Species delimitation reveals an underestimated diversity of Andean catfishes of the family Astroblepidae (Teleostei: Siluriformes)

Luz E. Ochoa1,2, Bruno F. Melo1, Jorge E. García-Melo3,4, Javier A. Maldonado-Ocampo5, Camila S. Souza6, Juan G. Albornoz-Garzón4,6, Cristhian C. Conde-Saldaña2, Francisco Villa-Navarro4, Armando Ortega-Lara7 and Claudio Oliveira2

Catfishes of the family Astroblepidae form a group composed by 82 valid species of the genus Astroblepus inhabiting high-gradient streams and rivers throughout tropical portions of the Andean Cordillera. Little has been advanced in the systematics and biodiversity of astroblepids other than an unpublished thesis, a single regional multilocus study and isolated species descriptions. Here, we examined 208 specimens of Astroblepus that apparently belong to 16 valid species from several piedmont rivers from northern Colombia to southern Peru. Using three single-locus approaches for species delimitation in combination with a species tree analysis estimated from three mitochondrial genes, we identified a total of 25 well-delimited lineages including eight valid and 17 potential undescribed species distributed in two monophyletic groups: the Central Andes Clade, which contains 14 lineages from piedmont rivers of the Peruvian Amazon, and the Northern Andes Clade with 11 lineages from trans- and cis-Andean rivers of Colombia and Ecuador, including the Orinoco, Amazon, and Magdalena-Cauca basins and Pacific coastal drainages. Results of species delimitation methods highlight several taxonomical incongruences in recently described species denoting potential synonymies.

Keywords: Andes, Catfishes, Delimitation, Ostariophysi, Systematics, Taxonomy.
Los bagres de la familia Astroblepidae son un grupo compuesto por 82 especies válidas del género *Astroblepus* que habitan quebradas y ríos de alto gradiente a través de la porción tropical en la Cordillera de los Andes. Poco se ha avanzado en la sistemática y biodiversidad de los astroblepidos aparte de una tesis no publicada, un único estudio multilocus regional y descripciones aisladas de especies. Aquí, examinamos 208 especímenes de *Astroblepus* que aparentemente pertenecen a 16 especies válidas provenientes de ríos de pie de monte de la cordillera de los Andes, desde el norte de Colombia hasta el sur de Perú. Utilizando tres metodologías de delimitación de especies para un único locus en combinación con análisis de un árbol de especies a partir de tres genes mitocondriales, identificamos un total de 25 linajes bien definidos que incluyen ocho especies válidas y 17 potenciales especies no descritas distribuidas en dos grupos monofiléticos: un clado de los andes centrales, que contiene 14 especies de los ríos de pie de monte de la Amazonía peruana y un clado de los andes del norte con 11 especies de los ríos trans y cisandinos de Colombia y Ecuador, incluyendo las cuencas del Orinoco, Amazonas y Magdalena-Cauca así como drenajes costeros del Pacífico. Los resultados de los métodos de delimitación de especies destacan varias incongruencias taxonómicas en especies recientemente descritas que denotan posibles sinonimias.

**Palabras clave:** Andes, Bagres, Delimitación, Ostariophysi, Sistemática, Taxonomía.

**INTRODUCTION**

The family Astroblepidae includes 82 valid species within the single genus *Astroblepus* Humboldt, 1805 (Fricke et al., 2020) characterized by a specialized head and body morphology, with the presence of a sucker-shaped mouth with fleshy and expanded lips and a dorsal opening to the gill chamber, between the dorsal margin of the opercle and the ventral edge of the pterotic (Schaefer, 2003). Species of this genus have bodies without bony plates and with a modified pelvic musculature completely separate from the hypaxial muscles, forming a pair of strings-like muscles between the pectoral and pelvic fins (the *protractor ischii* muscles; Shelden, 1937). Alternating contraction of the *protractor* and *ischii* muscles anterior of the pelvic girdle and the *retractor ischii* muscles posterior of the pelvic girdle, combined with adhesion to substrates using the oral disk results in a remarkable ability of *Astroblepus* to move forward against very strong currents and even scale vertical surfaces (Shelden, 1937; De Crop et al., 2013), thus giving this clade the name “climbing catfishes” (Schaefer, Arroyave, 2010).

