Multilocus phylogeny and historical biogeography of *Hypostomus* shed light on the processes of fish diversification in La Plata Basin

SHORT RUNNING TITLE: *Hypostomus* diversification in La Plata Basin

Yamila P. Cardoso[1*†], Luiz Jardim de Queiroz[2†], Ilham A. Bahechar[2], Paula E. Posadas[1] & Juan I. Montoya-Burgos[2]

[1] Laboratorio de Sistemática y Biología Evolutiva, Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, Paseo del Bosque S/N, B1900FWA, La Plata, Buenos Aires, Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina.

[2] Department of Genetics and Evolution, University of Geneva, 30 quai Ernest Ansermet, 1211, Geneva 4, Switzerland.

[†] These authors contributed equally to this work.

*Author to whom correspondence should be addressed: Tel: +54 221 422-8451 int. 140, e-mail: yamilapcardoso@gmail.com, ORCID: https://orcid.org/0000-0003-3497-4359

Abstract

Distribution history of the widespread Neotropical genus *Hypostomus* to shed light on the processes that shaped species diversity. We inferred a calibrated phylogeny; ancestral habitat preference, ancestral areas distribution, and the history of dispersal and vicariance events of this genus. The phylogenetic and distributional analyses indicate that *Hypostomus* species inhabiting La Plata Basin do not form a monophyletic clade, suggesting that several unrelated ancestral species colonized this basin in the Miocene (~17 Mya). Dispersal to other rivers of La Plata Basin started about 8 Mya, followed by habitat shifts and an

increased rate of cladogenesis. Amazonian *Hypostomus* species colonized La Plata Basin several times in the Middle Miocene, probably via the Upper Paraná and the Paraguay rivers that acted as biogeographic corridors. During the Miocene, La Plata Basin experienced marine incursions; and geomorphological and climatic changes that reconfigured its drainage pattern, driving the dispersal and diversification of *Hypostomus*. The Miocene marine incursion was a strong barrier and its retraction triggered *Hypostomus* dispersal, increased speciation rate and ecological diversification. The timing of hydrogeological changes in La Plata Basin coincides well with *Hypostomus* cladogenetic events, indicating that the history of this basin has acted on the diversification of its biota.

Keywords Ancestral areas, biogeographic corridor, dispersal, divergence times, freshwater fishes, molecular phylogenetics, Neotropics, Paraná River, vicariance.

Introduction

During the Neogene (23 to 2.6 Mya), Neotropical river networks underwent a complex evolution due to marked changes in geomorphology and environmental conditions, most likely driving great biological diversification in freshwater fishes [1–4]. However, considerable uncertainty remains about the precise links between the region's geomorphological history and the evolutionary origin, geographical distribution and diversification pattern of modern Neotropical fish fauna. Among Neotropical freshwater fishes, armored catfishes of the family Loricariidae represent an excellent model to study the effects of landscape evolution on fish lineage diversification. This family is the most species-rich among Siluriformes and has members in all Neotropical basins. Over more than a quarter-century, investigations of the molecular phylogeny of loricariids have resulted in important discoveries and changes in their classification [5–10]. However, there are still several unresolved phylogenetic relationships in this family, such as in lineages with fast diversification, or due to the under-representation of species coming from particular river basins. These remaining challenges limit the investigation of important questions regarding on the evolution of these Neotropical freshwater fishes.

The genus *Hypostomu*s is one of the most diverse and widespread members of the family Loricariidae, and inhabits most of the Neotropical Realm [11]. The phylogeny and systematics of this group has attracted much interest [10,12–17]. Particularly, in La Plata Basin (which includes the rivers Paraguay, Paraná, Uruguay and Río de La Plata) this genus is represented by more than 50 species, becoming one of the most species-rich genera of this large basin [11,18,19]. This feature is very relevant since La Plata Basin is the third most speciose basin in South America, after the Amazon and the Orinoco basins [20]. Thus, *Hypostomus* is an excellent model to comprehend the events that have given rise to the current diversity of fish species in La Plata Basin. However, understanding the causes of this species richness requires detailed species-level knowledge of geographic distribution, ecologically relevant traits and phylogenetic relationships.

Understanding why members of a clade have dispersed to some places and not to others is essential to recognize large-scale patterns of clade distribution. Historical biogeography studies have become narrowly focused on using phylogenies to discover the history of connections among regions. The geographical distribution of a clade will be determined by (i) the ancestral ecological niche and the opportunities for ecological niche shifts that are afforded to species by their geographical locations; (ii) the ancestral distribution area and the limitations (both physical barriers and/or biotic conditions) to dispersal; and (iii) the amount of time during which niche shifts and dispersal could have occurred [21].

For freshwater fishes, distribution patterns and dispersal dynamics must be interpreted in the context of the geological and geographic histories of river basins [20]. Unlike vagile terrestrial organisms, the ability of strictly freshwater fish to move between drainages is limited by the hydrological connectivity. To understand how the complex geomorphological history of La Plata Basin may have affected the geographical distribution of fish species, it is crucial to infer the phylogenies of widely distributed fish lineages. These phylogenies contain information to trace past physical connections and disconnections within and between

freshwater habitats[1,3,22]. Paleogeographically structured biogeographic patterns may inform about species divergence times especially when no fossils have been assigned to the clade[23].

Geomorphological changes in La Plata Basin Approximately 30 Mya, the Michicola Arc uplifted generating the division between the proto Amazonas-Orinoco and La Plata Basin[24]. However, the boundary between both basins has been locally remodeled since then, providing dispersal opportunities between basins for freshwater organisms. For instance, [25] proposed a notable river capture event between the Pilcomayo River (La Plata Basin) and the Upper Rio Grande Basin (Amazon Basin) 3.5 Mya, which shifted the divide between the two basins farther north. Also, [26] suggested that the Amazon-Paraguay divide is a complex semipermeable geomorphological structure shaped by the Andean orogeny, the megafan environment, and climatic factors, which allowed the migration of several taxa from tributaries of the Amazon to La Plata Basin.

Marine transgressions occurred at least twice in La Plata Basin during the Miocene 15-5 Mya. The first maximum flooding event occurred between 15 and 13 Mya [27,28] and the second between 10 and 6.8 Mya [29–31], [32]. The two marine transgressions covered most of the current Lower Paraná River Basin [28], with a considerable impact on the distribution of freshwater organisms.

The pattern of tributaries of La Plata Basin has undergone numerous modifications through time. Although there is no consensus on the different positions taken by the course of the Paraná River over time, [33,34] proposed that in its initial configuration, the Upper Paraná River was pouring its waters directly into the Uruguay River. Later, about 7 Mya, the lower section of the Upper Paraná changed its course towards the west and connected to the Paraguay river (forming the Lower Paraná), resulting in its current configuration. However, according to [35], the Upper Paraná River was initially turning eastward and headed towards the Patos lagoon (Brazil). Then it took its current configuration about 2.5 Mya. The current hydrographical pattern of the Paraná and Uruguay rivers indicate that they are exclusively connected via the Río de la Plata

estuary. However, several authors [36–38] have proposed contemporaneous water pathways between the Middle Paraná and Middle Uruguay rivers during flood periods.

On the east side of the La Plata Basin, river headwater capture events between the Upper Paraná and the coastal rivers of southern Brazil were documented [39]. According to [40], these capture events have been very numerous between 28-15 Mya and may explain the origin of the fish diversity of the coastal rivers of southern Brazil through dispersal processes [41].

In this work, we use the species-rich catfish genus *Hypostomus* (Loricariidae) as a model group to determine from where and when Amazonian species colonized La Plata Basin, and how they diversified in this basin that has undergone intense orogenic changes. We addressed these questions by inferring a time-calibrated phylogeny for *Hypostomus* with an almost complete representation of species inhabiting La Plata Basin. Using our chronogram together with dated changes in the hydrological pattern of the basin such as the marine incursions, we reconstructed the ancestral distribution ranges and estimated the ancestral preferred habitats for *Hypostomus* species, allowing us to propose an integrated view of the history of diversification of this genus in La Plata Basin.

Materials and methods

Taxon sampling and data collection We analysed 52 species of *Hypostomus* (32 of which were collected in La Plata Basin), representing most of the species described to date in La Plata Basin and covering all the rivers of this basin. Specimens were euthanized immediately after collection in a solution containing a lethal dose of Eugenol (clove oil). The statement on the ethical use of animals is in the supplementary file. Tissue samples for genetic studies were preserved in 96% ethanol and frozen at -20 °C. The voucher specimens were fixed in formalin 4% and deposited at MHNG, IPLA, and MACN (the institutional abbreviations are as listed in Ferraris, 2007) and at Fundación de Historia Natural Félix de Azara, Buenos Aires, Argentina

(CFA-IC). We included five species of closely related genera as outgroups.

Ethics statement In argentina, fish were collected with the permission of the local authorities: Ministerio de Ecologíıa y Recursos Naturales Renovables de Misiones (Disp. 2015 and 2016); Ministerio de Producción y Ambiente de Formosa (N° 040/2015); Secretaría de Medio Ambiente y Desarrollo Sustentable de Santa Fe (Res. 081/2015); Ministerio de Asuntos Agrarios de Buenos Aires (Res. 355/10); Dirección de Fauna y Areas Naturales Protegidas de Chaco (Cons. Aut. 2011); Dirección de Flora, Fauna Silvestre y Suelos de Tucumán (Res. 223/2015); Secretaría de Medio Ambiente y Desarrollo Sustentable de Salta (Res. 091/2015) and Dirección General de Bosques y Fauna de Santiago del Estero (Ref. 17461/2015). Our Institutions do not provide the statements on ethical approval of the study. Notwithstanding, this study was approved by the National Council of Scientific and Technical Research of Argentina (exp. 7879/14) and it is a requirement of this institution to follow the guidelines of its "Comité de Etica" (https://www.conicet.gov.ar/wp-content/uploads/OCR-RD-20050701-1047.pdf) and its biological sampling guide (https://proyectosinv.conicet.gov.ar/solicitud- scientific-collection /). Also, fish handling during sampling was performed following guidelines of the UFAW Handbook on the Care and Management of Laboratory Animals (http://www.ufaw.org.uk). Fish were anesthetized and killed using water containing a lethal dose of eugenol (clove oil). Other used specimens and tissues are deposited in Museu de Ciências e Tecnologia da PUCRS; Muséum d'histoire naturelle - Ville de Genève; and Laboratório de Ictiologia e Pesca of the Universidade Federal de Rondonia (Porto Velho, Brazil).

