Western Amazon was a center of Neotropical fish dispersal, as evidenced by the continental-wide time-stratified biogeographic analysis of the hyper-diverse *Hypostomus* catfish

Running title: Amazon is the center of fish dispersal

Luiz Jardim de Queiroz[1], Xavier Meyer[2], Yamila P. Cardoso[3], Ilham A. Bahechar[1], Raphaël Covain[4], Thiago E. Parente[5], Gislene Torrente-Vilara[6], Paulo A. Buckup[7], Juan I. Montoya-Burgos[1]

[1] Department of Genetics and Evolution, University of Geneva, Switzerland

[2] Department of Integrative Biology, University of California, Berkeley, USA

3 Laboratorio de Sistemática y Biología Evolutiva. Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata. CONICET. Argentina

[4] Department of Herpertology and Ichthyology, Museum of Natural History of Geneva, Switzerland

[5] Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro, Brazil

[6] Departamento de Ciências do Mar, Universidade Federal de São Paulo, Brazil

[7] Museu Nacional, Universidade Federal do Rio de Janeiro, Brazil

[8] iGE3 Institute of Genetics and Genomics of Geneva

Correspondence: Juan I. Montoya-Burgos. Department of Genetics and Evolution, Faculty of Sciences, University of Geneva. Quai Ernest Ansermet 30, CH-1205 Geneva, Switzerland. E-mail: Juan.Montoya@unige.ch.

Abstract

The Amazon is probably the most diverse realm on Earth, and is considered to be the primary source of diversity and a center of dispersal for Neotropical terrestrial organisms. Yet, the assumption that the Amazon basin is a primordial place of fish species origination and dispersal into other drainages still need to be tested. We addressed this issue by inferring a time-stratified biogeographic history and reconstructing the ancestral habitat preference of *Hypostomus*, a continentally widespread and species-rich Neotropical genus. We found that *Hypostomus* emerged in the Western Amazon (~14.7 Ma), when the Western Amazon River was flowing northwards and disconnected from the Eastern Amazon. We show that dispersal events in the first half of *Hypostomus* evolution occurred from the Western Amazon into adjacent basins, initiating its Neotropical radiation. The ancestral preferred habitat consisted in small rivers with running waters, a predominant habitat in river headwaters. Because of strong niche conservatism in the early evolution of *Hypostomus*, we suggest that most of the out-of-Western-Amazon dispersal occurred via headwater captures. The radiation of *Hypostomus* was further promoted by major reconfigurations of river basins, which opened dispersal opportunities into new drainages. Diversification in habitat preference coincided with colonization of basins already occupied by congenerics, indicative of niche shifts triggered by inter-specific competition and species coexistence. By analyzing the evolutionary history of *Hypostomus*, we show that Western Amazon was the main center of fish dispersal in the Neotropical Region from Middle Miocene to the present, supporting the cradle hypothesis of fish origination and dispersal.

Keywords: Amazon River, Lake Pebas, Biogeography, Ancestral ecology, Niche conservatism

Introduction

The Neotropical Region, a large biogeographic realm that covers an area from Southern Mexico to Argentina (Cox, 2001; Leroy et al., 2019), harbors a vast variety of tropical biomes and habitats, and an unparalleled biotic diversity (Hughes, Pennington, & Antonelli, 2013). The processes that have given rise to this diversity are complex, and the many hypotheses of tempo and mode of diversification are still controversial (reviewed in Antonelli, Ariza, et al., 2018). Yet, there is a consensus on the importance of the dynamic geology and climatic variations in shaping the diversity observed nowadays (Antonelli, Ariza, et al., 2018; Antonelli & Sanmartín, 2011). Therefore, taking these processes into account is central for a better comprehension of the origin of the Neotropical biodiversity and the geographical patterns of distribution.

One of the most striking features along the history of the Neotropical Region was the uplift of the Andes, which was the result of a gradual thickening of the Earth's crust due to the interaction between the Nazca and South American plates (Jordan et al., 1983; Pilger, 1984). The Andes is a continuous mountain range along the western edge of South America, extending over more than 7,000 km in length. This mountain range has been uplifting slowly over the last 30 million years, and the modern configuration and elevation were achieved about 14–10 million years ago (Evenstar, Stuart, Hartley, & Tattitch, 2015; C. Hoorn et al., 2010; Pilger, 1984). One of the most emblematic landscape reconfigurations driven by the Andes uplift was the modification of the course of the Amazon River (C. Hoorn, Guerrero, Sarmiento, & Lorente, 1995). Until around 10 Ma, the drainages of the current Amazon and Orinoco rivers formed together the proto-Amazon-Orinoco system (PAO), which was the most important basin that existed from the early Oligocene to the middle-Miocene in the Neotropical Region. PAO flowed northwards into the Caribbean Sea through the outlet of the contemporary Lake Maracaibo (C. Hoorn et al., 2010; Jaramillo et al., 2017). During the

Miocene, marine transgressions changed the landscape of part of the Neotropical Region, including PAO, where a vast saline swamp-like environment was formed, the Pebasian Sea (Wesselingh, Guerrero, Rasanen, Romero Pittman, & Vonhof, 2006). With the progressive uplift of the Northern Andes, the outlet of PAO was gradually closed, blocking large masses of water in the inland. Concomitantly, the rise of a topographic elevation in Northwestern South America, the Vaupés Arch, caused the division of the POA into the western Amazonian basin and the Orinoco basin (C. Hoorn et al., 2010). Then, with the breach of the Purus Arch, which was the drainage divide separating current western and eastern Amazonian drainages, the Amazon River eventually reached the Atlantic ocean and achieved its modern configuration (James S. Albert et al., 2018; C. Hoorn et al., 1995; Carina Hoorn et al., 2017).

Following these drastic modifications of the Neotropical landscape during the Miocene, Amazonian plant and animal lineages experienced an explosive diversification (C. Hoorn et al., 2010). There is strong evidence that the Amazon region is likely the most important source of diversity of terrestrial lineages worldwide (e.g. flowering plants, ferns, birds, mammals and reptiles), playing a pivotal role in shaping the biodiversity across the American continent (Antonelli, Zizka, et al., 2018; Musher, Ferreira, Auerbach, McKay, & Cracraft, 2019). The reasons why Amazonia is a primary source of terrestrial biodiversity are probably multiple, and some interdependent features have been evoked such as the high level of environment heterogeneity spanning over a very large area, a dynamic landscape where diversification has been kept at relatively high levels and constant over time, with low extinction rates, and a high connectivity with neighboring biomes, facilitating dispersal of terrestrial lineages (Musher et al., 2019).

As for fish, the Amazon Basin is the biogeographic core of the Neotropical system, and by far the most species-rich river network (J. S. Albert & Reis, 2011; Böhlke et al., 1978), with more than 7’000 species estimated (Lévêque, Oberdorff, Paugy, Stiassny, & Tedesco,

2008; Reis, Kullander, & Ferraris, 2003). The Amazon Basin has also been suggested to be a cradle of fish diversity, an area where species have originated and immigrated into other basins (J. S. Albert & Reis, 2011; James S. Albert et al., 2018; Fontenelle, Marques, Kolmann, & Lovejoy, 2021; Oberdorff et al., 2019), but this hypothesis still needs further testing. Whether the many and profound landscape changes that occurred in the Amazon Basin were instrumental in fostering the origination, dispersal and diversification of modern Neotropical fish lineages is an assumption that needs to be verified too.

The increasing knowledge about the historical modifications of Neotropical basins offers a timely opportunity to infer freshwater fish ancestral distributions and dispersal routes integrating the changes in watershed configurations through time, but this has been explicitly explored in surprisingly few biogeographic reconstructions of Neotropical fish (e.g. Fontenelle et al., 2021; Wendt, Silva, Malabarba, & Carvalho, 2019). To take advantage of the accumulated knowledge of basins connectivity through time, it is important that reconstructions use the finest possible biogeographic partitions of the hydrological systems, but this implies extensive knowledge about species distribution as well as the inclusion of a comprehensive taxonomic sampling. When based on a limited number of biogeographic areas, as in most of the biogeographical studies of Neotropical fish (Mariguela, Roxo, Foresti, & Oliveira, 2016; e.g. Roxo et al., 2014; Silva et al., 2016), ancestral distribution ranges tend to be unrealistically large, some covering half of South America, thus losing credibility and biological relevance.

In the present work, we investigated of the role played by the various Neotropical river basins, integrating their historical changes in connectivity, as centers of origin and diversification of new fish lineages and as dispersal platforms from which neighboring basins were colonized. To this aim, we reconstructed the biogeographic history of one of the most species-rich genera in the Neotropical Region, the armored catfish genus *Hypostomus*

(Loricariidae). *Hypostomus* are bottom-dwelling fishes with an herbivorous, detritivorous or xylophagous diet, and are generally territorial and poor swimmers. By considering a large fraction of the species diversity of this genus and by including representatives from most South American drainages thanks to the comprehensive phylogeny of *Hypostomus* and closely related genera we recently published [26], we assembled a dataset with one of the finest geographical resolution to date, at the Neotropical level. To work efficiently at the finest spatial resolution, we developed a new tool for the program RevBayes (Höhna et al., 2016) which is an interactive environment for Bayesian statistical analyses with a phylogenetic focus. The new tool we present here allows the user to disregard unrealistic distribution ranges when performing biogeographic analyses, accelerating the analyses and allowing the inclusion of more and finer geographical areas.

Taking advantage of our comprehensive dataset of the exceptionally diverse and widespread genus *Hypostomus*, and using it a model Neotropical fish group, we tested whether (i) the genus *Hypostomus* originated in the Amazon Basin, and (ii) the Amazon Basin was the most important source of diversity for this lineage in the Neotropical Region. Furthermore, we investigated how habitat preference evolved along the phylogeny of *Hypostomus* and how it might have influenced the dispersal pattern of this genus across the Neotropical river basins.

Material and Methods

Phylogenetic reconstruction and tree calibration

To reconstruct the phylogenetic tree of *Hypostomus*, we used the dataset published by (Jardim de Queiroz et al., 2020), which includes six loci: the mitochondrial (i) Cytochrome c oxidase subunit I protein-coding gene (*COI*) and (ii) *D-loop*, and the nuclear (iii) Gene encoding Teneurin transmembrane protein 3 (*Hodz3*), (iv) a *Hypostomus* anonymous marker,

149 possible gene *ZBTB10* intron 3 (HAM-*ZBTB10-3*), (v) the *Recombination activating gene 1*
150 (*RAG1)* and (vi) the Fish *Reticulon 4* (*RTN4*). This dataset contained 206 *Hypostomus*
151 individuals organized in 108 putative species and 49 species belonging to closely related
152 genera (outgroups). We completed this dataset with three additional species: two samples of
153 *Hypostomus* sp. 'Araguaia' (Tocantins-Araguaia Basin, Brazil), one of *Hypostomus* sp.
154 'Aqua75' (Paraguay River, Brazil) and one of *Hypostomus annectens* (Rio Dagua, a Pacific
155 coastal river of Colombia). The primers listed in (Jardim de Queiroz et al., 2020) were used to
156 amplify the desired six loci for the new samples. Genbank accession numbers of the
157 additional samples can be found in Supporting file 3 (Table S3).

158 The tree inference and calibration was performed using BEAST 1.8.1 (Drummond &
159 Rambaut, 2007), while the input file (an Extensible Markup Language, XML, file) was built
160 in BEAUTi 1 (Drummond, Suchard, Xie, & Rambaut, 2012a). We partitioned the data by
161 gene (six partitions), and substitution models and clock models were unlinked among the
162 partitions. A Yule process model was used as a species tree prior. We ran PartitionFinder 2
163 (Lanfear, Frandsen, Wright, Senfeld, & Calcott, 2016) to identify the best nucleotide
164 substitution model based on Bayesian information criterion (BIC): TRN+I+G for *COI* and *D-*
165 *loop*, HKY+G for *Hodz3* and HAM-*ZBTB10-3*, HKY+G+I for *RTN4*, and K80+G for *RAG1*.
166 For the six partitions, we assumed a lognormal relaxed clock (Drummond, Ho, Phillips, &
167 Rambaut, 2006) using a lognormal distribution model with log mean of 0.01 and log standard
168 deviation of 1. We ran five independent runs (15×10^7 generations), sampling the parameters
169 values and trees every 10,000 generations (Supporting file 1).

