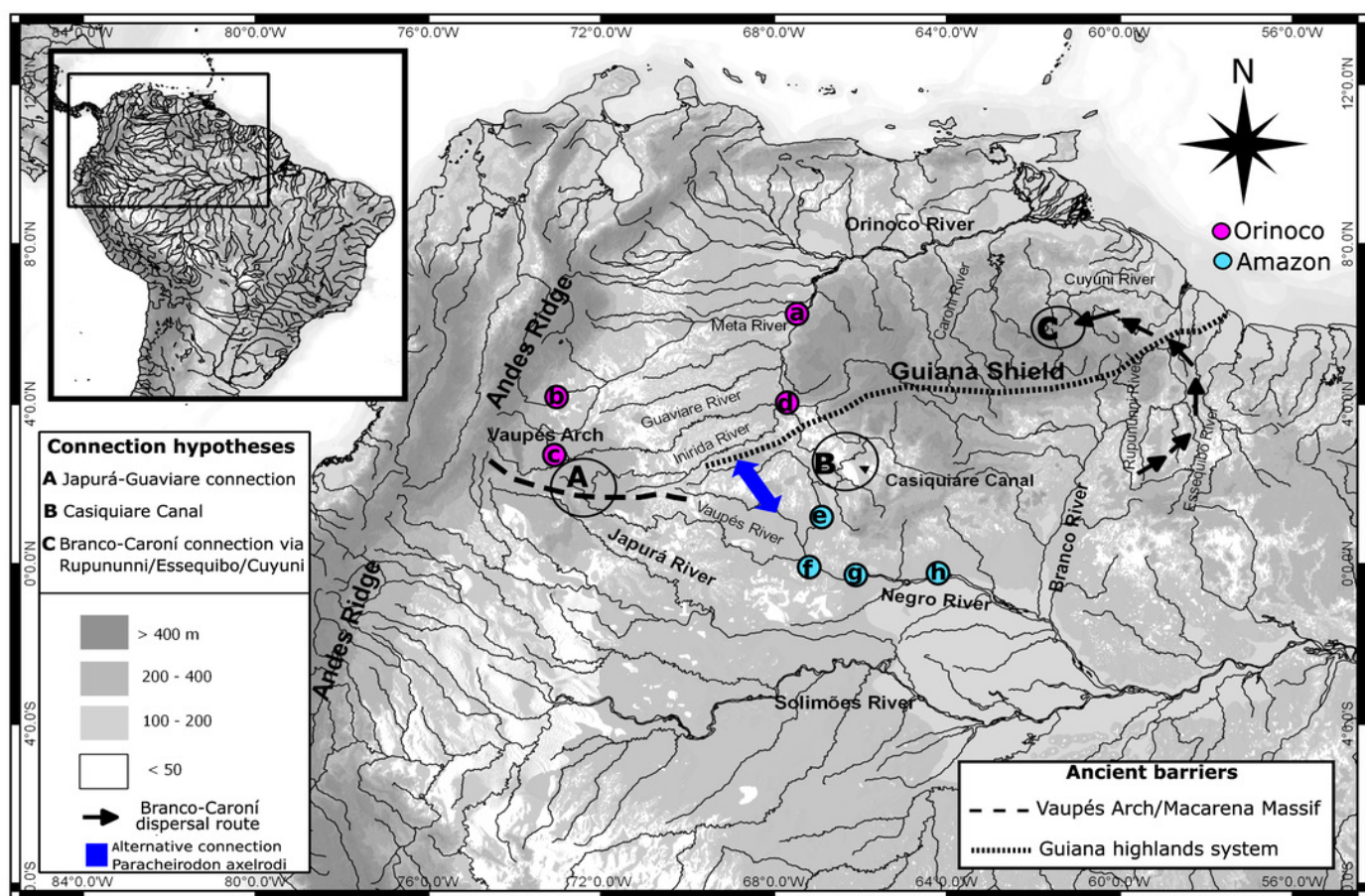
934

935    **Fig. S3**: Bayesian time-calibrated Maximum Credibility Tree, showing the phylogenetic
936    reconstruction and ancestral relationship of *Paracheirodon axelrodi*, obtained from the
937    spatiotemporal analysis by the Relaxed Random Walk method in BEAST v.1.8.2. Samples from
938    Orinoco River basin [(a): Puerto Carreño; (b): Puerto Gaitán; (c): San José del Guaviare; (d):
939    Inírida] and Negro River basin [(e): Cucuí; (f): São Gabriel da Cachoeira; (g): Santa Isabel; (h):
940    Barcelos], using COI and MYH6 exonic region, estimated under a HKY+GAMMA and
941    GTR+GAMMA models of molecular evolution, respectively. The observed times are median
942    heights into 95% credibility intervals, calibrated with the widely accepted mtDNA substitution