DNA isolation, PCR, and sequencing Genomic DNA was extracted using the salt-extraction protocol[42]. Polymerase chain reaction (PCR) was used to obtain two mitochondrial DNA (mtDNA) markers: 592 bp of Control Region (D-loop) and 826 bp of Cytochrome Oxidase Subunit I (COI), and partial sections of three nuclear markers [43] : 1345 bp of Recombinase Activating gene-1 (RAG1), 1059 bp of Hodz3 gene and 1144 bp of HAMzbtb10-3 gene. Hodz3 (*Hypostomus* ortholog of the Odz3 gene in vertebrates) encodes a transmembrane protein with a conserved role in eye development in vertebrates [44]. The anonymous marker HAMzbtb10-3 is most probably a fragment of intron number 3 of the gene zbtb10 (zinc finger and BTB

domain containing 10), as identified in the genome of the catfish Ictalurus punctatus (Source:NCBI gene;Acc:108257261). According to SINE Base (http://sines.eimb.ru/), it includes a sequence of 313 bp that is a putative SINE of the Ras1 type [45].

Each PCR amplification was performed in a total volume of 50 µl, containing 5 µl of 10x Taq DNA Polymerase reaction buffer, 1 µl of deoxyribonucleoside triphosphate (dNTP) mix at 10 mM each, 1 µl of each primer at 10 µM, 0.2 µl of Taq DNA Polymerase equivalent to 1 unit of Polymerase per tube, and 1 µl of genomic DNA. Cycles of amplification were programmed as follows: (1) 3 min. at 94 °C (initial denaturation), (2) 30 sec. at 94 °C, (3) 30 sec. at 52–56 °C, (4) 1 min. at 72 °C, and (5) 5 min. at 72 °C (final elongation). Steps 2 to 4 were repeated 40 times. All used primers are listed in Table 2. The PCR products were purified and sequenced by the company MAGROGEN (Korea). A full list of analysed specimens and their respective sequence GenBank accession numbers is given in Table S1.

Sequence alignment, congruence among markers and phylogenetic analysis The editing and alignment of the sequences were performed using BioEdit v.7.0.1 [46]. For each marker and for the alignment of five concatenated markers, we implemented JModelTest2 v.2.1.10 [47] to estimate appropriate substitution models with the Akaike information criterion (AIC).

To assess whether the five used loci could be combined in a joint analysis, we used the Congruence Among Distance Matrices (CADM) test [48] implemented in the R library ape [49]. A Kendall's coefficient of concordance W value of 0 signifies complete incongruence among loci, while a value of 1 indicates complete concordance. The null hypothesis of complete incongruence among DNA sequence distance matrices was tested with 999 permutations. We also assessed congruence among gene phylogenies by running individual gene tree analyses and calculating gene (gCF) and site (sCF) concordance factors in IQtree [50] and following [51]. In these analyses the gCF for each branch is the fraction of decisive gene trees concordant with this branch and the sCF for each branch is the fraction of decisive alignment sites supporting that branch.

We used the concatenated matrix, partitioned by gene, to perform two phylogenetic reconstruction methods . First, maximum likelihood (ML) reconstructions were performed using RAxML [52]. Confidence values for internal edges of the ML tree were computed via bootstrapping [53], based on 1,000 replicates. Second, a Bayesian Inference (BI) analysis was conducted in MrBayes 3.2 [54,55]. Four chains were run simultaneously (three heated, one cold) for 10 million generations, with tree space sampled every 100th generation. After a graphical analysis of the evolution of the likelihood scores, the first 300,000 generations were discarded as burn-in. The remaining trees were used to calculate the consensus tree. Both reconstruction (ML and BI) methods were run on the CIPRES Science Gateway computing cluster [56].

Time-calibrated Tree A time-calibrated phylogenetic reconstruction based on the concatenated data was performed using BEAST v.1.7.5 [57,58]. We used a different substitution model for each marker as selected through Jmodeltest2, with the Yule speciation process and a log normal-relaxed molecular clock. The analysis was run for 20 million generations and sampled every 1,000th generation. The convergence of the runs and ESS values were investigated using Tracer v.1.7.5 [59]. A conservative burn-in of 25% was applied after checking the log-likelihood curves. Finally, a maximum credibility tree with median ages and the 95% highest posterior density (HPD) was subsequently generated using TreeAnnotator v.1.7.5[58]. Due to the absence of an appropriate fossil record, we used two calibration time points based on dated hydrogeological changes. The first point was proposed by [3] and consists in attributing 8 Mya to the speciation between *Hypostomus hondae* and its closest relative, *H. plecostomoides*. *Hypostomus hondae* is distributed only in the Lago Maracaibo and Magdalena basins, while *H. plecostomoides* is known only from the Orinoco basin and some localities in the upper Amazon. These distribution patterns match the vicariant episode that separated Lago Maracaibo system from the Amazon and Orinoco basins 8 Mya [24]. We incorporated a second and new calibration point. Our phylogenetic analyses indicate that *H. formosae*, distributed in the Pilcomayo River (draining into La Plata Basin), is the closest relative to *Hypostomus* sp. BO14-052, which comes from the Rio Grande, a tributary of the Rio Madeira Basin (draining into the Amazon Basin). Since these distribution

patterns match the headwater capture proposed by [25] 3.5 Mya (Table 1), it is reasonable to attribute this age to the allopatric speciation event that gave rise to *H. formosae* and *Hypostomus* sp. BO14-052.

Inference of Ancestral Areas To infer the biogeographic history of *Hypostomus* in La Plata Basin, we used the Dispersal-Extinction-Cladogenesis model (DEC)[60] in the software RevBayes[61].
As our interest focuses on La Plata Basin, we have decided to simplify the hydrological ecoregions of the Amazon, Orinoco, Guiana Shield, and Pacific Coastal drainages in only one region, called C (Fig.1). For La Plata Basin, we used the hydrological ecoregions proposed by Abell et al. (2008): Upper Paraná Basin (E), Paraguay Basin (B), Middle and Lower Paraná Basin and the Endorreic Andean basins (A), and Uruguay Basin (D). We also used other nearby regions such as the Northeastern Brazilian drainages, which also encompass the São Francisco River (G) and the Southeastern Brazilian drainages (F). Distribution areas of the *Hypostomus* species analysed here were taken from many primary-source publications and the collected specimens of CFA, MCP, MACN, MHNG.

In our analyses, we constrained rates of dispersal between ecoregions that are separated (non-adjacent) by one intercalated region to 1/10 of the dispersal rate between adjacent ecoregions. Moreover, due to Miocene marine transgression (10-6.8 Mya) that flooded the Middle and Lower Paraná section with saltwater, and because *Hypostomus* species live in freshwater, we excluded dispersal events crossing the Middle and Lower Paraná Basin (A) during the time window of this marine transgression [time-stratified constraint (MT)]. The tables of pairwise area connectivity constraints are given in Fig. S3 for cases without marine transgression (MT-) and with marine transgression (MT+). As a null model (non time-stratified model), we also ran the same analysis disregarding the changes in the connectivity among ecoregions during the marine transgression. Bayes factor was used for the comparison between both models [23]. Both models were run based on default values as described in the RevBayes tutorial (http://revbayes.github.io/tutorials.html), with a total of 10,000 generations. Since the range size o the most widespread species in our phylogeny covers three

ecoregions, we constrained the ancestral range reconstruction to include a maximum of three discrete ecoregions.

Ancestral habitat reconstruction Habitat preference for each species was classified according to one or more of the following five freshwater habitat categories, taken from [62]. Habitat preference was taken from the original species description, catalogues, checklists and the authors' personal field notes:

1. Upland streams and small rivers in Andean piedmont: swiftly flowing, well oxygenated and rich in dissolved minerals. The substrate comprises rocks and stones; usually between 800 m and 1200 m above sea level.
2. Upland streams and small rivers in shield escarpments: also fast flowing, well oxygenated, but carrying little suspended sediments and dissolved minerals; generally between 200 m and 1000 m above sea level.
3. Lowland *terra-firme* streams and medium rivers lying above the extent of seasonal flooding and over Tertiary formations; usually below 200 m contour. Water with suspended sediments and nutrients and tea-like blackwater coloration. Flow rates are typically low; the substrate is comprised of sand and clay.
4. Deep river channels: deep, swiftly flowing rivers with a seasonal flood pattern; typically exceeding around 3-5 m maximum depth during a typical low water. Turbid water, rich in suspended sediments and nutrients, which imparts a *café au lait* coloration.
5. Tidal estuarine systems with freshwater.

To infer ancestral habitat (discrete traits) across time in the *Hypostomus* tree, we ran the rayDISC function in the R package corrHMM [63] which is a maximum-likelihood method that allows multi-state characters and polymorphic characters. Species showing polymorphic habitat states were assigned a priori to occur in distinct habitats. We ran three different models: (i) one-parameter equal rates (ER) model, in which a single rate is estimated for all possible transitions; (ii) symmetric (SYM) model, in which forwards and reverse transitions between states are constrained to be equal; and (iii) all rates different matrix (ARD) model, in

which all possible transitions between states receive distinct values. The corrected-Akaike information criterion (AICc) was assessed to choose the best model representing the evolution of habitat preference in *Hypostomus* phylogeny.

