Extreme genetic divergence in the endemic fish *Chirostoma humboldtianum*: implications for its conservation

Divergencia genética extrema en el pez endémico *Chirostoma humboldtianum*: implicaciones para su conservación

García-Martínez Rosa María[1, 2], Omar Mejía[3], Francisco Javier García-De León[4] and Irene De Los Angeles Barriga-Sosa[2]

[1]Doctorado en Ciencias Biológicas y de la Salud, División CBS, Universidad Autónoma Metropolitana Unidad Iztapalapa, Av. San Rafael Atlixco 186 Col. Vicentina, Iztapalapa, D. F. 09340. México
[2] Laboratorio de Genética y Biología Molecular de la Planta Experimental de Producción Acuícola Departamento de Hidrobiología Universidad Autónoma Metropolitana Unidad Iztapalapa, Av. San Rafael Atlixco 186 Col. Vicentina, Iztapalapa, D. F. 09340. México
[3] Laboratorio de Variación Biológica y Evolución Departamento de Zoología Escuela Nacional de Ciencias Biológicas-IPN, Miguel Hidalgo, D.F. 11340. México
[4] Laboratorio de Genética para la Conservación Centro de Investigaciones Biológicas del Noroeste, S.C. La Paz, B.C.S. 23090, México
email: ibs@xanum.uam.mx

García-Martínez R. M., H. O. Mejía Guerrero, F. J. García-De León and I. A. Barriga-Sosa. 2015. Extreme genetics divergence in the endemic fish *Chirostoma humboldtianum*: implications for its conservation. *Hidrobiológica* 25 (1): 95-106

ABSTRACT

Chirostoma humboltianum is an endemic species widely distributed in isolated basins of Central México. However, habitat alteration had reduced the range of distribution and led to the local extinction of the species in some basins. During the Miocene these basins were connected, allowing dispersal and colonization of new hydrological systems. Later on, tectonic, volcanic and climatic events of the Plio-Pleistocene promoted continuous periods of isolation and reconnection allowing the species evolve through continuous cycles of expansion and contraction of its distribution. Therefore it is expected that these events have left signals in the geographical distribution and genetic diversity and divergence of existing populations. Although the analysis of genetic diversity and genetic structure in the population becomes an important factor for the conservation of a species, few studies have been made in this taxon. In this study we used a 341pb segment of the domain I of the hypervariable region of the mitochondrial control region to analyze the genetic diversity and their distribution in 20 individuals of each one of six lakes located in central México. The values of haplotypic (0 - 0.938) and nucleotide (0 - 0.0352) diversity suggested continuous periods of expansion and population contraction related with the formation of the lakes during the Pleistocene which is supported by the BSP and mismatch analysis, and recent anthropogenic factors. In addition, the large number of exclusive haplotypes (66%) and the highly significant genetic differentiation among populations suggests that each one of the population must be conserved because each one is an important component in the evolutionary legacy of the species.

Key words: *Chirostoma humboldtinaum*, control region, genetic diversity, population genetics.

RESUMEN

Chirostoma humboltianum es una especie endémica ampliamente distribuida en cuencas aisladas del Centro de México. Sin embargo, la alteración del hábitat ha reducido drásticamente el área de distribución y llevado a la extinción local de la especie en algunas cuencas. Durante el Mioceno estas cuencas estuvieron conectadas, permitiendo la colonización y dispersión en nuevos sistemas hidrológicos. Posteriormente, los eventos tectónicos, volcánicos y climáticos del Plio-Pleistoceno promovieron continuos periodos de aislamiento y reconexión, permitiendo que la especie evolucionara a través de ciclos de expansión y contracción de su distribución. Por lo tanto, se espera que estos eventos hayan dejado huella en la distribución geográfica y diversidad genética de las poblaciones existentes. Si bien, el análisis de la diversidad genética en una población llega a ser un factor importante para la conservación de una especie, pocos estudios han sido realizados en este taxón. En este trabajo usamos un segmento de 341pb del dominio I hipervariable de la región control mitocondrial, para analizar la diversidad genética y su distribución para 20 individuos de cada uno de los seis lagos localizados en la región central de México. Los valores de diversidad haplotípica (0-0.938) y nucleotídica (0-0.0352) sugieren continuos periodos de expansión y contracción poblacional relacionado con la formación de los lagos en el Pleistoceno lo cual es soportado por los análisis BSP y mismatch; y de factores antropogénicos recientes. En adición, la gran cantidad de haplotipos exclusivos (66%) y la alta diferenciación genética significativa entre las poblaciones, sugiere que cada una de las poblaciones debe de ser conservada porque cada una es un componente importante en el legado evolutivo de la especie.

Palabras clave: *Chirostoma humboldtinaum*, diversidad genética, genética poblacional, región control.

INTRODUCTION

The genus *Chirostoma* (Pisces: Atherinopsidae) (Saeed *et al.*, 1994; Dyer & Chernoff, 1996) is a group of endemic fishes inhabiting lotic and lentic systems in Central México (Barbour, 1973a; Miller *et al.*, 2005). Although the taxonomic status of the group has not fully resolved, Barbour (1973b) considers that there are 18 species and 6 subspecies. *C. humboldtianum* (Valenciennes, 1853) is one of the species with the widest geographic distribution. This species is found in geographically isolated lakes and ponds of both clear and turbid water from the valley of Mexico to the Pacific Ocean following the Lerma - Santiago basin system (Miller *et al.*, 2005). Although currently the distribution of the species occurs in isolated basins, lacustrine deposits of the Miocene suggested that the Mesa Central of México was once drained by the vast Lerma - Santiago system (Miller & Smith, 1986) that in turn was connected with other reservoirs that had greater extensions (Barbour, 1973a; Israde–Alcántara, 1997; Moncayo-Estrada *et al.*, 2001). Therefore, it has been pointed that these connections allowed *C. humboldtianum* populations to disperse and colonize new hydrological systems from the east to the west. The formation of the Mesa Central in the Pliocene, coupled with process of tectonic led to the compartmentalization of the basins, which were intensified by the formation of the Transvolcanic Belt that in turn promoted vicariant events in various aquatic organisms (Webb, 2004; Mulcahy & Mendelson, 2000) and a strong divergence from intraspecific to the interespecific level (Echelle & Echelle, 1984; Barbour, 1973a). Likewise, geological activity, coupled with severe climatic oscillations promoted continuous periods of isolation and reconnections that generated a complex hydrologic system (Moncayo-Estrada *et al.*, 2001; Israde-Alcantara, 1997). The above mentioned events of fragmentation and colonization are expected to leave signatures in the geographical distribution and in the genetic diversity of extant populations. Genetic diversity and geographical distribution often reflects process that occurred during historical time, as well as processes and environmental changes over contemporary time related mainly with the anthropogenic activities (Avise, 2000; Bernatchez & Wilson, 1998; Hewitt, 1996).

In *C. humboldtianum*, some factors related to human activity such as habitat loss, pollution and overfishing (Álvarez & Navarro, 1957), as well as introduction of non-native fishes (Barbour, 1973a) have recently played important roles in the decline or even disappearance (local extinction) of populations. In spite of this situation documented over almost six decades, the species has not been appointed as threatened in the Official Mexican Standard Norms (NOM-059-Ecol-2001) nor in the IUCN Red List.

The aim in this study is to evaluate genetic diversity and differentiation of *C. humboldtianum* populations and to elucidate how historical and current factors have influenced its distribution, in order to develop a rational management and conservation program for the species due that the level of genetic diversity reflects the evolutionary potential of a species, which in turn is necessary to enable populations to cope with future environmental changes (Frankham, 2005).

