

Supporting Information

Appendix S1. Extended Methods.

Appendix S1A. Details on molecular data and laboratory protocols. Sequences from the previous studies included five mitochondrial (*ATP synthase subunits 8 and 6*, *cytochrome b*, and ribosomal *12S* and *16S*) and one nuclear (*recombination activating gene* [*rag*] *2*; 73 species only) gene fragments (Betancur-R. *et al.* 2007; Betancur-R. 2009). New sequences generated here included *rag2* for 49 (out of 54) species not examined previously, *rag1* (encompassing part of *exon 1*, complete *intron 1*, and part of *exon 2*; 1073 bp) for 105 species, and partial *cardiac muscle myosin heavy chain 6 alpha* (639 bp) for 121 species. Amplification protocols via nested PCR and primers for *rag2* and *myh6* followed Betancur R. *et al.* (2007) and Li *et al.* (2007), respectively. Primers for *Rag1 exons 1-2* (*rag1-F74*, *rag1-R1333*, and *rag1-F89I*) are listed in Sullivan *et al.* (2006), with one additional designed specifically for Ariidae for the second (nested) PCR (*rag1-R1054*.Ari: 5'-CTGTACTTTTGGATGCAGGTGCG-3'). To improve estimates of divergence times at the root of the ingroup, *rag1 exon 3* sequences (1494 bp) from GenBank were also obtained for five ariids (representing major lineages) and 19 outgroup taxa. The final alignment length included 7190 sites and 11 gene fragments (five mitochondrial and six nuclear). The material examined is listed in Appendix S1B.

Appendix S1B. Material examined and GenBank Accession numbers (see XLS file; after Betancur-R. 2009). New sequences generated here are given in bold. Two letter country

codes follow ISO-3166. NV, not vouchered; NCA, not catalogued.

Appendix S1C. Phylogenetic reconstructions and calibrations used for divergence time estimations. Maximum likelihood (ML) searches were conducted in RAxML (Stamatakis 2006; Stamatakis *et al.* 2008) and TreeFinder v. 2008 (Jobb 2008); Bayesian inferences (BI) were conducted in BEAST 1.6.1 (Drummond *et al.* 2006; Drummond & Rambaut 2007). Sequence alignment was performed in Muscle (Edgar 2004). The datasets were partitioned into eight *a priori* categories as follows: (1–3) two mitochondrial-protein-coding genes by codon position (respectively), (4–6) four nuclear exons by codon position (respectively), (7) two mitochondrial ribosomal genes, and (8) a nuclear intron. Best-fit models were estimated for each partition separately (i.e., unlinked) using the Akaike information criterion (AIC), implemented in TreeFinder, as follows: GTR + Γ (partitions 1, 4–6), J1 (= TIM) + Γ (partition 2), J2 (= TIM) + Γ (partitions 3, 7), and HKY (partition 8). The ML analyses in TreeFinder were run with these settings, and model parameters obtained from that analysis were set as priors for the BEAST analysis. Fifty alternative RAxML searches were conducted under the GTRGAMMA model, and branch support was assessed using the rapid bootstrapping algorithm with automatic estimation of runs.

Divergence times were estimated in BEAST using the uncorrelated log-normal (UCLN) clock-model (Drummond *et al.* 2006). A starting chronogram that satisfied all priors (e.g., monophyly and initial divergence times) was generated under penalized likelihood in r8s v1.71 (Sanderson 2003) using the RAxML tree. Clock and tree priors were linked across partitions, and followed a birth-death tree prior for speciation. The

Markov chain Monte Carlo (MCMC) analyses were run in triplicate for 100 million generations, sampling every 1000 generations. Convergence in the marginal distributions was analyzed using the Effective Sample Size (ESS) statistic and by visually inspecting the likelihood plots in the program Tracer v 1.5 (Drummond *et al.* 2007). Stationarity was assumed for parameters with ESS values >200 as well as for asymptotic plots (Drummond *et al.* 2006).