Diversification through time Analyses based on the time-calibrated phylogenetic tree were performed to quantify the evolutionary tempo and mode of diversification in *Hypostomus* within La Plata Basin. Firstly, using our time-calibrated tree we created lineages through-time (LTT) plot implemented in the R package ape[49] based on only *Hypostomus* from La Plata Basin. To do this we pruned from the *Hypostomus* phylogenetic tree the specimens not included in the desired dataset. Second, to quantify historical rates of species accumulations within La Plata Basin (i.e., tests that diversification has changed over time), we compared our LTT plot with 1000 simulated trees under a constant-rate birth-death model, with identical root age and number of terminals. The simulations were performed using the function 'ltt.plot' (script developed by Hallinan & Matzke, available at http://ib.berkeley.edu/courses/ib200b/labs/lab12/rbdtree.n3.R), and values of speciation and extinction rates to feed this function were estimated using the R package 'diversitree' v. 0.9.10 [64]. Louca and Pennell [65] revealed that an infinite number of pairs of speciation and extinction rates could give rise to any given LTT plot, and it is thus unclear how to determine the correct diversification rates. Therefore, phylogenies based only on extant species may not provide reliable information about past dynamics of diversification. Then, to circumvent the potential problem that underlies LTT plots, we followed Louca and Pennell's [65] recommendation by calculating the 'pulled diversification rate' (PDR). PDR is a variable that is based on a set of congruent diversification scenarios that are compatible with a single LTT, and describes the relative change in diversification rate through time[65]. To do this analysis, we used the function 'fit_hbd_pdr_on_grid' of the R package 'castor' [66]. Also, we graphically represented the rate of speciation of the *Hypostomus* genus only within La Plata Basin. This analysis shows the number of speciation for every 1 Mya time windows. We plotted each principal clade separately and the total cladogenetic events for the entire basin. Unlike most of the LTT plots used, our graph is not cumulative.

Results

Molecular data set The final concatenated alignment comprised 4966 bp, including two mitochondrial (D-loop and COI) and three nuclear markers (RAG1, Hodz3, and HAMzbtb10-3). A total of 52 *Hypostomus* species were represented in the alignment, 32 of which are distributed in La Plata Basin. The concatenated data set has only 1.7% of missing character data. Kendall's coefficient of concordance (W) rejected the incongruence among the matrices of the five loci, globally (W = 0.796, p = 0.003). Thus, the results of the CADM test indicated that all loci could be combined in a single phylogenetic analysis. The calculated gCf and sCF for each branch of the concatenated tree are shown in Fig. S1. Table S2 shows the best nucleotide substitution models for each partition evaluated in JModelTest2 and used in the phylogenetic analyses.

Phylogenetic relationships Phylogenetic analyses based on the five loci provided substantial resolution of relationships among species of *Hypostomus* (Fig. 1). Both maximum likelihood tree (ML) and Bayesian Inference (BI) retrieved the same topology, with well-supported clades [>70 bootstraping (BS) and >0.8 posterior probability (PP)]. Our results supported the organization of *Hypostomus* into four groups, named D1, D2, D3 and D4 (all of them with BS=100 and PP=1), as initially proposed by Montoya-Burgos [3], retrieved by Cardoso [14–16] and recently by Jardim de Queiroz [43]. However, our analyses showed a different relationship among these clades (D1(D2(D3+D4))) with high BS and PP values. Clade D1, forming the *H. cochliodon* group, comprised only one species inhabiting La Plata Basin (*H. cochliodon*). Clade D2, comprising the *H. plecostomus* group, contained 12 species from La Plata Basin. Clades D3 and D4, forming the *H. regani* group, were comprised of 19 species from La Plata Basin.

Relaxed clock, ancestral area and ancestral habitat estimation According to our time calibrated tree reconstruction, *Hypostomus* started diversifying about 20.89 Mya (14.54-28.02 Mya 95% HPD). The ancestral range inference based on RevBayes shows that the genus originated somewhere in the Amazon-Orinoco-Guyane (hydrological ecoregion C in Fig. 1), and occupied upland streams and small rivers in shield escarpments (habitat 2 in Fig. 2; see also Table 3 that shows the AICc values to choose the best model used in the habitat inference).

The earliest split event in the phylogeny of *Hypostomus* divided the group *H. cochliodon* (Clade D1) from the rest of the genus. The diversifications of the Clade D1 started 15.56 Mya (10.85-20.6 Mya 95% HPD) and occurred mostly in the Amazon-Orinoco-Guyane (ecoregion C in Fig. 1). In Clade D1, RevBayes analysis estimated a single dispersal event at 8.34 Mya (6.33-12.56 Mya 95% HPD) from ecoregion C to B and A along the branch leading to *H. cochliodon*. Most *Hypostomus* species in Clade D1 occupied upland streams and small rivers in escarpments (habitat 2 in Fig. 2).

The ancestor of Clade D2 was also hypothesized to have originated in the Amazon-Orinoco-Guyane, approximately 17.01 Mya (11.86-22.78 Mya 95% HPD). The first split in clade D2 generated on one side *Hypostomus watwata* (from coastal Guyana region) and a large clade including 12 species inhabiting the La Plata Basin. This later clade shows on one side *H. formosae* is nested within a lineage of Amazon-Orinoco-Guyane species; on another side 11 species are grouped in a lineage composed principally of species from La Plata Basin. The origin of the later lineage was estimated to 11.73 Mya (7.86-15.67 Mya 95% HPD) in the Upper Paraná (ecoregion E in Fig. 1) after a dispersal event from Amazon-Orinoco-Guyane ecoregion. Our analysis of ancestral habitat inference showed that this ancestor inhabited small rivers (habitat 2 in Fig. 2). Habitat shifts and dispersal from Upper Paraná to other rivers of La Plata Basin as well as to the southern coastal rivers of Brazil were estimated to be younger than 7.05 Mya.

Clade 3 originated in the Paraguay ecoregion (B) after a dispersal event from the Amazon basin, while Clade 4 appeared in the initial ecoregion C (Fig.1). Members of both clades where living in small rivers (habitat 2 in Fig. 2). The first diversification event within Clade 3 was inferred as a vicariance event (but may correspond to the closure of a dispersal portal) between Amazon-Orinoco-Guyane and Upper Paraná approximately 10.38 Mya (6.87-14.2 Mya 95% HPD). Clade 3 includes nine species from La Plata Basin. Most ancestors of these species had the Upper Paraná (ecoregion E in Fig. 1) as ancestral area and occupied small and medium-sized rivers (habitat 2 in Fig. 2). Within this clade, dispersal events to other regions of La

Plata Basin started 8.25 Mya. A lineage from São Francisco River (ecoregion G) is nested within the Clade D3 with a splitting age of 3.31 Mya (1.66-5.22 Mya 95% HPD)

Clade D4 contains ten species from La Plata Basin. Its first diversification event started in the Paraguay River (ecoregion B in Fig. 1), in small and medium rivers, approximately 10.01 Mya (6.69-13.72 Mya 95% HPD). Dispersal events to other rivers of La Plata Basin and shifts in habitat preference were estimated to have occurred very early in the history of this lineage, concomitant with the beginning of the diversification of the D4 lineage.

Diversification through time In this analysis we plotted the numbers of speciation events through time. The LTT plot of *Hypostomus* from La Plata Basin showed an exponential growth in the lineage accumulation over time, but with a significant upshift in the cumulative number of lineages after 5 Ma (Fig. 3). The simulated LTT plots rejected a model of constant net speciation (γ= 1.226787e-01) and extinction (μ=1.985507e-06). Until 7 Mya, the shape of the LTT plot is just under the 75% confidence interval of the simulated LTT data under the constant speciation. Then, its slope increases and become passes the upper 95% confidence level at -5 Mya, higher than expected (above the 99 % confidence intervals at -4 Mya). The pulled diversification rates (PDR) of birth-death models that best explains the phylogenetic tree of *Hypostomus* from La Plata confirmed the pattern observed with the LTT plot, indicating a relative increase of lineage diversification at 5 Mya (Fig 3). We plotted the numbers of speciation events through time only within La Plata Basin (Fig. S4). This figure indicates the number of splitting events for each clade in a time window of one Mya, as well as for the total speciation of the *Hypostomus* genus within La Plata Basin. We also plotted the trend line of each data. The speciation process within this basin peaked between 3 and 5 Mya, with a total of 12 speciation events in this time window. The trend line of the total data showed a positive tendence, thus the most speciation events were young. Our results also show the relative contribution of the three clades with species from La Plata Basin (D2, D3 and D4) in the species diversification in this basin (Fig. S4). The clades D2 and D3 showed also a trend line with positive tendence, strongly influencing

the total data. Conversely, the clade D4 showed a trend line with almost no slope through time, thus the number of speciation events was constant within this clade.

Discussion

In this study, we generated the most thoroughly-sampled, species-level phylogenetic hypotheses for the *Hypostomus* species inhabiting La Plata Basin, to date. The recovery of identical topologies under BI and ML methods based on five loci, with nodes supported with high bootstrap and PP values, suggests that strong phylogenetic signal is present in the data and that the relationships have been accurately reconstructed. Our robust *Hypostomus* phylogeny provides the opportunity to investigate the processes of evolutionary diversification in one of the most species-rich Neotropical fish genera. Previous studies included fewer taxa or were based on fewer molecular markers, resulting in trees with lower statistical support and/or with polytomies within the inter-species relationships [3,8,10,14]. Nevertheless, the main lineages within the genus are similar to those found in [3,10,14–16,43], where four main Clades D1, D2, D3, and D4 were identified.

The South American fish fauna is the most diverse freshwater fish assemblage in the world and currently comprises more than 5000 described species, with many more still undescribed [67]. In their phylogeny, many South American fish groups show sister clade relationships between current basins, most likely reflecting major hydrogeological changes in the continent [68]. Indeed, there is substantial evidence that South America has undergone marked changes over the last 90 Mya, including marine incursions, weathering of ancient shields, uplift of the Andes and of palaeo-arches, significant reconfigurations in the size and drainage patterns of the principal river basins [24,69]

Our phylogenetic results show that the *Hypostomus* species inhabiting La Plata Basin today do not form a monophyletic group. This result indicates that the species richness in this basin has not been entirely formed through the *in situ* diversification of a single ancestor, but is attributed in part to multiple colonization events. [70] analysed several freshwater fish clades under BPA and showed that the fish species pool of modern La

Plata Basin were assembled during the rise of the Michicola Arch ~24 Mya. Our chronology places the origin and the first split within the *Hypostomus* genus in the early Miocene (~20 Mya), and geographically outside from the current La Plata Basin. The polyphyletic origin of *Hypostomus* species from La Plata Basin, combined with the middle Miocene age (~17 Mya) of the three independent colonization events, imply that *Hypostomus* is not part of the 'Old Species of La Plata Basin'. However, this younger age of first colonisation agrees with others studies on *Hypostomus* [3,10,14] and also on other subfamilies of Siluriformes (i.e, Callichthyinae [71]).