MATERIALS AND METHODS

Sampling sites and sequencing of DNA. One hundred and twenty specimens of *C. humboldtianum* were collected from six locations in Central Mexico (see Fig. 1 & Table 1) and stored in 98% alcohol for mtDNA analysis.

Twenty individuals per locality were used for DNA extraction from ethanol preserved muscle using the salt extraction protocol of Aljanabi & Martinez (1997). A fragment of 360 bp of the hypervariable section I of the mtDNA control region was amplified by PCR using the primers described by Pérez-Ramírez (2005), DloopF (Forward) 5'-

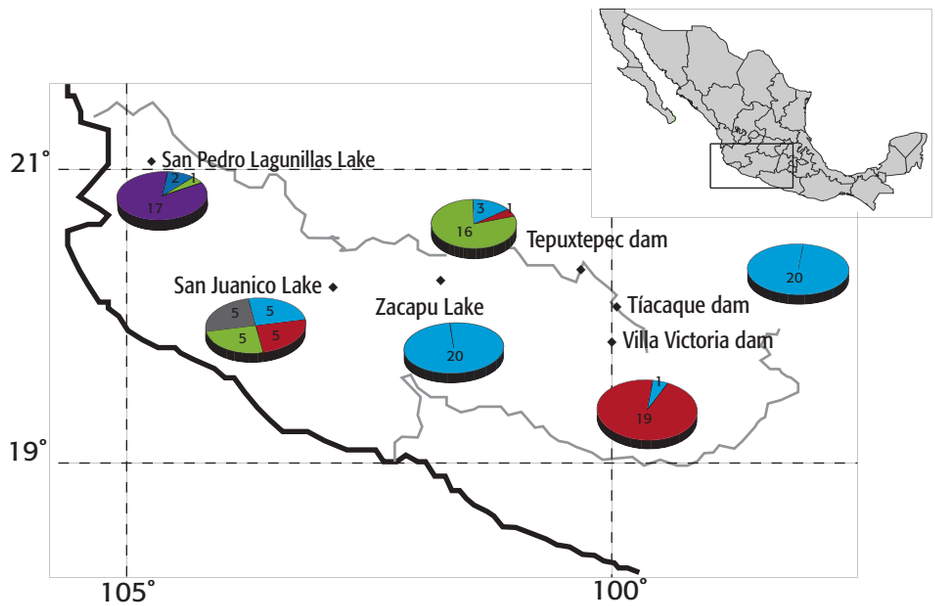

Figure 1. Map showing collected localities from *Chirostoma humboldtianum* specimens. The graphs represents haplotypes groups (Group A = blue, Group B = red, Group C = green, Group D = purple and Group E = gray), the numbers on the graphs represents the individuals of each population by each group of haplotypes.

Table 1. Summary of genetic diversity parameters in six Mexican lacustrine samples of *Chirostoma humboldtianum* based on mtDNA control region sequences. It shows the number of specimens (n), haplotype number (HN), nucleotide and haplotype diversities (π, h, ±SD).

Collecting site	ID	Geographic coordenates	Collecting year	n	HN	*h*	*π*
Las Tazas dam, (Tiacaque), Edo. de México	T	19º38'29"N-99º42'27"O 2540 msnm	2002	20	1	0	0
Villa Victoria dam, Edo. de México	VV	19°26'N-100°00'O 2570 msnm	2009	20	9	0.789±0.086	0.0123±0.0029
Tepuxtepec dam, Michoacán	Tx	19°59'42"N-100°13'33"O 2361 msnm	2010	20	11	0.932±0.030	0.0261±0.0033
Zacapú lagoon, Michoacán	Z	19°49'26"N-101°46'45" 1980 msnm	2003	20	9	0.832±0.063	0.0098±0.0014
San Juanico lagoon, Cotija, Michoacán	SJ	19°51'16"N-102°40'34"O 1625 msnm	2010	20	13	0.958±0.025	0.0384±0.0020
San Pedro Lagunillas lagoon, Nayarit	SP	21°12'48"N-104°44'12"O 1248 msnm	2010	20	17	0.979±0.024	0.0186±0.0038
Total				120	60	0.938±0.016	0.0352±0011

GCT CTA ACT CCC AGG AAA ATT-3 ' and DloopR1 (Reverse) 5'-CAC CCC GAT TGC AAC TGT CC-3 '. PCR reactions were carried out in a final volume of 25 µl and using a Biorad MyCycler Thermal Cycler (USA). PCR reactions contained 50-100 ng DNA, 1X PCR buffer, 2.5 mM $MgCl_2$, 2.5 mM dNTP mixture, 10 nmol of "forward" and "reverse" primers and 1U of enzyme Promega GoTaq PCR. The PCR program used includes an initial denaturation step at 96°C for 2 min, followed by 30 cycles of 96°C for 45s, 59°C for 30s, and 72°C for 45s, with a final extension at 72°C for 5min. The PCR products were visualized in agarose gels stained with ethidium bromide. Amplified products were purified using Wizard SV Genomic DNA Purification System Promega kit. The purified PCR products were sequenced in both directions using the PCR primers "forward" and "reverse" in a 10 µl reaction containing 2 µl BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems), 2µl 2X buffer, 1µl of each primer (10 µm), 3µl DNA template (containing approximately 50 ng of DNA) and 2µl of ddH2O water. The thermocycler amplification conditions were: 35 cycles at 96 °C for 30 s, 59.2 °C for 15 s and 60 °C for 4 min. Sequencing reactions were resolved in the automatic sequencer ABI Prism 3100 Avant (Applied Biosystems) at Laboratorio Divisional de Biología Molecular de la UAM-Iztapalapa.

Genetic diversity and population structure. Compiled DNA sequences were edited in BioEdit (Hall, 1999), aligned with Clustal X (Thompson *et al.*, 1997) and checked by eye. Genetic diversity within lakes was determined by assessing the number of haplotypes, polymorphic sites, haplotype (*h*) and nucleotide diversity (*π*), using DnaSP *ver.* 5.0 (Librado & Rozas, 2009).

Population structure was determined using BAPS (Corander *et al.*, 2003; 2004). The Bayesian method implemented in the BAPS software was run under spatial model with the maximal number of groups (K) set from 1 to 20. Each run was replicated six times, and the results were averaged according to the resultant likelihood scores. Additionally the spatial distribution of genetic variation was examined using a hierarchical Analysis of Molecular Variance (AMOVA, Excofier *et al.*, 1992) with significance levels set at $\alpha = 0.05$ and 10,000 random permutations as implemented in Arlequin 3.5.1.2 (Excoffier & Lischer, 2010). Total genetic variation was partitioned according to the number of clusters defined in BAPS software. On the other hand, genetic differentiation between pairs of population were evaluated a *posteriori* by means of the pairwise F_{ST} (Excoffier *et al.*, 1992). The significance of the F_{ST} value was evaluated by performing a randomization test with 10000 replications with a level of significance of $\alpha = 0.05$. The diversity indices and F_{ST} values were calculated using the Tamura-Nei (1993) model with gamma shape distribution parameter of 0.473, which was determined by Modeltest (Posada & Crandall, 1998) as the best evolutionary model for our sequences.

Relationships among populations. The genetic relationships among the resolved haplotypes were reconstructed using the neighbor-joining (NJ) method (Saitou & Nei, 1987) implemented in PAUP (Swofford, 1998). Genetic distances were generated for phylogenetic reconstructions using resolved models of substitution suggested by Modeltest (Posada & Crandall, 1998). The best fit for the control region data was TRN model (Tamura & Nei, 1993) with invariable sites and gamma shape parameters (TRN + I + Γ, I = 0.642, Γ = 0.473, with bases frequencies: A: 0.3582, C: 0.1983, G: 0.1630, T: 0.2806). A bootstrap analysis with 1000 replicates was used to evaluated the support for genetics relationships (Felsenstein, 1985).

