Genetic diversity shaped by historical and recent factors in the live-bearing twoline skiffia *Neotoca bilineata*

C. P. Ornelas-García*†‡, F. Alda*§, E. Díaz-Pardo∥, A. Gutiérrez-Hernández∥ and I. Doadrio*

*Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales, CSIC, José Gutiérrez Abascal 2, Madrid, Spain, †Centro Tlaxcala de Biología de la Conducta, Universidad Autónoma de Tlaxcala, Tlaxcala, Mexico, §Smithsonian Tropical Research Institute, Apartado 0843-03092 Balboa, Ancón, Republic of Panama and ∥Laboratorio de Biología Acuática, Licenciatura en Biología, Facultad de Ciencias Naturales, Universidad Autónoma de Querétaro, Querétaro, Mexico

(Received 20 July 2011, Accepted 4 August 2012)

The endangered twoline skiffia *Neotoca bilineata*, a viviparous fish of the subfamily Goodeinae, endemic to central Mexico (inhabiting two basins, Cuitzeo and Lerma-Santiago) was evaluated using genetic and habitat information. The genetic variation of all remaining populations of the species was analysed using both mitochondrial and microsatellite markers and their habitat conditions were assessed using a water quality index (*I*~WQ~). An 80% local extinction was found across the distribution of *N. bilineata*. The species was found in three of the 16 historical localities plus one previously unreported site. Most areas inhabited by the remaining populations had *I*~WQ~ scores unsuitable for the conservation of freshwater biodiversity. Populations showed low but significant genetic differentiation with both markers (mtDNA *φ*~ST~ = 0·076, *P* < 0·001; microsatellite *F*~ST~ = 0·314, *P* < 0·001). Borbollon, in the Cuitzeo Basin, showed the highest level of differentiation and was identified as a single genetic unit by Bayesian assignment methods. Rio Grande de Morelia and Salamanca populations showed the highest genetic diversity and also a high migration rate facilitated by an artificial channel that connected the two basins. Overall, high genetic diversity values were observed compared with other freshwater fishes (average *N*~a~ = 16 alleles and loci and mean ± s.d. *H*~o~ = 0·63 ± 0·10 and nucleotide diversity *π* = 0·006). This suggests that the observed genetic diversity has not diminished as rapidly as the species’ habitat destruction. No evidence of correlation between habitat conditions and genetic diversity was found. The current pattern of genetic diversity may be the result of both historical factors and recent modifications of the hydrological system. The main threat to the species may be the rapid habitat deterioration and associated demographic stochasticity rather than genetic factors.

INTRODUCTION

Predicting whether a species is likely to become extinct is a complex issue that requires information regarding its genetic attributes as well as the ecological and evolutionary properties of its populations (Lande, 1988; Frankham, 2005). As a
general rule, it is assumed that reduced genetic diversity limits the capacity of a population to respond to environmental change and could in turn lead to the expression of deleterious alleles through inbreeding depression, increasing the risk of extinction (Saccheri et al., 1998; Westemeier et al., 1998; Brook et al., 2002; Reed & Frankham, 2003; Spielman et al., 2004; Frankham, 2005). Some authors, however, propose that ecological or demographic factors can drive threatened species to extinction before genetic factors have time to affect them (Lande, 1988; Elgar & Clode, 2001). This demographic stochasticity can also influence the rate of genetic diversity loss (Lippe et al., 2006), which itself depends on life history, mating systems and lifespan of the species (Avise et al., 2002; England et al., 2003). Hence, the demographic stochasticity is determined mainly by the destruction and loss of natural habitats, which is currently considered the greatest threat to the conservation of biodiversity (Lin & Liu, 2006).

In North America, Mexico is the region with the greatest freshwater fish diversity, including c. 506 species, 200 of which are endemic (Miller, 2005; Lévéque et al., 2008). One important region for conservation is the basin of the Lerma-Santiago River, which flows east to west from the Mesa Central of Mexico (Fig. 1), where nearly 70% of its fauna is endemic (SEMARNAT, 2002), including the live-bearing species of the fish subfamily Goodeinae. The Lerma-Santiago Basin, with the Cuitzeo system, is also one of the most populated and industrialized areas of Mexico and has consequently suffered serious degradation mainly from the industrial and urban wastewater discharged along the main river channel (Sedeno-Díaz & López-López, 2007).

Goodeinae comprise between 37 and 41 species of 16 to 19 genera (Doadrio & Domínguez, 2004; Miller, 2005), including twoline skiffia *Skiffia bilineata* (Bean 1887), although phylogenetic studies have not recovered the species as monophyletic within the genus *Skiffia* (Doadrio & Domínguez, 2004). Moreover, different molecular markers have shown high genetic differentiation between *S. bilineata* and other members of the genus *Skiffia*, similar or even higher than the differentiation between genera of Goodeinae (*mt-coI* > 10% *p*-distances and *mt-cyb* > 10% *p*-distances, between species within *Skiffia*) (Doadrio & Domínguez, 2004; Webb et al., 2004). Therefore, the species is recognized in its original genus *Neotoca* (Doadrio & Domínguez, 2004), and is referred to as twoline skiffia *Neotoca bilineata* (Bean 1887). Further studies, including morphology and a more extensive molecular analysis, are necessary to elucidate the taxonomic status of this genus and will contribute to a better understanding of the evolutionary history of the group.

The species *N. bilineata* is characterized, among goodeins, by its small size, reaching 50 mm standard length (*L*_S) in males and 60 mm *L*_S in females, a marked sexual dimorphism and a male:female sex ratio of 1:4. In captivity, female reproductive recruitment is achieved at 30 mm *L*_S with a single annual reproduction period from May to July and fecundity of five to 20 litters per female. A negative relationship between temperature and reproduction has been described, with the optimum temperature in the range of 19–22°C (SEMARNAT, 2002).

Currently, *N. bilineata* is reported in a few localities of the Lerma River and Cuitzeo system, in the middle Lerma-Santiago Basin and in the Rio Grande de Morelia (Fig. 1). The remaining populations of *N. bilineata* are fragmented and scarce and occupy areas with severe habitat disturbances. The outlook for the conservation of *N. bilineata* is not favourable, and it is considered endangered (SEMARNAT, 2002).
Among its main threats, as for other goodein species, are discharge of industrial and urban wastewaters, the introduction of exotic species and the loss of freshwater habitat occurring in the Lerma-Santiago Basin during the past 25 years (Lyons et al., 1995; SEMARNAT, 2002; De La Vega-Salazar et al., 2003; Valero et al., 2008).