Interestingly, our results suggest that the ancestors of the *Hypostomus* lineages from La Plata Basin colonized this basin via the Upper Paraná and Paraguay rivers and their descendants remained confined in this region until the end of the marine incursion which operated as a barrier in the Lower Paraná River (~ 7 Mya).

The predictions of the principle of vicariance assume that multiple lineages will be affected by the raising of the putative barrier, giving origin to several sister-species pairs with similar ages [72]. Conversely, the disappearance of a barrier does not imply the opportunity to disperse for all taxa. Certain geomorphological and climatic processes operating on continental platforms displace barriers to organismal dispersal and probably affect all three processes of macroevolutionary diversification: speciation (species lineage splitting), dispersal (species geographic range expansion), and extinction (species termination) [73]. However, the dispersal also depends on the species ancestral ecological niche and its opportunities for ecological niche shifts to live in new geographical locations.

The marine incursion was an obvious dispersal barrier during late Miocene; however, its impact on freshwater fish dispersal may have persisted, at least partially, after the marine water receded, because such events can permanently alter soil properties and river water characteristics [74]. Therefore, the dispersal of the *Hypostomus* lineages, after this marine incursion, depended on the ancestral ecological niche. The ancestral

habitat and the geographical starting point for dispersal of all *Hypostomus* ancestors of La Plata Basin were estimated by our analyses of area and ancestral habitat inferences as follows. The ancestors of Clades D2, D3 and D4 occupied upland streams and small rivers in shield escarpments (habitat 2). The adaptation to lowland *terra firme* streams and deep river channels (habitats 3 and 4) and the colonization of other ecoregions started ~6.5 Mya, short after the regression of the sea. This result indicates that the ancestral species that crossed the lower parts of La Plata Basin after marine water receded had the correct habitat preference for dispersal through this basin, and thus were adapted to live in larger rivers like Lower Paraná. In contrast, the environmental effect of the marine incursion and the ancestral habitat preference of the Clade D1 could explain why the *H. cochliodon* group showed a very low rate of diversification and dispersion in La Plata Basin. Clade D1 gave rise to a single species that occupies highland streams and small rivers in escarpments in this basin, which it colonized later (~8 Mya) than the other clades. In fact, the *H. cochliodon* group (formerly the genus *Cochliodon* Kner) and the genus *Panaque* Eigenmann are capable of eating wood, an exceptional diet among fishes [75] and have large, spoon-shaped teeth. In contrast, the other *Hypostomus* species inhabiting in La Plata Basin have the typical diet of loricariids, they are algivorous or detritivorous. The habitat preference and the diet differentiation reflects a particular ecological niche of La Plata Basin representatives of Clade D1 that could have played a key role in delaying the dispersion of this group in La Plata Basin.

Once the Upper Paraná was occupied by *Hypostomus* ancestral species of Clades D2 and D3 , their descendants could have disperse throughout the La Plata Basin after the period of sea level regression, a process opening speciation opportunities. Our analysis of LLT plots rejected a model of constant net speciation and extinction and showed an upshift of speciation rate after 5 Mya and other lines of evidence [10] suggest that the Paraná River acted as "biogeographical corridor" for the genus *Hypostomus* to colonize and speciate in other rivers belonging to La Plata Basin today (Paraguay, Uruguay and Río de la Plata). Even the LTT plots were strongly criticized recently [65,76] , our LTT plot agrees with the PDR and the increase of the number of speciation events in our tree after 7 Mya. Consequently, the *Hypostomus* species of these rivers,

which have a tropical origin, contrast with the xerophilous forests and temperate steppes (grassland and savannas) through which most of the rivers run. The effect of the Paraná River as biogeographical corridor of tropical freshwater-related biota towards temperate latitudes is a pattern described by numerous authors in relation to a number of diverse plant and animal groups [77–79]. The existence of a relatively mesic microclimate on the riverbanks and the development of riparian humid forests and wetlands facilitate the dispersion of tropical floristic and faunistic species in temperate latitudes of the Plata Basin, in this way promoting new diversification.

La Plata Basin is a well-connected system today. Thus, in the absence of geographical barriers to dispersal, this basin may have promoted the mixing of populations at the regional level and reduced the likelihood of allopatric speciation. However, ancestral species of *Hypostomus*, subsequent to the multiple independent colonization events from the Amazon, diversified in La Plata Basin, showing both ecological and morphological diversification. The rate of species diversification within La Plata Basin markedly increased between 5 to 3 Mya. The current species diversity of *Hypostomus* in La Plata Basin is lower, but species are more phenotypically diverse than their congeners inhabiting other basins in the Neotropical Realm (as shown by [10]. The heterogeneity of the water environment may be promoting the diversification observed in La Plata Basin. However, this basin was affected by numerous tectonic and climatic events during the last 5 Mya that did not leave obvious geographic features in the contemporary landscape. As shown for other regions of the Neotropics, our study suggests that the current drainage structure in La Plata Basin might be a poor predictor of *Hypostomus* biodiversity.

The colonization of La Plata Basin has not been carried out exclusively through the biogeographic corridor of Upper Paraná River. In the D1 clade, the *H. cochliodon* species from La Plata Basin is nested within the Amazon Basin species (Fig. 2), and our calibrated tree indicates a dispersion event from the Amazon to the Paraguay River (CB ecoregion) ~ 8 Mya, by highland streams and small rivers (habitats 2). Moreover, the ancestral of clade D4, occupied the Paraguay River ~ 10 Mya and the recent formation of *H. formosae* in the

Paraguay River indicate anone other connection about 3.5 Mya [26] showed that the Paraguay and Upper Madeira rivers share most of same species. These authors suggested that the headwaters of some tributaries of the Amazon Basin and the Paraguay River (tributary of La Plata Basin) could interconnect during the rainy season, functioning as portals for the dispersal of fishes possessing the appropriate headwater ecology. However, this suggestion is less supported nowadays, as many of the cases mentioned by Carvalho & Albert [26] and others [80] have recently been recognized as distinct species by their genetics and/or morphology. For example, *Hoplias malabaricus*, considered widespread in South America, has been divided into 16 vicariant species [81]. Inter-basin dispersal of fishes via exceptional portals followed by allopatric speciation has been documented in other regions of South America [82,83], but these flood events would have been older than thought. Because Neotropical ichthyology remains a dynamic field, with many new species described every year [84], the hypothesis of contemporary inter-basin dispersal portals opening during the rainy season will probably be challenged more time.

Our study of species diversification in the genus *Hypostomus* highlights the key role played by river captures. Here, we used a well-documented river capture between the Amazon and Paraguay River [25] to add a calibration point to the *Hypostomus* tree. Moreover, in our analysis of ancestral distribution range, we observed two signatures of river capture involving La Plata Basin. First, in Clade D2, we observed a vicariance event at ~5 Mya between La Plata Basin and the coastal rivers of southern Brazil (ecoregion G). This split was very likely due to the past capture of the Upper Paraná headwaters by coastal drainages [85]. Our results and previous ones [3,10] support the invasion by ancestral lineages of *Hypostomus* from La Plata Basin into the coastal rivers of southern Brazil. Furthermore, different lineages of Neotropical fish show similar patterns of dispersion and colonization. For instance, the time-trees of *Hoplosternum* and *Callichthys* show that both Callichthyidae catfish genera reached the coastal rivers of southern Brazil from the Upper Paraná, and that dispersal occurred from the Amazon to the Upper Paraná but apparently never vice-versa [71]. In contrast, other Loricariidae (Neoplecostominae) and other freshwater fishes (Cynolebiini) had an ancestral origin in the coastal rivers of southern Brazil and dispersed later to La Plata Basin [5,10,41,86]. According to [40], all

these river capture events were very numerous between 28 and 15 Mya; however, our results and others [5,10,41,71] give evidence that more recent river captures have affected fish dispersal and diversity. Our results indicate a second river capture in Clade D3, in which a very well supported lineage shows a São Francisco species nested within La Plata Basin lineages. This result suggests that the Upper Paraná River was not only a "corridor" permitting the entrance of Amazonian species into La Plata Basin, but acted also as a passageway from the La Plata Basin to the São Francisco and to the coastal rivers of southern Brazil, at least for the *Hypostomus* genus.

Phylogenetic and biogeographical knowledge continues to be essential for the elucidation of species evolution, diversification, and the understanding of distribution patterns. However, it is necessary to note that the reconstruction of area cladograms do not depend solely on the history of geological connections among areas, but also on the history of connections among suitable habitats. The inference of a robust phylogenetic hypothesis further enhances the value of a given group of organisms as a model system for understanding Neotropical biodiversity. Our approach, incorporating different data sets (molecular, ecological, and geographical) to analyse the diversification history of a densely sampled fish lineage in the La Plata Basin (Fig. 4), has yielded thorough results for understanding the evolution, ancestral biogeographical ranges, and habitat preferences of *Hypostomus* species in this basin.

Acknowledgements

Numerous colleagues and students participated in the collection of materials and helped in this project, including Sergio Bogan and Juan M. Meluso (Fundación Félix de Azara); Raphael Covain and Sonia Fischer-Muller (Muséum d'histoire naturelle - Ville de Genève); Florencia Brancolini and Gustavo Somoza (Universidad Nacional de San Martín); Carlos Lucena (Museu de Ciências e Tecnologia da PUCRS); Lucila Protogino and Ariel Paracampo (Instituto de Limnología "Dr. Raúl A. Ringuelet"); Carlos Rivera-Rivera (University of Geneva); Adriana Almirón and Jorge Casciotta (Museo de La Plata) and Facundo Vargas (Fauna y Áreas Naturales Protegidas, Provincia del Chaco). We thank Nureen Ghuznavi (George

Washington University) for English revision. Funding for this project was supported by PICT2014-0580 and PICT2017-1330 (YPC); Seed Money Latin America 2015 (JMB/YPC); BOLD-CONICET (YPC); Brazilian–Swiss Joint Research Programme 2015 (MJB/LJQ) and CNPq/Science without Borders (LJQ).

