Population genetics of the freshwater fish *Prochilodus magdalenae* (Characiformes: Prochilodontidae), using species-specific microsatellite loci (#33631)

First revision

Guidance from your Editor

Please submit by **20 Jul 2020** for the benefit of the authors .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Custom checks
Make sure you include the custom checks shown below, in your review.

Author notes
Have you read the author notes on the guidance page?

Raw data check
Review the raw data. Download from the materials page.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the materials page.

1 Tracked changes manuscript(s)
1 Rebuttal letter(s)
4 Figure file(s)
10 Table file(s)
4 Other file(s)

! Custom checks

DNA data checks

! Have you checked the authors data deposition statement?

! Can you access the deposited data?

! Has the data been deposited correctly?

! Is the deposition information noted in the manuscript?

Field study

! Have you checked the authors field study permits?

Are the field study permits appropriate?

For assistance email peer.review@peerj.com

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

📄 You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- *i* Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.

- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	Example
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Population genetics of the freshwater fish *Prochilodus magdalenae* (Characiformes: Prochilodontidae), using species-specific microsatellite loci

Ricardo M Landínez-García [1], **Juan Carlos Narvaez** [2], **Edna J. Márquez** [Corresp. 1]

[1] Facultad de Ciencias, Escuela de Biociencias, Laboratorio de Biología Molecular y Celular, Universidad Nacional de Colombia - Sede Medellín, Medellín, Colombia

[2] Grupo de Biodiversidad y Ecología Aplicada, Laboratorio de Genética Molecular, Universidad del Magdalena, Santa Marta, Magdalena, Colombia

Corresponding Author: Edna J. Márquez
Email address: ejmarque@unal.edu.co

Prochilodus magdalenae is a freshwater fish endemic to the Colombian hydrographic Magdalena-Cauca and Caribe basins. The genetic structure patterns of populations of different members of *Prochilodus* and the historic restocking of its depleted natural stocks suggest that *P. magdalenae* exhibits genetic stocks that coexist and co-migrate throughout the rivers Magdalena, Cauca, Cesar, Sinú, and Atrato. To test this hypothesis and explore the levels of genetic diversity and population demography of 725 samples from the studied rivers, we developed a set of 11 species-specific microsatellite loci using next-generation sequencing, bioinformatics, and experimental tests of the levels of diversity of the microsatellite loci. The results evidenced that *P. magdalenae* exhibits high genetic diversity, significant inbreeding coefficient ranging from 0.162 to 0.202, and signs of erosion of the genetic pool. Additionally, the population genetic structure constitutes a mixture of genetic stocks heterogeneously distributed along the studied rivers, and moreover, a highly divergent genetic stock was detected in Chucurí, Puerto Berrío, and Palagua that may result from restocking practices. This study provides molecular tools and a wide framework regarding the genetic diversity and structure of *P. magdalenae,* which is crucial to complement its baseline information, diagnosis and monitoring of populations, and to support the implementation of adequate regulation, management, and conservation policies.

1　Population genetics of the freshwater fish *Prochilodus magdalenae* (Characiformes: Prochilodontidae),

2　using species-specific microsatellite loci

3

4　Ricardo M. Landínez-García,[1] Juan C. Narvaez[2] and Edna J. Márquez[1]

5

6　[1] Facultad de Ciencias, Escuela de Biociencias, Laboratorio de Biología Molecular y Celular, Universidad

7　Nacional de Colombia - Sede Medellín, Medellín, Colombia.

8　[2] Grupo de Biodiversidad y Ecología Aplicada, Laboratorio de Genética Molecular, Universidad del

9　Magdalena, Santa Marta, Magdalena, Colombia.

10

11

12　Author correspondence: Edna J. Marquez, Cra. 65 No. 59A-110, Bloque 19 A Laboratorio 310, Medellín,

13　050034 Colombia; ejmarque@unal.edu.co; ejmarque@gmail.com.

14

16 **Abstract**

17 *Prochilodus magdalenae* is a freshwater fish endemic to the Colombian hydrographic Magdalena-Cauca

18 and Caribe basins. The genetic structure patterns of populations of different members of *Prochilodus* and

19 the historic restocking of its depleted natural stocks suggest that *P. magdalenae* exhibits genetic stocks

20 that coexist and co-migrate throughout the rivers Magdalena, Cauca, Cesar, Sinú, and Atrato. To test this

21 hypothesis and explore the levels of genetic diversity and population demography of 725 samples from

22 the studied rivers, we developed a set of 11 species-specific microsatellite loci using next-generation

23 sequencing, bioinformatics, and experimental tests of the levels of diversity of the microsatellite loci. The

24 results evidenced that *P. magdalenae* exhibits high genetic diversity, significant inbreeding coefficient

25 ranging from 0.162 to 0.202, and signs of erosion of the genetic pool. Additionally, the population genetic

26 structure constitutes a mixture of genetic stocks heterogeneously distributed along the studied rivers, and

27 moreover, a highly divergent genetic stock was detected in Chucurí, Puerto Berrío, and Palagua that may

28 result from restocking practices. This study provides molecular tools and a wide framework regarding the

29 genetic diversity and structure of *P. magdalenae,* which is crucial to complement its baseline information,

30 diagnosis and monitoring of populations, and to support the implementation of adequate regulation,

31 management, and conservation policies.

32

33

34

35

36 **INTRODUCTION**

37

38 The family Prochilodontidae (Teleostei: Characiformes) comprises the genera *Prochilodus,*

39 *Semaprochilodus*, and *Ichthyoelephas*, and encompasses 21 Neotropical freshwater fish species in the

40 main river basins of South America (Castro & Vari, 2004). Most of the prochilodontids exhibit large body

41 sizes, high fecundities, and abundances, representing around 50–80% of the biomass caught by the

42 subsistence and commercial fisheries in some regions of their distribution area (Barroca et al., 2012b;

43 Melo et al., 2016a). Furthermore, some members of Prochilodontidae constitute a potential resource for

44 fish farming due to certain characteristics such as their fast growth and weight increase, rustic

45 management, and high economic value (Flores-Nava & Brown, 2010; DellaRosa et al., 2014; Roux et al.,

46 2015).

47 In addition to the economic importance, Prochilodontidae plays an important trophic role in aquatic

48 ecosystems. These detritivorous and migratory fishes contribute to the nutrient cycling, distribution,

49 equilibrium, and maintenance of energetic flows and support a wide trophic network for a great number

50 of predators (Flecker, 1996). Hence, the adequate management of fisheries is crucial for the maintenance

51 of high productivity and permanent resource availability, as well as to guarantee the stability and

52 continuity of the aquatic ecosystems (Taylor, Flecker, & Hall, 2006; Batista & Lima, 2010).

53 The bocachico *Prochilodus magdalenae* Steindachner 1878 is the most representative endemic species of

54 the Colombian ichthyofauna, considered the emblematic fishery resource of the Magdalena-Cauca Basin,

55 with an estimated unload for the Magdalena Basin of 2,182.67 metric tons in 2013 (Colombian fishing

56 statistical service: SEPEC). However, between 1978 and 2012, this species experienced drastic decreases

57 in its population densities, catches (approx. 85%), and mean catch sizes. These effects resulted from

58 overfishing during migratory periods, violations of legislation related to mean catch sizes, and habitat

59 disturbances including deforesting, floodplain lake desiccations, agrochemical or chemical contamination

60   derived from farming and mining activities, sedimentation, and dam/hydropower construction, among

61   others (Cortes Millan, 2003; Lasso et al., 2010; Mojica et al., 2012).

62   To counteract this detrimental situation, several state regulations were implemented for the management

63   and conservation of *P. magdalenae* (Usma et al., 2009; Lasso et al., 2010; Mojica et al., 2012).

64   Specifically, this fish resource was catalogued as critically endangered in 2002 and as vulnerable since

65   2012 in the Colombian Red List of freshwater fishes (Mojica et al., 2012). Additionally, national

66   regulations of territorial entities and autonomous corporations focused their efforts on the restocking of

67   natural stocks threatened in the last 20 years (INPA regulation 531-1995; ANLA, INCODER, AUNAP

68   regulation 2838-2017). However, these last-mentioned activities are not based on knowledge of the

69   population genetics of *P. magdalenae* and their ecological, genetic, and sanitary impacts are unknown due

70   to the lack of programmatic monitoring and regulation of fish farming (Povh et al., 2008; FAO, 2011).

71   Moreover, population genetic studies of *P. magdalenae* are recent, scarce, and fragmented (López-Macías

72   et al., 2009; Aguirre-Pabón, Narváez-Barandica, & Castro García, 2013; see Mancera-Rodríguez,

73   Márquez, & Hurtado-Alarcón, 2013; Orozco-Berdugo & Narváez-Barandica, 2014; Hernández, Navarro,

74   & Muñoz, 2017), and most of the required information regarding the origin, genetic diversity, and

75   structure of juveniles used for restocking of natural stocks remains unavailable. Hence, natural stocks of

76   *P. magdalenae* are highly susceptible to experiencing disturbances of their genetic background resulting

77   from artificial mixtures of genetic stocks with different evolutionary histories or, alternatively, from the

78   high competition for resources among different stocks.

79   Since *P. magdalenae* performs long longitudinal migrations (around 1,224 km; velocity: 50.6 km/day)

80   (López-Casas et al., 2016), it is reasonable to think that its natural stocks experience extensive gene flow.

81   However, the observation that *Prochilodus lineatus* (Godoy, 1959) and *Prochilodus argenteus* (Godinho

82   & Kynard, 2006) show fidelity to spawning sites ("homing") suggests that *P. magdalenae* may exhibit a

83   population genetic structure even in the absence of physical barriers.

84 Furthermore, previous genetic studies have found the population structure and/or coexistence of multiples

85 stocks along the Magdalena River and several of its tributaries (López-Macías et al., 2009; see Mancera-

86 Rodríguez, Márquez, & Hurtado-Alarcón, 2013; Orozco-Berdugo & Narváez- Barandica, 2014).

87 Although this structure may result from the unregulated restocking of the natural stocks, it could also

88 reflect a natural behavior of *P. magdalenae* since similar patterns of genetic population structure have

89 been found in other congeners such as *P. reticulatus* (López-Macías et al., 2009), *P. argenteus* (Hatanaka

90 & Galetti Jr., 2003; Hatanaka, Henrique-Silva, & Galetti Jr., 2006; Barroca et al., 2012a), *P. lineatus*

91 (Ramella et al., 2006; Rueda et al., 2013; Gomes et al., 2017), and *P. costatus* (Barroca et al., 2012a,b).

92 This study tested the hypothesis that *P. magdalenae* exhibits genetic stocks that coexist and co-migrate

93 throughout the rivers, tributaries, and floodplain lakes of the different Colombian hydrographic areas

94 Magdalena-Cauca and Caribe. Additionally, we compare the genetic diversity and structure with those of

95 five sites (Pijiño, Llanito, Mompox, Palomino, and San Marcos) previously studied by Orozco-Berdugo

96 & Narváez-Barandica (2014). To test this hypothesis, we developed species-specific microsatellite loci

97 due to their advantages in the studies of population genetics (Fernandez-Silva et al., 2013; Putman &

98 Carbone, 2014).

99

100

101 **MATERIALS AND METHODS**

102

103 **Sample collection**

104 This study analyzed a total of 725 muscle tissues of *P. magdalenae* from the river mainstream and

105 floodplain lakes along the different Colombian hydrographic areas Magdalena-Cauca and Caribe (Fig. 1;

106   Table S1) that included 40 juveniles from a local fish hatchery. The samples preserved in 70% ethanol

107   were provided by Integral S.A. through two scientific cooperation agreements (September 19, 2013;

108   Grant CT-2013-002443). Sampling collection was performed by Integral S.A., framed under an

109   environmental permit from Ministerio de Ambiente, Vivienda y Desarrollo Territorial de Colombia #

110   0155 (January 30, 2009) for Ituango hydropower construction. Samples previously studied by Orozco-

111   Berdugo & Narváez-Barandica (2014) were collected during project 111752128352 of COLCIENCIAS

112   under collection permit #1293 (2013) of the Universidad del Magdalena.

