Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 43 (2007) 111–123
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Genetic variation and phylogeography of the swordtail fish *Xiphophorus cortezi* (Cyprinodontiformes, Poeciliidae)

Carla Gutiérrez-Rodríguez [a,*], Molly R. Morris [a], Natalie S. Dubois [a], Kevin de Queiroz [b]

[a] *Department of Biological Sciences, Ohio University, Athens, OH 45701, USA*
[b] *Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA*

Received 17 March 2006; revised 27 September 2006; accepted 20 October 2006
Available online 6 November 2006

Abstract

Swordtail fish have been studied extensively in relation to diverse aspects of biology; however, little attention has been paid to the patterns of genetic variation within and among populations of swordtails. In this study, we sequenced the mtDNA control region from 65 individuals and 10 populations of *Xiphophorus cortezi* to investigate the genetic variation within and among populations, including tests for correlations between genetic and geographic distances and tests for species monophyly. We found low gene and nucleotide diversity within populations and high degrees of genetic differentiation among populations. Significant and positive correlations between genetic distance and both river and straight-line geographic distance indicate that genetic differentiation among *X. cortezi* populations can be explained, to some extent, by an isolation-by-distance model and provide evidence of stream capture. Phylogenetic analyses suggest that *X. cortezi* is paraphyletic relative to *X. malinche*, raising questions concerning the status of these taxa as separate species.

Keywords: Genetic differentiation; Gene flow; Monophyletic; Mitochondrial DNA; Swordtail; *Xiphophorus cortezi*

1. Introduction

Quantification of genetic variation within and among populations is crucial to understanding the evolutionary processes that promote and maintain their biodiversity (Moritz, 2002) as well as establishing genetic relationships among populations in order to reconstruct their evolutionary history. Within a species, the distribution of genetic variation within and among populations is influenced by several factors such as gene flow and genetic drift. Limited dispersal of individuals among populations due to the presence of natural barriers translates into restricted gene flow and can result in genetic differentiation among populations. In river systems, factors that can influence the migration of individuals between populations include the topography of the streams, habitat fragmentation and patterns of flow from one stream to another (Carvalho, 1993).

In this study, we investigate the genetic variation of the swordtail *Xiphophorus cortezi*, a freshwater fish species endemic to the Pánuco river system of Mexico. Swordtails and platyfishes (*Xiphophorus*) have been studied extensively in relation to their systematics, biogeography, genetics, oncology, and behavior (reviewed in Meffe and Snelson, 1989) and are considered a model system for studies in behavioral ecology (Ryan and Rosenthal, 2001). While phylogenetic relationships among species have been examined extensively (Rosen, 1979; Rauchenberger et al., 1990; Meyer et al., 1994; Borowski et al., 1995; Meyer, 1997; Morris et al., 2001), little attention has been paid to patterns of genetic variation within and among populations of swordtails, particularly in a phylogenetic and geographic

* Corresponding author. Present address: Departamento de Biología Evolutiva, Instituto de Ecología, A.C. Km 2.5 antigua carretera a Coatepec No. 351, Congregación El Haya Xalapa, Veracruz 91070, México. Fax: +52 228 818 7809.
E-mail address: carla.gutierrez@inecol.edu.mx (C. Gutiérrez-Rodríguez).

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.10.022

context. Morris et al. (2001) included several populations of *X. cortezi* in their phylogenetic study of the northern swordtails, and in some of their analyses, these populations did not form a monophyletic group. However, relationships among *X. cortez*i populations were not specifically addressed or well supported by their allozyme data.

Xiphophorus cortezi is distributed throughout the southern portion of the Pánuco river basin of eastern Mexico. This river enters the Gulf of Mexico at Tampico and drains the Sierra Madre Oriental. *X. cortezi* inhabits relatively still pools in small streams with rocky bottoms but is missing from larger rivers with sandy bottoms (Rauchenberger et al., 1990). The physical distribution of *X. cortezi* in small streams across three larger river drainages suggests that the populations might have evolved independently from each other. Because the rivers that connect the populations become quite large and therefore are not suitable habitat for this fish, the river habitats downstream may act as physical and ecological barriers among the populations in the headwater streams. Therefore, we expect *X. cortezi* populations to be genetically differentiated. In addition, some of the populations found in the headwaters of different rivers are geographically closer than the populations found within the same river system. As the river systems in this area are known to be relatively instable (Rauchenberger et al., 1990), we suspect that stream capture (i.e., the process by which a river or stream erodes through a divide so that its flow is diverted into a neighboring drainage system) may have played a role in the current distribution of this species. The purpose of our study is to assess the distribution of genetic variation within and among populations of *X. cortezi*, establish the geographical relationships among populations, test for cases of stream capture, and test the hypothesis of monophyly for this species.

2. Materials and methods

2.1. Samples

We collected 65 individuals from 10 different sites located in three drainages of the Pánuco River system in Mexico, thereby sampling the known distribution of the species (Fig. 1 and Table 1). A fin-clip was collected from the caudal fin of each fish and preserved in salt-saturated 20% dimethyl sulphoxide solution (Seutin et al., 1991). Fin-clips from single individuals of swordtail species *X. birchmanni*, *X. malinche*, *X. montezumae* and *X. multilineatus* from Atlapexco, Soyatla, Río Frío, and Río Coy, respectively, were included as presumptive outgroups for the phylogenetic analyses.

2.2. Amplification, sequencing and alignment

Total DNA was extracted from fin-clips using the DNeasy tissue kit (Qiagen Inc.) following the manufacturer's instructions. The entire mitochondrial control region was amplified in 25 or 50 μl reactions using the polymerase chain reaction (PCR) and the previously published primers K (5′ AGCTCAGCGCCAGAGCGCCGGTCTTGTAAA 3′) and G (5′ CGTCGGATCCCATCTTCAGTGTTAT GCTT 3′) (Lee et al., 1995) according to standard methods. Reactions were performed on a MJ Research PTC-100 thermocycler (MJ Research Inc.) under the following conditions: 2 min at 94°C, followed by 35 cycles at 90 or 94 °C for 30 s, 55–63 °C for 30 s and 72 °C for 45 s, with a final extension period of 7 min at 72 °C. PCR products were purified using the UltraClean PCR Clean-up DNA Purification Kit (MoBio Laboratories Inc.) or the QIAquick Gel Extraction Kit (Qiagen Inc.) and sequenced in both directions to check the validity of the sequence data using the BigDye terminator cycle sequencing kit. Sequences were visualized with an ABI 310 or 3730 automated sequencers (Applied Biosystems) at the DNA Analysis Facility at Ohio University and the Plant-Microbe Genomics Facility at Ohio State University respectively.