In addition, genealogical relationships were examined by means of haplotype networks using the median joining algorithm implemented in Network software (Bandelt *et al.*, 1999).

Demographic patterns. Considering that both historical and contemporary process of basin had contributed to the genetic composition of the species, we tested for possible changes in demographic patter-

ns using several approaches. First, in order to quantify the significant departure from mutation-drift equilibrium, we evaluated the neutrality estimators Tajima's *D* (Tajima, 1989) and Fu´s *Fs* (Fu, 1997) in Arlequin 3.5.1.2 (Excoffier & Lischer, 2010). Likewise, we performed a mismatch distribution analysis (MMD) of pairwise substitution differences among haplotypes to investigate the demographic history of population by comparing the observed distribution with that expected under the assumption of an expansion model. A unimodal mismatch distribution indicates a recent range expansion; multimodal, including bimodal, mismatch distribution indicates diminishing population size or structured, in contrast a population that has been stationary for a long time become ragged and erratic (Excoffier and Schneider , 1999; Roger and Harpending, 1992; Excoffier *et al.*, 1992), however, the multimodal distributions may also indicate that the population is influenced by migration, is subdivided and/or has undergone a greater number of recent coalescent events (historical contractions) (Ray, *et al.* 2003; Marjoram & Donnelly, 1994). The smoothness of the observed distribution was quantified by the sum-of-squared deviations and Harpending´s raggedness index *R* (Harpending, 1994) as implemented in Arlequin ver. 3.5.1.2 (Excoffier & Lischer, 2010).

Secondly, we estimated the demographic expansion parameters Γ (MMD age expansion parameter), Θ_0 (pre-expansion population size), Θ_1 (post-expansion population size), using the generalized non-linear least-square approach. The absolute time of population expansion (*t)* was calculated through the relationship $t = \Gamma/2u$, where Γ measures the time in unit of ½ μ generations, μ is the mutation rate per sequence under study per generation and u represents the mutation rate per sequence under study per year (Rogers & Harpending, 1992). We used a time of generation of 1 year and a mutation rate of 3.5% as estimated for the same region in the silverfish *Pleuragramma antarcticum* (Zane *et al.*, 2006).

As a final step, the demographic history was reconstructed using the Bayesian skyline plot method to detected changes in the effective population size through time using Beast 1.5.1 (Drummond *et al.*, 2005). Four independent Markov chains assuming a strict molecular clock and a coalescent Bayesian skyline model were run for 30 million generations with sampling performed every 10,000 steps. The results of the four independent chains were combined in LogCombiner 1.5.4. (Drummond & Rambaut, 2007) and the Bayesian skyline plots for each population were generated in Tracer 1.5 (Rambaut & Drummond, 2009).

RESULTS

Genetic diversity and population structure. We sequenced and analyzed 341 pb for the control region (CR) in 120 specimens, and recovered 57 polymorphic sites (42 parsimony informative) and 55 haplotypes (Accession KF651987 to KF652041), with almost 66% of unique haplotypes. No insertions or deletions were detected. Overall haplotype diversity was $h = 0.938$ and nucleotide diversity $\pi = 0.0352$ (Table 1). Most haplotypes were defined by substitutions at a single site. Only two haplotypes were shared among individuals from different localities. One of them was also the most common haplotype, found in 27 individuals from the sites Tiacaque (20), Villa Victoria dam (1), Tepuxtepec dam (3), Zacapu lake (1) and San Juanico lake (2), and the other haplotype was found in two individuals from San Juanico lake and San Pedro Lagunillas lake. Seventeen of the haplotypes were shared among individuals within the same localities. The remaining 36 haplotypes (66%) were presented only in single individuals. Haplotype diversity values (*h*) were high, ranging from 0.789 to 0.979 (Table 1). The exception was the Tiacaque population for which only one haplotype was resolved. Nucleotide diversity (π) analyzed per population showed values ranging from 0.0000 to 0.0384, with the highest value observed in San Juanico (0.0384) and the lowest in Tiacaque (0.00) (Table 1).

BAPS resolved five genetic clusters, with log marginal likelihoods of -1408.6458. The analysis showed a mixture of individuals from different geographic locations in the same clusters. Group A is the most heterogeneous and it is constituted by individuals from Tiacaque dam (20 individuals), Zacapu Lake (20), Tepuxtepec dam (3), San Juanico Lake (5), San Pedro Lagunillas Lake (2) and Villa Victoria dam (1). Group B is conformed by individuals from Villa Victoria dam (19), San Juanico Lake (5) and Tepuxtepec dam (1). Group C include individuals from Tepuxtepec dam (16), San Juanico Lake (5) and San Pedro Lagunillas Lake (1). Group D is formed of individuals from San Pedro Lagunillas Lake (17) and group E was constituted with the specimens from San Juanico Lake (5).

An Analysis of Molecular Variance (AMOVA) revealed that 70.30% of the total variation corresponds to variation unshared among BAPS defined groups, 10.15% is variation unshared among population within a group and 19.55% is variation distributed within populations. The fixation indices shown significant differences in the three levels after 10000 randomizations ($F_{CT} = 0.030$, $P < 0.001$; $F_{SC} = 0.3418$, $P < 0.001$; $F_{ST} = 0.8045$, $P < 0.001$) (Table 2).

Table 2. AMOVA results based on samples of *Chirostoma humboldtianum* from six locations of Central region at Mexico. The significance values associated to the variance components were obtained after 10000 randomizations.

Source of variation	df	Sum squares	Variance Components	Variance %	*P*
Among groups ($F_{CT} = 0.7030$)	4	502.142	5.2843	70.30	< 0.001
Among populations within groups ($F_{SC} = 0.3418$)	9	53.857	0.7631	10.15	< 0.001
Within populations ($F_{ST} = 0.8045$)	106	155.743	1.4693	19.55	<0.001
Total	119	711.742	7.5166		

The F_{ST} paired values were high and significant, and ranged from 0.1712 for San Juanico lake – Tepuxtepec dam up to 0.8288 for Villa Victoria dam – Tiacaque (Table 3).

Relationships among populations. The generated NJ of the 55 resolved haplotypes allowed the identification of the same five groups recovered by BAPS, and analyzed in AMOVA, with bootstrap support higher than 60 % (Fig. 2).

The median joining network exhibits a pattern consistent with a complex demographic history and little geographic structure (Fig. 3). Haplotype A01 is widely distributed, occurring at high frequency in Tiacaque dam and also being present in four of the five basins (Villa Victoria dam, Tepuxtepec dam, San Juanico and Zacapu lakes). Haplotype A10 was shared between two individuals from different localities (San Juanico and San Pedro Lagunillas), 17 of the 55 haplotypes were shared by individuals of the same location and the remaining 36 haplotypes were found each at a single individual. The five groups recovered in the Bayesian analysis and NJ was also resolved in the haplotypes network.

Haplogroups A and B are separated by seven mutational steps, and those groups are separated from C, D and E by nine nucleotide substitutions (Fig. 3). Groups C and D are separated each one only by four mutational steps, while group C and E are separated by ten nucleotide substitutions. Haplotypes within groups were mostly separated by one or two nucleotide substitutions; however, groups C and D, shows some haplotypes separated by six substitutions.

Table 3. Pairwise F_{ST} values for the *Chirostoma humboldtianum* samples from six locations in Central region of Mexico (below the diagonal) and its significance value (above the diagonal).