Astroblepids inhabit moderate to high elevation streams on eastern and western flanks of the Andes Mountains in Panama, Colombia, Venezuela, Ecuador, Peru and Bolivia (Schaefer, 2003). Species of astroblepids have generally restricted geographic distributions, being limited to small portions of isolated drainages between 400 and 4000 m, above sea level, but less frequent above 3,000 m (Schaefer, 2003; Schaefer, Arroyave, 2010). Distribution and ecological niche modeling analyses have demonstrated that astroblepids are physiologically confined to drainage islands in the Andes, where elevation and temperature oscillation strongly influence their geographical
distribution and endemism (Vélez-Espino, 2006; Schaefer, Arroyave, 2010). Ecological specializations and limited habitat distributions make Astroblepus vulnerable to many environmental impacts such as water pollution, deforestation, species introductions and urbanization (Anderson, Maldonado-Ocampo, 2011), e.g., Astroblepus vaillanti (Regan, 1904) in Ecuador (Alexiades, Encalada, 2017).

Much of the research on Astroblepus has been focused on either their ecology and life history (Vélez-Espino, 2003; Briñez-Vásquez et al., 2005; Vélez-Espino, Fox, 2005; Schaefer, Arroyave, 2010), cytogenetic characterization (Conde-Saldaña et al., 2019), functional morphology (De Crop et al., 2013), or their molecular and physiological evolution (Castiglione et al., 2017). However, a few studies have focused on the species-level systematics and taxonomic diversity of astroblepids, such as the single taxonomic revision of the genus (Regan, 1904), an unpublished dissertation describing the osteology of some Colombian species (Buitrago-Suárez, 1995), a molecular study of endemics from southern Peru and northern Bolivia (Schaefer et al., 2011), and various isolated species descriptions (Cardona, Guerao, 1994; Ardila-Rodriguez, 2011a,b, 2012, 2013a,b, 2014, 2015a,b). However, taxonomic limits among species of Astroblepus are questionable and problematic. High levels of inter and intraspecific variation in body shape, fin sizes and body coloration further contribute to the taxonomic complexity of the group (Schaefer et al., 2011). Most of the early descriptions of Astroblepus species were brief on descriptive details (Regan, 1904; Eigenmann, 1917) and diagnoses in the last years were based in a single-character or a combination of distinct and labile pigmentation patterns. Morphological characters, such as meristic features, coloration patterns and presence or absence of adipose fin do not allow to distinguish astroblepids, and it has been assumed that future taxonomic revisions including additional characters will reduce the overestimated number of valid species recognized nowadays (Schaefer, 2003). Although specimens of this group are abundant in ichthyology collections in the Andean region, a large proportion of the individuals are misidentified, unidentified, or identified with taxonomic uncertainty (e.g., species cataloged as ‘sp.’, ‘aff.’ and ‘cf.’) (Hernández-Zapata et al., 2018).

The most recent phylogenetic study combined mitochondrial and nuclear loci to investigate nine morphospecies of Astroblepus from northern Bolivia and southern Peru (Schaefer et al., 2011). Results supported seven species and the authors argued that allopatric species from distant drainages are more related to each other than to syntopic species or species from adjacent drainages (i.e., disjunct distribution of sister taxa). Despite relatively low species coverage (~11%), their study produced a valuable molecular dataset that can be used to incorporate additional species from other Andean regions. Similarly, a growing number of molecular studies have examined species-level diversity in Neotropical freshwater fishes (e.g., Pereira et al., 2013; Costa-Silva et al., 2015; Machado et al., 2018; García-Melo et al., 2019), however none have focused on astroblepids. Besides, various multilocus, coalescent-based approaches have yielded better resolution of species limits in morphologically cryptic radiations across different organisms (e.g., Barley et al., 2013; Myers et al., 2013; Satler et al., 2013; Welton et al., 2013) including freshwater fishes (Niemiller et al., 2012; Bagley et al., 2015).

Here, we combined newly generated mitochondrial sequences with those previously generated by Schaefer et al. (2011) to study species diversity, delimit species boundaries within Astroblepus from central and northern Andes in Colombia, Ecuador and Peru,
and to test Schaefer et al. (2011) hypothesis of the disjunct distribution of sister taxa along the Andean Cordillera.