113   **Microsatellite loci development**

114   Low-coverage sequencing of the genomic library of one specimen of *P. magdalenae* from the middle

115   section of the Magdalena River was performed using the Illumina MiSeq v2 platform (Illumina,

116   California, USA), the "whole genome shotgun" strategy and the Nextera library preparation kits

117   (Illumina, California, USA) for the sequence reads. All steps regarding the read cleaning, contig

118   assemblage, identification of microsatellite loci, primer design, in silico alignment of primers using

119   electronic PCR (ePCR), PCR optimization, and polymorphism analysis of 52 microsatellites were

120   performed following the methodology described by Landínez-García & Márquez (2016). A set of 21

121   polymorphic microsatellite loci were selected and fluorescently labeled for genotyping of 88 randomly

122   chosen samples. Then, a subset of 11 loci were selected for further evaluation of genetic diversity and

123   structure in 725 samples because they satisfied the criteria of clearly defined peaks, reproducibility and

124   consistency of amplifications, absence of stutter bands, specific bands, correct motif sizes, levels of

125   heterozygosity, and high polymorphism information content (PIC) values, among others parameters

126   required to validate new microsatellite primers (Neff, Garner, & Pitcher, 2011; Fernandez-Silva et al.,

127   2013; Schoebel et al., 2013).

128   **Genotyping of samples**

129   The PCRs were conducted in a volume of 10 µl, which contained 2–4 ng/µl of template DNA isolated

130   with the GeneJET Genomic DNA purification kit (Thermo Scientific, Karlsruhe, Germany) following the

131   manufacturer´s instructions, 1 × buffer (Invitrogen, California, USA), 0.2 mM dNTPs (Thermo Scientific,

132   Massachusetts, USA), 0.05 U/µl Platinum™ Taq DNA Polymerase (Invitrogen, , California, USA), 2.5

133   mM $MgCl_2$, 2% formamide (Sigma-Aldrich, Steinheim, Germany), 0.35 pmoles/µl labeled forward

134   primer (either FAM6, VIC, NED, or PET, Applied Biosystems, California, USA), and 0.5 pmoles/µl

135   reverse primer (Macrogen, Seoul, Korea). The PCRs were performed on a T100 thermocycler (BioRad,

136   California, USA) with an initial denaturation step of 95 °C for 3 min followed by 32 cycles consisting of

137   a denaturation step of 90 °C for 22 s and an annealing step for 18 s using the annealing temperatures

138   described for each primer in Table 1. The extension step and a final elongation were absent in this thermal

139   profile. Finally, the PCRs were submitted to electrophoresis on an automated sequencer ABI 3730 XL

140   (Applied Biosystems, California, USA) using GeneScan 500 LIZ dye size standard (Applied Biosystems,

141   California, USA) as the internal molecular size. Allelic fragments were denoted according to their

142   molecular size and scored using GeneMapper v4.0 software (Applied Biosystems, California, USA; Table

143   S2). Before the statistical analysis, Micro-Checker v.2.2.3 (van Oosterhout et al., 2004) was run to detect

144   potential genotyping errors.

145   **Statistical analysis**

146   Tests for Hardy-Weinberg and Linkage equilibria, observed (H_O) and expected (H_E) heterozygosities and

147   inbreeding coefficient (F_{IS}) were estimated using Arlequin v3.5.2.2 software (Excoffier, Laval, &

148   Schneider, 2005). The sequential Bonferroni correction was applied to adjust the statistical significance in

149   multiple comparisons (Holm, 1979; Rice, 1989). The average number of alleles per locus and the PIC

150   (Botstein et al., 1980) for each microsatellite locus were calculated with GenAlEx v6.503 software

151   (Peakall & Smouse, 2006) and Cervus v3.0.7 software (Marshall et al., 1998), respectively.

152   The genetic differentiation among geographical samples was calculated using the standardized statistics

153   F'_{ST} (Wright, 1943, 1965; Meirmans, 2006), Jost's D'_{est} (Jost, 2008; Meirmans & Hedrick, 2011) and

154   analysis of molecular variance (AMOVA) (Meirmans, 2006) with 10,000 permutations and bootstraps

155   included in GenAlEx v6.503 software (Peakall & Smouse, 2006). Furthermore, the diploid genotypes of

156   11 loci (22 variables) in 725 individuals were submitted to discriminant analysis of principal components

157   (DAPC) using the R-package Adegenet (Jombart, 2008).

158   To examine other groupings of the samples, genetic differentiation among samples was tested using the

159   Bayesian analysis of population partitioning with Structure v2.3.4 software (Pritchard, Stephens &

160   Donnelly, 2000). Parameters included 350,000 Monte Carlo Markov Chain steps and 50,000 iterations as

161   burn-in, the admixture model, correlated frequencies, and the LOCPRIOR option for detecting relatively

162   weak population structure (Hubisz et al., 2009). Each analysis was repeated 20 times for each simulated K

163   value, which ranged from 1 to n + 3 (n, number of populations compared). For a best estimation of

164   genetic stocks (K), STRUCTURESELECTOR web-based software (Li & Liu, 2018) was used to

165   calculate the ΔK ad hoc statistic (Evanno, Regnaut, & Goudet, 2005), the estimators MEDMEANK,

166   MAXMEANK, MEDMEDK, and MAXMEDK (Puechmaille, 2016), and to generate the graphical

167   representation of results using the integrated Clumpak software (Kopelman et al., 2015). Based on the

168   coancestry coefficients provided by Structure and Clumpp, the individuals were reorganized by genetic

169   stock in sample sites that showed multiple stocks and were later submitted to the genetic analyses

170   described above.

171   Additionally, the occurrence of recent genetic bottlenecks of populations was evaluated by calculating the

172   levels of heterozygosity and the M ratio using Bottleneck v1.2.02 software (Piry, Luikart, & Cornuet,

173   1999) and Arlequin v3.5.2.2 (Excoffier, Laval, & Schneider, 2005), respectively. Excess heterozygosity

174   was assessed by employing the Wilcoxon sign-rank test (Luikart & Cornuet, 1998). The M ratio – the

175   mean ratio of the number of alleles compared to the range of allele size – indicates that the population has

176   experienced a recent and severe reduction in population size when its values are smaller than 0.680

177   (Garza & Williamson, 2001).

178   To explore non-neutral evolutionary forces acting on the microsatellite loci, a scanning analysis was

179   performed using the BayeScan v2.1 software (Foll & Gaggiotti, 2008) to detect candidate loci under

180   selection. Parameters for BayeScan analyses included 10:1 prior odds for the neutral model and 20 pilot

181   runs consisting of 5,000 iterations each followed by 250,000 iterations with a burn-in length of 50,000

182   iterations (Foll & Gaggiotti, 2008).

183   **Phylogenetic relationships among genetic groups**

184   To explore the phylogenetic relationships among individuals sampled along the basin, partial fragments of

185   the mitochondrial *cox1* gene (~650 bp) were amplified in a subset of samples using primers and PCR

186   conditions previously described by Ward et al. (2005) and Ivanova et al. (2007). PCR products were

187   sequenced by the Sanger method using an automated sequencer, ABI 3730 XL (Applied Biosystems). The

188   best-fit evolutionary model was determined based on the Bayesian information criterion as implemented

189   in the jModelTest v2.1.7 software (Posada & Crandall, 1998). Phylogenetic relationships were determined

190   by Bayesian inference using the MrBayes v3.2.6 software (Ronquist & Huelsenbeck, 2003). For this

191   purpose, we performed two independent runs of 20 million generations sampled each 1,000 generations

192   using 25% as burn-in. The remaining values were left as default. The convergence of each parameter was

193   checked based on a potential scale reduction factor nearing 1, an average standard deviation of the split

194   frequencies lower than 0.010, and the visualization of the resulting trees was performed with FigTree

195   v1.4.3 software (Rambaut, 2012). Finally, the pair-wise divergences of *P. magdalenae* and *P. reticulatus*

196   haplotype sequences were estimated using the Kimura 2-parameters model in MEGA v10.1.8 software

197   (Tamura et al., 2013).

198

199 **RESULTS**

200

201 **Microsatellite loci development**

202 Genomic sequencing of the Illumina shotgun library of *P. magdalenae* (0.115 GB) generated 277,133

203 reads and 14,124 of 50,404 that contained microsatellite loci, were flanked by suitable PCR priming sites.

204 The dinucleotides (47.758%) were the most abundant repeat motifs, followed by tetranucleotide

205 (28.353%), trinucleotide (16.193%), pentanucleotide (5.146%), and hexanucleotide (2.549%) repeats. The

206 most common motifs found were AC (29.661%), TC (18.905%), ATT (4.811%), and AAAT (4.794%).

207 The sequences of contigs containing the microsatellite loci obtained in the present study are provided in

208 supplementary files S3 and S4.

209 A total of 21 of the 52 microsatellite loci evaluated were polymorphic and showed Hardy-Weinberg

210 disequilibrium (Table 1) and Linkage equilibrium (Table S5). The number of alleles per locus ranged

211 from 11 to 37, with an average number of 20.619 alleles/locus, the average values of observed and

212 expected heterozygosities were Ho = 0.589 and He = 0.876 and the PIC values ranged from 0.399 to

213 0.949 (average 0.867) (Table 1). A total of 10 loci failed to satisfy the selection criteria, showing low

214 values of PIC (Pma32), dropout and stuttering (Pma32, Pma08), inconsistent amplifications (Pma17,

215 Pma47, Pma57), or low-definition peaks (Pma42, Pma56, Pma26, Pma50). Consequently, only 11

216 (Pma39, Pma25, Pma02, Pma35, Pma01, Pma40, Pma46, Pma36, Pma18, Pma13, and Pma14) satisfied

217 most of the parameters required to validate the new microsatellites primers described previously.

218 **Genetic diversity, population demography, and outlier loci screening**

219 Comparisons among rivers revealed that 8 of 11 loci satisfied the Hardy-Weinberg equilibrium

220 expectations in at least one case (Table 2). However, the analysis across loci showed significant

221 departures from Hardy-Weinberg equilibrium expectations in all rivers evaluated (Table 2). The average

222   number of alleles per locus was higher in Cauca (22.455) and Magdalena (19.455), followed by Nare

223   (15.636), Sinú (15.273), the fish hatchery (14.818), and Atrato (14.636) and was lowest in San Jorge

224   (13.545) and Cesar (13.364). Additionally, the highest values of observed and expected heterozygosities

225   were found in San Jorge (Ho: 0.809; He: 0.884) and Cesar (Ho: 0.782; He: 0.873), followed by Sinú (Ho:

226   0.767; He: 0.882), Magdalena (Ho: 0.758; He: 0.896), and Cauca (Ho: 0.725; He: 0.898) and were lowest

227   in Atrato (Ho: 0.718; He: 0.879), the fish hatchery (Ho: 0.691; He: 0.880), and Nare (Ho: 0.659; He:

228   0.876) (Table 2).