	T	VV	Tx	Z	SJ	SP
T	0	0.00	0.00	0.00	0.00	0.00
VV	0.8288*	0	0.00	0.00	0.00	0.00
Tx	0.6512*	0.5572*	0	0.00	0.00	0.00
Z	0.5436*	0.6876*	0.5419*	0	0.00	0.00
SJ	0.5334*	0.3296*	0.1712*	0.4342*	0	0.00
SP	0.7684*	0.6698*	0.2901*	0.6740*	0.3421*	0

T = Tiacaque; VV = Villa Victoria dam; Tx = Tepuxtepec dam; Z = Zacapu lagoon; SJ = San Juanico lagoon; SP = San Pedro Lagunilla lagoon. *P < 0.05 values

Demographic patterns. The histograms of the mismatch distribution (MMD) of the entire population showed a bimodal distribution that deviated from the expected distribution under the sudden expansion model (Fig. 4). This outcome is supported by lack of significance of the Tajima's *D* and Fu's *Fs* tests. On the other hand, the MMD in Group A showed L-shaped distribution, which represent a balanced distribution. Groups B, C and D showed an unimodal distribution that suggests an expansion process (Fig. 4). The MMD for group E was not performed because this group is only comprised by five individuals and demographic analyzes lose their validity under a very low sample size. In all cases, raggedness indices (*r*) were not significant, thus not allowing us to reject the null hypothesis of stationarity (Table 4). The examined populations displayed negative non-significant Tajima's *D* values. Similarly, the calculated Fu's *Fs* values were negative, and only group D showed a significant value with a time since expansion of 225,178 years ago (τ= 2.6875) (Table 4). We estimated that during the expansion, the population increased from Θ_0 (1.4641) to Θ_1 (225). The Bayesian Skyline Plot analysis for group D also indicates a population expansion during the last 300,000 years. Group A, B show a stable trend in population size in the last 200,000 years, and C resolved also as stable since the last 250,000 years (Fig. 5). There is not graph for group E for the reasons previously mentioned.

DISCUSSION

Ward *et al*. (1994) suggested that freshwater fish in general terms exhibit lower levels of genetic diversity in comparison to marine fish using allozyme markers. Other works using allozymes markers had shown low genetic diversity levels for this and other species of the genus *Chirostoma*, although such results might be related to the low polymorphism exhibited by those markers (Barriga-Sosa *et al*., 2004) and to the limited number of specimens analyzed (6-8 individuals) (Barriga-Sosa *et al*., 2002). In contrast to the result reported by Ward *et al.* (1994), we recovered a relatively high number of haplotypes (55) with a mitochondrial marker in 120 analyzed specimens, with almost 66% of unique haplotypes, and high average values of both haplotype and nucleotide diversity (h = 0.938 and π = 0.035). For other species of freshwater fishes, lower values of genetic diversity have been reported using the same mitochondrial region. For instance, in fish living in cenotes, Vázquez-Domínguez *et al*. (2009) resolved 5 haplotypes in 56 individual and average haplotype and nucleotide diversity of 0.15 and 0.001 respectively; in the atherinomorphs, *Atherinomorus endrachtensis* (Quoy & Gaimard, 1825), Gotoh *et al*. (2011) found 72 haplotypes in 205 individuals, h = 0.151-0.593 and π = 0.00026 – 0.0029, whereas for the close relative *Odontestes argentinensis* (Valenciennes, 1835), Beherengaray & Sunnucks (2001) reported what they considered high values of h = 0.85 - 0.94 and π = 0.014 – 0.018.

The analysis of genetic diversity in the examined population of *C. humboldtianum* exhibited greatly variable values. Grant & Bowen (1998) have proposed that high values of nucleotide diversity (π) and haplotype diversity (h) found in San Juanico lake, Tepuxtepec dam, Villa Victoria dam and San Pedro Lagunillas lake and Zacapu lake (Table 1), are characteristic of a stable population with long evolutionary history or population admixture of differentiated lineages. However, the genealogical analysis does not support a long evolutionary history for

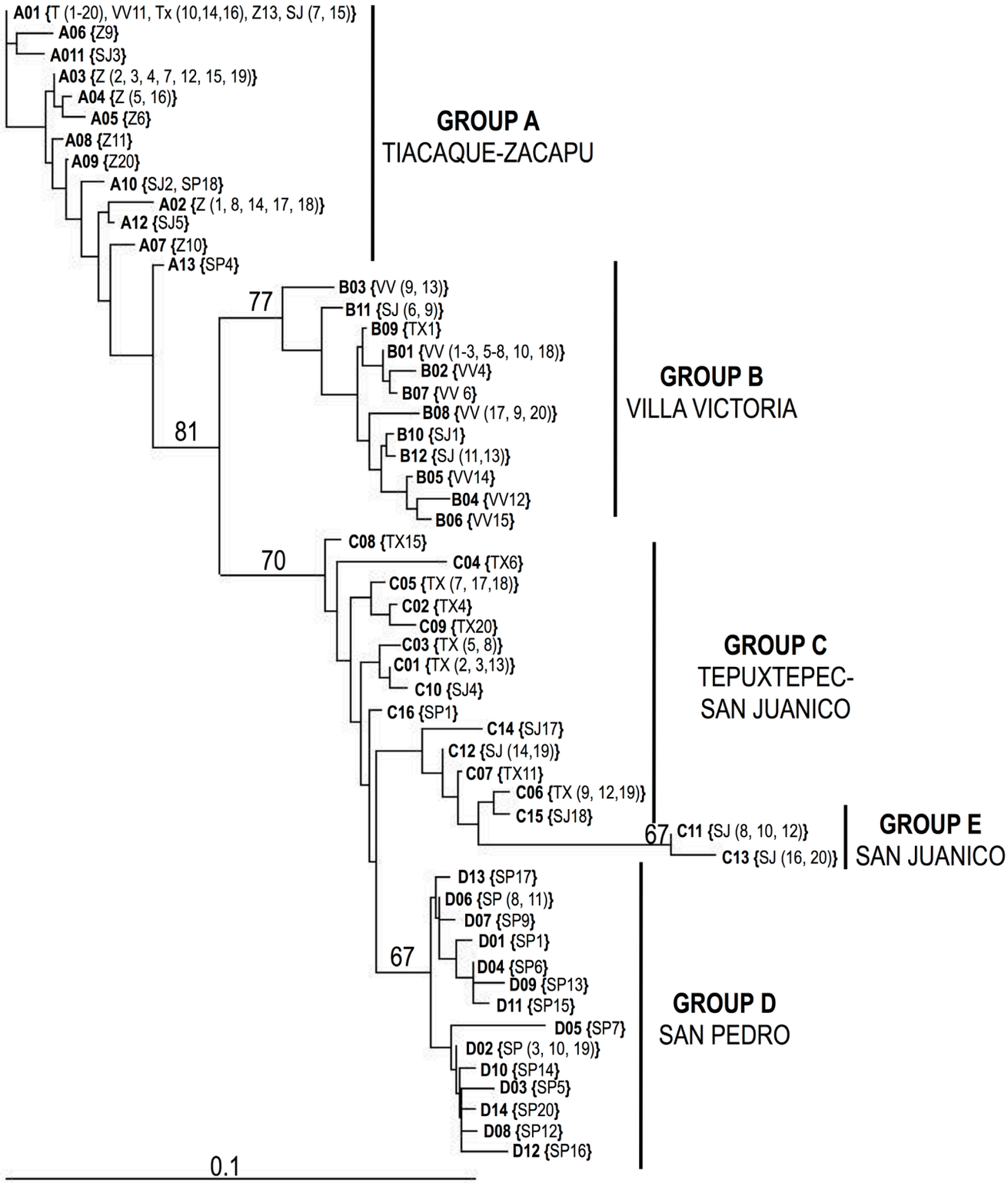

Figure 2. Unrooted NJ tree using the substitution model TRN with invariable sites and gamma distribution (TRN + I + Γ, I = 0.642, Γ = 0.473) for the 55 resolved haplotypes of *Chirostoma humboldtianum*. Bootstrap supports >60% and after 1000 replicates are shown in the upper part of the clades.