MATERIAL AND METHODS

Taxon sampling. Sequences of a total of 172 specimens of *Astroblepus* were newly generated, in which 77 were previously identified in 16 valid species representing 20% of species diversity and, the remaining 95 specimens were identified at the genus level, as “aff.” or “cf.”. Additionally, we included 36 sequences available from GenBank for eight morphospecies identified with the letters A-I as delimited by the molecular study of Schaefer et al. (2011). Four species of Loricariidae were included as outgroup taxa: *Farlowella nattereri* Steindachner, 1910, *Lamontichthys stibaros* Isbrücker & Nijssen, 1978, *Loricaria similima* Regan, 1904, and *Pterygoplichthys multiradiatus* (Hancock, 1828); specimens of Loricariidae were chosen due to previous phylogenies indicating the sister relationship with Astroblepidae (de Pinna 1998; Lujan et al., 2015; Roxo et al., 2019). Specimens were either collected during the course of this study or available in museum collections. Voucher identifications were determined according to original descriptions and geographic distributions, with the exception of eight morphospecies proposed by Schaefer et al. (2011), whose specimens were unavailable for examination in this study. Unfortunately, samples from intervening or adjacent regions (e.g., Ecuador, Bolivia) were also limited or unavailable here. Vouchers are deposited in various ichthyology collections: the Academy of Natural Sciences of Drexel University, Philadelphia, USA (ANSP), Auburn University Museum of Natural History, Auburn, USA (AUM), Colección Zoológica de la Universidad del Tolima, Ibagué, Colombia (CZUT-IC), Instituto de Ciencias Naturales, Museo de Historia Natural, Universidad Nacional de Colombia, Bogotá, Colombia (ICNMHN), Instituto de Investigación Recursos Biológicos Alexander von Humboldt, Villa de Leyva, Colombia (IAvH-P), Laboratório de Biologia e Genética de Peixes, Universidade Estadual Paulista, Botucatu, Brazil (LBP), Museo Javeriano de Historia Natural “Lorenzo Uribe”, Bogotá, Colombia (MPUJ) and Royal Ontario Museum, Toronto, Canada (ROM). Catalog numbers, localities and GenBank accession numbers of all terminals are summarized in S1. Photographs for representative species are included in the Fig. 1, and geographical distribution of *Astroblepus* species and samples included in this study are shown in the Fig. 2.

DNA extraction and sequencing. DNA was extracted from tissues preserved in 95% EtOH using a silica-based method (Ivanova et al., 2006). Partial sequences of three mitochondrial genes 16S rRNA (16S), cytochrome oxidase c subunit I (COI) and cytochrome b (Cytb) were amplified by polymerase chain reaction (PCR) with the primers described in S2. Amplifications were performed in a total volume of 12.5 μl with 1.25 μl of 10X buffer (10 mM Tris–HCl+15 mM MgCl2), 0.5 μl dNTPs (200 nM of each), 0.5 μl each 5 mM primer, 0.05 μl Platinum® Taq Polymerase (Invitrogen), 1 μl genomic DNA (10–50 ng), and 8.7 μl ddH2O. The thermo-cycler profile consisted of an initial denaturation (4 min at 95°C) followed by 30 cycles of chain denaturation (30 s at 95°C), primer hybridization (30–60 s at 52°C for 16S and Cytb, 54°C for COI), nucleotide extension (30–60 s at 72°C), and a final extension (10 min at 72°C). We
FIGURE 1 | Species of *Astroblepus* included in this study, A. *A. ardiaduartei* (LBP 26696 topotype live, 4.54 mm SL), B. *A. cachara* (LBP 26712 topotype live, 4.23 mm SL), C. *A. caquetae* (CZUT-IC 18464 topotype of museum, 7.84 mm SL), D. *A. curitensis* (LBP 97118 topotype live, 5.92 mm SL), E. *A. homodon* (CZUT-IC 18390, 6.15 mm SL), F. A. gr. *grixalvii* (LBP24242 topotype live, 11.70 mm SL); F’. A. gr. *grixalvii* (CZUT-IC 18498 specimen of Magdalena basin 6.01 mm SL); F”. A. gr. *grixalvii* (CZUT-IC 18320 specimen of Cauca basin, 15.25 mm SL), G. *A. itae* (topotype live, 3.58 mm SL), H. *A. latidens* (topotype live, 13.40 mm SL), I. *A. onzaagenensis* (topotype live, 7.82 mm SL), J. *A. pradai* (topotype live, 4.53 mm SL), K. *A. trifasciatus* (topotype of museum, 9.65 mm SL), K’. *A. trifasciatus* (topotype of museum, 9.01 mm SL), L. *A. aff. trifasciatus* (specimen of Magdalena basin, 7.94 mm SL), M. *A. verai* (topotype live, 3.51 mm SL).
observed heterozygosity in seven individuals in both forward and reverse sequences of COI indicating a putative co-amplification of a nuclear pseudogene. In this case, the primers L6252-Asn and H7271-COXI (Melo et al., 2011) were used and increased annealing temperature (56°C) in the PCR reactions.

All PCR products were first visually identified on 1% agarose gel and purified using ExoSap-IT® (USB Corporation) following manufacturer instructions. The purified PCR products were sequenced using the Big DyeTM Terminator v 3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems), purified again by ethanol...
precipitation and loaded onto an automatic sequencer 3130-Generic Analyzer (Applied Biosystems) in the Instituto de Biociências, Universidade Estadual Paulista, Botucatu, Brazil.