Nineteen calibration points (fossils or geologic events) on four deep osteichthyan nodes, three non-siluriform ostariophysan nodes, five non-ariid siluriform nodes, and seven ariid nodes were set as age constraints (using median ages for fossil calibrations). Prior distributions for each calibration were chosen according to degree of completeness/certainty (see Ho & Phillips 2009; Pyron 2010), ranging from log-normal (geologic events or complete fossil record) to uniform (poor fossil record). For instance, the estimated origin of siluriforms according to molecular evidence (175-130 mya; Briggs 2005; Peng *et al.* 2006; Lundberg *et al.* 2007) is much older than predicted by the oldest fossil dated from the Late Campanian-Early Maastrichtian (73–68 mya), suggesting a stratigraphic gap. Thus, priors for most catfish (as well as for many other ostariophysan) nodes were constrained under uniform distributions using minimum-age offset values. Priors for the BEAST analysis are given in Table S1; calibrations points are as follows:

(1) Osteichthyes (= Teleostomi). MRCA: *Latimeria*, *Danio*. Lower bound (418 my): *Andreolepis hedei* (Marss 2001; Hurley *et al.* 2007). Upper bound (505 my): Burgess

Shale deposits (Santini *et al.* 2009). Additional references for calibration: Santini *et al.* (2009); Pyron (2010).

(2) Actinopterygii. MRCA: *Polypterus*, *Danio*. Lower bound (392 my): stem actinopterans (Gardiner 1993). Upper bound (418 my): stem *Andreolepis hedei* (Marss, 2011; Hurley *et al.*, 2007). Additional references for calibration: Hurley *et al.* (2007).

(3) Actinopteri. MRCA: *Polyodon*, *Danio*. Lower bound (345 my): *Cosmoptychius* (Coates 1999). Upper bound (392 my): stem actinopterans (Gardiner 1993). Additional references for calibration: Hurley *et al.* (2007).

(4) Neopterygii. MRCA: *Amia*, *Danio*. Lower bound (284 my): *Brachydegma caelatum* (Hurley *et al.* 2007). Upper bound: (370 my): based on maximum estimations by Hurley *et al.* (2007). Additional references for calibration: Hurley *et al.* (2007).

(5) Ostariophysi. MRCA: *Chanos*, *Danio*. Lower bound (146 my): *Tischlingerichthys vohli* (oldest Ostariophysi; Arratia 1997) and *Rubiesichthys gregalis* (oldest gonorrhynchiform; Potayo-Ariza 1996). Upper bound: (284 my; poorly defined): *Brachydegma caelatum* (Hurley *et al.* 2007). Note: it is difficult to assign non-uniform distributions given the incompleteness of the ostariophysan record.

(6) Cyprinoidea + Cobitoidea. MRCA: *Cobitis*, *Danio*. Lower bound (67 my): Paleocene stem catostomids from North America (Cavender 1986). Upper bound: (284 my; poorly

defined): *Brachydegma caelatum* (Hurley *et al.* 2007). Note: it is difficult to assign non-uniform distributions given the incompleteness of the ostariophysan record.

(7) Characiformes. MRCA: *Distichodus*, *Chalceus*. Lower bound (97 my): characiform teeth (Malabarba & Malabarba 2010). Upper bound: (284 my; poorly defined): *Brachydegma caelatum* (Hurley *et al.* 2007). Notes: it is unclear whether the fossil teeth are stem or crown characiforms; stem-based calibration was used (“includeStem” statistic in BEAST). It is difficult to assign non-uniform distributions given the incompleteness of the ostariophysan record.

(8) Doradoidea + Aspredinidae. MRCA: *Ageneiosus*, *Pterobunocephalus*. Lower bound (67 my): doradoid (Lundberg *et al.* 2007). Upper bound: (146 my): *Tischlingerichthys viohli* (Arratia 1997). Additional references for calibration: Lundberg *et al.* (2007). Note: it is difficult to assign non-uniform distributions given the incompleteness of the ostariophysan record.

(9) Ictaluridae. MRCA: *Ameiurus*, *Ictalurus*. Lower bound (36 my): *Ameiurus pectinatus* (references in Lundberg *et al.* 2007). Upper bound: (64 my): stem ictalurid, *Astephus* sp. (References in Lundberg *et al.* 2007). Additional references for calibration: Lundberg *et al.* (2007). Note: it is difficult to assign non-uniform distributions given the incompleteness of the ostariophysan record.

(10) Ictaluridae + Cranoglanidae. MRCA: *Ictalurus*, *Cranoglanis*. Lower bound (64 my): stem ictalurid, *Astephus* sp. (References in Lundberg *et al.* 2007). Upper bound: (146 my): *Tischlingerichthys viohli* (Arratia 1997). Additional references for calibration: Lundberg *et al.* (2007). Note: it is difficult to assign non-uniform distributions given the incompleteness of the ostariophysan record.

