Phylogeography of the Mayan cichlid *Mayaheros urophthalmus* (Teleostei: Cichlidae) in the Yucatan peninsula based on mitochondrial markers CYTB and COI

J. Barrientos-Villalobos · **J. J. Schmitter-Soto**

Abstract The Yucatan Peninsula (YP) suffered several marine transgressions and regressions during the Quaternary, thus molding the distribution of its present biota, especially its freshwater fish fauna. The Mayan cichlid (*Mayaheros urophthalmus* Günther) is a euryhaline fish native to the Atlantic slope of Mexico and northern Central America, including the YP; it is one of the most widespread freshwater species in the region. Herein we discuss a phylogeographic scenario by which the Mayan cichlid may have reached its current distribution in the YP. A Bayesian analysis and minimum spanning network were inferred from two partial mitochondrial genes, Cytochrome b (CYTB) and Cytochrome c Oxidase I (COI). The two fragments showed genetic differentiation among populations ($Fst = 0.31$, p value <0.001). Tajima's D and Fu´s F revealed a tendency to the expansion of some populations. A consistent ordination of north vs south populations was observed. A spatial analysis of molecular variance (SAMOVA) was performed to recognize putative barriers among populations of *M. urophthalmus*. A secondary molecular calibration located the window time in which the dispersal event may have occurred during the Pleistocene, around 1 Mya. We determined that a Quaternary dispersal around the old coastlines from the south explains the current distribution of the Mayan cichlid.

Introduction

The Quaternary period (starting 2.58 Mya) began with a major extensive glaciation and was marked by periods of large eustatic changes and alternation of about 35 glacial and interglacial periods since 1.8 Mya, with many species migrating to refuges, especially towards the maximum cold at the Late Pleistocene, a progression punctuated by brief warm interglacial periods (Ogg et al. 2016). These geological and climatic changes left their imprint on the colonization history and current geographical distribution of freshwater fish species (Wilkens 1982).

The karstic nature of YP has impeded the formation and existence of surface rivers, but it has favored the formation of cenotes (sinkholes). The most isolated, inland cenotes, have been colonized by only a few fish species (Wilkens 1982).

Two alternative mechanisms (vicariance and dispersal) have been proposed to explain the current and past distribution of species (Matamoros et al. 2014), but the latter is likely to be the principal mechanism by which the cichlids have colonized YP from South America, because the northern part of YP emerged only around 1.8 Mya, whereas the southern half was already available to colonization 5.5 Mya (Fig. 1; Bautista et al. 2011). Three waves of dispersal across the isthmian bridge of Panama during the Quaternary have been hypothesized to explain

J. Barrientos-Villalobos (✉) · J. J. Schmitter-Soto
El Colegio de la Frontera Sur, A.P. 424, 77000 Chetumal, Mexico
e-mail: celacanto50@hotmail.com

contemporary primary fish distribution in Central America during the Great American Biotic Interchange (Martin and Bermingham 1998; Woodburne 2010). Cichlid fishes colonized Middle America from South America during the Oligocene, 29-24 Mya, in the course of the Greater Antilles–Aves Ridge land bridge (GAARlandia) event (Říčan et al. 2013).

The Mayan Cichlid (*Mayaheros urophthalmus* Günther) is a Neotropical fish very frequent and abundant in brackish and freshwater environments, but also tolerant of marine salinity (Miller et al. 2009), which facilitates its dispersion along the coast. It ranges from southern Veracruz, Mexico, to Nicaragua, inhabiting rivers, lakes, ponds, and karstic sinkholes (cenotes) in the YP, including coastal lagoons in the islands of Contoy and Mujeres. It has been introduced to southern Florida (Harrison et al. 2014) and southeastern Asia (Nico et al. 2007). The species is often consumed locally and even cultured as an alternative to tilapia (Martínez-Palacios and Ross 1988). Although its commercial importance is modest, as a predator (an omnivore with carnivorous tendency: Miller et al. 2009), it is ecologically very relevant in most water bodies of the region, especially in the faunistically impoverished inland cenotes of the YP (Schmitter-Soto et al. 2002).

The species displays such a variety of body shape and coloration along its distribution range in YP, that Hubbs (1936) was able to describe eight subspecies. Barrientos-Medina (2005) provided partial diagnoses for them based solely on morphometrics and proposed raising these taxa to species level; however, the validity of these has been questioned (Barrientos-Villalobos et al. 2018), the phenotypic variability being ascribed to the type of aquatic environment in which the fish live, i.e., river, cenote or lagoon; this conclusion was based on a molecular analysis of two concatenated fragments of mitochondrial DNA (mtDNA), cytochrome b (CYTB) and cytochrome oxidase subunit I (COI).

Although Barrientos-Villalobos et al. (2018) concluded that "the patterns of the body shape of *M. urophthalmus* are more consistent with ecophenotypic variation than with genetic differentiation due to geographic isolation by distance", they also found low genetic divergence among populations (and elevated gene flow, which has diluted any speciation process). In another analysis, based on cytb as well, but also on nuclear microsatellites, Harrison et al. (2014) found genetic subdivision in YP and explained it through "pre-Columbian human transportation,"

because their results showed that northern YP was more similar to the Guatemalan Petén than to southern YP. Part of the same dataset was interpreted earlier by Razo-Mendívil et al. (2013) as having "two groups with low divergence and with no correspondence with geographical distribution," concluding, nevertheless, that "historical long-distance dispersal and drought periods during the Holocene, with significant population size reductions and fragmentations, are factors that could have shaped the genetic structure of the Mayan cichlid."

Following the geological reconstruction of the past coastline of YP by Urrutia-Fucugauchi et al. (2008, Fig. 1) and the geological timing inferred for the YP by Bautista et al. (2011), herein we reanalyze the data from Barrientos-Villalobos et al. (2018) in order to explore a scenario of colonization of the YP by the Mayan Cichlid by means of a phylogeographic analysis. Our hypothesis (argued previously by Schmitter-Soto 2002, among others) was that the genetic structure of the Mayan Cichlid would reflect the north-south pattern and would be explainable by marine transgressions and regressions from Pleistocene, alternatively allowing and hindering colonization of freshwaters in YP.