**Alignment and species delimitation analyses.** Consensus sequences were assembled and edited in Geneious 7.1.4 (Kearse et al., 2012) and aligned using the MUSCLE algorithm (Edgar, 2004). We estimated the index of substitution saturation (Iss) (Xia et al., 2003; Xia, Lemey, 2009) in DAMBE v6 (Xia, 2017). Three independent matrices were assembled for each gene and nucleotide substitution models were selected according to the Bayesian Information Criterion (BIC) in the test module of MEGA v7 (Kumar et al., 2016).

Because the morphology-based identification of *Astroblepus* species is uncertain and difficult, we identified genetic lineages using first a single locus COI applying three species delimitation approaches. The first method correspond with a Bayesian implementation of the Poisson Tree Process model (bPTP, Zhang et al., 2013), using the best maximum likelihood (ML) tree as input file in the PTP web server (http:// species.h-its.org) with default settings. The best ML tree was estimated with RAxML PTHREADS–SSE3 implemented in RAxML v8.019 (Stamatakis, 2014), which executed five searches with five randomized maximum-parsimony topologies using the GTRGAMMA model and 1,000 bootstrap pseudoreplicates by the autoMRE bootstopping criterion (see Pattengale et al., 2010).

The Generalized Mixed Yule–Coalescent model (GMYC) was the second method implemented (Pons et al., 2006; Fujisawa, Barraclough, 2013) that aims to discern stochastic birth–death processes between species, from neutral coalescent processes within species, analyzing time intervals between branching events in a single time-calibrated gene tree (Pons et al., 2006). For this analysis, we estimated a calibrated tree using the lognormal relaxed molecular clock model, which assumes that the rates of molecular evolution are uncorrelated, but log-normally distributed among lineages as implemented in BEAST v1.8.2 (Drummond, Rambaut, 2007). We obtained an ultrametric gene tree with two independent runs of 30 million generations with one tree sampled every 1000th generation. BEAST log files were examined in Tracer v1.6 (Rambaut et al., 2018) to access both stationary phase and parameter convergence. A maximum clade credibility tree was estimated in TreeAnnotator v1.8.1 based on the sampled trees after discarding the first 25% generations as a burn-in procedure. All likelihood and Bayesian analyses were performed on 2x 40 CPU 128GB Brycon server at LBP/UNESP. The ultrametric gene tree was used in the GMYC analysis conducted with standard parameters (interval = c (0,10)) using a single threshold via the splits package in R v1.3.1056 (http://r–forge.r– project.org/projects/splits) on a single gene tree.

The third method correspond with the Automatic Barcode Gap Discovery (ABGD; Puillandre et al., 2012) that estimates divergence gaps in the interspecific comparison relative to the amount of intraspecific variation. We used a Pmax at 0.15 and a relative gap width of 1.5, with the Jukes–Cantor model (JC69) and other parameters at default in the ABGD webserver (wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html). In summary, the COI lineages identified by these methods were used in downstream concatenated analyses.

Using a concatenated matrix, we estimated a maximum likelihood tree in RAxML v8.019 (Stamatakis, 2014) with GTRGAMMA model and 1,000 bootstrap
Pseudoreplicates using the autoMRE bootstrapping criterion (Pattengale et al., 2010). Afterward, we assigned COI lineages to species groups and estimated a species tree using the multispecies coalescent method *BEAST* implemented in BEAST2 (Bouckaert et al., 2014). We performed two BEAST2 runs of 350 million generations each, sampling every 100,000 generations and with other parameters at default. Log files from each run were checked for convergence and mixing and ensured that ESS scores were >200 in Tracer v1.4. Then, we implemented a species delimitation analysis using a concatenated matrix (COI, 16S and Cytb) to test this groupings on a multispecies coalescent model in the program Bayesian Phylogenetics and Phylogeography (BP&P) v3.2 (Rannala, Yang, 2013; Yang, Rannala, 2010, 2014; Yang, 2015). The BP&P method incorporates both the coalescent theory and the phylogenetic uncertainty, accounts for gene tree variance and incomplete lineage sorting (ILS), and calculates mutation-scaled population size (θ) and divergence time (τ) estimates; it also assumes that no gene flow occurs following speciation, an analogy to the biological species concept (Mayr, 1942). The algorithm explores different species delimitation models and different species phylogenies with fixed specimen assignments to populations. We ran the A10 analysis of BP&P, a species delimitation using a user-specified guide tree (Yang, 2015), based on the concatenated dataset partitioned by gene, used the BEAST2 species tree as a guide tree, and specifying a Dirichlet distribution (α = 2) to account for variation in mutation rates among loci. We conducted runs using three different combinations of gamma-distributed priors for ancestral θ with a default prior θ ~ G(3, 0.004), a diffuse prior θ ~ G(3,0.002) and an estimated prior from sequences θ ~ G(3, 0.003) and the same values were used for (τ_0). The program implements two different τjmCMC algorithms tested for each species delimitation hypothesis, and the program auto-adjusted the fine-tuning parameters. We conservatively accepted daughter lineages from nodes with speciation probabilities >0.95 across all three priors as strongly supported species.

