Comparative phylogeography of two codistributed species of the genus *Herichthys* (Actinopterygii: Cichliformes: Cichlidae) in northeastern Mexico

Fabian PÉREZ-MIRANDA[1,2], Omar MEJÍA[1], Benjamín LÓPEZ[3], Eduardo SOTO-GALERA[1], Amairany BERNAL-PORTILLO[1], Wilfredo A. MATAMOROS[2,4]

1   *Departamento de Zoología, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Ciudad de México, Mexico*

2   *Instituto de Ciencias Biológicas, Universidad de Ciencias y Artes de Chiapas, Tuxtla Gutiérrez, Chiapas, Mexico*

3   *Universidad Michoacana de San Nicolás Hidalgo, Michoacán de Ocampo, Mexico*

4   *Field Museum of Natural History, Chicago, IL, USA*

https://zoobank.org/1302CFAD-D245-496B-AFB1-B6E48E90A786

Corresponding author: Fabian Pérez-Miranda (monodactilo@hotmail.com)

Citation: Pérez-Miranda F, Mejía O, López B, Soto-Galera E, Bernal-Portillo A, Matamoros WA (2023) Comparative phylogeography of two codistributed species of the genus *Herichthys* (Actinopterygii: Cichliformes: Cichlidae) in northeastern Mexico. Acta Ichthyologica et Piscatoria 53: 227–242. https://doi.org/10.3897/aiep.53.112183

Abstract

Phylogeographic patterns of freshwater fishes in coastal regions are highly susceptible to eustatic sea level changes associated with Pleistocene glaciations. In this context, the Plain Coastal Gulf in northeastern Mexico represents an ideal study area due to its low elevation. Herein, we compare the phylogeographic structures of two cichlid species of the genus *Herichthys* Baird et Girard, 1854 widely distributed in the Pánuco–Tamesí system in northeastern Mexico using two mitochondrial markers. The species studied were: *Herichthys carpintis* (Jordan et Snyder, 1899) and *Herichthys pantostictus* (Taylor et Miller, 1983). We estimate their genetic diversity, gene flow, and demographic history and perform biogeographic reconstructions using a Bayesian computation approach and environmental niche modeling. The biogeographic reconstruction suggests a different history for each species. Environmental niche modeling indicates that both species experienced a demographic expansion during the Pleistocene but responded differently to Pleistocene climatic changes. In summary, their current sympatric distribution could be the outcome of contemporary and not historical processes reflecting a pseudo-incongruent pattern.

Keywords

Bayesian computation, gene flow, niche modeling, paleohydrography

Introduction

Evolutionary biogeography seeks to reconcile dispersal and vicariance paradigms through five stages that reconstruct a geobiotic scenario that attempts to explain biotic component evolution based on geological data (Morrone 2007). One stage involves dating the vicariant events using molecular clocks and phylogeographic studies. Phylogeography deals with how genetic lineages are arranged through geographic space (Avise 2009). Therefore, comparative phylogeography studies are relevant since they enable us to interpret how different cenocrons (biotic elements) have been integrated into a horobiota (a snapshot of a biota in a particular time) (Morrone 2020). In a comparative phylogeographic study, three different scenarios could arise. First, in a concerted response, the codistributed species respond similarly to geological and climatic events leading to a congruent pattern. Second, in an

independent response, the codistributed species show independent responses to simultaneous regional processes. Third, in a multiple response, the species could show similar spatial congruence but different temporal frames leading to a pseudo-congruent pattern or different responses both in space and time leading to a pseudo-incongruent pattern (Bagley and Johnson 2014). The cooling and heating periods experienced by the planet during the Pleistocene led to speciation and extinction events and changes in the distributional patterns of species worldwide (Hewitt 1996, 2000, 2004). However, Pleistocene glaciations also affect populations, leaving genetic signatures that are traceable through phylogeographic studies (Hewitt 2003; Lister et al. 2005; Comte and Grenouillet 2015).

These phenomena are more evident in freshwater ecosystems due to climate changes and geological factors, such as volcanism and orogenesis, playing a fundamental role in drainages rearrangement, leading to isolation, reconnection, and formation of new rivers and lakes (Bermingham and Martin 1998; Waters and Wallis 2000; Rincon-Sandoval et al. 2019). Their effects are particularly apparent in species distributed toward the coastal shoreline (Abreu et al. 2020; Pio and Carvalho 2021) due to the Pleistocene's eustatic sea level changes leading to the exposition and covering of the continental shelf, promoting both population connection and isolation (Hewitt 2004; Lambeck and Chappell 2001).

Due to Mexico's complex geologic and paleohydrographic history, its freshwater fishes are ideal for phylogeographic studies (Marshall and Liebherr 2000; Morrone 2004; Domínguez-Domínguez and Pérez-Ponce de León 2009), their study has already led to several published papers in selected taxonomical groups, including atherinopsids (Bloom et al. 2009; García-Martínez et al. 2020), characids (Strecker et al. 2004; Ornelas-García et al. 2008; Hausdorf et al. 2011; Coghill et al. 2014), cyprinids (García-Andrade et al. 2021), and poecilids (Mateos et al. 2002; Chen and Borowsky 2004; Mateos 2005; Gutiérrez-Rodríguez et al. 2007; Bagley et al. 2013).

Cichlids are one of the most diverse clades of freshwater fishes in Mexico. However, despite their species abundance, phylogeographic studies with them are scarce (e.g., Barluenga and Meyer 2010; Bagley et al. 2017; McMahan et al. 2017). The genus *Herichthys* Baird and Girard, 1854 represents an excellent model for evolutionary and biogeographic studies since it includes species of wide and restricted geographic distributions that have been extensively studied in recent years, from systematics to molecular clocks and biogeography among others (Pérez-Miranda et al. 2018, Pérez-Miranda et al. 2020).

In this study, we compared the phylogeographic structures of two cichlid species of the genus *Herichthys* that have a wide geographic distribution in the Pánuco–Tamesí system: *Herichthys carpintis* (Jordan et Snyder, 1899) and *Herichthys pantostictus* (Taylor et Miller, 1983). The latter species was included in the past in the genus *Nosferatu* (see De la Maza-Benignos et al. 2015), however, as pointed out by Říčan et al. (2016), the shape of the teeth

that are the diagnostic character of the genus is a plesiomorphic state for the *Theraps–Parannetroplus* clade and it is also present in some species of the genus *Herichthys* turning the genus *Nosferatu* in a paraphyletic and not natural group. Therefore, all species previously placed in the genus *Nosferatu* by De la Maza-Benignos et al. (2015) should now be included in the genus *Herichthys* (see Fricke et al. 2023; Froese and Pauly 2023; WoRMS Editorial Board 2023). The Pánuco–Tamesí system drains part of the Gulf Coastal Plain in northeastern Mexico covering a surface of near 157 752 km^2 and comprises 11 sub-basins (FAO 2022). This vast region is characterized by altitudes ranging from 0 to 600 m above sea level and was prone to the effects of marine transgressions and regressions during the Pleistocene glaciation periods (Álvarez 1961; Bagley et al. 2013). While both species have a sympatric distribution throughout most of their geographic range, their current geographic distribution patterns are likely the result of different evolutionary and biogeographic histories (Pérez-Miranda et al. 2020). While *H. carpintis*' sister species, *Herichthys tepehua* De la Maza-Benignos, Ornelas-García, Lozano-Vilano, García-Ramírez et Doadrio, 2014 is distributed towards the coastline, *H. pantostictus*' sister group, comprising *Herichthys bartoni* (Bean, 1982) and *Herichthys labridens* (Pellegrin, 1903), is distributed inland (Pérez-Miranda et al. 2018).

Therefore, we expected these species to show a pseudo-incongruent pattern due to different past distributions and recent community assemblies with a scarce or null shared history between them. We test our prediction by evaluating the effect of Pleistocene glaciations on the colonization and connectivity of *H. carpintis* and *H. pantostictus* populations by determining the numbers and ages of their genetic populations, the gene flow among them, and their demographic history and colonization processes using two mitochondrial markers, COI, and D-loop.

Material and methods

Sampling and genetic analysis. The specimens of *Herichthys carpintis* and *Herichthys pantostictus* used were collected between 2000 and 2016 and covered the known geographic distribution of both species (Suppl. material 1). Tissue samples were obtained from 96 *H. carpintis* individuals and 60 *H. pantostictus* individuals. DNA extraction was performed according to the protocol of Aljanabi and Martinez (1997) for amplifying mitochondrial markers cytochrome oxidase subunit 1 (COI) and D-loop. For the mitochondrial COI marker, we complemented our *H. carpintis* data set with an additional 104 previously generated sequences available in two Barcode of Life Data (BOLD) projects (FFPTR and HBGM).

Polymerase chain reactions (PCRs) had a final volume of 25 µL. For the COI marker, we used the primers reported by Ward et al. (2005) with the conditions reported by León-Romero et al. (2012) to amplify a fragment of 589 bp for the D-loop marker, we designed FPM-F

(5′-CTTTGGGAGTTAGGGGTGA-3′) and FPM-R (5′-CACTGAAGATGTTAAGACGG-3′) primers to amplified a 687 bp fragment in a reaction mix comprising 1× PCR buffer, 3 mM MgCl$_2$, 200 μM dNTPs, 0.15 μM of each primer, 40 ng of DNA template, and 1 U of Go-Taq (Invitrogen) using the following conditions: initial preheating at 95°C for 5 min, followed by 35 cycles of denaturing at 96°C for 1 min, annealing at 60°C for 1.5 min, and extension at 72°C for 1 min, with a final 5 min extension at 72°C. PCR products were purified, and both strands were sequenced at the Laboratorio Nacional de Genómica para la Biodiversidad campus (Irapuato, México).

Sequences were aligned with Clustal X versión 2.0 (Larkin et al. 2007) and edited in Seaview (Gouy et al. 2010). The final numbers of available sequences were 195 for COI and 73 for D-loop in *H. carpintis* and 41 for COI and 54 for D-loop in *H. pantostictus*.

Population genetics. We are aware of the potential caveats of dealing with sample size and their effect on genetic diversity and other population genetics estimators. For the aforementioned, we evaluate if the number of analyzed individuals was enough using the function HACSim s implemented in the R library HACSim to build an accumulation curve and to estimate the R-value that represents the estimated fraction of species haplotype diversity captured from sampling (Phillips et al. 2020).

After that, we assessed the number of genetic clusters (*k*) (hereinafter referred to as populations) using the "optimise.baps" option in the *fastbaps* library (Tonkin-Hill et al. 2019) of the R statistical software version 4.0.4 (RStudio Team 2020). Each population's diversity parameters, such as haplotypic (*h*) and nucleotidic (π) diversity, were estimated using Arlequin v.3.5.2 (Excoffier and Lischer 2010).

