






Molecular systematics of the *Awaous banana* complex (River gobies; Teleostei: Oxudercidae)

Caleb D. McMahan¹  | Diego J. Elías²  | Yue Li² |
 Omár Domínguez-Domínguez³ | Sheila Rodríguez-Machado² |
 Alejandra Morales-Cabrera⁴ | Diana Velásquez-Ramírez⁴ | Kyle R. Piller⁵  |
 Prosanta Chakrabarty²  | Wilfredo A. Matamoros⁶ 

¹Field Museum of Natural History, Chicago, Illinois

²LSU Museum of Natural Science, Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana

³Laboratorio de Biología Acuática, Facultad de Biología, Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Mexico

⁴Escuela de Biología, Facultad de Ciencias Químicas y Farmacia, Universidad de San Carlos de Guatemala, Guatemala City, Guatemala

⁵Department of Biological Sciences, Southeastern Louisiana University, Hammond, Louisiana

⁶Colección de Ictiología, Instituto de Ciencias Biológicas, Universidad de Ciencias y Artes de Chiapas, Tuxtla Gutiérrez, Mexico

Correspondence

Caleb D. McMahan, Field Museum of Natural History, 1400 S. Lake Shore Dr., Chicago, IL 60605, USA.

Email: cmcmahan@fieldmuseum.org

Abstract

Diadromous fishes can exhibit interesting evolutionary and population-level patterns given their use of freshwater and marine environments as part of their life histories. The River goby genus *Awaous* are prominent members of riverine ichthyofaunas and occur throughout Atlantic and Pacific slopes of the Americas from the southern United States to Ecuador and Brazil. Here we study the widespread and polymorphic *Awaous banana* complex to assess phylogeographic patterns and test previous hypotheses that all populations of this species in the Americas belong to the same species. Analysis of sequence data based on the mitochondrial cytochrome oxidase I gene shows multiple clades within the Atlantic and Pacific basins, which correspond to previously described species. Additionally, haplotype analysis demonstrates unique and unconnected networks between these species. Within these clades we document biogeographic patterns that are congruent with results of other co-occurring diadromous species, as well as a novel biogeographic pattern for the region. Our results support the recognition of distinct species of *Awaous* in the Atlantic (*A. banana* and *A. tajasica*) and Pacific (*A. transandeanus*) basins. These results are concordant with previously established morphological characters permitting the separation of these species.

KEYWORDS

biogeography, goby, Middle America, Neotropics, phylogeography, rivers

1 | INTRODUCTION

Geographically widespread species are frequently the basis of studies to test hypotheses related to systematics, phylogeography and taxonomy. This has certainly been the case for freshwater fishes in Middle America, the region comprising Mexico, Central America and the Greater Antilles, and has included studies of catfishes (Perdices *et al.*, 2002), livebearers (Marchio & Piller, 2013) and cichlids (McMahan *et al.*, 2017). Compared to stenohaline or euryhaline fishes, diadromous fishes can exhibit interesting patterns of evolutionary

history and population structure given their occupancy of both freshwater and marine systems as integral parts of their life histories. These structured biogeographic patterns have been observed in several diadromous taxa in Middle America, including sleepers (family Eleotridae, Guimarães-Costa *et al.*, 2017 and Galván-Quesada *et al.*, 2016) and mullets (family Mugilidae, Díaz-Murillo *et al.*, 2017 and McMahan *et al.*, 2013). Intraspecific divergence was recovered across all of these taxa, with different lineages present in Atlantic and Pacific-slope rivers, as well as multiple lineages within each slope.

TABLE 1 Accession and locality data for tissue samples from specimens sequenced for this study

	Tissue catalogue no.	Accession no.	Country	Locality	Latitude	Longitude
1	STRI 14696	MG496094	Nicaragua	Río Escalante	11.52800	-86.16470
2	UMSNH 10674	MZ130163	Mexico	Estero Barra de Pichi	18.97486	-102.32623
3	UMSNH 11104	MZ130164	Mexico	Lázaro Cárdenas	18.03361	-102.55917
4	UMSNH 11105	MZ130165	Mexico	Lázaro Cárdenas	18.03361	-102.55917
5	UMSNH 11365	MZ130166	Mexico	Huahua	18.17769	-103.00731
6	UMSNH 11474	MZ130167	Mexico	Cachan	18.23486	-103.24728
7	UMSNH 12523	MZ130168	Mexico	Mexcalhuacán	18.05603	-102.65836
8	SLU-TC 244	MZ130169	Mexico	Río Mascota	20.72667	-105.16111
9	SLU-TC 1051	MZ130170	Mexico	Río Ayuquila	19.68188	-104.08447
10	LSUMZ-F 2268	MZ130171	El Salvador	Río Cangrejera	13.47428	-89.18172
11	LSUMZ-F 2265	MZ130172	El Salvador	Río Cangrejera	13.47428	-89.18172
12	UMSNH 22680	MZ130173	El Salvador	Mizata	13.51105	-89.59551
13	LSUMZ-F 2269	MZ130174	El Salvador	Río Cangrejera	13.47428	-89.18172
14	UMSNH 11375	MZ130175	Mexico	Huahua	18.17769	-103.00731
15	SLU-TC 1004	MZ130176	Mexico	Río Sinaloa	25.95550	-109.05369
16	LSUMZ-F 2960	MZ130177	Nicaragua	Río Soledad	12.13144	-86.64275
17	UMSNH 22722	MZ130178	El Salvador	Mizata	13.51105	-89.59551
18	UMSNH 11364	MZ130179	Mexico	Huahua	18.17769	-103.00731
19	UMSNH 20900	MZ130180	Mexico	Estero Chucutitan	18.01226	-102.45879
20	SLU-TC 250	MZ130181	Mexico	Río Tehuantepec	16.32958	-95.23583
21	SLU-TC 1699	MZ130182	Mexico	Río Colotepec	15.83771	-97.02761
22	LSUMZ-F 2959	MZ130183	Nicaragua	Río Soledad	12.13144	-86.64275
23	UMSNH 11597	MZ130184	Mexico	Río Coahuayana	18.68478	-103.73717
24	SLU-TC 857	MZ130185	Mexico	Río Ayuquila	19.68188	-104.08447
25	LSUMZ-F 2395	MZ130186	El Salvador	Río Banderas	13.58772	-89.73486
26	LSUMZ-F 2398	MZ130187	El Salvador	Río Banderas	13.58772	-89.73486
27	UMSNH 20901	MZ130188	Mexico	Estero Chucutitan	18.01226	-102.45879
28	SLU-TC 2314	MZ130189	Mexico	Río Ayuquila	19.66594	-104.08539
29	UMSNH 8547	MZ130190	Mexico	Estero Teolán	18.07231	-102.73108
30	STRI 11168	MG936712	Panama	Río Santa Maria	8.41320	-81.04850
31	STRI 11209	MG936725	Panama	Río Bayano	9.22360	-79.09220
32	LSUMZ-F 1404	MZ130159	Costa Rica	Río Claro	8.68025	-83.00700
33	LSUMZ-F 1403	MZ130160	Costa Rica	Río Claro	8.68025	-83.00700
34	LSUMZ-F 1822	MZ130161	Panama	Río Bayano	8.61931	-78.16981
35	STRI 11008	MG936713	Panama	Río Cocle del Sur	8.62100	-80.44900
36	STRI 4407	MG936717	Panama	Río Tuira	8.12330	-77.54390
37	STRI 2091	MG496095	Costa Rica	Río Pirris	9.51990	-84.32500
38	STRI 92	MG936715	Panama	Río Chiriqui	8.20380	-81.58610
39	LSUMZ-F 1821	MZ130162	Panama	Río Bayano	8.61931	-78.16981
40	STRI 164	MG936716	Panama	Río Chiriqui	8.27030	-81.86320
41	STRI 18730	MG936719	Panama	Río Tuira	8.66090	-77.79330
42	STRI 6973	MG936714	Panama	Río Tuira	7.74250	-77.88400
43	LPB 37064	FBCRB235-09	Brazil	Río Escuro	-23.44200	45.19100
44	LPB 37091	FBCRB240-09	Brazil	coastal São Paulo	-23.54700	45.08700
45	LBP 38263	FBCRB-254-09	Brazil	Río Escuro	-23.44200	45.19100
46	MCNIP 1478	MUCU 148-14	Brazil	Minas Gerais, Nanuque	-18.09970	-40.44780