(11) Callichthyidae. MRCA: *Callichthys*, *Corydoras*. Lower bound (58 my): *Corydoras revelatus* (references in Lundberg *et al.* 2007). Upper bound: (146 my): *Tischlingerichthys viohli* (Arratia 1997). Additional references for calibration: Lundberg *et al.* (2007). Note: it is difficult to assign non-uniform distributions given the incompleteness of the ostariophysan record.

(12) Pimelodidae. MRCA: *Steindachneridion*, *Leiarius*. Lower bound (25 my): *Steindachneridion* spp. (Santos 1973). Upper bound: (146 my): *Tischlingerichthys viohli* (Arratia 1997). Additional references for calibration: Hardman and Lundberg (2006). Note: it is difficult to assign non-uniform distributions given the incompleteness of the ostariophysan record.

(13) Ariidae. MRCA: *Galeichthys*, *Arius*. Lower bound (70 my): ariids remains from North and South America; Gayet and Meunier, 2003. Upper bound: (146 my): *Tischlingerichthys viohli* (Arratia 1997). Additional references for calibration: Lundberg *et al.* (2007); Betancur-R. and Armbruster (2009); Betancur-R. (2009).

(14) *Cathorops*. MRCA: *C. dasycephalus*, *C. mapale*. Lower bound (20 my): (*Cathorops* sp.; Aguilera & De Aguilera 2004). Upper bound: (70 my): ariids remains from North and South America (Gayet & Meunier 2003). Additional references for calibration: Betancur-R. and Armbruster (2009); Betancur-R. (2009). Note: Probably stem within the subgenus *Cathorops*; it was assigned to the divergence of the subgenera *Cathorops* and *Precathorops*.

(15) *Potamosilurus*. MRCA: *Potamosilurus macrorhynchus*, *P. velutinus*. Lower and upper bounds (4.0–8.0 my): uplift of New Guinea Central mountain range (van Ufford & Cloos 2005; Betancur-R. 2010).

(16) *Cathorops mapale* + *C. fuerthii*. MRCA: *Cathorops mapale*, *C. fuerthii*. Lower and upper bounds (2.8–3.1 my): last rising of Panama isthmus (Coates & Obando 1996). Additional references for calibration: Lessios (2008); Betancur-R. (2009).

(17) *Notarius* aff. *kessleri* + *N. neogranatensis*. MRCA: *Notarius* aff. *kessleri*, *N. neogranatensis*. Same as 16.

(18) *Ariopsis seemanni* + *Ariopsis* sp. MRCA: *Ariopsis seemanni*, *Ariopsis* sp. Same as 16.

(19) *Cathorops mapale* + *C. wayuu*. MRCA: *Cathorops* sp., *C. mapale*. Lower and upper bounds (0.8 my): northward displacement of Santa Marta Massif (Betancur-R. *et al.* 2010).

Appendix S1D. Habitat-dependent diversification. To test for differences in diversification rates and dispersal between freshwater and marine environments, we used the Geographic-state Speciation and Extinction (GeoSSE) model (Goldberg *et al.* 2011) and implemented in the R package 'diversitree.' We coded euryhaline, marine, and freshwater species as 0, 1, and 2, respectively. We used the sampling fraction function (sampling.f) to account for missing taxa in the phylogeny (0= 0.824, 1= 0.788, 2= 0.614). The GeoSSE model implements parameters for speciation in states 0, 1, and 2, extinction in states 1 and 2, and dispersal from 1 to 2 and vice versa, for a total of seven parameters. We tested a set of 10 distinct models using the time-calibrated tree described above. We first tested a model in which all parameters were free to vary, and compared this against a model in which all habitat-specific rates were equal and among-habitat speciation was zero, all habitat-specific rates were equal and among-habitat speciation was allowed, speciation, extinction, and dispersal rates were equal, both alone and in combination, while allowing among-habitat speciation. This represents the full range of models.