229   Furthermore, comparisons among sites within each river showed similar high levels of genetic diversity

230   (Table 3). The highest value of genetic diversity was found in the floodplain lake Palagua in the

231   Magdalena River (Na: 17.182 alleles/locus; He: 0.895; Ho: 0.792), whereas the lowest was observed in

232   Beté, a site of the Atrato River (Na: 9.273 alleles/locus; He: 0.791; Ho: 0.711). In addition, all sites

233   exhibited a highly significant deficit of heterozygosity (Table 3) with Mata de Palma and Samaná Norte

234   River showing the lowest and highest heterozygosity deficits, respectively. Inbreeding coefficients (F_{IS})

235   per site in main rivers of the different Colombian hydrographic areas were significant and ranged from

236   0.120 to 0.255 (Table 3). Although decreased in magnitude, the inbreeding coefficients (Table 3)

237   remained significant even after comparing the genetic diversity according to genetic stocks in Chucurí,

238   Puerto Berrío, and Palagua and among the Magdalena River and tributaries.

239   Results of the genetic bottleneck (Table 4) were significant for all populations under the infinite alleles

240   model (IAM) and for most populations under the two-phase model (TPM), whereas they were non-

241   significant under the stepwise mutation model (SMM). As it is thought that few loci follow the strict

242   SMM (Piry, Luikart, & Cornuet, 1999), the best estimation of expected heterozygosity at mutation-drift

243   equilibrium is expected under a combination of IAM and TPM. Additionally, all values of the M ratio

244   were substantially smaller than 0.680, indicating that all populations have experienced recent and severe

245   reductions in population size (Table 4).

246   In contrast to other samples that did not show evidence of selection, BayeScan analysis revealed that 8 of

247   11 loci (Pma39, Pma25, Pma02, Pma35, Pma40, Pma36, Pma13, and Pma14) exhibit substantial evidence

248   of selection in the Magdalena River (Table 5).

249   **Genetic structure and phylogenetic relationships among the samples studied**

250   At regional scale, the Bayesian analysis showed the presence of two genetic stocks (ΔK = 2; MEDMEDK

251   = 2; MEDMEANK = 2), one in the Magdalena River (Chucurí + Puerto Berrio + Palagua) and the other

252   one in the remaining evaluated rivers (Fig. 2A), which is concordant with DAPC (Fig. 2B) and AMOVA

253   ($F_{ST(7, 1407)}$ = 0.009; P = **0.000**). Together with Chucurí + Puerto Berrio + Palagua, a predominant genetic

254   stock with different levels of genetic admixture in Sinú and Atrato rivers was revealed in the clusters

255   suggested by the MAXMEAN and MAXMED statistics (K = 3). The additional clustering patterns (K = 4

256   – 8; Fig. 2A) examined to compared them with other approaches, showed genetic admixture with other

257   stocks absent in the examined rivers. Likewise, pairwise comparisons of the standardized statistics F'_{ST}

258   (Meirmans, 2006) and Jost´s D'_{est} (Meirmans, & Hedrick, 2011) showed genetic differences among

259   Atrato, the fish hatchery, Sinú, and the remaining rivers (Table 6) as well as among the Magdalena River

260   and its tributaries, Cauca and Nare.

261   However, excluding samples that exhibit loci putatively under selection (Chucurí + Puerto Berrio +

262   Palagua), comparisons among sites within each river revealed a genetic admixture of two stocks (ΔK = 2;

263   MEDMEDK = 2; MEDMEANK = 2) homogenously distributed in Magdalena River and its tributaries

264   (Figs. 3A and 3B; Tables 6, 7). Additionally, this analysis revealed a genetic substructure in Sinú (ΔK =

265   2; MEDMEDK = 2; MEDMEANK = 2; Figs. 3C and 3D; $F_{ST(1, 67)}$ = 0.033; P = **0.000**; F'_{ST} = 0.027; P =

266   **0.004**; D'_{est} = 0.149; P = **0.005**). In Atrato River, the Bayesian analysis showed a single genetic stock

267   (ΔK = 2; MEDMEDK = 1; MEDMEANK = 1; Fig. 3E) although remaining analysis showed genetic

268   differentiation among sites (Fig. 3F; $F_{ST(1, 57)}$ = 0.045; P = **0.000**; F'_{ST} = 0.047; P = **0.000**; D'_{est} = 0.330; P

269   = **0.000**).

270   Finally, the Bayesian tree using the *cox1* gene clustered our samples (GenBank accession numbers

271   MK330430 to MK330494; Supplementary file S6) with sequences of *P. magdalenae* and *P. reticulatus*

272   deposited in public databases and in a different group, *Prochilodus marie* and *Prochilodus nigricans* (Fig.

273   4). Moreover, Kimura-2-parameters genetic distances (Table S7) were larger among haplotypes of *P.*

274   *magdalenae* (0.002 – 0.014) than among *P. magdalenae* and *P. reticulatus* haplotypes (0.002 – 0.010).

275

276

277   **DISCUSSION**

278

279   **Microsatellite loci development**

280   This work developed species-specific microsatellite loci using next-generation sequencing and

281   bioinformatic analysis. Although a total of 21 of 52 microsatellite loci with tri- and tetra-nucleotide motifs

282   were polymorphic in *P. magdalenae*, the consistency in the amplification in a larger sample, allelic size

283   class distribution, and high definition peaks allowed the selection of only 11 microsatellite loci for further

284   population genetic analysis. Most of the loci showed departures from Hardy-Weinberg equilibrium and

285   significant heterozygosity deficit, which may be related to the significant levels of inbreeding as well as

286   the genetic structure of the samples by the coexistence of two genetic stocks (see below).

287   Although the levels of genetic diversity measured by the expected heterozygosities were similar, the

288   levels of observed heterozygosity as well as the average number of alleles per locus found in this study

289   were substantially greater than those found by Orozco-Berdugo & Narváez-Barandica (2014). These

290   results support the idea that the heterologous microsatellite loci used by these authors may be limited by

291   the presence of null alleles or genotyping errors related to their dinucleotide motifs because a higher

292   variability is commonly expected in shorter motifs (e.g. 2mers, Orozco-Berdugo & Narváez-Barandica,

293   2014) compared with longer motifs (3mers and 4mers, this study). However, despite these differences,

294   both heterologous and species-specific microsatellite loci revealed a general deficit of heterozygotes in all

295   samples, indicating that its causes are biological rather than technical. In this context, the species-specific

296   microsatellite loci developed in this study seem to provide a good approach to studying the population

297   genetics of *P. magdalenae* considering that the levels of heterozygosity constitute a parameter used to

298   estimate the genetic diversity of the populations.

299   **Genetic diversity and population demography**

300   Microsatellite data revealed average values of genetic diversity (He: 0.737) among the highest values

301   found in other Prochilodontidae species, only surpassed by those reported for *P. costatus* (Melo et al.,

302   2013) and *P. argenteus* (Coimbra et al., 2017) (0.747 and 0.753 respectively). Similarly, the average

303   levels of expected heterozygosity were higher than that found in *P. magdalenae* measured by

304   heterologous microsatellites (He: 0.877; Orozco-Berdugo & Narváez-Barandica, 2014) and Neotropical

305   Characiforms (He: 0.675 ± 0.160; see review by Hilsdorf & Hallerman, 2017).

306   Additionally, this study found levels of observed heterozygosity higher than those found by Orozco-

307   Berdugo & Narváez-Barandica (2014). However, the use of species-specific microsatellite loci developed

308   in this study revealed similar values of expected heterozygosity among samples analyzed by Orozco-

309   Berdugo & Narváez-Barandica (2014) and the remaining samples analyzed, indicating that differences

310   between the two studies are related to the type of microsatellite loci utilized (heterologous vs. species-

311   specific microsatellite loci).

312   The significant deficit of heterozygosity in all studied samples corroborates the previous findings for *P.*

313   *magdalenae* from Magdalena River (Orozco-Berdugo & Narváez-Barandica, 2014); however, the

314   magnitude of the heterozygosity deficit as well as the inbreeding coefficient were substantially lower

315   (0.075 – 0.239) than those previously reported (0.624 – 0.788). Following Franklin (1980) and Soulé

316   (1980), the values above 10% of the inbreeding coefficient indicate that these populations require careful

317   management to avoid future detrimental effects on its populations. This point is important since it has

318   been recommended recently that any inbreeding coefficient higher than zero will usually have an adverse

319   fitness effect (Frankham, Bradshaw, & Brook, 2014).

320   Another non-excluding alternative is plausible considering that the significant deficit of heterozygosity

321   observed in all sites analyzed may be also explained by the coexistence of genetic stocks (Wahlund

322   effect) as this was evidenced by the genetic structure analysis (see below). Another biological cause of

323   heterozygosity deficit, assortative mating, does not seem to explain the results found in this study because

324   *P. magdalenae* is iteroparous and characterized by total spawning (Jaramillo-Villa & Jiménez-Segura,

325   2008) as described in its congeners, *P. costatus* (Carolsfield et al., 2004) and *P. lineatus* (Roux et al.,

326   2015).

327   On the other hand, this study also provided evidence for a population bottleneck, suggesting that *P.*

328   *magdalenae* shows signs of erosion of the genetic pool, likely by the constant pressure from fishing and

329   other anthropogenic activities exerted on its populations. Although paradoxical to the heterozygosity

330   deficit found in all populations evaluated, this outcome is plausible considering that the Bottleneck

331   algorithm tests not for an excess of heterozygotes (Ho > He) but rather for an excess of heterozygosity

332   (He > He at mutation-drift equilibrium) (Piry, Luikart, & Cornuet, 1999). Besides, the combination of a

333   population bottleneck and a heterozygosity deficit may result from population growth in a closed system,

334   population genetic structure, or admixture (Barson, Cable, & Oosterhout, 2009). Considering the lengths

335   of the rivers studied, population growth in a closed system is unlikely but the last two alternatives may

336   explain our results due to the coexistence of genetic stocks in the samples studied and the continuous

337   restocking of natural stocks using juveniles from fish hatcheries, which may create an apparent excess of

338   novel alleles and an incomplete allele frequency distribution. Similar results have also been found in

339   guppies, *Poecilia reticulata*, in Trinidad and Tobago (Barson, Cable, & Oosterhout, 2009).

340   **Genetic structure**

341   This study tested the hypothesis that *P. magdalenae* exhibits genetic stocks that coexist and co-migrate

342   along sections of the main channel and some tributaries of the Magdalena River (Cauca, San Jorge, and

343   Cesar), Sinú, and Atrato. Before testing this hypothesis, we compared the genetic structure at regional

344   scale, finding two spatially structured populations: one in the Magdalena River (Puerto Berrío and the

345   floodplains Chucurí and Palagua) and the other in the remaining rivers evaluated.

346   The geographical genetic structure may result from taxonomic differences among stocks due to the lack

347   of regulations on the restocking of natural stocks of *P. magdalenae*. The phylogenetic analysis using

348   partial sequences of *cox1* gene indicates that samples do not correspond to species such as *P. marie* or *P.*

349   *nigricans* because this genetic stock is clustered with previously published sequences of *P. magdalenae*

350   (Aguirre-Pabón, Narváez-Barandica, & Castro-García, 2013). However, it remains to be seen whether

351   they represent artificial mixtures of *P. magdalenae* and *P. reticulatus* because the current mitochondrial

352   phylogenetic analysis of Prochilodontidae does not allow the two species to be discriminated (Melo et al.,

353   2016b, 2018). Moreover, the morphological and molecular similitudes have led to the proposal that *P.*

354   *magdalenae* and *P. reticulatus* represent only one species with probable allopatric differentiation resulting

355   from the uplift of the Sierra del Perijá (Melo et al., 2016b). Thus, a separated clustering of mitochondrial

356   sequences of those stocks is not expected in the phylogenetic analysis even though they represent

357   allopatric populations.

358   An alternative explanation is that the genetic differences result from eight outlier loci that are putatively

359   under selection in three sites of the Magdalena River, suggesting that *P. magdalenae* experiences

360   natural/artificial selection or local adaptation, although testing of these hypotheses is out of the scope of

361   the present study. The explanation that outlier loci represent false positives resulting from the inclusion of

362   severely bottlenecked populations (Teshima, Coop, & Przeworski, 2006; Foll & Gaggiotti, 2008) seems

363   unlikely because the significant excess of heterozygosity and small values of the M ratio were found even

364 in populations that do not exhibit outlier loci. Thus, considering that those sites have been exposed to

365 restocking since 20 years ago and since most microsatellite loci are not transcriptionally active, the outlier

366 loci found in this study may reflect hitchhiking selection resulting from restocking using juveniles

367 selected artificially by fish hatcheries. Alternatively, the outlier loci may result from asymmetric gene

368 flow by unidirectional migration from hatchery stocks to wild populations. Similar results were found in

369 Denmark in populations of three brown trout, which have been significantly admixtured with stocked

370 hatchery trout (Hansen, Meier, & Mensberg, 2010).