TABLE 1 (Continued)

	Tissue catalogue no.	Accession no.	Country	Locality	Latitude	Longitude
47	LBP 38264	FBCRB-255-09	Brazil	Rio Escuro	-23.44200	-45.19100
48	LBP 37088	FBCRB237-09	Brazil	coastal São Paulo	-23.54700	-45.08700
49	LBP 38262	FBCRB-253-09	Brazil	Rio Escuro	-23.44200	-45.19100
50	LBP 37090	FBCRB239-09	Brazil	coastal São Paulo	-23.54700	-45.08700
51	LBP 37065	FBCRB236-09	Brazil	Rio Escuro	-23.44200	-45.19100
52	LBP 38261	FBCRB252-09	Brazil	Rio Escuro	-23.44200	-45.19100
53	LBP 37089	FBCRB238-09	Brazil	coastal São Paulo	-23.54700	-45.08700
54	MCNIP 1478	MUCU 135-14	Brazil	Minas Gerais, Nanuque	-18.09970	-40.44780
55	UMSNH 5530	MZ130154	Mexico	Río Pantepec	20.73200	-98.02269
56	UMSNH 3468	MZ130155	Mexico	Avila Camacho	20.55128	-97.87250
57	UMSNH 5594	MZ130156	Mexico	Puente Cazones	20.63536	-97.39925
58	SLU-TC 4903	MZ130157	Mexico	Río Tecolutla	20.43722	-97.16554
59	SLU-TC 4902	MZ130158	Mexico	Río Tecolutla	20.43722	-97.16554
60	ANC 12.2.454	MZ130131	Cuba	Río Cabagán	21.83878	-80.10561
61	STRI 13919	MG496096	Nicaragua	Río Coco	13.51290	-85.80990
62	STRI 1732	MG936718	Panama	Río Acla	8.84590	-77.68820
63	STRI 2970	MG936721	Panama	Río Cascajal	9.54640	-79.60620
64	LSUMZ-F 2146	MZ130134	Panama	Río Changuinola	9.04067	-82.29089
65	LSUMZ-F 2175	MZ130135	Panama	Río Changuinola	9.25069	-82.41044
66	LSUMZ-F 3258	MZ130136	Honduras	Río Motagua	15.66678	-88.20626
67	SLU-TC 1876	MZ130137	Jamaica	Swift River	18.19447	-76.57952
68	UMSNH 16512	MZ130138	Venezuela	Puente el Encanto	10.48873	-66.11419
69	ANC 12.2.454	MZ130139	Cuba	Río Cabagán	21.83878	-80.10561
70	ANC 12.2.454	MZ130140	Cuba	Río Cabagán	21.83878	-80.10561
71	STRI 13918	MG496093	Nicaragua	Río Coco	13.51290	-85.80990
72	STRI 3717	MG736722	Panama	Río Chagres	9.41170	-78.64580
73	LSUMZ-F 2145	MZ130141	Panama	Río Changuinola	9.04067	-82.29089
74	LSUMZ-F 4000	MZ130142	Honduras	Río Patuca	14.28901	-85.12000
75	LSUMZ-F 2115	MZ130143	Panama	Río Garamo	8.90861	-82.18800
76	USNM 447326	MT455852	United States	NC: Bouge Sound	34.72310	-76.75030
77	STRI 4986	NC036224	Panama	Bocas del Toro	9.36000	-82.59000
78	LSUMZ-F 1959	MZ130152	Panama	Río Chagres	9.57658	-79.48342
79	ANC 12.2.453	MZ130153	Cuba	Río San Sebastian	22.29839	-83.80737
80	ANC 12.2.454	MZ130144	Cuba	Río Cabagán	21.83878	-80.10561
81	LSUMZ-F 3256	MZ130145	Honduras	Río Motagua	15.66678	-88.20626
82	LSUMZ-F 1722	MZ130146	Costa Rica	Río Sixaola	9.65453	-82.76372
83	SLU-TC 2132	MZ130147	Jamaica	Milk River	17.88060	-77.34528
84	SLU-TC 3261	MZ130148	Belize	Temash River	16.04068	-89.02927
85	LSUMZ-F 3257	MZ130149	Honduras	Río Motagua	15.66678	-88.20626
86	LSUMZ-F 4143	MZ130150	Honduras	Río Patuca	14.25099	-86.16664
87	LSUMZ-F 6340	MZ130151	Honduras	Río Patuca	14.34151	-85.49107
88	STRI 1354	MG936723	Panama	Río Cocolé del Norte	8.81870	-80.55300
89	USNM 447327	MT455549	United States	NC: Bouge Sound	34.72310	-76.75030
90	USNM 447328	MT455993	United States	NC: Bouge Sound	34.72310	-76.75030
91	STRI 6858	MG936720	Panama	Bocas del Toro	9.04060	-82.28580
92	USNM 447325	MT455114	United States	NC: Bouge Sound	34.72310	-76.75030