Appendix S1E. Morphological data. We quantified morphological disparity using data for traditional morphometrics (body shape), maximum total length (MTL), and meristics (counts; see below). Measurements were taken with either a ruler and recorded to the nearest millimeter or centimeter (large specimens) or with dial calipers and recorded to

the nearest 0.1 mm. RBR, APM, and AMS examined a total of 912 ariid specimens from over 20 museums, representing 118 (out of 123) species in the phylogeny. Forty morphometric variables were recorded as specified in Betancur-R. and Willink (2007) and Marceniuk (2007), as follows: total length*, standard length, head width*, head depth*, snout length, mouth width, upper-lip width, maxillary barbels, mandibular barbels, mental barbels, distance between lateral ethmoids, distance between frontals, distance between supracleitra, anterior internarial distance, posterior internarial distance, interorbital distance, eye diameter, postorbital length*, width of supraoccipital process*, length of supraoccipital process*, head length, body depth*, body width, predorsal-fin length, dorsal-fin base*, dorsal-fin spine height, distance between dorsal and adipose fins*, preadipose-fin length, adipose-fin base, adipose-fin height*, prepectoral-fin length, pectoral-fin base*, pectoral-fin spine length, prepelvic-fin length, pelvic-fin base*, pelvic-fin length*, preanal-fin length, anal-fin base, anal-fin height, caudal peduncle depth (those characters indicated by an asterisk “*” were further removed; see below).

We vetted the morphometric matrix by regressing SL against the remaining variables and obvious outliers (putatively errors in measurement) were removed. To reduce the effects of allometric variation, sexual dimorphism, lack of consistency among data collectors, and missing data, 246 specimens (mostly juveniles) and 13 variables were discarded from the matrix. The final morphometric matrix included 666 specimens from 118 species (1–33 specimens per species, mean = 5.6, mode = 2), 27 variables, and 7.4% missing data. Following logarithmic transformation of the data, missing entries were imputed using the expectation maximization multivariate method (Strauss *et al.* 2003) in the package Mclust v3.4 (Fraley & Raftery 2006) implemented in R v.2.13 (R

Development Core Team 2011). This strategy produces much lower Type I error rates than typical strategies, such as mean replacement (Strauss *et al.* 2003). Imputation was desirable, as downstream analyses were based on mean trait values for species, and without imputation sample sizes would differ among traits, potentially skewing results.

MTL and meristic data were obtained from the museum specimens or compiled from the literature. In cases where maximum standard length (SL) rather than MTL was reported, we conducted regression analyses for data interpolation. Meristic data included modal values for each species for number of rakers on first gill arch and anal-fin rays. Preliminary disparity analyses (see text) conducted on the morphometric, MTL, and meristic data separately vs. all data combined yielded similar results; for simplicity and to capture as much variation as possible, we report analyses on the combined dataset only. Log-transformed MTL data and square-root-transformed meristic data were appended to the morphometric data matrix after estimating mean values for each species. The final morphologic matrix included 118 species and 30 variables (file available from the Dryad repository doi:10.5061/dryad.35d85c0p).

To reduce dimensionality in the morphologic matrix, a principal component analysis (PCA) was performed on the transformed data using the correlation matrix. We conducted the phylogenetically corrected PCA method (implemented in R) to account for evolutionary dependence (Revell 2009). The first PC axis captures mostly size variation; the remaining components explain mostly size-free disparity in shape (Jolliffe 2002), for which we used PC2 and PC3.

Appendix S2. Phylogeny, divergence times, and habitat optimization. The relationships among ariid taxa inferred under ML and relaxed Bayesian criteria here as well as those obtained by the previous study (Betancur-R. 2009) are all highly congruent. As compared to the previous study, however, our analyses (using more nuclear genes, more outgroup taxa, and more calibrations) resulted in younger divergence time estimates for the origin of ariids and sub-clades (although in most cases the 95% highest posterior density [HPD] overlap). While discrepancies in absolute dates may have implications for interpretation of the historical biogeography of the group (see Betancur-R. 2009 for a discussion), they are unlikely to impact relative rates of diversification.

To test this, we plotted the branching times in our tree against those from the previous uncorrelated lognormal chronogram and found a significantly strong correlation ($r^2 = 0.98$; $P < 0.000$). We repeated this procedure using the upper HPD bounds for the nodal times (i.e., oldest estimates) and obtained similar results ($r^2 = 0.99$; $P < 0.000$); this suggests that our analyses are robust to the observed differences in absolute dates. Our current estimates date the origin of crown ariids at ca. 46 my. These estimations imply that ariid fossils predating the divergence of the crown group are stem taxa (e.g., the Late Campanian-Early Maastrichtian fossils; see Gayet & Meunier 2003).

Table S1. Calibration points for divergence time estimations and priors for the BEAST analysis. MRCA, most recent common ancestor. L, lower bound; U, upper bound. See also Appendix S1C.