The genetic diversity and conservation status of goodeins have been the topic of recent studies, because of their precarious situation. For some species, these works have revealed genetic bottlenecks and direct relationship between habitat loss and genetic diversity (Ritchie et al., 2005; Bailey et al., 2007; Domínguez-Domínguez et al., 2007; Hamill et al., 2007). This situation may also be expected for
N. bilineata, considering its currently reduced and isolated populations that may have been affected by habitat disturbances. Thus, it was hypothesized that if the remaining populations of N. bilineata have suffered recent reductions and fragmentation, a pattern of restricted gene flow and genetic drift should be observed. Adverse habitat conditions for the remnant populations could contribute to the decline of the species.

The main goals of this study were to examine and identify the historical and recent events shaping the genetic diversity of the remaining populations of N. bilineata and to determine the factors representing the greatest threat to the conservation of the species.

MATERIALS AND METHODS

SAMPLE COLLECTION AND DNA EXTRACTION

Historical records of the occurrence of N. bilineata were obtained from the Instituto Politécnico Nacional (IPN) and Universidad Autónoma de Querétaro (UAQ). To determine the current distribution of N. bilineata and assess environmental threats, 17 localities were surveyed, 16 of which were included in the historical record.

Survey sites in the Cuitzeo system and Lerma-Santiago Basin (Fig. 1) were sampled at least twice from August 2005 to March 2006. Seines (2 m × 10 m with a mesh of 0.5 cm) and electrofishing with the aid of hand-nets (50 cm × 60 cm with a mesh of 0.3 cm) were used to collect N. bilineata and to obtain a representation of the fish community at all sampling sites. Fishes were captured and released unharmed with the exception of exotic species, which were removed and deposited in the fish collection of the UAQ. Fin clips were obtained from 144 specimens of N. bilineata. Sample sizes ranged from 30 to 39 individuals per locality (Table I). Total cellular DNA was isolated from all samples by standard proteinase K and by phenol or chloroform extraction (Sambrook et al., 1989).

Habitat characterization

In order to assess the potential risk of extinction of the species, water conditions were evaluated in three of the four localities where the species was found (Table I). Water quality data were obtained from the National Water Commission data bank (Table SI, Supporting Information). Although no data were available for Borbollon, an analysis of the other three populations was conducted. Borbollon is a spring used for human consumption and, consequently, the potential risk of extinction at this site is not expected to be related to water quality. Data were obtained from years coincident with the sampling period, 2005–2007.

The status of the aquatic habitat was evaluated by estimating the multiplicative weighted water quality index ($I_{WQ}$) proposed by Dinius (1987): $I_{WQ} = \prod_{i=1}^{n} I_{Wi}^W$, where $I_{WQ}$, water quality index, is a number ranging from 0 to 100; $I_i$, subindex of variable, is an integer between 0 and 100; $W_i$, unit weight of variable, ranging from 0 to 1; and $\sum_{i=1}^{n} W_i = 1$, where $n$ = number of variables.

Nine variables were considered for water quality characterization: dissolved oxygen (% of saturation), nitrate (mg l$^{-1}$), alkalinity (mg l$^{-1}$ CaCO$_3$), hardness (mg l$^{-1}$ CaCO$_3$), chloride (mg l$^{-1}$), conductivity (μS cm$^{-1}$), pH, total temperature [the difference between ambient ($T_a$) and water ($T_s$) temperature] and colour (Pt-Co units). Because three of the 12 variables originally included by Dinius (1987) were not available, it was necessary to re-scale the weight of the variables (Table SII, Supporting Information). The individual annual subindices were calculated using the mean annual values of each variable. MANOVA and Tukey post hoc tests were performed to compare $I_{WQ}$ scores among localities and years.

MANOVA analyses were used to compare physicochemical variables among localities and years after log$_{10}$ normalization of the following variables: NO$_3$, chloride, conductivity and colour. All calculations were carried out in Statistica v6 (Statistica; www.statsoft.com).
Table I. *Neotoca bilineata* populations and main biotic and physicochemical monitoring stations

<table>
<thead>
<tr>
<th>Population</th>
<th>Basin</th>
<th>Type of habitat</th>
<th>Habitat modifications</th>
<th>Exotic species present</th>
<th>Monitoring station Code of the NWC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borbollon (BOR)</td>
<td>Cuitzeo</td>
<td>Spring</td>
<td>Water used for human consumption</td>
<td>None</td>
<td>No data available</td>
</tr>
<tr>
<td><em>n</em> = 36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rio Grande de Morelia (RGM)</td>
<td>Cuitzeo</td>
<td>River</td>
<td>Urban wastewater spills</td>
<td><em>Oreochromis mossambicus</em>, <em>Cyprinus carpio</em></td>
<td>00MI12GA0030001</td>
</tr>
<tr>
<td><em>n</em> = 39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rio Turbio Corralejo (RTC)</td>
<td>Lerma-Santiago</td>
<td>River</td>
<td>Agriculture and cattle raising</td>
<td><em>O. mossambicus</em></td>
<td>00GU12BE0310003</td>
</tr>
<tr>
<td><em>n</em> = 30</td>
<td></td>
<td></td>
<td>Distilling Industry spills</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salamanca (SAL)</td>
<td>Lerma-Santiago</td>
<td>River</td>
<td>Refinery wastewater spills</td>
<td><em>O. mossambicus</em>, <em>Xiphophorus variatus</em>, <em>Poecilia reticulata</em></td>
<td>00GU12HC0270002</td>
</tr>
<tr>
<td><em>n</em> = 39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*n*, sample size; NWC, National Water Commission.
mtDNA sequencing and analysis

DNA samples from 48 individuals representing all populations were amplified for two overlapping fragments of the mitochondrial cytochrome b gene (total of 1140 pb) via PCR. The primers used for cytochrome b were those described in the study by Machordom & Doadrio (2001). PCR mixture and conditions were conducted as described by Doadrio & Domínguez (2004). PCR products were checked in 1% agarose gels and sequenced using an ABI PRISM 3700 DNA analyser (www.secugen.es). Chromatograms and alignments were visually examined and verified. All sequences were deposited in GenBank under accession numbers JQ769330–JQ769376.

Population genetic statistics such as the number of polymorphic sites $S$, haplotype diversity $H_d$ (Nei, 1987), nucleotide diversity $\pi$ (Nei, 1987) and the average number of pair-wise nucleotide differences $k$ (Tajima, 1983) were calculated using DnaSP 4.0 (Rozas et al., 2003).