(Continues)

TABLE 1 (Continued)

	Tissue catalogue no.	Accession no.	Country	Locality	Latitude	Longitude
93	ANC 12.2.454	MZ130132	Cuba	Río Cabagán	21.83878	−80.10561
94	LSUMZ-F 1958	MZ130133	Panama	Río Chagres	9.57658	−79.48342
Outgroups:						
95	<i>A. grammepomus</i>	MH721183	Vietnam	–	–	–
96	<i>A. ocellaris</i>	KC959856	Philippines	–	–	–
97	<i>A. ocellaris</i>	JQ431473	French Polynesia	–	–	–
98	<i>A. grammepomus</i>	KU692309	Indonesia	–	–	–

Note. Acronyms follow Sabaj (2020). Accession numbers refer to GenBank or BOLD databases; BOLD accession numbers are italicized.

Gobies (order Gobiiformes) are an incredibly species-rich group of fishes, with an extraordinary range of sizes, behaviours and morphologies; they occupy diverse micro-habitats in primarily marine environments but also some estuarine and freshwater habitats (Tornabene *et al.*, 2013). Diadromous riverine gobies of the genus *Awaous* are widespread in tropical and subtropical systems and currently comprise some 20 species (Fricke *et al.*, 2021). These gobies are a prominent component of the ichthyofauna in rivers throughout the Atlantic and Pacific slopes of the Americas, from the southern United States (in the Atlantic) to Brazil and Ecuador (Watson, 1992).

Species-level identifications of members of this genus have been difficult. Several previously recognized species of *Awaous* from the Americas are now considered synonyms of highly variable taxa such as *A. banana* (Valenciennes, 1837), *A. transandeanus* (Günther, 1861) and *A. tajasica* (Lichtenstein, 1822). These three species have most frequently been used and treated as valid by authors. *Awaous flavus* (Valenciennes, 1837) is the additional member of the genus occurring in the Western Atlantic in South America. However, this species is not a member of the *A. banana* complex and not included in the present study. *Awaous flavus* is the sole member of the monotypic subgenus *Euctenogobius* and readily distinguished from congeners (Lasso-Alcalá & Lasso, 2008). The most recent systematic review of the *A. banana* complex was conducted by Watson (1996), who revised the subgenus *Chonophorus* and concluded there was insufficient evidence to support the existence of separate species of *Awaous* on Atlantic and Pacific slopes of the Americas, recognizing *A. transandeanus* as a synonym of *A. banana* and restricting *A. tajasica* to Brazil south of the mouth of the Amazon River. Since then, some authors have followed these designations; however, other authors have maintained separate Pacific and Atlantic species (Bussing, 1998; Miller, 2005) given that populations in the two basins could be morphologically distinguished. Furthermore, recent work has provided evidence of divergence between the mitochondrial genomes of *Awaous* specimens from the Pacific and Atlantic slopes of Panama based on comparison of a single individual from each coast (Alda *et al.*, 2018). Therefore, the objective of this study was to assess phylogeographic structure within the *A. banana* complex and test Watson's (1996) hypothesis that all Middle American populations belong to the same species.

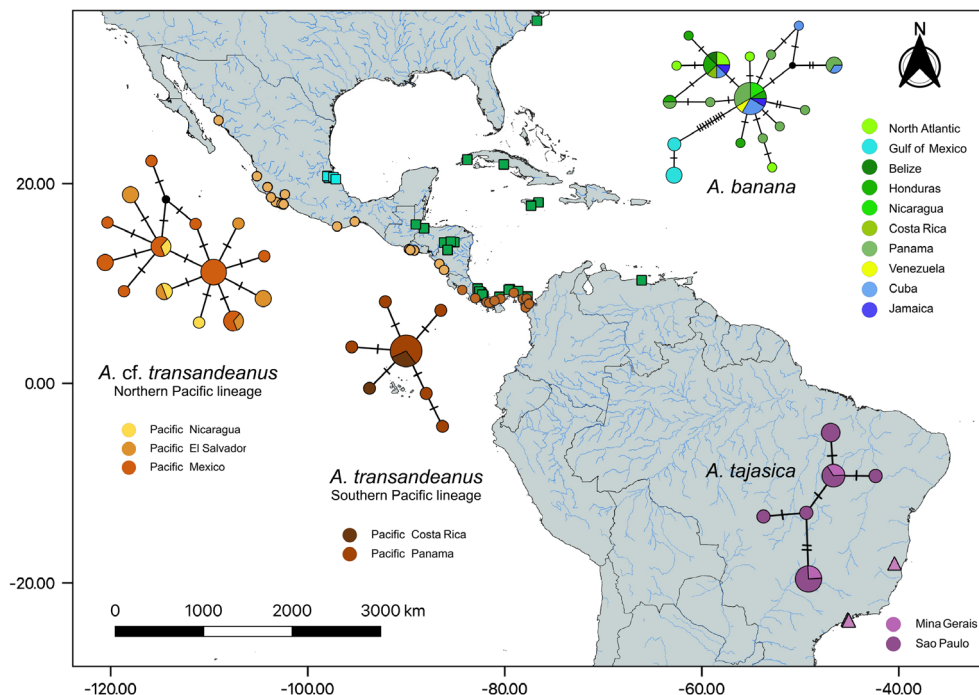
2 | MATERIALS AND METHODS

2.1 | Molecular data

Specimens and tissue samples of the *Awaous banana* complex were collected throughout its distribution (Table 1 and Figure 1). The collection of specimens complied with all international and local permitting regulations across all countries and museums (Panama, Autoridad Nacional del Ambiente permit SC/A-17-11; Costa Rica, Museo de Zoología, Universidad de Costa Rica; Nicaragua, Ministerio del Ambiente y Recursos Naturales permit DGPN/DB/DAP-IC-0008-2010; Honduras, Instituto Nacional de Conservación y Desarrollo Forestal, Áreas Protegidas, y Vida Silvestre permit DVS-ICF-03302009). Tissue samples were taken from fin clips and/or muscle from the right side of specimens. Specimens were then preserved in 10% formalin, transferred to 70% ethanol and deposited in the LSU Museum of Natural Science (LSUMZ), Southeastern Louisiana University Vertebrate Museum (SLU), Universidad Michoacana de San Nicolas Hidalgo (UMSNH) and Acuario Nacional de Cuba (ANC). A list of tissue samples and locality information is provided in Table 1.