170 To time-calibrate the tree, we overcame the absence of *Hypostomus* fossils by
171 applying an approach based on biogeographic calibrations. Ho and colleagues (Ho et al.,
172 2015) listed two key assumptions for the application of such calibrations: i) a significant
173 impact on population or species, and ii) available information on the age. We used four dated

hydrogeological changes as calibration points that meet these assumptions: (1) Uplift of the Cordillera de Mérida (mean age of 8.0 Ma ± 0.08 at the crown age of *H. hondae* + *H. plecostomoides*); (2) Uplift of the Ecuadorian Andes (10 Ma ± 0.07 at the crown age of *Aphanotorulus* + *Isorineloricaria*); (3) Disconnection of the Tocantins-Araguaia River from the Amazon Basin (2.6 Ma ± 1 at the crown age of *Hypostomus* sp. 'gr. *cochliodon*-Tar' + *Hypostomus*. sp. 'gr. *cochliodon*-Xin2') (4) Boundary displacement between the Amazon and the Rio Paraguay systems (1 Ma ± 0.6 at the crown age of *Hypostomus* sp. 'Rio Grande' + *Hypostomus* sp. cf. *borelli,*). For more details regarding the calibration points, refer to the Supporting file 3. The inclusion of multiple calibrations may reduce the influence of eventual erroneous calibration and improve the robustness of the time tree (Duchêne, Lanfear, & Ho, 2014).

To evaluate the effect of the use of the aforementioned biogeographic events as calibration points in the *Hypostomus* time-tree, we performed three additional inferences (Supporting file 1). In each of these new reconstructions, one of the new calibration points was not used. We then assessed if important disparities in the time estimates among the distinct reconstruction are observed. Furthermore, assuming that the use of calibration points during phylogenetic reconstructions may improve accuracy of the phylogenetic reconstructions (Drummond et al., 2006), we also verified if major changes in the topology of the tree were observed. For these alternative time-tree reconstructions, we used the same parameters and priors as described before.

Ancestral distribution range

To infer the biogeographic history of *Hypostomus*, we estimated the ancestral distribution range of all the nodes of the phylogenetic tree using the Dispersal–Extinction–Cladogenesis (DEC) model (Ree et al., 2005; Ree & Smith, 2008) implemented in RevBayes

198 (Höhna et al., 2016) (Supporting file 2). The DEC model was originally described — and has
199 been often used — as a maximum likelihood approach (Ree et al., 2005; Ree & Smith, 2008),
200 but RevBayes offers a Bayesian implementation, It allows one to assess parameter uncertainty
201 and to calculate Bayes Factor to compare competitive hypotheses (Landis, Freyman, &
202 Baldwin, 2018).

203 Discrete biohydrogeographic regions (BHG) were defined based on the areas proposed
204 by Abell et al. (2008), and on the areas of Neotropical freshwater fish endemism (Hubert &
205 Renno, 2006; Montoya-Burgos, 2003; Vari, 1988). These areas were adapted to fit with the
206 BHG regions that remained unfragmented during the historical changes in basin connectivity
207 from the Miocene to the present. In this way, we defined 12 BHG regions (Figure 7A): (BHG
208 region A) *Western Amazon* covers the upper portion of the Amazon Basin to the Middle
209 Amazon River, including the sub-basins of Madeira and Negro rivers; (B) *Eastern Amazon*
210 includes the lower portion of the Amazon River and tributaries draining in the Guiana and
211 Brazilian shields, such as the Tapajós, Xingu and Trombetas rivers; (C) *Tocantins-Araguaia*
212 River Basin; (D) *Orinoco* River Basin and the island of Trinidad; (E) *Guianas–Essequibo*
213 includes the Guianese coastal rivers draining the Guyana Shield and flowing into the Atlantic
214 Ocean; (F) *Paraguay* River System covers the Pilcomayo and Paraguay rivers and the lower
215 section of the Paraná River; (G) *Upper Paraná* River covers the upper portion of the Paraná
216 River; (H) *Uruguay* River Basin and *Lagoa dos Patos* System; (I) *Coastal Atlantic* drainages
217 of the South- and Northeastern Brazil from the south of Santa Catarina state in Brazil to the,
218 but not including, São Francisco River mouth; (J) *São Francisco* River Basin; (K) *Coastal*
219 *Atlantic* drainages of Northern Brazil includes rivers flowing into the Atlantic Ocean between
220 the Tocantins-Araguaia and São Francisco mouths; (L) trans-*Andean* drainages includes the
221 Lake Maracaibo, Magdalena River Basin (which drain into the Caribbean Sea) and Pacific
222 drainages of Ecuador and Colombia.

We defined two stationary biogeographic models, with constant constraints in dispersal possibilities for the complete time period of our study (25 Ma to present). The first model (M1) assumes full permeability between all pairs of BHG regions (connectivity value = 1), irrespective of whether they are adjacent to each other or not. The second model (M2) is semi-permeable, in which dispersion is possible but constrained between adjacent BHG regions (connectivity value = 0.1), and dispersion is not allowed between non-adjacent BHG regions (connectivity value = 0). One of the advantages of the DEC model is the possibility of combining *a priori* different temporal connectivity patterns among BHG regions by using a pairwise matrix of connectivity. Hence, we defined a third time-stratified model (M3) by taking into account historical changes in connectivity across BHG regions. For this aim, we used the following connectivity values: (i) a value of 1 was given to a pair of BHG regions for the period of time in which they displayed a full hydrological continuity (i.e. they were parts of a same basin); (ii) a value of 0.5 to a pair of BHG regions for the period of time in which they displayed partial hydrological continuity, for example through a permanent local connection such as the Casiquiare River currently connecting the Western Amazon with the Orinoco basins; (iii) a value of 0.1 to a pair of adjacent BHG regions showing no particular long standing water connection; and (iv) 0 to non adjacent BHG regions. We established these rules because fish, being limited to the aquatic environment, cannot disperse from non-adjacent regions in a single step. The M3 considered four time windows (Supporting file 3, Table S2), each one with a determined connectivity matrix according to the known paleo-landscape changes. For details, see Supporting file 3.

To run the time-stratified (M3) and stationary models (M1 and M2), we estimated a single average value of allopatry and subset sympatry rates. Moreover, the maximum ancestral range was set to three BHG regions, which corresponds to the maximum number of regions occupied by extant *Hypostomus* species included in our study. The original functions

in RevBayes were originally designed to explore all the combinations of discrete BHG regions (Höhna et al., 2016). In our case, with a total of 12 BHG regions and setting the limit of ancestral areas to combinations of at most three BHG regions, we would have 298 possible combinations to compute. To optimize the analyses and reduce computation time, we modified the RevBayes script to disregard combinations of areas that have no biological meaning in terms of species distribution range, such as very distant areas with no present or past connectivity (*e.g. ABK*, *AGH* or *CDE*), reducing the number of possible ranges to 67. This new feature was key to this study since the biogeographic analyses defined using default RevBayes functionalities could not be successfully run for our dataset due to their excessive computational cost (i.e., requiring more than 32GB of RAM and several minutes per iteration). With our modification, the analyses did not suffer from memory limitations and ran faster (i.e., 17 seconds per iteration). A version of the helper script for biogeographic reconstructions in RevBayes can be found in the Supporting file 2 as well as in the GIT repository https://bitbucket.org/XavMeyer/biogeographyrevscript/src/master/, where improvements may be available. The results obtained with the different biogeographic models were compared and ranked by calculating the Bayes Factor, which is a typical model selection approach in the Bayesian framework (Baele, Lemey, & Vansteelandt, 2013). To do so, we ran 40'000 iterations sampling every 10 iterations. Convergence was assessed by plotting log parameters in Tracer.

Habitat preference evolution

In order to investigate how ancestral ecology played a role on the biogeographic history of *Hypostomus*, we evaluated the pattern of evolution of the habitat preference across the phylogenetic tree of *Hypostomus*. We reconstructed the ancestral habitat preference by

defining the nine categories of habitats that were modified from those according to (2011), as listed below (Figure 6):

i) *Slow flowing small streams* (SFSS): streams with maximum 50 m width, such as headwater streams or small tributaries, being localized either in uplands or in lowlands, generally muddy, with substrate comprising mostly of clay and sand.

ii) *Medium to fast flowing small streams* (FFSS): as the previous category, but substrate comprised mainly of rocks or small stones.

iii) *Slow flowing medium-sized rivers* (SFMR): rivers with *c.* 50–1,000 m width, either placed in uplands or lowlands, with muddy and/or sandy substrate.

iv) *Medium to fast flowing medium-sized rivers* (FFMR): as in SFMR, but substrate mainly composed of gravel, stones and rocks. Medium-sized rapids and waterfalls are included in this category.

v) *Slow flowing large rivers* (SFLR): large rivers with more than 1 km width, which include all the rivers or portions of them with very slow flowing waters, with substrate composed mainly of clay and sand.

vi) *Medium to fast flowing large rivers* (FFLR): as in SFLR, but with gravel, rocky and stony substrates, including large rapids and waterfalls.

vii) *Floodplains and ponds* (FLPO): areas that are seasonally inundated by the main rivers, including *várzea* and *igapó* lakes, ponds, flooded forests, flooded savannahs and flooded grasslands.

viii) *Freshwater estuarine systems* (FRES): areas of estuary of very large rivers, where the salinity is still very low. In this category we included only the Río de La Plata, comprising specifically the mouth of Uruguay and Paraguay rivers.

ix) *Brackish water* (BRWA): estuary and coastal regions characterized by relatively high salinity. A single *Hypostomus* species lives in this very atypical habitat in the coast of the Guyanas.

We used BEAST 1.8.1 (Drummond, Suchard, Xie, & Rambaut, 2012b) to reconstruct the ancestral habitat state (Supporting file 1). We employed the aforementioned time-calibrated tree, by fixing topology and branch lengths. We assumed a symmetric model for discrete state reconstructions. A strict clock was used, thereby enforcing a homogeneous rate of trait evolution across all branches, and the prior on mean rate of habitat evolution was set as a gamma distribution. Three independent runs of 5×10^7 generations, sampled every 5,000th generation, were ran. Convergence of the runs and ESS >200 were confirmed with Tracer 1.6. The post-burn-in samples from the three runs were combined using LogCombiner 2.4.8 to produce a consensus tree in TreeAnnotator 2.4.8.

Results

A robust calibrated Phylogeny of Hypostomus

We inferred a time-calibrated phylogeny of *Hypostomus* using four calibration points (calibration scenario 1) indicating that the *Hypostomus* lineage trace back to 14.7 Ma (HPD 95% = 17.8–11.4 Ma), which corresponds to the split between *Hypostomus* and its closest relatives *Hemiancistrus fuliginosus* + *Hemiancistrus punctulatus* (Figure 1). The estimated age of the Most Recent Common Ancestor (MRCA) of *Hypostomus* is 12.1 Ma (14.3–10.1 Ma) while the ages of divergence between the main *Hypostomus* lineages are presented in our calibrated phylogeny (Figure 1).

To evaluate the robustness of our calibrated phylogeny, we assessed the compatibility of each calibration point with regards to the three others by running three new calibrated phylogenetic inferences with the exclusion of one of the four calibrations at a time

(calibration scenarios 2 to 4 in Table 1). The resulting topologies were very similar among each other and also with the results obtained with the full set of calibration points (Supporting file 1). The super-groups of *Hypostomus*, as previously defined in the literature (Jardim de Queiroz et al., 2020), where reconstructed with strong posterior probability (PP ≥ 0.97) in all calibration scenarios (Supporting file 1) . The only noticeable difference was the age retrieved for the calibration point 2 consisting in the split between *Isorineloricaria* and *Aphanotorulus* (IA split) attributed to the uplift of the Ecuadorian Andes. When this calibration point was not used as a prior (calibration scenario 4), the age of this split was estimated to be 19.6 Ma (27.3–13 Ma), that is, older than the uplift of the Ecuadorian Andes, the hypothesized vicariant event causing the split. We also noticed that the splitting ages at the deepest nodes were in general more ancient when the IA split was not considered as a calibration point (calibration scenario 4). However, in all these cases, there is an important overlap of the HPD 95% across the different calibration scenarios (Table 1). Therefore, we used the scenario with four calibration points (calibration scenario 1) in the downstream analyses.