943    rate for poikilotherm vertebrates in Martin & Palumbi (1993) of 0.68x10-8 mutations per site per

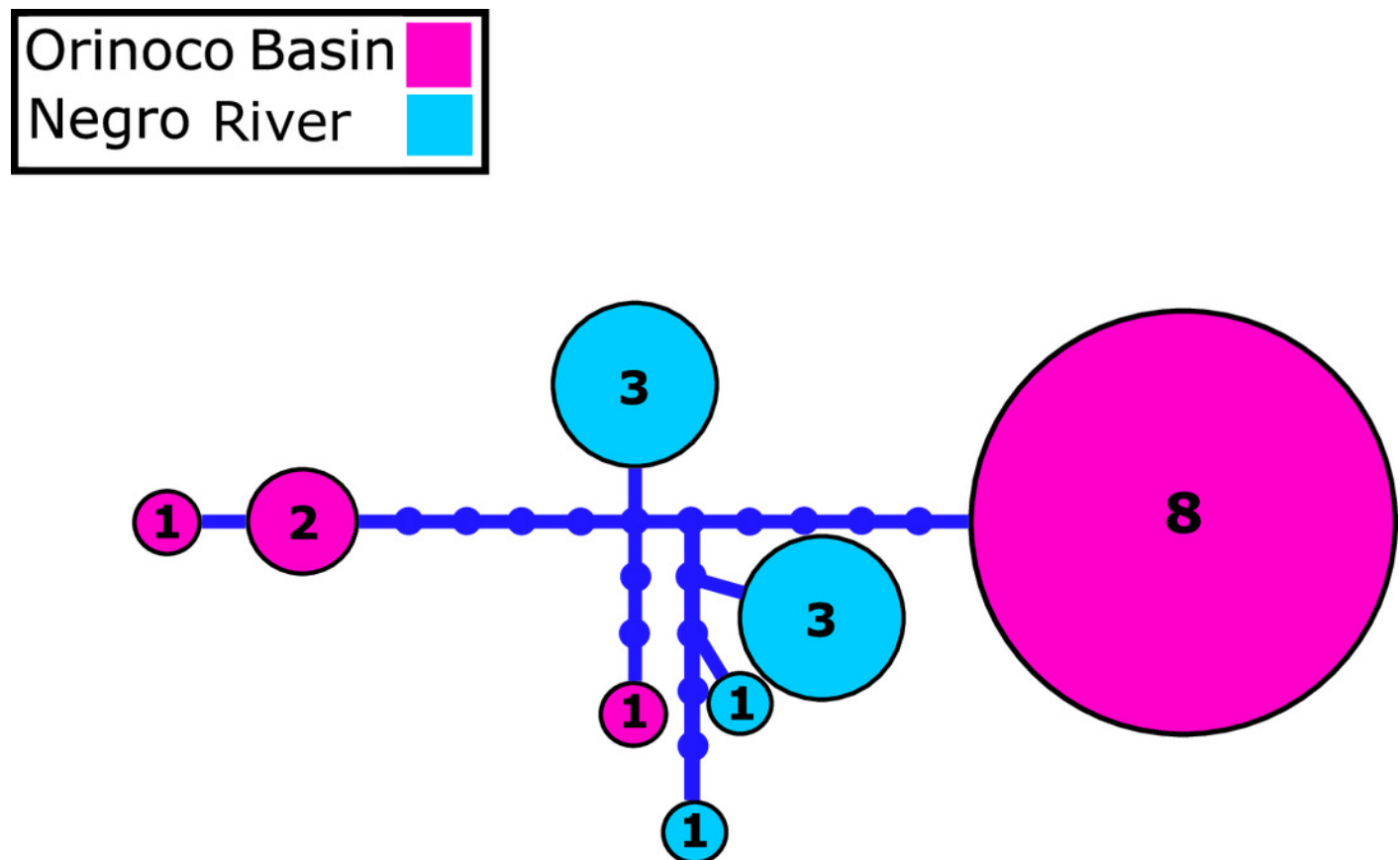
944    year for COI, and 1.0x10-9 for MYH6 (Freeland, 2005). Circles indicate nodes with supports
945    ≥0.95 of posterior probability.
946
947    **Fig. S4**: Visualization of the optimal number of clusters using multiple methods in the web-
948    based software StructureSelector, after testing K=1-8 in STRUCTURE, for 163 samples of
949    *Paracheirodon axelrodi* from eight populations from Orinoco and Negro River (Amazon basin).
950    (A) Evanno method (ΔK), identifying K=2 as the most likely number of clusters. (C) Puechmille
951    method estimators (MaxMed and MaxMean) showing K=6 as the most likely number of clusters.
952
953
954