371 Although the above reasoning might explain the genetic differences between stocks, an additional

372 justification is required to explain the restricted distribution of one genetic stock in only three sites of the

373 Magdalena River considering the migratory abilities of these species/allopatric populations. Thus, this

374 genetic structure seems to result from recent restocking before reproductive/feeding migrations, use of

375 artificial barriers to avoid migration of the fish, clogging by sedimentation or vegetation, or the

376 desiccation of access to floodplain lakes or may be a product of the intensive anthropic intervention in

377 these territories characterized by the exploitation of hydrocarbons and livestock. This idea is concordant

378 with the fact that degradation of preferred habitat and barriers that impede dispersal contribute to the

379 degree of genetic differentiation among populations (Faulks, Gilligan, & Beheregaray, 2011).

380 Furthermore, the results found here provide support for the hypothesis that *P. magdalenae* exhibits

381 genetic stocks that coexist and co-migrate along sections of the rivers Magdalena, Cauca, Cesar

382 (tributaries of the Magdalena River), Sinú, and Atrato. Since similar patterns of genetic structure are

383 found in *P. reticulatus* (López-Macías et al., 2009), *P. marggravii* (Hatanaka & Galetti Jr., 2003), *P.*

384 *argenteus* (Sanches et al., 2012), *P. costatus* (Barroca et al., 2012a), *P. magdalenae* (Orozco Berdugo &

385 Narváez Barandica, 2014; Hernández, Navarro, & Muñoz, 2017), and *Ichthyoelephas longirostris*

386 (Landínez-García & Márquez, 2016), this outcome supports the idea that this genetic structure is a

387 generalized tendency within the family Prochilodontidae.

388   Excluding the genetic stock of Puerto Berrío and the floodplains Chucurí and Palagua, each river showed

389   the coexistence of at least two genetic stocks. Homogeneous and non-homogeneous distributions of these

390   genetic stocks along the rivers explain similarities (Cauca, Magdalena, San Jorge, Cesar, and Nare) as

391   well as geographical differences among the rivers analyzed (within Magdalena, including Puerto Berrío

392   and the floodplains Chucurí and Palagua, Sinú, and Atrato). This genetic structure also explains the

393   significant heterozygosity deficit observed in all sites analyzed (Wahlund effect) as discussed above.

394   Similar evidence of the Wahlund effect has been documented in the congener *P. costatus*, which

395   exhibited genetic differences resulting from temporal isolation (Braga-Silva & Galetti Jr., 2016).

396   Although sampling in this study was not designed to detect temporal genetic structuring, genetic

397   similarities among samples collected in different years suggest that the Wahlund effect must be more

398   spatial than temporal. It remains to be seen whether this behavior is natural or artificial, considering that

399   the restocking activities have been widely implemented along different Colombian rivers.

400

401   **CONCLUSIONS**

402   This study provides evidence that *P. magdalenae* exhibits high genetic diversity, significant inbreeding

403   levels between 0.162 to 0.202 per genetic stock, and signs of erosion of the genetic pool and conforms a

404   mixture of genetic stocks heterogeneously distributed along the rivers studied. Additionally, this study

405   developed a set of 11 microsatellite loci that allows the detection of reliable levels of genetic diversity,

406   providing a tool for monitoring changes in the genetic diversity of the species, brood stocks, and juveniles

407   used for supportive breeding and for measuring the efficacy of current population restocking activities.

408   Management and conservation strategies need to be implemented at the level of the basins Magdalena-

409   Cauca, Sinú, and Atrato concordantly with their genetic population structure.

410

411 **ACKNOWLEDGMENTS**

415

416 **REFERENCES**

417

418 Aguirre-Pabón J, Narváez-Barandica J, Castro-García L. 2013. Mitochondrial DNA variation of the

419   bocachico *Prochilodus magdalenae* (Characiformes, Prochilodontidae) in the Magdalena River

420   Basin, Colombia. *Aquatic Conservation: Marine and Freshwater Ecosystems* 23:594–605. DOI:

421   10.1002/aqc.2339.

422 Barroca TM, Arantes FP, Magalhães BF, Siqueira FF, Horta CCR, Pena IF, Dergam JA. Kalapothakis E.

423   2012a. Genetic diversity and population structure of *Prochilodus costatus* and *Prochilodus*

424   *argenteus* preceding dam construction in the Paraopeba River, São Francisco River Basin, Minas

425   Gerais, Brazil. *Open Journal of Genetics* 02:121–130. DOI: 10.4236/ojgen.2012.22017.

426 Barroca TM, Santos GB, Duarte NVR, Kalapothakis E. 2012b. Evaluation of genetic diversity and

427   population structure in a commercially important freshwater fish *Prochilodus costatus*

428   (Characiformes, Prochilodontidae) using complex hypervariable repeats. *Genetics and Molecular*

429   *Research* 11:4456–4467. DOI: 10.4238/2012.September.27.4.

430 Barson NJ, Cable J, Oosterhout CVAN. 2009. Population genetic analysis of microsatellite variation of

431   guppies (*Poecilia reticulata*) in Trinidad and Tobago: Evidence for a dynamic source – sink

432   metapopulation structure, founder events and population bottlenecks. *Journal of Evolutionary*

433   *Biology* 22:485–497. DOI: 10.1111/j.1420-9101.2008.01675.x.

434 Batista VS, Lima LG. 2010. In search of traditional bio-ecological knowledge useful for fisheries co-

435   management: The case of jaraquis *Semaprochilodus* spp. (Characiformes, Prochilodontidae) in

436   Central Amazon, Brazil. *Journal of Ethnobiology and Ethnomedicine* 6:15. DOI: 10.1186/1746-

437   4269-6-15.

438 Botstein D, White LR, Skolnick M, Davis RW. 1980. Construction of a genetic linkage map in man using

439   restriction fragment length polymorphisms. American Journal of Human Genetics 32:314–

440   331.Braga-Silva A, Galetti Jr. PM. 2016. Evidence of isolation by time in freshwater migratory fish

441   *Prochilodus costatus* (Characiformes, Prochilodontidae). *Hydrobiologia* 765:159–167. DOI:

442   10.1007/s10750-015-2409-8.

443 Carolsfield J, Harvey B, Ross C, Baer A. 2004. *Migratory Fishes of South America: Biology, Fisheries,*

444   *and Conservation Status*. Canada: World Fisheries Trust, World Bank, IDRC.

445 Castro RMC, Vari RP. 2004. Detritivores of the South American fish family Prochilodontidae (Teleostei:

446   Ostariophysi: Characiformes): A phylogenetic and revisionary study. *Smithsonian Contributions to*

447   *Zoology* 622:1–190. DOI: 10.5479/si.00810282.622.

448 Coimbra MRM, Lima APS, Oliveira KKC, Severi W. 2017. Microsatellite assessment of the genetic

449   diversity in indigenous populations of curimba (*Prochilodus argenteus*) in the São Francisco River

450   (Brazil). *Conservation Genetics* 18:965–975. DOI: 10.1007/s10592-017-0947-5.

451 Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population

452   bottlenecks from allele frequency data. *Genetics* 144:2001–2014.

453 Cortés Millán GA. 2003. Guía para el manejo, cría y conservación del bocachico *Prochilodus*

454   *magdalenae* Steindachner 1878. Bogotá, DC: Convenio Andrés Bello.

455 DellaRosa P, Roux JP, Sánchez S, Ortiz JC, Domitrovic HA. 2014. Productividad del sábalo (*Prochilodus*

456   *lineatus*) cultivado en estanques con diferentes tipos de fondo. *Revista Veterinaria* 25:126–130.

457 Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software

458   Structure: A simulation study. *Molecular Ecology* 14:2611–2620. DOI: 10.1111/j.1365-

459   294X.2005.02553.x.

460 Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): An integrated software package for

461   population genetics data analysis. *Evolutionary Bioinformatics* 1:47–50. DOI: 10.4137/EBO.S0.

462   FAO. 2011. Taller sobre repoblamiento de cuerpos de agua continentales en américa latina y el caribe. In:

463   *Serie de Acuicultura y Pesca en América Latina*. Cuernavaca, México: FAO Organisation, 21–24.

464   Faulks LK, Gilligan DM, Beheregaray LB. 2011. The role of anthropogenic vs. natural in-stream

465   structures in determining connectivity and genetic diversity in an endangered freshwater fish,

466   Macquarie perch (*Macquaria australasica*). *Evolutionary Applications* 4:589–601. DOI:

467   10.1111/j.1752-4571.2011.00183.x.

468   Fernandez-Silva I, Whitney J, Wainwright B, Andrews KR, Ylitalo-Ward H, Bowen BW, Toonen RJ,

469   Goetze E, Karl SA. 2013. Microsatellites for next-generation ecologists: A post-sequencing

470   bioinformatics pipeline. *PLoS One* 8:e55990. DOI: 10.1371/journal.pone.0055990.

471   Flecker AS. 1996. Ecosystem engineering by a dominant detritivore in a diverse tropical stream. *Ecology*

472   77:1845–1854.

473   Flores-Nava A, Brown A. 2010. *Peces nativos de agua dulce de América del Sur de interés para la*

474   *acuicultura: una síntesis del estado de desarrollo tecnológico de su cultivo*. Roma, Italia: FAO.

475   Foll M, Gaggiotti O. 2008. A genome-scan method to identify selected loci appropriate for both dominant

476   and codominant markers: A Bayesian perspective. *Genetics* 180:977 LP-993. DOI:

477   10.1534/genetics.108.092221.

478   Frankham R, Bradshaw C, Brook B. 2014. *Genetics in conservation management: Revised*

479   *recommendations for the 50/500 rules, Red List criteria and population viability analyses*. DOI:

480   10.1016/j.biocon.2013.12.036.

481   Franklin IR. 1980. Evolutionary change in small populations. In: Soulé ME, Wilcox BA eds.

482   *Conservation biology: an evolutionary-ecological perspective*. Sunderland, MA: Sinauer Associates,

483   135–149.

484   Garza JC, Williamson EG. 2001. Detection of reduction in population size using data from microsatellite

485   loci. *Molecular Ecology* 10:305–318.

486   Godinho AL, Kynard B. 2006. Migration and spawning of radio-tagged zulega *Prochilodus argenteus* in

487   a dammed Brazilian river. *Transactions of the American Fisheries Society* 135:811–824. DOI:

488   10.1577/T04-176.1.

489   Godoy MP. 1959. Age, growth, sexual maturity, behaviour, migration, tagging and transplantation of the

490   curimbata, *Prochilodus scrofa*, Steindachner, 1881, of the Mogi Guassu River, Sao Paulo State,

491   Brazil. *Anais da Academia Brasileira de Ciencias* 31:447–477.

492   Hansen MM, Meier K, Mensberg K. 2010. Identifying footprints of selection in stocked brown trout

493   populations: a spatio-temporal approach. *Molecular Ecology* 19:1787–1800. DOI: 10.1111/j.1365-

494   294X.2010.04615.x.

495   Hartl DL, Clark AG. 1997. *Principles of population genetics*. Massachusetts: Sinauer Associates.

496   Hatanaka T, Galetti Jr. PM. 2003. RAPD markers indicate the occurrence of structured populations in a

497   migratory freshwater fish species. *Genetics and Molecular Biology* 26:19–25. DOI: 10.1590/S1415-

498   47572003000100004.

499   Hatanaka T, Henrique-Silva F, Galetti Jr. PM. 2006. Population substructuring in a migratory freshwater

500   fish *Prochilodus argenteus* (Characiformes, Prochilodontidae) from the São Francisco River.