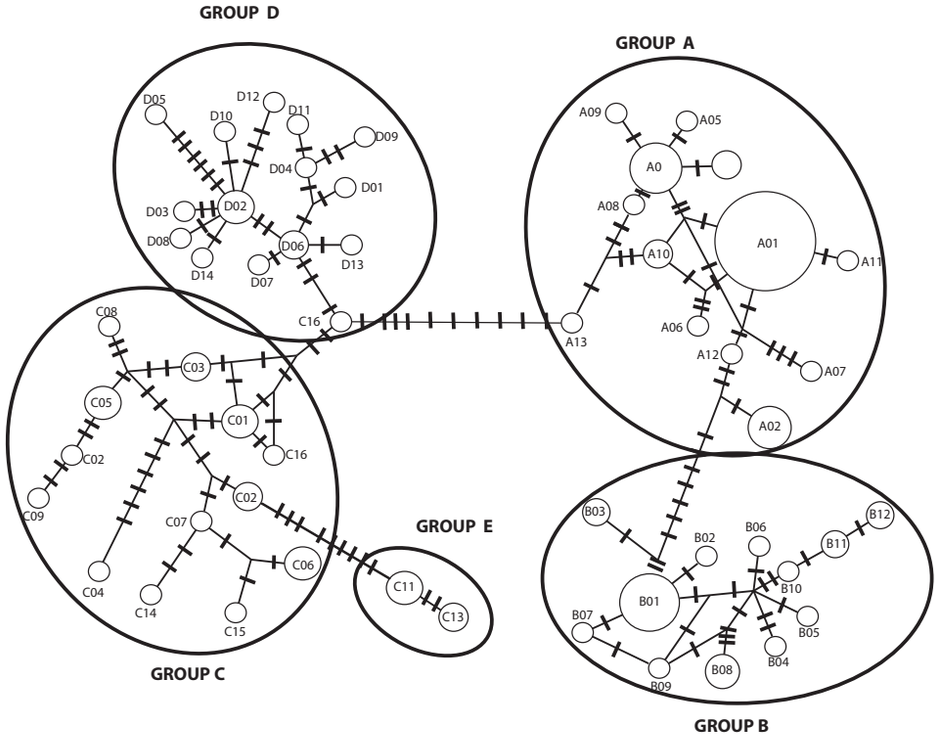

Figure 3. Haplotype median-network resolved for the 55 haplotypes of the CR and the previously resolved groups A, B, C, D & E. The sizes of the circles are proportional to the frequency of the haplotypes. Perpendicular bars to the vertical lines that connect the haplotypes represent the number of nucleotide substitutions between them.

the species, as can be observed by the absence of deep branching among haplotypes (Fig. 2). This result indicates that *C. humboldtianum* is a recent species (Moncayo-Estrada et al., 2001, Echelle & Echelle, 1984) that diversified ca. 0.52 mya (Bloom *et al.*, 2013). Furthermore, we believe that the admixture hypothesis is more likely for the species, since each geographic population consists of a mixture of haplotypes that do not show geographic pattern (Fig. 1).

The complex geological history of the Mesa Central in the Plio-Pleistocene caused several periods of isolation and reconnection among basins (Domínguez-Domínguez *et al.*, 2008; Miller *et al.*, 2005; Moncayo-Estrada *et al.*, 2001; Israde-Alcántara & Garduño-Monroy, 1999; Barbour, 1973a; Tamayo & West, 1964). In first instance, the separation of basins produced geographic isolation. This separation is well documented for the late Pleistocene and early Holocene (Álvarez & Moncayo, 1976; Barbour, 1973a). This geographic isolation is reflected in the high and significant genetic differentiation found among basins/groups in this study. On the other hand, the periods of reconnection favored gene flow among previously isolated populations. Additionally, climatic oscillations related to Pleistocene glaciations caused many cycles of contraction of geographic ranges during cold periods and expansion during subsequent warming (Bradbury, 2000; Israde-Alcántara & Garduño-Monroy, 1999; Ferrari *et al.*, 1999). In this sense, both analyses the mismatch distribution and Bayesian Skyline plot support the hypotesis of a rapid population growth of *C. humboldtianum* during this period. Likewise, the global MMD suggest connections and gene flow among the basins in the past or multiple events of population expansion.

The hypothesis of sudden expansion of group D is also supported by the negative Fu's *Fs* values and the star-shaped haplotype network (Fig. 3). This type of network is characteristic in populations that have experienced a recent population expansion from a smaller number of founders (Avise, 2000). The Bayesian skyline plot showed that the group D experienced an expansion process 200,000 ~ 300,000 years ago (Fig. 5); furthermore the expansion was estimated to be ~ 225,178 years ago from the sudden expansion model (τ = 2.26875) (Table 4). This estimation suggests that the expansion occurred in the middle Pleistocene (between 126,000 and 781,000) when volcanic and tectonic activity were more intense in the Mesa Central causing the overflow of the Chapala basin towards the ancestral Río Grande de Santiago and conducting to the isolation of Río Lerma and the formation of modern Ameca and Lerma-Santiago rivers (Miller *et al.*, 2005, Barbour, 1973a).

Our results of MMD suggest that the expansion time of the group C and group B was during middle Pleistocene (594,855 and 458,215 years ago, respectively) (τ = 2.6875) (Table 4). Previous studies and historical inferences based on the fish fauna of this region have reported a series of paleolakes that joined the Valle de México and Río Lerma (Domínguez-Domínguez *et al.*, 2008; Moncayo-Estrada *et al.* 2001; Israde-Alcántara & Garduño-Monroy, 1999; Barbour, 1973a; Tamayo & West, 1964). However, during the Plio–Pleistocene these connections were broken by tectonic and volcanic activity (Doadrio & Domínguez, 2004; Israde-Alcántara & Garduño-Monroy, 1999). During late Pleistocene to early Holocene a successive reconnection between the Valle de México and Río Lerma led to a new contact between species and to a genetic exchange through NE-SW and E-W faulting, which cuts the Sie-