We aligned the forward and reverse sequences using the computer software SeqMan II (DNAStar Inc.). Contiguous sequences were initially aligned using Clustal X (Thompson et al., 1997) with the default settings, followed by manual alignment with the program Sequence Alignment Editor v2.0a11. All sequences were deposited in Genbank (Table 1).

2.3. Population genetics analyses

All population genetic analyses were performed using ARLEQUIN (Schneider et al., 2000). Levels of intra-population variation were estimated by calculating the mean gene diversity (h), the probability that two randomly chosen haplotypes are different (Nei, 1987) and the nucleotide diversity (π), the average number of differences between all pairs of haplotypes (Tajima, 1983; Nei, 1987). An analysis of molecular variance (AMOVA) (Excoffier et al., 1992) was used to test whether sequences from previously defined groups were significantly different from each other. We defined groups corresponding to the three drainages (Tampeon, Moctezuma and Tempoal) from which the samples were collected (Table 1). An AMOVA without defined groups was also conducted. We used 16,000 permutations to test for statistical significance of both AMOVAs. AMOVA compares the similarity within and among groups using genetic distance as the measure of similarity. We used the Tamura-Nei (TrN) model (Tamura and Nei, 1993) with the transition/transvertion (ti/tv) ratio = 3.0232 and the gamma shape parameter (α) = 0.8982 (see below) to calculate the genetic distance between the different sequences. HKY, which is the model that best fitted our data (see below), is not an option in ARLEQUIN 2.0 and TrN is the closest available approximation to HKY. We calculated the degree of genetic differentiation between all the sampled populations with pairwise F-statistics (Weir and Cockerham, 1984) and their significance by performing 1000 permutations.

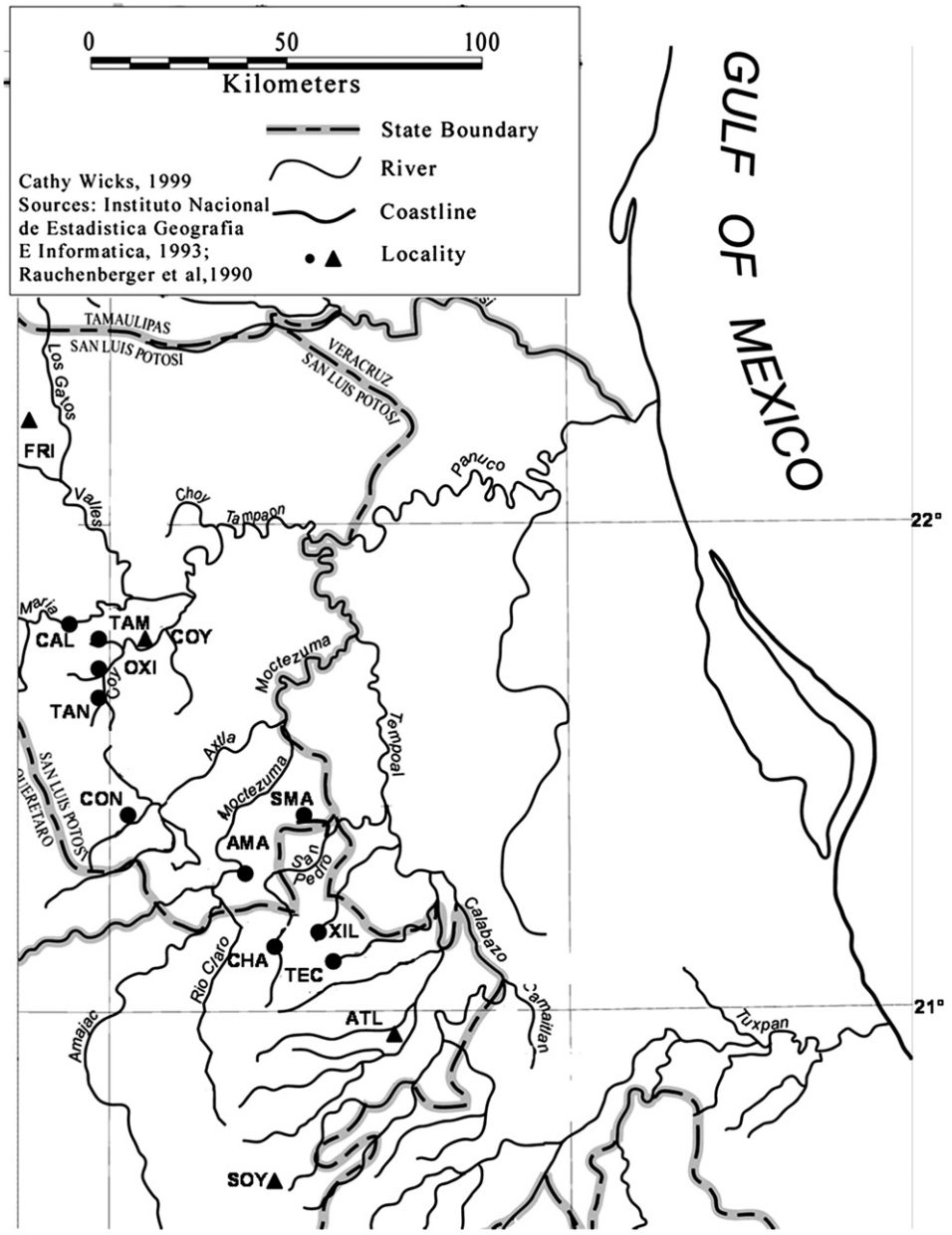

Fig. 1. Map showing the different collection sites of *X. cortezi* (●) and four outgroup taxa (▲).

Two Mantel tests with 10,000 permutations were performed to test for a correlation between genetic and geographic distance using the computer program IBD (Bohonak, 2002). Genetic distances were calculated using the method of Nei and Li (1979) and 1000 permutations in both tests. Nei and Li's measure of genetic distance (1979) is an estimate of the number of nucleotide differences between populations that incorporates information on both haplotype frequencies and the number of substitutional differences between haplotypes. Geographic distances were measured using river distances (the paths along the water courses connecting two localities) and straight-line distances (the minimum great circle distance between two localities). Both distances were calculated using the software ExpertGPS 1.3.7 (Topografix, 2003).