Western Amazon was a centre of origin and dispersal

We reconstructed the biogeographic history of the genus *Hypostomus* using three biogeographic models in RevBayes. The first two were stationary models in which the connectivity between biohydrogeographic (BHG) regions remained unchained along time: the first with full permeability between all pairs of BHG regions (Model 1, M1), and the second combining (i) limited permeability between adjacent BHG regions and (ii) no permeability between non-adjacent ones (Model 2, M2). The third model (M3) assumed a time-stratified scenario with connectivity patterns adapted to four time windows, which were defined based on the most important hydrological changes documented for South America. The time-stratified model M3 showed the highest score relative to the two stationary models (Bayes

342 Factor difference > 19; Supporting file 3, Table S1), providing the best approximation of
343 ancestral ranges throughout the radiation of *Hypostomus,* which we used in our further
344 analyses. Accordingly, the most recent common ancestor (MRCA) of the lineage comprising
345 *Hypostomus* and its sister *Hemiancistrus* group lived in the Western Amazon + Paraguay
346 (start state = region AF; posterior probability = 0.44; Supporting file 1a). This ancestor then
347 experienced a vicariant event (AF → A|F), by which the ancestral population isolated in the
348 Western Amazon (A) gave origin to the first ancestral *Hypostomus*. While living in Western
349 Amazon, this ancestral *Hypostomus* split into consecutive lineages being at the origin of the
350 four recognized super-groups of *Hypostomus* (*Hyp. cochliodon* super-group, *Hyp. hemiurus*
351 super-group, *Hyp. auroguttatus* super-group and *Hyp. plecostomus* super-group). The
352 divergences among these super-groups followed a pattern where one of the descendants
353 remained in Western Amazon, while its sister descendent colonized a new BHG region
354 (Figure 1). In this way, a population from Western Amazon colonized the Guianas–Essequibo
355 region (A→AE) giving rise to the ancestor of the *Hyp. hemiurus* super-group (HHsg).
356 Subsequently, stemming from the ancestral species that remained in Western Amazon, a
357 population entered into the Paraguay System (A→AF), initiating the *Hyp*. *nematopterus* +
358 *Hyp. auroguttatus* super-group (HAsg). Finally, the ancestral species still residing in Western
359 Amazon gave rise to the *Hyp*. *plecostomus* super-group (HPsg).
360 Similarly, other Neotropical main basins where gradually colonized by *Hypostomus*
361 representatives stemming from Western Amazon. The first colonization of the Orinoco Basin
362 (BHG region D) took place around 10.5 Ma with the arrival from Western Amazon (A→D) of
363 the ancestor of *Hyp*. *sculpodon* (a member of the HCsg lineage; Figure 2). The Lower
364 Amazon Basin (Eastern Amazon, BHG region B) was also colonized first by emigrants
365 stemming from Western Amazon at 9.9 Ma (A→BC; Figure 2). The *trans*-Andean area (BHG
366 region L) was the only region adjacent to the Western Amazon that was not initially colonized

by Western Amazon ancestors, but by *Hypostomus* stemming from the Orinoco Basin (D → DL; Figure 2).

According to our biogeographic reconstructions, the Lower Amazon Basin (BHG region B), the Orinoco Basin (BHG region D), and the Guianas-Essequibo System (BHG region E) were repeatedly colonized by *Hypostomus* species coming from Western Amazon. For example, the Orinoco Basin was independently colonized by Western Amazon ancestors at least in three instances: once in the lineage leading to *Hyp. hemicochliodon* (A→ABD; 1.3 Ma, Figure 2), once in the lineage leading to *Hyp. hondae* + *Hyp. plecostomoides* (A→AD; 8.8 Ma, Figure 2), and once at the root of the *Hyp. robinii* group (A→AD; 7.1 Ma, Figure 5). The Guianas-Essequibo System, besides the early colonization by the ancestor of the HHsg, was independently colonized two additional times by Western Amazon ancestors, the first leading to the monotypic lineage of *Hyp. nematopterus* (A→E; 9.2 Ma, Figure 4) and the second giving rise to the *Hyp. watwata* group (A→AE; 7.3 Ma, Figure 5). The re-colonization of the Lower Amazon by Western Amazon emigrants can be exemplified by the arrival of the ancestor of today's *Hyp.* sp. 'Paru' (A→AB; 5 Ma, Figure 3).

Hypostomus radiation into Southern and Eastern South American basins

The colonization of the Upper Paraná (region G) started with the ancestor of the HAsg at 9.2 Ma (Figure 4), stemming from the pre-existing Paraguayan lineage (F→FG). Subsequently, also within the HAsg, the first colonization of the Uruguay System/Lagoa dos Patos (region H) took place at about 4.6 Ma, with colonists emigrating from the Paraguay System (F→FH, Figure 4). Moreover, we recovered two independent colonization of the São Francisco River (region J), both within the *Hyp. auroguttatus* super-group (HAsg). The first gave origin to the *Hyp. asperatus* group (C→CJ; 6.4 Ma,), while the second was the ancestor

of todays *Hyp.* sp. 'BR98751' (G→GJ; 1.9–1.4 Ma), a member of the *Hyp. regani* group (Figure 4).

A single colonization of the coastal Atlantic drainages of Northern Brazil (region K) was inferred and involved the ancestor of *Hyp. pusarum* + *Hyp.* sp. "Pindaré" (B→BCK; 4.2 Ma, Figure 5). On the other hand, the coastal Atlantic rivers (region I) were colonized four independent times. In two cases the ancestral colonists stemmed from the Upper Paraná: (i) the ancestor of *Hyp. auroguttatus* (belonging to the HAsg) at 5.7 Ma (G→I, Figure 4) and (ii) the ancestor of *Hyp.* sp. "gr. *ancistroides* 4" (HPsg) (G→GI) at 3.6 Ma (Figure 5); while in the third case (iii) the colonists came from a region composed of the Paraguay plus the Upper Paraná, at about 3.2 Ma, which led to a small monophyletic lineage comprising five species of our dataset, including *Hyp. affinis* and relatives (FG → GI, Figure 5); and in the fourth case (iv) the ancestor stemmed from the São Francisco Basin at around 1.8 Ma and led to the group of species *Hyp.* sp. "AZ4" + *Hyp.* sp. "Contas" + *Hyp.* aff. *francisci* (J→JI, Figure 4).

Habitat preference conservatism with recent shifts

The ancestral habitat preference reconstructed with BEAST revealed that the MRCA of *Hypostomus* was living in small to medium size streams and rivers with medium–fast flowing waters (EF, Figure 6). This type of habitat remained the preferred one along most of the evolutionary history of the genus and still persists in many representatives of all main lineages. This result is indicative of marked niche conservatism in this genus. Among the changes in habitat preference, the most ancient shift is observed in the stem branch of the lineage including *Hyp. hoplonites* + *Hyp. carinatus* about 7.9 Ma, as this lineage adapted to more lentic environments (floodplains and slow flow medium-large rivers, ACD). The next shift in habitat preference occurred at about 6.1 Ma within the HCsg lineage, with a change towards more lentic environments (floodplains + slow flow medium–large rivers, ACD; and

floodplains + slow flow small–medium rivers, ABC). All the remaining shifts in habitat preference occurred quite recently, starting at about 4.5 Ma. Today, the proportion of species inhabiting lentic habitats is approximately equal to the proportion living in lotic habitats (Figure 6).

Discussion

Although the Amazon region has been recognized as the major source of diversity of terrestrial organisms in the Neotropical realm [e.g. 15, 16], an equivalent role for freshwater lineages has not been thoroughly investigated. In the present study, we have addressed this issue by analyzing *Hypostomus* as a model group due to its remarkable species-richness, its continental-wide distribution, and because its species have colonized all Neotropical freshwater habitats. We assembled information about the distribution, ecology and habitat preference of 111 *Hypostomus* species from the main basins of South America and we calibrated a large phylogenetic tree. We also took advantage of the growing knowledge related to the history of river basins and landscape changes in Tropical South America (J. S. Albert & Carvalho, 2011; J. S. Albert & Reis, 2011; Carvalho & Albert, 2011; e.g. C. Hoorn et al., 2010; John G. Lundberg, Pérez, Dahdul, & Aguilera, 2011; Tagliacollo, Roxo, Duke-Sylvester, Oliveira, & Albert, 2015) to define time windows for each configuration of river network and connectivity, leading to more realistic biogeographic reconstructions. To improve our biogeographic reconstructions, we developed a new tool for RevBayes allowing for computationally tractable analyses with more precise geographic divisions by exploiting *a priori* biologically meaningful distribution ranges.

Western Amazon was the centre of origin of Hypostomus

Deciphering the precise distribution range of a distant ancestor based on its current day descendants depends largely on the completeness of the species sampling, the knowledge

438 of the species distribution range and the size and number of the geographic areas considered
439 in the analysis. Previous attempts to infer the ancestral range of *Hypostomus* used restricted
440 taxonomic sampling, and resulted in excessively large ancestral distribution range for the first
441 ancestral *Hypostomus*, encompassing the entire proto Amazon-Orinoco System (Cardoso et
442 al., 2021) or even the proto-Amazon-Orinoco plus Upper Paraná (Silva et al., 2016). With our
443 comprehensive taxonomic sampling of *Hypostomus* and close relatives, we found that the
444 MRCA of *Hypostomus* inhabited the Western Amazon, an area that nowadays corresponds to
445 the Upper Amazon Basin including some important tributaries, such as the Madeira, Negro,
446 Purus and Japurá rivers. Our calibrated phylogeny based on four cross-tested calibration
447 points (Figs. 1–5; Supporting file 1: Time tree 1) indicates that the first ancestral *Hypostomus*
448 emerged in the Middle-Miocene, approximately from 14.7 Ma, and started to diversify at
449 approximately 12.1 Ma. These temporal results are in agreement with previous hypotheses (J.
450 I. Montoya-Burgos, 2003; Silva et al., 2016), but contrast with the Oligo-Miocene origin (~25
451 Ma) proposed recently (Cardoso et al., 2021).

452 Our findings indicate that the first ancestral *Hypostomus* originated and lived in the
453 Western Amazon Region, at a time where this region was mostly occupied by the Lake Pebas.
454 This mega basin was composed of a mosaic of lentic water bodies and wetlands that
455 dominated the lowlands of western Amazonia, bordered by the slopes of the Andes on the
456 western side and of the Brazilian and Guyana shields on the eastern side (C. Hoorn et al.,
457 2010; Lovejoy, Bermingham, & Martin, 1998; Wesselingh et al., 2006).

458 Due to marine transgressions, the lentic waters of the Lake Pebas basin are suggested
459 to have been influenced by the sea, increasing its salinity (Lovejoy et al., 1998; Wesselingh et
460 al., 2006). It is unlikely that the ancestral *Hypostomus* lived in this brackish-like lentic habitat,
461 as we found that adaptation to brackish waters occurred only once and recently in the
462 evolutionary history of *Hypostomus*, with *Hypostomus watwata* inhabiting the estuaries of

Guianese rivers. Moreover, our habitat preference reconstructions indicate that the ancestral *Hypostomus* was adapted to small rivers with medium-to-fast flowing waters. It is therefore most likely that the first *Hypostomus* was distributed in small rivers draining the slopes of the mountains surrounding Lake Pebas, either on the Andean side or on the slopes of the Purus Arch.

According to our biogeographic analysis, the Western Amazon is the main biohydrogeographic region from which most of the dispersal events originated during the early radiation of *Hypostomus* (Figure 7). In total, the Western Amazon accounted for 30% of all the reconstructed dispersal events along the evolutionary history of *Hypostomus*, a proportion only reached by the Paraguay System (~31%). However, when we consider only the first half of the *Hypostomus* radiation period (12.1–6 Ma), the Western Amazon was the dominant center of dispersal, accounting for 60% of these events.