**RESULTS**

The molecular dataset included 208 sequences of COI (534 bp), 201 sequences of 16S (495 pb) and 177 sequences of Cytb (489 pb). All matrices did not present insertions, deletions, stop codons or sequencing errors due to cross-contamination. The Iss.c values were greater than Iss values, indicating the absence of saturation in all three datasets. The concatenated dataset included a total of 215 terminals and 1,517 pb.

The single-locus approaches revealed a combined result of 42 COI lineages of which 11 belong to valid species, eight correspond to the morphospecies reported by Schaefer et al. (2011) and the 23 remaining represent unidentified lineages (Fig. 3). The ABGD analysis delimited 42 lineages (genetic distances 0.010–0.220), the tree-based bPTP analysis delimited 36 lineages (25% of species >0.95 bootstrap), and the GMYC analysis delimited 88 lineages ($L_0=1988.87$, a value not significantly better than the null model of coalescent branching rates $L_0=1977.87$). The later method delimited some entities represented by one or two individuals with very low genetic distances (0.000). Fourteen clusters were coincident among all methods, nine were coincident only between ABGD and PTP, 13 between ABGD and GMYC and two clusters between bPTP and GMYC (Fig. 3). According to these results, the preliminary species
delimitation included 14 clusters coincident between three methods and 38 clusters recovered in at least two methods that, when compared to alternative groupings in the third approach, continue with interspecific genetic distance equal to 0. The overall mean genetic distance between the putative species was 0.108±0.009 and interspecific genetic distance values ranged from 0.009 ± 0.004 between *Astroblepus* sp. 28 and *A. micrescens* Eigenmann, 1918 to 0.187±0.022 between *Astroblepus* sp. A and *A. trifasciatus* (Eigenmann, 1912). The final number of 42 COI lineages of *Astroblepus* was used as an input prior for the multigene analysis.

Maximum likelihood tree of the concatenated matrix with nodal support is shown in S3 and the species tree (Fig. 4) presented highest posterior probability (PP=1) (ln L = -19303.083, ESS=211) for 73% of the COI lineages with 17 showing low support.

**FIGURE 3** Results of single-locus approaches using cytochrome oxidase c subunit I (COI) for developing preliminary species delimitation hypothesis with 42 lineages. Results are represented on the ultrametric gene tree with collapsed nodes. All nodal support values were PP>0.95. Blocks at right of the tree represent hypothesized species groups and the values in the middle indicate the number of clusters identified by ABGD, bPTP and GMYC analyses for every collapsed node. COL: Colombia, ECU: Ecuador, PER: Peru.
The BP&P analysis supported only 59.5% of these lineages with a total of 25 species (PP = 1) recovering seven of 11 valid species and 18 from 23 unidentified lineages delimited in the singles-locus analysis (Fig. 4). The lineages supported were consistent in every set of priors used (S4, S5 and S6). The topology clearly shows two strongly-supported monophyletic groups that are well-defined biogeographically: the herein named “Northern Andes clade” with 23 lineages predominantly from Colombia, and the “Central Andes clade” with 19 lineages from Peru and Ecuador identified in the COI analysis (Fig. 4).


**FIGURE 4** Species tree inferred from the concatenated dataset of mitochondrial genes (COI, Cytb, and 16S). Nodal support values are Bayesian posterior probabilities. Non-significant speciation probabilities identified in BP&P analysis algorithm A10 (PP<0.95) are indicated by black circles and species supported with asterisk.
Astroblepus sp. 22, 23, 25, 26, 27, 28, and 29, potential new species (Fig. 4, black circles). In the Central Andes Clade, the analysis supported 14 lineages and reveal low support for morphospecies A and I, as well as Astroblepus sp. 9, 12, and 13 from piedmont rivers of the Peruvian Amazon.