Methods

To test our hypothesis, we based the phylogeographic analyses on the genetic descriptors, AMOVA test and Bayesian phylogeny described by Barrientos-Villalobos et al. (2018), who used two partial coding genes of mitochondrial DNA, CYTB and COI. Because cytb contains regions with evolutionary rates both slow and fast as well as more conserved regions, it is used at very different levels, from population to deep phylogeny (Farias et al. 2001), whereas COI is much more conserved, so it has been employed in the barcoding approach (Ward et al. 2009); both genes are widely used in phylogeographic scenarios to answer how species reached their present genetic and geographic configuration (Avise 2000; Lovejoy and de Araújo 2000).

Mayan cichlids were surveyed at 15 sites within their distribution range (Fig. 1, Appendix 1) in the Mexican and Belizean parts of the YP. The sampling size varied from nine individuals at Cenote Popolvuh and Ría Celestún to just one individual at Silvituc lake, where the species is scarce, probably because of the introduction of the alien tilapia *Oreochromis niloticus* (Linnaeus). Dorsal fins were clipped from fish caught
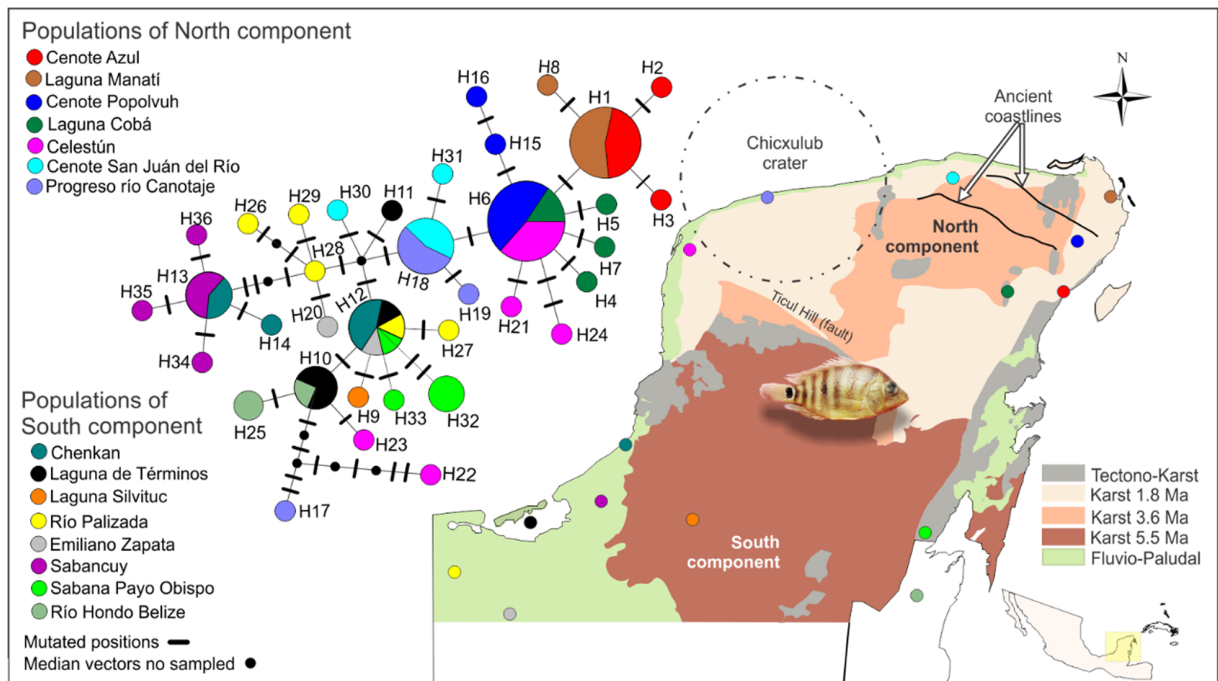
Fig. 1 Tissue-sampled localities of *M. urophthalmus* for the phylogeographic study, and geologic time, Ticul hill and the ancient coastlines of the Peninsula de Yucatan. The coastlines are based on Urrutia-Fucugauchi et al. (2008), and the geological times for the YP at (Bautista et al. 2011) coupled with topology of haplotypic relationships inferred at Network and Statistical Parsimony based on two mitochondrial protein gene fragments (CYTB and COI, 1677 bp) and showing the North and south components. The haplotype network figure is adapted and modified by Barrientos-Villalobos et al. (2018). Circles represent the frequency of each haplotype and the color denotes their locality

using cast nets and double-cone traps and stored in 90% ethanol for transport to the laboratory. A subsample of the captured fishes was vouchered at the ichthyological collection of El Colegio de la Frontera Sur, Chetumal, Mexico (ECO-CH) (Appendix 1).

Molecular data

For total genomic DNA extraction, a small amount of tissue of dorsal fin (i.e. 2 mm^3 approximately) was ground in DNeasy Blood & Tissue Kit (QIAGEN, Cat No. /ID: 69504) following the protocol of the manufacturers. The CYTB mtDNA oligonucleotides used for the amplification were described by Irwin et al. (1991); Farias et al. (2001) and Chakrabarty (2006) described the primers of COI.

DNA amplification was conducted in Peltier-effect thermocyclers (MultiGene OptiMax) under the following parameters: one initial cycle at 95 °C for 120 s, followed by 35 cycles of 95 °C for 20 s, 50 °C for 20 s, 72 °C for 60 s, with one final cycle at 72 °C for 240 s. All PCR reactions were conducted along with positive and negative controls.

The amplifications of PCR were visualized on 2% agarose gels stained with ethidium bromide and successful amplifications were purified using Wizard® SV Gel and PCR Clean-Up System (Promega). Purified products of PCR were sent for sequencing in both directions to South Korea (Macrogen). Sequence files were analyzed with the aid of the program BioEdit Sequence Alignment v7.0.9 (Hall 1999). All sequences were deposited in GenBank under the next accession: MF741939 to MF741974, and MF776666 to MF776701.