Whole genomic DNA was extracted using DNeasy Tissue Kits (Qiagen, Inc., Valencia, CA, USA) following manufacturer protocols. The mitochondrial marker cytochrome oxidase I (COI) was amplified using the primers BOL-F and BOL-R, and polymerase chain reaction (PCR) followed protocols from Ward *et al.* (2005). The COI “barcode” (Ward *et al.*, 2005, 2009) has been widely used to address questions at population and species levels in freshwater and marine fishes (e.g., Pereira *et al.*, 2013; Rees *et al.*, 2020; Weigt *et al.*, 2012). The PCR products were visualized on 1.0% agarose gels and sequenced at the Beckman Coulter Genomics Facility (Danvers, MA, USA) and the Pritzker Lab at the Field Museum of Natural History. Sequence data were generated for a total of 60 individuals. Chromatograms were visually examined and low-quality base pairs calls were removed in the software Geneious version 10.0.9 (Kearse *et al.*, 2012). An additional 34 ingroup sequences were added to the dataset based on published records in GenBank and BOLD (Table 1) to include additional populations in our analyses. *Awaous ocellaris* (Broussonet, 1782) and *A. grammepomus* (Bleeker, 1849) were included as outgroup taxa. We used the Muscle algorithm (Edgar, 2004) with default parameters

FIGURE 1 Map showing sample localities for specimens and haplotype networks based on analysis of sequences from the mitochondrial COI gene for *Awaous* in the Eastern Pacific and Western Atlantic. Colours for sample localities correspond to recovered clades. Colours of individual haplotypes correspond to geographic localities



implemented in Geneious to generate a multiple sequence alignment that consisted of 98 sequences.

2.2 | Phylogenetic analyses

We evaluated the best model of nucleotide evolution for the COI dataset using PartitionFinder 2 (Lanfear *et al.*, 2012). The best model was selected using the corrected Akaike Information Criterion (AICc; Hurvich and Tsai, 1989). We inferred a phylogenetic hypothesis under a Bayesian framework in the software Mr. Bayes 3.2.6 (Huelsenbeck *et al.*, 2001) in the CIPRES Science gateway portal (Miller *et al.*, 2010). Bayesian analyses were run for 12,000,000 generations sampling every 6000 generations. Sampling stationary was visually inspected in Tracer 1.7 (Rambaut *et al.*, 2018) and we checked for effective sample size (ESS > 200) and observed average standard deviation of split frequencies (<0.01). We discarded the first 25% of sampled trees as burn-in, and bayesian posterior probabilities (BPP) were calculated using the post burn-in trees to assess support for clades. Three independent analyses were performed to assess topological congruence.

2.3 | Haplogroups and molecular diversity

We tested for the presence of distinct haplogroups across the distribution of the *A. banana* complex. Haplogroups were inferred using statistical parsimony analysis (Templeton *et al.*, 1992) in the software TCS v.1.21 (Clement *et al.*, 2000) with a 95% confidence limit for connection. If unconnected haplogroups were identified by the TCS analysis, the haplotype networks for each haplogroup were independently inferred. Haplotype networks were constructed using the median-

joining network (Bandelt *et al.*, 1999) implemented in the software POPART (Leigh & Bryant, 2015). Uncorrected sequence divergence (*p*-distances) between haplogroups were calculated in the software MEGA v.10.1.8 (Kumar *et al.*, 2018), and summary statistics of genetic diversity (*e.g.*, nucleotide and haplotypic diversity) were calculated for each independent haplogroup in the software DnaSP v.6.1 (Rozas *et al.*, 2017).

3 | RESULTS

3.1 | Phylogenetic analyses

Our final alignment consisted of 94 ingroup samples and four outgroups with an alignment length of 621 base pairs. The Bayesian inference recovered phylogenetic structure across the distribution of the *Awaous banana* complex (Figure 2). Our phylogenetic hypothesis recovered four well-supported clades across populations of *Awaous* in the Eastern Pacific and Western Atlantic basins, as well as a single clade formed by all samples of *A. tajasica* in the Atlantic coast of Brazil (Figure 2).

Two clades of *Awaous* were recovered from rivers in the Eastern Pacific basin (Figure 2). The two Pacific clades were 7.13% divergent based on COI sequences (Table 2) and were not recovered as each other's closest relatives based on phylogenetic analysis of COI sequences. The Northern Pacific clade extended from northern Mexico to just north of the Nicoya Peninsula in Costa Rica and was recovered as the sister group to the other *Awaous* clades (Figures 1 and 2). Populations of *Awaous* south of the Nicoya Peninsula formed the Southern Pacific clade. This clade was recovered as the sister group to a clade inclusive of all samples and species of *Awaous* distributed in the western Atlantic basin (Figure 2). The Southern Pacific clade was