A time-calibrated phylogenetic tree was estimated for each species and molecular marker in BEAST v 1.7.5 (Drummond et al. 2205), to date the trees, we used secondary calibration points previously estimated for *H. carpintis* and *H. pantostictus* (see Pérez-Miranda et al. 2020) which used different calibration points including fossil record and vicariant events (but see also Říčan et al. 2013); for *H. carpintis*, we used the split with her sister species *H. tepehua* (outgroup in the phylogeny) occurred 2.5 Ma, meanwhile, for *H. pantostictus* we used its divergence time against her sister clade (*H. bartoni* + *H. labridens*) (outgroup in the phylogeny) occurred 6.5 Ma. Four independent runs of 10 million generations, sampling every 10 000 generations, were performed assuming a GTR substitution model, a strict molecular clock, and a Yule speciation model; convergence among chains was assumed if the ESS values were higher than 200, then, the trees were summarized using LogCombiner v.1.7.4 (Drummond and Rambaut 2007), and a consensus tree was constructed after a 25% burn-in using TreeAnotator v2.6.6 (Drummond and Rambaut 2007).

To evaluate gene flow levels among populations, we used the Bayesian approach implemented in migrate v.4.4.2 (Beerli 1998; Beerli and Felsenstein 2001) using a static heating scheme with four temperature chains (1, 1.5, 3, and 1 000 000), with each analysis comprising 10 000 000 genealogies sampled every 1000 generations after a 10% burn-in. We used the full matrix model as a null hypothesis for each species and molecular marker and postulated several gene flow models following Miller et al. (2005). Since *H. pantostictus* is mainly distributed inland, its colonization process should be from inland towards the coastline. In contrast, since *H. carpintis* is mainly distributed along the coastal shoreline, its colonization process should be from the coastline towards the inland.

Ten gene flow models were evaluated for *H. pantostictus* (five per molecular marker), and 11 models were evaluated for *H. carpintis* (six for COI and five for D-loop; see Suppl. material 2 for complete details). We compared gene flow models using a Bayes Factor test with the Bezier approach's marginal likelihood (Beerli et al. 2019) using the BF function in R's *mtraceR* library (Pacioni et al. 2015).

Effective population size changes were inferred using Tajima's *D* and Fu's F_s tests in Arlequín v.3.5.2 (Excoffier and Lischer 2010). In addition, effective population size changes through time were inferred from Bayesian skyline plots (BSPs) created using BEAST v1.7.5 (Drummond et al. 2005) with a lumping approach since coalescent-based tests are extremely sensitive to sample size (Heller et al. 2013). The BSPs plots were created using four Markov chains of 10 000 000 generations, sampled every 1000 generations, and a strict molecular clock with the same calibration points mentioned above. The chains' results were combined using LogCombiner v1.7.4 after a 25% burn-in (Drummond and Rambaut 2007), and the BSPs were plotted in Tracer v1.5 (Rambaut et al. 2018).

Biogeographic scenarios and niche modeling. We tested two biogeographic scenarios for the colonization routes of each species using the approximate Bayesian computation (ABC) approach implemented in the DIYABC software (Cornuet et al. 2010). Scenario one assumed a coastal-to-inland colonization process. Scenario two assumed an inland-to-coastal colonization process. The scenarios were compared using reference tables simulating 1×10^6 datasets based on haplotype numbers and the Tajima and Fu test values. First, considering the reference tables' first 10 000 scenarios, we used a principal component analysis to evaluate whether the generated dataset's distribution approached that of the observed dataset. Then, a normalized Euclidean distance between the simulated and the observed datasets was calculated to determine the most plausible scenario. Finally, considering the 1% of generated datasets closest to the observed datasets, a direct and logistic regression was used to estimate the posterior probability and the type I and type II errors for each scenario with a 95% highest posterior distribution (HPD). Therefore, the most probable scenario was chosen based on the highest posterior probability and the absence of overlap in the HDP intervals (Cornuet et al. 2010).

Finally, we used an environmental niche model (ENM) approach to evaluate the possible effects of Pleistocene glaciations on both species' demographic history. The suitability areas for each species were determined using the maximum entropy algorithm implemented in Maxent v.3.2.19 (Phillips et al. 2006) applied to the Mejía et al. (2022) collecting dataset. First, the *spThin* R library (Aiello-Lammens et al. 2015) was used to reduce spatial autocorrelation by pruning collection records that are <1 kilometer, leaving a total of 205 collection records for *H. carpintis* and a total of 124 collection records for *H. pantostictus*. Then, the geographic space *M* for each species was defined with the Pfafstetter HydroBasin levels 6 and 7 (Lehner and Grill 2013) in QGIS v.3.16.5 (QGIS Development Team 2009). The 19 WorldClim bioclimatic variables (Hijmans et al. 2005) were downloaded for the current period. The paleoclimatic projections used the MPI-ESM-P general circulation model in three temporal frames: the last interglacial period (LIG; 120 ka), the last glacial maximum (LGM; 21 ka), and the mid-Holocene (6 ka). Additionally, three topographic variables (aspect, topographic position index, and slope) were calculated based on a digital elevation model from Hydrosheds (Lehner et al. 2008) using the terrain function of the *raster* R library (Hijmans and Van Etten 2012). The number of variables in the ENM was reduced using the variance inflation factor (VIF) with the *vifcor* and *vifstep* functions in the *usdm* R library (Naimi et al. 2014).

The ENMs were constructed using the *dismo* (Hijmans et al. 2020), *ENMeval* (Muscarella et al. 2014), *rmaxent* (Baumgartner et al. 2017), and *kuenm* (Cobos et al. 2019) R libraries. We used 30% of the collection records for model construction and the remaining 70% for model training with several combinations of feature classes (linear, quadratic, product, and threshold) and regularization multipliers (1, 2, 5, 10, 15, and 20; Warren and Seifert 2011). Finally, the best ENM models were selected using the Akaike information criteria (Akaike 1974) and partial receiver operating characteristic (ROC) curve values over 1000 bootstraps (Peterson and Nyari 2008).

Results

Population genetics. The new sequences of *Herichthys carpintis* and *Herichthys pantostictus*, generated for this study, were deposited in GenBank under the accession numbers OP738881–OP738896; OP738385–OP738395, and OP751419–OP751525 (Suppl. material 3). The R values recovered from the HACSim curves were higher than 0.9 for both species and molecular markers (Suppl. material 4) and suggest that the number of individuals analyzed was enough to recover the genetic diversity of both species.

The *fastbaps* analysis recovered three genetic clusters (populations) for each molecular marker. In *H. pantostictus*, based on the COI marker, population one (Guayalejo) comprised five individuals and had a geographic centroid

near the coastal shoreline at the northern limit of its distribution. Population two (Pánuco) comprised 20 individuals with a geographic centroid located near the center of its geographic distribution. Population three (Tanconchín) comprised 16 individuals with a geographic centroid in the south of its distribution (Fig. 1). Based on the D-loop marker, population four (Tamesí) comprised 23 individuals and had a geographic centroid toward the north of its distribution. Population five (Valles) comprised 19 individuals and had a geographic centroid at the center of its distribution. Population 6 (Naranjos) comprised 12 individuals and had a geographic centroid toward the South of its distribution (Fig. 1).

In *H. carpintis*, based on the COI marker, population one (Ozuluama) comprised 90 individuals and had a geographic centroid in the south of its distribution. Population two (Adjuntas) comprised 15 individuals and had a geographic centroid in the center of its distribution. Population three (Mante) comprised 90 individuals and had a geographic centroid in the north of its distribution. For the D-Loop marker, population four (Jaumave) comprised 50 individuals and had a geographic centroid in the north of its distribution. Population five (Tempoal) comprised 12 individuals and had a geographic centroid in the center of its distribution. Population six (Gallinas) comprised 11 individuals and had a geographic centroid in the center of its distribution toward the inland (Fig. 1).

The *h* values in *H. pantostictus* populations ranged from 0.542 in Tanconchín to 0.900 in Guayalejo for the COI marker and from 0.544 in Valles to 0.573 in Tamesí for the D-loop. In *H. carpintis* populations, *h* values were low for the COI marker, ranging from 0.391 in Mante to 0.664 in Ozuluama, but high for the D-loop marker, ranging from 0.818 in Gallinas to 1.000 in Tempoal. The π values were low in the majority of populations, ranging from 0.001 for the COI marker in the *H. carpintis* Adjuntas and Mante populations to 0.026 for the D-loop marker in the *H. carpintis* Jaumave population (Table 1).