Phylogenetic reconstruction of the sequences obtained was performed by neighbour-joining (NJ) analysis with transversion model (TVM) $+\Gamma^I$ distances, which was selected by ModelTest 3.07 under the AIC as the model that best fitted the data (Posada & Crandall, 1998). Robustness of NJ was assessed by bootstrapping, with 1000 replicates in PAUP 4.0b10 (Swofford, 1998). Sequence data were also analysed by Bayesian inference using Mr. Bayes v. 3.1.2 software (Huelsenbeck & Ronquist, 2001). The Bayesian tree was inferred by eight Markov chain Monte-Carlo (MCMC) for 5 million generations sampling every 100 trees. Posterior probability values were used as support for the Bayesian topology after burn-in of 10% of the generations to ensure stability of the likelihood values.

To test specific phylogeographical hypotheses underlying the observed haplotype distribution pattern, a haplotype network analysis based on mutational differences and a nested cladistic analysis (NCA) (Templeton, 1998, 2004) were conducted. The most parsimonious un-rooted haplotype network was estimated using TCS 1.18 (Clement et al., 2000). The resulting cladogram was converted to a nested clade design following the method described by Templeton et al. (1987) employing the special modifications of these rules (Templeton & Sing, 1993). To check for significant association between clades and geographical locations, nested contingency analyses (Templeton & Sing, 1993) were performed using the programme GeoDis 2.2 (Posada et al., 2000). The AUTOINFER 1.0 (Zhang et al., 2006) software package was then used to infer the most suitable population structure model and historical scenario for the observed geographical associations.

Pair-wise $\phi_{ST}$ values were calculated among all populations as an estimate of genetic differentiation. This parameter was calculated using TrN (Tamura & Nei, 1993) with gamma correction distances, as this was the available model with the highest probability under the AIC. AMOVA was performed (Excoffier et al., 1992) to test the significance of genetic differentiation under varying fragmentation scenarios, based on the current geographical distribution of the species in the two river basins (Cuitzeo and Lerma-Santiago) and on the genetic clusters inferred by the Bayesian assignment method for the microsatellite data. Correlation between riparian geographic distances (i.e. along currently existing ways) and genetic distances was determined using a Mantel test (10 000 permutations) (Mantel, 1967). All analyses were performed using Arlequin 3.1 (Schneider et al., 2000).

Microsatellite analysis

Five microsatellite loci (ZT1.2, ZT1.6, ZT1.7, ZT1.9 and ZT1.43) were amplified in 144 N. bilineata individuals using primers and conditions described elsewhere (Boto & Doadrio, 2003) with minor modifications. All PCR products were analysed using an ABI PRISM 3700 DNA analyser and genotyped using GeneScan 3.5.1 and GenoTyper 3.6 (Applied Biosystems; www.appliedbiosystems.com).

Data were examined for null alleles, stuttering and allele dropouts using DROPOUT (McKelvey & Schwartz, 2005). Linkage disequilibria within populations were tested for each pair of loci using GenePop 3.4 (Raymond & Rousset, 1995). Deviations from Hardy–Weinberg equilibrium were also assessed by applying the exact tests (Guo & Thompson, 1992) in GenePop 3.4. The number of alleles ($N_a$), allelic richness ($A_R$), observed and expected heterozygosities ($H_o$ and $H_e$) and $F_{IS}$ were calculated using FSTAT 2.9.3 (Goudet, 1995).

To check for recent population bottlenecks, two methods were used. First, two statistic indices were determined: $\Delta H$, which is the difference between gene diversity ($H_e$) of the
sample and the heterozygosity expected for a population in Wright–Fisher equilibrium \( (H_A) \) (Cornuet & Luikart, 1996), and \( M \), which is the ratio of the number of alleles over the range in allele size (Garza & Williamson, 2001). Following a population reduction, a positive \( \Delta H \) would be expected (i.e. heterozygosity excess), as allelic diversity is reduced faster than gene diversity. \( M \) should decrease since, under genetic drift, alleles are lost more rapidly than the range in allele size at a locus. Garza & Williamson (2001) demonstrated both empirically and using simulated data that \( M \) values were always >0.68 for populations in Wright–Fisher equilibrium. BOTTLENECK 1.2.02 (Piry et al., 1999) was used to calculate \( \Delta H \), employing the two-phased model with the 90% stepwise mutation model, and the Wilcoxon rank test to assess its significance. AGARst (Harley, 2001) was used to calculate \( M \).

Genetic structure was assessed through pair-wise \( F_{ST} \) and AMOVA as explained for the mtDNA data, in Arlequin 3.1. Recent migration among populations was investigated using BayesAss 1.3 (Wilson & Rannala, 2003). To ensure convergence of the results, 10 runs of \( 9 \times 10^6 \) iterations, including a burn-in of \( 3 \times 10^6 \) iterations, were conducted, with a sampling frequency of 2000. Delta values were tuned individually to obtain acceptance rates within 40 and 60% of the total. The effect of geographic distance along hypothesized waterways over the genetic distance, calculated as \( F_{ST}(1 - F_{ST})^{-1} \), was assessed by a Mantel test with 10 000 permutations using Arlequin 3.1.

Because of the uncertainty that the geographical assignment of individuals to populations will represent biologically significant entities, a Bayesian analysis was conducted using the programme Structure 2.1. Ten independent runs from \( K = 1–8 \) populations were conducted, assuming correlated allele frequencies and an admixture model (Falush et al., 2003). Putative population information was not included in the analysis. For each value of \( K \), MCMC was run with a burn-in period of \( 5 \times 10^5 \) steps and chain length of \( 1 \times 10^6 \). Multiple runs were performed to assess convergence of \( \ln L \). Mean log probabilities were used to calculate \( \Delta K \) (i.e. a quantity based on the second-order rate of change of the log probability of data between successive \( K \) values) and to find the true \( K \), following the method of Evanno et al. (2005).

## RESULTS

### POPULATION STATUS AND HABITAT MODIFICATIONS

Although \( N. \) bilineata is considered relatively tolerant to habitat disturbance (SEMARNAT, 2001), the species was found in only three of the 16 localities where it had been previously reported. This translates to 80% local extinction in the past 30 years, although some reports suggest that this rate could be lower, c. 60% (De La Vega-Salazar, 2006). The remaining populations are distributed across two hydrological systems, the Cuitzeo and middle Lerma-Santiago Basins, each of which includes two populations: Borbollon (BOR) and Rio Grande de Morelia (RGM) in the Cuitzeo system and Rio Turbio Corralejo (RTC) and Salamanca (SAL) in the Lerma-Santiago Basin (Table I). The SAL population is a new record of the species in Salamanca city (Fig. 1).