Competing interests: The authors declare no competing interests.

Authors' contributions: YPC and JIMB conceived the project ; YPC, LJD, IAB, JIMB performed the laboratory work, conducted the fieldwork and collected the data with additional material from collaborators; YPC and LJD analysed the data; YPC drafted the manuscript; YPC, LJD, IAB, PEP and JIMB prepared the final version of the manuscript.

Biosketch: YPC is broadly interested in the biogeography of Neotropical Region. This work represents a component of her PhD work at University of La Plata on the systematics and diversity of Loricariidae from La Plata Basin. LJD works on the evolutionary history of Neotropical fishes and on the processes underlying their tremendous diversity. This work is part of his PhD work at University of Geneva, Switzerland.

FIGURES

Fig. 1 Time-calibrated multilocus phylogeny for the *Hypostomus* genus, illustrating ancestral areas estimations performed using RevBayes analysis. For the ecoregion codes and the D1 – D4 clades, see text. Grey vertical bars are the maximum flowing events of the Miocene Marine ingression. Lower bootstrap values and posterior probabilities are indicated.

Fig. 2 Time-calibrated multilocus phylogeny for the *Hypostomus* genus, illustrating ancestral habitat inferences performed using a Bayesian discrete phylogeographic approach. Numbers next to branches show their posterior probability. Habitat preference codes: upland streams and small rivers in Andean piedmont (1), upland streams and small rivers in shield escarpments (2), lowland *terra-firme* streams and medium rivers (3), deep river channels (4), and tidal estuarine systems (5). Grey vertical bars are the maximum flowing events of the Miocene Marine ingression.

Fig. 3 Tempo and mode of evolutionary diversification in *Hypostomus* within La Plata Basin. The black line shows the lineage through time (LTT) curve of species accumulation over time. Colors are the confidence intervals of a neutral model. Blue line is the Pulled Diversification Rate (PDR) over time. Grey vertical bars are the maximum flowing events of the Miocene Marine ingression.

Fig. 4 Graphic summary of the main colonizations in the evolutionary history of the genus *Hypostomus* in La Plata Basin. Major dispersals are shown with arrows. All dates are included and taken from our time-calibrated tree (Fig. 1).

TABLES

Table 1. Documented past hydrogeological events in La Plata Basin.

Hydrogeological events	Mya	References
Boundary displacements between the proto Amazonas-Orinoco and La Plata Basin	30-20	[24]
River headwater capture events between the Upper Paraná and the coastal rivers of southern Brazil	28-15	[26]
First maximum flooding event of marine incursion in La Plata Basin	15-13	[28]
Second maximum flooding event of marine incursion in La Plata Basin	10-6.8	[28,29,31,32]
Boundary displacements between Paraná and Uruguay rivers	7-2.5	[33,35,37]
River headwater capture events between the Rio Grande and Pilcomayo River	3.5	[25]

Table 2. Primers used to amplify the five markers, the sequences length of the amplified DNA fragment (bp) and the number of phylogenetically informative bp (Infor bp) .

Mitochondrial marker	Primers name and sequences (5'-3')	References	bp/ Infor bp
Control region	DLAIII: TATTTAAAGRCATAATCTCTTGAC	[14]	592/197
	HgDL-R: WTGCKARTATGTGCCGYYTG	[14]	
Cytochrome oxidase I	SilCOI-D: GGTCAACAAATCATAAAGATATTGG	This study	825/157
	SilCOI-R: TAAACTTCAGGGTGACCAAAAATCA	This study	
Nuclear marker			
Rag 1	F74: TTTCGGAATGGAAGTTTAAGCTsTTTCG	[87]	1345/62
	R1333: GTCAAACACACAGACTTCACATC	[87]	
	F74-iD: TGATTCTTCAGCTCTCTCAMC	This study	
	R1333-ir: CAAGTGAGAGAGGGTAAGCAG	This study	
CDC27	CDC27-D5: CCACAATATTCTAGATKCCAC	This study	1148/200
	CDC27-R3: GCATTAGGAAATGAGCTCTTC	This study	

	CDC27-id: CCTTTCAGWTATAKWTCTTRTTG	This study	
	CDC27-ir: TATAAARAGCAWGTGGAKMTACAC	This study	
Hodz3	Hodz3-F: GGAGATCAGCAGYGGWGAGG	This study	1060/80
	Hodz3-R: GCCACYTTRATGCAGTCCTC	This study	
	Hodz3-hF: CGGATATTAGCACCTGGAC	This study	
	Hodz3-hR: ACTCCTTTRCCAATCAGGG	This study	

Table 3. Nucleotide substitution models for each partition as evaluated by JModelTest and models used in the Maximum Likelihood (ML) and Bayesian Inference (IB) analyses.

Marker	**JModelTest best n**	**Model used in Bea**	**Model used in MrBayes and RAx**
Control Region	TIM2+I+G	HKY+I+G	GTR+I+G
Cytochrome oxidase I	TVM+G	GTR+G	GTR+I+G
Rag 1	TVM+G	GTR+G	GTR+I+G
HAMzbtb10-3	TVM+G	GTR+G	GTR+I+G
Hodz3	GTR+I+G	GTR+I+G	GTR+I+G

Bibliography

1. Lundberg, J. G. *et al.* in *Phylogeny and classification of Neotropical fishes* (eds. Malabarba, R., Reis, R., Vari, R., Lucena, M. & Lucena, S.) 13–48 (Edipucrs, 1998).
2. Schaefer, S. A. in *Phylogeny and classification of Neotropical fishes* (eds. Malabarba, L. R., Reis, R. E., Vari, R. P., Lucena, Z. M. & S., L. C. A.) 375–400 (Edipucrs, 1998).
3. Montoya-Burgos, J. I. Historical biogeography of the catfish genus Hypostomus (Siluriformes: Loricariidae), with implications on the diversification of Neotropical ichthyofauna. *Mol. Ecol.* **12,** 1855–1867 (2003).
4. Pereira, E. H. L. Resurrection of Pareiorhaphis. *Neotropical Ichthyology* **3,** 271–276 (2005).
5. Chiachio, M. C., Oliveira, C. & Montoya-Burgos, J. I. Molecular systematic and historical biogeography of the armored Neotropical catfishes Hypoptopomatinae and Neoplecostominae (Siluriformes: Loricariidae). *Mol. Phylogenet. Evol.* **49,** 606–617 (2008).

6. Cramer, C. A., Bonatto, S. L. & Reis, R. E. Molecular phylogeny of the Neoplecostominae and Hypoptopomatinae (Siluriformes: Loricariidae) using multiple genes. *Mol. Phylogenet. Evol.* **59,** 43–52 (2011).

7. Fisch-Muller, S., Mol, J. H. A. & Covain, R. An integrative framework to reevaluate the Neotropical catfish genus Guyanancistrus (Siluriformes: Loricariidae) with particular emphasis on the Guyanancistrus brevispinis complex. *PLoS ONE* **13,** e0189789 (2018).

8. Lujan, N. K., Armbruster, J. W., Lovejoy, N. R. & López-Fernández, H. Multilocus molecular phylogeny of the suckermouth armored catfishes (Siluriformes: Loricariidae) with a focus on subfamily Hypostominae. *Mol. Phylogenet. Evol.* **82 Pt A,** 269–288 (2015).

9. Montoya-Burgos, J. I., Weber, C., Pawloski, J. & Muller, S. in *Phylogeny and classification of Neotropical fishes* (eds. Malabarba, L. R., Reis, R. E., Vari, R. P., Lucena, Z. M. & Lucena, C. A. S.) 363–374 (Edipucrs, 1998).

10. Silva, G. S. C. *et al.* Transcontinental dispersal, ecological opportunity and origins of an adaptive radiation in the Neotropical catfish genus Hypostomus (Siluriformes: Loricariidae). *Mol. Ecol.* **25,** 1511–1529 (2016).

11. Weber, C. in *Check List of the freshwater fishes of South and Central America* (eds. Reis, R. E., Kullander, S. O. & Ferraris, J., C.) 351–372 (Edipucrs, 2003).

12. Armbruster, J. W. Phylogenetic relationships of the suckermouth armoured catfishes (Loricariidae) with emphasis on the Hypostominae and the Ancistrinae. *Zool. J. Linn. Soc.* **141,** 1–80 (2004).

13. Cardoso, Y. P., Brancolini, F., Protogino, L. & Lizarralde, M. Actinopterygii, Siluriformes, Loricariidae, Hypostomus aspilogaster (Cope, 1894). Distribution extension and first record for Argentina. *Check List* **7,** 596–598 (2011).

14. Cardoso, Y. P. *et al.* Origin of species diversity in the catfish genus Hypostomus (Siluriformes: Loricariidae) inhabiting the Paraná river basin, with the description of a new species. *Zootaxa* **83,** 69–83 (2012).

15. Cardoso, Y. P. *et al.* Hypostomus formosae, a new catfish species from the Paraguay river basin with

redescription of H. boulengeri (Siluriformes: Loricariidae). *Ichthyological Exploration of Freshwaters* **27,** 9–23 (2016).