No.	Node	L (my)	U (my)	Distribution	UCLN priors	
					Mean	St. Dev.
1	Osteichthyes (= Teleostomi)	418	505	Lognormal	6.1300	0.0482
2	Actinopterygii	392	418	Lognormal	6.0034	0.0164
3	Actinopteri	345	392	Lognormal	4.6196	0.1857
4	Neopterygii	284	370	Lognormal	5.7946	0.0743
5	Ostariophysi	146	284	Uniform	–	–
6	Cyprinoidea + Cobitoidea	67	284	Uniform	–	–
7	Characiformes	97	284	Uniform	–	–
8	Doradoidea + Aspredinidae	67	146	Uniform	–	–
9	Ictaluridae	36	64	Uniform	–	–
10	Ictaluridae + Cranoglanidae	64	146	Uniform	–	–
11	Callichthyidae	58	146	Uniform	–	–
12	Pimelodidae	25	146	Uniform	–	–
13	Ariidae	70	146	Exponential	24.5000	–
14	<i>Cathorops</i>	20	70	Exponential	16.8600	–
15	<i>Potamosilurus</i>	4	8	Lognormal	1.7329	0.1768
16	<i>Cathorops mapale</i> + <i>C. fuerthii</i>	2.8	3.1	Lognormal	1.0805	0.0260
17	<i>Notarius</i> aff. <i>kessleri</i> + <i>N. neogranatensis</i>	2.8	3.1	Lognormal	1.0805	0.0260
18	<i>Ariopsis seemanni</i> + <i>Ariopsis</i> sp.	2.8	3.1	Lognormal	1.0805	0.0260
19	<i>Cathorops mapale</i> + <i>C. wayuu</i>	0.8	0.8	Lognormal	-0.2410	0.0968

Table S2. Putative clades/lineages involved in independent events of freshwater colonization (see Fig. 1 for clade numbers), including euryhaline lineages. Net diversification rates are calculated using the method-of-moments estimator on stem ages for freshwater clades including >3 species (in bold).

No.	Clade/taxon	Stem age	Total spp. (including missing taxa)	Habitat	Net diversification rates
1	<i>N. bonillai/N. cookei</i>	3	2	Freshwater	—
2	<i>C. festae</i>	2	1	Freshwater	—
3	<i>C. aguadulce/melanopus/kailolae</i>	12	3	Freshwater	0.06
4	<i>C. tuyra/C. hypophthalmus</i>	7	2	Euryhaline	—
5	<i>C. agassizii/C. nuchalis</i>	1.6	2	Freshwater	—
6	<i>S. dowii</i>	10	1	Euryhaline	—
7	<i>Potamarius</i>	16	4	Freshwater	0.06
8	Madagascar plus eastern Africa	16	5	Freshwater	0.07
9	<i>C. melanochir/H. stormii</i>	15	7	Freshwater	0.09
10	<i>A. gagora</i>	10	1	Euryhaline	—
11	<i>C. borneensis</i>	7	1	Freshwater	—
12	Australia-New Guinea minus <i>N. midgleyi/N. paucus</i>	16	26	Freshwater	0.16
13	<i>N. midgleyi/N. paucus</i>	7	2	Freshwater	—
14	<i>P. magatensis</i>	<16*	1	Freshwater	—
15	<i>A. assimilis</i>	<16*	1	Freshwater	—
16	<i>N. phyrigiatus</i>	<8*	1	Euryhaline	—
17	<i>A. labiatus</i>	<25*	1	Freshwater	—
18	<i>C. gigas</i>	<20**†	1	Freshwater	—

*Maximum age based on presumed phylogenetic position; single sp./colonization event not represented in the phylogeny

Table S3. Relative rates of morphological evolution for multiple pairwise comparisons (see Fig. 2 for graphical visualization). Values >1 indicate higher rates in assemblage 1 (numerator) and <1 higher rates in Assemblage 2 (denominator); Δ AICc values greater than 4 (bold) suggest substantial statistical support. Values in parentheses are number of taxa on each assemblage. Euryhaline species were coded as marine and freshwater under criteria 1 and 2, respectively.