According to the historical and recent records reviewed (CONABIO, IPN, UNAM, UMSNH in Morelia and UAQ), the species is not currently present in the Chapala Basin, where it was reported by Miller (2005). This is the only record in the literature for \( N. \) bilineata in Chapala; thus, it cannot be confirmed nor determined if this constitutes another local extinction or a possible misidentification of the species.

As for other freshwater fauna of the region, the current critical situation of \( N. \) bilineata is mainly the consequence of a greatly deteriorated water quality and modifications to river flow. Environmental changes primarily consist of over-extraction of water for agricultural purposes as well as extensive habitat transformation into recreational areas. Other localities, such as Cuitzeo Lake, where the species

© 2012 The Authors
was not found, RTC and RGM have suffered serious deterioration of water quality related to the high levels of eutrophication in recent years (Alfaro et al., 2002; M. Velázquez-Bucio, unpubl. data).

The means ± s.d. of $I_{WQ}$ scores per locality were 42.27 ± 4.31, 57.40 ± 3.95 and 63.12 ± 6.73 for RTC, RGM and SAL, respectively (Table II). The MANOVA analysis showed significant differences among localities ($F_{2,4} = 10.846$, $P < 0.05$). The post hoc mean comparisons (Tukey’s HSD test) showed significant differences between RTC and SAL ($P = MS = 22.949$, d.f. = 4, $P < 0.05$). In contrast, the $I_{WQ}$ values were not significantly different between years ($P > 0.05$).

According to the general rating scale for water quality from Dinius (1987), the $I_{WQ}$ values obtained were classified as not acceptable for a public water supply, doubtful for recreational uses, in need of extensive treatment for most agriculture and industrial uses and suitable for tolerant fish fauna. Additionally, MANOVA analysis of the physicochemical variables showed highly significant differences among localities ($F_{2,4} = 8.96$, $P < 0.001$) and among years ($F_{2,4} = 2.106$, $P < 0.05$).

Exotic species were observed in all the localities where $N. bilineata$ was found, with the exception of BOR, and constituted the dominant element. The exotic species collected were goldfish $Carassius auratus$ (L. 1758), common carp $Cyprinus carpio$ L. 1758, tilapia $Oreochromis mossambicus$ (Peters 1852) and ornamental poeciliids: green swordtail $Xiphophorus helleri$ Heckel 1848, platyfish $Xiphophorus variatus$ (Meek 1904) and guppy $Poecilia reticulata$ Peters 1859 (Table I).

mtDNA phylogenetic and diversity analyses

The entire sequence of the cytochrome ($cyt$) $b$ gene was obtained for 48 individuals from four localities. The total alignment comprised 1140 bp, of which 68 bp were variable and 21 bp were parsimony informative. Overall nucleotide diversity ($\pi$) was 0.006 (Table III).
Table III. Genetic variation of *Neotoca bilineata* at five microsatellite loci and mitocondrial cytochrome *b* (*mt-cyb*)

<table>
<thead>
<tr>
<th>Population</th>
<th>Locus</th>
<th>ZT1.2</th>
<th>ZT1.6</th>
<th>ZT1.7</th>
<th>ZT1.43</th>
<th>ZT1.9</th>
<th>ALL ± s.d.</th>
<th>mt-cyb (mtDNA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGM</td>
<td>(N_a)</td>
<td>11</td>
<td>26</td>
<td>11</td>
<td>34</td>
<td>28</td>
<td>22,000 ± 10,464</td>
<td>n/h</td>
</tr>
<tr>
<td>(n = 39)</td>
<td>(A_R)</td>
<td>8.183</td>
<td>22.012</td>
<td>9.333</td>
<td>14.326</td>
<td>14.901</td>
<td>13.751 ± 5.486</td>
<td>(S)</td>
</tr>
<tr>
<td></td>
<td>(H_o)</td>
<td>0.615</td>
<td>0.750</td>
<td>0.487</td>
<td>0.710</td>
<td>0.948</td>
<td>0.702 ± 0.171</td>
<td>(H_d) ± s.d.</td>
</tr>
<tr>
<td></td>
<td>(H_e)</td>
<td>0.650</td>
<td>0.944</td>
<td>0.542</td>
<td>0.951</td>
<td>0.953</td>
<td>0.808 ± 0.197</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(F_{IS})</td>
<td>0.067</td>
<td>\textbf{0.219}</td>
<td>0.114</td>
<td>\textbf{0.266}</td>
<td>0.017</td>
<td>\textbf{0.144 ± 0.104}</td>
<td>(\pi) ± s.d.</td>
</tr>
<tr>
<td>BOR</td>
<td>(N_a)</td>
<td>3</td>
<td>17</td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>8.200 ± 5.805</td>
<td>n/h</td>
</tr>
<tr>
<td>(n = 36)</td>
<td>(A_R)</td>
<td>2.639</td>
<td>14.905</td>
<td>3.512</td>
<td>5.672</td>
<td>10.327</td>
<td>7.411 ± 5.139</td>
<td>(S)</td>
</tr>
<tr>
<td></td>
<td>(H_o)</td>
<td>0.500</td>
<td>0.485</td>
<td>0.444</td>
<td>0.575</td>
<td>0.666</td>
<td>0.534 ± 0.087</td>
<td>(H_d) ± s.d.</td>
</tr>
<tr>
<td></td>
<td>(H_e)</td>
<td>0.511</td>
<td>0.898</td>
<td>0.447</td>
<td>0.725</td>
<td>0.742</td>
<td>0.665 ± 0.184</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(F_{IS})</td>
<td>0.036</td>
<td>\textbf{0.471}</td>
<td>0.020</td>
<td>\textbf{0.221}</td>
<td>0.118</td>
<td>\textbf{0.210 ± 0.184}</td>
<td>(\pi) ± s.d.</td>
</tr>
<tr>
<td>RTC</td>
<td>(N_a)</td>
<td>4</td>
<td>25</td>
<td>7</td>
<td>6</td>
<td>19</td>
<td>15 ± 9.027</td>
<td>n/h</td>
</tr>
<tr>
<td>(n = 30)</td>
<td>(A_R)</td>
<td>3.995</td>
<td>23.959</td>
<td>6.843</td>
<td>18.028</td>
<td>12.591</td>
<td>13.083 ± 8.132</td>
<td>(S)</td>
</tr>
<tr>
<td></td>
<td>(H_o)</td>
<td>0.321</td>
<td>0.692</td>
<td>0.633</td>
<td>0.769</td>
<td>0.478</td>
<td>0.578 ± 0.179</td>
<td>(H_d) ± s.d.</td>
</tr>
<tr>
<td></td>
<td>(H_e)</td>
<td>0.456</td>
<td>0.950</td>
<td>0.664</td>
<td>0.907</td>
<td>0.921</td>
<td>0.779 ± 0.214</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(F_{IS})</td>
<td>0.312</td>
<td>\textbf{0.290}</td>
<td>0.063</td>
<td>\textbf{0.171}</td>
<td>0.497</td>
<td>\textbf{0.276 ± 0.163}</td>
<td>(\pi) ± s.d.</td>
</tr>
<tr>
<td>SAL</td>
<td>(N_a)</td>
<td>8</td>
<td>24</td>
<td>10</td>
<td>31</td>
<td>27</td>
<td>20 ± 10.368</td>
<td>n/h</td>
</tr>
<tr>
<td></td>
<td>(H_o)</td>
<td>0.568</td>
<td>0.688</td>
<td>0.590</td>
<td>0.711</td>
<td>0.868</td>
<td>0.684 ± 0.119</td>
<td>(H_d) ± s.d.</td>
</tr>
<tr>
<td></td>
<td>(H_e)</td>
<td>0.678</td>
<td>0.936</td>
<td>0.718</td>
<td>0.945</td>
<td>0.942</td>
<td>0.843 ± 0.134</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(F_{IS})</td>
<td>0.176</td>
<td>\textbf{0.280}</td>
<td>\textbf{0.191}</td>
<td>0.260</td>
<td>0.092</td>
<td>\textbf{0.202 ± 0.075}</td>
<td>(\pi) ± s.d.</td>
</tr>
<tr>
<td>ALL</td>
<td>(N_a)</td>
<td>12</td>
<td>39</td>
<td>14</td>
<td>37</td>
<td>34</td>
<td>27.200 ± 13.103</td>
<td>n/h</td>
</tr>
<tr>
<td>(n = 144)</td>
<td>(A_R)</td>
<td>6.074</td>
<td>22.886</td>
<td>8.112</td>
<td>23.024</td>
<td>22.263</td>
<td>16.472 ± 8.596</td>
<td>(S)</td>
</tr>
<tr>
<td></td>
<td>(H_o)</td>
<td>0.514</td>
<td>0.651</td>
<td>0.535</td>
<td>0.689</td>
<td>0.774</td>
<td>0.632 ± 0.108</td>
<td>(H_d) ± s.d.</td>
</tr>
<tr>
<td></td>
<td>(H_e)</td>
<td>0.624</td>
<td>0.954</td>
<td>0.698</td>
<td>0.951</td>
<td>0.947</td>
<td>0.834 ± 0.161</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(F_{IS})</td>
<td>0.132</td>
<td>0.315</td>
<td>0.110</td>
<td>0.236</td>
<td>0.147</td>
<td>\textbf{0.200 ± 0.085}</td>
<td>(\pi) ± s.d.</td>
</tr>
</tbody>
</table>