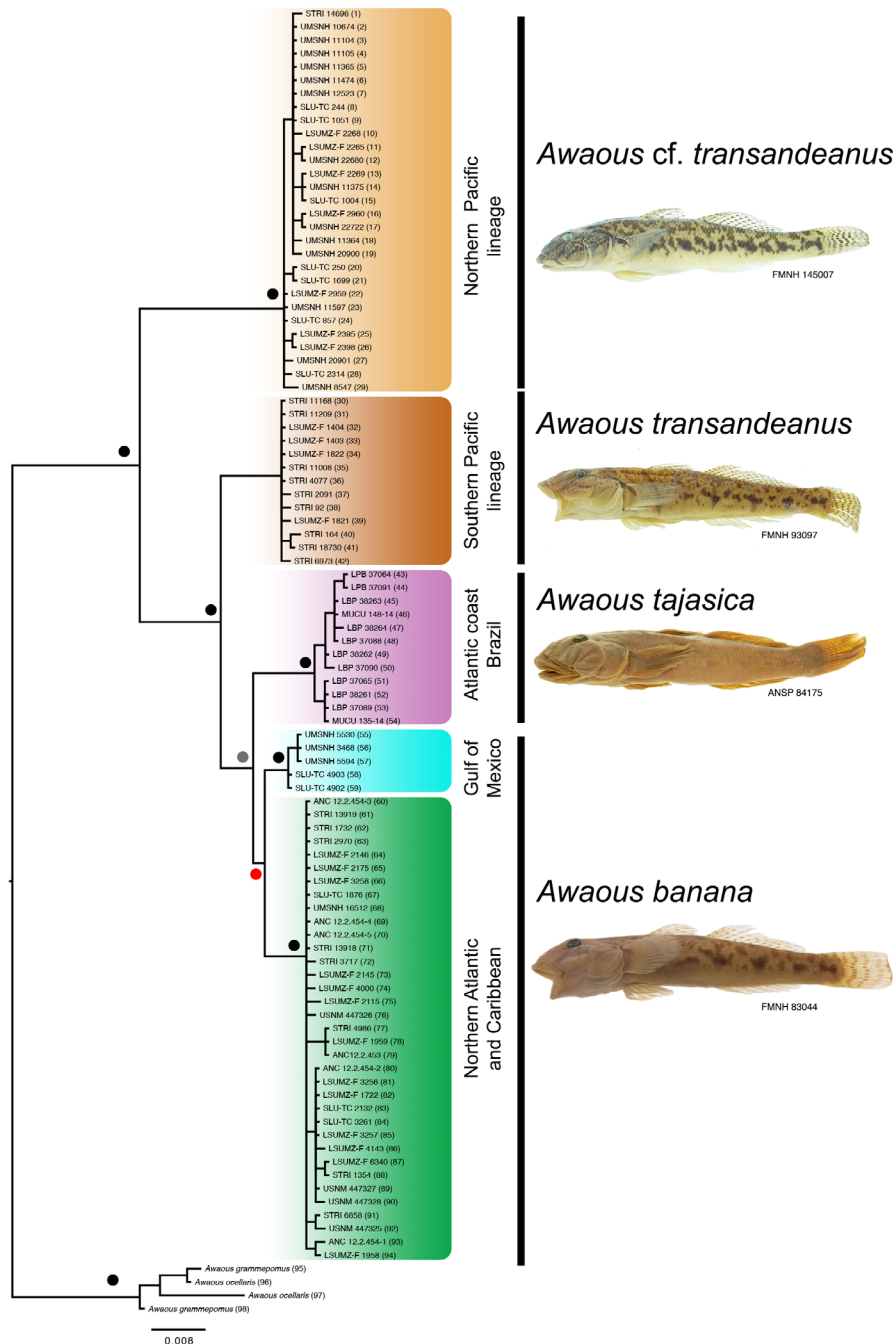


FIGURE 2 Bayesian phylogeny for *Awaous* from the Eastern Pacific and Western Atlantic based on mitochondrial COI sequences. Tip labels indicate tissue catalogue numbers; numbers in parentheses correspond to Table 1. Colours for clades correspond to sample localities in Figure 1. Bayesian posterior probabilities: (●) $\text{bpp} \geq 0.95$; (●) $0.95 < \text{bpp} < 0.90$; (●) $\text{bpp} < 0.90$

TABLE 2 Uncorrected p-distances as percentages (below diagonal) and standard deviation (above diagonal) between *Awaous tajasica* and the four recovered clades in the *A. banana* complex

	<i>A. tajasica</i>	Caribbean	Gulf of Mexico	Southern Pacific	Northern Pacific
<i>A. tajasica</i>	–	0.65	0.66	0.80	1.02
Caribbean	3.10	–	0.54	0.70	1.00
Gulf of Mexico	2.68	1.86	–	–	–
Southern Pacific	4.12	3.62	3.08	–	1.07
Northern Pacific	7.14	6.79	6.58	7.13	–

between 3.08% and 3.62% divergent from the Caribbean and Gulf of Mexico clades, respectively. The genetic divergence between the Southern Pacific clade and *A. tajasica* was 4.12% (Table 4).

All samples of the Atlantic members of *Awaous* were recovered as monophyletic (Figure 2), with a sister relationship (with low BPP support) between Caribbean and Gulf of Mexico lineages of *A. banana*

(1.86% sequence divergence; Table 2) and *A. tajasica* recovered as the sister group to those two clades (2.68%–3.10% sequence divergence; Table 2). Intra-clade genetic divergence was low within all recovered clades (Table 3).

3.2 | Haplogroups and molecular diversity

To avoid missing data in the estimation of haplogroups and molecular diversity, two sequences, *Awaous tajasica* (MUCU135-14) and *A. banana* (MG936713), were excluded from the dataset given short sequence reads. Furthermore, the alignment was truncated to a final length 615 base pairs. The TCS analysis recovered four unconnected haplogroups that were largely congruent with the clades recovered in our phylogenetic analysis. Two independent haplogroups were recovered in the Eastern Pacific basin that were congruent with the northern and southern Pacific lineages (Figure 1). Fourteen haplotypes were identified within the Northern Pacific haplogroup and this haplogroup possessed the highest haplotypic diversity ($hd = 0.906$; Table 4) in the region. The Southern Pacific haplogroup possessed seven haplotypes (Figure 1) and relatively lower haplotypic diversity ($hd = 0.773$; Table 4) in the Eastern Pacific basin.

In the Atlantic basin, two haplogroups were recovered by the TCS analysis. One haplogroup was congruent with the *A. tajasica* clade (Figures 1 and 2) and comprised all samples from the Atlantic Coast of Brazil. Six haplotypes were identified within *A. tajasica* (Figure 1) with relatively high haplotypic diversity ($hd = 0.873$; Table 4). The other haplogroup identified in the Atlantic basin comprised all samples from

the Caribbean lineage and the Gulf of Mexico lineage. Seventeen haplotypes were identified within the Caribbean and Gulf of Mexico clades (Figure 1 and Table 4). The Caribbean clade possessed a relatively high haplotypic diversity ($hd = 0.847$; Table 4) in contrast with the haplotypic diversity observed in the Gulf of Mexico clade ($hd = 0.600$; Table 4).