The molecular clock analysis inferred similar root ages for both markers in *H. pantostictus* (6.9 Ma, HPD 5.3–8.4 Ma). Population ages inferred based on the COI marker were 0.69 Ma (HPD 0.2–1.1 Ma) for Guayalejo, 1.91 Ma (0.8–3.0 Ma) for Pánuco, and 2.68 Ma (HPD 1.1–4.6 Ma) for Tanconchín. Population ages inferred from the D-loop marker were 2.55 Ma (HPD 1.2–3.8 Ma) for Naranjos, 2.15 Ma (HPD 1.0–3.3 Ma) for Valles, and 2.53 Ma (HPD 1.2–3.8 Ma) for Tamesí (Fig. 2). However, they provided opposing ancient population geographic locations. The youngest population was located toward the coastline for the COI marker (Guayalejo) but inland for the D-loop marker (Valles). In *H. carpintis*, the inferred root ages were 1.9 Ma (HPD 1.5–2.3 Ma) with the COI marker and 2.3 Ma (HPD 1.6–2.9 Ma) with the D-loop marker. Population ages inferred based on the COI were 1.47 Ma (HPD 0.9–1.9 Ma) for Ozuluama, 0.82 Ma (HPD 0.3 1–3 Ma) for Adjuntas, and 1.24 Ma (0.7–1.7 Ma) for Mante. Population ages inferred based on the D-loop marker were 1.9 Ma (HPD 1.5–2.3 Ma) for
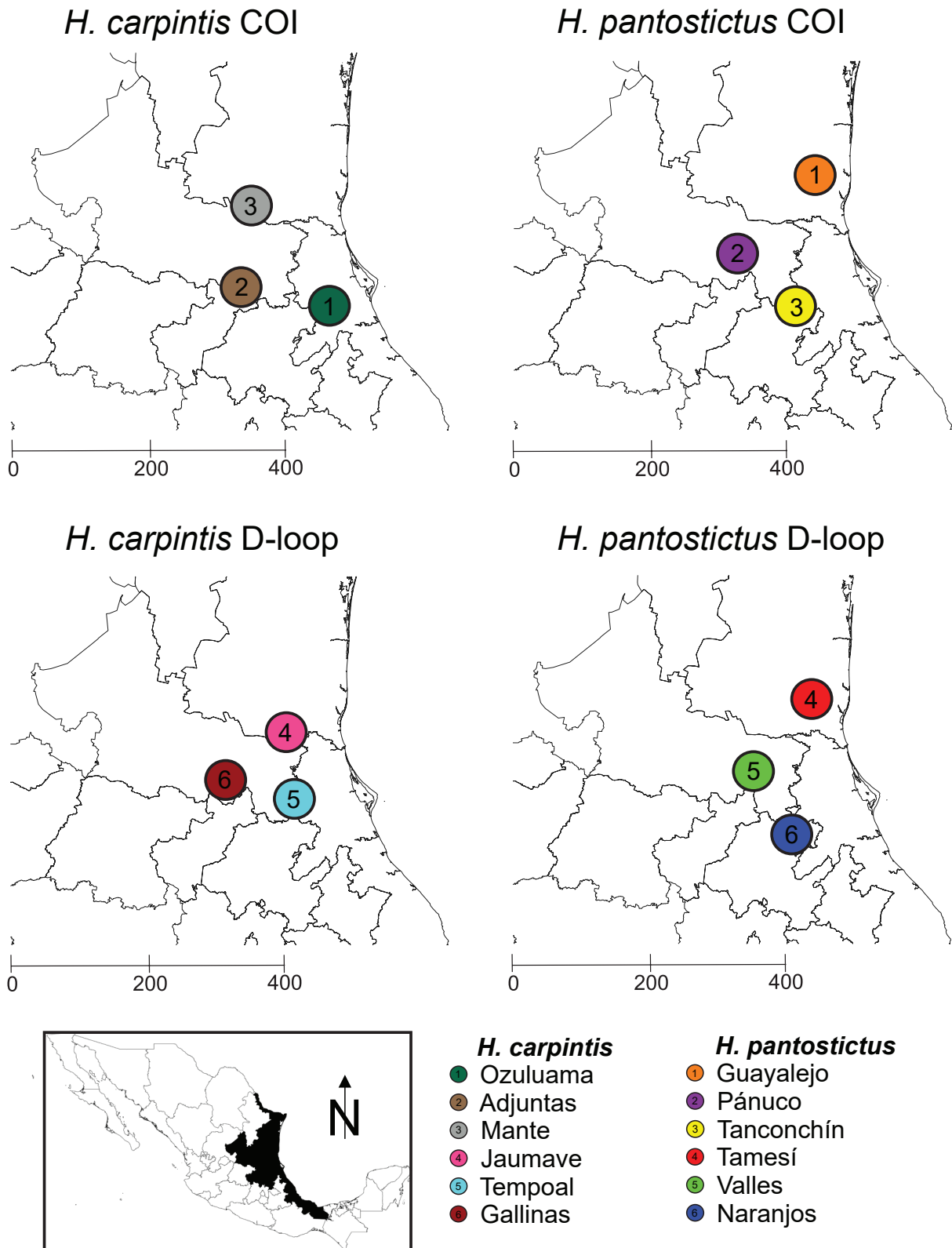
H. carpintis COI

H. pantostictus COI

H. carpintis D-loop

H. pantostictus D-loop

H. carpintis
1 Ozuluama
2 Adjuntas
3 Mante
4 Jaumave
5 Tempoal
6 Gallinas

H. pantostictus
1 Guayalejo
2 Pánuco
3 Tanconchín
4 Tamesí
5 Valles
6 Naranjos

Figure 1. Geographic distribution of genetics populations of *Herichthys carpintis* and *Herichthys pantostictus* recovered by fast-baps. Each circle represents the geographic centroid of the localities that contribute to the formation of the genetic group

Jaumave, 1.2 Ma (HPD 0.8–1.6 Ma) for Tempoal, and 0.7 Ma (0.3–1.0 Ma) for Gallinas (Fig. 2). Therefore, the youngest *H. carpintis* population was toward the inland with both markers: Adjuntas with the COI marker and Gallinas with the D-loop marker (Fig. 2).

The migrate analysis, which estimated gene flow, suggested that the most probable model for *H. pantostictus* based on the COI marker was colonization from inland towards the coastline (maximum likelihood [ML] = −1101.44, $P = 0.964$; Suppl. material 5), with Tanconchín

Table 1. Summary of the genetic diversity statistics recovered in *Herichthys pantostictus* and *Herichthys carpintis* for the mitochondrial molecular markers COI and D-loop.

Species	Marker	*k*	*n*	*h*	*π*	*D*	*F*$_s$
H. pantostictus	COI	Guayalejo	5	0.900	0.003	−1.094	−3.578[P2]
		Pánuco	20	0.447	0.002	−1.719[P1]	−34.080[P2]
		Tanconchín	16	0.542	0.003	−2.003[P1]	−27.681[P2]
	D–loop	Tamesí	23	0.573	0.003	−2.006[P1]	−27.762[P2]
		Valles	19	0.544	0.003	−1.894[P1]	−28.311[P2]
		Naranjos	12	0.848	0.007	0.483	−11.383[P2]
H. carpintis	COI	Ozuluama	90	0.664	0.004	1.209	1.530
		Adjuntas	15	0.562	0.001	0.139	−10.727[P2]
		Mante	90	0.391	0.001	−0.886	−14.130[P2]
	D–loop	Jaumave	50	0.995	0.026	−1.099	−24.227[P2]
		Tempoal	12	1.000	0.018	0.487	−4.241[P1]
		Gallinas	11	0.818	0.007	−0.712	−2.855[P1]

[P1] = $P < 0.05$, [P2] = $P < 0.001$; *k* = number of genetic clusters recovered by *fastbaps*, *n* = number of individuals examined, *h* = haplotypic diversity, *π* = nucleotidic diversity, *D* = Tajima *D* test values, *F*s = Fu test values.
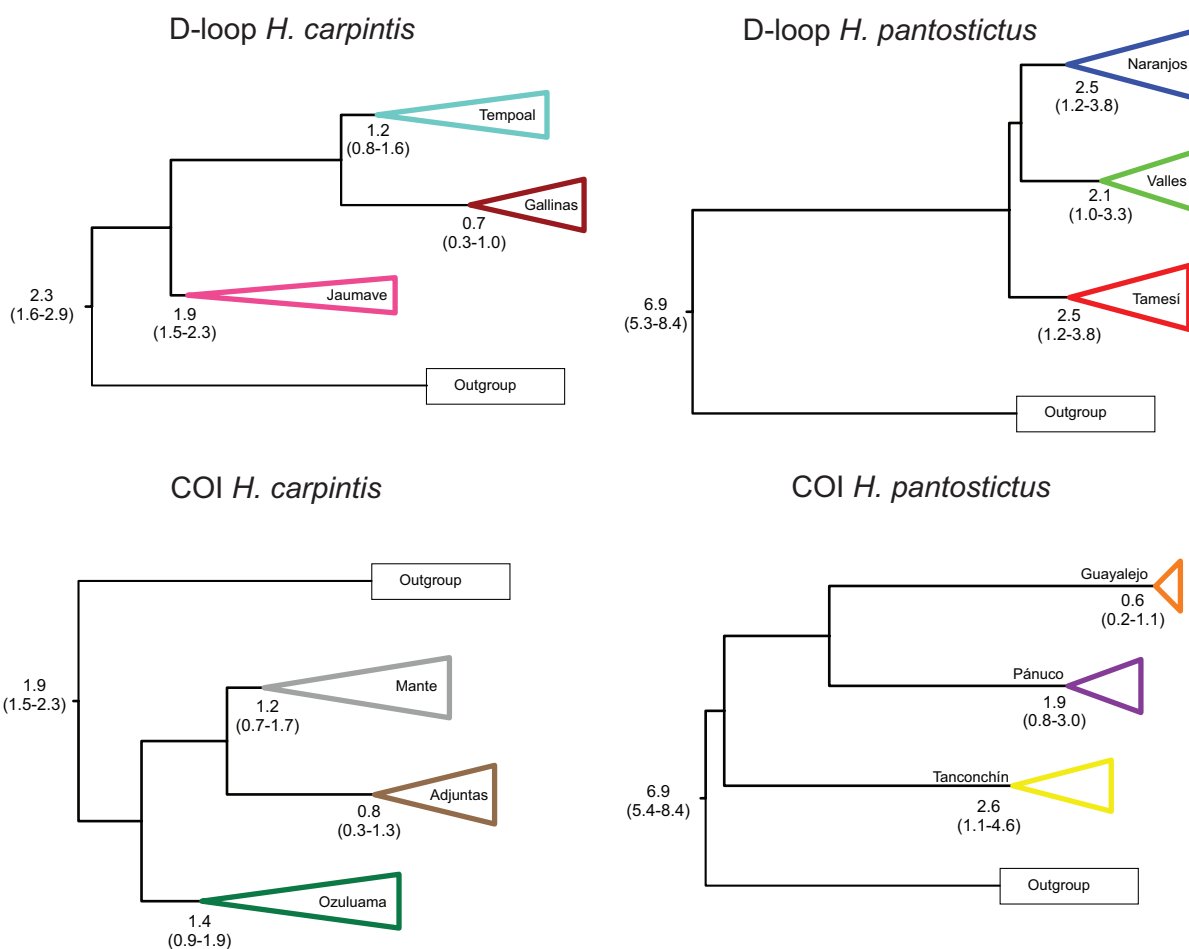
Figure 2. Molecular dated phylogeny of the genetic clusters recovered in *Herichthys carpintis* and *Herichthys pantostictus* for the mitochondrial markers COI and D-loop. The median divergence time (Ma) and the HPD intervals (Ma) are shown below the nodes. The colors represent the genetic clusters depicted in Fig. 1

giving rise to Pánuco (*M* = 138; 95% HPD: 0–420), followed by Pánuco giving rise to Guayalejo (*M* = 138; 95% HPD: 0–937; Fig. 3). However, for the D-loop marker, the colonization process occurred from the distribution's center to the North and then to the South (ML = −1040.74, *P* = 0.668; Suppl. material 5), with Valles giving rise to Tamesí

(*M* = 322; 95% HPD: 83–357), followed by Tamesí giving rise to Naranjos (*M* = 415; 95% HPD: 0–483; Fig. 3).