We uncovered that Western Amazon was the center of origin of *Hypostomus*, a result that is in agreement with the few available findings on Neotropical fish. In a recent study of Amazonian fish biodiversity, the Western Amazon area has been suggested to be the main geographic region of origination of the Amazonian fish fauna, with a downstream colonization of the lower Amazon basin (Oberdorff et al., 2019). However, this study was geographically restricted to the Amazon basin only. At a wider spatial scale, detailed fish biogeographic analyses revealed that the entire Amazon basin, taken as a single BHG region, was the center of origin and dispersal into other Neotropical basins for the catfish subfamily Hypoptopomatinae (Chiachio, Oliveira, & Montoya-Burgos, 2008) and for the characiform genus *Triportheus* (Mariguela et al., 2016). Recently, a more refined biogeographic reconstruction also revealed the Western Amazon as the origin of marine-derived Neotropical freshwater stingrays (Fontenelle et al., 2021), yet the marine origin of this group may not be representative of other primary freshwater lineages. Therefore, our results bring new evidence

in support of the hypothesis that Western Amazon was a primary center of freshwater fish origination.

The role of ecology in Hypostomus *dispersal*

The preferred ancestral habitat of *Hypostomus* consisted in small-to-medium streams with medium-to-fast flowing waters (Figure 6). This habitat preference predominated in the early evolution of *Hypostomus*, while the colonization of new habitats, such as lakes, floodplains, large rivers and estuaries occurred more recently. According to this eco-evolutionary hypothesis, the first half of the *Hypostomus* radiation was characterized by a strong niche conservatism, which may have played a role on the dispersal and speciation modes of this genus. The preferred ancestral habitat of *Hypostomus* is very common in river headwaters, which are portions of drainages prone to be captured by adjacent basins through evolutionary time. Indeed, the geomorphological history of South America has been marked by many events of headwater captures (James S. Albert et al., 2018; Lavarini, Magalhães Júnior, Oliveira, & Carvalho, 2016; Stokes, Goldberg, & Perron, 2018), their incidence being increased by the active tectonics of this continent and by river meandering and erosion in relatively flat relief in several watershed divides.

Relying on our findings of *Hypostomus* ancestral habitat preference and strong niche conservatism along most of their early radiation, we propose that much of the dispersal events into new drainages occurred through the process of headwater captures, promoting range expansion and subsequent allopatric speciation due to geographic isolation. In support of this hypothesis, the results of the ancestral habitat reconstructions indicate that ~61% of dispersal events along the evolution of *Hypostomus* took place in species preferring small or medium rivers with fast-flowing waters (lotic habitats). This is an expected outcome of dispersal by headwater captures, and could hardly be explained by dispersal through alternative

512 environments, such as river delta connections during low sea-level periods (Cardoso &
513 Montoya-Burgos, 2009; A. T. Thomaz, Malabarba, & Knowles, 2017), as the habitat
514 preference of the dispersing species does not match the size of the water body or the stillness
515 of the water found in such environments. The role of headwater captures in dispersal and
516 speciation in South American freshwater fishes has been often suggested, and the
517 hydrogeological hypothesis of diversification (J. I. Montoya-Burgos, 2003) was coined to
518 incorporate this phenomenon. Evidence of headwater captures have been found in drainage
519 boundaries all over Tropical South America, such as among Guianese rivers (e.g. Cardoso &
520 Montoya-Burgos, 2009), between the Amazon Basin and Guianese rivers (Hubert & Renno,
521 2006), between the Amazon and the Orinoco basins (J. G. Lundberg et al., 1998), between the
522 Amazon Basin and the Paraguay System (Hubert & Renno, 2006; J. G. Lundberg et al., 1998;
523 J. I. Montoya-Burgos, 2003; Tagliacollo et al., 2015), between the Upper Paraná and the São
524 Francisco river (Machado, Galetti, & Carnaval, 2018), among Atlantic coastal rivers (M. C. S.
525 L. Malabarba, 1998; Ribeiro, 2006; Roxo et al., 2012), and among many others.

526 However, river headwaters are not always located in mountainous landscapes and
527 temporary connections in basin divides located in flatlands might have also been dispersal
528 routes. For instance, the headwaters of the Guaporé River (Amazon Basin) and Paraguay are
529 located in a flat landscape (around 500 m in altitude) and are regularly interconnected during
530 rainy seasons (Carvalho & Albert, 2011). However, this hypothesis fits better with dispersals
531 later in the radiation of *Hypostomus*, when species shifted their habitat preference for more
532 lentic environments.

533 Every dispersal event into a new basin might have been an ecological opportunity for
534 *Hypostomus* to occupy its preferred habitat, with no congeneric competitors. Although more
535 distant relatives were likely present in the newly colonized basins, we can speculate that a
536 perfect niche overlap between the residing distant relatives and the new *Hypostomus* colonists

is unlikely, although our data did not allow us to test this hypothesis. The hypothetical absence of strong competitors in basins colonized for the first time by *Hypostomus* would explain the marked niche conservatism characterizing the early evolutionary history of this genus, a hypothesis that can be tested in future studies.

We hypothesize that when basins were colonized for the second time, *Hypostomus* started to shift habitat preference, possibly due to competition between resident and colonist congeneric species. This hypothesis is supported by the finding that *Hypostomus* species started to diversify their habitat preference relatively recently along their evolutionary history, when the new dispersers where invading basins already occupied by congenerics. We suggest that the diversity in habitat preference among extant *Hypostomus* species (Figure 6) is most likely the outcome of growing intra-generic competition triggered by the repeated colonization of occupied rivers. According to radiation theory, niche shift is a response to inter-specific competition that allows the co-existence of closely related species, increasing local species-richness (Brew, 1982; Schluter, 2000). Consequently, our findings provide instrumental evidence for explaining the remarkable species richness, species coexistence and the wide distribution range of *Hypostomus* in the Neotropics. The eco-evolution of *Hypostomus* diversity, ecology and distribution may serve as evolutionary model for the entire fish community to which *Hypostomus* belongs, and our results may be informative for understanding the evolutionary history of fish taxa with comparable biological and ecological traits.

Western Amazon as a fish center of dispersal

Elaborating on our findings, we propose that Western Amazon is a center of fish dispersal for five main reasons. (I) It is the center of origin of *Hypostomus*, and this is probably the case for many other species-rich Neotropical fish lineages. (II) Western Amazon

561 is geographically located at the heart of Tropical South America, sharing drainage divides
562 with almost all other main river basins with numerous episodic ichthyologic interchanges
563 through headwater captures. (III) During the Miocene to the present, Western Amazon was
564 the BHG region that had the largest amount of long lasting hydrological interchanged with
565 neighboring basins (Supporting file 3, Table S2). (IV) Western Amazon encompasses the
566 largest Neotropical freshwater system, and is the most fish species-rich BHG region
567 nowadays and was already remarkably rich during the Neogene (John G. Lundberg et al.,
568 2011). Together, these characteristics help to explain why Western Amazon exported
569 repeatedly colonists into its neighboring biogeographic areas all along the evolutionary
570 history of *Hypostomus*. We argue that these characteristics may also be valid for many other
571 Tropical South American fish lineages that had Miocene representative in Western Amazon,
572 as this region acted as a distribution platform, boosting their dispersion throughout
573 Neotropical freshwaters.

574 **Conclusions**

575 In the present work, we assessed the assumption that the Amazon basin was a major
576 center of fish dispersal, spreading new species into neighboring Neotropical river basins from
577 the Miocene to the present. Our detailed time-stratified biogeographic reconstruction of
578 *Hypostomus*, one of the most species-rich genera in the Neotropical Region, indicated that
579 this group emerged in the Western Amazon in the Middle Miocene, when the palaeolandscape
580 was completely different from what we observe nowadays. While new *Hypostomus* species
581 were gradually accumulating in Western Amazon, their ancestral habitat preference enabled
582 them to colonize niches devoid of closely related competitors, and through headwater
583 captures, multiplied the ecological opportunities to spread and diversify into newly colonized
584 basins. Dispersal out of Western Amazon was also boosted by the reconfiguration of the

585 paleo-Amazon-Orinoco watershed, with the disconnection of the Orinoco basin and the
586 junction with Eastern Amazon. The central geographic location of Western Amazon, its
587 longstanding connectivity with other basins, its large extension composed of many sub-basins
588 hosting growing number of species, and the headwater captures spreading *Hypostomus*
589 species in adjacent basins, altogether designate Western Amazon as a center of origin and
590 dissemination for *Hypostomus*. As these characteristics might hold true for many fish lineages
591 with Miocene representatives in Western Amazon, the pattern we observed here for
592 *Hypostomus* may also be valid for a significant fraction of the Neotropical fish diversity.
593 This scenario is supported by the fact that the Western Amazon is an extensive area located at
594 the heart of Tropical South America, with long lasting hydrological connections with
595 neighboring basins, as well as multiple episodic fish exchanges via headwater captures.

596 **Acknowledgments**

597 We thank Margaret Zur and Mary Burridge from the Royal Ontario Museum (Canada) for
598 sample donating (*Hypostomus annectens)*, Jaime Sarmiento and Soraya Barrera from the
599 Museo Nacional de Historia Natural, La Paz, for sample sharing, and Mike R. May for
600 RevBayes-related advises. Part of the analyses was performed on the Baobab cluster, a high
601 performance-computing cluster of the University of Geneva. Funding were provided by
602 *Conselho Nacional de Desenvolvimento Científico e Tecnológico*, Brazil (Program Sciences
603 without Borders, CNPq 229237/2013-4, granted to LJQ; CNPq 423526/2018-9 and
604 312801/2017-3, granted to PAB; CNPQ 307775/2018-6 and 424668/2018-1, granted to TEP);
605 Brazilian–Swiss Joint Research Programme, Switzerland (BSJRP S18794, granted to JIMB,
606 GTV and LJQ); *Comissão de Aperfeiçoamento do Pessoal de Nível Superior,* Brazil
607 (CNPq/CAPES 88882.156885/2016-01 granted to PAB); *Donation Claraz*, Switzerland
608 (granted to JIMB); *Fundação Carlos Chagas Filho de Apoio à Pesquisa do Estado do Rio de*

Janeiro (FAPERJ 200.063/2019 granted to PAB); Swiss Seed Money Grants Latin America 2015, Switzerland (granted to JIMB and YPC); Swiss National Science Foundation (SNSF 3100A0-104005, granted to JIMB; SNSF P400PB 186777, granted to XM).

References

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., … Petry, P. (2008). Freshwater Ecoregions of the World: A New Map of Biogeographic Units for Freshwater Biodiversity Conservation. *BioScience*, *58*(5), 403–414. doi: 10.1641/B580507

Albert, J. S., & Carvalho, T. P. (2011). Neogene Assembly of Modern Faunas. In *Historical Biogeography of Neotropical Freshwater Fishes* (pp. 119–136). Berkeley, Los Angeles and London: University of California Press. Retrieved from https://www.researchgate.net/publication/279241320_Continental-Scale_Tectonic_Controls_of_Biogeography_and_Ecology

Albert, J. S., & Reis, R. E. (2011). Introduction to Neotropical Freshwaters. In J. Albert & R. E. Reis (Eds.), *Historical Biogeography of Neotropical Freshwater Fishes* (pp. 2–19). Berkeley, Los Angeles and London: University of California Press.