\(n\), number of individuals analysed; \(N_a\), number of alleles; \(A_R\), allelic richness; \(H_o\), observed heterozygosity; \(H_e\), expected heterozygosity; \(F_{IS}\), inbreeding index; \(n\), number of haplotypes; \(S\), number of polymorphic sites; \(H_d\), gene diversity; \(k\), average pair-wise nucleotide differences.

Significant departures from Hardy–Weinberg equilibrium are shown in bold.
Forty-one haplotypes were found, and all populations except BOR shared haplotypes. The population with the highest number of haplotypes was RGM. Overall haplotype diversity was high, with a mean ± s.d. $H_d = 0.971 ± 0.043$. The RTC population showed the lowest $H_d$ value. The population with the highest number of polymorphic sites was RGM followed by SAL ($S = 34$ and 20, respectively; Table III).

Independent phylogenetic analyses of the mitochondrial data, based on NJ and Bayesian inference, yielded similar topologies where four groups were recovered (Fig. 2). Groups I and IV comprised three or all four populations, while the other
groups were more restricted geographically. Group II included only the populations of RGM and SAL and group III only individuals from RTC (Fig. 2).

NCA generated two four-step clades. One contained exclusively haplotypes belonging to populations in the Cuitzeo system (clade 4-1), and the other included haplotypes from both the Cuitzeo system and the middle Lerma-Santiago Basin (clade 4-2). Haplotypes from BOR appeared only in clade 4-1 (Fig. 3). The analysis implemented in GEODIS rejected the null hypothesis of non-geographical association of the haplotypes in two clades: clade 3-5 ($\chi^2 = 23.415, P < 0.001$) and clade 4-2 ($\chi^2 = 17.199, P < 0.01$). In both cases, the GEODIS inference chain suggested restricted gene flow and isolation by distance.

Genetic structure and population differentiation

The AMOVA analysis for the mtDNA data detected significant genetic structure among all populations ($\phi_{ST} = 0.314, P < 0.001$). When the existing arrangement between Cuitzeo v. Lerma-Santiago Basins (BOR, RGM v. RTC, SAL) was tested, no significant structure was detected (Table IV).

Mitochondrial pair-wise $\phi_{ST}$ values were significant for all comparisons except between RGM and SAL. Highest values were obtained between RTC and BOR sites (Table V). A significant and positive correlation was also detected between genetic and riparian geographic distances based on mtDNA ($r^2 = 0.88, P < 0.05$; Fig. 4).

Nuclear microsatellite loci

Five microsatellite genotypes were obtained for 144 individuals of *N. bilineata*. Overall, no significant linkage disequilibrium was observed ($P > 0.05$ for each pair.
Table IV. Analysis of molecular variance for the mitochondrial and microsatellite markers of *Neotoca bilineata* (see Fig. 1)

<table>
<thead>
<tr>
<th>Structure tested</th>
<th>Variance among groups (%)</th>
<th>$F_{ST}$</th>
<th>$F_{SC}$</th>
<th>$F_{CT}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>nDNA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall (RGM, BOR, RTC, SAL)</td>
<td>10.81</td>
<td><strong>0.108</strong></td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>River basins (RGM, BOR) (RTC, SAL)</td>
<td>−2.79</td>
<td><strong>0.068</strong></td>
<td><strong>0.093</strong></td>
<td>−0.028</td>
</tr>
<tr>
<td>Genetic clusters (BOR) (RGM, RTC, SAL)</td>
<td>5.78</td>
<td><strong>0.103</strong></td>
<td><strong>0.048</strong></td>
<td>0.058</td>
</tr>
<tr>
<td><strong>mtDNA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall (RGM, BOR, RTC, SAL)</td>
<td>31.46</td>
<td><strong>0.314</strong></td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>River basins (RGM, BOR) (RTC, SAL)</td>
<td>19.43</td>
<td><strong>0.358</strong></td>
<td><strong>0.204</strong></td>
<td>0.194</td>
</tr>
<tr>
<td>Genetic clusters (BOR) (RGM, RTC, SAL)</td>
<td>19.94</td>
<td><strong>0.386</strong></td>
<td><strong>0.234</strong></td>
<td>0.199</td>
</tr>
</tbody>
</table>

$F_{ST}$, variation among localities within all; $F_{SC}$, variation among localities within clusters; $F_{CT}$, variation among clusters within all.