**Northern and Central Andes clades.** The Northern Andes clade includes 11 species from Colombia, except *Astroblepus* sp. 20 from the Río Almendro in the northern Peru and *Astrobleopus* cf. *regani* from the Río Wiñacay in Ecuador. This clade also includes one species from piedmont rivers of the Orinoco basin (*Astrobleopus* sp. 21), one species from the Caquetá river, upper Amazon basin (*A. caquetae* Fowler, 1943), one species from the Río Dagua, a coastal Pacific river (*A. trifasciatus*), and six species from the Magdalena–Cauca basin (*A. grixalvii* Humboldt, 1805, *A. cacharas* Ardila–Rodriguez, 2011, *A. curitensis* Ardila–Rodriguez, 2015, *A. homodon* (Regan, 1904), *A. nicefori* Myers, 1932, and *Astrobleopus* sp. 24). The *A. grixalvii* group includes toptotypes of the type species of the genus, *A. grixalvii* from Río Palace (upper Río Cauca) and specimens from several Colombian drainages such as upper Magdalena basin (Ríos Guarapas, Saldaña, Coello, Opia, Totare, and Sumapáz), middle Magdalena basin (Río Samaná), upper Cauca basin (Río Quindío), coastal Pacific rivers (Río San Juan and Río Patía), and the cis–Andean piedmont rivers of the Orinoco basin (Río Negro). The *A. pradai* Ardila–Rodriguez, 2015 lineage contained three nominal species (*A. pradai*, *A. curitensis*, and *A. floridablancaensis* Ardila–Rodriguez, 2016) from tributaries of Río Sogamoso and Río Lebrija (middle Magdalena basin); the *A. cacharas* lineage includes two nominal species (*A. cacharas* and *A. verai* Ardila–Rodriguez, 2015) from tributaries of Río Lebrija and Río Sogamoso (middle Magdalena basin), and the *A. itae* lineage includes *A. itae* and *A. ardidahuartii* Ardila–Rodriguez, 2015 from tributaries of Río Lebrija (middle Magdalena basin). The species tree recovered *A. trifasciatus* (Río Dagua, Pacific drainage) as sister to the remaining taxa, and then *A. caquetae* and *Astrobleopus* sp. 20 from the upper Amazonas (Río Caqueta and Río Almendro, respectively), with this clade being sister–group to all other remaining species of the Northern Andes clade (Fig. 4).

The Central Andes clade includes 14 species: two from the Madre de Dios drainage, a tributary of the Madeira basin (morphospecies B and C), 10 species from the Marañon–Ucayali system, including the ríos Apurimac, Urubamba, Chorobamba, and Pozuzo (morphospecies D, E, F–G, H and *Astrobleopus* sp. 10, 11, 14, 15, 16, and 17), and two species from Río Wiñacay, a Pacific coastal drainage of Ecuador (*Astrobleopus* sp. 18 and 19). The nine morphospecies (A–I) of Schaefer et al. (2011) represented eight delimited species (i.e., their morphospecies F and G were delimited as a single species). *Astrobleopus* cf. *regani* and *Astrobleopus* sp. 20 (Northern Andes clade) and *Astrobleopus* sp. 13, 17, 18, and 19 (Central Andes clade) occur either in southern Ecuador or northern Peru, representing the boundaries of the two major clades, and likely an intersection region for the diversification of *Astrobleopus* (Fig. 4).

Additionally, most specimens that could not be allocated into valid species (“sp.”) or those showing morphological similarity with valid species identified in the categories as “aff.” and “cf.” were successfully grouped with valid species and potential species with high support values for the clusters and low genetic divergences (S3). From a total of 208 specimens included in this study, the species tree and species validation in a coalescent multilocus analysis allowed the identification of 97 individuals assigned to the categories
“sp.”, “aff.” and “cf.”, as well as supported the eight morphotypes identified previously by Schaefer et al. (2011).

**DISCUSSION**

The present study identified 25 well delimited lineages of *Astroblepus*, eight of which being allocated to valid species (9.7% of species diversity) and the other 17 corresponding to distinct genetic lineages and potential undescribed species. These results indicate an underestimated molecular diversity and represent a first step to improve the taxonomic, phylogenetic and biogeographic understanding of astroblepids. Results represent an expansion of the taxon sampling relative to the previous multilocus analysis of *Astroblepus* (e.g., Schaefer et al., 2011), which proposed eight morphotypes of *Astroblepus* from the piedmont of the upper Río Madre de Dios and tributaries of the upper Marañon–Ucayali system, all major tributaries of the Amazon basin in Peru. Our results agree with the relationships among those morphotypes.

Although astroblepids are morphologically unmistakable regarding to other catfishes by having specialized mouth morphology and pelvic-fin musculature, the species identification is very complex. Regan (1904) already highlighted morphological differences in the adipose fin due to size, individual variation and preservation of specimens. Morphological characters currently used to distinguish species of *Astroblepus*, such as tooth type, presence/absence of adipose spine, body shape, fin size and pigmentation pattern have extreme intraspecific variation (Schaefer et al., 2011). Most described species in the last decade are supported by single-character diagnoses or color pattern. These ambiguous diagnoses with excessive intraspecific variation have generated confusion due to the fact that species were described without a comprehensive taxonomic or phylogenetic support (Ardila-Rodriguez, 2011a,b, 2012, 2013a,b, 2014, 2015a,b).