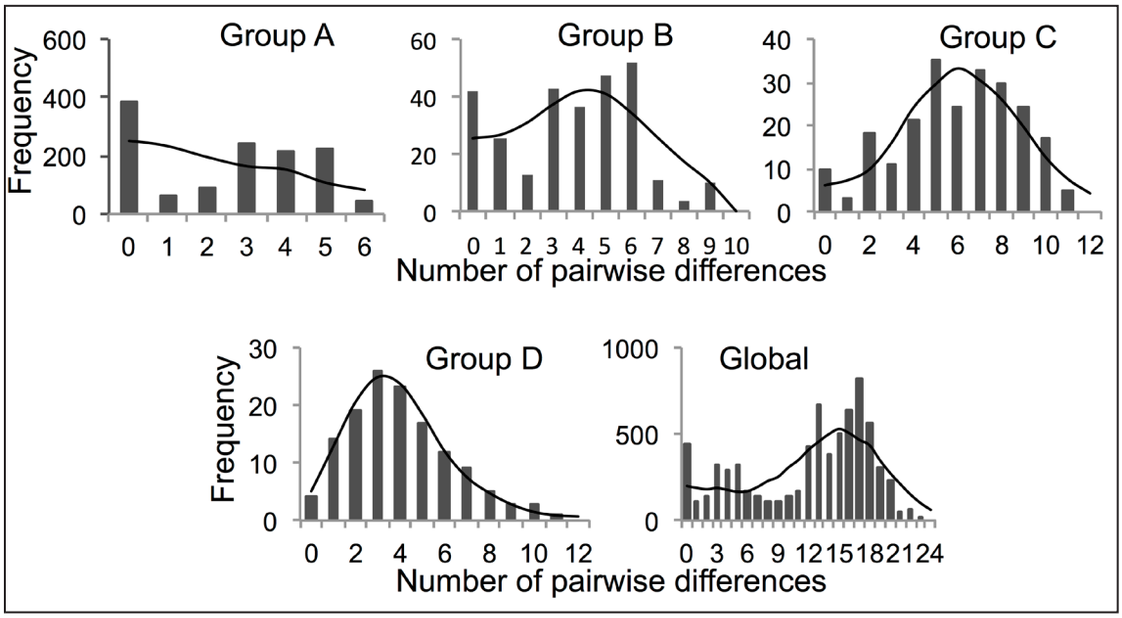

Figure 4. Mismatch distribution for four groups of *C. humboldtianum* and total population. Grey bars indicate the observed values and black lines show the expected distribution under the sudden expansion model.

Table 4. Demographic parameter estimates and neutrality test within the six populations samples in the central region of México.

Populations	*Fu's Fs*	*Tajima's D*	Mismatch distribution				
	Fs	*D*	τ	Θ_0	Θ_1	*SDD*	*Hri*
Group A	-3.1147^{ns}	-0.6758	4.7344	0	4.3066	0.0465^{ns}	0.1012^{ns}
Group B	-2.6818^{ns}	-0.4700^{ns}	5.4688	0.0018	11.1186	0.0255^{ns}	0.0615^{ns}
Group C	-3.3212^{ns}	-0.1819^{ns}	7.0996	0.088	34.4629	0.0051^{ns}	0.0203^{ns}
Group D	-8.2647**	-1.2522^{ns}	2.6875	1.4641	225	0.0006^{ns}	0.0150^{ns}
Group E	1.6876^{ns}	1.4588^{ns}	2.4199	0	3.8189	0.2335^{ns}	0.8800^{ns}
Total	-3.1389^{ns}	-0.2242^{ns}	4.4820	0.2949	55.7415	0.0622^{ns}	0.2156^{ns}

Fs, Fu's statistic; *D*, Tajima's D test; Γ, time since expansion expressed in units of mutation times; Θ_0, pre-expansion population size; Θ_1, post-expansion population size; *SDD*, sum of squared deviations, *Hri, Harpending's* raggedness index. $^{*}p< 0.05$, ** $p < 0.01$, ns = not significant.

rra de las Cruces (Israde- Alcantara & Garduño-Monroy, 1999; Barbour, 1973a). The above mentioned is supported by the fossil record of fishes of the Valle de México, which currently are not distributed in this basin (Álvarez & Moncayo, 1976).

The complex recent geological processes of Central México region were decisive in shaping the genetic structure of this species, however, complexity increased by human activity. For instance, it is documented that during the 70's translocations of species of "peces blancos" and minnows were conduced into several water reservoirs of México including Michoacán and Estado de México (Rosas, 1976). In this regard, Álvarez (1963) and Barbour (1973a) referred to *Chirostoma consosium consocium* and *C. c. reseratum* as native species of San Juanico lake, however, E. Soto-Galera (per.com. Laboratorio de Ictiología y Limnología) reports the current presence of *C. humboldtianum* in the basin. Furthermore and although the presence of the species in Zacapu Lake is referred by several authors as part of its original distribution (Paulo-Maya *et al.*, 2000; Barbour, 1973a; Álvarez, 1963), Medina Nava (1997) mentions its introduction in this basin, however, she did not mention possible year(s) of introduction. The species has also been reported in lakes Chapala, Pátzcuaro and Zirahuén (Barriga-Sosa *et al.*, 2002; Alaye, 1993; Alvarez & Navarro, 1957; Barbour, 1973b).

The present study allows concluding that the genetic structure of *C. humboldtianum* can relate to geological and climatic events of the Plio-Pleistoceno, although we cannot discard that translocations could

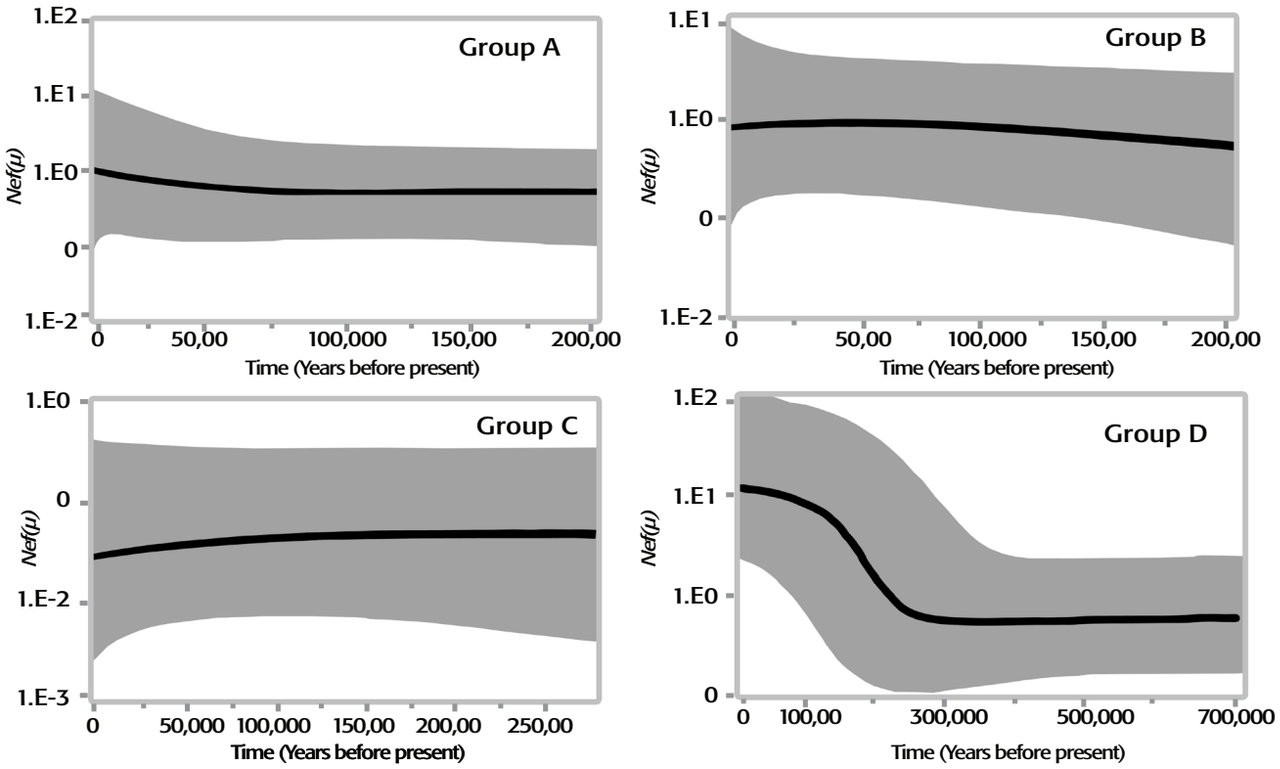

Figure 5. Bayesian Skyline Plot for four populations of *C. humboldtianum*, including the 95% highest probability density (HPD) interval.

explain the presence of the shared haplotypes in different geographical location. Thus, our results prompts for futher genetic population studies designed to determine the level of integrity and/or genetic introgression of the species and closely related species along its range of distribution.