Comparisons	Untransformed branches			Transformed branches		
	PC1	PC2	PC3	PC1	PC2	PC3
Marine (79)/freshwater (39) – criterion 2	2.4	4.8	1.8	2.8	4.1	1.1
Δ AICc	0.2	25.7	5.1	9.0	21.5	8.2
Marine (92)/freshwater (26) – criterion 1	2.8	3.5	2.4	2.6	3.2	1.2
Δ AICc	3.7	12.9	4.6	5.0	11.4	3.8
Freshwater AU-NG (24)/ Marine – rest (79)	3.3	6.2	2.7	2.8	5.3	1.5
Freshwater – rest (15)/ Marine – rest (79)	0.8	2.5	0.3	2.4	2.2	0.3
Freshwater – AU-NG (15)/ Freshwater - rest (24)	4.0	2.4	8.3	1.2	2.4	4.3
Δ AICc	4.8	26.6	7.8	6.7	21.7	12.0

Table S4. Faunal composition of freshwater ariids from 12 basins compared to the remaining fish species in each basin. We selected rivers hosting ariid species for which ichthyofaunal inventories exist. To maintain the scale of comparison, only smaller rivers with less than 200 fish species were considered. Remaining fish species are calculated as total number of fish species (including primary freshwater and peripheral species) reported from these studies, minus the ariid species. For most rivers, estimation of ariid species follows our taxonomic expertise.

River/basin (country/region)	Total ariids	Remain. fishes	Main reference
Fly (Papua New Guinea)	16	70	Swales <i>et al.</i> (2000)
Sepik (Papua New Guinea)	5	52	Allen and Coates (1990)
Usumacinta (Guatemala and Mexico)	3	105	Miller (1966)
Tuyra (Panama)	2	71	Rodriguez-Olarte <i>et al.</i> (2011)
Magdalena (Colombia)	1	158	Rodriguez-Olarte <i>et al.</i> (2011)
Atrato (Colombia)	1	119	Rodriguez-Olarte <i>et al.</i> (2011)
Lago Izabal (Guatemala)	3	79	Pérez (2004)
Maracaibo (Venezuela)	1	114	Rodriguez-Olarte <i>et al.</i> (2011)
Northwestern basin (Madagascar)	4	67	Ng and Sparks (2003); Sparks and Stiassny (2003) http://wwf.panda.org/about_our_earth/ecoregions/salween_river.cfm
Salween (Myanmar)	2	138	Eigenmann (1921)
Guayas (Ecuador)	2	86	http://fish.mongabay.com/data/ecosystems/Volta.htm
Volta (west Africa)	1	163	tm

Table S5. Top five loadings for each principal component in the phylogenetic PCA.

No.	PC1		PC2		PC3	
	Variable	Loading value	Variable	Loading value	Variable	Loading value
1	Standard length	0.991	Number of anal-fin rays	-0.672	Number of gill rakers	-0.841
2	Preadipose-fin length	0.990	Upper lip width	0.503	Number of anal-fin rays	0.451
3	Preanal-fin length	0.988	Mandibulatory barbel length	-0.485	Maxillary barbel length	0.286
4	Prepelvic-fin length	0.987	Maxillary barbel length	-0.463	Upper lip width	0.256
5	Predorsal-fin length	0.982	Mental barbels length	-0.462	Snout length	-0.223

References

- Aguilera O. & De Aguilera D.R. (2004). Amphi-American Neogene sea catfishes (Siluriformes, Ariidae) from northern South America. In: *Fossils of the Miocene Castillo Formation, Venezuela: Contributions on Neotropical Palaeontology* (eds. Sánchez-Villagra MR & Clack JA). Special Papers in Palaeontology 71, pp. 29–48.
- Allen G.R. & Coates D. (1990). An ichthyological survey of the Sepik River, Papua New Guinea. *Rec. West. Aust. Mus.*, 34, 31-116.
- Arratia G. (1997). Basal teleosts and teleostean phylogeny. *Palaeo Ichthyologica*, 7, 1-168.
- Betancur-R. R. (2009). Molecular phylogenetics and evolutionary history of ariid catfishes revisited: a comprehensive sampling. *BMC Evol. Biol.*, 9, 175.
- Betancur-R. R. (2010). Molecular phylogenetics supports multiple evolutionary transitions from marine to freshwater habitats in ariid catfishes. *Mol. Phylogenet. Evol.*, 55, 249-258.
- Betancur-R. R., Acero P A., Duque-Caro H. & Santos S.R. (2010). Phylogenetic and morphologic analyses of a coastal fish reveals a marine biogeographic break of terrestrial origin in the Southern Caribbean. *PLoS ONE*, 5, e11566.
- Betancur-R. R., Acero P. A., Bermingham E. & Cooke R. (2007). Systematics and biogeography of New World sea catfishes (Siluriformes: Ariidae) as inferred from mitochondrial, nuclear, and morphological evidence. *Mol. Phylogenet. Evol.*, 45, 339-357.
- Betancur-R. R. & Armbruster J.W. (2009). Molecular clocks provide new insights into the evolutionary history of galeichthyine sea catfishes. *Evolution*, 63, 1232–1243.
- Betancur-R. R. & Willink P.W. (2007). A new freshwater ariid (Otophysi: Siluriformes) from the Río Usumacinta basin. *Copeia*, 2007, 818–828.
- Briggs J.C. (2005). The biogeography of otophysan fishes (Ostariophysi : Otophysi): a new appraisal. *J. Biogeogr.*, 32, 287-294.
- Cavender T.M. (1986). Review of the fossil history of North American freshwater fishes. . In: *The Zoogeography of North American Freshwater Fishes* (eds. Hocutt CH & Wiley EO). Wiley & Sons New York, pp. 700-724.
- Coates A.G. & Obando J.A. (1996). Geologic evolution of the Central American Isthmus. In: *Evolution and Environments in Tropical America* (eds. Jackson JB, Budd AF & Coates AG). University of Chicago Press Chicago, pp. 21-56.
- Coates M.I. (1999). Endocranial preservation of a Carboniferous actinopterygian from Lancashire, UK, and the interrelationships of primitive actinopterygians. *Phil. Trans. R. Soc. Lond. B.*, 354, 435-462.
- Drummond A.J., Ho S.Y.W., Phillips M.J. & Rambaut A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol.*, 4, 699-710.
- Drummond A.J., Ho S.Y.W., Rawlence N. & Rambaut A. (2007). A rough guide to BEAST 1.4. In. University of Edinburgh Edinburgh.