The gene flow estimation for *H. carpintis* suggests that the colonization process occurred from the coastline toward the inland for the COI marker (ML = −3099.68, *P* = 1.000), with Ozuluama giving rise to Mante (*M* = 82;
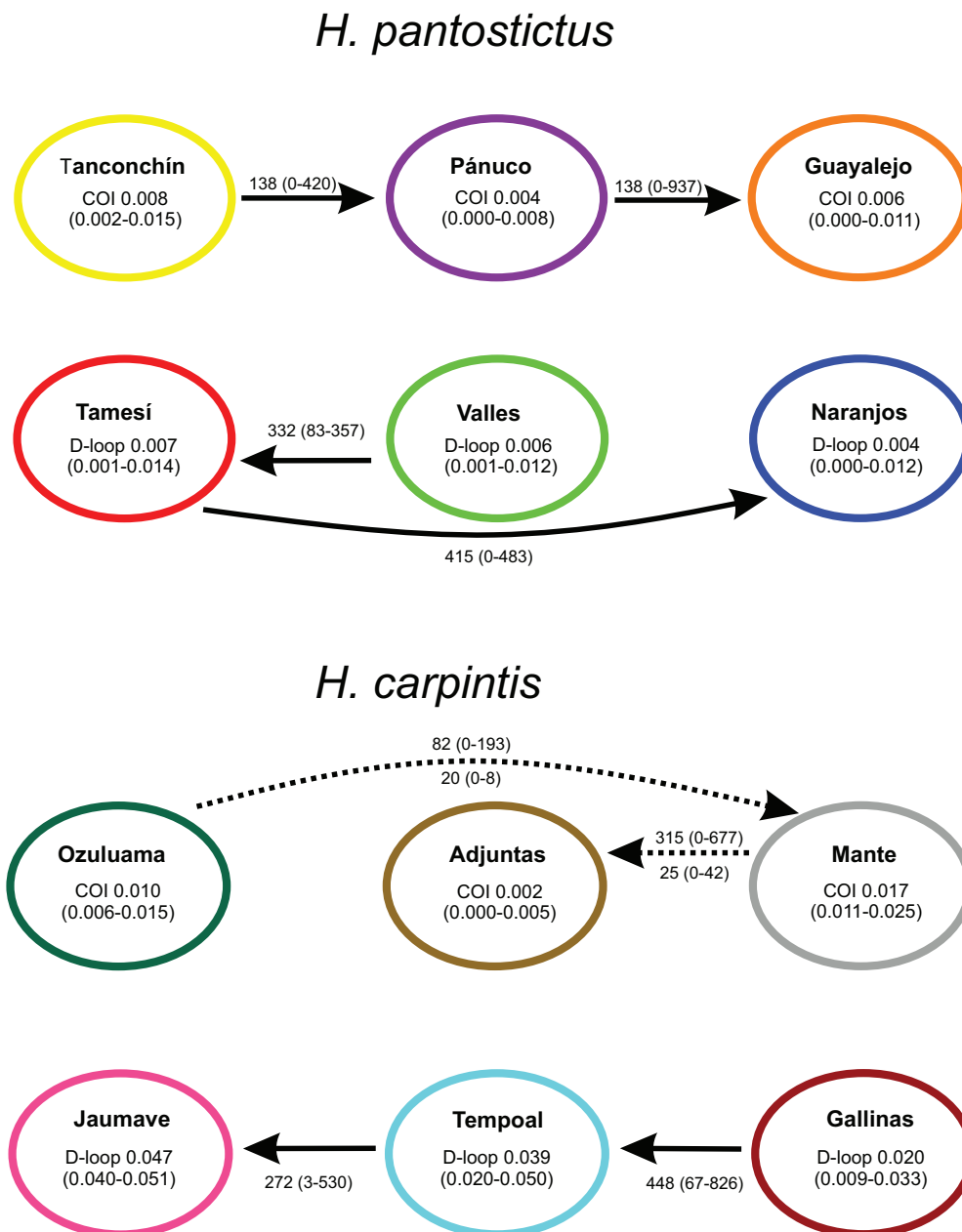
H. pantostictus

H. carpintis

Figure 3. Gene flow among populations of *Herichthys carpintis* and *Herichthys pantostictus* according to the most probable model recovered in Bayes Factor analysis (see Suppl. materials 2, 4 for complete details). Theta values and 95% HPD intervals are shown within each genetic cluster. The values on the arrows show the number of migrants from one population to another with the minimum and maximum values maximum in parentheses; solid arrows indicate a single gene flow event, dashed arrows indicated recurrent gene flow (values below), after the initial migrant event (values above).

95% HPD: 0–193), followed by Mante giving rise to Adjuntas (*M* = 315; 95% HPD: 0–677; Fig. 3). Finally, for the D-loop, the colonization process occurred from inland to the coastline (ML = −4002.55, *P* = 1.000), with Gallinas giving rise to Tempoal (*M* = 448; 95% HPD: 67–286), followed by Tempoal giving rise to Jaumave (*M* = 275; 95% HPD: 3–530; Fig. 3).

The Tajima's D and Fu's F tests used to evaluate effective population size changes suggest a demographic expansion of the *H. pantostictus* populations, except the Tajima tests for both COI and D-loop markers in the Naranjos and Guayalejo populations (Table 1). Both

markers' BSPs provided similar results (Fig. 4). For the COI marker, a slight demographic decrease occurred at ~1 Ma, followed by a sudden expansion at ~250 kya. Similarly, for the D-loop marker, a demographic expansion occurred at ~250 kya (Fig. 4). However, in *H. carpintis*, the Tajima test was non-significant in all populations with both markers. In contrast, the Fu test suggested an expansion in all populations with both markers, except for Ozuluama with the COI marker (Table 1). Finally, both markers' BSPs suggested a demographic expansion starting at 100 kya for the COI marker and 1 Ma for the D-loop marker (Fig. 4).

Biogeographic scenarios and niche modeling. The ABC analysis inferred opposing biogeographic histories for *H. pantostictus* with each molecular marker. For the COI marker, the most probable scenario suggests colonization from the coastline to inland populations (scenario one; logistic posterior probability = 0.596; 95% HPD: 0.587–0.604). In contrast, scenario two was the most plausible for the D-loop marker, suggesting an inland to coastline colonization (*P* = 0.804; 95% HPD = 0.797–0.813; Table 2; Suppl. material 6). However, for *H. carpintis*, scenario one was inferred as the most probable scenario for both molecular markers, suggest-

ing colonization from the coastline to inland (Table 2; Suppl. material 6).

Finally, in the ENM, eight bioclimatic (BIO2, BIO3, BIO5, BIO9, BIO13, BIO14, BIO15, and BIO18) and three topographic variables were retained after the VIF test. The potential geographic distribution models for *H. pantostic-tus* and *H. carpintis* showed good performance (area under the ROC curve [AUC] >0.8; AUC [partial ROC curve] >1.1; Table 3) without over-adjustment (AUC difference = 0.143–0.190 for *H. pantostictus* and 0.159–0.178 for *H. carpintis*; Table 3). For both species, most of the observed variance was explained by the warmest month's maximum
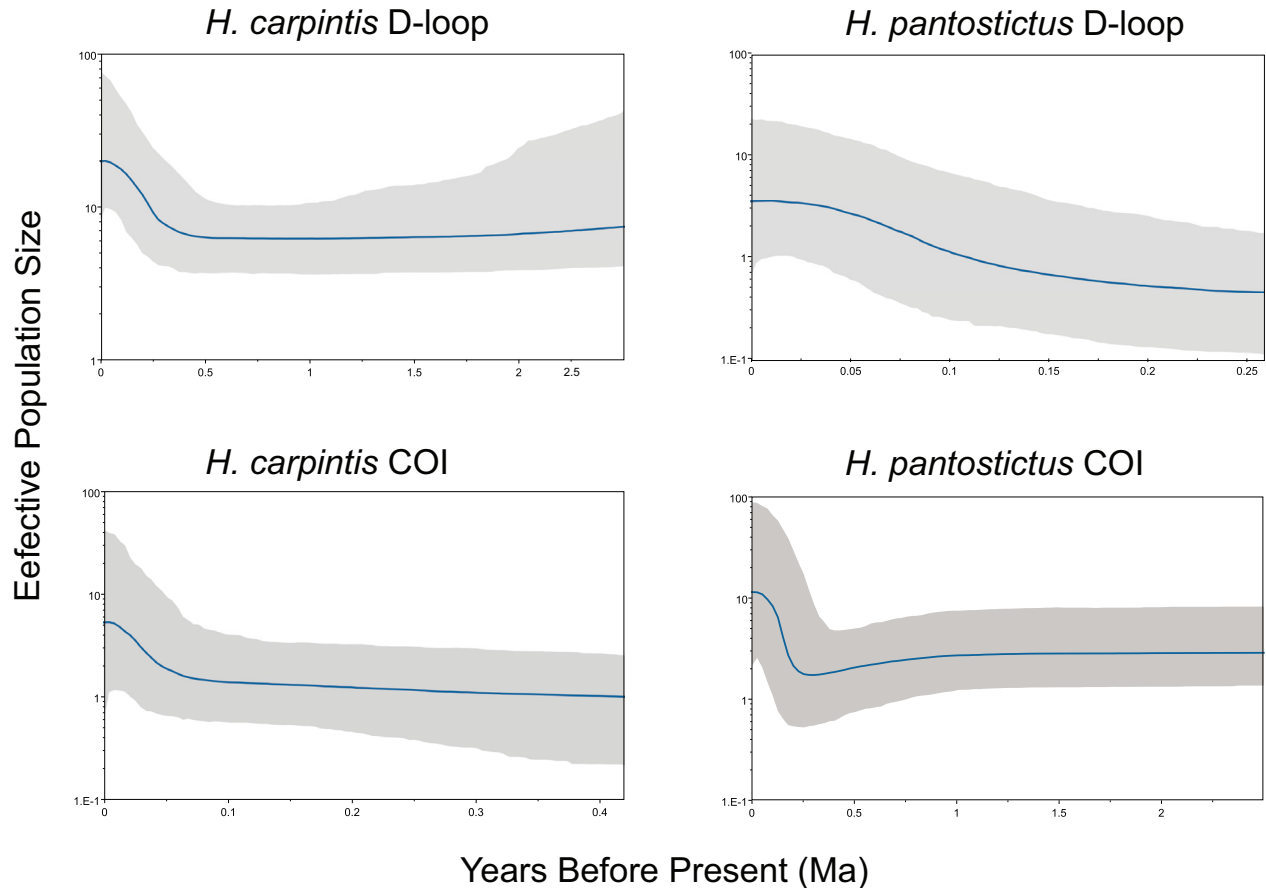

Figure 4. Bayesian skyline plots (BSP) of *Herichthys carpintis* and *Herichthys pantostictus* recovered from the analysis of the lumping populations of each of the molecular markers analyzed in this study.

Table 2. Scenarios used in the Approximate Bayesian computation (ABC) for the biogeographic history of *Herichthys pantostictus* and *Herichthys carpintis* for the two molecular markers used in this study. The posterior probability for the direct and the logistic regression as well as probability of type I and type II errors are indicated for each one of the postulated scenarios. Selected scenarios are set in bold typeface.