16. Cardoso, Y. P. *et al.* An integrated approach clarifies the cryptic diversity in Hypostomus Lacépède 1803 from the Lower La Plata Basin. *An. Acad. Bras. Cienc.* **91,** e20180131 (2019).
17. Montoya-Burgos, J. I., Weber, C. & Le Bail, P. Y. Phylogenetic relationships within Hypostomus (Siluriformes:Loricariidae) and related genera based on Mmitochondrial D-loop sequences. *Revue Suisse de Zoologie* **109,** 369–382 (2002).
18. Koerber, S. & Weber, C. The Hypostominae (Siluriformes : Loricariidae) of Argentina. *Ichthyological Contributions of PecesCriollos* **29,** 1–10 (2014).
19. Litz, T. & Koerber, S. Check List of the Freshwater Fishes of Uruguay (CLOFF-UY). *Ichthyological Contributions of PecesCriollos* **28,** 1–40 (2014).
20. Reis, R. E. *et al.* Fish biodiversity and conservation in South America. *J. Fish Biol.* **89,** 12–47 (2016).
21. Wiens, J. J. & Donoghue, M. J. Historical biogeography, ecology and species richness. *Trends Ecol Evol (Amst)* **19,** 639–644 (2004).
22. Dias, M. S. *et al.* Global imprint of historical connectivity on freshwater fish biodiversity. *Ecol. Lett.* **17,** 1130–1140 (2014).
23. Landis, M. J., Freyman, W. A. & Baldwin, B. G. Retracing the Hawaiian silversword radiation despite phylogenetic, biogeographic, and paleogeographic uncertainty. *Evolution* **72,** 2343–2359 (2018).
24. Lundberg, J. G. in *Phylogeny and classification of Neotropical fishes* (eds. Malabarba, L. R., Reis, R. E., Vari, R. P., Lucena, Z. M. & Lucena, C. A. S.) 49–68 (Edipucrs, 1998).
25. Kennan, L., Lamb, S. H. & Hoke, L. High-altitude palaeosurfaces in the Bolivian Andes: evidence for late Cenozoic surface uplift. *Palaeosurfaces: Recognition, Reconstruction and Palaeoenvironmental Interpretation* **120,** 307–323 (1997).
26. Carvalho, T. & Albert, J. in *Historical Biogeography of Neotropical Freshwater Fishes* (eds. Albert, J. & Reis, R. E.) 193–202 (University of California Press, 2011).
27. Aceñolaza, F. G. Paleobiogeografía de la Región Mesopotámica. *INSUGEO, Miscelánea* **12,** 25–30

(2004).

28. Hernández, R. M. *et al.* Age, distribution, tectonics, and eustatic controls of the Paranense and Caribbean marine transgressions in southern Bolivia and Argentina. *Journal of South American Earth Sciences* **19,** 495–512 (2005).

29. Candela, A. M., Bonini, R. A. & Noriega, J. I. First continental vertebrates from the marine Paraná Formation (Late Miocene, Mesopotamia, Argentina): Chronology, biogeography, and paleoenvironments. *Geobios* **45,** 515–526 (2012).

30. Cione, A. L., Cabrera, D. A. & Barla, M. J. Oldest record of the Great White Shark (Lamnidae, Carcharodon; Miocene) in the Southern Atlantic. *Geobios* **45,** 167–172 (2012).

31. Pérez, L. in *El Neógeno de la Mesopotamia Argentina* (eds. Brandoni, D. & Noriega, J. I.) 7–12 (2013).

32. Cione, A., Cabrera, D., Azpelicueta, M. de las M., Casciotta, J. & Barla, M. J. in *El Neógeno de la Mesopotamia argentina* (eds. Brandoni, D. & Noriega, J. I.) 71 (2013).

33. Popolizio, E. *Los cambios de posición del valle del Río Paraná a lo largo de su historia geomorfológica*. 2–5 (2001).

34. Popolizio, E. El Paraná, un río y su historia geomorfológica. *Revista Geográfica* **140,** 79–90 (2006).

35. Castellanos, A. *Historia hidrogeológica del río Corrientes*. 27 (1959).

36. Casciotta, J., Almirón, A. & Bechara, J. *Peces del Iberá - Hábitat y diversidad*. 244 (2005).

37. Iriondo, M. & Krohling, D. *Cambios ambientales en la cuenca del río Uruguay. Desde dos millones de años hasta el presente*. 360 (2008).

38. Weber, C. Hypostomus microstomus sp. nov. et autres poissons-chats cuirasées du Rio Parana (Pisces, Siluriformes, Loricariidae). *Revue Suisse de Zoologie* **40,** 273–284 (1987).

39. Ribeiro, A. C. Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: an example of faunal evolution associated with a divergent continental margin. *Neotrop. ichthyol.* **4,** 225–246 (2006).

40. Albert, J. & Reis, R. *Historical Biogeography of Neotropical Freshwater Fishes*. 308 (University of California Press, 2011).

41. Roxo, F. F. *et al.* Evolutionary and biogeographic history of the subfamily Neoplecostominae (Siluriformes: Loricariidae). *Ecol. Evol.* **2,** 2438–2449 (2012).

42. Aljanabi, S. M. & Martinez, I. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* **25,** 4692–4693 (1997).

43. Jardim de Queiroz, L. *et al.* Evolutionary units delimitation and continental multilocus phylogeny of the hyperdiverse catfish genus Hypostomus. *Mol. Phylogenet. Evol.* **145,** 106711 (2019).

44. Aldahmesh, M. A., Mohammed, J. Y., Al-Hazzaa, S. & Alkuraya, F. S. Homozygous null mutation in ODZ3 causes microphthalmia in humans. *Genet. Med.* **14,** 900–904 (2012).

45. Ogiwara, I., Miya, M., Ohshima, K. & Okada, N. V-SINEs: a new superfamily of vertebrate SINEs that are widespread in vertebrate genomes and retain a strongly conserved segment within each repetitive unit. *Genome Res.* **12,** 316–324 (2002).

46. Hall, T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acid Syposium Series* **41,** 95–98 (1999).

47. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52,** 696–704 (2003).

48. Campbell, V., Legendre, P. & Lapointe, F.-J. The performance of the Congruence Among Distance Matrices (CADM) test in phylogenetic analysis. *BMC Evol. Biol.* **11,** 64 (2011).

49. Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20,** 289–290 (2004).

50. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32,** 268–274 (2015).

51. Minh, B. Q., Hahn, M. & Lanfear, R. New methods to calculate concordance factors for phylogenomic datasets. *BioRxiv* (2018). doi:10.1101/487801

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

53. Felsenstein, J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* **39,** 783 (1985).

54. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17,** 754–755 (2001).

55. Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61,** 539–542 (2012).

56. Miller, M. A., Pfeiffer, W. & Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. in *2010 Gateway Computing Environments Workshop (GCE)* 1–8 (IEEE, 2010). doi:10.1109/GCE.2010.5676129

57. Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4,** e88 (2006).

58. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7,** 214 (2007).

59. Rambaut, A., Drummond, A., Xie, D., Baele, G. & Suchard, M. Tracer v1.7.5. (2018).

60. Ree, R. H. & Smith, S. A. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* **57,** 4–14 (2008).

61. Höhna, S. *et al.* RevBayes: Bayesian Phylogenetic Inference Using Graphical Models and an Interactive Model-Specification Language. *Syst. Biol.* **65,** 726–736 (2016).

62. Crampton, W. in *Historical biogeography of Neotropical freshwater fishes* (eds. Albert, J. & Reis, R.) 165–189 (2011).

63. Beaulieu, J. M. Package 'corHMM'. Analysis of Binary Character Evolution. (2017). at <https://cran.r-project.org/web/packages/corHMM/corHMM.pdf>

64. FitzJohn, R. G. Diversitree : comparative phylogenetic analyses of diversification in R. *Methods Ecol. Evol.* **3,** 1084–1092 (2012).

65. Louca, S. & Pennell, M. W. Extant timetrees are consistent with a myriad of diversification histories. *Nature* **580,** 502–505 (2020).

66. Louca, S. & Doebeli, M. Efficient comparative phylogenetics on large trees. *Bioinformatics* **34,** 1053–1055 (2018).

67. Reis, R., Kullander, S. & Ferraris, C. J. *Check list of freshwater fishes of South and Central America (CLOFFSCA).* 729 (Edipucrs, 2003).

68. Lovejoy, N. R. & De Araújo, M. L. Molecular systematics, biogeography and population structure of neotropical freshwater needlefishes of the genus Potamorrhaphis. *Mol. Ecol.* **9,** 259–268 (2000).

69. Hoorn, C. *et al.* Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* **330,** 927–931 (2010).

70. Albert, J. & Carvalho, T. in *Historical biogeography of Neotropical freshwater fishes* (eds. Albert, J. & Reis, R.) 119 (2011).

71. Mariguela, T. C., Alexandrou, M. A., Foresti, F. & Oliveira, C. Historical biogeography and cryptic diversity in the Callichthyinae (Siluriformes, Callichthyidae). *J. Zoological System.* (2013). doi:10.1111/jzs.12029

72. Crisp, M. D., Trewick, S. A. & Cook, L. G. Hypothesis testing in biogeography. *Trends Ecol Evol (Amst)* **26,** 66–72 (2011).

73. Albert, J. S., Schoolmaster, D. R., Tagliacollo, V. & Duke-Sylvester, S. M. Barrier displacement on a neutral landscape: toward a theory of continental biogeography. *Syst. Biol.* **66,** 167–182 (2017).

74. Tuomisto, H. Interpreting the biogeography of South America. *J. Biogeogr.* **34,** 1294–1295 (2007).

75. Armbruster, J. W. Peckoltia sabaji, a new species from the Guyana Shield (Siluriformes: Loricariidae). *Zootaxa* **344,** 1–12 (2003).

76. Pagel, M. Evolutionary trees can't reveal speciation and extinction rates. *Nature* 461–462 (2020).

77. Arzamendia, V. & Giraudo, A. R. Influence of large South American rivers of the Plata Basin on distributional patterns of tropical snakes: a panbiogeographical analysis. *J. Biogeogr.* **36,** 1739–1749 (2009).

78. Claramunt, S. Phylogenetic relationships among Synallaxini spinetails (Aves: Furnariidae) reveal a new biogeographic pattern across the Amazon and Paraná river basins. *Mol. Phylogenet. Evol.* **78,** 223–231

(2014).

79. Kay, R. F. Biogeography in deep time - What do phylogenetics, geology, and paleoclimate tell us about early platyrrhine evolution? *Mol. Phylogenet. Evol.* **82 Pt B,** 358–374 (2015).