Molecular analyses

The best-fitted model of molecular evolution under the Akaike information criterion (AIC) for our concatenated sequences data (CYTB and COI) were determined with the aid of software jModeltest v.2.1.10 (Guindon and Gascuel 2003; Darriba et al. 2012). A Bayesian analysis with the software MrBayes 3.2 (Ronquist and

Huelsenbeck 2003; Ronquist et al. 2012) were executed for inferred the evolutionary relationships among haplotypes, under the next parameters: Lset Nst = 6, revmatpr = (0.3049, 7.4737, 1.0000, 0.3049, 2.5344, and 1), statefreqpr = (0.2395, 0.3215, 0.1525, and 0.2866), pinvarpr = (0.9620), as suggested by jModeltest. Ten million generations searches were implemented with four parallel chains and sampling every 1000th generation, after a 25% burn-in value; finally, to estimate posterior probabilities for each node a majority-rule consensus tree was implemented. To root the phylogeny, under the outgroup criterion, the species *Petenia splendida* (Günther) and *Darienheros calobrensis* (Meek & Hildebrand) were used (with the next accession numbers from GenBank: AF370679, EU751899, AY843381, GU817255). In addition, the evolutionary relationships among haplotypes of *M. urophthalmus* were not constrained as monophyletic.

A phylogenetic minimum spanning network based on statistical parsimony was inferred for the haplotypes (Templeton et al. 1992). This unrooted network was recovered using the program TCS v. 1.21 with a 95% connection probability limit, treating gaps as missing data (Clement et al. 2000). The loops were resolved according to coalescent theory criteria proposed by Pfenninger et al. (2002). Alternatively, a median-joining (MJ) network algorithm based on genetic distances was calculated with an epsilon value = 10 using software NETWORK v.5.0.0 (Bandelt et al. 1999); The maximum parsimony criterion (MP) was implemented to purge superfluous links and median vectors from the network as a post-process (Polzin and Daneshmand 2003).

To detect probable genetic barriers on geographical space, a spatial analysis of molecular variance (SAMOVA) by means of the Monmonier algorithm was implemented in software SAMOVA v.2.0 (Monmonier 1973; Stenico et al. 1998; Simoni et al. 1999). By means of this analysis, geographical and genetically homogeneous groups of populations were defined through the maximization of total genetic variance due to differences between populations of F_{ST} genetic distances. A meta-heuristic search of simulated annealing procedure to find a global optimum configuration was implemented, with no groups defined a priori. The transitions, transversions, and deletion weights were = 1, letting the software identify the best genetic structure without geographical constraint, with 10,000 steps in each process. In the end, a new AMOVA

based on the best genetic-geographical configuration determined by the analysis of SAMOVA software was recalculated, based on pairwise differences of F_{ST} with 1000 permutations on the new configuration determined by the spatial analysis.

Finally, the dating of putative barriers was performed by means of a secondary Bayesian relaxed lognormal clock under an uncorrelated model, which does not require rates to be heritable and, therefore, allows lineage-specific rate heterogeneity (Thorne et al. 1998; Aris-Brosou and Yang 2002). Two independent Montecarlo Markovian chains (MCMC) analyses of 40 million of generations each, sampling every 1000th generation with BEAST-BEAUti v1.8.2 (Drummond et al. 2012), were followed by visualization with Tracer v1.6 (Rambaut et al. 2014), the statistics and convergence of model parameter values, such as Effective Sample Size (ESS), and estimates of optimal posterior distributions of node heights.

For secondary dating, we used two points of calibrations based on estimates from Říčan et al. (2013) used as priors under normal distributions as recommended by Ho and Phillips (2009): 1. - the divergence time between *D. calobrensis* and *P. splendida* (mean 12.9 Mya, s.d. 1.5), and 2. - the divergence time between *M. urophthalmus* and *P. splendida* (mean 6.6 Mya. s.d. 1.5). The trees and log files from the runs were combined using the software LogCombiner v. 1.8.2 (Rambaut et al. 2014). A maximum-credibility tree representing the maximum a posteriori topology with mean heights and 95% higher posterior densities (HPD) of age estimates were obtained from the overall outputs after removing burn-in trees (10%) using the software TreeAnnotator v. 1.8.2, and the final topology was visualized and edited by means of package FigTree v. 1.4.3 (Rambaut 2016).

Results

The concatenated matrix included 1677 aligned positions from 81 individuals from 15 natural populations of *M. urophthalmus* (Fig. 1, Appendix 1); 1053 pb corresponded to CYTB and 618 bp to COI. The posterior Basic Local Alignment Search Tool (BLAST) inspection corroborated that the fragments of both genes employed in this study correspond to the species. The best-fit model selected by the AIC criterion as prior for

the execution of the analyses was TIM3 + I p-inv = 0.9620.

The number of polymorphic sites observed at the concatenated mitochondrial sequences was 31, of which 24 were transitions and 7 were transversions; all mutations were substitutions, and no indels were observed. A total of 36 haplotypes (H) were recovered for 15 populations, of which 28 were singletons. H12 and H6 were the most widespread and the most common haplotypes respectively (Fig. 1). As commonly observed in fishes, the relative nucleotide composition observed in the CYTB and COI fragments were characterized by an antiguanine bias: cytosine, 33.6%; 30.12%; thymine, 28.3%, 29.33; adenine 24.3%, 22.4%, and guanine, 13.7%, 18%, respectively.

There was an excess of rare mutations, as well as the large negative Fs values in the neutrality tests, that suggest population expansion, in the populations of Cobá ($Fs = -1.937$, $p = 0.011$), Río Palizada ($Fs = -2.517$, $p = 0.017$), and Sabancuy ($Fs = -1.812$, $p = 0.02$). According to the AMOVA, the largest proportion of molecular variance occurred within populations (68.11%), whereas the fixation index F_{ST} was conservatively intermediate to low ($F_{ST} = 0.318$). However, the absolute number of migrants per generation (m) was sufficient for diluting speciation process.

Phylogeographic interrelationships

The Bayesian tree supports the monophyletic relationships among haplotypes of *M. urophthalmus* with high posterior probability, whereas the internodes had low support values, and the topology was poorly resolved. However, a recurrent ordination of populations was observed on the map, corresponding to a north-south pattern reflected at both Network and Bayesian phylogenies (Figs. 1, 2).