- Drummond A.J. & Rambaut A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.*, 7, 8.
- Edgar R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, 32, 1792-1797.
- Eigenmann C.H. (1921). The Nature and Origin of the Fishes of the Pacific Slope of Ecuador, Peru and Chili. *Proc. Am. Philos. Soc.*, 60, 503-523.
- Fraley C. & Raftery A.E. (2006). MCLUST Version 3 for R: Normal Mixture Modeling and Model-Based Clustering. In: *Technical Report no. 504*. Department of Statistics, University of Washington.
- Gardiner B.G. (1993). Osteichthyes: Basal Actinopterygians. In: *Fossil Record II* (ed. Benton MJ). Chapman and Hall London, pp. 611-619.
- Gayet M. & Meunier F.J. (2003). Paleontology and palaeobiogeography of catfishes. In: *Catfishes* (eds. Arratia G, Kapoor BG, Chardon M & Diogo R). Science Publishers Enfield, NH, pp. 491-522.
- Goldberg E.E., Lancaster L.T. & Ree R.H. (2011). Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Syst. Biol.*, 60, 451-65.
- Hardman M. & Lundberg J.G. (2006). Molecular phylogeny and a chronology of diversification for "phractocephaline" catfishes (Siluriformes : Pimelodidae) based on mitochondrial DNA and nuclear recombination activating gene 2 sequences. *Mol. Phylogenet. Evol.*, 40, 410-418.
- Ho S.Y.W. & Phillips M.J. (2009). Accounting for Calibration Uncertainty in Phylogenetic Estimation of Evolutionary Divergence Times. *Syst. Biol.*, 58, 367-380.
- Hurley I.A., Mueller R.L., Dunn K.A., Schmidt E.J., Friedman M., Ho R.K., Prince V.E., Yang Z.H., Thomas M.G. & Coates M.I. (2007). A new time-scale for ray-finned fish evolution. *Proc. R. Soc. B-Biol. Sci.*, 274, 489-498.
- Jobb G. (2008). TREEFINDER. *Munich, Germany: Distributed by the author.*
- Jolliffe I.T. (2002). *Principal Component Analysis*. 2nd edn. Springer-Verlag, New York.
- Lessios H.A. (2008). The Great American Schism: Divergence of Marine Organisms After the Rise of the Central American Isthmus. *Annu. Rev. Ecol. Evol. Syst.*, 39, 63-91.
- Li C., Orti G., Zhang G. & Lu G. (2007). A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evol. Biol.*, 7, 44.
- Lundberg J.G., Sullivan J.P., Rodiles-Hernández R. & Hendrickson D.A. (2007). Discovery of African roots for the Mesoamerican Chiapas catfish, *Lacantunia enigmatica*, requires an ancient intercontinental passage. *Proc. Acad. Nat. Sci. Phila.*, 156, 39-53.
- Malabarba L.R. & Malabarba M.C. (eds.) (2010). *Biogeography of Characiformes: an evaluation of the available information of fossil and extant taxa*. Verlag Dr. Friedrich Pfeil, München, Germany.
- Marceniuk A.P. (2007). Revalidação de *Cathorops arenatus* (Valenciennes, 1840) e *Cathorops agassizii* (Eigenmann & Eigenmann, 1888) (Siluriformes, Ariidae),