80. Briñoccoli, Y., Bogan, S., Meluso, J. M. & Cardoso, Y. Update of the distribution of Rhamdella aymarae (Siluriformes: Heptapteridae). *MACN* **20,** 323–332 (2018).

81. Cardoso, Y. P. *et al.* A continental-wide molecular approach unraveling mtDNA diversity and geographic distribution of the Neotropical genus Hoplias. *PLoS ONE* **13,** e0202024 (2018).

82. Buitrago-Suárez, U. & Burr, B. M. Taxonomy of the catfish genus Pseudoplatystoma Bleeker (Siluriformes: Pimelodidae) with recognition of eight species. *Zootaxa* **38,** 1–38 (2007).

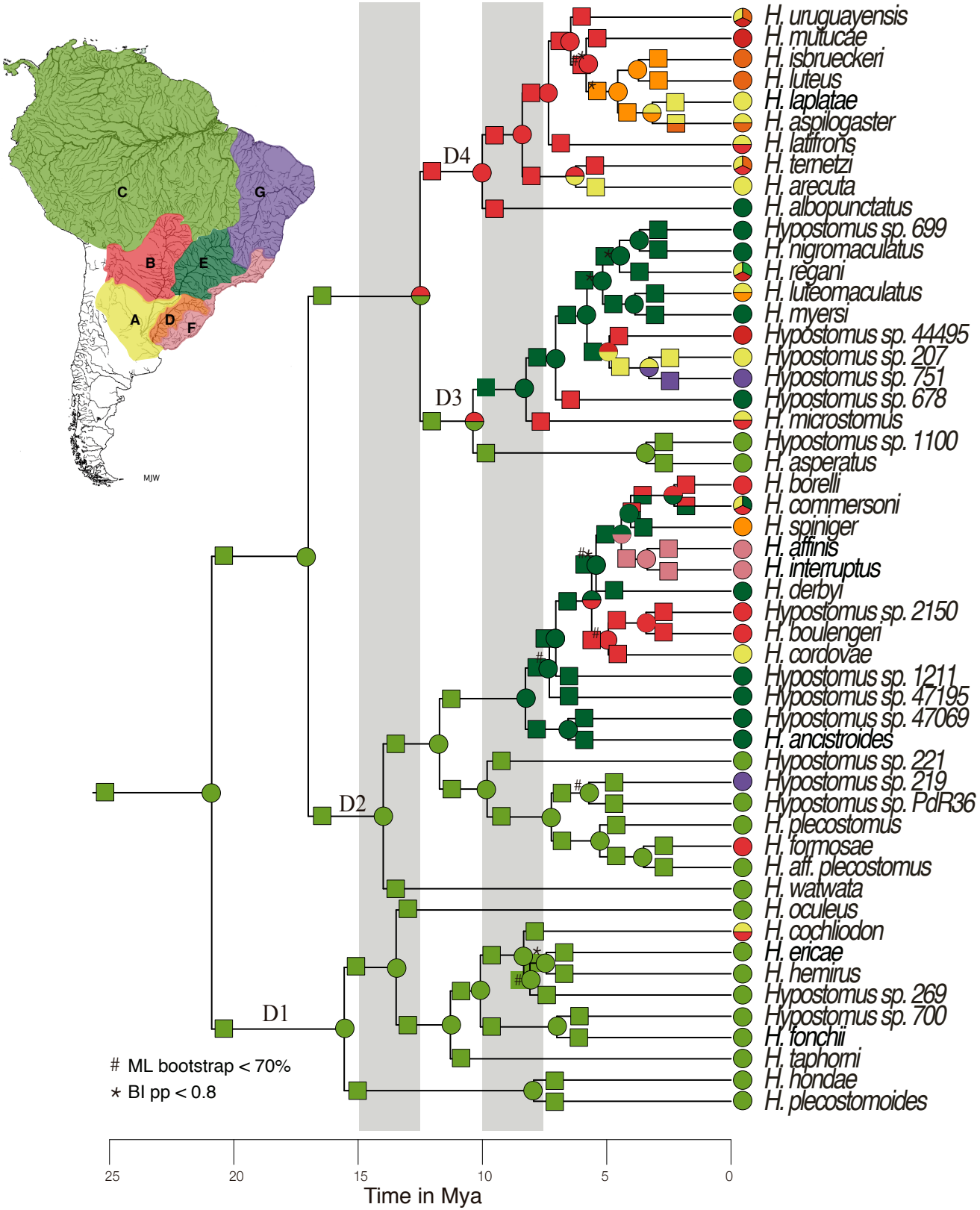
83. Winemiller, K. O., López-Fernández, H., Taphorn, D. C., Nico, L. G. & Duque, A. B. Fish assemblages of the Casiquiare River, a corridor and zoogeographical filter for dispersal between the Orinoco and Amazon basins. *J. Biogeogr.* **35,** 1551–1563 (2008).

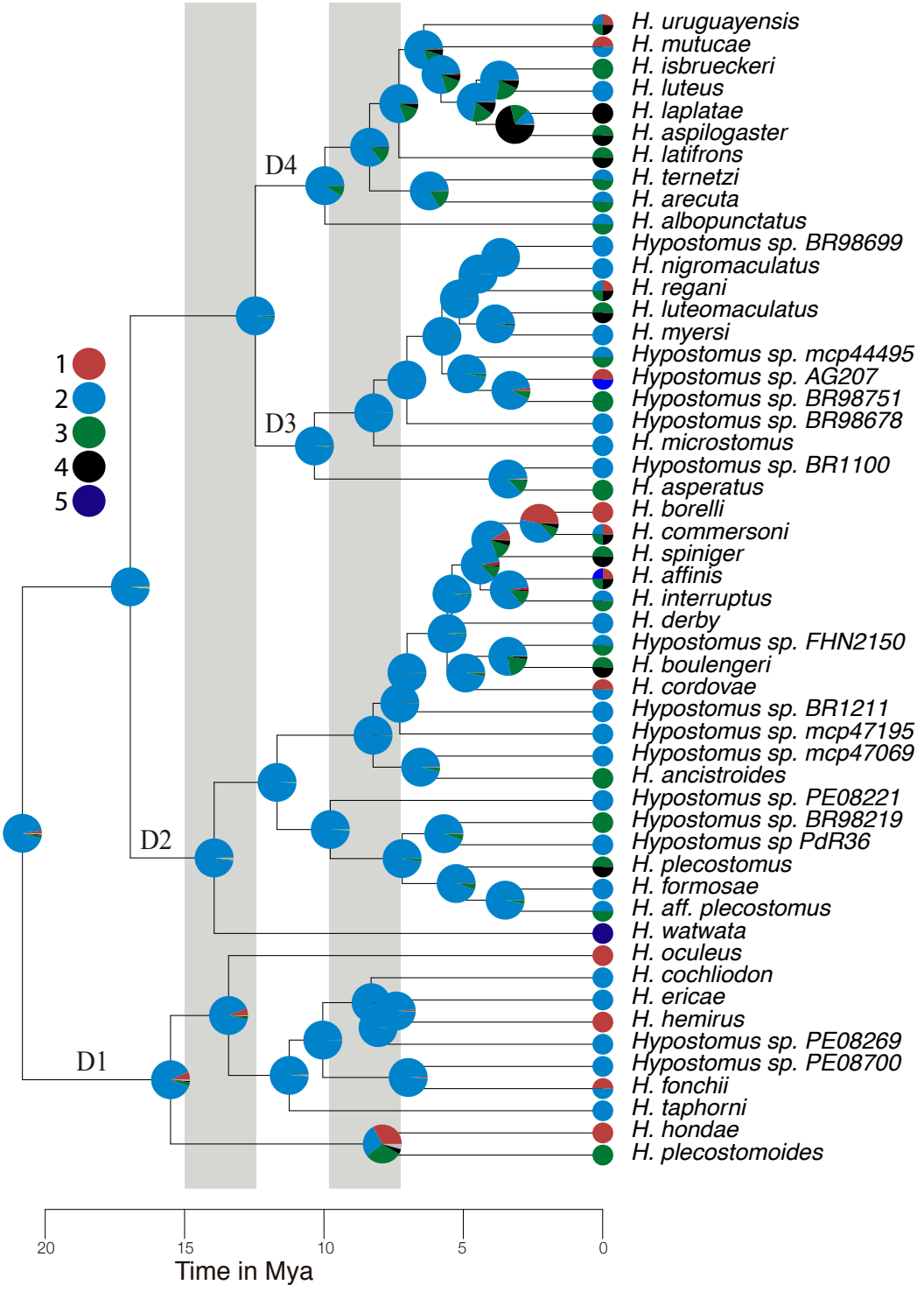
84. Eschmeyer, W. N., Fricke, R. & Laan, R. van der. Catalog of Fishes: Genera, Species, References. (2020). at <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp).>