501   *Genetica* 126:153–159. DOI: 10.1007/s10709-005-1445-0.

502   Hernández H. D, Navarro M. O, Muñoz F. J. 2017. Diversidad genética del bocachico *Prochilodus*

503   *magdalenae* en el departamento de Sucre. *Revista Colombiana de Ciencia Animal* 9:99–106. DOI:

504   10.24188/recia.v9.nS.2017.527.

505   Hilsdorf AWS, Hallerman EM. 2017. *Genetic resources of Neotropical fishes*. Switzerland: Springer

506   International Publishing. DOI: 10.1007/978-3-319-55838-7.

507   Holm S. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of*

508   *Statistics* 6:65–70.

509   Hubisz MJ, Falush D, Stephens M, Pritchard JK. 2009. Inferring weak population structure with the

510   assistance of sample group information. *Molecular Ecology Resources* 9:1322–1332. DOI:

511   10.1111/j.1755-0998.2009.02591.x.

512   Ivanova NV, Zemlak TS, Hanner RH, Hebert PDN. 2007. Universal primer cocktails for fish DNA

513   barcoding. *Molecular Ecology Notes* 7:544–548. DOI: 10.1111/j.1471-8286.2007.01748.x.

514   Jaramillo-Villa U, Jiménez-Segura LF. 2008. Algunos aspectos biológicos de la población de *Prochilodus*

515   *magdalenae* en las ciénagas de Tumaradó (Río Atrato), Colombia. *Actualidades Biológicas* 30:55–

516   66.

517   Jombart T. 2008. ADEGENET: A R package for the multivariate analysis of genetic markers.

518   *Bioinformatics* 24:1403–1405. DOI: 10.1093/bioinformatics/btn129.

519   Jost LOU. 2008. G_{ST} and its relatives do not measure differentiation. *Molecular ecology* 17(18): 4015-

520   4026.

521   Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. CLUMPAK: a program for

522   identifying clustering modes and packaging population structure inferences across K. *Molecular*

523   *Ecology Resources* 15:1179–1191. DOI: 10.1111/1755-0998.12387.

524   Landínez-García RM, Márquez EJ. 2016. Development and characterization of 24 polymorphic

525   microsatellite loci for the freshwater fish *Ichthyoelephas longirostris* (Characiformes:

526   Prochilodontidae). *PeerJ* 4:e2419. DOI: 10.7717/peerj.2419.

527   Lasso C, Agudelo E, Jimenez-Segura L, Ramírez-Gil H, Morales-Betancourt M, Ajiaco-Martínez R,

528   Gutiérres F de la P, Usma S, Muñoz S, Sanabria-Ochoa A. 2010. *Catálogo de los recursos*

529   *pesqueros continentales de Colombia*. Bogotá, D. C, Colombia: Instituto de Investigación de

530   Recursos Biológicos Alexander von Humboldt (IAvH).

531   Li YL, Liu JX. 2018. STRUCTURESELECTOR: A web-based software to select and visualize the

532   optimal number of clusters using multiple methods. *Molecular Ecology Resources* 18:176–177.

533   DOI: 10.1111/1755-0998.12719.

534   López-Casas S, Jiménez-Segura LF, Agostinho AA, Pérez CM. 2016. Potamodromous migrations in the

535   Magdalena River Basin: Bimodal reproductive patterns in Neotropical rivers. *Journal of Fish*

536   *Biology* 89:157–171. DOI: 10.1111/jfb.12941.

537   López-Macías JN, García Vallejo F, Rubio E, Rincón E, Castillo Giraldo A, Cerón F, López-Macias JN,

538   Garcia Vallejo F, Rúbio Rincón E, Castillo Giraldo A, Cerón F. 2009. Diversidad Genética del

539   Bocachico (*Prochilodus reticulatus*) de la Cuenca Alta del Río Cauca (Colombia). *Acta Biológica*

540   *Paranaense* 38:113–138. DOI: 10.5380/abpr.v38i0.16928.

541   Luikart G, Cornuet JM. 1998. Empirical evaluation of a test for identifying recently bottlenecked

542   populations from allele frequency data. *Conservation Biology* 12:228–237.

543   Mancera-Rodríguez NJ, Márquez E, Hurtado-Alarcón JC. 2013. Uso de la citogenética y técnicas

544   moleculares en estudios de diversidad genética en peces colombianos. In: López HA, ed. *Biología*

545   *molecular aplicada a la producción animal y la conservación de especies silvestres*. Medellín:

546   Universidad Nacional de Colombia, 237–279.

547   Marshall TC, Slate J, Kruuk LE, Pemberton JM. 1998. Statistical confidence for likelihood-based

548   paternity inference in natural populations. *Molecular Ecology* 7:639–655. DOI: 10.1046/j.1365-

549   294x.1998.00374.x.

550   Meirmans PG. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation

551   measure. *Evolution* 60:2399–2402. DOI: 10.1111/j.0014-3820.2006.tb01874.x.

552   Meirmans PG, Hedrick PW. 2011. Assessing population structure: F_{ST} and related measures. *Molecular*

553   *Ecology Resources* 11:5–18. DOI: 10.1111/j.1755-0998.2010.02927.x.

554   Melo BF, Dorini BF, Foresti F, Oliveira C. 2018. Little divergence among mitochondrial lineages of

555   *Prochilodus* (Teleostei, Characiformes). *Frontiers in Genetics* 9:107. DOI:

556   10.3389/fgene.2018.00107.

557   Melo BF, Ochoa LE, Vari RP, Oliveira C. 2016a. Cryptic species in the Neotropical fish genus

558   *Curimatopsis* (Teleostei, Characiformes). *Zoologica Scripta* 45:650–658. DOI: 10.1111/zsc.12178.

559   Melo BF, Sato Y, Foresti F, Oliveira C. 2013. The roles of marginal lagoons in the maintenance of

560   genetic diversity in the Brazilian migratory fishes *Prochilodus argenteus* and *P. costatus*.

561   *Neotropical Ichthyology* 11:625–636. DOI: 10.1590/S1679-62252013000300016.

562   Melo BF, Sidlauskas BL, Hoekzema K, Frable BW, Vari RP, Oliveira C. 2016b. Molecular phylogenetics

563   of the Neotropical fish family Prochilodontidae (Teleostei: Characiformes). *Molecular*

564 *Phylogenetics and Evolution* 102:189–201. DOI: 10.1016/j.ympev.2016.05.037.

565 Mojica JI, Usma JS, Álvarez-León R, Lasso CA. 2012. *Libro rojo de peces dulceacuícolas de Colombia*.

566 Bogotá, DC: Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Instituto

567 de Ciencias Naturales de la Universidad Nacional de Colombia, WWF Colombia y Universidad de

568 Manizales.

569 Neff BD, Garner SR, Pitcher TE. 2011. Conservation and enhancement of wild fish populations:

570 preserving genetic quality versus genetic diversity. *Canadian Journal of Fisheries and Aquatic*

571 *Sciences* 68:1139–1154. DOI: 10.1139/f2011-029.

572 van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. Micro-Checker: Software for

573 identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*

574 4:535−538. DOI: 10.1111/j.1471-8286.2004.00684.x.

575 Orozco-Berdugo G, Narváez-Barandica JC. 2014. Genetic diversity and population structure of

576 bocachico *Prochilodus magdalenae* (Pisces, Prochilodontidae) in the Magdalena River Basin and

577 its tributaries, Colombia. *Genetics and Molecular Biology* 37:37–45.

578 Peakall R, Smouse PE. 2006. Genalex 6: Genetic analysis in Excel, population genetic software for

579 teaching and research. *Molecular Ecology Notes* 6:288−295. DOI: 10.1111/j.1471-

580 8286.2005.01155.x.

581 Piry S, Luikart G, Cornuet J. M. 1999. BOTTLENECK: A computer program for detecting recent

582 reductions in the effective population size using allele frequency data. *Heredity* 90:502–503. DOI:

583 10.1093/jhered/90.4.502.

584 Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics*

585 14:817–818.

586 Povh JA, Lopera Barrero NM, Ribeiro RP, Lupchinski Jr. E, Gomes PC, Lopes TS. 2008. Monitoreo

587 genético en programas de repoblamiento de peces mediante marcadores moleculares. *Ciencia e*

588 *Investigación Agraria* 35:5–15. DOI: 10.4067/S0718-16202008000100001.

589 Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype

590   data. *Genetics* 155:945–959.

591   Puechmaille SJ. 2016. The program structure does not reliably recover the correct population structure

592   when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular*

593   *Ecology Resources* 16:608–627. DOI: 10.1111/1755-0998.12512.

594   Putman AI, Carbone I. 2014. Challenges in analysis and interpretation of microsatellite data for

595   population genetic studies. *Ecology and Evolution 4*:4399–4428. DOI: 10.1002/ece3.1305.

596   Rambaut A. 2012. FigTree. http://tree.bio.ed.ac.uk/software/figtree/.

597   Rice WR. 1989. Analyzing tables of statistical tests. *Evolution* 43:223–225.

598   Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models.

599   *Bioinformatics* 19:1572–1574. DOI: 10.1093/bioinformatics/btg180.

600   Roux JP, González AO, Ortiz J, Sánchez S, Comolli J. 2015. Larvicultura intensiva de sábalo

601   (*Prochilodus lineatus*) con diferentes densidades de cría. *Revista Veterinaria* 26:143–146.

602   Sanches A, Galetti Jr. PM, Galzerani F, Derazo J, Cutilak Bianchi B, Hatanaka T. 2012. Genetic

603   population structure of two migratory freshwater fish species (*Brycon orthotaenia* and *Prochilodus*

604   *argenteus*) from the Sao Francisco River in Brazil and its significance for conservation. *Latin*

605   *American Journal of Aquatic Research* 40:177–186. DOI: 10.3856/vol40-issue1-fulltext-17.

606   Schoebel CN, Brodbeck S, Buehler D, Cornejo C, Gajurel J, Hartikainen H, Keller D, Leys M, Ríčanová

607   S, Segelbacher G, Werth S, Csencsics D. 2013. Lessons learned from microsatellite development for

608   nonmodel organisms using 454 pyrosequencing. *Journal of Evolutionary Biology* 26:600–611. DOI:

609   10.1111/jeb.12077.

610   Soulé ME. 1980. Thresholds for survival: Maintaining fitness and evolutionary potential. In: Soulé ME,

611   Wilcox BA eds. *Conservation biology: an evolutionary-ecological perspective*. Sunderland, MA:

612   Sinauer Associates, 151–169.

613   Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics

614   Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549.

615   Taylor BW, Flecker AS, Hall RO. 2006. Loss of a harvested fish species disrupts carbon flow in a diverse

616   tropical river. *Science* 313:833 LP-836. DOI: 10.1126/science.1128223.

617   Teshima KM, Coop G, Przeworski M. 2006. How reliable are empirical genomic scans for selective

618   sweeps? *Genome Research* 16:702–712. DOI: 10.1101/gr.5105206.

619   Usma JS, Valderrama M, Escobar MD, Ajiaco-Martínez RE, Villa-Navarro FA, Castro F, Ramírez-Gil H,

620   Sanabria AI, Ortega-Lara A, Maldonado-Ocampo JA, Alonso JC, Cipamocha C, Usma-Oviedo JS,

621   Valderrama M, Escobar MD, Ajiaco-Martínez RE, Villa-Navarro FA, Castro F, Ramírez-Gil H,

622   Sanabria-Ochoa AI, Ortega-Lara A, Maldonado-Ocampo JA, Alonso JC, Cipamocha C. 2009. *Peces*

623   *dulceacuícolas migratorios en Colombia*. Bogota, DC; Colombia, DC; Colombia: Colombia:

624   Ministro de Ambiente, Vivienda Y Desarrollo Territorial & WWF Colombia.

625   Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish species.

626   *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 360:1847–

627   1857 DOI 10.1098/rstb.2005.1716.