The Bayesian topology was also highly consistent with the relationships among haplotype nets inferred by means of statistical parsimony at TCS (data not shown) and the median-joining results. These two networks were identical, with a high proportion of singletons, which means that most coalescent events occurred near the root and few occurred later (Fig. 1). A north-south distribution pattern was revealed in the haplotype net reconstructions and phylogeny, although the Mantel's test between linearized (F_{ST}) values and geographic distances was not significant, discarding an effect of

isolation by distance among populations ($r^2 = -0.0028$, $p = 0.523$).

The spatial analyses with no groups defined a priori suggested the existence of two population groups (K = 2), and a putative geographic barrier between these groups was suggested by SAMOVA with highest F_{CT} value (0.43, $p = 0.0001$). This pattern represents the better genetic-geographic configuration, which is in accord with the Bayesian and network analyses. The groups corresponds to northern (1, 2, 3, 7, 8, 10, and 13) vs. southern (4, 5, 6, 9, 11, 12, 14, and 15) populations (Figs. 1, 2).

The a posteriori hierarchical AMOVA on the new configuration determined by SAMOVA revealed that 43.8% of variance was due to differences between these two groups (the north populations versus south populations), 24.7% was due to differences among populations within groups, and 31.4% was due to variance within populations. However, when K is incremented a priori to constrain to 3, 4, and 5 groups, these two groups are disgregated and the new groups of subpopulations identified by SAMOVA do not clearly any geographic barrier, while the fixation index Fsc and F_{ST} values decline, and Fct values increase. So the recognition by SAMOVA of two groups (the north populations versus south populations) and a barrier that separates them represents the statistically most robust configuration.

Dating the incipient north-south pattern

We attempted to date this incipient but highly consistent pattern recognized both by the Bayesian and network analyses, as well as the SAMOVA suggestion of a putative barrier among two groups of populations (North-South distribution pattern).

A coalescent growth population size was considered because the general pattern suggested by Fu's test for *M. urophthalmus* was the demographic expansion (Barrientos-Villalobos et al. 2018). The other priors were: a general time reversible (GTR) model, Lset Nst = 5, revmatpr = fixed (0.3049, 7.4737, 0.3049, 2.5344, 10,000), statefreqpr = fixed (0.2395, 0.3215, 0.1525, 0.2866), and pinvarpr = fixed (0.9620), as suggested by jModeltest. The MCMC analyses resulted in acceptable mixing of the statistics and convergence of parameter values.

Our dating of the incipient divergence among north-south components corresponds to the Pleistocene (mean 1 Mya), with high support value (0.9 posterior Bayesian
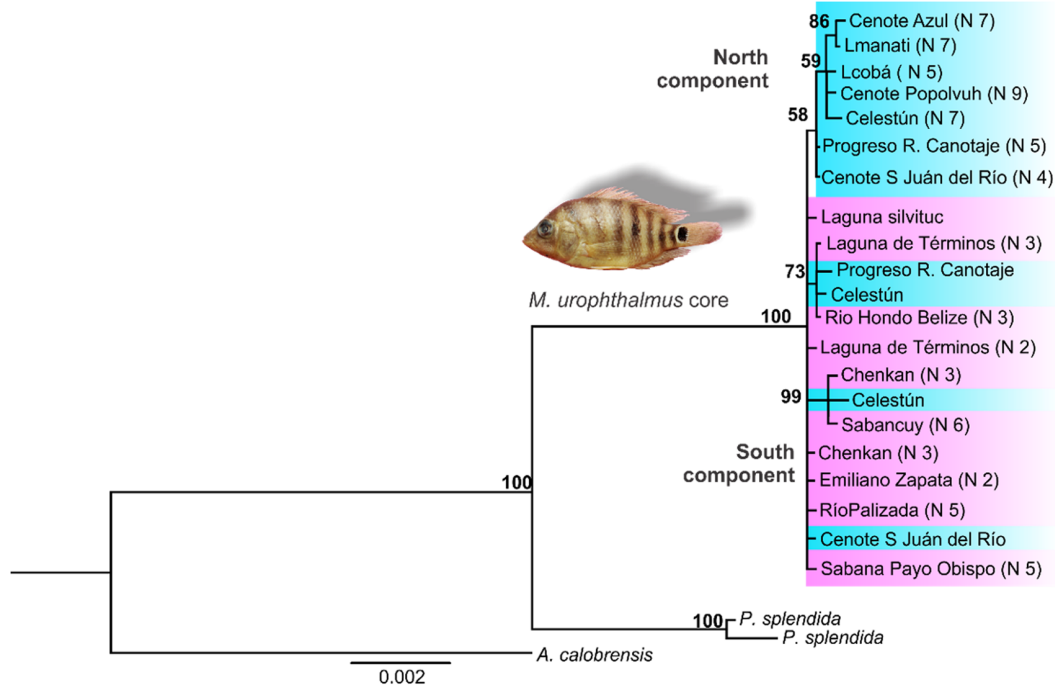
Fig. 2 The Bayesian 50% majority rule tree of *M. urophthalmus* core based on two mitochondrial protein gene fragments (CYTB and COI, 1677 bp). Bayesian Posterior Probabilities (PPB) support at the base of internodes. The figure is based and modified by Barrientos-Villalobos et al. (2018). The principal nodes were collapsed for more clear visualization. The N corresponds to the number of organisms that share the same haplotype

probability), and 95% HPD interval 1.7-0.063 Mya (Fig. 3). This places the colonization event of the north of the YP in the window of time corresponding to the Quaternary, when the last eustatic changes of the Caribbean Sea occurred in consonance with glaciations alternated with warm periods. In addition, the emergence of the north of YP has been dated around 1.8 Mya, leaving as the only possibility of colonization of the north of YP the dispersal of organisms from the south of the peninsula during the last eustatic changes.

Discussion

The genetic flow and number of migrants per generation are congruent with the haplotypic diversity and effective population size of females observed on lentic aquatic systems relative to others for these molecular markers referred by Barrientos-Villalobos et al. (2018, Table 1). On the other hand, the high to moderate gene flow among most of the other populations of *M. urophthalmu*s and the genetic differentiation (F_{ST} = 0.31) observed by us is similar to the 0.4% (in CYTB)

reported by Razo-Mendivil et al. (2013). Furthermore, the observed pattern of the network, with many single-tons, may be indicative of a population that recently began to increase from a few founders or a bottleneck previously to populational expansion (Slatkin and Hudson 1991); this star-like pattern was found as well by Razo-Mendívil et al. (2013).