2.4. Relationships among haplotypes

We assessed genetic differences between pairs of haplotypes using both the total number of base differences and maximum-likelihood distances (HKY model, ti/tv ratio and base frequencies estimated from the data, values for proportion of invariant sites and gamma shape parameter as obtained from MODELTEST software, see below). Distance calculations were performed using PAUP* 4.0b10 (Swofford, 2002).

We analyzed the phylogenetic relationships among the haplotypes using a variety of inference methods. Phylogenetic trees of the *X. cortezi* and outgroup haplotypes were inferred using PAUP* 4.0b10 (Swofford, 2002) under parsimony, maximum-likelihood and least-squares distance

Table 1
Populations of *X. cortezi* sampled in this study, measures of their genetic diversity and GenBank Accession Nos.

Locality	Species	Abbreviation	Drainage	*N*	Nh	*h*	π	GenBank Accession Nos.
Oxitipa	*X. cortezi*	OXI	Tampaón	3	1	0.000	0.0000	DQ445674
Caldera	*X. cortezi*	CAL	Tampaón	5	1	0.000	0.0000	DQ445674
Tambaque	*X. cortezi*	TAM	Tampaón	5	1	0.000	0.0000	DQ445674
Tanute	*X. cortezi*	TAN	Tampaón	8	1	0.000	0.0000	DQ445674
La Conchita	*X. cortezi*	CON	Moctezuma	12	2	0.530	0.0006	DQ445669-70
Amacuzac	*X. cortezi*	AMA	Moctezuma	1	1	—	—	DQ445671
San Martín	*X. cortezi*	SAM	Tempoal	10	1	0.000	0.0000	DQ445671
Chalpuhuacanita	*X. cortezi*	CHA	Tempoal	10	2	0.467	0.0005	DQ445672-73
Tecolutlo	*X. cortezi*	TEC	Tempoal	4	1	0.000	0.0000	DQ445675
Xiliatl	*X. cortezi*	XIL	Tempoal	7	2	0.476	0.0011	DQ445676-77
Atlapexco	*X. birchmanni*	ATL	Tempoal	1	1	—	—	DQ445678
Soyatla	*X. malinche*	SOY	Tempoal	1	1	—	—	DQ445679
Río Frío	*X. montezumae*	FRI	Tampaón	1	1	—	—	DQ445680
Río Coy	*X. multilineatus*	COY	Tampaón	1	1	—	—	DQ445681

N = number of sequences (individuals), Nh = number of haplotypes, h = mean gene diversity, π = nucleotide diversity, — = calculations are not possible because only one sample was collected from the site.

criteria. For the parsimony analysis, we conducted a branch and bound search, treating gaps as a "fifth base". For the likelihood analyses, we selected a nucleotide substitution model using hierarchical likelihood ratio tests as implemented in MODELTEST 3.06 (Posada and Crandall, 1998). We estimated the optimal tree using a successive approximation approach (Swofford et al., 1996; Sullivan et al., 2005) with PAUP* 4.0b10 (Swofford, 2002). This method consists of a series of successive likelihood analyses, with the parameter values used in a given analysis estimated on the optimal tree from the previous analysis and reiterated until the topology, the likelihood score for the tree, and the parameter values match those from the previous analysis. Each iteration used a heuristic search with starting trees obtained by random stepwise addition (100 replicates) and TBR branch swapping. For the distance analysis, we used unweighted least squares as the optimality criterion and maximum-likelihood (HKY85 + G + I) distances, with the parameters estimated from the neighbor-joining tree used in the hierarchical likelihood ratio tests. Gaps were treated as missing data in a heuristic search with starting trees obtained by random stepwise addition (100 replicates) and TBR branch swapping. Nodal support was assessed under the likelihood criterion using non-parametric bootstrap resampling (1000 replicates) and heuristic searches with starting trees obtained by random stepwise addition (10 replicates) and TBR branch swapping.

Relationships among the haplotypes were also inferred using the statistical parsimony method of Templeton et al. (1992). Unlike the previous methods, relationships inferred by the statistical parsimony method are not constrained to take the form of a tree (minimally connected graph), allowing them to take the form of more extensively connected graphs in which additional connections represent alternative, equally parsimonious, mutational pathways from one haplotype to another. TCS v1.21 (Clement et al., 2000) was used to estimate the statistical parsimony network (SPN), with the default 0.95 probability connection limit and treating gaps both as a fifth state and as missing data.

Because different outgroups suggested two different root positions, we conducted three analyses to determine whether our data could distinguish between the alternative root placements with confidence. First, we constrained the *X. cortezi* plus *X. malinche* haplotypes to form a monophyletic group and used the Kishino-Hasegawa (KH; Kishino and Hasegawa, 1989) and Shimodaira-Hasegawa (SH; Shimodaira and Hasegawa, 1999) tests to determine whether the maximum-likelihood trees in the presence versus absence of this constraint were significantly different in their ability to explain the data. Second, we searched for optimal trees (1) using only *X. montezumae* and *X. birchmanni* as outgroups (excluding *X. multilineatus*) and (2) using only *X. mulitlineatus* as the outgroup (excluding *X. montezumae* and *X. birchmanni*). We compared the maximum-likelihood trees inferred in each of these analyses (using the successive approximation approach described above) with the trees using the same outgroup(s) attached to the branch identified as the root by the alternative outgroup(s) using the KH and SH tests. For these tests, we used PAUP* 4.0b10 with the same model as in the maximum-likelihood tree searches, estimating the transition/transversion ratios, the nucleotide base frequencies, the proportion of variable sites and the shape parameter (α), and using RELL bootstrap resampling (1000 replicates).

To assess the strength of the support for the derivation of *X. malinche* from within *X. cortezi*, we constrained *X. cortezi* to be monophyletic and, because of the different root positions suggested by the different outgroups, we conducted two separate analyses. We estimated maximum-likelihood trees using the successive approximations approach described above with the same model, first using only *X. montezumae* and *X. birchmanni* as outgroups (excluding *X. multilineatus*) and then using only *X. multilineatus* as the outgroup (excluding *X. montezumae* and *X. birchmanni*). We then compared the maximum-likelihood trees inferred in these analyses with those inferred

for the same set of taxa in the absence of the monophyly constraint using KH and SH tests. For these tests, we used PAUP* 4.0b10 with the same model as in the maximum-likelihood tree searches, estimating the transition/transversion ratios, the nucleotide base frequencies, the proportion of variable sites and the shape parameter (α), and using RELL bootstrap resampling (1000 replicates).