628   Wright S. 1943. Isolation by distance. *Genetics* 28(2): 114.

629   Wright S. 1965. The interpretation of population structure by F-statistics with special regard to systems of

630   mating. *Evolution* 19(3): 395–420.

631

Table 1(on next page)

Primer sequences, characteristics, polymorphism levels, and genetic diversity of 21 species-specific microsatellite loci in 88 individuals of *Prochilodus magdalenae* randomly chosen from the whole sample.

Ta: annealing temperature standardized in PCRs, Na: number alleles per locus; Ra: allelic size range (bp); PIC: polymorphism information content; H_O and H_E: observed and expected heterozygosities, respectively; P: statistical significance (values in bold represent significance at $P < 0.05$). [a] Satisfied selection criteria, [b] inconsistent amplifications, [c] low definition peaks, [d] dropout, [e] stuttering, [f] low value of PIC.

Locus	Primer sequence (5′−3′)	Motif	Ta(°C)	Na	Ra	PIC	H_O	H_E	P
Pma39[a]	F: CCAATGACCTGTTTTCTACATTTGG R: AATCTACTACCCGGATGGCG	(ATCT)n	58	14	231 - 283	0.860	0.671	0.878	**0.002**
Pma25[a]	F: AAGGGGAAAGAAATCCAGGC R: ATCCTGGGTTCATACCGACG	(AAGGC)n	60	12	174 - 229	0.816	0.795	0.840	**0.003**
Pma02[a]	F: CGACATTCAACATGACAGTGC R: CACCAAATTGATGCAAACTGC	(ATCT)n	58	19	231 - 307	0.917	0.816	0.927	**0.019**
Pma35[a]	F: GCAGTCTGGCATTTTAGTGGC R: ACCACATCTCGCATCACTGG	(ATCT)n	58	21	269 - 353	0.935	0.536	0.944	**0.000**
Pma56[c]	F: ATTTGGTGCCTGTAGCTGGG R: ACGGTCGGTGCACTAATTCC	(ATT)n	60	37	132 - 279	0.949	0.670	0.956	**0.000**
Pma01[a]	F: TTGTCATTTCCCGGTTTTCC R: TGGCCCAGCTGTAATTTGG	(ATCT)n	58	25	216 - 344	0.938	0.753	0.947	**0.000**
Pma40[a]	F: CTGGTTACCCACCACTGTCG R: CACATTGCCATTTGGAGACG	(ATCT)n	58	25	236 - 344	0.932	0.686	0.941	**0.000**
Pma46[a]	F: TTGATGTAAACATCTCATTGCCG R: TTGCTGGAGGTTCTGTCCG	(ATCT)n	56	19	126 - 198	0.918	0.830	0.929	**0.005**
Pma36[a]	F: TCATGATGAAATGCCACACC R: TGCACGTGAACTTAGGCACC	(ATCT)n	58	24	119 - 219	0.925	0.674	0.935	**0.000**
Pma18[a]	F: ACTGAGACAAAACCCGGAGG R: CTTCATACACCCACCATCAGG	(ATT)n	62	13	209 - 251	0.728	0.471	0.755	**0.000**
Pma13[a]	F: CCGAAGCTATTTACCCAGCG R: TGAAATATGCTCGTGCTCCC	(AAAT)n	62	11	154 - 194	0.815	0.670	0.841	**0.007**
Pma14[a]	F: GTTCAGGGTCCTGCTGTTCC R: TTTCGGTGTTGGAACATTGC	(TTC)n	58	21	146 - 209	0.907	0.605	0.919	**0.000**
Pma42[c]	F: TTACACAGCGTCCCAATTCC R: GCTGCAGGGATTGTCCTACC	(ATCT)n	58	25	146 - 254	0.933	0.759	0.942	**0.000**
Pma26[c]	F: TGATGTTTCCTCCCCTCACC R: GTGTTTCCTGCTCTCTGCCC	(ATCTC)n	58	20	141 - 281	0.888	0.553	0.902	**0.000**
Pma34[d,e]	F: GAGTGCCGATGACAGAGACG R: CAAGATGCCCTGTAGTGCCC	(ATCT)n	58	24	202 - 406	0.919	0.363	0.930	**0.000**
Pma50[c]	F: GATTCCTTCCTACCGGAGCC R: ATGAGCACCACCCTCAATCC	(ATCT)n	58	30	171 - 299	0.942	0.565	0.950	**0.000**
Pma32[f]	F: GAAAAGACACAACAGCGCCC R: GTCGCTAATAGCCATGCCG	(ATCT)n	58	13	146 - 294	0.399	0.375	0.430	**0.006**
Pma57[b]	F: ATGGCAATGGTTAAGGGTCG	(AAC)n	58	11	191 - 230	0.838	0.306	0.861	**0.000**

	R: CTGAAAGCCCCTGTTTGTGC								
Pma08[b,e]	F: TTTTATTATTCCCCATTTTCTCCC R: TGGGTTTTGAGCTGTTCTGC	(AAAG)n	58	12	254 - 298	0.833	0.257	0.856	**0.000**
Pma17[b]	F: CTGTGGGCAGCAAAGTGC R: CTTTGAGCCACTTCAAACGG	(ATT)n	58	36	151 - 346	0.892	0.595	0.904	**0.000**
Pma47[b]	F: TGGCTGCTAAATTAAATCCTTTGG R: AAGCAAAACCGTTCCACAGC	(ATCT)n	58	21	176 - 280	0.915	0.413	0.928	**0.000**
Across loci				20.619	119 - 353	0.867	0.589	0.876	**0.000**

1

Table 2(on next page)

Genetic diversity of *Prochilodus magdalenae* in main rivers of the distribution range of the species in Colombian hydrographic areas

N: sample size; Na: number alleles per locus; H_O and H_E: observed and expected heterozygosity, respectively; P: statistical significance for tests of departure from Hardy Weinberg equilibrium. Values in bold represent significance at $P < 0.05$.

River (N)	Diversity	Pma39	Pma25	Pma02	Pma35	Pma01	Pma40	Pma46	Pma36	Pma18	Pma13	Pma14	Across loci
Cauca	Na	19.000	15.000	25.000	25.000	34.000	28.000	21.000	25.000	17.000	13.000	25.000	22.455
(308)	H_O	0.662	0.805	0.883	0.591	0.756	0.708	0.818	0.688	0.552	0.821	0.685	0.725
	H_E	0.889	0.855	0.935	0.935	0.941	0.944	0.920	0.932	0.775	0.842	0.926	0.898
	P	**0.000**	**0.000**	**0.002**	**0.000**	**0.000**	**0.000**	**0.001**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Magdalena	Na	15.000	12.000	21.000	22.000	31.000	26.000	18.000	21.000	15.000	12.000	21.000	19.455
(232)	H_O	0.664	0.891	0.861	0.642	0.781	0.679	0.818	0.745	0.599	0.854	0.803	0.758
	H_E	0.874	0.865	0.930	0.941	0.943	0.944	0.925	0.926	0.784	0.833	0.929	0.896
	P	**0.000**	**0.001**	0.510	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.058	**0.002**	**0.000**
San Jorge	Na	10.000	11.000	16.000	19.000	16.000	18.000	14.000	14.000	9.000	9.000	13.000	13.545
(20)	H_O	0.850	1.000	0.950	0.700	0.950	0.750	0.900	0.800	0.850	0.700	0.450	0.809
	H_E	0.881	0.878	0.947	0.951	0.947	0.942	0.918	0.914	0.831	0.851	0.912	0.884
	P	0.650	0.299	0.645	**0.000**	0.638	**0.002**	0.531	0.307	**0.009**	0.318	**0.000**	**0.000**
Cesar	Na	10.000	9.000	15.000	16.000	21.000	15.000	13.000	17.000	9.000	8.000	14.000	13.364
(20)	H_O	0.500	0.950	1.000	0.750	1.000	0.650	1.000	0.800	0.600	0.800	0.550	0.782
	H_E	0.867	0.874	0.940	0.949	0.954	0.940	0.924	0.927	0.815	0.776	0.883	0.873
	P	**0.000**	0.890	0.947	**0.033**	0.208	**0.002**	0.484	0.148	0.097	0.846	**0.000**	**0.000**
Nare	Na	13.000	13.000	19.000	18.000	25.000	19.000	14.000	20.000	8.000	8.000	15.000	15.636
(41)	H_O	0.610	0.780	0.902	0.415	0.780	0.439	0.927	0.805	0.341	0.756	0.488	0.659
	H_E	0.887	0.877	0.931	0.930	0.952	0.931	0.912	0.934	0.708	0.781	0.912	0.876
	P	**0.002**	0.200	0.619	**0.000**	**0.011**	**0.000**	0.792	**0.001**	**0.000**	0.357	**0.000**	**0.000**
Sinú	Na	13.000	12.000	19.000	19.000	23.000	18.000	14.000	15.000	8.000	10.000	17.000	15.273
(34)	H_O	0.441	0.912	0.912	0.647	0.647	0.824	0.824	0.882	0.735	0.824	0.794	0.767
	H_E	0.916	0.867	0.939	0.919	0.936	0.906	0.884	0.921	0.827	0.823	0.904	0.882
	P	**0.000**	0.064	0.129	**0.000**	**0.000**	**0.004**	0.074	**0.004**	0.143	0.089	**0.036**	**0.000**
Atrato	Na	11.000	9.000	17.000	20.000	22.000	21.000	15.000	15.000	7.000	6.000	18.000	14.636
(30)	H_O	0.300	0.933	0.900	0.600	0.700	0.700	0.900	0.933	0.667	0.500	0.767	0.718
	H_E	0.817	0.849	0.933	0.945	0.946	0.946	0.912	0.920	0.849	0.788	0.933	0.879
	P	**0.000**	0.409	0.257	**0.000**	**0.000**	**0.000**	0.511	0.995	**0.010**	**0.003**	**0.002**	**0.000**
Fish Hatchery	Na	11.000	9.000	19.000	16.000	23.000	18.000	14.000	18.000	9.000	8.000	18.000	14.818
(40)	H_O	0.750	0.750	0.925	0.500	0.800	0.625	0.725	0.625	0.600	0.675	0.625	0.691

H_E	0.887	0.825	0.940	0.925	0.943	0.919	0.927	0.922	0.795	0.799	0.920	0.880
P	**0.030**	**0.014**	0.625	**0.000**	**0.007**	**0.000**	**0.001**	**0.000**	**0.000**	0.137	**0.000**	**0.000**

1

2

3

4

Table 3(on next page)

Genetic diversity and inbreeding coefficient of *Prochilodus magdalenae* per site and per genetic stock suggested by Structure in the main rivers of the distribution range of the species in Colombian hydrographic areas.

N: sample size; Na: average number of alleles per locus; H_O and H_E: observed and expected heterozygosity, respectively; F_{IS}: inbreeding coefficient; P: statistical significance for tests of departure from Hardy Weinberg equilibrium. Values in bold represent significance at P < 0.05. [1] Sampling site on the main stream. [2] Sampling site on floodplain lakes . S6a: Floodplain Lake Grande, S6b: Floodplain Lake Caimanera F, S6c: Guaranda.