Species	Marker	Set	Scenario	Posterior probability				Type I error		Type II error	
				Direct	95% HPD	Logistic	95% HPD	Direct	Logistic	Direct	Logistic
H. pantostictus	COI	**Coast–Inland**	**1**	**0.504**	**0.066, 0.942**	**0.596**	**0.587, 0.604**	**0.457**	**0.463**	**0.512**	**0.542**
		Inland–Coast	2	0.496	0.058, 0.934	0.404	0.396, 0.413	0.524	0.514	0.645	0.487
	D-loop	Coast–Inland	1	0.376	0.000, 0.801	0.195	0.187, 0.203	0.588	0.642	0.745	0.789
		Inland–Coast	**2**	**0.624**	**0.199, 1.000**	**0.805**	**0.797, 0.813**	**0.489**	**0.514**	**0.564**	**0.561**
H. carpintis	COI	**Coast–Inland**	**1**	**0.81**	**0.466, 1.000**	**0.997**	**0.997, 0.998**	**0.401**	**0.402**	**0.654**	**0.641**
		Inland–Coast	2	0.19	0.000, 0.534	0.003	0.002, 0.003	0.506	0.48	0.895	0.985
	D-loop	**Coast–Inland**	**1**	**0.522**	**0.084, 0.960**	**0.699**	**0.713, 0.764**	**0.284**	**0.236**	**0.689**	**0.657**
		Inland–Coast	2	0.478	0.040, 0.916	0.301	0.237, 0.287	0.311	0.343	0.716	0.764

Table 3. Best Maxent model selected for *Herichthys pantostictus* and *Herichthys carpintis* for each one of the four periods considered in this study. LIG (Last interglacial 120 ky), LGM (Last Glacial Maximum (21 ky), mid Holocene (6 ky), Current. Best fit MaxEnt model select for each of species for the four periods considered.

Period	Parameter						
	AUC	Diff. AUC	AICc	pROC	*w*	Area [km²]	Area [%]
H. pantostictus							
LIG	0.898	0.172	2912.548	1.458	0.750	148262.6	90.0
LGM	0.873	0.190	2954.457	1.267	0.895	129617	80.0
Mid Holocene	0.876	0.143	2941.143	1.555	0.889	109585	68.0
Current	0.885	0.174	2942.630	1.294	0.817	154917	96.0
H. carpintis							
LIG	0.912	0.161	4079.705	1.246	1.000	119371.4	81.4
LGM	0.871	0.167	4129.669	1.395	0.968	118783	81.0
Mid Holocene	0.878	0.159	4118.810	1.294	0.998	113358	77.3
Current	0.859	0.178	4138.734	1.142	0.970	117648	80.3

ROC = receiver operating characteristic, AUC = area under the curve, Diff. AUC = mean difference in the AUC values between the training data and the test data, AICc = Akaike information criterion, pROC = partial ROC values, *w* = Akaike weight, Area = total suitability area available for the taxon in km² and expressed in km² or as a percentage of the total geographic extent (*M*).

temperature (BIO5 = 14.441% for *H. pantostictus* and 15.506% for *H. carpintis*), the warmest quarter's precipitation (BIO18 = 44.390% for *H. pantostictus* and 20.216% for *H. carpintis*), and altitude (13.361% for *H. pantostictus* and 50.890% for *H. carpintis*).

The geographic extension of suitable climatic areas decreased by 22% from the LIG to the mid-Holocene for *H. pantostictus* (Table 3; Fig. 5 A–C). However, while a similar pattern was inferred for *H. carpintis*, it had a smaller decrease (Table 3; Fig. 5 E–G). Finally, an increase in suitable areas was inferred for both species from the mid-Holocene (6 kya) to the present, which was small (3%) for *H. carpintis* and larger (28%) for *H. pantostictus*.

Discussion

Population structure and diversity. The number of genetic populations recovered in a phylogeographic study depends on several factors, such as the choice of molecular marker; the size of the fragment; sample size; evolutionary processes such as selection, mutation, and genetic drift; and the species' intrinsic attributes related to habitat and life history traits (Gavrilets 2003; Sofia et al. 2006; Matsumoto and Hilsdorf 2009). In cichlids, we must expect that species distributed in lakes will show a near panmictic with a low population structure, and species distributed in rivers, such as the species analyzed in this study, will show a highly structured population. This pattern does not always occur. For example, there are riverine species without genetic structure, such as the South American *Gymnogeophagus setequedas* Reis, Malabarba et Pavanelli, 1992 (see Souza-Shibatta et al. 2018) and other Neotropical cichlids found in rivers in South Mexico (Elías et al. 2020). However, like this study's results, structured populations in riverine environments have been recovered in species distributed in Mexico, such as *Trichromis salvini* (Günther, 1862) (see Elías et al. 2020)

and other South American species (Abreu et al. 2020; Pio and Carvalho 2021).

One possible explanation for the presence of genetic structure in this study could be associated with the Pleistocene glacial periods with more arid environments, leading to water body desiccation and a lack of connectivity and gene flow (Strecker et al. 2004; Gutiérrez-Rodríguez et al. 2007). Then, during the interglacial period, wetter conditions increase water levels, habitat connectivity, and gene flow (Elías et al. 2020). Another possible explanation could be that both species analyzed in this study had populations distributed toward the coastline. In other South American tropical fish species, it has been postulated that high sea level fluctuation due to Pleistocene climatic changes led to sea level transgression and regression. Therefore, when the sea level rose, the mouth of the rivers was displaced inland, promoting population isolation. In contrast, when the sea level dropped, the influence of the freshwater could extend into the continental shelf, promoting connections between previously isolated populations and gene flow (Abreu et al. 2020, Pio and Carvalho 2021).

The potential effect of sea level fluctuations and Pleistocene climatic changes could be masked by ENM caveats. Most ENM algorithms identify suitability areas through correlations with species occurrence data, ignoring factors such as habitat saturation, response to environmental changes, and biotic interactions (Cordellier and Pfenninger 2009; Wiens et al. 2009). Moreover, they cannot deal with aquatic organisms since climatic variables are optimized for terrestrial organisms, leading to suitability area overestimations (Domisch et al. 2011; Elith et al. 2011). Nevertheless, a potential Pleistocene climatic change effect was effectively detected in *Herichthys pantostictus* through the 12% reduction in its suitability area from LGM to mid-Holocene compared to the 4% reduction for *Herichthys carpintis* in the same period (Table 3). Indeed, a close inspection of Fig. 5 showed that
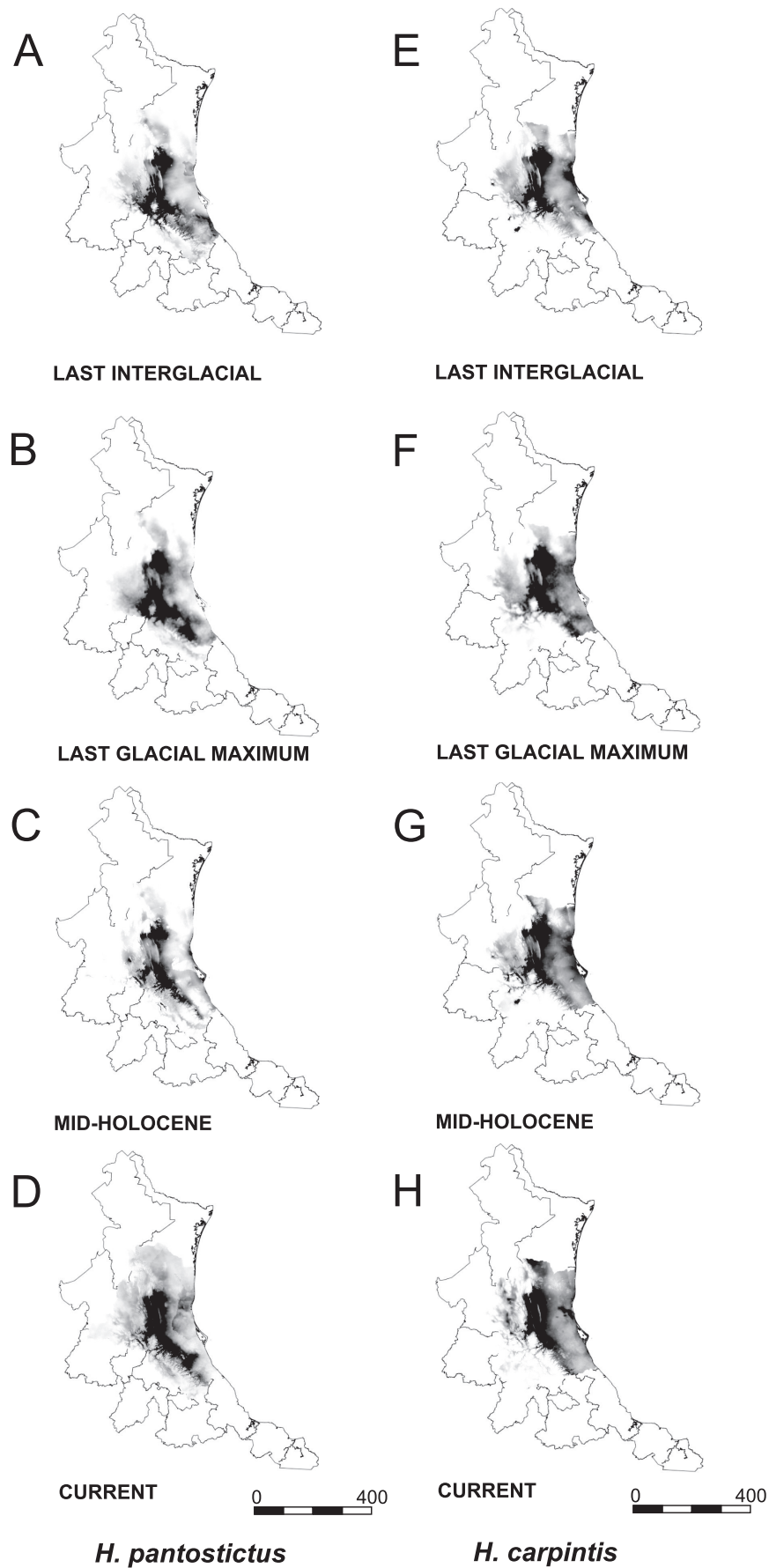
Figure 5. Potential geographic distribution of *Herichthys carpintis* and *Herichthys pantostictus* identified using an ecological niche modelling under current bioclimatic conditions (1950–2000), as well as projections to three paleoclimatic periods (LIG, 120kya; LGM; 21kya; Mid Holocene, 6 kya). Darker areas represent the pixels of major suitability conditions for the presence of the taxon.

most of *H. pantostictus*' lost areas were located toward the coastline. Therefore, despite sympatry across most of their geographic distribution, both species showed different responses to environmental variables, as has been previously proposed based on the inequivalence in their environmental niches (see table S3 in Mejía et al. 2022).