BOR, Borbollon; RGM, Rio Grande de Morelia; RTC, Rio Turbio Corralejo; SAL, Salamanca (see Fig. 1).

Significant values are shown in bold ($P < 0.05$).

of loci across the sample). All the populations showed loci with significant departures from the Hardy–Weinberg equilibrium. All were, however, in equilibrium when the populations were considered together (Table III). No evidence of genotyping errors were observed using DROP OUT. Further, repeated failure to amplify alleles at any locus did not occur in any individual, suggesting there were no null alleles (Selkoe & Toonen, 2006).

The genetic variability found in *N. bilineata* was high (Table III). In all populations, all loci were polymorphic and exhibited a total of 136 alleles (mean ± s.d. $N_a = 27.20 ± 13.10$). The number of alleles per locus ranged from $N_a = 3$ for ZT1.2 in BOR to $N_a = 34$ for ZT1.43 in RGM. In terms of mean ± s.d. $H_o$, greatest genetic diversity was found in RGM ($H_o = 0.70 ± 0.17$) with the lowest in BOR ($H_o = 0.53 ± 0.09$).

All the populations showed positive $F_{IS}$ values, denoting a significant heterozygosity deficit (Table III). The Wilcoxon rank test for heterozygosity excess in BOTTLENECK detected no evidence of a recent bottleneck event in any of the

Table V. Pair-wise $F_{ST}$ (below diagonal) and $\phi_{ST}$ (above diagonal) values among populations of *Neotoca bilineata* (see Fig. 1)

<table>
<thead>
<tr>
<th></th>
<th>RGM</th>
<th>BOR</th>
<th>RTC</th>
<th>SAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGM</td>
<td>—</td>
<td><strong>0.188</strong></td>
<td><strong>0.305</strong></td>
<td>0.079</td>
</tr>
<tr>
<td>BOR</td>
<td><strong>0.124</strong></td>
<td>—</td>
<td><strong>0.622</strong></td>
<td>0.407</td>
</tr>
<tr>
<td>RTC</td>
<td><strong>0.079</strong></td>
<td><strong>0.085</strong></td>
<td>—</td>
<td><strong>0.225</strong></td>
</tr>
<tr>
<td>SAL</td>
<td>0.012</td>
<td><strong>0.104</strong></td>
<td><strong>0.056</strong></td>
<td>—</td>
</tr>
</tbody>
</table>

BOR, Borbollon; RGM, Rio Grande de Morelia; RTC, Rio Turbio Corralejo; SAL, Salamanca (see Fig. 1).

Significant $P$ values after Bonferroni correction are shown in bold.
GENETIC DIVERSITY OF *N. BILINEATA* 1975

![Graph showing correlation between riparian geographic distances and genetic distances](image)

Fig. 4. Correlation between riparian geographic distances and genetic distances between all populations of *Neotoca bilineata* (see Fig. 1). The mtDNA \([\phi_{ST}(1 - \phi_{ST})^{-1}] (\bullet)\) and microsatellite (nDNA) \([F_{ST}(1 - F_{ST})^{-1}] (\circ)\) distances were plotted against the geographical distances. mtDNA: the curve was fitted by \(y = 1.388x - 3.159 (r^2 = 0.88, P < 0.05)\). nDNA: the relationship was not significant \((r^2 = 0.22, P > 0.05)\). RGM, Rio Grande de Morelia; SAL, Salamanca (see Fig. 1).

populations. Conversely, the \(M\) ratio was <0.680 in BOR \((M = 0.555)\), RTC \((M = 0.540)\) and marginally lower in RGM \((M = 0.654)\), possibly indicating population bottlenecks.

Pair-wise \(F_{ST}\) values were significant for all population comparisons except between SAL and RGM. Comparisons involving BOR rendered the highest \(F_{ST}\) values (Table V). Overall, the genetic structure of *N. bilineata* was significantly different among populations \((F_{ST} = 0.108, P < 0.001)\). The structure hypotheses tested in the AMOVA analysis, according to the geographical distribution of the populations, and the genetic groups inferred by Structure, however, were not significant (Table IV). No significant correlation was observed between microsatellite-inferred genetic distances and along-river distances \((r^2 = 0.22, P > 0.05; Fig. 4)\).

**Table VI.** Migration rates \((m)\) between *Neotoca bilineata* populations (see Fig. 1) inferred from BayesAss for the microsatellite data. Values in columns are migration rates from right into left populations. Values along the diagonal represent proportions of non-migrant individuals.

<table>
<thead>
<tr>
<th></th>
<th>RGM</th>
<th>BOR</th>
<th>RTC</th>
<th>SAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGM</td>
<td><strong>0.71803633</strong></td>
<td>0.0242333</td>
<td>0.01070277</td>
<td>0.247027</td>
</tr>
<tr>
<td>BOR</td>
<td>0.00559359</td>
<td><strong>0.98493933</strong></td>
<td>0.00385065</td>
<td>0.00561663</td>
</tr>
<tr>
<td>RTC</td>
<td>0.00881736</td>
<td>0.03746977</td>
<td><strong>0.94218933</strong></td>
<td>0.0115235</td>
</tr>
<tr>
<td>SAL</td>
<td>0.04075653</td>
<td>0.03060663</td>
<td>0.01584207</td>
<td><strong>0.91279467</strong></td>
</tr>
</tbody>
</table>

BOR, Borbollon; RGM, Rio Grande de Morelia; RTC, Rio Turbio Corralejo; SAL, Salamanca (see Fig. 1).

Significant values are shown in bold.

© 2012 The Authors
Estimates of recent migration obtained by BayesAss indicated no significant migration between most of the populations (\(m < 0.833, 0.675–0.992\)). The only population that showed a significant migration rate was RGM, which was estimated to have received individuals from SAL at a mean rate of \(m = 0.247\) (Table VI).

The Bayesian analysis implemented in Structure to detect the genetic structure of \(N. \text{bilineata}\) based on deviations from Hardy–Weinberg and linkage equilibriums clearly assigned maximum \(\ln L\) and \(\Delta K\) values to \(K = 2\). Cluster 1 contained 92.2% of the individuals from BOR, whereas cluster 2 contained 83.7 to 93.1% of the individuals from RTC, SAL and RGM.