Conservation implications. México has 375 freshwater fish species, 60% native (Miller, 1986). Anthropogenic factors such as habitats destruction, introduction of exotic species, pollution and eutrophication threaten wildlife inhabiting freshwater reservoirs (Miller *et al*. 2005).

A study performed by Soto-Galera *et al*. (1998) in 116 sites of the Río Lerma basin showed that of 44 native species recorded in the period 1885-1975, three were extinct, and 23 species had a large reduction in rank or health. In contrast, the same author found that in the period 1985-1993 more than half of the study sites disappeared or were so polluted that they could no longer support fish fauna. In addition, although *C. humboldtianum* had been designated as stable specie, it has been extirpated from some basins (Lyons *et al*. 1998).

The results of the present study might contribute to the development of conservation strategy plans for *C. humboldtianum*, as has been previously proposed by Barriga-Sosa *et al*. (2002), whom reported a strong genetic population differentiation in the species as it is also resolved in the present study (see pairwise F_{ST} values).

Although we are aware of the possible evolutionary stochastic effects and of the limitation of defining conservation units only on mtDNA results, we propose that five of the six analyzed populations could be proposed as candidates to be preserved under legal protection, since all of them contain a portion of the total variation of the species. Each of these geographic locations can be tentatively defined as an ESU (Evolutionary Significance Unit), since each ones is a segment of the population or group of stocks that are reproductive isolated from other con-specific, or unrelated population and represent an important component of the evolutionary legacy of the species (Waples, 1995). However, further studies using other molecular markers as microsatellites or SNPs, can aid in corroborating this proposal.

ACKNOWLEDGEMENTS

This study is part of the PhD thesis of the first author at Doctorado en Ciencias Biológicas y de la Salud at Universidad Autónoma Metropolitana– Iztapalapa. Fellowship CONACYT-224707 was granted to Rosa María García Martínez. Financial support was received from Consejo Nacional de Ciencia y Tecnología (CONACYT-CB-2009-01-130220) and UAM.147.07.03/147.09.0. Mexican government kindly issued permit number DGOPA.07343.310810.4128 to conduct this research. We thank the Biologist Ramón Cisneros, M. en B. Laura Georgina Núñez García and the fishermen of each collection site for support in the capture and maintenance of organisms. We also thank Dr. Joel Paulo Maya for their help in specimens identification.

REFERENCES

Alaye, R. N. 1993. Hematología de Aterínidos de aguas dulces: Género Chirostoma spp. del Lago de Pátzcuaro, Michoacán. *Ciencia Pesquera* 10: 97-109.

Aljanabi, S. M. & I. Martinez. 1997. Universal and rapid salt–extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* 25 (22): 4692–4693.

Álvarez, J. 1963.IctiologíaMichoacana III. Los peces de San Juanico y de Tocumbo, Michoacán. *Anales de la Escuela Nacional de Ciencias Biologicas* 12 (1-4): 111-138.

Álvarez, J. & M. E. Moncayo. 1976. Contribución a la paleoictiología de la Cuenca de México. *Anales del Instituto Nacional de Antropología e Historia* 6: 191-242.

Álvarez, J & L. Navarro. 1957. *Los peces del valle de México.* Secretaría de Marina, Instituto Nacional de Pesca, Dirección de Pesca e Industrias Conexas, México, D.F. 62 p.

Avise, J. C. 2000. *Phylogeography*: *The history and formation of species*. Harvard University Press, Cambridge, Massachusetts. 447 p.

Bandelt, H. J., P. Forster & A. Röhl. 1999. Median – joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16 (1): 37-48.

Barbour, C. D. 1973a. A Biogeographical history of *Chirostoma* (Pisces: Atherinidae): A Species Flock from the Mexican Plateau. *Copeia* 1973 (3): 533-556.

Barbour, C. D. 1973b.The systematic and evolution of the genus *Chirostoma* Swaison (Pisces: Atherinidae). *Tulane Studies in Zoology and Botany* 18 (3): 97-141.

Barriga-Sosa, I. D. L. A., A. L. Ibáñez-Aguirre & J. L. Arredondo Figueroa. 2002. Morphological and genetic variation in seven species of the endangered *Chirostoma "humboldtianum* species group" (Atheriniformes: Atherinopsidae). *Revista de Biología Tropical* 50 (1): 199-216.

Barriga-Sosa, I. D. L. A., L. E. Eguiarte & J. L. Arredondo Figueroa. 2004. Low but Significant subdivision among Populations of *Chirostoma grandocule* from Lake Patzcuaro, México. *Biotropica* 36 (1): 85-98.

Beheregaray L. B. & P. Sunnucks. 2001. Fine-scala genetic structure, estuarine colonization and incipient speciation in the marin silverside fish. *Molecular Ecology* 10: 2849 – 2866.

Bernatchez, L. & C. C. Wilson. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology* 7: 431- 452.

Bloom, D. D., J. T. Weir, K. R. Piller & N. R. Lovejoy. 2013. Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analysis and molecular phylogenetics of New Word silversides (Atherinopsidae). *Evolution* 67(7): 2040 – 2057.

Bradbury, J. P. 2000.Limnologic history of Lago de Patzcuaro, Michoacan, Mexico for the past 48000 years: Impacts of climate and man. *Palaeogeography, Palaeoclimatology, Palaeoecology* 163(1–2): 69-95.

Corander J, P. Waldmann & M.J. Sillanpaa. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics* 163: 367–374.

Corander J., P. Waldmann, P. Marttinen & M.J. Sillanpaa. 2004. BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* 20: 2363–2369.

Doadrio, I. & O. Domínguez. 2004. Phylogenetic relationships within the fish family Goodeidae based on cytocrome b sequence data. *Molecular Phylogenetics and Evolution* 31 (2): 416-430.

Domínguez–Domínguez, O., F. Alda, G. Pérez–Ponce de León, J.L. García–Garitagoitia & I. Doadrio. 2008. Evolutionary history of the endangered fish *Zoogoneticus quitzeoensis* Bean, 1898 (Cyprinodontiformes: Goodeidae) using a sequential approach to phylogeography base on mitochondrial and nuclear DNA data. BMC *Evolutionary Biology* 8: 161.

Drummond, A. J. & A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC *Evolutionary Biology* 7: 214. doi: 10.1186/1471-2148-7-214.

Drummond, A. J., A. Rambaut, A. Shapiro & O. G. Pybus. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* 22 (5): 1185-1192.

Dyer, B. S. & B. Chernoff. 1996. Phylogenetic relationships among atheriniform fishes (Teleostei, Atherinomorpha). *Zoological Journal Linnean Society, London* 117: 1-69.

Echelle, A. A. & A. F. Echelle. 1984. Evolutionary genetics of a species flock: Aterinid fishes on the Mesa Central of México. In: Evoloution of fish species flocks edited by: Echelle A. A. & I. Kornfield. Orono, University of Maine, pp. 93-110.

Excoffier, L. & H. E. L. Lischer. 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.

Excoffier, L. & S. Schneider. 1999. Why hunter-gatherer populations do not show sign of Pleistocene demographic expansions. *Proceding of the Naturalist Academy of Sciences*. USA 96 (10): 10597-10602.

Excoffier, L., P. E. Smouse & J. M. Quatro. 1992. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human mitochondrial DNA Restriction Data. *Genetics* 131 (2): 479-491.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39 (4): 783-791.