3. Results

3.1. Sequence variation

We obtained 881 bp sequences for the mitochondrial DNA control region of 65 *X. cortezi* individuals from 10 populations as well as of four individuals of closely related species. For the *X. cortezi* and four outgroup haplotypes together, 67 variable sites out of the 881 were detected (7.6%) and for the *X. cortezi* haplotypes alone, 16 variable sites were detected (1.8%). The minimum number of mutations for the *X. cortezi* haplotypes alone was 16 (none of the variable sites for the nine haplotypes observed in this species had more than two bases). Although four sites (69–72) could not be aligned unambiguously, we retained them in our analyses. Three of the haplotypes (H7, H8, and H9) differ only in the ambiguously aligned region yet are clearly different from one another. We thought that it was preferable to retain the information present in the identification of three distinct haplotypes even though the specific substitutions separating those haplotypes are ambiguous. In any case, analyses excluding these positions yielded qualitatively similar results (not presented).

3.2. Genetic diversity and population structure

Nine haplotypes were identified among the 65 sequenced *X. cortezi* individuals and four among the outgroup taxa (Table 2). Haplotypes H3 and H6 were shared by individuals from different populations, and haplotype H6 was both the most widespread (four localities) and the most common (21 individuals). The remainder of the haplotypes were only shared by individuals from the same locality (Table 2). Three of the populations (Chalpuhuacanita, La Conchita, Xiliatl) exhibited two different haplotypes; all other populations exhibited only a single haplotype. Mean gene diversity (h) was very low to moderate among *X. cortezi* populations (Table 1). Nucleotide diversity (π) was very low in all populations, ranging from 0 to 0.0011 (Table 1) indicating a high degree of similarity of sequences within populations.

The analysis of variance, in which we specified three different groups based on river drainages (Tampaón, Moctezuma and Tempoal), showed significant genetic structure at all hierarchical levels. Most of the variation (53.71%) was explained by differences among the three drainages. Differences among the populations within each drainage were also large and explained 42.07% of the variance. Only 4.22% of the variance was attributed to differences among individuals within populations (Table 3). The AMOVA performed without groupings also showed high and significant degrees of genetic structure (Table 3). Most of the variation (95.03%) was explained by differences among populations, and only 4.97% by differences within populations.

Genetic differentiation between pairs of populations is presented in Table 4. The majority of the F_{ST} values were high ($\geqslant 0.4232$) and significantly different from zero. The main exceptions were the comparisons between pairs of populations located in the Tampaón drainage, which were all zero and thus not significant. F_{ST} values between populations from different drainages were generally higher than values between populations from the same drainage, although San Martín and Chalpuhuacanita (Tempoal drainage) exhibited low values in comparison to La Conchita (Moctezuma drainage) relative to other between-drainage comparisons and relatively high values in comparison with Tecolutlo and Xiliatl (Tempoal drainage) relative to other within-drainage comparisons.

Table 2
Number and geographic distribution of haplotypes from *X. cortezi* and four putative outgroup species

Species	Location	Haplotype												
		H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13
X. cortezi	OXI						3							
X. cortezi	CAL						5							
X. cortezi	TAM						5							
X. cortezi	TAN						8							
X. cortezi	CON	5	7											
X. cortezi	AMA			1										
X. cortezi	SAM			10										
X. cortezi	CHA				7	3								
X. cortezi	TEC							4						
X. cortezi	XIL								5	2				
X. birchmanni	ATL										1			
X. malinche	SOY											1		
X. montezumae	FRI												1	
X. multilineatus	COY													1
Total		5	7	11	7	3	21	4	5	2	1	1	1	1

Table 3
Hierarchical analysis of molecular variance (AMOVA) for *X. cortezi*

Source of variation	df	Sum of squares	Variance components	Percentage of variation
AMOVA with drainage grouping				
Among groups	2	99.700	1.82654(Va)	53.71*
Among populations within groups	7	57.428	1.43058(Vb)	42.07*
Within populations	55	7.899	0.14362(Vc)	4.22*
Total	64	165.028	3.40074	
AMOVA without grouping				
Among populations	9	157.129	2.74359(Va)	95.03*
Within populations	55	7.899	0.14362(Vb)	4.97
Total	64	165.028		

* Significant values at $P < 0.05$.

Table 4
Pairwise F_{ST} comparisons (below the diagonal) and Nei and Li (1979) genetic distances (above the diagonal) among *X. cortezi* populations

	OXI	CAL	TAM	TAN	CON	AMA	SAM	CHA	TEC
OXI		0	0	0	19.7771	19.1938	19.1938	22.5261	31.323
CAL	0.0000		0	0	19.7771	19.1938	19.1938	22.5261	31.323
TAM	0.0000	0.0000		0	19.7771	19.1938	19.1938	22.5261	31.323
TAN	0.0000	0.0000	0.0000		19.7771	19.1938	19.1938	22.5261	31.323
CON	0.9421*	0.9497*	0.9497*	0.9578*		2.2833	2.5833	5.3156	18.7771
AMA	1.0000	1.0000	1.0000	1.0000	0.6657		0	3.3323	20.1938
SAM	1.0000*	1.0000*	1.0000*	1.0000*	0.8186*	0.0000		3.3323	20.1938
CHA	0.9487*	0.9565*	0.9565*	0.9644*	0.7809*	0.6412	0.8206*		22.9261
TEC	1.0000*	1.0000*	1.0000*	1.0000*	0.9381*	1.0000	1.0000*	0.9554*	
XIL	0.9301*	0.9437*	0.9437*	0.9563*	0.9031*	0.8727	0.9486*	0.9190*	0.4232*

The significance of the F_{ST} values between Río Amacuzac and the rest of populations is not reported because only a single individual was sampled from the Río Amacuzac.
* $P < 0.05$ significance level.

Mantel tests revealed significant and positive correlations between genetic distance and river distance (Fig. 2A, $r = 0.766$, $P = 0.0004$) as well as between genetic distance and straight-line geographic distance (Fig. 2B, $r = 0.769$, $P = 0.0003$). All population comparisons fell within the 95% confidence intervals for both regressions, although eight comparisons in the analysis using river distances were much further than the others from the line of best fit. Four of these represent pairs of populations whose genetic distances are substantially greater than expected given their geographic distances; they involve comparisons between San Martín and Chalpuhuacanita, and comparisons between Tecolutlo and Xiliatl, all from the Tempoal drainage. The other four comparisons represent pairs of populations whose genetic distances were substantially smaller than expected given their geographic distances; they involve comparisons between San Martín and Chalpuhuacanita from the Tempoal drainage and La Conchita and Amacuzac from the Moctezuma drainage.