A biogeographic scenario for the colonization of the YP

Two alternative mechanisms, vicariance and dispersal, have been proposed to explain biological distributions. The first process takes place when a barrier emerges dividing a species and posteriorly the divergence occurs; during dispersal, a subgroup of individuals of a species colonize new areas or simply their distribution range grows (Brown and Lomolino 1988). The latter mechanism has marked the major biogeographical events at the history of the Central American cichlid fishes, since these migrated from South America during the Oligocene (between 29 and 24 Mya) through the GAARlandia event (Říčan et al. 2013). As few as two to four independent dispersal events to Middle America
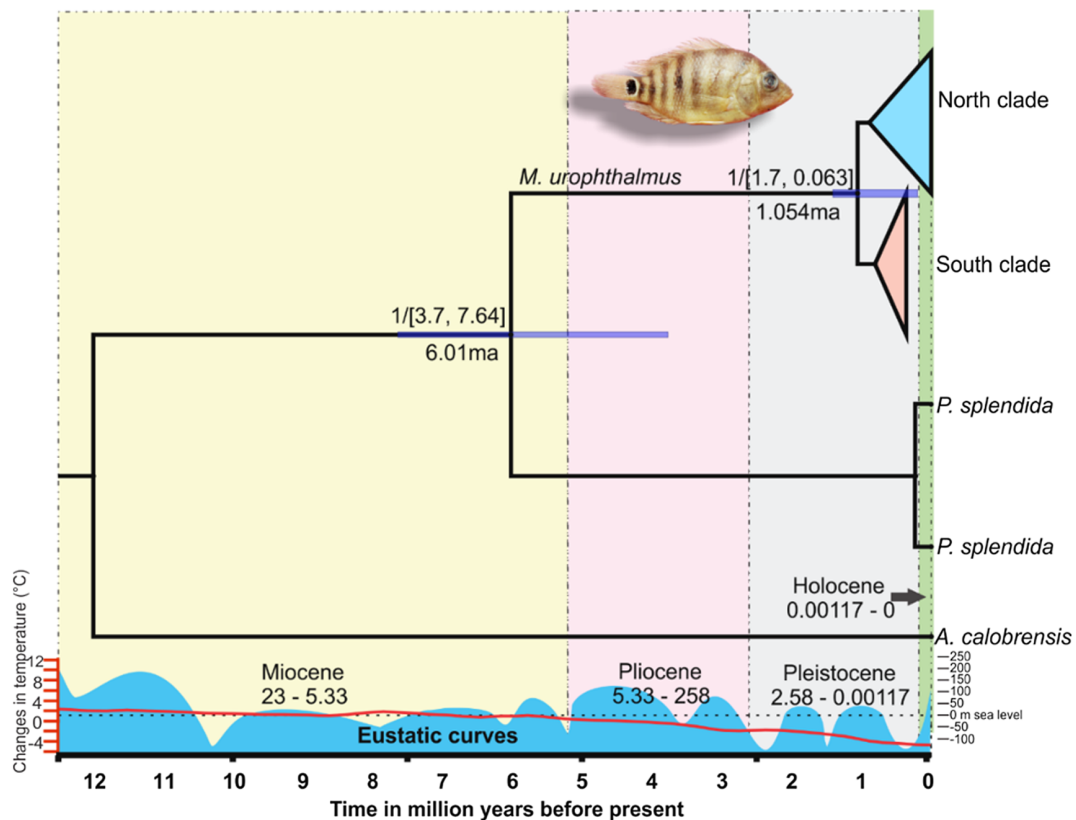
Fig. 3 Relative time of divergence among North and South subclades of *Mayaheros urophthalmus* based on two fragments of mtDNA genes. Blue bars represent the 95% High Probability Densities (HPD) of divergence times. Upper clade: Bayesian Posterior Probabilities / Intervals of divergence times estimates (95% of HPD) in million years before present. Down clade: mean of Posterior Credibility Value of absolute times of divergence. At the bottom of the graph at blue colour the eustatic curves of the sea Based on Haq et al. (1987), and, red line represent the geologic changes of temperature based on Condamine et al. (2016), and, relative times of epochs based on the international chronostratigraphic chart v2018-08 (www.stratigraphy.org)

occurred between 24 and 16 Mya from Central America by secondary and halotolerant fishes, such as poeciliids, rivulids, and cichlids (Concheiro Pérez et al. 2007). They colonized new areas along the coasts relatively easily: *M. urophthalmus* and *Gambusia yucatana* (Regan) are the more euryaceous fishes present in northern and inland YP, and both are derived from the ichthyofauna of coastal lagoons (Wilkens 1982). This hypothesis is supported also by the geologic history of YP: the peninsula emerged gradually at least from the Oligocene onwards, with the youngest part in the north emerged around 1.8 Mya (Fig. 1; Lugo-Hubp et al. 1992; Vázquez-Domínguez and Arita 2010).

Divergence times between *M. urophthalmus* and *P. splendida* have been estimated at around 6.6 Mya by Říčan et al. (2013); our dating of the incipient north-south event of colonization of *M. urophthalmus* (1 Mya, 95% HPD interval 1.7-0.063 Mya) corresponds to the

Quaternary and is in accord to the geologic history of northern YP. Populations of *M. urophthalmus* could have emerged relatively recently from the late Miocene to the Quaternary (3.6 Mya-1.8 kya, Figs. 1, 3; Gondwe et al. 2010), a period that started with acontinental glaciation ~2.58 Mya, albeit the major glacial/interglacial oscillations took place around 1.4 -0.4 Mya, with climatic cycles of ~41 Kyr, expanding to intervals of ~100 Kyr in Middle to Late Pleistocene (Head and Gibbard 2015). This cycle induced sea level fluctuations between 50 m above current sea level and −125 m below (Haq et al. 1987; Rohling et al. 2014), so the shorelines also changed (Urrutia-Fucugauchi et al. 2008), especially in the shallow north of YP (see old coastlines, Fig. 1).