Several of the species diagnosed by those labile characters appeared as a single genetic lineage in our study. For example, our analysis delimited as a single species, topotypes representatives of *Astroblepus curitiensi*, *A. pradai* and *A. floridablancaensis*, described from tributaries of the same hydrographic basin (Río Sogamoso, middle Magdalena basin) with similar pigmentation patterns of body and fins, barbel length and oral disc diameter but diagnosed by variations in the morphological structure of the adipose fin (a denticulate imbedded spine in *A. pradai* vs. a well-developed adipose fin without spine in *A. curitiensi* vs. an adipose fin composed of spine posteriorly attached by a membrane to the caudal peduncle in *A. floridablancaensis*) (Ardila–Rodriguez, 2015b, 2016). Likewise, our molecular results found a single species including *A. cacharas* from Río Cachira and *A. verai* from Quebrada del Medio, a tributary of Río Chucuri (Ardila–Rodriguez, 2011a, 2015b). These species were described by a combination of features with high intraspecific variation, such as configuration of the adipose fin, pigmentation pattern and meristic data. While some differences can be detected between specimens of these species, such as length of the maxillary barbel (extending beyond the oral disc in *A. cacharas* vs. short maxillary barbel not surpassing the oral disc in *A. verai*) and shape of the outer row of premaxillary teeth (10–11 conic and one–two bicuspid in *A. cacharas* vs. 7–8 conic and one bicuspid in *A. verai*), the coloration pattern and size of oral disc clearly overlap.
In addition to taxonomic issues detected in recently described species discussed above, our results strongly support the recognition of two biogeographically structured clades of *Astroblepus*. The Northern Andes clade contains species mostly from Magdalena-Cauca, Orinoco and Caquetá (upper Amazon) basins in Colombia, including the type species of the genus, *Astroblepus grixalvii*, along with specimens from several Colombian regions (upper and middle Magdalena basin, Patía basin and piedmont Orinoco). This clade composes the large *A. grixalvii* species complex, as better delimited herein. Morphological studies of *A. grixalvii* are limited to an unpublished dissertation (Briñez-Vásquez, 2011) in which the author identified morphological variation among isolated populations of the species in the Magdalena and Cauca rivers associated with discrepancies in habitat conditions. This result is supported by the presence of various lineages within the complex (S3). As one of the most widespread species of *Astroblepus*, it is expected that more species within the *A. grixalvii* complex will be recognized and described afterwards.

Although a detailed biogeographic analysis is beyond the scope of this paper, the structured pattern of Central and Northern Andes clades is an interesting result. Lujan et al. (2015) found a similar biogeographic pattern in loricariids of the *Chaetostoma* clade, that includes two northern Andean clades as sister to a group widely distributed in the central-southern Andes. Similarly, a biogeographic analysis using species distribution models for endemic birds in the tropical region, identified different events of isolation and diversification associated with geographical barriers, being one of the most important the Marañon River valley in the northern Peru. This barrier was responsible to isolate the Central and Northern Andean taxa since the Miocene splitting the areas in four bioregions: the Central Cordillera of Peru, the Eastern Side of the Central Andes, the Northern Central Cordillera and the Tumbesian Mountains (Hazzi et al., 2018). According to those authors, one of the primary drivers of the biogeographic distribution was the formation of sizeable warm river valleys during periods of high Andean uplift, isolating species in the highland ecosystems (Hazzi et al., 2018). A central-northern pattern was also observed in Neotropical plants of the family Rubiaceae (Antonelli et al., 2009). Therefore, astroblepids represent an excellent model to further study how these processes influenced the distribution of freshwater fishes in the Andean region.

The present study also supports the hypothesis that most astroblepids are endemic to single or adjacent rivers of the Andes and that their distribution generally does not cross major headwater divides (Schaefer et al., 2011), despite broader distributions of some lineages such as *Astroblepus* sp. D, E, and F and within the *A. grixalvii* species complex. Our results also corroborate the hypothesis of disjunct distribution of sister taxa of Astroblepidae (Schaefer et al., 2011). This can be clearly observed in species from the Pacific region of Ecuador and Colombia that are phylogenetically closer to species from Magdalena-Cauca, Orinoco, and western Amazon systems, forming a reticulate pattern of diversification, even separated by three mountain chains of the Andes: Western, Central and Eastern Cordilleras. Such pattern in the Andes was reported for trans-Andean (Géry, 1962) and Amazonian freshwater fishes (Hubert, Renno, 2006; Lima, Ribeiro, 2011; Ribeiro et al., 2013; Dagosta, de Pinna, 2017), and further genetic/morphological studies with genera having such distribution (e.g., *Creagrutus* Günther, 1864, *Trichomycterus* Valenciennes, 1832) might test this hypothesis. In addition, the apical position of the Pacific lineages in both clades and their non-monophyletic
pattern show the complex association between phylogeographic connections with the hydrographic history in the Andean region. In this case mediated probably by the formation of the arc of Western Cordillera (Lundberg et al., 1998; Albert, Reis, 2011) and the subduction of Caribbean and South America plates that drove the uplift of the Central Cordillera separating the Pacific drainages from the remaining systems in South America (Rincon-Sandoval et al., 2019).