# Figure 1

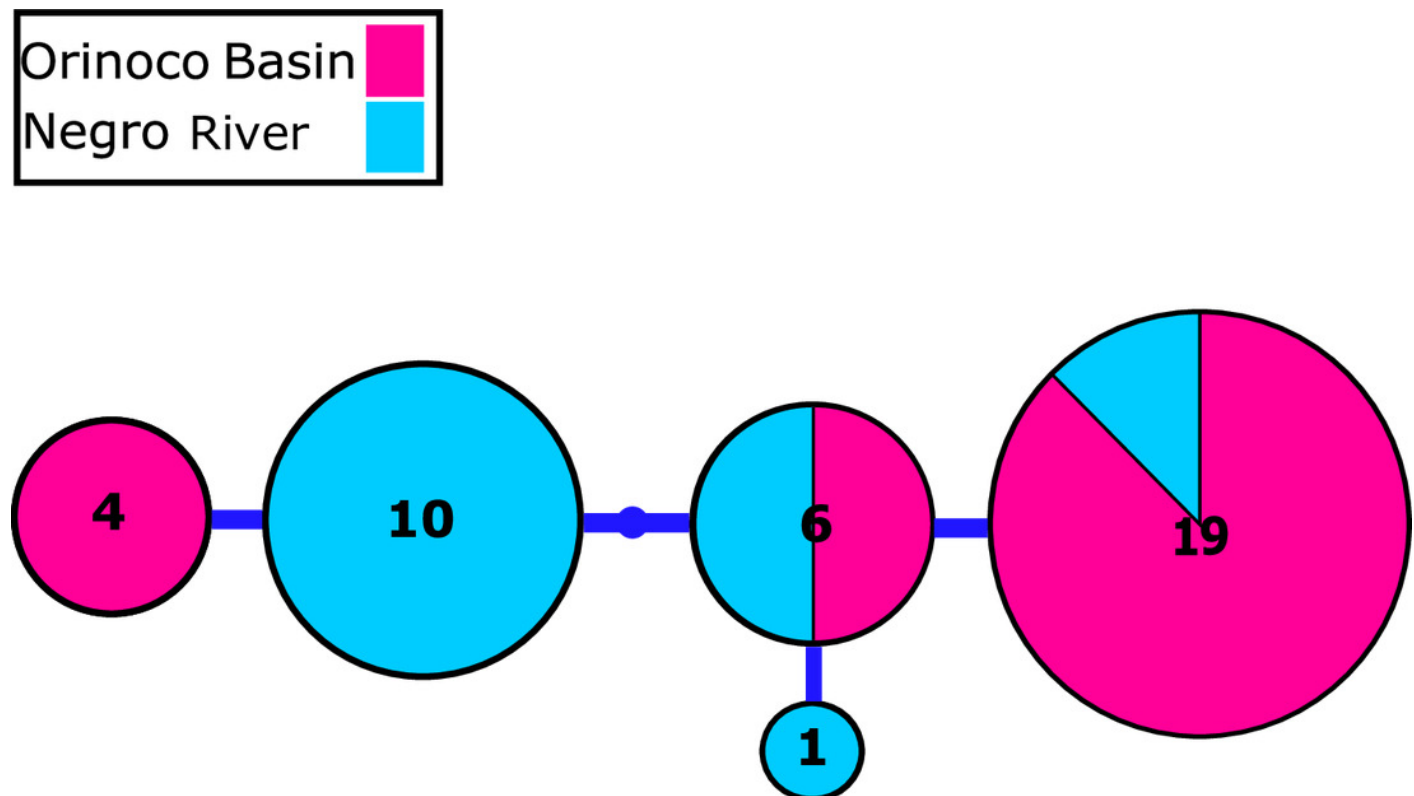
Map showing hydro-geological scenario and sampling areas