It has long been recognized that the Pliocene and the Pleistocene interval was characterized by an alternation of glacial and interglacial episodes, with several

Table 1 Conventional population pairwise F_{st} values from haplotype frequencies downcast and the absolute number of migrants per generation (m), ∞ (infinite) upper case. Significant differences (*) to level: $p < 0.05$. The table was modified from Barrientos-Villalobos et al. (2018), adding the absolute number of migrants per generation in the upper case of the table

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 Cenote Azul		1.124	∞	0.55	0.754	0.827	0.435	0.635	0.882	0.852	0.665	1.411	0.754	0.754	0.950
2 Laguna Cobá	0.307*		0.623	4.5	2	2.165	5.366	1.434	6.764	∞	2.07	9.5	2	2	2.78
3 Laguna Manatí	−0.043	0.445*		0.2	0.433	0.492	0.278	0.381	0.385	0.547	0.337	0.752	0.433	0.433	0.56
4 Laguna Silvituc	0.476	0.1	0.714		1.166	1.375	0.357	0.75	∞	1.3	1	∞	1.166	1.166	2
5 Laguna Términos	0.398*	0.2	0.535*	0.3		1.966	0.561	0.915	3.653	1.241	3.019	3.863	1.166	1.346	1.522
6 Chenkán	0.376*	0.187*	0.503*	0.266	0.202		0.619	1	∞	1.332	1.216	11.12	1.270	1.966	5.75
7 Cenote Popolvuh	0.534*	0.085	0.642*	0.583	0.471*	0.446*		0.489	0.576	∞	0.477	0.955	0.561	0.561	0.699
8 P. Río Canotaje	0.440*	0.258*	0.567*	0.4	0.353*	0.333*	0.505*		1.2	1	0.829	1.875	∞	0.915	1.166
9 R. Emiliano Zapata	0.361	0.068	0.564	0	0.12	−0.09	0.464	0.294		1.893	1.9	∞	1.862	3.653	3
10 Ría Celestún	0.369*	−0.032	0.477*	0.277	0.287*	0.27*	−0.002	0.331*	0.208		1.19	2.675	1.241	1.241	1.559
11 Río Hondo, Belze	0.428*	0.194	0.597*	0.333	0.142	0.291	0.511*	0.376*	0.207	0.295		3.157	1.096	1.096	1.5
12 Río Palizada	0.261*	0.050	0.399*	0	0.114	0.043	0.343*	0.210	−0.111	0.157	0.136		2.833	3.863	4.285
13 C. S. J. del Río	0.398*	0.200	0.535*	0.3	0.3*	0.282*	0.471*	−0.08	0.211	0.287*	0.313	0.15		1.166	1.522
14 S. Payo Obispo	0.398*	0.200	0.535*	0.3	0.27	0.202	0.471*	0.353*	0.12	0.287*	0.313	0.114	0.3*		1.522
15 Sabancuy	0.344*	0.152	0.471*	0.2	0.247*	0.08	0.416*	0.3*	0.142	0.242*	0.25	0.104	0.247*	0.247*	

droughts during the most intense glacial periods, an enhanced high primary production, sea level lowering, and more vigorous atmospheric circulation (Delmonte et al. 2007). Advances and recessions of glaciers provoked contrasting climatic changes, which contributed to mold the distribution of species (Hulsey and López-Fernández 2011). According to oxygen isotope records, about 35 major glaciations have occurred since 1.8 Mya (Montaggioni and Braithwaite 2009). These changes in sea level, and the karstic nature of the terrain, are responsible for the lack of rivers and the abundance of isolated sinkholes in northern YP (Humphries and Miller 1981; Hulsey and López-Fernández 2011); the open sinkholes were colonized during Pliocene-Pleistocene, probably coupled with marine regressions (Wilkens 1982; Wilkens and Strecker 2017).

The "sierrita" ("small hill range") of Ticul has been considered a limit between north and south YP; this limit corresponds with the ordination of two subgroups north and south recognized by SAMOVA. Northern YP is characterized by a low diversity of freshwater fishes, which increases closer to the coast (Wilkens 1982). Furthermore, fish associations in the south can be explained by ecological factors, like salinity, temperature, and distance from the coastline (Schmitter-Soto and Gamboa-Pérez 1996), whereas in the north associations can be due to both historical and ecological factors (Schmitter-Soto 1999).

Except for some endemisms, the vertebrate fauna of the north of YP is mostly a subset of the fauna at the base of YP, including Petén in Guatemala, suggesting that northern faunas were originated simply by dispersal from the south (Arita 1997; Vázquez-Domínguez and Arita 2010), and this dispersal presumably occurred during Post-Pliocene times, because marine transgressions submerged several times the north of YP (Morrone 2005).

This north-south difference and the decrease in species diversity towards the northern tip of YP, a "peninsula effect" (Simpson 1964), has been invoked to establish two biogeographical (sub) provinces, the Peten province in the south and the Yucatan province in the north (Udvardy 1975; Espadas Manrique et al. 2003).

To complement this biogeographic scenario based on *M. urophthalmus*, additional phylogeographic analyses and divergence dating from other species will be needed to make more conclusive the already quite suggestive patterns in the historical biogeography of the YP and the Caribbean Sea.

Conclusions

Is clear to us that dispersal was the only possible mechanism to explain the current distribution of *M. urophthalmus*. This dispersal could have occurred with displacements around the old coastlines from the south. The phylogeographic analyses, the SAMOVA, and geologic history of YP suggest so; in addition, the dating of the event of dispersal by some individuals from the south of the YP to North of the YP geologically makes sense.

On the other side, the establishment of the populations of *M. urophthalmus* on the northern of YP probably occurred during Post-Pliocene times, because marine transgressions submerged several times the north of YP, preventing this from happening before.

All the above agrees perfectly with the biogeographic history observed in other groups of vertebrates of YP.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

Statement on the welfare of animals This work did not include experiments, but fish were collected and preserved at the collection in ECOSUR. This was done under the proper permits from Mexican environmental authorities (No. PPF/DGOPA-053/15), and all applicable international, national, and institutional guidelines for the use of animals in research were followed. Such ethical standards include a humane euthanization of the fish (using ice) and the capture of just enough specimens to conduct the study.
The work did not include studies with human participants performed by any of the authors.