Low genetic diversity levels have been associated with the absence (Olivieri et al. 2008; Craul et al. 2009; McMahan et al. 2017) or the presence of genetic structure (Agbebi et al. 2016). In this study, *h* values ranged from medium (0.390) to high (1.000) with low π values (0.001–0.026), suggesting a sudden population expansion (Duftner et al. 2006; Barluenga and Meyer 2010; Ferreira et al. 2015; Azevedo et al. 2017) that could cause the absence of population genetic structure. However, this apparent absence of genetic structure can be counteracted by the cichlids' life history traits, such as territoriality, parental care, and philopatric behavior, which restrict the displacement of individuals and allow new genetic variants to arise (Budaev et al. 1999; Sofia et al. 2008; Pereira et al. 2009; Sefc 2011). Therefore, the genetic structures of *H. carpintis* and *H. pantostictus* we inferred in this study could be the result of populations that experienced a demographic expansion and life-history-related traits.

Demographic history and ENMs. The BSPs for both markers and species (Fig. 4) suggest a population expansion, consistent with other neotropical cichlids (Barluenga and Meyer 2010; Bagley et al. 2017; McMahan et al. 2017). The demographic expansion found in this study could be attributable to the Pleistocene's heating and cooling periods, as has been proposed for other neotropical fishes (Bagley et al. 2013; McMahan et al. 2017, Beltrán-López et al. 2018). Nevertheless, our study results do not correspond with the results found in the ENM, where a reduction in the available suitability areas was inferred for both species since the LIG (Table 3). In *H. carpintis*, the results suggest a slight reduction (4%) in the suitability area from LIG to mid-Holocene (Table 3) but a drastic reduction (22%) for *H. pantostictus* that must have impacted its demographic history. However, this pattern was not present in the BSP (Fig. 4).

These results could be attributable to different factors. For *H. carpintis*, the ENM showed that since the LIG, the most suitable areas for this species were toward the coastline and have existed up to the present (Fig. 5). Therefore, sea level invasions and regressions during glacial periods had a negligible effect on this species' effective population but did affect its spatial expansion. The gene flow analysis and DIYABC scenarios suggest a coast-to-inland colonization process for *H. carpintis*, consistent with that proposed for other cichlids in Middle America (Bagley et al. 2013; McMahan et al. 2017).

For *H. pantostictus*, the ENM suggested that the most suitable areas were inland since the LGM. Therefore, we expect that eustatic sea level changes must have impacted this species' population size over time. Dry conditions and low precipitation characterized the LGM in Central Mexico, with the evaporation of some water bodies (Caballero et al. 2010) explaining the drastic reduction in this species' available suitable areas from the LGM to mid-Holocene and the apparent slight reduction in its effective population size inferred in the BSP with the COI marker (Fig. 4). Another possible scenario is that the ancestral *H. pantostictus* population was on the coastline and the increase in sea level forced the species to move to inland, as suggested by the DIYABC results with the COI marker (Table 2). Therefore, we could suggest that, contrary to expectations, glacial periods did not significantly affect the demographic history of these cichlid species in Northeast Mexico, as has been previously documented in other codistributed species (Rocamontes-Morales et al. 2021).

Conclusions

This study's results suggest that both species (*Herichthys carpintis* and *Herichthys pantostictus*) arose and developed in different geographic areas: *H. carpintis* in lacustrine environments and low-flow rivers in coastal regions and *H. pantostictus* in clear water inland rivers with medium to high flows. The current sympatry of both species is more compatible with a pseudo-incongruence pattern resulting from dispersal events during the Pleistocene, supported by the DIYABC results for *H. carpintis* with both markers and *H. pantostictus* with the D-loop marker, than an independent response to paleoclimatic events, supported by the ENMs and the DIYABC results for *H. pantostictus* with the COI marker. However, further analysis with other molecular markers is required to disentangle both species' biogeographic histories and to fully understand the complex history of this region of northeastern Mexico.

Acknowledgments

References

Abreu JMS, Waltz BT, Albert JS, Piorski NM (2020) Genetic differentiation through dispersal and isolation in two freshwater fish species from coastal basins of northeastern Brazil. Neotropical Ichthyology 18(3): e190114. https://doi.org/10.1590/1982-0224-2019-0114

Agbebi OT, Echefu CJ, Adeosun IO, Ajibade AH, Adegbite EA, Adebambo AO, Ilori MB, Durosaro SO, Ajibike AB (2016) Mitochondrial diversity and time divergence of commonly cultured cichlids in Nigeria. British Biotechnology Journal 13(2): 1–7. https://doi.org/10.9734/BBJ/2016/25470

Aiello-Lammens ME, Boria RA, Radosavljevic A, Vilela B, Anderson RP (2015) spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography 38(5): 541–545. https://doi.org/10.1111/ecog.01132

Akaike H (1974) A new look at the statistical model identification. IEEE Transactions on Automatic Control 19(6): 716–723. https://doi.org/10.1109/TAC.1974.1100705

Aljanabi SM, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Research 25(22): 4692–4693. https://doi.org/10.1093/nar/25.22.4692

Álvarez M Jr (1961) Provincias fisiográficas de la República Mexicana. Boletín de la Sociedad Geológica Mexicana 24(2): 5–20. https://doi.org/10.18268/BSGM1961v24n2a1

Avise JC (2009) Phylogeography: Retrospect and prospect. Journal of Biogeography 36(1): 3–15. https://doi.org/10.1111/j.1365-2699.2008.02032.x

Azevedo RS, Bitencourt A, Silva DA, Amorim A, Mazzoni R, Carvalho EF, Amaral CRL (2017) Genetic diversity of *Geophagus brasiliensis* from the South American Atlantic Rainforest. Forensic Science International. Genetics Supplement Series 6: e433–e434. https://doi.org/10.1016/j.fsigss.2017.09.149

Bagley JC, Johnson JB (2014) Testing for shared biogeographic history in the lower Central American freshwater fish assemblage using comparative phylogeography: Concerted, independent, or multiple evolutionary responses? Ecology and Evolution 4(9): 1686–1705. https://doi.org/10.1002/ece3.1058

Bagley JC, Sandel M, Travis J, Lozano-Vilano MDL, Johnson JB (2013) Paleoclimatic modeling and phylogeography of least killifish, *Heterandria formosa*: Insights into Pleistocene expansion-contraction dynamics and evolutionary history of North American Coastal Plain freshwater biota. BMC Evolutionary Biology 13(1): 223. https://doi.org/10.1186/1471-2148-13-223

Bagley JC, Matamoros WA, McMahan CD, Tobler M, Chakrabarty P, Johnson JB (2017) Phylogeography and species delimitation in convict cichlids (Cichlidae: *Amatitlania*): implications for taxonomy and Plio–Pleistocene evolutionary history in Central America. Biological Journal of the Linnean Society. Linnean Society of London 120(1): 155–170. https://doi.org/10.1111/bij.12845

Barluenga M, Meyer A (2010) Phylogeography, colonization and population history of the Midas cichlid species complex (*Amphilophus* spp.) in the Nicaraguan crater lakes. BMC Evolutionary Biology 10(1): 326. https://doi.org/10.1186/1471-2148-10-326

Baumgartner J, Wilson P, Esperon-Rodriguez M (2017) Rmaxent: tools for working with Maxent in R. R package version 0.4.1.9000. https://rdrr.io/github/johnbaums/rmaxent/f/README.md

Beerli P (1998) Estimation of migration rates and population sizes in geographically structured populations. Pp. 39–54. In: Carvalho G (Ed.) NATO Advanced Science Institutes Series, Series A. Life Sciences 306: 39–54.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proceedings of the National Academy of Sciences of the United States of America 98(8): 4563–4568. https://doi.org/10.1073/pnas.081068098

Beerli P, Mashayekhi S, Sadeghi M, Khodaei M, Shaw K (2019) Population genetic inference with MIGRATE. Current Protocols in Bioinformatics 68(1): e87. https://doi.org/10.1002/cpbi.87

Beltrán-López RG, Domínguez-Domínguez O, Pérez-Rodríguez R, Piller K, Doadrio I (2018) Evolving in the highlands: the case of the Neotropical Lerma live-bearing *Poeciliopsis infans* (Woolman, 1894) (Cyprinodontiformes: Poeciliidae) in Central Mexico. BMC Evolutionary Biology 18(1): 56. https://doi.org/10.1186/s12862-018-1172-7

Bermingham E, Martin AP (1998) Comparative mtDNA phylogeography of neotropical freshwater fishes: Testing shared history to infer the evolutionary landscape of lower Central America. Molecular Ecology 7(4): 499–517. https://doi.org/10.1046/j.1365-294x.1998.00358.x

Bloom DD, Piller KR, Lyons J, Mercado-Silva N, Medina-Nava M (2009) Systematics and biogeography of the silverside tribe Menidiini (Teleostomi: Atherinopsidae) based on the mitochondrial ND2 gene. Copeia 2009(2): 408–417. https://doi.org/10.1643/CI-07-151

Budaev SV, Zworykin DD, Mochek AD (1999) Individual differences in parental care and behaviour profile in the convict cichlid: A correlation study. Animal Behaviour 58(1): 195–202. https://doi.org/10.1006/anbe.1999.1124

Caballero M, Lozano-García S, Vázquez-Selem L, Ortega B (2010) Evidencias de cambio climático y ambiental en registros glaciales y en cuencas lacustres del centro de México durante el último máximo glacial. Boletín de la Sociedad Geológica Mexicana 62(3): 359–377. https://doi.org/10.18268/BSGM2010v62n3a4

Chen KC, Borowsky RL (2004) Comparative phylogeography of *Xiphophorus variatus* and *Heterandria jonesi* (Poeciliidae) using RAPD data. Ichthyological Exploration of Freshwaters 15(1): 25–40.