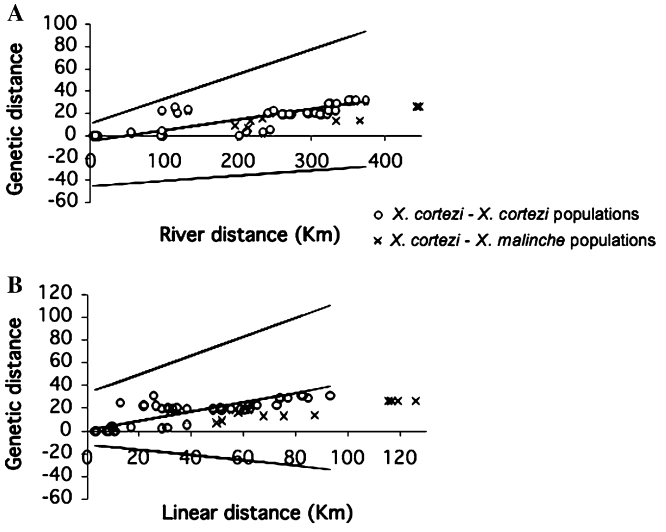

Fig. 2. Plots of (A) Nei and Li (1979) genetic distance compared to river distance for all possible pairs of populations, (B) Nei and Li (1979) genetic distance compared to straight-line geographic distance for the same pairs of populations. The regression line and the lines that bound the 95% confidence intervals are shown and were calculated only for the *X. cortezi* data. Comparisons between all *X. cortezi* populations to the *X. malinche* population are included and are denoted with a different symbol.

3.3. Relationships among haplotypes

Hierarchical likelihood ratio tests identified HKY85 + I + G with the following initial parameter estimates as the appropriate model for our data: transition/transversion (ti/tv) ratio = 3.0323, proportion of invariable

sites (i) = 0.8015, gamma shape parameter (α) = 0.8982, and base frequencies $A = 0.3091$, $C = 0.2274$, $G = 0.1391$, $T = 0.3243$.

The optimal trees derived from parsimony, distance and likelihood analyses were generally similar, although the strict consensus of the maximum-likelihood trees exhibited greater resolution (8/11 possible nodes) than those derived from distance (7/11) and parsimony (6/11) analyses. The maximum-likelihood tree with nodal support (bootstrap) values is presented in Fig. 3. The only group present in this tree that is contradicted by the results of analyses based on other optimality criteria is H3, H4, H6, H10 and H12. In the parsimony tree H3 and H4 group with H5, while in the least squares distance tree they group with H1, H2 and H5. None of these conflicting relationships is strongly supported (bootstrap proportions all $\leqslant$55%). Details of the results of all three analyses are presented in Appendix A.

In all optimal trees inferred by likelihood and distance analyses and some of those inferred by parsimony analysis, the trees cannot be rooted so that the haplotypes found in *X. cortezi* (and *X. malinche*) form a monophyletic group because different outgroups suggest different root positions (Fig. 3). The *X. mulitlineatus* haplotype roots the tree between the group H7–H9 + H11 (found in *X. cortezi* and *X. malinche* populations from the Tempoal river drainage) and the remaining haplotypes (Fig. 3A); whereas the *X. birchmanni* plus *X. montezumae* haplotypes root the tree on the branch to H6 (found in four populations from the Tampaón drainage) (Fig. 3B). The first root position (Fig. 3A) is inferred in those parsimony trees that can be rooted without disrupting monophyly of the *X. cortezi* (and *X. malinche*) haplotypes. Regardless of root position, the *X. malinche* haplotype (H11) falls out within those of *X. cortezi* and is most closely related to haplotypes from the Tempoal drainage (H7–9).

In the trees resulting from all of these analyses, the *X. cortezi* haplotypes fall into groups that correspond roughly to the river drainages in which they are found. All populations from the Tampaón drainage exhibit a single haplotype (H6). The haplotypes from the Moctezuma drainage (H1–3) form a para- or polyphyletic group relative to those from the Tempoal drainage, or the Tempoal and Tampaón drainages, depending on the position of the root. Finally, some of the haplotypes from the Tempoal drainage (H7–9) form a monophyletic group while the others (H3–5) are intermixed with haplotypes from the Moctezuma drainage (H1–3). Branch lengths leading to the various *X. cortezi* haplotypes are generally short, and only four groups in the maximum-likelihood tree are supported by bootstrap values >50% (Fig. 3).

The statistical parsimony networks (Fig. 4) postulate a number of hypothetical unsampled haplotyes (which differ depending on whether gaps are treated as a fifth state or as missing data) in addition to the nine *X. cortezi* and one *X. malinche* sampled haplotypes that were connected with probabilities $\geqslant$0.95 (ignoring the outgroup *X. multilineatus*