4 | DISCUSSION

4.1 | Taxonomic status

Watson (1996) considered there to be insufficient data supporting the existence of separate species of *Awaous* in Pacific and Atlantic rivers in North, Central and South America. However, subsequent authors rejected this hypothesis based on diagnosable differences between the two basins. Based on characters provided in Miller (2005), *A. banana* possesses 69–76 lateral scales and dark blotches along the sides of the body without black, vertical bars. *A. transandeanus* possesses 60–67 lateral scales and dark blotches along the sides of the body but with narrow, black vertical bars present. *A. tajasica* possess dark lateral blotches but fewer lateral scales than *A. banana* (61–66; Watson, 1996). Our phylogenetic and population-level results are concordant with these morphological differences that can be used to distinguish these species (Figure 2). Our results also support the restriction of *A. tajasica* to Brazil, south of the Amazon River (Watson, 1996).

Our results demonstrate that populations of *Awaous* on the Atlantic and Pacific slopes of the Americas are distinct lineages. While two clades were recovered within each ocean basin, additional data are necessary to determine if these lineages warrant species status. While available names for these lineages are tentatively available [*Gobius mexicanus* Günther, 1861 (Gulf of Mexico) and *A. nelsoni* Evermann, 1898 (northern Pacific)], the re-identification of diagnostic characters in type material and comparisons with our recovered clades will be needed to further substantiate the putative validity of these two currently synonymized species. Regardless, it is clear that at present treating *A. banana* and *A. transandeanus* as distinct species is the most robust taxonomic hypothesis and most congruent with phylogenetic and biogeographic evidence.

TABLE 3 Genetic variation and standard deviation (S.D.) within *Awaous tajasica* and the four recovered clades within the *A. banana* complex

	Genetic variation	S.D.
<i>A. tajasica</i>	0.36	0.15
Caribbean	0.29	0.11
Gulf of Mexico	0.00	0.00
Southern Pacific	0.18	0.077
Northern Pacific	0.29	0.095

TABLE 4 Genetic diversity observed for *Awaous tajasica* and the haplogroups recovered within the *A. banana* complex

	<i>n</i>	<i>vs</i>	<i>pis</i>	<i>H</i>	<i>hd</i>	π
<i>A. tajasica</i>	11	6	4	6	0.873 (0.071)	0.0034 (0.0005)
Caribbean	35	15	7	15	0.847 (0.047)	0.0029 (0.0004)
Gulf of Mexico	5	1	1	2	0.600 (0.175)	0.00098 (0.0003)
Southern Pacific	12	6	1	7	0.773 (0.128)	0.0018 (0.0005)
Northern Pacific	29	13	7	14	0.906 (0.039)	0.0028 (0.0003)

Standard deviation in parentheses.

Note. *H*, number of haplotypes; *hd*, haplotype diversity; *n*, number of sequences; *pis*, parsimony-informative sites; *vs*, variable size; π , nucleotide diversity.

The two clades that exist in the Atlantic basin are sister lineages with a connected haplotype network, and at present we consider these two lineages to represent genetically differentiated populations of *A. banana*. More work is needed to establish the existence or extent of gene flow between populations in the Caribbean and Gulf of Mexico. The two clades of *A. transandeanus* in the Pacific are not each other's closest relatives based on our analyses; however, this result could be due to limitations of the mitochondrial marker selected. More variable markers or expanded genomic coverage could lead to recovering the two Pacific populations as monophyletic, as well as potentially increasing genetic differences between Caribbean and Gulf of Mexico populations in the Western Atlantic. Nevertheless, the two clades are quite divergent and possess completely unique haplotypes. As with the Atlantic populations, at present we take a conservative approach and recognize the two Pacific populations as *A. transandeanus* (Southern Pacific clade) and *A. cf. transandeanus* (Northern Pacific clade), pending additional molecular and morphological data to substantiate distinctiveness of these lineages.

4.2 | Systematics and Biogeography

Gilmore (1992) noted that most larvae of *Awaous* hatching in or entering marine water likely re-enter parental streams, although some may be dispersed *via* ocean currents before returning to freshwater. However, the low genetic divergence and variability within each of the clades of *Awaous* offers the potential that most larvae making it to marine waters do not re-enter parental streams but instead drift in currents and enter other rivers. Widespread lineages across the Pacific basin of Middle America have been observed in other diadromous fishes such as the goby *Sicydium salvini* Ogilvie-Grant, 1884 (Chabarria & Pezold, 2013), the mullet *Dajaus monticola* (Bancroft, 1834; McMahan *et al.*, 2013) and the sleeper *Dormitator latifrons* (Richardson, 1844; Galván-Quesada *et al.*, 2016).

Phylogeographic studies of several diadromous fishes in Middle America have also demonstrated population-level divergence between the Caribbean and Gulf of Mexico (*Dajaus monticola*, McMahan *et al.*, 2013; *Dormitator maculatus* (Bloch, 1792), Galván-Quesada *et al.*, 2016; *Gobiomorus dormitor* Lacepède, 1800; Guimarães-Costa *et al.*, 2017). *Awaous banana* similarly exhibits this biogeographic pattern, with 1.86% COI divergence between these two clades (*i.e.*, Caribbean and Gulf of Mexico). The Loop Current appears to be a potential barrier around the southern tip of Florida between these two populations of *A. banana*, as well as the paucity of suitable rivers for habitat on the Yucatan Peninsula, which has also been hypothesized to separate populations of *D. monticola* (McMahan *et al.*, 2013), with samples of both species from North Carolina recovered as part of the Caribbean clade. These factors in conjunction likely promote the observed genetic differentiation.

While other phylogeographic studies of diadromous fishes in Middle America show the existence of two clades within the Pacific basin, the geographic distributions of these two clades differ across species. The split between populations of *A. transandeanus* in the Pacific basin

appear to be based around the Nicoya Peninsula, with localities north and south of the peninsula recovered in separate clades. The distribution of the Northern Pacific clade falls within the Mexican Tropical Pacific and Chiapas-Nicaragua ecoregions (*sensu* Spalding *et al.*, 2007) and the distribution of the southern Pacific clade lies within the Nicoya and Panama Bight ecoregions (*sensu* Spalding *et al.*, 2007). This could partially be explained by the well-documented seasonal cold-water upwellings around this area that coincide with changes in water temperature, ocean current patterns and resource availability such as food (Vargas, 2016). However, as far as we know this biogeographic pattern has not been documented in other widespread marine or diadromous fishes in the Eastern Pacific. Future work aimed at assessing gene flow, admixture and population limits around the Nicoya Peninsula will be important for more thoroughly investigating this pattern.