- bagres marinhos das regiões norte e nordeste da América do Sul. *Iheringia Ser. Zool.*, 97, 360-375.
- Marss T. (2001). *Andreolepis* (Actinopterygii) in the upper Silurian of northern Eurasia. *Proc. Eston. Acad. Sc. Geol.*, 50.
- Miller R. (1966). Geographical distribution of Central American freshwater fishes. *Copeia*, 773-802.
- Ng H.H. & Sparks J.S. (2003). The ariid catfishes (Teleostei: Siluriformes: Ariidae) of Madagascar, with the description of two new species. *Occas. Pap. Mus. Zool. Univ. Mich.*, 735, 1-21.
- Peng Z.G., He S.P., Wang J., Wang W. & Diogo R. (2006). Mitochondrial molecular clocks and the origin of the major Otocephalan clades (Pisces : Teleostei): A new insight. *Gene*, 370, 113-124.
- Pérez L. (2004). La ictiofauna del lago de Izabal y sus afluentes: composición, distribución y ecología. In: *Biología*. Universidad del Valle de Guatemala, p. 259.
- Potayo-Ariza F.J. (1996). The phylogenetic relationships of *Rubiesichthys gregalis* and *Gordichthys conquensis* (Ostariophysi, Chanidae), from the Early Cretaceous of Spain. In: *Systematics and Paleoecology* (eds. Arratia G & Viohl G). Verlag Dr. Friedrich Pfeil München, pp. 329-348.
- Pyron R.A. (2010). A Likelihood Method for Assessing Molecular Divergence Time Estimates and the Placement of Fossil Calibrations. *Syst. Biol.*, 59, 185-194.
- R Development Core Team (2011). R: a Language and environment for statistical computing. URL <http://www.r-project.org>.
- Revell L.J. (2009). Size-correction and principal components for interspecific comparative studies. *Evolution; international journal of organic evolution*, 63, 3258-68.
- Rodriguez-Olarte D., Mojica Corzo J.I. & Taphorn D.C. (2011). Northern South America - Magdalena and Maracaibo Basins. In: *Historical Biogeography of Neotropical Freshwater Fishes* (eds. Albert JS & Reis RE). The Regents of the University of California, pp. 243-257.
- Sanderson M.J. (2003). r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*, 19, 301-302.
- Santini F., Harmon L., Carnevale G. & Alfaro M. (2009). Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. *BMC Evol. Biol.*, 9, 194.
- Santos R.S. (1973). *Steindachneridion iheringi* (Woodward) um siluriforme da Bacia do Paraíba, Estado de São Paulo. *Anais da Academia Brasileira de Ciências*, 45, 667.
- Sparks J.S. & Stiassny M.L.J. (2003). Introduction to the freshwater fishes. In: *The Natural History of Madagascar* (eds. Goodman SM & Benstead JP). University of Chicago Press Chicago, pp. 849-882.
- Stamatakis A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22, :2688-2690.

- Stamatakis A., Hoover P. & Rougemont J. (2008). A rapid bootstrap algorithm for the RAxML web servers. *Syst. Biol.*, 57, 758 - 771.
- Strauss R.E., Atanassov M.N. & Oliveira J.A. (2003). Evaluation of the principal-component and expectation-maximization methods for estimating missing data in morphometric studies. *J. Vertebr. Paleontol.*, 23, 284–296.
- Sullivan J.P., Lundberg J.G. & Hardman M. (2006). A phylogenetic analysis of the major groups of catfishes (Teleostei : Siluriformes) using *rag1* and *rag2* nuclear gene sequences. *Mol. Phylogenet. Evol.*, 41, 636-662.
- Swales S., Storey A.W. & Bakowa K.A. (2000). Temporal and spatial variations in fish catches in the fly river system in Papua New Guinea and the possible effects of the Ok Tedi copper mine. *Environ. Biol. Fish*, 57, 75-95.
- van Ufford A.Q. & Cloos M. (2005). Cenozoic tectonics of New Guinea. *AAPG Bull.*, 89, 119-140.