Map showing the hydro-geological scenario of study and sampling localities of *Paracheirodon axelrodi.* Negro basin (Amazon) locations are in purple. Orinoco basin Locations are in green. Locations are (a) Puerto Carreño (*n*=20), (b) Puerto Gaitán (*n*=21), (c) San José del Guaviare (*n*=20), (d) Puerto Inirida (*n*=21), (e) Cucui (*n*=21), (f) São Gabriel da Cachoeira (n=20),(g) Santa Isabel (*n*=20), (h) Barcelos (*n*=20). Points A, B and C represent the three more famous hypotheticals connectivity routes between Orinoco and Amazon basins for the ichthyofauna (Lovejoy & Araujo, 2000; Winemiller & Willis, 2011; Escobar et al., 2015). Alternative routes suggested in this study for the Cardinal tetra are represented in blue.

# Figure 2

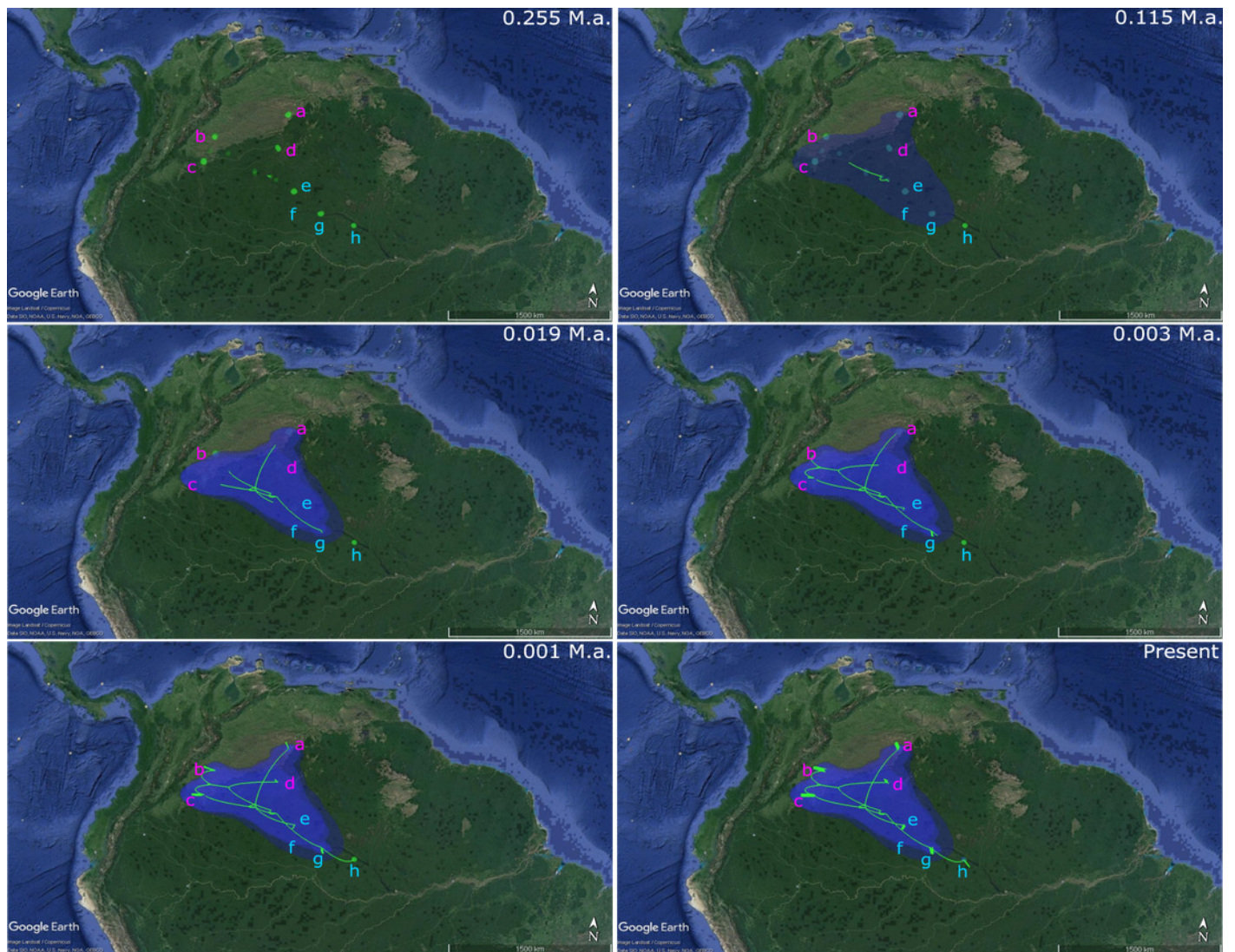
Mitochondrial COI haplotype network

Mitochondrial *Cytochrome oxidase I* (COI) haplotype network obtained from RAxML. The size circles reflect frequency of a particular haplotype in Orinoco and Negro basins. Inside every circle is shown the number of individuals sharing one haplotype. The parallel bars indicate a number of mutational steps.