ACKNOWLEDGEMENTS

We thank Susan Mochel and Kevin Swagel (FMNH) for assistance with this project. Portions of the molecular data were gathered in the Pritzker Molecular Lab at the Field Museum of Natural History, with support from the Pritzker Foundation. Liz Marchio (SLU) kindly provided information on specimens used in the study. We thank Kevin Feldheim, Erica Zahnle and Isabel DiStefano (FMNH) for assistance with molecular portions of this study. Additional portions of the molecular work were supported by a Smithsonian Tropical Research Institute Short-Term Fellowship awarded to S.R.M. We thank Harilaos Lessios (STRI) for advice and support, as well as Eyda Gómez, Ligia Calderón and Axel Calderón for invaluable lab assistance (STRI). We also thank the following colleagues for administrative or field assistance in the collection of specimens used in this study: Arturo Angulo Sibaja, Carlos Garita, Aaron Geheber, Parker House, Carlos Ramiro Mejía, Enrique Barraza, Veronica Esperanza Melara, Rigoberto Gonzalez, Norman Mercado-Silva and John Lyons.

AUTHOR CONTRIBUTIONS


C.D.M., W.A.M. and K.R.P led the study design. C.D.M., P.C., O.D-D. and K.R.P. led fieldwork efforts, and all authors participated in fieldwork and acquiring samples. C.D.M., S.R-M. and Y.L. participated in data collection. D.J.E., A.M-C. and D.V-R were involved in analyses. All authors participated in writing and reviewing the manuscript.

ORCID

Caleb D. McMahan  <https://orcid.org/0000-0003-2113-8554>

Diego J. Elías  <https://orcid.org/0000-0003-4215-0384>

Kyle R. Piller  <https://orcid.org/0000-0003-1289-9351>

Prosanta Chakrabarty  <https://orcid.org/0000-0003-0565-0312>

Wilfredo A. Matamoros  <https://orcid.org/0000-0002-6241-5354>

REFERENCES

- Alda, F., Adams, A. J., Chakrabarty, P., & McMillan, W. O. (2018). Mitogenomic divergence between three pairs of putative geminate fishes from Panama. *Mitochondrial DNA Part B*, 3, 1–5. <https://doi.org/10.1080/23802359.2017.1413288>.