Ferrari, L., M. López-Martínez, G. Aguirre-Díaz& G. Carrasco-Núñez. 1999. Space-time patterns of Cenozoic arc volcanism in central México: from the Sierra Madre Occidental to the Mexican Volcanic Belt. *Geology* 27 (4): 303-306.

Frankham, R. 2005. Genetics and extinction. *Biological conservation* 126: 131-140.

Fu, Y. X. 1997. Statistical test of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147 (2):915-925.

Gotoh, R. O., S. N. Chiba, T. V. Goto, H. B. Tamate & N. Hanzawa. 2011. Population genetic structure of the striped silverside, *Atherinomorus endrachtensis* (Atherinidae, Atheriniformes, Teleostei), inhabiting marin lakes and adjacent lagoon in Palau: marine lakes are "Islands" for marine species. *Genes and Genetic Systems* 86: 325-337.

Grant, B. R. & B. W. Bowen. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons of conservation. *The Journal of Heredity* 89 (5):415-426.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41:95-98.

Harpending, H. C. 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* 66 (4): 591-600.

Hewitt, G. 1996. Some genetic consequences of ice ages and their role in divergence and speciation. *Biological Journal of the Linnean Society* 58: 247-276.

Israde-Alcántara, I. & V. H. Garduño-Monroy. 1999. Lacustrine record in a volcanic intra-arc setting: the evolution of late Neogene Cuitzeo basin system (central-western México, Michoacán). *Palaeogeography, Palaeoclimatology, Palaeoecology* 151: 209-227.

Israde-Alcántara, I. 1997. Neogene diatoms of cuitzeo lake central sector of the Trans-Mexican Volcanic Belt and their relationship with the volcano-tectonic evolution. *Quaternary International* 43/44: 137-143.

Librado, P. & J. Rozas. 2009. DnaSp ver. 5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25 (11): 1451-1452.

Lyons J., G. González-Hernández, E. Soto-Galera & M. Guzmán-Arroyo. 1998. Decline of Freshwater fishes and fisheries in selected drainages of west-central Mexico. *Fisheries Managements* 23(4): 10-18.

Marjoram, P. & P. Donnelly. 1994. Pairwise comparisons of mitocondrial DNA sequences in subdivided populations and implications for early human evolution. *Genetics* 136 (2): 673 – 683.

Medina Nava, M. 1997.Ictiofauna de la subcuenca del Río Angulo cuenca del Lerma-Chapala, Michoacán. *Zoología Informa* 35: 25-52.

Miller, R. R. 1986. Composition and derivation of the freshwater fish fauna of México. *Anales de la Escuela Nacional de Ciencias Biológicas* 30: 121-155.

Miller, R. R. & M. L. Smith. 1986. Origin and geography of the fishes of Central México. *In*: C.H. Hocutt & E.O. Wiley (Eds.). *Zoogeography of North American Freshwater Fishes*. John Wiley & Sons, New York, pp. 487-517.

Miller, R. R., W. L. Minckley & S. M. Norris. 2005. *Freshwater Fishes of México*. The University of Chicago Press. Chicago. 490 p.

Moncayo-Estrada, R., I. Israde-Alcántara & V. H. Garduño-Monroy. 2001. La cherehuita *Hubbsina Turneri* De Buen (1941) (Pisces: Goodeidae). Origen, distribución y uso de la regionalización de la cuenca del Lerma. *Hidrobiológica* 11(1): 1-18.

Mulcahy, D. G. & J. R. Mendelson. 2000. Phylogeography and speciation of the morphologically variable, widespread species *Bufo valliceps*, based on molecular evidence from mtDNA. *Molecular Phylogenetics and Evolution* 17 (2): 173-189.

Official Mexican Standar NOM-059-Semarnat-2001.Diario Oficial de la Federación. Norma Oficial Mexicana, México. Diciembre. 77 p.

Paulo-Maya, J., G. Figueroa Lucero & M. Soria-Barreto. 2000. Peces dulceacuícolas mexicanos XIX. *Chirostoma humboldtianum* (Atheriniformes: Atherinopsidae). *Zoología Informa* 43: 59-74.

Pérez-Ramírez, M. Y. 2005.Discriminación morfológica, molecular y determinación de la variación genética de peces blancos(Pisces: Atherinopsidae).Tesis de Maestría en Biología, División de Ciencias Biológicas y de la Salud, UAM Iztapalapa, México. 88 p.

Posada, D. & K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.

Rambaut, A. & A. J. Drummond. 2009. Tracer v1.5. Available from http://tree.bio.ed.ac.uk/software/tracer/

Ray, N., M. Currat & L. Excoffier. 2003. Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology and Evolution* 20 (1): 76 – 86.

Rogers, A. & H. Harpending. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9 (3): 552-559.

Rosas, M. M. 1976. Datos biológicos de la Ictiofauna del Lago de Pátzcuaro con especial énfasis en la alimentación de sus especies. *Memorias del simposio sobre pesquerías en aguas continentales.* Tuxtla Gutiérrez, Chis. Tomo II. *Instituto Nacional de Pesca*. México, pp. 299-366.

Saitou, N. & M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4 (4):406-25.

Saeed, B., W. Ivantsoff & L. E. L. M. Crowley. 1994. Systematic relationships of atheriniform families within Division I of the Series Atherinomorpha (Acanthopterygii) with relevant historical perspectives. *Journal of Ichthyology* 34 (9): 27-72.

Soto-Galera, E., E. Díaz-Pardo, E. López-López & J. Lyons. 1998. Fish as indicators of environmental quality in the Río Lerma Basin, México. *Aquatic Ecosystem Health and Management* 1: 267-276.

Sowfford, D. L. 1998. *PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods)*. Ver. 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Tamayo, J. L. & R. C. West. 1964. Surface Configuration and Associated Geology of Middle America. *In*: Wauchope R. & R.C. West (Eds.). *Handbook of Middle American Indians.* Univ. Texas Press, Austin, pp. 84-121.

Tamura, K. & M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* 10 (3): 512-526.

Tajima, F. 1989. Statistical methods for testing the neutral hypothesis by DNA polymorphism. *Genetics* 123 (1): 253-262.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin & D. G. Higgins. 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 24: 4876-4882.

Vázquez-Domínguez, E., A. Hernández-Valdés, A. Rojas-Santoyo & L. Sambrano. 2009. Contrasting genetic structure in two codistributed freshwater fishes species of highly seasonal systems. *Revista Mexicana de Biodiversidad* 80: 181-192.

Waples, R. S. 1995. Evolutionary significant units and the conservation of biological diversity under the endangered species act. *American Fisheries Society Symposium* 17: 8-27.

Ward, R. D., M. Woodwark & D. O. F. Skibinski. 1994. A comparison of genetic diversity levels in marine, freshwater, and anadramous fishes. *Journal Fish Biology* 44: 213-232.

Webb, S. A., J. A. Graves, C. Macias-Garcia, A. E. Magurran, D. Ó Foighil & M. G. Ritchie. 2004. Molecular phylogeny of the livebearing Goodeidae (Cyprinodontiformes). *Molecular phylogenetics and Evolution* 30: 527-544.

Zane, L., S. Marcato, L. Bargelloni, E. Bortolotto, C. Papetti, M. Simonato, V. Varotto & T. Patarnello. 2006. Demographic history and population structure of the Antartic silverfish *Pleuragramma antarticum. Molecular Ecology* 15 (14): 4499-4511.

*Recibido:*21 octubre 2013.

*Aceptado:*29 septiembre 2014.