1

River	Sampling Site (N)	Na	H_O	H_E	P	F_{IS}	P
Cauca	S1 (33)	15.273	0.667	0.878	**0.000**	0.255	**0.000**
	S2[1] (30)	15.727	0.773	0.885	**0.000**	0.143	**0.000**
	S3 (28)	14.182	0.740	0.886	**0.000**	0.182	**0.000**
	S4 (38)	14.818	0.732	0.885	**0.000**	0.186	**0.000**
	S5[1] (40)	15.636	0.700	0.885	**0.000**	0.221	**0.000**
	S6a[2] (34)	14.455	0.706	0.864	**0.000**	0.197	**0.000**
	S6b[2] (26)	14.364	0.752	0.881	**0.000**	0.165	**0.000**
	S6c (34)	15.364	0.719	0.879	**0.000**	0.196	**0.000**
	S8[2] (45)	15.909	0.743	0.887	**0.000**	0.173	**0.000**
Magdalena	Pijiño[2] (19)	12.273	0.780	0.865	**0.000**	0.125	**0.000**
	Mompox[1] (19)	13.091	0.770	0.882	**0.000**	0.154	**0.000**
	Palomino[1] (20)	13.182	0.759	0.869	**0.000**	0.152	**0.000**
	Río Viejo[2] (24)	13.909	0.739	0.883	**0.000**	0.184	**0.000**
	Llanito[2] (31)	15.000	0.774	0.879	**0.000**	0.135	**0.000**
	Barrancabermeja[1] (24)	13.636	0.727	0.872	**0.000**	0.186	**0.000**
	Chucurí (Ch)[2] (32)	15.000	0.699	0.882	**0.000**	0.223	**0.000**
	Puerto Berrío (B)[1] (28)	14.818	0.714	0.883	**0.000**	0.208	**0.000**
	Palagua (P)[2] (35)	17.182	0.792	0.895	**0.000**	0.129	**0.000**
	ChBP Stock1 (28)	13.000	0.698	0.851	**0.000**	0.198	**0.000**
	ChBP Stock2 (48)	18.636	0.759	0.895	**0.000**	0.162	**0.000**
	ChBP Stock3 (14)	9.909	0.695	0.833	**0.000**	0.202	**0.000**
Cauca + Magdalena-(ChBP)	Stock1 (241)	21.182	0.723	0.893	**0.000**	0.192	**0.000**
	Stock2 (285)	21.727	0.742	0.895	**0.000**	0.172	**0.000**
San Jorge	San Marcos River	13.364	0.782	0.873	**0.000**	0.130	**0.000**
Cesar	Mata de Palma	13.545	0.809	0.884	**0.000**	0.110	**0.000**
Nare	Samaná Norte River	15.630	0.659	0.876	**0.000**	0.260	**0.000**
Sinú	Caño Grande[1] (16)	11.000	0.744	0.845	0.000	0.151	0.000
	Doctrina[1] (18)	11.545	0.788	0.867	0.000	0.120	0.001
Atrato	Palo Blanco[1] (19)	12.727	0.722	0.869	**0.000**	0.195	**0.000**
	Beté[1] (11)	9.273	0.711	0.791	**0.000**	0.149	**0.000**

2

Table 4 (on next page)

Tests to detect recent genetic bottleneck in *Prochilodus magdalenae* populations.

IAM: infinite alleles model; SMM: stepwise mutation model; TPM: two-phase model. Wilcoxon test probability (one tail for H excess) (Luikart & Cornuet, 1998) calculated by Bottleneck v1.2.02 (Piry, Luikart, & Cornuet, 1999). M ratio value (Garza Williamson, 2001), calculated by Arlequin v3.5.2.2 (Excoffier, Laval, & Schneider, 2005).

River/Stock	IAM	SMM	TPM	M ratio value
Cauca (C)	**0.000**	0.958	**0.027**	0.254 ± 0.037
Magdalena (M)	**0.000**	0.517	**0.008**	0.219 ± 0.032
Sinú	**0.000**	0.183	**0.000**	0.155 ± 0.026
Atrato	**0.000**	0.584	0.062	0.151 ± 0.022
Fish Hatchery	**0.000**	0.382	**0.001**	0.173 ± 0.022
Chucurí (Ch)	**0.000**	0.232	**0.001**	0.156 ± 0.067
Puerto Berrío (B)	**0.000**	0.074	**0.000**	0.154 ± 0.067
Palagua (P)	**0.000**	0.740	**0.005**	0.175 ± 0.051
ChBP Stock1	**0.000**	0.958	0.103	0.160 ± 0.239
ChBP Stock2	**0.000**	0.551	**0.000**	0.228 ± 0.050
ChBP Stock3	**0.002**	0.551	0.160	0.126 ± 0.021
CM Stock1	**0.000**	0.997	**0.027**	0.240 ± 0.044
CM Stock2	**0.000**	0.966	**0.003**	0.245 ± 0.025

1

Table 5(on next page)

Parameters estimated using Bayesian likelihood method for searching candidate loci under selection

P: posterior probability of the model including selection; $Log_{10}(PO)$: the logarithm of posterior odds to base 10 for the model including selection; q-value: minimum false discovery rate at which a locus may become significant; Alpha: locus-specific component shared by all populations using a logistic regression, indicating the strength and direction of the selection; F_{ST}: coefficient to measure the difference in allele frequency between the common gene pool and each subpopulation, calculated as a posterior mean using model averaging.

1

Locus	P	$Log_{10}(PO)$	q-value	Alpha	F_{ST}
Pma39	0.883	0.880	0.017	-1.470	0.008
Pma25	0.987	1.890	0.002	-2.062	0.004
Pma02	0.999	3.220	0.000	-2.002	0.004
Pma35	0.998	2.660	0.000	-1.862	0.005
Pma01	0.122	-0.860	0.141	0.078	0.028
Pma40	1.000	1000	0.000	1.210	0.082
Pma46	0.048	-1.300	0.215	0.000	0.026
Pma36	1.000	1000	0.000	-2.589	0.002
Pma18	0.599	0.170	0.059	0.416	0.039
Pma13	1.000	1000	0.000	1.384	0.095
Pma14	1.000	1000	0.000	-2.116	0.004

2

Table 6(on next page)

Pairwise Jost's D_{est} (upper diagonal) and F'_{ST} (below diagonal) of *Prochilodus magdalenae* samples among rivers of the distribution range of the species in Colombian hydrographic areas.

Values in bold denote statistical significance after Bonferroni correction ($P < 0.002$).

River/Deme	1	2	3	4	5	6	7	8	9
1. Cauca		**0.065**	0.009	0.010	-0.003	0.020	**0.146**	**0.146**	**0.105**
2. Magdalena	**0.004**		**0.033**	0.052	0.047	**0.086**	**0.219**	**0.182**	**0.134**
3. Magdalena-ChBP	0.002	**0.003**		0.019	-0.007	0.013	**0.152**	**0.134**	**0.103**
4. Cesar	0.008	0.010	0.009		-0.010	0.025	**0.104**	**0.139**	0.042
5. San Jorge	0.008	0.010	0.008	0.014		0.007	**0.108**	**0.156**	**0.114**
6. Nare	0.005	**0.009**	0.006	0.013	0.012		**0.156**	**0.132**	**0.097**
7. Sinú	**0.013**	**0.016**	**0.014**	**0.017**	**0.018**	0.017		**0.202**	**0.209**
8. Atrato	**0.014**	**0.015**	**0.014**	**0.020**	**0.021**	0.017	0.021		**0.149**
9. Fish Hatchery	**0.010**	**0.011**	**0.011**	0.013	**0.018**	**0.014**	0.020	0.018	

1

2

3

Table 7(on next page)

Pairwise Jost's D_{est} (upper diagonal) and F'_{ST} (below diagonal) of *Prochilodus magdalenae* samples among sites of the rivers Cauca and Magdalena.

S6a: Floodplain Lake Grande, S6b: Floodplain Lake Caimanera F, S6c: Guaranda. Values in bold denote statistical significance after Bonferroni correction (Cauca: $P < 0.0005$; Magdalena: $P < 0.0001$).

River	Sampling site	1	2	3	4	5	6	7	8	9
Cauca	1. S1		0.023	0.069	0.050	0.000	0.066	0.014	0.056	0.023
	2. S2	0.011		0.059	0.006	0.007	0.020	0.020	0.036	-0.003
	3. S3	0.014	0.013		0.023	0.018	**0.096**	0.043	0.060	0.001
	4. S4	0.012	0.009	0.010		0.007	0.062	0.045	0.056	0.018
	5. S5	0.009	0.009	0.010	0.008		0.050	0.013	0.021	0.015
	6. S6a	0.013	0.010	**0.016**	0.012	0.011		0.052	0.038	0.044
	7. S6b	0.011	0.012	0.013	0.012	0.010	0.013		0.073	0.002
	8. S6c	0.013	0.011	0.013	0.012	0.010	0.011	0.014		0.003
	9. S8	0.009	0.008	0.009	0.008	0.008	0.010	0.009	0.008	
Magdalena	1. Pijiño		0.046	0.081	0.092	0.039	0.038	**0.414**	**0.387**	**0.312**
	2. Mompox	0.018		0.014	0.027	-0.001	0.006	**0.325**	**0.358**	**0.216**
	3. Palomino	0.020	0.016		0.082	0.006	-0.019	**0.416**	**0.373**	**0.273**
	4. Rio Viejo	0.019	0.015	0.018		-0.005	0.013	**0.400**	**0.411**	**0.277**
	5. Llanito	0.014	0.012	0.012	0.011		-0.041	**0.381**	**0.395**	**0.245**
	6. Barrancabermeja	0.016	0.014	0.012	0.013	0.008		**0.356**	**0.350**	**0.238**
	7. Chucurí	**0.036**	**0.029**	**0.035**	**0.032**	**0.031**	**0.031**		0.018	0.059
	8. Puerto Berrío	**0.035**	**0.031**	**0.033**	**0.033**	**0.032**	**0.031**	0.011		-0.006
	9. Palagua	**0.028**	**0.022**	**0.026**	**0.024**	**0.022**	**0.023**	0.012	0.009	

1

Figure 1

Prochilodus magdalenae sampling sites (stars) in the Colombian Magdalena-Cauca and Caribe hydrographic areas.