**DISCUSSION**

**HABITAT MODIFICATIONS AND CONSERVATION STATUS OF THE POPULATIONS OF \(N. \text{bilineata}\)**

The high rate of habitat degradation in the Lerma region presents a challenge for goodein conservation (De La Vega-Salazar *et al.*, 2003; De La Vega-Salazar, 2006; Domínguez-Domínguez *et al.*, 2006, 2007; Hamill *et al.*, 2007). Decrease in numbers of \(N. \text{bilineata}\) reflects habitat loss and fragmentation, where, according to this study, the species has suffered 80% local extinction (Fig. 1). Despite some reports suggesting that this rate could be less pronounced, c. 60% (De La Vega-Salazar, 2006), the species still suffers a high degree of habitat loss, mainly as the outcome of overexploitation, including river drainage for agricultural purposes, habitats converted to recreational areas, degradation by water pollution from anthropogenic sources and increase in eutrophication (Alfaro *et al.*, 2002; De La Vega-Salazar *et al.*, 2003; M. Velazquez-Bucio, unpubl. data.). Exotic species were present in most localities, and particularly in those where \(N. \text{bilineata}\) was not found (e.g. Querendaro River). Hence, it is important to draw attention to the possible negative effects of exotic species on \(N. \text{bilineata}\). Negative interactions between \(P. \text{reticulata}\) and \(N. \text{bilineata}\) (co-occurring in SAL site) as a consequence of courtship and forced copulation attempts have been reported (Valero *et al.*, 2008), along with other examples of negative effects of alien species on native freshwater fish fauna in this region (Mercado-Silva *et al.*, 2002, 2006).

Significant differences in \(I_{WQ}\) scores among population localities, especially for those in the same basin (RTC and SAL), reflected fragmentation of \(N. \text{bilineata}\) habitats. Similarly, when physicochemical variables among populations were compared, RTC was the most differentiated population (RTC and SAL, \(P < 0.05\)). According to the Dinius (1987) scale, \(I_{WQ}\) scores denoted non-suitable conditions for the maintenance of freshwater fish fauna diversity. In general terms, \(I_{WQ}\) scores were associated with land use such as cattle farming and distillery and oil spills in most localities (Fig. 1).

The RTC locality showed a high degree of eutrophication, as evidenced by the highest values of nitrate concentration (127 mg l\(^{-1}\)), which exceeded the maximum permissible by Mexican law (SEMARNAT, 1996), high water temperature (26° C) and anoxic water conditions. Although habitat conditions in SAL were significantly better than in RTC, they were far from optimal for biodiversity conservation (Dinius, 1987) as demonstrated by the low pH and high conductivity values (2.81 and
Furthermore, anoxic water conditions were recorded during the dry season. The SAL sampling point was located near one of the most productive oil refineries in Mexico, and multiple studies have reported its negative effects, which include anoxic conditions in a river stretch of c. 15 km during most of the year (CERCAS, unpubl. data). Such conditions could act as a barrier for population dispersal in the Lerma-Santiago Basin. Rio Grande de Morelia is one of the most important tributaries to Cuitzeo Lake; however, after passing through Morelia, one of the largest cities in Mexico, the $I_{WQ}$ score indicates water quality inadequate for the establishment of fish fauna communities (Table II), registering nearly anoxic conditions, with a mean oxygen concentration of 1.033 mg l$^{-1}$.

Degradation of water quality, such as that observed in these surveyed localities, has been shown to be associated with high mortality rates in fish populations, which may eventually lead to increased demographic stochasticity and likelihood of extinction (Lande, 1988; Frankham & Franklin, 1998; Frankham, 2005). This was supported by some of the bottleneck tests performed, which showed the lowest $M$ value and hence highest evidence of population reduction for RTC, which is also the locality with the most disturbed habitat. The water temperatures exceeding 25°C recorded for most of the surveyed areas are likely to affect species reproduction (SEMARNAT, 2001).

HIGH GENETIC DIVERSITY IN BOTH MT DNA AND MICROSATELLITE MARKERS

A widely debated topic in conservation biology is the relative contribution of genetic and demographic factors to the extinction risk of a species (Lande, 1988, 1995; Frankham, 1996, 2001, 2005; Frankham & Franklin, 1998). It is widely accepted that a species’ genetic diversity and its extinction risk are negatively correlated (Soulé, 1976; Soulé & Wilcox, 1980; Frankham, 2005). Other studies have shown that some species do not necessarily follow this rule; however, demographic stochasticity is the primary determinant of their risk of extinction (Saillant et al., 2004; Ellis et al., 2006; Hamill et al., 2007; Song & Mitchell-Olds, 2007; Vega et al., 2007).

In this study, *N. bilineata* showed high levels of genetic diversity according to both microsatellite and mtDNA markers. mtDNA inferred diversity with a mean of $H_d = 0.971$ and a range of $\pi = 0.003–0.006$. These values are much higher than those reported for other endangered freshwater fishes such as the amago salmon *Oncorhynchus masou* (Brevoort 1856), which has values of $H_d = 0.38–0.68$ and $\pi = 0.0003–0.0013$ (Kawamura et al., 2007), and closer to those recorded for non-endangered freshwater fish populations (Bernardi et al., 2007; Ngamsiri et al., 2007). The population showing the highest mtDNA diversity was RGM in the Cuitzeo system.

Microsatellite data revealed a similar pattern of high genetic and allelic diversity ($H_o = 0.632 \pm 0.108$ and $N_a = 27.2$), higher than values reported for other freshwater fishes ($H_o = 0.46 \pm 0.34$ and $N_a = 7.5$) (DeWoody & Avise, 2000) or the other critically endangered goodeins, bold characodon *Characodon audax* Smith & Miller 1986, tequila splitfin *Zoogoneticus tequila* Webb & Miller 1998 and splitfin *Zoogoneticus quitzeoensis* (Bean 1898) (Boto & Doadrio, 2003; Domínguez-Domínguez et al., 2007, 2008; Hamill et al., 2007). In contrast, the diversity values obtained for *N. bilineata* were similar to those observed for other non-endangered and more
widely distributed species of the Goodeinae subfamily (with ranges of $H_0 = 0.5–0.7$ and $N_a = 14–21$ alleles) (Boto & Doadrio, 2003; Hamill et al., 2007; Domínguez-Domínguez et al., 2008). This is not the first endangered goodein species displaying high levels of genetic variability. The single wild population of butterfly splitfin *Ameca splendens* Miller & Fitzsimmons 1971, described as critically endangered, showed moderate values of genetic diversity ($H_e$ ranging from 0.34 to 0.94 and mean $N_a = 4.1$) (Hamill et al., 2007). High levels of genetic diversity have been reported for other endangered freshwater fishes such as the copper redhorse *Moxostoma hubbsi* Legendre 1952 ($H_o = 0.77 ± 0.08$ and mean $N_a = 12.5$) (Lippe et al., 2006) and the Cape Fear shiner *Notropis mekistocholas* Snelson 1971 ($H_o = 0.703$ and $N_a = 2–14$) (Saillant et al., 2004).