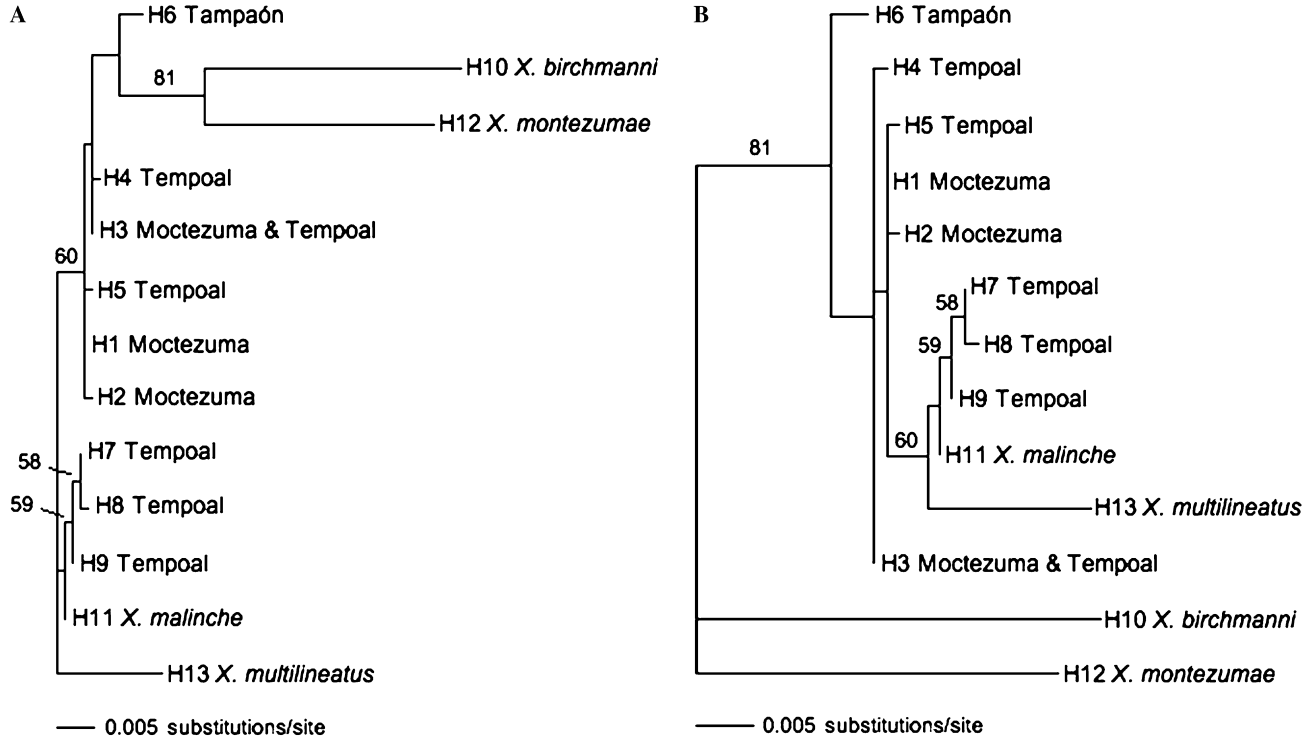

Fig. 3. Maximum-likelihood bootstrap tree for mitochondrial DNA control region haplotypes from *X. cortezi* (H1–9) and four putative outgroups (H10–13). Branch lengths are proportional to the number of substitutions per site. Numbers above branches indicate bootstrap support if greater than 50%. Names of the outgroup species and drainage name for the *X. cortezi* haplotypes are also indicated. (A) Rooted with the outgroup *X. multilineatus*. (B) Rooted with the outgroups *X. montezumae* and *X. birchmanni*.

in the network in which gaps were treated as missing data, Fig. 4B). Outgroup weights (Castelloe and Templeton, 1994) based on both haplotype frequencies and positions in the graph (internal versus terminal) suggest that either haplotype H4, found in the Chalpuhuacanita population, or H1, found in the La Conchita population, is ancestral (Fig. 4), depending on the interpretation of gaps.

The *X. cortezi* haplotypes form three groups of closely connected haplotypes (1–2 mutations between nearest neighbors) that are separated from one another by more distant connections (at least five mutations when treating gaps as a fifth state). One contains haplotypes from the Tampaón drainage (H6), a second has haplotypes from the Moctezuma and Tempoal drainages (H1–5), and a third is formed by haplotypes from the Tempoal drainage (H7–9). The haplotype from *X. malinche* (H11) is positioned in the networks 1 or 2 mutational steps from *X. cortezi* haplotype (H9) from the Tempoal drainage. In contrast, the *X. birchmanni* (H10) and *X. montezumae* (H12) haplotypes, and, when gaps are treated as a fifth base, that of *X. multilineatus* (H13), are not linked to the networks, indicating that the number of nucleotide differences between those *Xiphophorus* species and *X. cortezi* exceeds the connection limit of 12 steps, which corresponds to a probability of 0.95 that the number of mutations is accurately estimated by parsimony. For the *X. cortezi* haplotypes alone, 19 mutations were estimated by both tree and network methods that treated gaps as a fifth state (20 when the *X. malinche* haplotype is included).

Tests of the alternative root placements suggested by different outgroups (Fig. 3A and B) did not reveal significant differences. The maximum-likelihood tree in the absence of any constraints did not exhibit a significantly better fit to the data than the maximum-likelihood tree constrained so that the *X. cortezi* plus *X. malinche* haplotypes formed a monophyletic group according to both the KH test ($P = 0.552$) and the SH test ($P = 0.309$). This constrained tree (not shown) was rooted in the position indicated by *X. multilineatus* (see Fig. 3A). In addition, the maximum-likelihood tree including only *X. montezumae* and *X. birchmanni* as outgroups did not exhibit a significantly better fit to the data than a tree in which these two outgroups were attached in the position where *X. multilineatus* was attached in the analysis including all three outgroups (KH test, $P = 0.532$; SH test, $P = 0.246$). Similarly, the maximum-likelihood tree including only *X. multilineatus* as the outgroup did not exhibit a significantly better fit to the data than a tree in which *X. multilineatus* was attached in the position where *X. montezumae* and *X. birchmanni* were attached in the analysis including all three outgroups (KH test, $P = 0.156$; SH test, $P = 0.097$).

Although derivation of *X. malinche* from within *X. cortezi* (paraphyly of *X. cortezi* relative to *X. malinche*) is implied by the optimal trees from phylogenetic analyses under diverse optimality criteria (Fig. 3 and Appendix A), our data do not reject the alternative hypothesis of monophyly for the *X. cortezi* haplotypes. Comparison of the maximum-likelihood trees in the absence of any constraints to the maximum-likelihood trees under a constraint of *X. cortezi* monophyly did not reject the hypothesis of *X. cortezi* monophyly whether *X. montezumae* and *X. birchmanni* were used as outgroups (KH test,