Albert, J. S., Val, P., & Hoorn, C. (2018). The changing course of the Amazon River in the Neogene: Center stage for Neotropical diversification. *Neotropical Ichthyology*, *16*(3). doi: 10.1590/1982-0224-20180033

Antonelli, A., Ariza, M., Albert, J., Andermann, T., Azevedo, J., Bacon, C., … Edwards, S. V. (2018). Conceptual and empirical advances in Neotropical biodiversity research. *PeerJ*, *6*. doi: 10.7717/peerj.5644

Antonelli, A., & Sanmartín, I. (2011). Why are there so many plant species in the Neotropics? *TAXON*, *60*(2), 403–414. doi: 10.1002/tax.602010

633 Antonelli, A., Zizka, A., Carvalho, F. A., Scharn, R., Bacon, C. D., Silvestro, D., &
634 Condamine, F. L. (2018). Amazonia is the primary source of Neotropical biodiversity.
635 *Proceedings of the National Academy of Sciences*, *115*(23), 6034–6039. doi:
636 10.1073/pnas.1713819115
637 Baele, G., Lemey, P., & Vansteelandt, S. (2013). Make the most of your samples: Bayes
638 factor estimators for high-dimensional models of sequence evolution. *BMC*
639 *Bioinformatics*, *14*(1), 85. doi: 10.1186/1471-2105-14-85
640 Böhlke, J. E., Weitzman, S. H., Menezes, N. A., Böhlke, J. E., Weitzman, S. H., & Menezes,
641 N. A. (1978). Estado atual da sistemática dos peixes de água doce da América do Sul.
642 *Acta Amazonica*, *8*(4), 657–677. doi: 10.1590/1809-43921978084657
643 Brew, J. S. (1982). Niche shift and the minimisation of competition. *Theoretical Population*
644 *Biology*, *22*(3), 367–381. doi: 10.1016/0040-5809(82)90050-8
645 Cardoso, Y. P., Jardim de Queiroz, L., Bahechar, I. A., Posadas, P. E., & Montoya-Burgos, J.
646 I. (2021). Multilocus phylogeny and historical biogeography of Hypostomus shed
647 light on the processes of fish diversification in La Plata Basin. *Scientific Reports*,
648 *11*(1), 5073. doi: 10.1038/s41598-021-83464-x
649 Cardoso, Y. P., & Montoya-Burgos, J. I. (2009). Unexpected diversity in the catfish
650 Pseudancistrus brevispinis reveals dispersal routes in a Neotropical center of
651 endemism: The Guyanas Region. *Molecular Ecology*, *18*(5), 947–964. doi:
652 10.1111/j.1365-294X.2008.04068.x
653 Carvalho, T. P., & Albert, J. S. (2011). The Amazon-Paraguay Divide. In *Historical*
654 *Biogeography of Neotropical Freshwater Fishes* (pp. 193–202). Berkeley: University
655 of California Press. Retrieved from
656 http://oxfordindex.oup.com/view/10.1525/california/9780520268685.003.0011,
657 //oxfordindex.oup.com/view/10.1525/california/9780520268685.003.0011

658 Chiachio, M. C., Oliveira, C., & Montoya-Burgos, J. I. (2008). Molecular systematic and
659 historical biogeography of the armored Neotropical catfishes Hypoptopomatinae and
660 Neoplecostominae (Siluriformes: Loricariidae). *Molecular Phylogenetics and*
661 *Evolution*, *49*(2), 606–617. doi: 10.1016/j.ympev.2008.08.013
662 Cox, B. (2001). The biogeographic regions reconsidered. *Journal of Biogeography*, *28*(4),
663 511–523. doi: 10.1046/j.1365-2699.2001.00566.x
664 Crampton, W. G. R. (2011). An Ecological Perspective on Diversity and Distributions. In J.
665 Albert & R. E. Reis (Eds.), *Historical Biogeography of Neotropical Freshwater*
666 *Fishes* (pp. 165–189). Berkeley, Los Angeles and London: University of California
667 Press.
668 Drummond, A. J., Ho, S. Y. W., Phillips, M. J., & Rambaut, A. (2006). Relaxed
669 Phylogenetics and Dating with Confidence. *PLoS Biol*, *4*(5), e88. doi:
670 10.1371/journal.pbio.0040088
671 Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by
672 sampling trees. *BMC Evolutionary Biology*, *7*(1), 214. doi: 10.1186/1471-2148-7-214
673 Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012a). Bayesian Phylogenetics
674 with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–
675 1973. doi: 10.1093/molbev/mss075
676 Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012b). Bayesian Phylogenetics
677 with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–
678 1973. doi: 10.1093/molbev/mss075
679 Duchêne, S., Lanfear, R., & Ho, S. Y. W. (2014). The impact of calibration and clock-model
680 choice on molecular estimates of divergence times. *Molecular Phylogenetics and*
681 *Evolution*, *78*, 277–289. doi: 10.1016/j.ympev.2014.05.032

682 Evenstar, L. A., Stuart, F. M., Hartley, A. J., & Tattitch, B. (2015). Slow Cenozoic uplift of
683 the western Andean Cordillera indicated by cosmogenic 3He in alluvial boulders from
684 the Pacific Planation Surface. *Geophysical Research Letters*, *42*(20), 8448–8455. doi:
685 10.1002/2015GL065959
686 Fontenelle, J. P., Marques, F. P. L., Kolmann, M. A., & Lovejoy, N. R. (2021). Biogeography
687 of the neotropical freshwater stingrays (Myliobatiformes: Potamotrygoninae) reveals
688 effects of continent-scale paleogeographic change and drainage evolution. *Journal of*
689 *Biogeography*, *n/a*(n/a). doi: https://doi.org/10.1111/jbi.14086
690 Ho, S. Y. W., Tong, K. J., Foster, C. S. P., Ritchie, A. M., Lo, N., & Crisp, M. D. (2015).
691 Biogeographic calibrations for the molecular clock. *Biology Letters*. (world). doi:
692 10.1098/rsbl.2015.0194
693 Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, B. R., … Ronquist,
694 F. (2016). RevBayes: Bayesian Phylogenetic Inference Using Graphical Models and
695 an Interactive Model-Specification Language. *Systematic Biology*, *65*(4), 726–736.
696 doi: 10.1093/sysbio/syw021
697 Hoorn, C., Bogotá-A, G. R., Romero-Baez, M., Lammertsma, E. I., Flantua, S. G. A., Dantas,
698 E. L., … Chemale, F. (2017). The Amazon at sea: Onset and stages of the Amazon
699 River from a marine record, with special reference to Neogene plant turnover in the
700 drainage basin. *Global and Planetary Change*, *153*, 51–65. doi:
701 10.1016/j.gloplacha.2017.02.005
702 Hoorn, C., Guerrero, J., Sarmiento, G. A., & Lorente, M. A. (1995). Andean tectonics as a
703 cause for changing drainage patterns in Miocene northern South America. *Geology*,
704 *23*(3), 237–240. doi: 10.1130/0091-7613(1995)023<0237:ATAACF>2.3.CO;2
705 Hoorn, C., Wesselingh, F. P., Steege, H. ter, Bermudez, M. A., Mora, A., Sevink, J., …
706 Antonelli, A. (2010). Amazonia Through Time: Andean Uplift, Climate Change,

Landscape Evolution, and Biodiversity. *Science*, *330*(6006), 927–931. doi: 10.1126/science.1194585

Hubert, N., & Renno, J.-F. (2006). Historical biogeography of South American freshwater fishes. *Journal of Biogeography*, *33*(8), 1414–1436. doi: 10.1111/j.1365-2699.2006.01518.x

Hughes, C. E., Pennington, R. T., & Antonelli, A. (2013). Neotropical Plant Evolution: Assembling the Big Picture. *Botanical Journal of the Linnean Society*, *171*(1), 1–18. doi: 10.1111/boj.12006

Jaramillo, C., Romero, I., D'Apolito, C., Bayona, G., Duarte, E., Louwye, S., … Wesselingh, F. P. (2017). Miocene flooding events of western Amazonia. *Science Advances*, *3*(5), e1601693. doi: 10.1126/sciadv.1601693

Jardim de Queiroz, L., Cardoso, Y., Jacot-des-Combes, C., Bahechar, I. A., Lucena, C. A., Rapp Py-Daniel, L., … Montoya-Burgos, J. I. (2020). Evolutionary units delimitation and continental multilocus phylogeny of the hyperdiverse catfish genus Hypostomus. *Molecular Phylogenetics and Evolution*, *145*, 106711. doi: 10.1016/j.ympev.2019.106711

Jordan, T. E., Isacks, B. L., Allmendinger, R. W., Brewer, J. A., Ramos, V. A., & Ando, C. J. (1983). Andean tectonics related to geometry of subducted Nazca plate. *GSA Bulletin*, *94*(3), 341–361. doi: 10.1130/0016-7606(1983)94<341:ATRTGO>2.0.CO;2

Landis, M. J., Freyman, W. A., & Baldwin, B. G. (2018). Retracing the Hawaiian silversword radiation despite phylogenetic, biogeographic, and paleogeographic uncertainty. *Evolution*, *72*(11), 2343–2359. doi: 10.1111/evo.13594

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2016). PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and

731 Morphological Phylogenetic Analyses. *Molecular Biology and Evolution*, *34*(3), 772–
732 773. doi: 10.1093/molbev/msw260
733 Lavarini, C., Magalhães Júnior, A. P., Oliveira, F. S. de, & Carvalho, A. de. (2016).
734 Neotectonics, river capture and landscape evolution in the highlands of SE Brazil.
735 *Mercator (Fortaleza)*, *15*(4), 95–119. doi: 10.4215/rm2016.1504.0007
736 Leroy, B., Dias, M. S., Giraud, E., Hugueny, B., Jézéquel, C., Leprieur, F., … Tedesco, P. A.
737 (2019). Global biogeographical regions of freshwater fish species. *Journal of*
738 *Biogeography*, *46*(11), 2407–2419. doi: 10.1111/jbi.13674
739 Lévêque, C., Oberdorff, T., Paugy, D., Stiassny, M. L. J., & Tedesco, P. A. (2008). Global
740 diversity of fish (Pisces) in freshwater. In E. V. Balian, C. Lévêque, H. Segers, & K.
741 Martens (Eds.), *Freshwater Animal Diversity Assessment* (pp. 545–567). Dordrecht:
742 Springer Netherlands. doi: 10.1007/978-1-4020-8259-7_53
743 Lovejoy, N. R., Bermingham, E., & Martin, A. P. (1998). Marine incursion into South
744 America. *Nature*, *396*(6710), 421–422. doi: 10.1038/24757
745 Lundberg, J. G., Marshall, L. G., Guerrero, J., Horton, B., Malabarba, L. M. C. S., &
746 Wesselingh, F. (1998). The stage for neotropical fish diversification: A history of
747 tropical South American rivers. In L. R. Malabarba, R. E. Reis, R. Vari, L. C. A.
748 Lucena, & Z. M. S. Lucena (Eds.), *Phylogeny and Classification of Neotropical*
749 *Fishes* (pp. 13–48). Porto Alegre: EDIPUCRS.
750 Lundberg, J. G., Pérez, M. H. S., Dahdul, W. M., & Aguilera, O. A. (2011). The Amazonian
751 Neogene Fish Fauna. In *Amazonia: Landscape and Species Evolution* (pp. 281–301).
752 John Wiley & Sons, Ltd. doi: 10.1002/9781444306408.ch17
753 Machado, C. B., Galetti, P. M., & Carnaval, A. C. (2018). Bayesian analyses detect a history
754 of both vicariance and geodispersal in Neotropical freshwater fishes. *Journal of*
755 *Biogeography*, *45*(6), 1313–1325. doi: 10.1111/jbi.13207

Malabarba, M. C. S. L. (1998). Phylogeny of Fossil Characiformes and Paleobiogeography of the Tremembé Formation, São Paulo, Brazil. In *Phylogeny and Classification of Neotropical Fishes* (pp. 69–84). Porto Alegre: PUCRS.