Overall, our results represent an important step aiming to better delimit and recognize the underestimated species diversity in Astroblepidae and it opens a window for future systematic research on these now well-defined clades. Future directions to investigate the species diversity in this family involve (1) inclusion of unsampled species from Panama and other regions, (2) the formal description of species based on robust and unambiguous morphological diagnoses improving the samples size and ontogenetic variation, (3) the phylogenetic inference based on a time-calibrated genome-scale analysis, and (4) further test our hypothesis of biogeographically-structured species and whether ancestral astroblepids diversified before or after the uplift of the Andean Cordillera.

ACKNOWLEDGMENTS

We are grateful to Jorge Enrique García Díaz who provided logistic support and to José Luis Lozano who kindly helped us in the ANDESCOL2 expedition to Colombian Andes. We also thank the CaVFish Colombia project-Catálogo Visual de Peces de Agua Dulce de Colombia for providing all pictures. This study is dedicated to our colleague Javier A. Maldonado-Ocampo (in memoriam), who was always interested in discovering and protecting the Andean freshwater fishes and for providing a large amount of material of this study collected along many years. The ANDESCOL expeditions and Brycon server used for analyses were funded by FAPESP 14/26508-3 (CO). Authors were funded by grants CNPq 409551/2017-1, FAPESP 14/06853-8, FAPESP 18/23883-9, FAPESP 2016/19075-9 (LEO), FAPESP 16/11313-8 (BFM), FAPESP 17/06551-0 (CSS), 18/09767-6 (CCCS), and FAPESP 14/26508-3 (CO).

REFERENCES


• Ardila-Rodriguez CA. Astroblepus tamboensis sp. nov. - A. chinchaoensis sp. nov. y A. mayoensis sp. nov. Tres nuevas especies de los Andes del Perú. Barranquilla (Atlántico). 2014; 1–18.


• Hazzi NA, Moreno JS, Ortiz-Movlia C, Palacio RD. Biogeographic regions and events of isolation and diversification of the endemic biota of the tropical Andes. Proc Natl Acad Sci U S A. 2018; 115(31):7985–90. https://doi.org/10.1073/pnas.1803908115


**Species delimitation of astroblepid catfishes**


**AUTHOR’S CONTRIBUTION**

Luz E. Ochoa: Conceptualization, Formal analysis, Investigation, Writing—original draft.

Bruno F. Melo: Formal analysis, Investigation, Writing—original draft, Writing—review and editing.

Jorge E. García-Melo: Conceptualization, Funding acquisition, Methodology, Visualization, Writing—review and editing.

Javier A. Maldonado-Ocampo (in memoriam): Conceptualization, Data curation.

Camila S. Souza: Data curation.

Juan G. Albornoz-Garzón: Methodology, Writing—review and editing.

Cristhian C. Conde-Saldaña: Formal analysis, Writing—review and editing.

Francisco Villa-Navarro: Data curation, Investigation, Writing—review and editing.

Armando Ortega-Lara: Data curation, Writing—original draft.

Claudio Oliveira: Conceptualization, Funding acquisition, Investigation, Resources.

**ETHICAL STATEMENT**

The procedures used for the sampling, maintenance and analysis of the tissue fishes samples are in agreement with Colombian laws under the license for scientific collection for purposes of noncommercial scientific research held by Pontificia Universidad Javeriana and granted by the Ministerio de Ambiente y Desarrollo Sostenible of the Republic of Colombia according to Decree 1376 of June 23, 2013. Also the processes are in accordance with Brazilian law regulated by the National Council for the Control of Animal Experimentation (CONCEA) approved by the protocol 1058/2017 and ethical principles in animal research formulated by the Brazilian Society of Science in Laboratory Animal and authorized by the Bioscience Institute/UNESP Ethics Committee on the Use of Animals (CEUA). Fishes under study are not protected by wildlife conservation laws, and no experimentation was conducted on live specimens.

**COMPETING INTERESTS**

The authors declare no competing interests.

**HOW TO CITE THIS ARTICLE**