Appendix 1

Collection sites	Catalogue numbers	N
1. Cenote Azul	ECO-CH 7820	7
2. L. Manati	ECO-CH 7645	7
3. L. Cobá	ECO-CH 7821	5
4. Popolvuh	ECO-CH 7819	9
5. Celestún	ECO-CH 7534	9
6. Cenote San Juan Del Río	ECO-CH- 7531	5
7. Progreso río Canotaje	ECO-CH 7532	6

Collection sites	Catalogue numbers	N
8. Laguna de Términos	ECO-CH 7547	5
9. P. Chenkan	ECO-CH (uncat.)	6
10. Río Palizada	ECO-CH 7549	5
11. Sabana Payo Obispo	ECO-CH 7535	5
12. L. Silvituc	ECO-CH 7546	1
13. Emiliano Zapata	ECO-CH 7548	2
14. Río Hondo, Belize	ECO-CH 7570, 7585(1)	3
15. Sabancuy Pucteito	ECO-CH 7573	6

Acronym: ECO-CH, Colección Ictiológica de El Colegio de la Frontera Sur, Unidad Chetumal. The N corresponds to the size of the sample

References

Aris-Brosou S, Yang Z (2002) Effects of models of rate evolution on estimation of divergence dates with special reference to the metazoan 18S ribosomal RNA phylogeny. Syst Biol 51: 703–714

Arita HT (1997) The non-volant mammal fauna of Mexico: species richness in a megadiverse country. Biodivers Conserv 6: 787–795

Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Barrientos-Medina RC (2005) Estado taxonómico de la mojarra rayada *"Cichlasoma" urophthalmus* Günther, 1862 (Teleostei: Cichlidae). M.Sc. Thesis, El Colegio de la Frontera Sur, Chetumal, Mexico. El Colegio de la Frontera Sur

Barrientos-Villalobos J, Schmitter-Soto JJ, Espinosa de los Monteros A (2018) Several subspecies or phenotypic plasticity? A geometric morphometric and molecular analysis of variability of the Mayan cichlid *Mayaheros urophthalmus* in the Yucatan. Copeia 106:268–278. https://doi.org/10.1643/ci-17-657

Bautista F, Palacio-Aponte G, Quintana P, Zinck JA (2011) Spatial distribution and development of soils in tropical karst areas from the peninsula of Yucatan, Mexico. Geomorphology 135: 308–321. https://doi.org/10.1016/j.geomorph.2011.02.014

Brown JH, Lomolino MV (1988) Biogeography. Sinauer, Sunderland

Chakrabarty P (2006) Systematics and historical biogeography of greater Antillean Cichlidae. Mol Phylogenet Evol 39:619–627. https://doi.org/10.1016/j.ympev.2006.01.014

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9:1657–1659

Concheiro Pérez GA, Říčan O, Ortí G et al (2007) Phylogeny and biogeography of 91 species of heroine cichlids (Teleostei: Cichlidae) based on sequences of the cytochrome b gene. Mol Phylogenet Evol 43:91–110. https://doi.org/10.1016/j.ympev.2006.08.012

Condamine FL, Clapham ME, Kergoat GJ (2016) Global patterns of insect diversification: towards a reconciliation of fossil and molecular evidence? Sci Rep 6. https://doi.org/10.1038/srep19208

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9:772

Delmonte B, Robert Petit J, Basile-Doelsch I et al (2007) Late Quaternary interglacials in East Antarctica from ice-core dust records. Dev Quat Sci 7:53–73. https://doi.org/10.1016/S1571-0866(07)80031-5

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969–1973. https://doi.org/10.1093/molbev/mss075

Espadas Manrique C, Durán R, Argáez J (2003) Phytogeographic analysis of taxa endemic to the Yucatán peninsula using geographic information systems, the domain heuristic method and parsimony analysis of endemicity. Divers Distrib 9: 313–330. https://doi.org/10.1046/j.1472-4642.2003.00034.x

Farias IP, Ortí G, Sampaio I et al (2001) The cytochrome b gene as a phylogenetic marker: the limits of resolution for analyzing relationships among cichlid fishes. J Mol Evol 53:89–103. https://doi.org/10.1007/s002390010197

Gondwe BRN, Lerer S, Stisen S et al (2010) Hydrogeology of the South-Eastern Yucatan peninsula: new insights from water level measurements, geochemistry, geophysics and remote sensing. J Hydrol 389:1–17. https://doi.org/10.1016/j.jhydrol.2010.04.044

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704. https://doi.org/10.1080/10635150390235520

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: Nucleic acids symposium series. [London]: Information Retrieval Ltd., c1979-c2000., pp 95–98

Haq BU, Hardenbol J, Vail PR (1987) The new chronostratigraphic basis of Cenozoic and Mesozoic sea level cycles. In: Timing and depositional history of eustatic sequences: constraints on seismic stratigraphy. Cushman Foundation for Forminiferal Research, Special Publications, pp 7–13

Harrison E, Trexler JC, Collins TM et al (2014) Genetic evidence for multiple sources of the non-native fish *Cichlasoma urophthalmus* (Günther; Mayan cichlids) in southern Florida. PLoS One 9:e104173. https://doi.org/10.1371/journal.pone.0104173

Head MJ, Gibbard PL (2015) Formal subdivision of the quaternary system/period: past, present, and future. Quat Int 383:4–35. https://doi.org/10.1016/j.quaint.2015.06.039

Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst Biol 58:367–380

Hubbs CL (1936) Fishes of the Yucatan peninsula. Carnegie Inst Wash Publ 17:157–287

Hulsey CD, López-Fernández H (2011) Nuclear Central America. In: Albert JS, dos Reis RE (eds) Historical biogeography of Neotropical freshwater fishes. University of California Press, Oakland, pp 279–292

Humphries JM, Miller RR (1981) A remarkable species flock of pupfishes, genus *Cyprinodon*, from Yucatan, Mexico. Copeia 1981:52–64

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome b gene of mammals. J Mol Evol 32:128–144. https://doi.org/10.1007/BF02515385

Lovejoy NR, de Araújo MLG (2000) Molecular systematics, biogeography and population structure of Neotropical freshwater needlefishes of the genus *Potamorrhaphis*. Mol Ecol 9:259–268. https://doi.org/10.1046/j.1365-294X.2000.00845.x