Mariguela, T. C., Roxo, F. F., Foresti, F., & Oliveira, C. (2016). Phylogeny and biogeography of Triportheidae (Teleostei: Characiformes) based on molecular data. *Molecular Phylogenetics and Evolution*, *96*, 130–139. doi: 10.1016/j.ympev.2015.11.018

Montoya-Burgos, J. I. (2003). Historical biogeography of the catfish genus Hypostomus (Siluriformes: Loricariidae), with implications on the diversification of Neotropical ichthyofauna. *Molecular Ecology*, *12*(7), 1855–1867. doi: 10.1046/j.1365-294X.2003.01857.x

Musher, L. J., Ferreira, M., Auerbach, A. L., McKay, J., & Cracraft, J. (2019). Why is Amazonia a 'source' of biodiversity? Climate-mediated dispersal and synchronous speciation across the Andes in an avian group (Tityrinae). *Proceedings of the Royal Society B: Biological Sciences*, *286*(1900), 20182343. doi: 10.1098/rspb.2018.2343

Oberdorff, T., Dias, M. S., Jézéquel, C., Albert, J. S., Arantes, C. C., Bigorne, R., … Zuanon, J. (2019). Unexpected fish diversity gradients in the Amazon basin. *Science Advances*, *5*(9), eaav8681. doi: 10.1126/sciadv.aav8681

Pilger, R. H. (1984). Cenozoic plate kinematics, subduction and magmatism: South American Andes. *Journal of the Geological Society*, *141*(5), 793–802. doi: 10.1144/gsjgs.141.5.0793

Ree, R. H., Moore, B. R., Webb, C. O., Donoghue, M. J., & Crandall, K. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, *59*(11), 2299–2311. doi: 10.1554/05-172.1

779 Ree, R. H., & Smith, S. A. (2008). Maximum Likelihood Inference of Geographic Range
780 Evolution by Dispersal, Local Extinction, and Cladogenesis. *Systematic Biology*,
781 *57*(1), 4–14. doi: 10.1080/10635150701883881
782 Reis, R. E., Kullander, S. O., & Ferraris, C. J. (2003). *Check List of the Freshwater Fishes of*
783 *South and Central America*. EDIPUCRS.
784 Ribeiro, A. C. (2006). Tectonic history and the biogeography of the freshwater fishes from
785 the coastal drainages of eastern Brazil: An example of faunal evolution associated
786 with a divergent continental margin. *Neotropical Ichthyology*, *4*(2), 225–246. doi:
787 10.1590/S1679-62252006000200009
788 Roxo, F. F., Albert, J. S., Silva, G. S. C., Zawadzki, C. H., Foresti, F., & Oliveira, C. (2014).
789 Molecular Phylogeny and Biogeographic History of the Armored Neotropical Catfish
790 Subfamilies Hypoptopomatinae, Neoplecostominae and Otothyrinae (Siluriformes:
791 Loricariidae). *PLOS ONE*, *9*(8), e105564. doi: 10.1371/journal.pone.0105564
792 Roxo, F. F., Zawadzki, C. H., Alexandrou, M. A., Costa Silva, G. J., Chiachio, M. C., Foresti,
793 F., & Oliveira, C. (2012). Evolutionary and biogeographic history of the subfamily
794 Neoplecostominae (Siluriformes: Loricariidae). *Ecology and Evolution*, *2*(10), 2438–
795 2449. doi: 10.1002/ece3.368
796 Schluter, D. (2000). Ecological Character Displacement in Adaptive Radiation. *The American*
797 *Naturalist*, *156*(S4), S4–S16. doi: 10.1086/303412
798 Silva, G. S. C., Roxo, F. F., Lujan, N. K., Tagliacollo, V. A., Zawadzki, C. H., & Oliveira, C.
799 (2016). Transcontinental dispersal, ecological opportunity and origins of an adaptive
800 radiation in the Neotropical catfish genus Hypostomus (Siluriformes: Loricariidae).
801 *Molecular Ecology*, *25*(7), 1511–1529. doi: 10.1111/mec.13583
802 Stokes, M. F., Goldberg, S. L., & Perron, J. T. (2018). Ongoing River Capture in the Amazon.
803 *Geophysical Research Letters*, *45*(11), 5545–5552. doi: 10.1029/2018GL078129

804 Tagliacollo, V. A., Roxo, F. F., Duke-Sylvester, S. M., Oliveira, C., & Albert, J. S. (2015).
805 Biogeographical signature of river capture events in Amazonian lowlands. *Journal of*
806 *Biogeography*, 2349–2362. doi: 10.1111/jbi.12594
807 Thomaz, A. T., Malabarba, L. R., & Knowles, L. L. (2017). Genomic signatures of
808 paleodrainages in a freshwater fish along the southeastern coast of Brazil: Genetic
809 structure reflects past riverine properties. *Heredity*, *119*(4), 287–294. doi:
810 10.1038/hdy.2017.46
811 Vari, R. P. (1988). The Curimatidae, a lowland neotropical fish family (Pisces:
812 Characiformes): distribution, endemism, and phylogenetic biogeography. *Neotropical*
813 *Distribution Patterns: Proceedings of a Workshop. Rio de Janeiro: Academia*
814 *Brasileira de Ciencias.*, 343–377.
815 Wendt, E. W., Silva, P. C., Malabarba, L. R., & Carvalho, T. P. (2019). Phylogenetic
816 relationships and historical biogeography of Oligosarcus (Teleostei: Characidae):
817 Examining riverine landscape evolution in southeastern South America. *Molecular*
818 *Phylogenetics and Evolution*, *140*, 106604. doi: 10.1016/j.ympev.2019.106604
819 Wesselingh, F. P., Guerrero, J., Rasanen, M. E., Romero Pittman, L., & Vonhof, H. B. (2006).
820 Landscape evolution and depositional processes in the Miocene Amazonian Pebas
821 lake/wetland system: Evidence from exploratory boreholes in northeastern Peru.
822 *Instituto Geológico, Minero y Metalúrgico – INGEMMET*. Retrieved from
823 http://repositorio.ingemmet.gob.pe/handle/ingemmet/664

824

825 **Data Availability Statemanet**

826 The new DNA sequences used in this study are available from the NCBI

827 (Supporting file 3, Table S3). Complete concatenated alignment (except new sequences
828 generated in the present work) is available from Jardim de Queiroz and colleagues (Jardim de
829 Queiroz et al., 2020) DOI: 10.17632/wccvm8p5gx.1. XML files to reconstruct and calibrate
830 the phylogeny, and to reconstruct habitat preference are available from Supporting file 1.
831 Custom codes used to reconstruct the biogeography with RevBayes are available from
832 Supporting file 2 and from the GIT repository
833 https://bitbucket.org/XavMeyer/biogeographyrevscript/src/master/.

834 **Author Contribution**

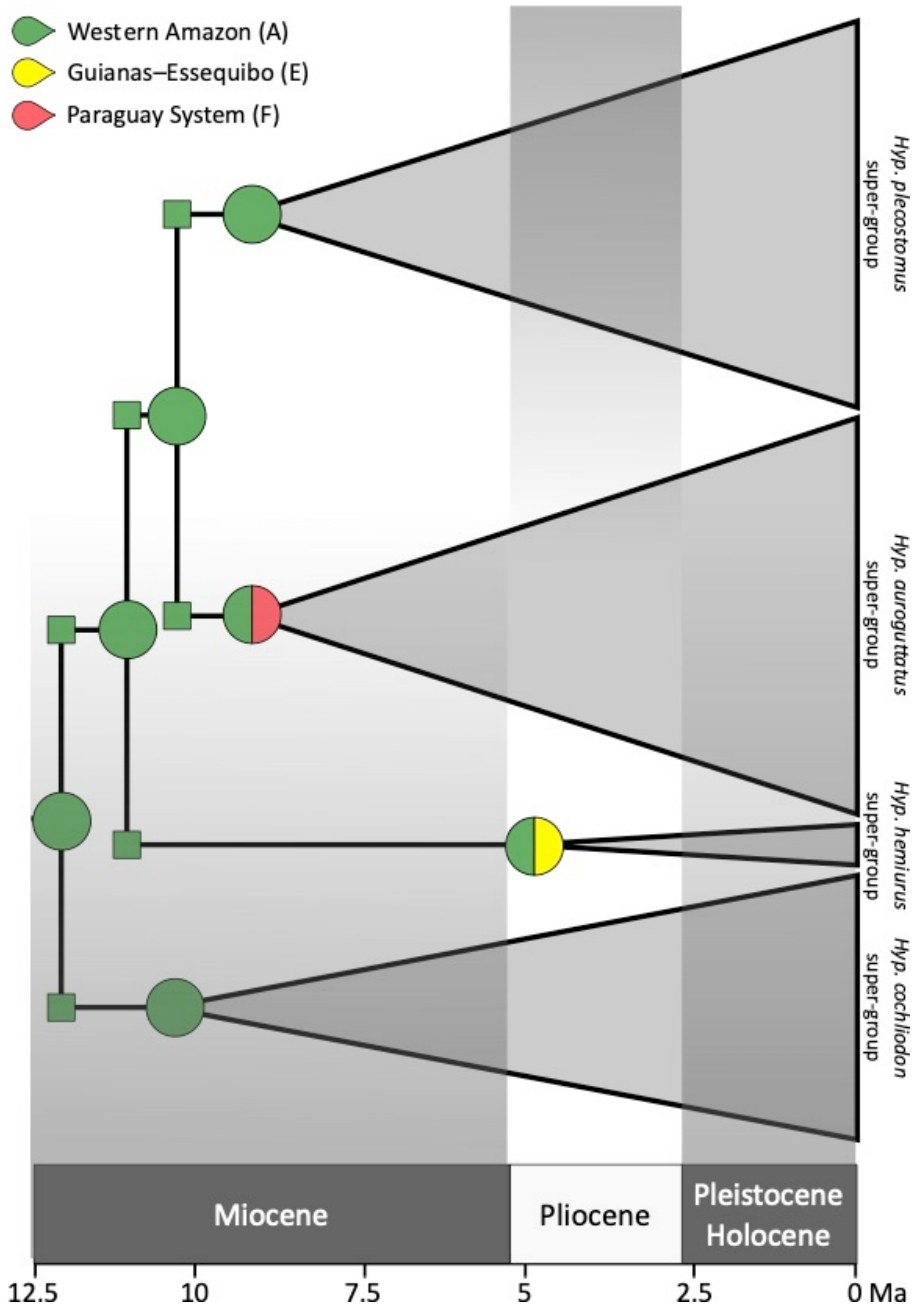
835 LJQ and JIMB designed the research, analyzed and interpreted the results and drafted the
836 manuscript. LJQ assembled the data and ran the analyses. IAB conducted part of the wet-
837 laboratory routine. XM wrote the helper script specifying RevBayes analyses, helped to run
838 the biogeographic analyses and to interpret the results. YC, RC, GTV, TEP and PAB
839 provided samples and/or laboratorial infrastructure to generate part of the data. All authors
840 read, contributed to the content and approved the final version.
841

842 Table 1. Retrieved ages and Highest Posterior Density (HPD) 95% (between brackets) of six representative
843 nodes of the *Hypostomus* phylogenetic tree. Four calibration analyses were performed (Supporting file 1). In
844 *scenario 1*, from which the results of this work are based on, all four-calibration points were used as priors; in
845 *scenario 2,* the calibration point 3 (Tocantins–Amazon disconnection) was not included; in *scenario 3*, the
846 calibration point 4 (Amazon–Pilcomayo headwater capture) was not included; and in *scenario 4*, calibration
847 point 2 (Uplift of Ecuadorian Andes) was not included. MRCA: most recent common ancestor.