# Figure 3

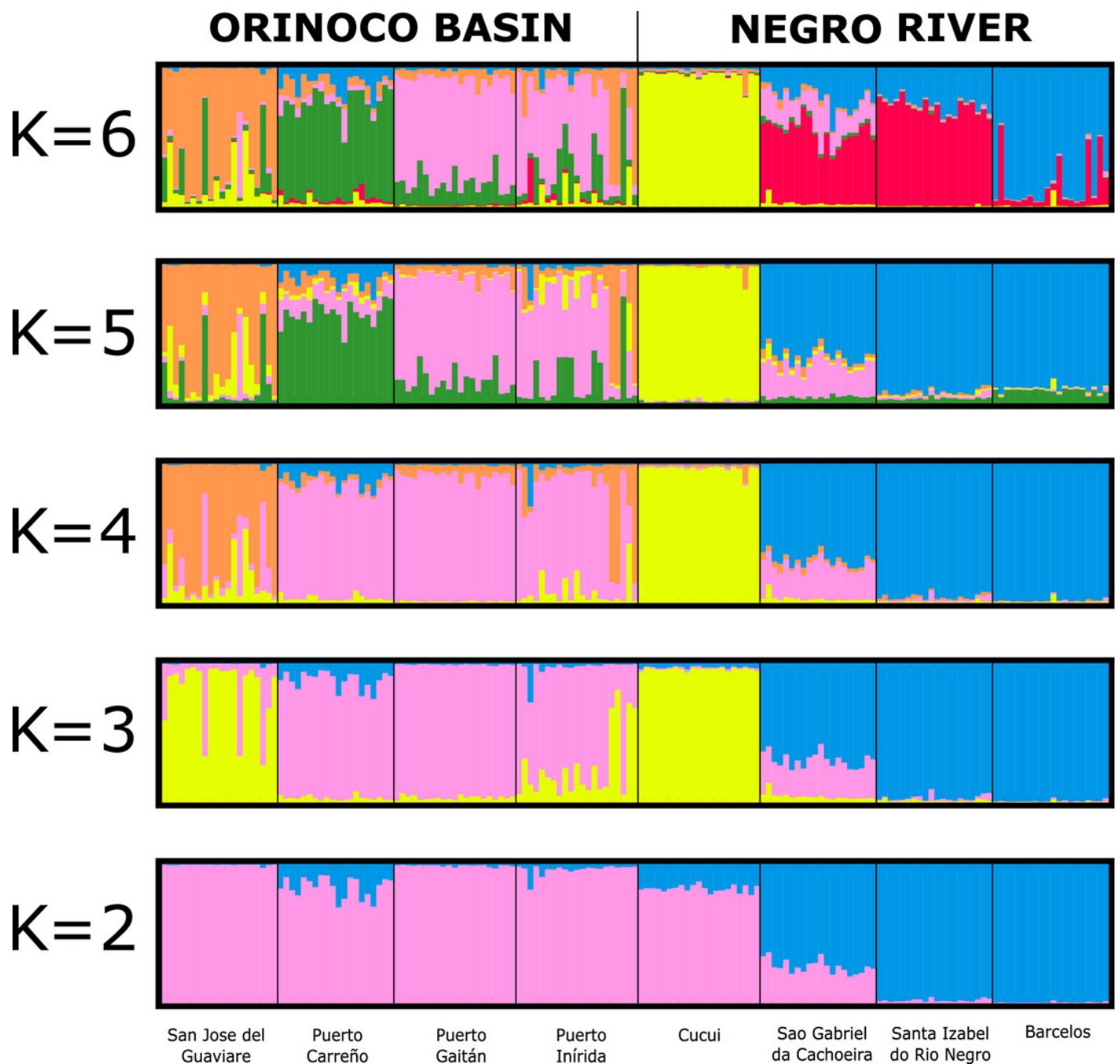
Nuclear MYH6 gene allele network

Nuclear *myosin heavy polypeptide 6 cardiac muscle alpha* (MYH6) gene alleles network obtained from RaxML. The size circles reflect frequency of a particular haplotype/allele in Orinoco and Negro basins. Inside every circle is shown the number of individuals sharing one haplotype. The parallel bars indicate a number of mutational steps.

# Figure 4

Phylogeographic reconstruction

Phylogeographical reconstruction analysis using non-concatenated dataset of the mitochondrial COI and the nuclear MYH6 gene, based on a relaxed random walk in continuous space and time method for *Paracheirodon axelrodi* from Orinoco and Negro basin, inferred in BEAST v.1.8.2 and SPREAD3 and visualized in GoogleMaps. Graph is showing the hypothetical spatial patterns and times of colonization for *P. axelrodi* in each basin, locating the ancestral population from all of them was originated. All the observed colonization times are median heights into 95% credibility intervals, which can be consulted in the supporting information (Fig. S1).

# Figure 5

Bayesian assignment

Bayesian assignment for each individual of *Paracheirodon axelrodi* inferred by the software Structure from K=2 to K=6, using eight microsatellite loci. Vertical lines indicate each of the individuals and the colors represent the probability membership coefficient of that individual to each genetic cluster. GV: San José del Guaviare; CA: Puerto Carreño; PG: Puerto Gaitán; IN: Inírida; CUC: Cucuí; SGC: São Gabriel da Cachoeira; SI: Santa Isabel; BAR: Barcelos.

# Figure 6

Patterns of genetic structure by DAPC

Patterns of genetic structure by discriminant analysis of principal components (DAPC) *(Adegenet package in R),* using eight microsatellite loci for samples of *Paracheirodon axelrodi* from (a): Puerto Carreño; (b): Puerto Gaitán; (c): San José del Guaviare; (d): Inírida; (e): Cucuí; (f): São Gabriel da Cachoeira; (g): Santa Isabel; (h): Barcelos.