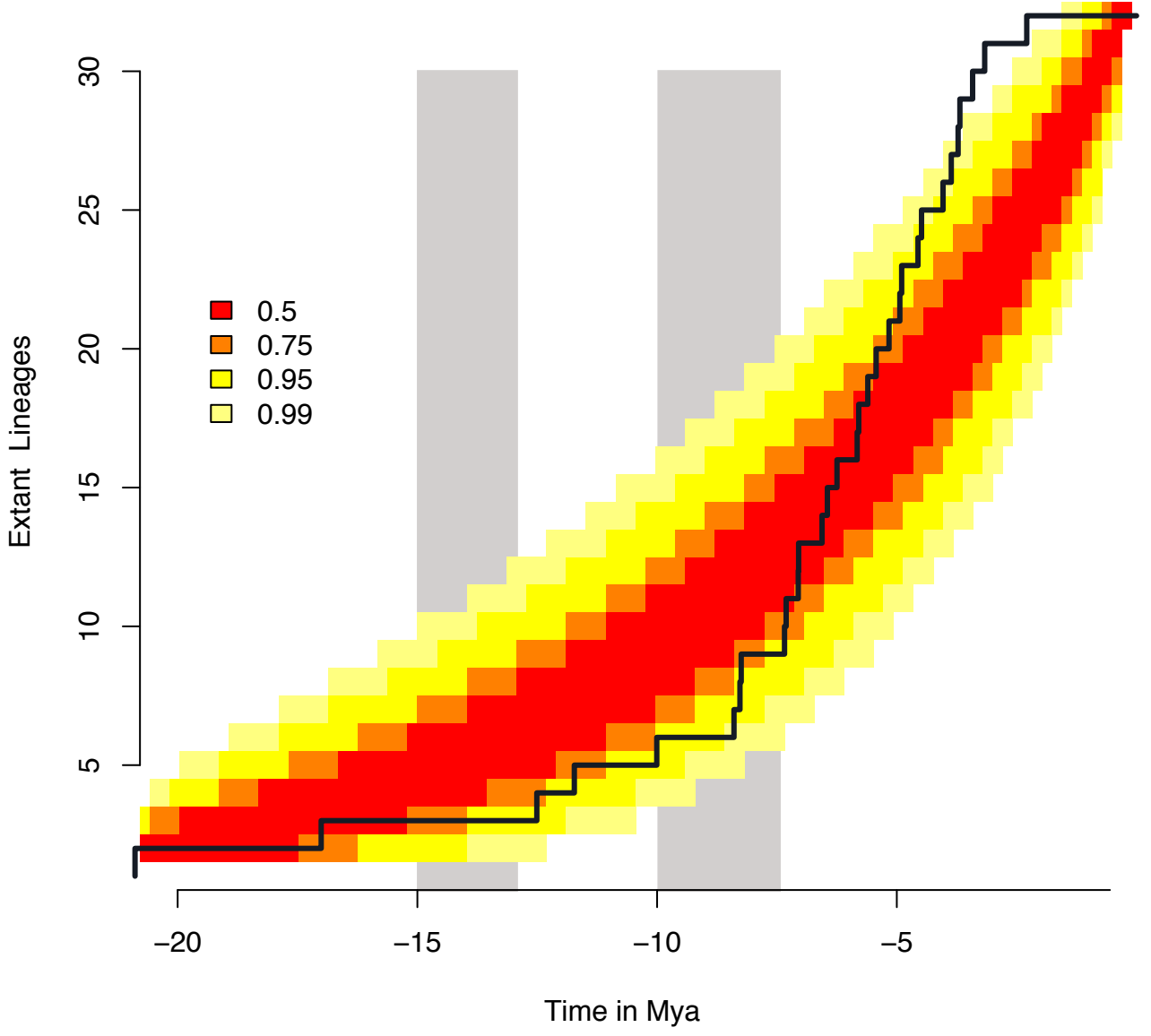
85. Buckup, P. A. in *Historical biogegraphy of Neotropical freshwater fishes* (eds. Albert, J. & Reis, R.) 203 (2011).

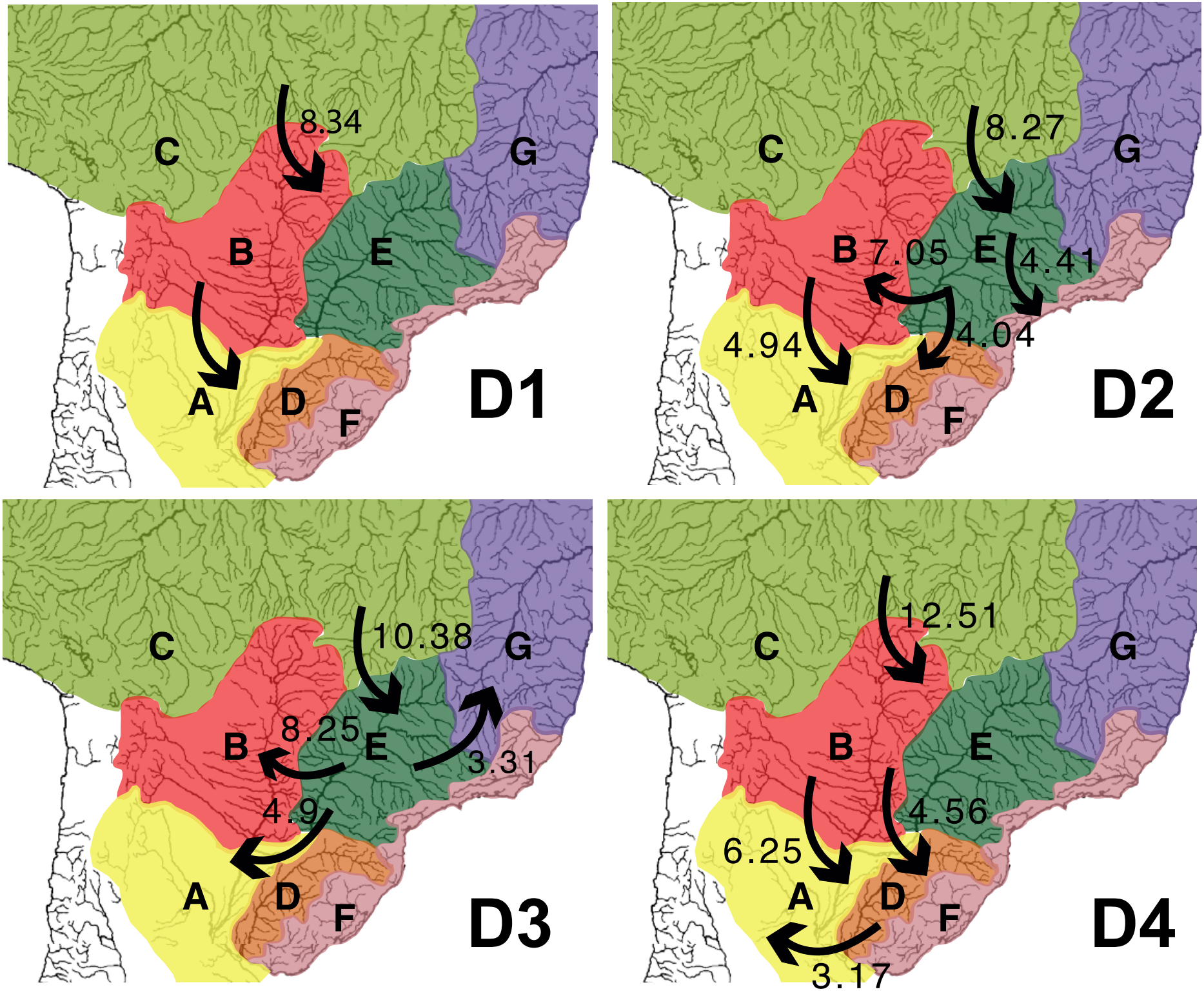
86. Costa, W. J. E. M., Amorim, P. F. & Mattos, J. L. O. Molecular phylogeny and timing of diversification in South American Cynolebiini seasonal killifishes. *Mol. Phylogenet. Evol.* **116,** 61–68 (2017).

87. Sullivan, J. P., Lundberg, J. G. & Hardman, M. A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear gene sequences. *Mol. Phylogenet. Evol.* **41,** 636–662 (2006).


C
G
B
E
A
D
F
MJW
D4
D3
D2
D1
H. uruguayensis
H. mutucae
H. isbrueckeri
H. luteus
H. laplatae
H. aspilogaster
H. latifrons
H. ternetzi
H. arecuta
H. albopunctatus
Hypostomus sp. 699
H. nigromaculatus
H. regani
H. luteomaculatus
H. myersi
Hypostomus sp. 44495
Hypostomus sp. 207
Hypostomus sp. 751
Hypostomus sp. 678
H. microstomus
Hypostomus sp. 1100
H. asperatus
H. borelli
H. commersoni
H. spiniger
H. affinis
H. interruptus
H. derbyi
Hypostomus sp. 2150
H. boulengeri
H. cordovae
Hypostomus sp. 1211
Hypostomus sp. 47195
Hypostomus sp. 47069
H. ancistroides
Hypostomus sp. 221
Hypostomus sp. 219
Hypostomus sp. PdR36
H. plecostomus
H. formosae
H. aff. plecostomus
H. watwata
H. oculeus
H. cochliodon
H. ericae
H. hemirus
Hypostomus sp. 269
Hypostomus sp. 700
H. fonchii
H. taphorni
H. hondae
H. plecostomoides
ML bootstrap < 70%
* BI pp < 0.8
25
20
15
10
5
0
Time in Mya


H. uruguayensis
H. mutucae
H. isbrueckeri
H. luteus
H. laplatae
H. aspilogaster
H. latifrons
H. ternetzi
H. arecuta
H. albopunctatus
Hypostomus sp. BR98699
H. nigromaculatus
H. regani
H. luteomaculatus
H. myersi
Hypostomus sp. mcp44495
Hypostomus sp. AG207
Hypostomus sp. BR98751
Hypostomus sp. BR98678
H. microstomus
Hypostomus sp. BR1100
H. asperatus
H. borelli
H. commersoni
H. spiniger
H. affinis
H. interruptus
H. derby
Hypostomus sp. FHN2150
H. boulengeri
H. cordovae
Hypostomus sp. BR1211
Hypostomus sp. mcp47195
Hypostomus sp. mcp47069
H. ancistroides
Hypostomus sp. PE08221
Hypostomus sp. BR98219
Hypostomus sp PdR36
H. plecostomus
H. formosae
H. aff. plecostomus
H. watwata
H. oculeus
H. cochliodon
H. ericae
H. hemirus
Hypostomus sp. PE08269
Hypostomus sp. PE08700
H. fonchii
H. taphorni
H. hondae
H. plecostomoides
D1
D2
D3
D4
1
2
3
4
5
20
15
10
5
0
Time in Mya


0.5
0.75
0.95
0.99
Extant Lineages
5
10
15
20
25
30
−20
−15
−10
−5
Time in Mya


C
G
B
E
A
D
F
8.34
D1
C
G
B
E
A
D
F
8.27
7.05
4.41
4.94
4.04
D2
C
G
B
E
A
D
F
10.38
8.25
3.31
4.9
D3
C
G
B
E
A
D
F
12.51
4.56
6.25
3.17
D4