Node	Biogeographical event and age	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Origin of *Hypostomus* + closest relatives *Hemiancistrus*	–	14.7 [17.8–11.4]	14.6 [17.7–11.7]	14.8 [18.0–12.0]	20.3 [27.4–14.2]
Hypostomus MRCA	–	12.1 [14.3–10.1]	12.1 [14.4–10.0]	12.2 [14.5–10.2]	16.3 [21.6–11.8]
Calibration point 1: *Hyp. hondae* + *Hyp. plecostomoides*	Uplift of Cordillera de Mérida (8 Ma)	7.3 [7.9–7.0]	7.3 [7.9–7.0]	7.4 [7.9–7.0]	7.7 [8.6–7.0]
Calibration point 2: *Isorineloricaria* + *Aphanotorlus*	Uplift of Ecuadorian Andes (10 Ma)	10.8 [12.0–9.8]	10.9 [12.0–9.8]	10.9 [12.0–9.8]	19.6 [27.3–13.0]
Calibration point 3: *Hyp.* sp.'xin2' + *Hyp.* sp. 'Tar'	Tocantins–Amazon disconnection (3.6–1.8 Ma)	2.6 [3.9–1.3]	2.7 [4.8–1.0]	2.6 [5.3–1.3]	3.0 [4.4–1.9]
Calibration point 4: *Hyp.* cf. *borelli* + *Hyp.* sp.'Rio Grande'	Amazon–Pilcomayo headwater capture (~1 Ma)	1.2 [1.7–0.6]	1.2 [1.7–0.7]	1.3 [2.0–0.7]	1.5 [2.3–0.8]

848

849



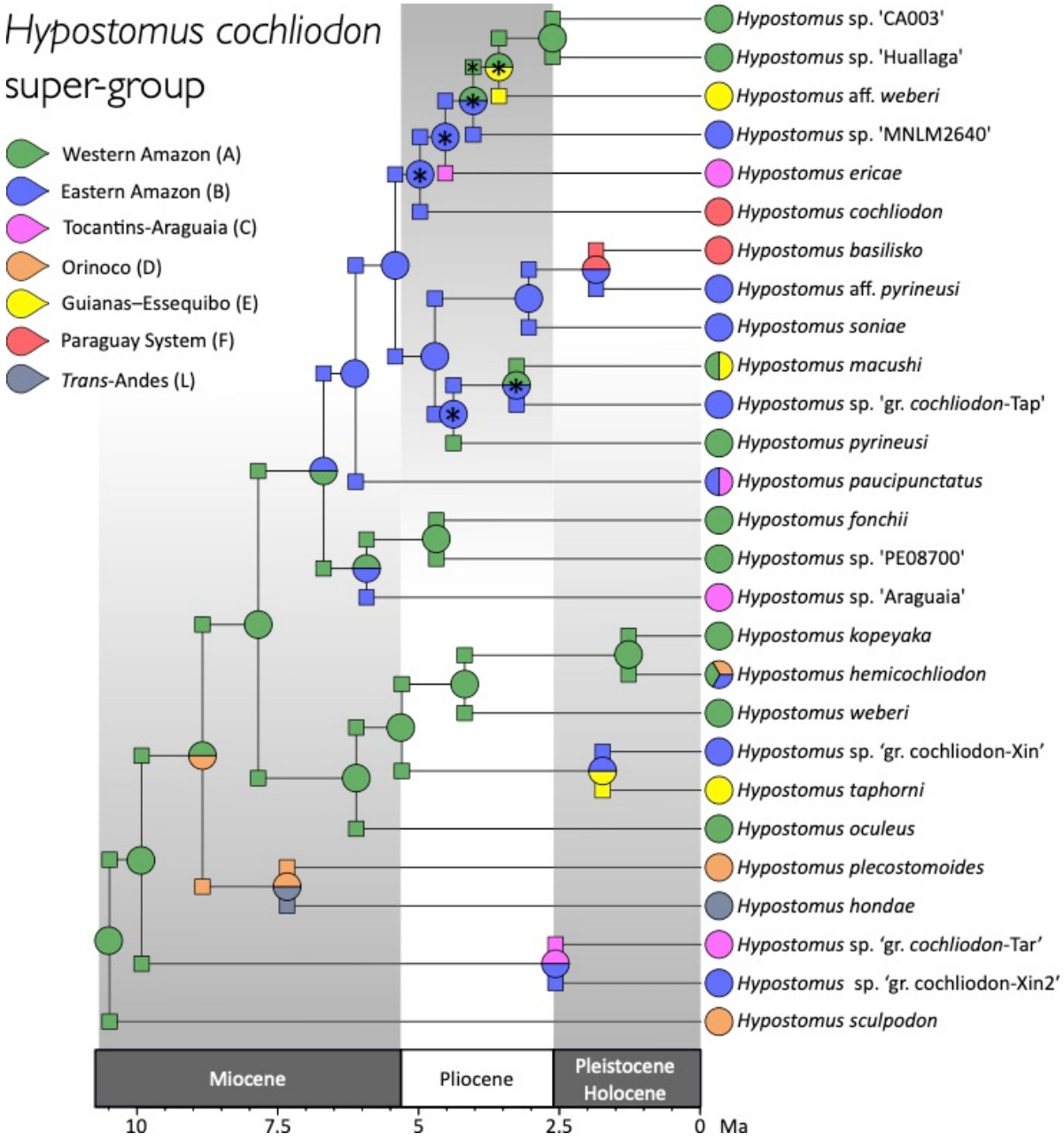
850

851 Figure 1 Ancestral range reconstruction of the first ancestral *Hypostomus* and of the ancestors of the four main

852 internal lineages. For detailed results, see Figs. 2–5. The posterior probabilities of the ancestral range

853 reconstructions shown here are ≥ 0.55.

854

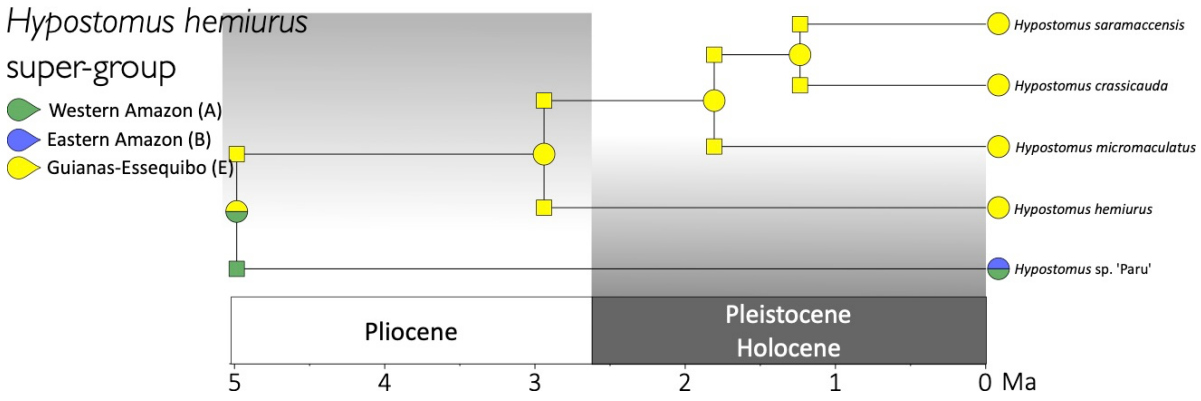


855

856 Figure 2 Ancestral range reconstruction of the *Hypostomus cochliodon* super-group. (* posterior probability of

857 ancestral range reconstruction < 0.5).