Lugo-Hubp J, Aceves-Quesada F, Espinasa-Pereña R (1992) Rasgos geomorfológicos mayores de la península de Yucatán. Rev del Inst Geol 10:143–150

Martin AP, Bermingham E (1998) Systematics and evolution of lower central American cichlids inferred from analysis of cytochrome b gene sequences. Mol Phylogenet Evol 9:192–203

Martínez-Palacios CA, Ross LG (1988) The feeding ecology of the central American cichlid *Cichlasoma urophthalmus* (Gunther). J Fish Biol 33:665–670

Matamoros WA, McMahan CD, Chakrabarty P, et al (2014) Derivation of the freshwater fish fauna of Central America revisited: Myers' hypothesis in the twenty-first century. Cladistics 1–12

Miller RR, Minckley WL, Norris SM (2009) Peces dulceacuícolas de México. CONABIO, SIMAC, ECOSUR, DFC, Mexico City

Monmonier MS (1973) Maximum-difference barriers: an alternative numerical regionalization method. Geogr Anal 5:245–261. https://doi.org/10.1111/j.1538-4632.1973.tb01011.x

Montaggioni LF, Braithwaite CJR (2009) Quaternary coral reef systems: history, development processes and controlling factors. Elsevier, Amsterdam

Morrone JJ (2005) Hacia una síntesis biogeográfica de México. An del Inst Biol Univ Nac Autónoma México, Ser Botánica 76:207–252

Nico LG, Beamish WH, Musikasinthorn P (2007) Discovery of the invasive Mayan cichlid fish *"Cichlasoma" urophthalmus* (Günther, 1862) in Thailand, with comments on other introductions and potential impacts. Aquat Invasions 2:197–214. https://doi.org/10.3391/ai.2007.2.3.7

Ogg JG, Ogg GM, Gradstein FM (2016) The concise geologic time scale. Elsevier, Amsterdam

Pfenninger M, Posada D, Shaw K (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. Evolution (N Y) 56:1776–1788

Rambaut A (2016) FigTree vers. 1.4.3

Rambaut A, Suchard MA, Xie DD (2014) Tracer vers. 1.6

Razo-Mendívil U, Vázquez-Domínguez E, Pérez-Ponce de León G (2013) Discordant genetic diversity and geographic patterns between *Crassicutis cichlasomae* (Digenea: Apocreadiidae) and its cichlid host, *"Cichlasoma" urophthalmus* (Osteichthyes: Cichlidae), in middle

America. J Parasitol 99:978–988. https://doi.org/10.1645/13-225.1

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

Rohling EJ, Foster GL, Grant KM et al (2014) Sea-level and deep-sea-temperature variability over the past 5.3 million years. Nature 508:477–482. https://doi.org/10.1038/nature13230

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Ronquist F, Teslenko M, Van Der Mark P et al (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542

Schmitter-Soto JJ (2002) Ictiogeografía de Yucatán, México. In: Lozano-Villano M de L (ed) Libro Jubilar en Honor al Doctor Salvador Contreras Balderas. Universidad Autónoma de Nuevo León, Monterrey, pp 103–116

Schmitter-Soto JJ (1999) Distribution of continental fishes in northern Quintana Roo, Mexico. Southwest Nat 44:166–172

Schmitter-Soto JJ, Comín FA, Escobar-Briones EG et al (2002) Hydrogeochemical and biological characteristics of cenotes in the Yucatan peninsula (SE Mexico). Hydrobiologia 467:215–228. https://doi.org/10.1023/A:1014923217206

Schmitter-Soto JJ, Gamboa-Pérez HC (1996) Composición y distribución de peces continentales en el sur de Quintana Roo, Península de Yucatán, México. Rev Biol Trop 44:199–212

Simoni M, Bakker E, Eurlings MCM et al (1999) Laboratory guidelines for molecular diagnosis of Y-chromosomal microdeletions. Int J Androl 22:292–299. https://doi.org/10.1046/j.1365-2605.1999.00193.x

Simpson GG (1964) Species density of north American recent mammals. Syst Zool 13:57–73. https://doi.org/10.2307/2411825

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129:555–562

Stenico M, Nigro L, Barbujani G (1998) Mitochondrial lineages in ladin-speaking communities of the eastern Alps. Proc R Soc Lond B Biol Sci 265:555–561

Templeton AR, Crandall KA, Sing CF (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

Thorne JL, Kishino H, Painter IS (1998) Estimating the rate of evolution of the rate of molecular evolution. Mol Biol Evol 15:1647–1657. https://doi.org/10.1093/oxfordjournals.molbev.a025892

Polzin T, Daneshmand SV, (2003) On Steiner trees and minimum spanning trees in hypergraphs. Operations Research Letters 31(1):1220

Udvardy MDF (1975) A classification of the biogeographical provinces of the world. IUCN, Morges

Urrutia-Fucugauchi J, Chávez-Aguirre JM, Pérez-Cruz L, de la Rosa JL (2008) Impact ejecta and carbonate sequence in the

eastern sector of the Chicxulub crater. Compt Rendus Geosci 340:801–810. https://doi.org/10.1016/j.crte.2008.09.001

Vázquez-Domínguez E, Arita HT (2010) The Yucatan peninsula: biogeographical history 65 million years in the making. Ecography (Cop) 33:212–219. https://doi.org/10.1111/j.1600-0587.2009.06293.x

Ward RD, Hanner R, Hebert PDN (2009) The campaign to DNA barcode all fishes, FISH-BOL. J Fish Biol 74:329–356. https://doi.org/10.1111/j.1095-8649.2008.02080.x

Wilkens H (1982) Regressive evolution and phylogenetic age: the history of colonization of freshwaters of Yucatan by fish and crustacea. Texas Meml Museum Bull 28:237–243

Wilken H, Strecker U (2017) Evolution in the dark. Darwin´s loss without selection. Springer, Berlin, Heidelberg https://doi.org10.1007/978-3-662-54512-6

Woodburne MO (2010) The great American biotic interchange: dispersals, tectonics, climate, sea level and holding pens. J Mamm Evol 17:245–264. https://doi.org/10.1007/s10914-010-9144-8