Cobos ME, Peterson AT, Barve N, Osorio-Olvera L (2019) kuenm: An R package for detailed development of ecological niche models using Maxent. PeerJ 7: e6281. https://doi.org/10.7717/peerj.6281

Coghill LM, Hulsey CD, Chaves-Campos J, de Leon FJG, Johnson SG (2014) Next generation phylogeography of cave and surface *Astyanax mexicanus*. Molecular Phylogenetics and Evolution 79: 368–374. https://doi.org/10.1016/j.ympev.2014.06.029

Comte L, Grenouillet G (2015) Distribution shifts of freshwater fish under a variable climate: Comparing climatic, bioclimatic and biotic velocities. Diversity and Distributions 21(9): 1014–1026. https://doi.org/10.1111/ddi.12346

Cordellier M, Pfenninger M (2009) Inferring the past to predict the future: Climate modelling predictions and phylogeography for the freshwater gastropod *Radix balthica* (Pulmonata, Basommatophora). Molecular Ecology 18(3): 534–544. https://doi.org/10.1111/j.1365-294X.2008.04042.x

Cornuet JM, Ravigné V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1. 0). BMC Bioinformatics 11(1): 401. https://doi.org/10.1186/1471-2105-11-401

Craul M, Chikhi L, Sousa V, Olivieri GL, Rabesandratana A, Zimmermann E, Radespiel U (2009) Influence of forest fragmentation on an endangered large-bodied lemur in northwestern Madagascar. Biological Conservation 142(12): 2862–2871. https://doi.org/10.1016/j.biocon.2009.05.026

De la Maza-Benignos M, Ornelas-García CP, Lozano-Vilano MDL, García-Ramírez ME, Doadrio I (2015). Phylogeographic analysis of genus *Herichthys* (Perciformes: Cichlidae), with descriptions of *Nosferatu* new genus and *H. tepehua* n. sp. Hydrobiologia 748: 201–231. https://doi.org/10.1007/s10750-014-1891-8

Domínguez-Domínguez O, Pérez-Ponce de León G (2009) ¿La mesa central de México es una provincia biogeográfica? Análisis descriptivo basado en componentes bióticos dulceacuícolas. Revista Mexicana de Biodiversidad 80(3): 835–852. https://doi.org/10.22201/ib.20078706e.2009.003.178

Domisch S, Jaehnig SC, Haase P (2011) Climate-change winners and losers: Stream macroinvertebrates of a submontane region in Central Europe. Freshwater Biology 56(10): 2009–2020. https://doi.org/10.1111/j.1365-2427.2011.02631.x

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7(1): 214. https://doi.org/10.1186/1471-2148-7-214

Drummond AJ, Rambaut A, Shapiro BETH, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular Biology and Evolution 22(5): 1185–1192. https://doi.org/10.1093/molbev/msi103

Duftner N, Sefc KM, Koblmueller S, Nevado B, Verheyen E, Phiri H, Sturmbauer C (2006) Distinct population structure in a phenotypically homogeneous rock-dwelling cichlid fish from Lake Tanganyika. Molecular Ecology 15(9): 2381–2395. https://doi.org/10.1111/j.1365-294X.2006.02949.x

Elías DJ, McMahan CD, Matamoros WA, Gómez-González AE, Piller KR, Chakrabarty P (2020) Scale (s) matter: Deconstructing an area of endemism for Middle American freshwater fishes. Journal of Biogeography 47(11): 2483–2501. https://doi.org/10.1111/jbi.13941

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17(1): 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10(3): 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

FAO (2022) Hydrological basins of the world. https://data.apps.fao.org/catalog/dataset/f2615a41-6383-4aa4-aa21-743330eb03ae

Ferreira DG, Galindo BA, Frantine-Silva W, Almeida FS, Sofia SH (2015) Genetic structure of a Neotropical sedentary fish revealed by AFLP, microsatellite and mtDNA markers: A case study. Conservation Genetics 16(1): 151–166. https://doi.org/10.1007/s10592-014-0648-2

Fricke. R, Eschmeyer WN, van der Laan R (Eds.) (2023) Eschmeyer's catalog of fishes: genera, species, references. California Academy of Sciences, San Francisco, CA, USA. [Electronic version accessed 13 November 2023] http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp

Froese R, Pauly D (Eds.) (2023) FishBase. [Version 10/2023] http://www.fishbase.org

García-Andrade AB, Beltrán-Lopez RG, Pérez-Rodríguez R, Domínguez-Domínguez O, Mejía-Mojica H, Doadrio I (2021) Evolutionary history of the Aztec shiner Aztecula sallaei (Günther, 1868) (Teleostei: Cyprinidae): An endemic and monotypic species of Mexico. Journal of Zoological Systematics and Evolutionary Research 59(8): 2103–2118. https://doi.org/10.1111/jzs.12555

García-Martínez RM, Barriga-Sosa I, López B, Mejía O (2020) Colonization routes and demographic history of Chirostoma humboldtianum in the central Mexican plateau. Journal of Fish Biology 97(4): 1039–1050. https://doi.org/10.1111/jfb.14462

Gavrilets S (2003) Perspective: models of speciation: what have we learned in 40 years? Evolution; International Journal of Organic Evolution 57(10): 2197–2215. https://doi.org/10.1111/j.0014-3820.2003.tb00233.x

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multi-platform graphical user interface for sequence alignment and phylogenetic tree building. Molecular Biology and Evolution 27(2): 221–224. https://doi.org/10.1093/molbev/msp259

Gutiérrez-Rodríguez C, Morris MR, Dubois NS, de Queiroz K (2007) Genetic variation and phylogeography of the swordtail fish Xiphophorus cortezi (Cyprinodontiformes, Poeciliidae). Molecular Phylogenetics and Evolution 43(1): 111–123. https://doi.org/10.1016/j.ympev.2006.10.022

Hausdorf B, Wilkens H, Strecker U (2011) Population genetic patterns revealed by microsatellite data challenge the mitochondrial DNA based taxonomy of Astyanax in Mexico (Characidae, Teleostei). Molecular Phylogenetics and Evolution 60(1): 89–97. https://doi.org/10.1016/j.ympev.2011.03.009

Heller R, Chikhi L, Siegismund HR (2013) The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. PLoS One 8(5): e62992. https://doi.org/10.1371/journal.pone.0062992

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnean Society. Linnean Society of London 58(3): 247–276. https://doi.org/10.1006/bijl.1996.0035

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. Nature 405(6789): 907–913. https://doi.org/10.1038/35016000

Hewitt GM (2003) Ice ages: their impact on species distributions and evolution. Pp. 339–361. In: Rothschild LJ, Lister AM (Eds.) Evolution on planet Earth, 1st edn. Academic Press, New York, NY, USA. https://doi.org/10.1016/B978-012598655-7/50045-8

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 359(1442): 183–195. https://doi.org/10.1098/rstb.2003.1388

Hijmans RJ, Van Etten J (2012) raster: Geographic analysis and modeling with raster data. R package version 2.0-12. http://CRAN.R-project.org/package=raster

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25(15): 1965–1978. https://doi.org/10.1002/joc.1276

Hijmans RJ, Phillips S, Leathwick J, Elith J (2020) dismo: Species Distribution Modeling. R package version 1.3-3. http://cran.r-project.org/web/packages/dismo/index.html

Lambeck K, Chappell J (2001) Sea level change through the last glacial cycle. Science 292(5517): 679–686. https://doi.org/10.1126/science.1059549

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics (Oxford, England) 23(21): 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Lehner B, Grill G (2013) Global River hydrography and network routing: Baseline data and new approaches to study the world's large river systems. Hydrological Processes 27(15): 2171–2186. https://doi.org/10.1002/hyp.9740

Lehner B, Verdin K, Jarvis A (2008) New global hydrography derived from spaceborne elevation data. Eos 89(10): 93–94. https://doi.org/10.1029/2008EO100001

León-Romero Y, Mejía O, Soto-Galera E (2012) DNA barcoding reveals taxonomic conflicts in the *Herichthys bartoni* species group (Pisces: Cichlidae). Molecular Ecology Resources 12(6): 1021–1026. https://doi.org/10.1111/1755-0998.12018

Lister AM, Sher AV, van Essen H, Wei G (2005) The pattern and process of mammoth evolution in Eurasia. Quaternary International 126: 49–64. https://doi.org/10.1016/j.quaint.2004.04.014

Marshall CJ, Liebherr JK (2000) Cladistic biogeography of the Mexican transition zone. Journal of Biogeography 27(1): 203–216. https://doi.org/10.1046/j.1365-2699.2000.00388.x

Mateos M (2005) Comparative phylogeography of livebearing fishes in the genera *Poeciliopsis* and *Poecilia* (Poeciliidae: Cyprinodontiformes) in central Mexico. Journal of Biogeography 32(5): 775–780. https://doi.org/10.1111/j.1365-2699.2005.01236.x

Mateos M, Sanjur OI, Vrijenhoek RC (2002) Historical biogeography of the livebearing fish genus *Poeciliopsis* (Poeciliidae: Cyprinodontiformes). Evolution; International Journal of Organic Evolution 56(5): 972–984. https://doi.org/10.1111/j.0014-3820.2002.tb01409.x

Matsumoto CK, Hilsdorf AWS (2009) Microsatellite variation and population genetic structure of a neotropical endangered Bryconinae species *Brycon insignis* Steindachner, 1877: Implications for its conservation and sustainable management. Neotropical Ichthyology 7(3): 395–402. https://doi.org/10.1590/S1679-62252009000300006

McMahan CD, Ginger L, Cage M, David KT, Chakrabarty P, Johnston M, Matamoros WA (2017) Pleistocene to Holocene expansion of the black-belt cichlid in Central America, *Vieja maculicauda* (Teleostei: Cichlidae). PLoS One 12(5): e0178439. https://doi.org/10.1371/journal.pone.0178439

Mejía O, Martínez-Méndez N, Pérez-Miranda F, Matamoros WA (2022) Climatic niche evolution of a widely distributed Neotropical freshwater fish clade. Biological Journal of the Linnean Society. Linnean Society of London 135(4): 839–855. https://doi.org/10.1093/biolinnean/blab153

Miller RR, Minckley WL, Norris SM (2005) Freshwater fishes of Mexico. University of Chicago Press, Chicago, IL, USA, 490 pp.