Contrary to what has been proposed for other goodein species (Domínguez-Domínguez et al., 2007), the populations with the least genetic diversity were not those with the lowest habitat quality, according to the $I_{WQ}$. For example, the highest genetic diversity for both types of molecular markers was observed in RGM, while the lowest genetic diversity was found in BOR, which is likely to have the best water quality, this is considered as the only locality with water used for human consumption. This suggests that habitat deterioration is unrelated to genetic diversity and, consequently, that these factors do not contribute in the same way to local extinctions.

Explanations for the high levels of genetic diversity in *N. bilineata* could include a gradual reduction in population size allowing the species to retain high levels of diversity or a long generation time limiting the loss of genetic diversity (Lippe et al., 2006). Migration has also been considered an important source of new alleles, limiting the loss of genetic diversity (Vilà et al., 2002; Vega et al., 2007). Hence, several factors could have determined the current genetic diversity and structure of *N. bilineata*.

**HISTORICAL AND RECENT PROCESSES SHAPING THE GENETIC STRUCTURE AND DIVERSITY OF N. BILINEATA**

Some hypotheses consider the existence during the Holocene of a paleolake, or Jalisco Lake (de Cserna & Alvarez, 1995), with the Cuitzeo and middle Lerma-Santiago Basins constituting a single system (Barbour, 1973; Tamayo, 1987). *Neotoca bilineata* is not the first species whose distribution supports this hypothesis, and other geological features and freshwater fauna similarities have suggested a strong relationship between the Cuitzeo and Lerma-Santiago Basins (López-López & Díaz Pardo, 1991; Díaz-Pardo et al., 1993; Moncayo-Estrada et al., 2001; Gesundheit & Macías Garcia, 2005; Domínguez-Domínguez et al., 2008).

With respect to *N. bilineata*, the main evidence for this theory may be found in the low genetic mtDNA differentiation of the two relatively well-supported clades identified by the NCAs (Fig. 3), with the populations from the Cuitzeo system (RGM and BOR) appearing in clade 4-1 (Fig. 3) and the populations from the middle Lerma (RTC and SAL) included in clade 4-2. It was in this latter clade that the RGM and SAL haplotypes appeared together. This is probably the outcome of a secondary contact between the systems (Cuitzeo and Middle Lerma) in the 1960s through the artificial La Cinta Channel. Additionally, $\phi_{ST}$ values were positively correlated with riparian geographical distances (Fig. 4).
In the phylogenetic and phylogeographical analyses, slight differentiation was observed between BOR in the Cuitzeo system and the other populations in the Cuitzeo system (RGM) and the Lerma-Santiago Basin (Figs 2 and 3). Nuclear data were consistent with the identification of BOR as a genetically differentiated population (Structure result: $K = 2$). Considering that BOR is currently a spring-fed pond isolated from the remaining populations, it is not surprising that this population showed the lowest genetic diversity values, lowest migration rates and highest differentiation level, probably as a consequence of genetic drift.

In the 1960s, an agricultural canal, flowing into the Lerma-Santiago Basin near SAL, connected Cuitzeo Lake and Yuriria Lake (Cinta Channel) (Fig. 1). Although this connection is now in disuse, it could have promoted the transfer of individuals between the two basins and particularly from SAL into RGM, explaining the high genetic diversity detected in these two populations as well as their genetic similarity and proportion of shared alleles. In addition, migration rates between SAL and RGM were greater than among populations within basins (Table VI). Unexpectedly, SAL and RTC, despite co-occurring in the Lerma-Santiago Basin, showed significant genetic differentiation and very low migration rates ($m = 0.24$; Table VI). This observation of restricted gene flow within the basin is probably due to habitat conditions that, as previously discussed, could act as a barrier.

Considering these issues, it could be suggested that the effect of recent migration events, along with the possibility of a longer generation time of the species as a result of the increased water temperature (SEMARNAT, 2001), could have diminished the genetic erosion of the populations. Further studies are needed to confirm this latter hypothesis and to better understand the dynamics of the remnant populations in situ.

The overall conservation status of *N. bilineata* is not encouraging, owing to, among other causes, the high degree of habitat degradation. To conserve the species, it is necessary to control the rapid habitat loss as well as its fragmentation, which could drive the species to extinction before its genetic variability would predict. It is necessary to regulate industrial and urban wastewaters to comply with current legislative limits on the permissible levels of xenobiotics such as heavy metals and pesticides that could reduce species fitness via sublethal toxicity (López-López *et al*., 2006; Arellano-Aguilar & Macías Garcia, 2008). Maintaining a minimum water flow in the habitats of the remaining populations and curtailing the input of nutrients to these water bodies could help decrease the physiological stress to the species. In particular, avoiding anoxic conditions could diminish the risk of high mortality rates provoking drastic demographic changes. Such measures are extremely important to preserve the few remaining *N. bilineata* populations and for the conservation of Central Mexico’s freshwater fauna in general.

The authors thank O. Domínguez for providing genetic material, L. Alcaraz for assistance with laboratory work and C. Pedraza for helping with an early version of the manuscript and field assistance. We thank J. V. Navarro from CONAGUA for his assistance and for providing the water quality data. This study was funded by projects: Ministerio de Educación (CGL 2006-1235BOS) - Modelo Integral para el Establecimiento de zonas de Conservación en la Mesa Central de México and CONACyT - SIHGO (2002020618) - Propuesta de conservación y manejo de la ictiofauna vivípara de las subcuencas Medio Lerma y Alto Pánuco. C.P.O.-G. was supported by grant 196833 from CONACyT – Fundación Carolina. F.A. benefitted from a FPU pre-doctoral grant from the Spanish Ministry of Education.
Supporting Information may be found in the online version of this paper:

**Table SI.** Physicochemical parameter values for the monitoring stations provided by the National Water Commission. RTC, Rio Turbio Corralejo; RGM, Rio Grande Morelia; SAL, Salamanca; $T_{m_a}$, air temperature; $T_{m_s}$, water temperature

**Table SII.** Subindices, weighted values of variables in the index of water quality ($IWQ$) and the descriptive values and correlation matrix of the variables used in the $IWQ$ calculation.

**References**


© 2012 The Authors


**Electronic Reference**