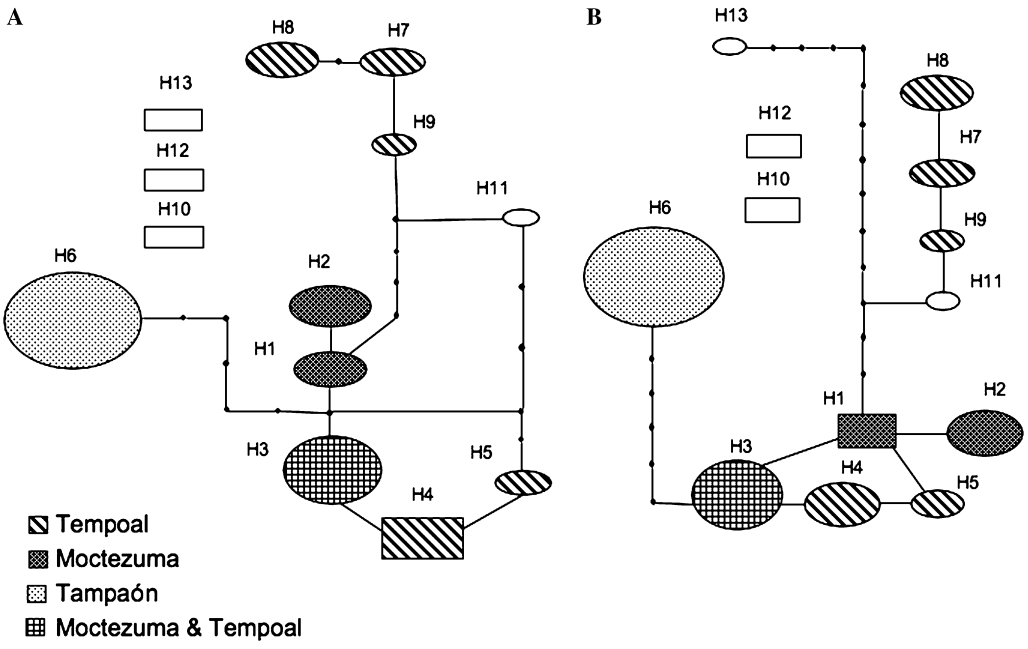

Fig. 4. Statistical parsimony network based on *X. cortezi* sequences of the mitochondrial DNA control region. (A) Treating gaps as a fifth state and (B) treating gaps as missing data. Haplotype designations are from Table 1. The rectangles represent the inferred ancestral haplotype. The size of the ovals and rectangles is proportional to the haplotype's frequency. Small black circles are hypothetical haplotypes necessary to connect the observed haplotypes in the network. Each of the lines between observed and/or hypothetical haplotypes represents one mutational step.

$P = 0.532$; SH test, $P = 0.246$) or *X. multilineatus* was used as the outgroup (KH test, $P = 0.783$; SH test, $P = 0.649$).

4. Discussion

4.1. Genetic variation

The results of this study indicate a low degree of genetic variation in the mtDNA control region of *X. cortezi* compared to other teleostean fishes (i.e., Fajen and Breden, 1992; Lee et al., 1995; Salzburger et al., 2003; Stefanni and Thorley, 2003; Aboim et al., 2005). Out of 881 aligned positions 16 (1.8%) were variable (the numbers are the same if the *X. malinche* sample is included) and out of 65 sequences nine (13.8%) distinct haplotypes were found (10 if the *X. malinche* sample is included). Gene diversity (h) was low to moderate and nucleotide diversity (π) was very low in *X. cortezi* populations. Although the number of genotypes found in each population should be interpreted with caution because of the limited sample sizes, similar low levels of within population variation in the mtDNA control region have been found in other poeciliids, such as the guppy species *Poecilia reticulata* (Fajen and Breden, 1992; Carvalho et al., 1996; Shaw et al., unpublished data) and the swordtail species *X. birchmanni* (Shearer et al., unpublished data).

The low genetic diversity within populations may reflect founder effects, as reductions in genetic polymorphisms have been observed in introduced populations of guppies (Shaw et al., 1992; Carvalho et al., 1996). Alternatively, the low values of nucleotide diversity (π) combined with moderate haplotype diversity (h) may suggest recent population expansion after a bottleneck or a founder event. However, a Tajima's D-test (Tajima, 1989) indicated this is not the case in *X. cortezi*, as all populations had positive D values that were not significantly different from zero (results not shown). These patterns of low genetic diversity within population are most likely the consequence of consistently small population sizes, which might have resulted in genetic drift and inbreeding (Nei et al., 1975; Wishard et al., 1984). Small population size could be a consequence of human activities, as there are indications of human contamination in these areas that are likely to contribute to fish mortality.

In contrast to the low variation within populations, we found high degrees of genetic differentiation among *X. cortezi* populations. The AMOVA with drainage grouping showed that the variation among populations within drainages explained a large proportion of the total variance, indicating genetic structure among populations located in the same drainage. This was supported by the significance of the AMOVA performed without groupings and by the high and significant values of the F_{ST} comparisons between most pairs of populations. High levels of genetic differentiation among populations have been reported in other freshwater fishes (Carvalho et al., 1991; Shaw et al., 1991, 1994; Fajen and Breden, 1992; Alves and Coelho, 1994; Coelho et al., 1997; Hänfling and Brandl, 1998a,b,c; Mesquita et al., 2001) and could be the result of discontinuities in suitable habitat. As a result, genetic differentiation might reflect the physical subdivision of populations (Carvalho, 1993). Sections of the streams where *X. cortezi* and other swordtails species are distributed dry seasonally and may be contaminated by human activities, resulting in destruction of suitable habitat. In addition, *X. cortezi* only occurs in small rivers, which are often more subdivided by barriers to fish movement than larger rivers.

The populations from the Tampaón drainage (Oxitipa, Caldera, Tambaque and Tanute) constitute an exception to the pattern of high genetic differentiation among populations from the same drainage. All the populations in Tampaón shared the same single haplotype (H6). Such genetic homogeneity among the sampling sites in this drainage suggests recent and/or historical migration throughout the streams of the drainage and could reflect the absence of barriers to gene flow. Alternatively, it could reflect the geographic proximity of the collection sites in this drainage. The geographic distances between the populations sampled from the Tampaón drainage are, in general, smaller than those between populations in the other drainages and do not include stretches of larger rivers with unsuitable *X. cortezi* habitat. The patterns of low intrapopulation genetic diversity combined with high genetic differentiation among populations found in this investigation have also been reported in studies of other freshwater fishes based on allozymes (Carvalho et al., 1991; Shaw et al., 1994; Hänfling and Brandl, 1998b,c) and mtDNA (Carvalho et al., 1996; Mesquita et al., 2001; Shaw et al., unpublished data).

4.2. Geographic patterns of genetic variation and evidence of stream capture

Our data exhibited significant correlations between genetic and geographic distances, regardless of whether geographic distance was measured along watercourses or as straight lines. These findings suggest that genetic differentiation among *X. cortezi* populations can be explained, to some extent, by an isolation-by-distance model. A strong geographic effect on genetic variation is also implied by the results of the AMOVA with drainage grouping, in which a significant amount of the total genetic variation was attributable to differences among groups of populations from different drainages. On the other hand, there does not appear to be a one-to-one correspondence between groups of related haplotypes and drainage groups. Neither the haplotype trees nor the haplotype networks exhibit groups that correspond precisely to the three drainage groups. In both cases, some haplotypes from the Tempoal drainage (H4–5) are closer to those from the Moctezuma drainage (H1–3), in terms of both common ancestry and mutational steps, than they are to other haplotypes from the Tempoal drainage (H7–9). Moreover, one of the haplotypes (H3) is found in both drainages.