- Bancroft, E. N. (1834). In G. Cuvier (Ed.), *The animal kingdom, arranged in conformity with its organization, by the Baron Cuvier, member of the Institute of France* (2nd ed., p. 680). London, UK: Whittaker and Co..
- Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, *16*, 37–48.
- Bleeker, P. (1849). Bijdrage tot de kennis der Blennioïden van den Soenda-Molukschen Archipel, met beschrijving van 42 nieuwe soorten. *Verhandelingen van het Bataviaasch Genootschap van Kunsten en Wetenschappen* (Vol. 22) (6), 1–40.
- Bloch, M.E. (1792). *Naturgeschichte der ausländischen Fische*. Sechster Theil. J. Morino & Co., Berlin, Germany 6.
- Broussonet, P. M. A. (1782). *Ichthyologia, sistens piscium descriptiones et icones* (p. 49). London, UK: Decas I.
- Bussing, W. A. (1998). *Peces de las aguas Continentales de Costa Rica* (p. 468). Costa Rica. Universidad de Costa Rica.
- Chabbarria, R. E., & Pezold, F. (2013). Phylogeography and historical demography of *Sicydium salvini* in the eastern Pacific. *Ichthyological Research*, *60*, 353–362.
- Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, *9*, 1657–1660.
- Díaz-Murillo, B. P., Ruiz-Campos, G., Piller, K. R., McMahan, C. D., García-de-León, F. J., & Camarena-Rosales, F. (2017). Assessing population-level morphometric variation of the mountain mullet *Agonostomus monticola* (Teleostei: Mugilidae) across its middle American distribution. *Neotropical Ichthyology*, *15*, e170036.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*, 1792–1797. <https://doi.org/10.1093/nar/gkh340>.
- Evermann, B. W. (1898). Notes on fishes collected by E.W. Nelson on the Tres Marias islands and in Sinaloa and Jalisco, Mexico. *Proceedings of the Biological Society of Washington*, *12*, 1–3.
- Fricke, R., Eschmeyer, W.N., & Van der Laan, R. (2021). Eschmeyer's catalog of fishes: Genera, species, references. Electronic version.
- Galván-Quesada, S., Doadrio, I., Alda, F., Perdices, A., Reina, R. G., García Varela, M., ... Domínguez-Domínguez, O. (2016). Molecular phylogeography and biogeography of the amphidromous fish genus *Dormitator* Gill 1861 (Teleostei: Eleotridae). *PLoS One*, *11*, e0153538.
- Gilmore, R. G. (1992). River goby, *Awaous tajasica*. In C. R. Gilbert (Ed.), *Rare and endangered biota of Florida. Volume 2, Fishes* (pp. 112–117). Gainesville, FL: University Press of Florida.
- Guimarães-Costa, A., Vallinoto, M., Giarizzo, T., Angulo, A., Ruiz-Campos, G., Schneider, H., & Sampaio, I. (2017). Exploring the molecular diversity of Eleotridae (Gobiiformes) using mitochondrial DNA. *Journal of Applied Ichthyology*, *2017*, 1–7.
- Günther, A. (1861). *Catalogue of the fishes in the British museum. Catalogue of the acanthopterygian fishes in the collection of the British museum. Gobiidae, Discoboli, Pediculati, Blenniidae, Labyrinthici, Mugilidae, Notacanthi* (Vol. 3), London, UK: British Museum (Natural History).
- Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, *76*, 297–307. <https://doi.org/10.1093/biomet/76.2.297>.
- Huelsenbeck, J. P., Ronquist, F., Nielsen, R., & Bollback, J. P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, *294*, 2310–2314. <https://doi.org/10.1126/science.1065889>.
- Kumar, S., Stecher, G., Li M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*, 1547–1549. <http://dx.doi.org/10.1093/molbev/msy096>.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.
- Lacepède, B.G.E. (1800). *Histoire naturelle des poissons*. vol. 2.
- Lasso-Alcalá, O. M., & Lasso, C. A. (2008). Revisión taxonómica del género *Awaous* Valenciennes 1837 (Pisces: Perciformes, Gobiidae) en Venezuela, con notas sobre su distribución y hábitat. *Memoria de la Fundación La Salle de Ciencias Naturales*, *168*, 117–140.
- Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). Partitionfinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, *29*, 1695–1701. <http://dx.doi.org/10.1093/molbev/mss020>.
- Leigh, J. W., & Bryant, D. (2015). Popart: Full-feature software for haplotype network construction. *Methods in Ecology & Evolution*, *6*, 1110–1116.
- Lichtenstein, M. H. C. (1822). Die Werke von Marcgrave und Piso Über die Naturgeschichte Brasiliens, erläutert aus den wieder aufgefundenen Original-Abbildungen. IV. Fische. *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin, 1820-1821*, 267–288.
- Marchio, E. A., & Piller, K. R. (2013). Cryptic diversity in a widespread live-bearing fish (Poeciliidae: *Belonesox*). *Biological Journal of the Linnean Society*, *109*, 848–860.
- McMahan, C. D., Davis, M. P., Domínguez-Domínguez, O., García-de-León, F. J., Doadrio, I., & Piller, K. R. (2013). From the mountains to the sea: Phylogeography and cryptic diversity within the mountain mullet, *Agonostomus monticola* (Teleostei: Mugilidae). *Journal of Biogeography*, *40*, 894–904.
- McMahan, C. D., Ginger, L., Cage, M., David, K. T., Chakrabarty, P., Johnson, M., & Matamoros, W. A. (2017). Pleistocene to Holocene expansion of the black-belt cichlid in Central America, *Vieja maculicauda* (Teleostei: Cichlidae). *PLoS One*, *12*, e0178439.
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES science gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*, *2010*, 1–8.
- Miller, R. R. (2005). *Freshwater fishes of México*. Chicago, IL: University of Chicago Press.
- Ogilvie-Grant, W. R. (1884). A revision of the fishes of the genera *Sicydium* and *Lentipes*, with descriptions of five new species. *Proceedings of the Zoological Society of London*, *1884*, 153–172.
- Perdices, A., Bermingham, E., Montilla, A., & Doadrio, I. (2002). Evolutionary history of the genus *Rhamdia* (Teleostei: Pimelodidae) in Central America. *Molecular Phylogenetics and Evolution*, *25*, 172–189.
- Pereira, L. H. G., Hanner, R., Foresti, F., & Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genetics*, *14*, 20.
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in bayesian phylogenetics using tracer 1.7. *Systematic Biology*, *67*, 901–904. <https://doi.org/10.1093/sysbio/syy032>.
- Richardson, J. (1844). Ichthyology part 1. In R. B. Hinds (Ed.), *The zoology of the voyage of H.M.S. Sulphur, under the command of Captain Sir Edward Belcher during the years 1836-42* (pp. 35–44). London, UK: Smith, Elder & Co.
- Rees, D. J., Poulsen, J. Y., Sutton, T. T., Costa, P. A. S., & Landaeta, M. F. (2020). Global phylogeography suggest eucosmopolitanism in mesopelagic fishes (*Maurollicus*: Sternoptychidae). *Scientific Reports*, *10*, 20544.
- Rozas, J., Ferrer-Mata, A., Sánchez-Del Barrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP v6: DNA sequence polymorphism analysis of large datasets. *Molecular Biology and Evolution*, *34*, 3299–3302.
- Sabaj, M. (2020). Codes for natural history collections in ichthyology and herpetology. *Copeia*, *108*, 593–669.
- Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., ... Robertson, J. (2007). Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience*, *57*, 573–583.
- Templeton, A. R., Crandall, K. A., & Sing, C. F. (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, *132*, 619–633. <https://doi.org/10.1093/genetics/132.2.619>.

- Tornabene, L., Chen, Y., & Pezold, F. (2013). Gobies are deeply divided: Phylogenetic evidence from nuclear DNA (Teleostei: Gobiodei: Gobiidae). *Systematics and Biodiversity*, 11, 345–361.
- Valenciennes, A. (1837). Chapitre IX: Des Gobies. In G. Cuvier & A. Valenciennes (Eds.), *Histoire naturelle des poissons. Tome douzième. Suite du livre quatorzième. Gobioides. Livre quinzième. Acanthoptérygiens à pectorales pédiculées* (Vol. 12, p. 1–138). Paris: Chez F. G. Levrault.
- Vargas, J. A. (2016). Chapter 6 The Gulf of Nicoya estuarine ecosystem. In M. Kappelle (Ed.), *Costa Rican ecosystems* (pp. 139–161). Chicago, IL: University of Chicago Press.
- Watson, R. E. (1992). A review of the gobiid fish genus *Awaous* from insular streams of the Pacific plate. *Ichthyological Exploration of Freshwaters*, 3, 161–176.
- Watson, R. E. (1996). Revision of the subgenus *Awaous* (*Chonophorus*) (Teleostei: Gobiidae). *Ichthyological Exploration of Freshwaters*, 7, 1–18.
- Weigt, L. A., Baldwin, C. C., Driskell, A., Smith, D. G., Ormos, A., & Reyier, E. A. (2012). Using DNA barcoding to assess Caribbean reef fish biodiversity: expanding taxonomic and geographic coverage. *PLoS One*, 7, e41059.
- Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 1847–1857. <https://doi.org/10.1098/rstb.2005.1716>.
- Ward, R. D., Hanner, R., & Hebert, P. D. N. (2009). The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology*, 74(2), 329–356. <https://doi.org/10.1111/j.1095-8649.2008.02080.x>.

How to cite this article: McMahan, C. D., Elías, D. J., Li, Y., Domínguez-Domínguez, O., Rodríguez-Machado, S., Morales-Cabrera, A., Velásquez-Ramírez, D., Piller, K. R., Chakrabarty, P., & Matamoros, W. A. (2021). Molecular systematics of the *Awaous banana* complex (River gobies; Teleostei: Oxudercidae). *Journal of Fish Biology*, 1–10. <https://doi.org/10.1111/jfb.14783>