Cauca River: Antioquia Department: Bolombolo (1), Puente Real (2), Gurimán (3), Espíritu Santo River (4), Puerto Valdivia (5), Cáceres (6), Man River (7), Margento (8). Bolívar Department: Floodplain Lakes Grande (9), Caimanera (10) and Panela (13), Achí (12). Sucre Department: Guaranda (11). **Magdalena River**: Bolívar Department: Palomino (14), Mompox (16). Magdalena Department: Pijiño Floodplain Lake (15). Santander Department: Barrancabermeja (18), Floodplain Lakes Llanito (17), Chucurí (19), Río Viejo (20). Antioquia Department: Puerto Berrío (21). Boyacá Department: Palagua Floodplain Lake (22). **San Jorge River**: San Marcos River, Sucre Department (23). **Cesar River**: Mata de Palma Floodplain Lake, El Paso, Cesar Department (24). **Nare River**: Samaná Norte River, Antioquia Department (25). **Sinú River**: Córdoba Department: Caño Grande (26), Doctrina (27). **Atrato River**: Antioquia Department: Palo Blanco (29). Chocó Department: Beté (28). **Dams** (triangles): D1: Urra I, D2: Ríogrande, D3: San Lorenzo, D4: Playas, D5: El Peñol, D6: La Fe, D7: Miel, D8: Muña, D9: Calima, D10: Río Prado, D11: Betania. Map image layer by Instituto Geográfico Agustín Codazzi - IGAC.
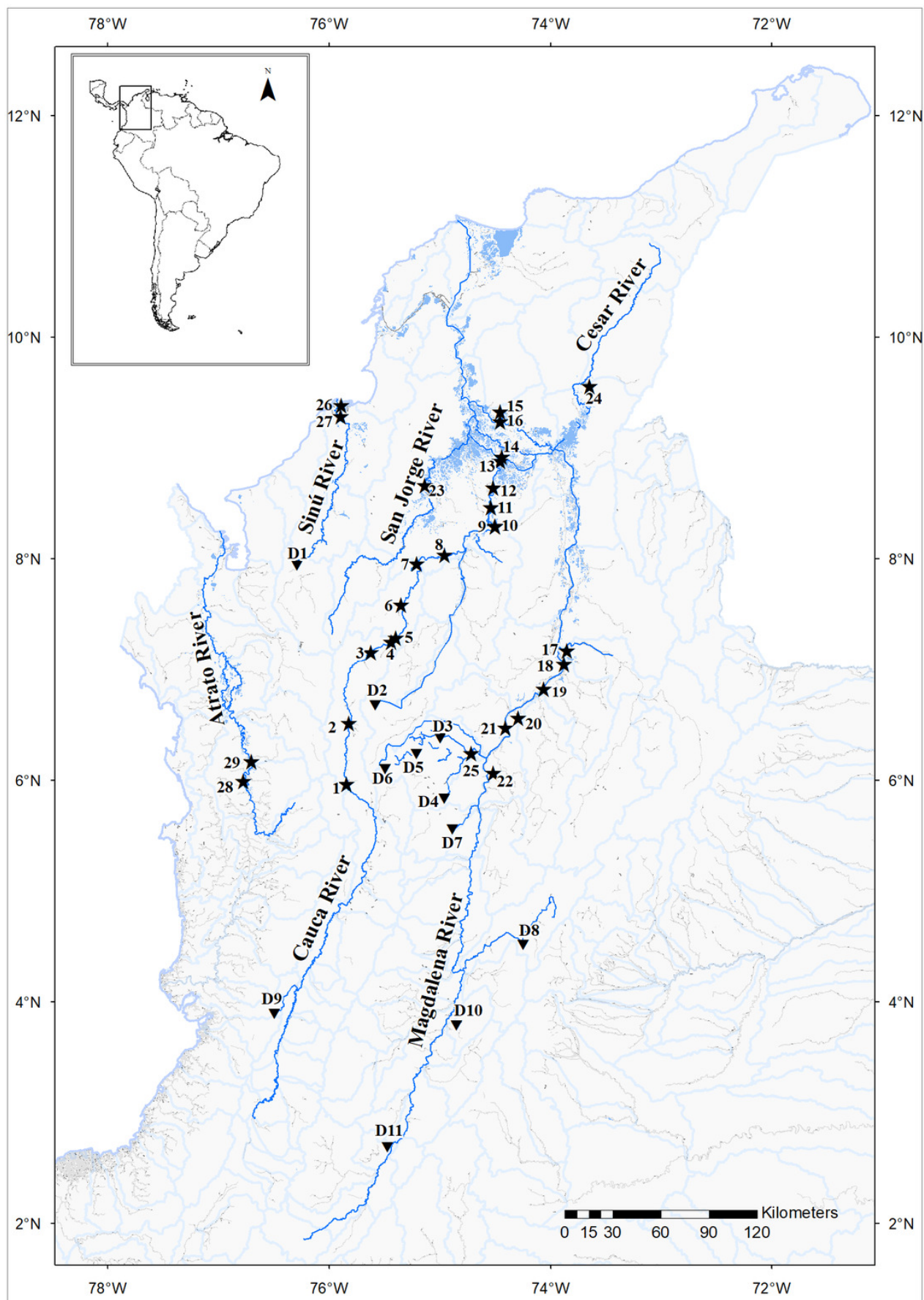
Figure 2

Bar plot of population ancestry coefficients as estimated by Structure (A) and discriminant analysis of principal components (B) of *Prochilodus magdalenae* from the Colombian hydrographic areas Magdalena-Cauca and Caribe.

CA: Cauca River; MG: Magdalena River; CS: Cesar River, SJ: San Jorge River; NA: Nare River; SU: Sinú River; AT: Atrato River; FH: fish hatchery. The numbers denote estimated genetic stocks (K).
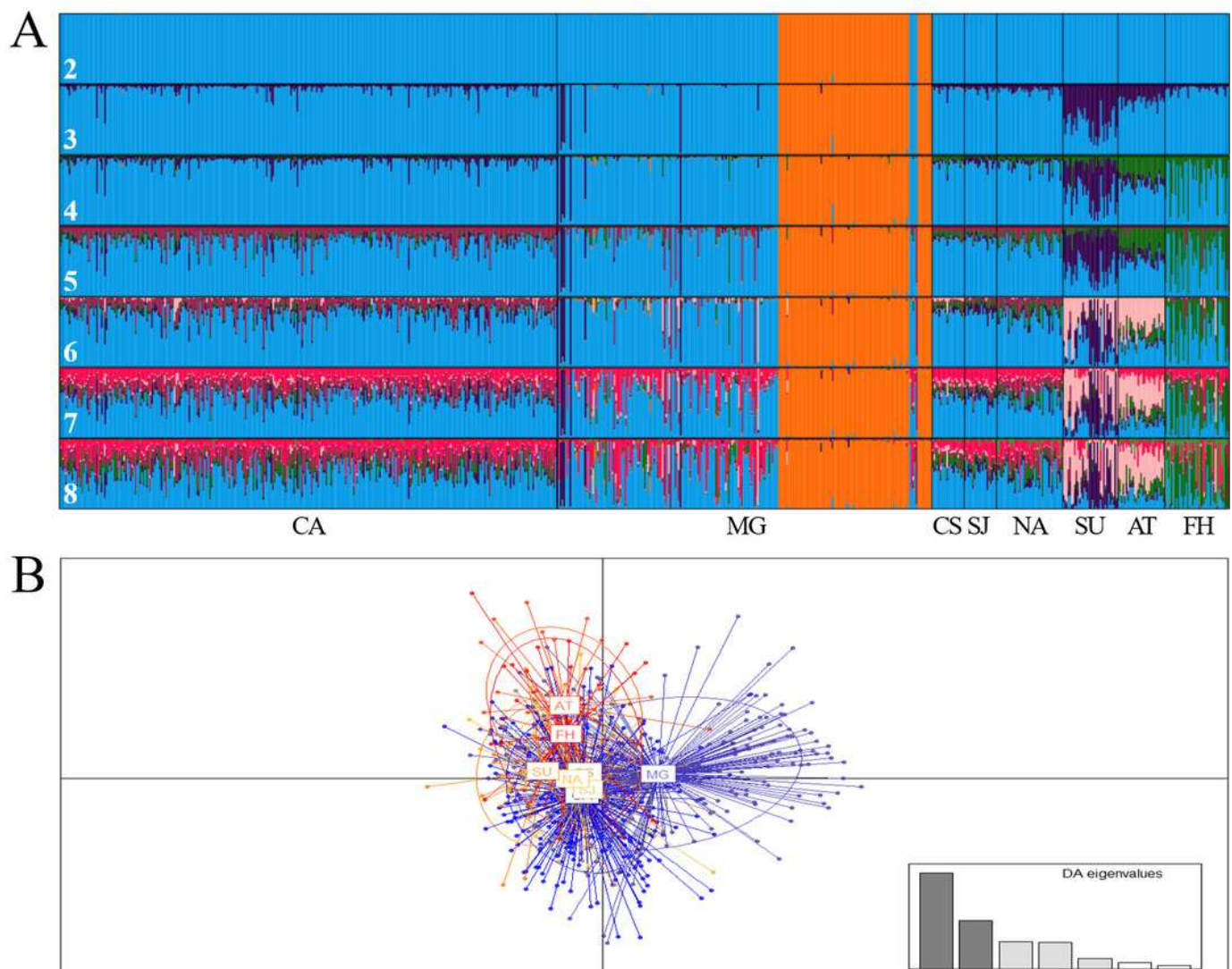
Figure 3

Bar plot of population ancestry coefficients as estimated by Structure and discriminant analysis of principal components of *Prochilodus magdalenae* in the Magdalena River and tributaries (A, B), Sinú River (C, D) and Atrato River (E, F).

CA: Cauca River; MG: Magdalena River; CS: Cesar River; SJ: San Jorge River; NA: Nare River.

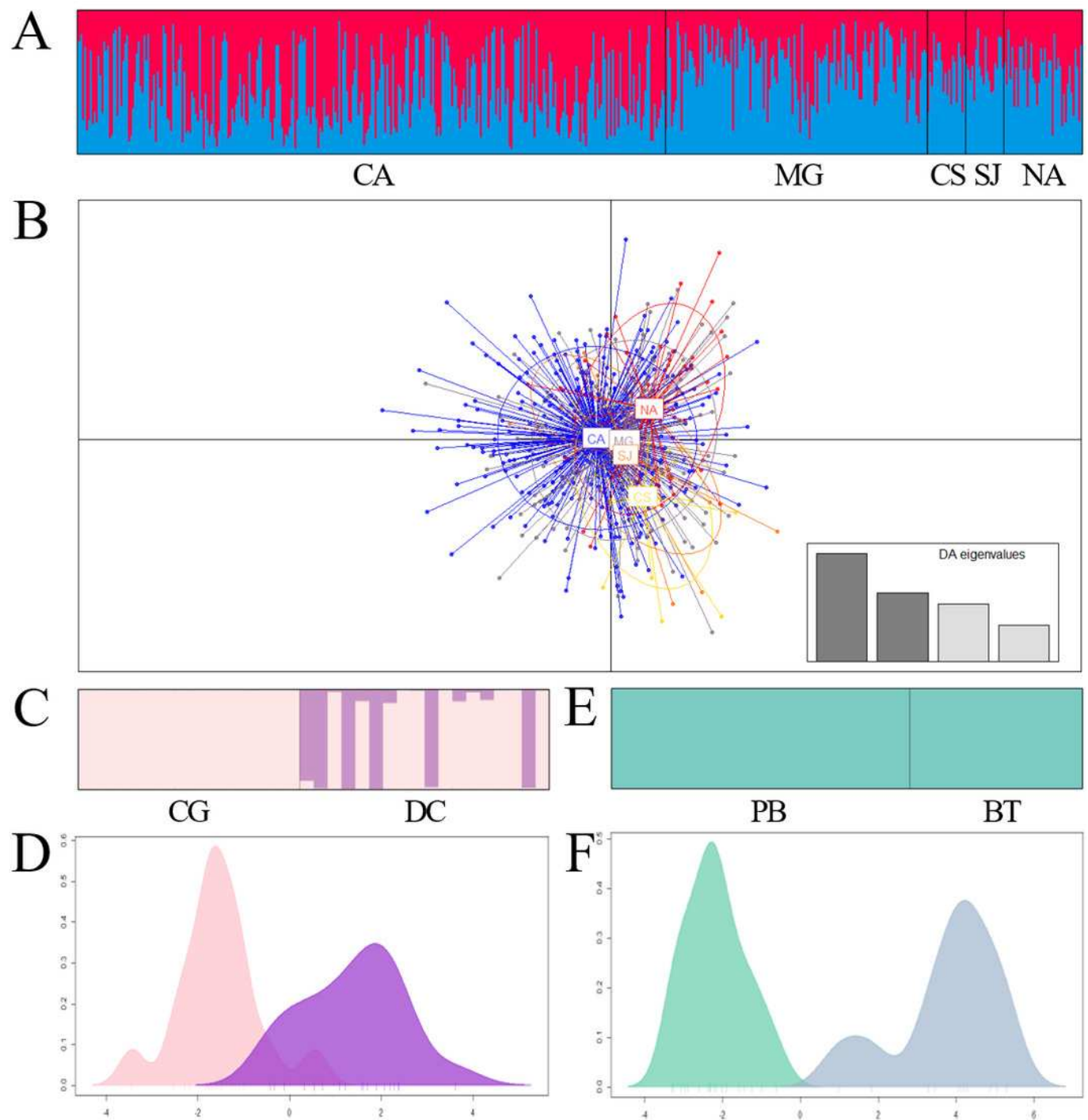CG: Caño Grande; DC: Doctrina; PB: Palo Blanco; BT: Beté.

Figure 4

Bayesian phylogenetic tree of *Prochilodus* based on partial sequences of *cox1* gene.

Node supports indicate posterior probability > 0.950. Red denote sequences from populations that exhibit outlier loci (Chucurí, Puerto Berrío, Palagua). Haplotypes are in parentheses. AT: Atrato River, CA: Cauca River; CS: Cesar River; MG: Magdalena River; SJ: San Jorge River; SU: Sinú River.