Morrone JJ (2004) Panbiogeografía, componentes bióticos y zonas de transición. Revista Brasileira de Entomologia 48(2): 149–162. https://doi.org/10.1590/S0085-56262004000200001

Morrone JJ (2007) Hacia una biogeografía evolutiva. Revista Chilena de Historia Natural 80(4): 509–520. https://doi.org/10.4067/S0716-078X2007000400011

Morrone JJ (2020) Biotic assembly in evolutionary biogeography: A case for integrative pluralism. Frontiers of Biogeography 12(4): e48819. https://doi.org/10.21425/F5FBG48819

Muscarella R, Galante PJ, Soley-Guardia M, Boria RA, Kass JM, Uriarte M, Anderson RP (2014) ENM eval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods in Ecology and Evolution 5(11): 1198–1205. https://doi.org/10.1111/2041-210X.12261

Naimi B, Hamm NA, Groen TA, Skidmore AK, Toxopeus AG (2014) Where is positional uncertainty a problem for species distribution modelling? Ecography 37(2): 191–203. https://doi.org/10.1111/j.1600-0587.2013.00205.x

Olivieri GL, Sousa V, Chikhi L, Radespiel U (2008) From genetic diversity and structure to conservation: Genetic signature of recent population declines in three mouse lemur species (*Microcebus* spp.). Biological Conservation 141(5): 1257–1271. https://doi.org/10.1016/j.biocon.2008.02.025

Ornelas-García CP, Domínguez-Domínguez O, Doadrio I (2008) Evolutionary history of the fish genus *Astyanax* Baird & Girard (1854) (Actinopterygii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. BMC Evolutionary Biology 8(1): 340. https://doi.org/10.1186/1471-2148-8-340

Pacioni C, Hunt H, Allentoft ME, Vaughan TG, Wayne AF, Baynes A, Haouchar D, Dorthc J, Bunce M (2015) Genetic diversity loss in a biodiversity hotspot: Ancient DNA quantifies genetic decline and former connectivity in a critically endangered marsupial. Molecular Ecology 24(23): 5813–5828. https://doi.org/10.1111/mec.13430

Pereira LHG, Foresti F, Oliveira C (2009) Genetic structure of the migratory catfish *Pseudoplatystoma corruscans* (Siluriformes: Pimelodidae) suggests homing behaviour. Ecology Freshwater Fish 18(2): 215–225. https://doi.org/10.1111/j.1600-0633.2008.00338.x

Pérez-Miranda F, Mejía O, Soto-Galera E, Espinosa-Pérez H, Piálek L, Říčan O (2018) Phylogeny and species diversity of the genus *Herichthys* (Teleostei: Cichlidae). Journal of Zoological Systematics and Evolutionary Research 56(2): 223–247. https://doi.org/10.1111/jzs.12197

Pérez-Miranda F, Mejia O, López B, Říčan O (2020) Molecular clocks, biogeography and species diversity in *Herichthys* with evaluation of the role of Punta del Morro as a vicariant brake along the Mexican Transition Zone in the context of local and global time frame of cichlid diversification. PeerJ 8: e8818. https://doi.org/10.7717/peerj.8818

Peterson AT, Nyari AS (2008) Ecological niche conservatism and Pleistocene refugia in the thrush-like mourner, *Schiffornis* sp., in the neotropics. Evolution 62(1): 173–183. https://doi.org/10.1111/j.1558-5646.2007.00258.x

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecological Modelling 190(3–4): 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips JD, Gillis DJ, Hanner RH (2020) HACSim: An R package to estimate intraspecific sample sizes for genetic diversity assessment using haplotype accumulation curves. PeerJ. Computer Science 6: e243. https://doi.org/10.7717/peerj-cs.243

Pio NL, Carvalho TP (2021) Evidence on the paleodrainage connectivity during Pleistocene: Phylogeography of a hypoptopomatine endemic to southeastern Brazilian coastal drainages. Neotropical Ichthyology 19(2): e200128. https://doi.org/10.1590/1982-0224-2020-0128

QGIS Development Team (2009) QGIS Geographic Information System. Open Source Geospatial Foundation. http://qgis.org

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67(5): 901–904. https://doi.org/10.1093/sysbio/syy032

Říčan O, Piálek L, Zardoya R, Doadrio I, Zrzavý J (2013) Biogeography of the Mesoamerican Cichlidae (Teleostei: Heroini): colonization through the GAARlandia land bridge and early diversification. Journal of Biogeography 40(3): 579–593. https://doi.org/10.1111/jbi.12023

Říčan O, Piálek L, Dragová K, Novák J (2016) Diversity and evolution of the Middle American cichlid fishes (Teleostei: Cichlidae) with revised classification. Vertebrate Zoology 66: 1–102. https://doi.org/10.3897/vz.66.e31534

Rincon-Sandoval M, Betancur-R R, Maldonado-Ocampo JA (2019) Comparative phylogeography of trans-Andean freshwater fishes based on genome-wide nuclear and mitochondrial markers. Molecular Ecology 28(5): 1096–1115. https://doi.org/10.1111/mec.15036

Rocamontes-Morales JA, Gutiérrez-Rodríguez C, Rios-Cardenas O, Hernandez-Romero PC (2021) Genetic and morphological differentiation in the green swordtail fish, *Xiphophorus hellerii*: The influence of geographic and environmental factors. Hydrobiologia 848(19): 4599–4622. https://doi.org/10.1007/s10750-021-04664-0

RStudio Team (2020) RStudio: integrated development for R. Rstudio Team, PBC, Boston, MA. https://www.rstudio.com/

Sefc KM (2011) Mating and parental care in Lake Tanganyika's cichlids. International Journal of Evolutionary Biology 2011: 1–20. https://doi.org/10.4061/2011/470875

Sofia SH, Silva CR, Galindo BA, Almeida FS, Sodre LM, Martinez CB (2006) Population genetic structure of *Astyanax scabripinnis* (Teleostei, Characidae) from an urban stream. Hydrobiologia 553(1): 245–254. https://doi.org/10.1007/s10750-005-1106-4

Sofia SH, Galindo BA, Paula FM, Sodré LM, Martinez CB (2008) Genetic diversity of *Hypostomus ancistroides* (Teleostei, Loricariidae) from an urban stream. Genetics and Molecular Biology 31(1 suppl): 317–323. https://doi.org/10.1590/S1415-47572008000200027

Souza-Shibatta L, Kotelok-Diniz T, Ferreira DG, Shibatta OA, Sofia SH, De Assumpcao L, Suelen F, Pini R, Makrakis S, Makrakis MC (2018) Genetic diversity of the endangered neotropical cichlid fish (*Gymnogeophagus setequedas*) in Brazil. Frontiers in Genetics 9: 13. https://doi.org/10.3389/fgene.2018.00013

Strecker U, Faúndez VH, Wilkens H (2004) Phylogeography of surface and cave *Astyanax* (Teleostei) from Central and North America based on cytochrome b sequence data. Molecular Phylogenetics and Evolution 33(2): 469–481. https://doi.org/10.1016/j.ympev.2004.07.001

Tonkin-Hill G, Lees JA, Bentley SD, Frost SD, Corander J (2019) Fast hierarchical Bayesian analysis of population structure. Nucleic Acids Research 47(11): 5539–5549. https://doi.org/10.1093/nar/gkz361

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD (2005) DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 360(1462): 1847–1857. https://doi.org/10.1098/rstb.2005.1716

Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: The importance of model complexity and the performance of model selection criteria. Ecological Applications 21(2): 335–342. https://doi.org/10.1890/10-1171.1

Waters JM, Wallis GP (2000) Across the southern Alps by river capture? Freshwater fish phylogeography in South Island, New Zealand. Molecular Ecology 9(10): 1577–1582. https://doi.org/10.1046/j.1365-294x.2000.01035.x

Wiens JA, Stralberg D, Jongsomjit D, Howell CA, Snyder MA (2009) Niches, models, and climate change: Assessing the assumptions and uncertainties. Proceedings of the National Academy of Sciences of the United States of America 106(Suppl. 2): 19729–19736. https://doi.org/10.1073/pnas.0901639106

WoRMS Editorial Board (2023). World Register of Marine Species. https://www.marinespecies.org [Accessed 13 November 2023] https://doi.org/10.14284/170

Supplementary material 1

List of specimens

Authors: Fabian Pérez-Miranda, Omar Mejía, Benjamín López, Eduardo Soto-Galera, Amairany Bernal-Portillo, Wilfredo A. Matamoros

Data type: xlsx

Explanation note: List of the specimens analysed in this study for the two mitochondrial markers. The specimens are deposited in the Colección Nacional de Peces Dulceacuícolas de la Escuela Nacional de Ciencias Biológicas.

Link: https://doi.org/10.3897/aiep.53.112183.suppl1

Supplementary material 2

Gene flow models used in this study

Authors: Fabian Pérez-Miranda, Omar Mejía, Benjamín López, Eduardo Soto-Galera, Amairany Bernal-Portillo, Wilfredo A. Matamoros

Data type: xlsx

Explanation note: Gene flow models used in this study for the analysis of mitchondrial markers COI and D-loop in *Herichthys pantostictus* and *Hericthys carpintis*. The selected models according with Bayes Factor (BF) are highlighted in bold. See Suppl. material 3: table S3 for further details.

Link: https://doi.org/10.3897/aiep.53.112183.suppl2

Supplementary material 3

GenBank numbers

Authors: Fabian Pérez-Miranda, Omar Mejía, Benjamín López, Eduardo Soto-Galera, Amairany Bernal-Portillo, Wilfredo A. Matamoros

Data type: xlsx

Explanation note: List of accession numbers in GenBank for the unique haplotypes generated in this study.

Link: https://doi.org/10.3897/aiep.53.112183.suppl3

Supplementary material 4

Accumulation curves

Authors: Fabian Pérez-Miranda, Omar Mejía, Benjamín López, Eduardo Soto-Galera, Amairany Bernal-Portillo, Wilfredo A. Matamoros

Data type: pdf

Explanation note: Haplotype accumulation curves to estimate the fraction of haplotype diversity captured from sample size used in this study. A) Haplotype accumulation curve for the COI marker in *H. pantostictus*. B) Haplotype accumulation curve for the D loop marker in *H. pantostictus*. C) Haplotype accumulation curve for the COI marker in *H. carpintis*. D) Haplotype accumulation curve for the D loop marker in *H. carpintis*. The R value represent the fraction of haplotype diversity recovered by sample size.

Link: https://doi.org/10.3897/aiep.53.112183.suppl4

Supplementary material 5

Bayes factor comparison among gene flow models

Authors: Fabian Pérez-Miranda, Omar Mejía, Benjamín López, Eduardo Soto-Galera, Amairany Bernal-Portillo, Wilfredo A. Matamoros

Data type: xlsx

Explanation note: Model comparison of the different gene flow models analysed in the populations of *Herichthys pantostictus* and *Herichthys carpintis* for the mitochondrial markers COI and D-loop. The models were compared through a Bayes Factor (BF) using the Bezier marginal likelihood obtained in Migrate. See Suppl. material 2: table S2 for complete details of models. lML (Bezier likelihood), LBF (Likelihood Bayes Factor), mod. rank (model ranking), mod.prob (model probability).

Link: https://doi.org/10.3897/aiep.53.112183.suppl5

Supplementary material 6

Competing scenarios

Authors: Fabian Pérez-Miranda, Omar Mejía, Benjamín López, Eduardo Soto-Galera, Amairany Bernal-Portillo, Wilfredo A. Matamoros

Data type: pdf

Explanation note: Schematic representation of the different competing scenarios considering to test the colonization routes of *H. carpintis* and *H. pantostictus* by Approximate Bayesian Computation (ABC) of the mitochondrial DNA markers COI and D-loop. In the first scenario a costal to inland colonization process was inferred, on the other hand, in the second scenario, a inland to coast colonization process was assumed.

Link: https://doi.org/10.3897/aiep.53.112183.suppl6