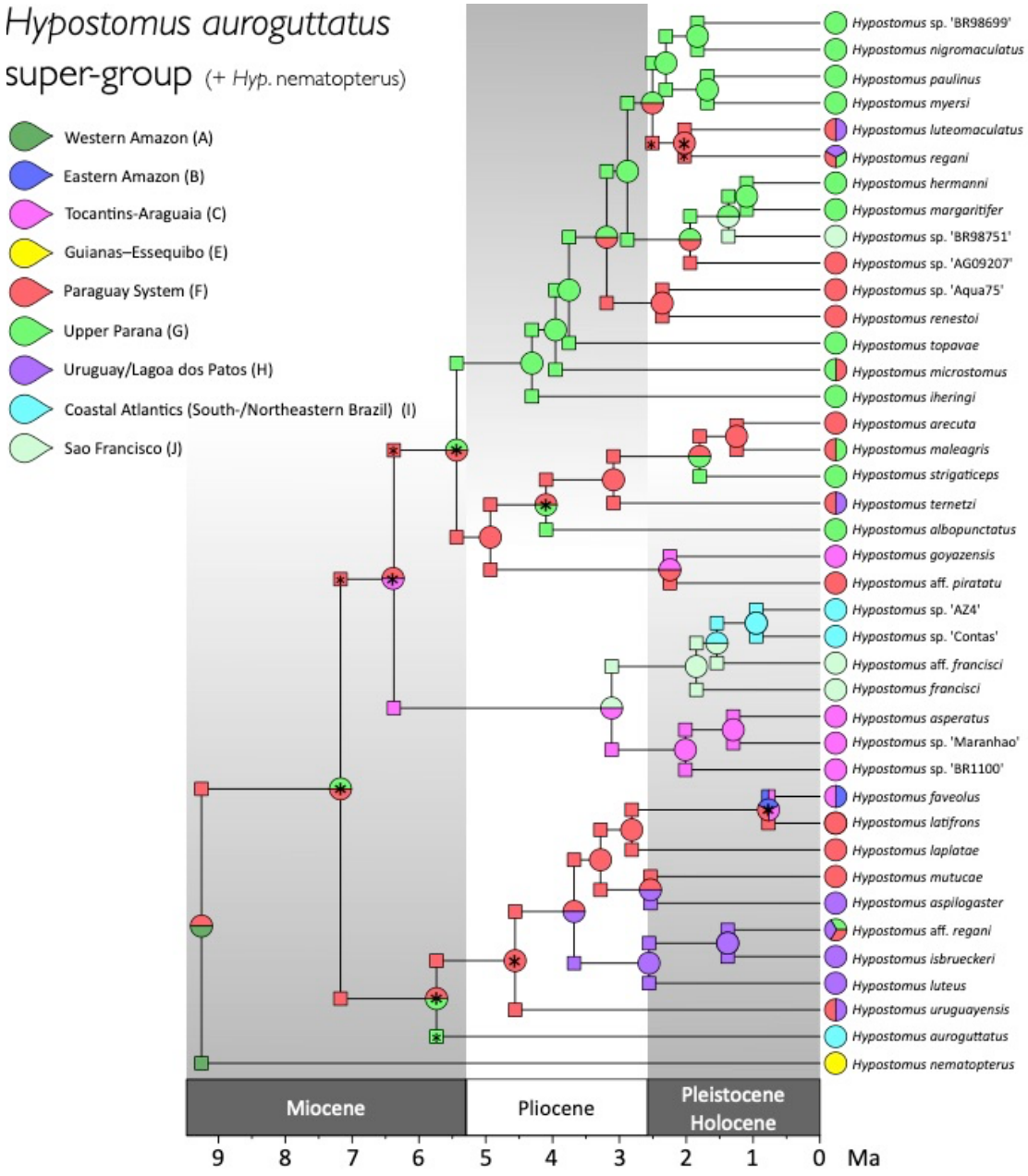
858



859

860 Figure 3 Ancestral range reconstruction of the *Hypostomus hemiurus* super-group. The posterior probabilities of

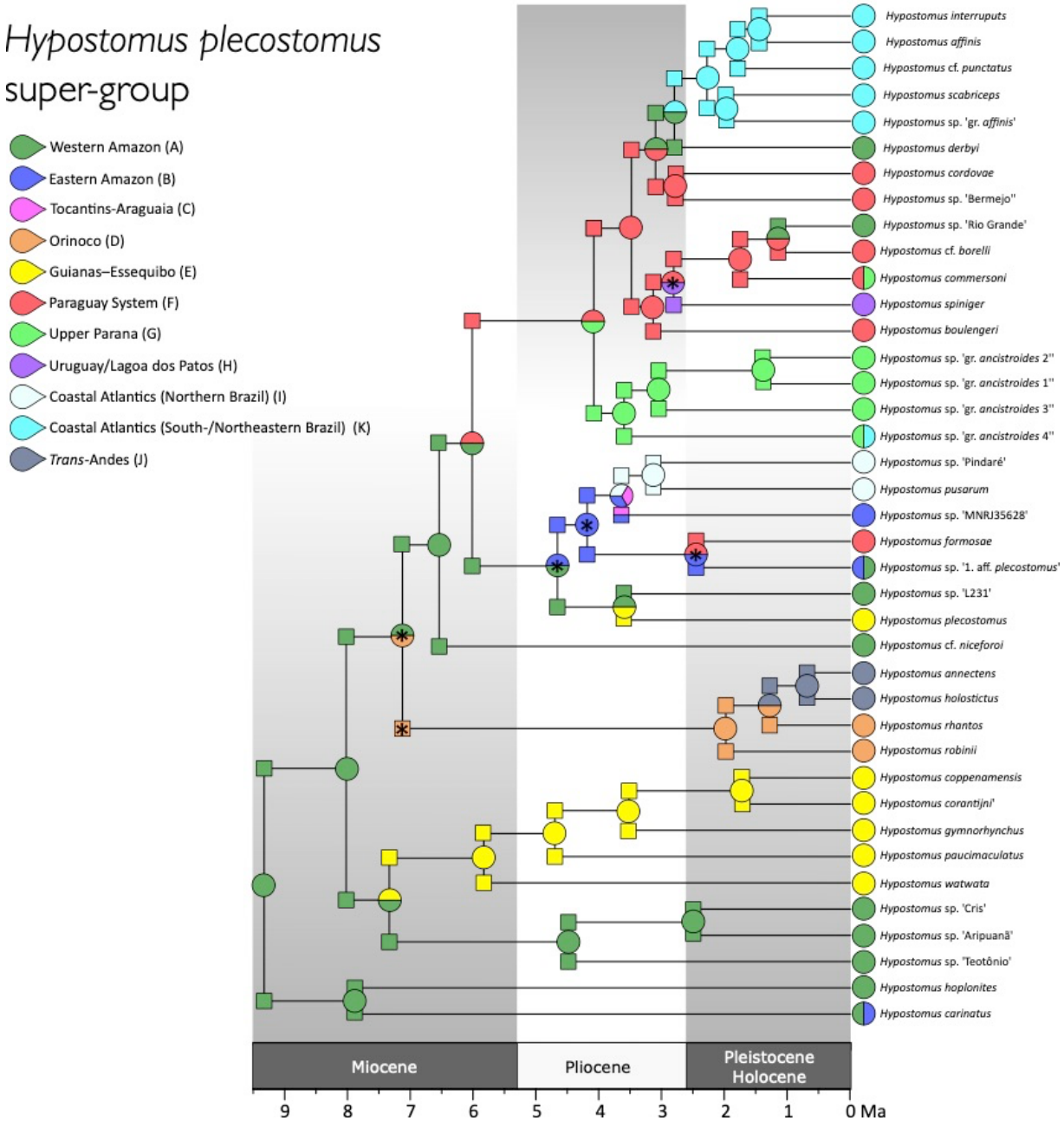
861 ancestral range reconstructions shown here are ≥ 0.54).



862

863 Figure 4 Ancestral range reconstruction of the clade that contains the *Hypostomus auroguttatus* super-group and

864 *Hypostomus nematopterus*. (* posterior probability of ancestral range reconstruction < 0.5).

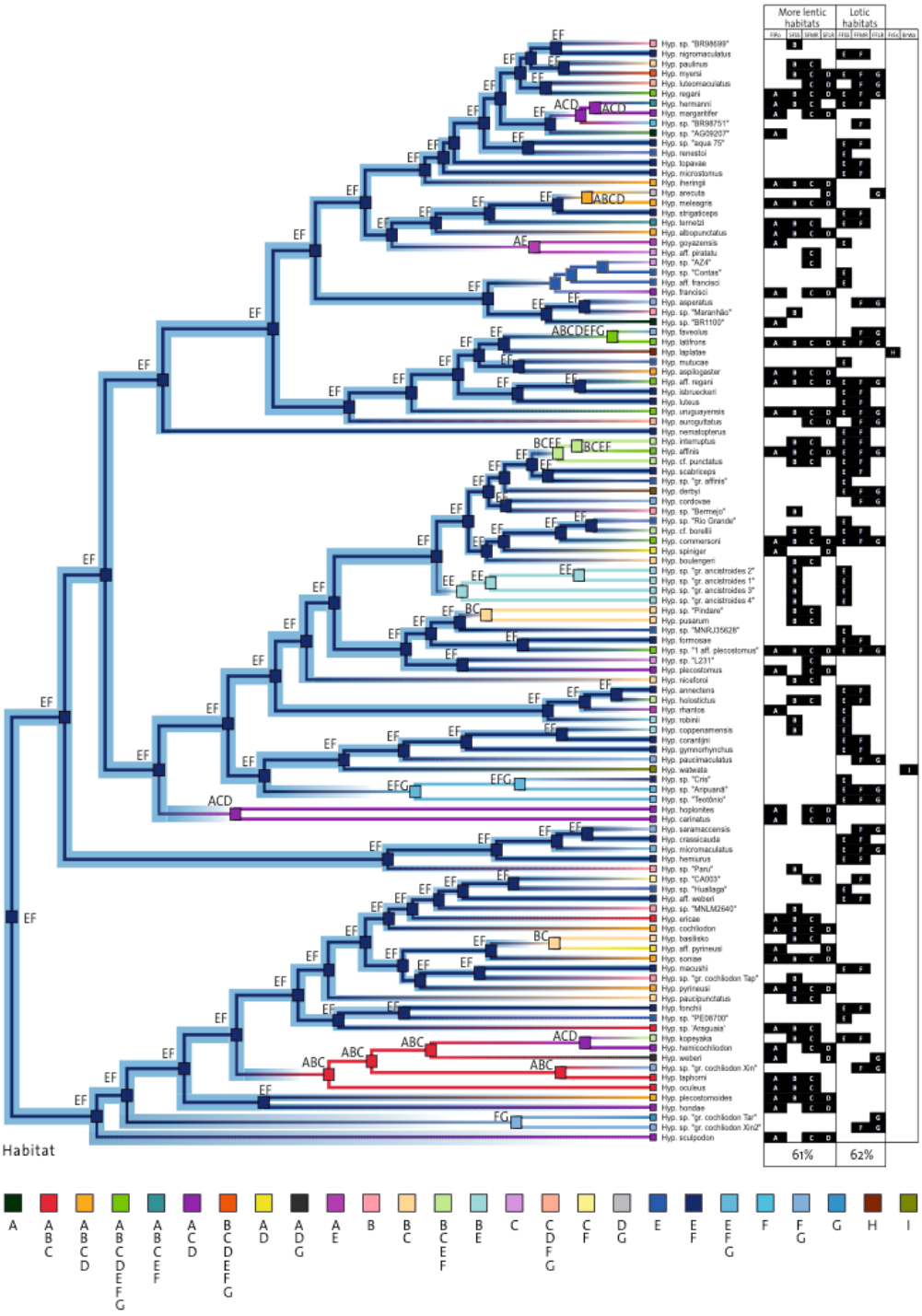


865

866 Figure 5 Ancestral range reconstruction of the *Hypostomus plecostomus* super-group. (* posterior probability of

867 ancestral range reconstruction < 0.5).

868



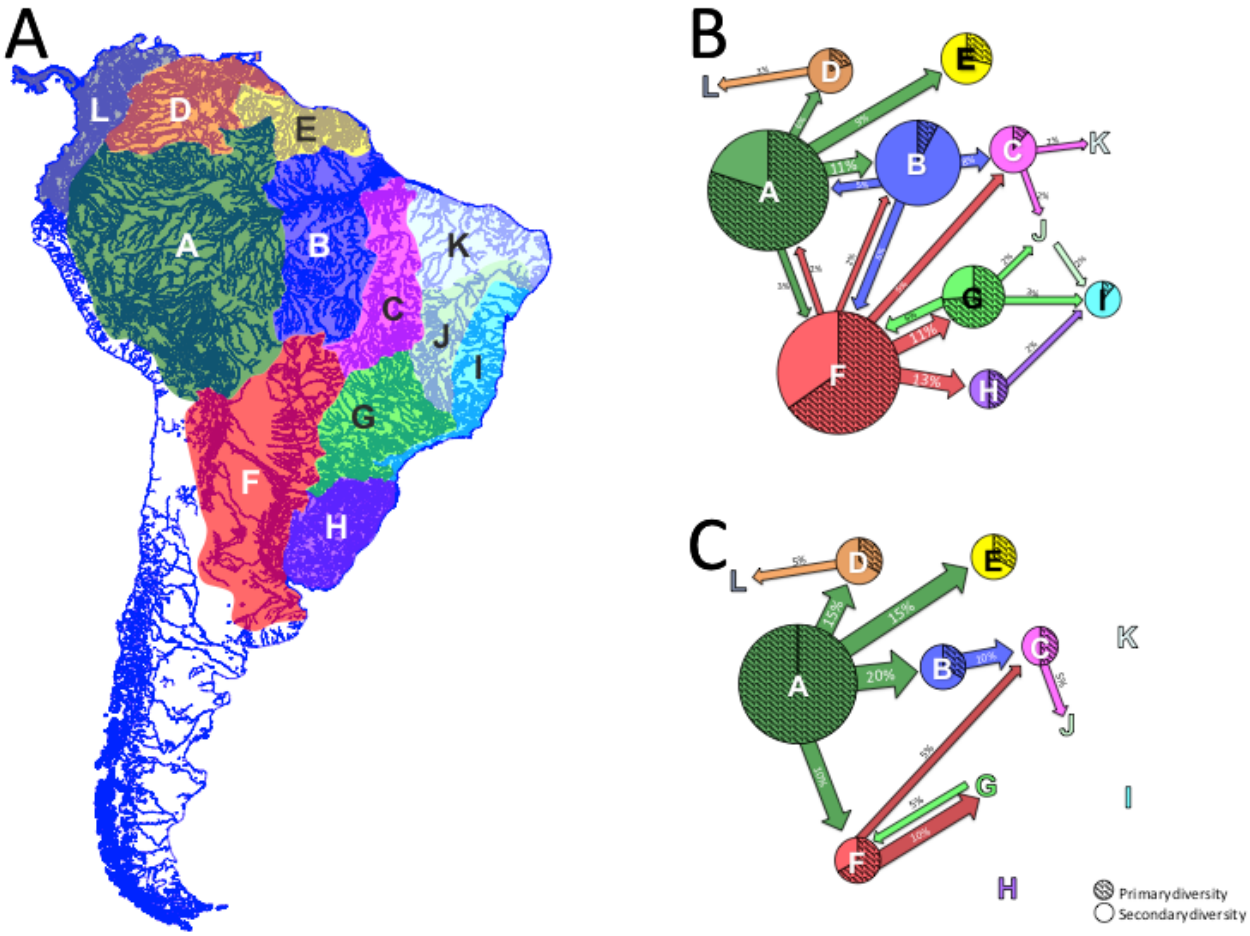
869

870 Figure 6 Ancestral habitat reconstruction of *Hypostomus*. A = floodplains and ponds (FlPo); B = slow flowing

871 small streams (SFSS); C = slow flowing medium rivers (SFMR); D = Slow flowing large rivers (SFLR); E =

872 medium to fast flowing small stream (FFSS); F = medium to fast flowing medium rivers (FFMR); G = medium

873 to fast flowing large rivers (FFLR); H = freshwater estuarine systems (FrEs); I = brackish water (BrWa). More
874 information, including habitat reconstruction of outgroups and posterior probabilities can be found in Supporting
875 file 1.



876

877 Figure 7 Dispersal events between biohydrogeographic regions (BHG). A) Map showing the division of South
878 America into BHG regions. B) Dispersal events inferred from 12.1 Ma to the present. C) Dispersal events
879 inferred from 12.1 to 6 Ma. Circle sizes are proportional to the number of extant *Hypostomus* species. Dashed
880 areas of the circles represent "primary diversity", i.e. the proportion of *Hypostomus* species that originated from
881 the first colonization of the region. Species that arose from secondary colonization events were considered as
882 "secondary diversity". As the Western Amazon (region A) was the place where *Hypostomus* originated, we
883 considered as primary diversity only the species that showed no ancestor distributed in another BHG region.
884

885 **Supporting information (doi:10.17632/2fh5rj2zvg.1)**

886 **Supporting file 1:** XML files for phylogenetic inference and tree calibration, and resulting
887 trees in nexus format. These files correspond to four different runs (four scenarios as
888 described in Table 1 in the main article). XML file for ancestral habitat reconstruction and
889 resulting tree in nexus format are also available.

890 **Supporting file 2:** RevBayes scripts that were modified to constraint a given list of
891 geographic ranges.

892 **Supporting file 3. Tables S1–S3. Table S1**. Marginal likelihoods obtained from the distinct
893 analyses conducted with RevBayes to reconstruct ancestral ranges. **Table S2.**
894 Biohydrogeographic regions and their connectivity through time periods, according to our
895 time-stratified model (M3). **Table S3.** Accession numbers of sequences generated in the
896 present work.