Several of our results provide evidence of stream capture. First, the Mantel tests revealed that genetic distance between populations is more strongly correlated with straight-line geographic distance than with geographic distance measured along river courses, suggesting that genetic differentiation is more strongly influenced by gene flow between populations in adjacent drainages (as might occur through stream capture) than by gene flow up and down river courses. Second, in the comparison of genetic distances versus river distances, the eight most obvious outliers all represent comparisons involving populations from the northwestern part of the Tempoal drainage. The populations in this drainage (represented in our study by San Martín and Chalpuhacanita) exhibit substantially larger genetic distances to the other populations in the Tempoal drainage (Tecolutlo and Xiliatl) and substantially smaller genetic distances to the populations in the Moctezuma drainage (La Conchita and Amacuzac) than expected given their river course geographic distances suggesting that these two streams may have drained into the Moctezuma in the past. Third, in both the haplotype trees and the statistical parsimony networks, the haplotypes from San Martín (H3) and Chalpuhuacanita (H4 and H5) from the Tempoal drainage are more closely related to haplotypes from the Moctezuma drainage (H1, H2 and H3, which is shared with San Martín) than they are to other haplotypes from the Tempoal drainage (H7, H8, and H9). The above findings, in conjunction with the fact that the northwestern branches of the Tempoal drainage are located geographically between the Moctezuma drainage and the southeastern branches of the Tempoal drainage, are consistent with a scenario in which the current northwestern branches of the Tempoal drainage were formerly part of the Moctezuma drainage and were later captured by the Tempoal drainage. The low elevation (less than 100 m) separating parts of the drainages also suggests the possibility of stream capture.

4.3. Root of the haplotype tree and ancestral haplotype

Although tests designed to establish the placement of the root of the haplotype tree could not conclusively distinguish between the alternatives, two lines of evidence favor the root position indicated by the outgroup *X. multilineatus* over that suggested by *X. birchmanni* and *X. montezumae*. First, some of the optimal trees under parsimony as well as the maximum-likelihood tree under the constraint of monophyly of the *X. cortezi* plus *X. malinche* haplotypes (trees not shown) were rooted in the position (branch) suggested by *X. multilineatus* on the maximum-likelihood tree in the absence of any constraints. Second, trees rooted using *X. birchmanni* and *X. montezumae* as outgroups were far from approaching significance in distinguishing between alternative root positions ($P = 0.246$–0.532), suggesting that these taxa are not particularly informative for discriminating between the two alternatives, while trees rooted using only *X. multilineatus* as the outgroup were closer to approaching significance in favoring the root position implied by that species over the position suggested by *X. birchmanni* and *X. montezumae* ($P = 0.097$–0.156).

The statistical parsimony networks indicated that either haplotype H4 (when gaps treated as fifth state) or H1 (when gaps treated as missing) is ancestral among the haplotypes from *X. cortezi* and *X. malinche*, an inference based on both the relative haplotype frequencies and numbers of connections to other haplotypes (Castelloe and Templeton, 1994; Clement et al., 2000). If the root position is assumed to occur at the node at which *X. multilineatus* attaches to the tree (see above), then the optimal phylogenetic tree suggests that the ancestral haplotype is located along the branch separating *X. cortezi* haplotypes H7–9 from haplotypes H1–6, which is only a few mutational steps from the haplotypes identified as ancestral by the analyses using TCS, particularly when gaps are treated as missing data (H1). In the results of that analysis (Fig. 4B) the outgroup *X. multilineatus* (H13) connects to the rest of the network in the same area. In addition to using different kinds of information to infer the ancestral haplotype (position and frequency versus outgroup) the minor differences between the results of these analyses may be attributable to the fact that the TCS analyses only identify observed haplotypes as ancestral while the phylogenetic analyses can identify an inferred haplotype as ancestral.

4.4. Monophyly versus paraphyly of *X. cortezi* haplotypes and the status of *X. malinche*

The phylogenetic analyses suggest that *X. cortezi* is not monophyletic. Regardless of the root position, *X. cortezi* haplotypes from the Tempoal drainage (H7–9) appear to be more closely related to the *X. malinche* haplotype (H11) than they are to other *X. cortezi* haplotypes (H1–6). In addition, the number of substitutions separating the *X. malinche* haplotype from the *X. cortezi* haplotypes are within the range for comparisons involving the *X. cortezi* haplotypes alone. Moreover, our sample of *X. malinche* (from Soyatla) and the haplotypes of *X. cortezi* to which it appears most closely related (from Xiliatl and Tecolutlo) all come from the Tempoal drainage. Although our data cannot reject the alternative hypothesis that the *X. cortezi* haplotypes form a monophyletic group exclusive of *X. malinche*, based on the phylogenetic relationships and geographic proximity of the *X. malinche* haplotypes to the *X. cortezi* haplotypes from Tecolutlo and Xiliatl we suspect that these haplotypes (and the populations in which they occur) are closely related and that failure to demonstrate this relationship with a high degree of confidence is a result of the limited power of the tests resulting from the small number of mutational differences between the haplotypes in question. The use of a more polymorphic marker would be necessary to conclusively resolve these relationships.

The finding that *X. cortezi* may be paraphyletic relative to *X. malinche* raises several issues concerning the status of these taxa as separate species. One possibility is that the

two taxa are separate species, with *X. cortezi* ancestral to *X. malinche*. Alternatively, *X. malinche* may not be a separate species but rather a group of *X. cortezi* populations that has evolved some distinctive features. This second possibility is supported by the fact that the genetic versus geographic distance comparisons between the *X. malinche* population and the various *X. cortezi* populations lie very close to (and below rather than above) the best fitting line representing the comparisons among the *X. cortezi* populations (see Fig. 2). A third possibility is that the taxon currently recognized as *X. cortezi* consists of several different species (from which *X. malinche* may or may not be specifically distinct). Although a large component of the total genetic variation is attributable to differentiation between the populations from different drainages, the differences also conform closely to an isolation-by-distance model, suggesting the possibility of ongoing gene flow. Moreover, genetic divergence within *X. cortezi* (Table 5) is not nearly as great as divergence between *X. cortezi* and populations that clearly represent different species. Nonetheless, it is possible that sets of populations in the different drainages are no longer exchanging genes but have not been isolated long enough for greater differentiation to have occurred.

4.5. Relationship of X. cortezi and X. malinche to X. birchmanni

Contrary to previous findings (Rauchenberger et al., 1990), which placed *X. malinche* and *X. birchmanni* as sister species, with *X. cortezi* as their sister group, the results of our phylogenetic analyses suggest that *X. malinche* is more closely related to *X. cortezi* than to *X. birchmanni*. In our phylogenetic trees and networks, the haplotype from *X. malinche* (H11) is very closely related to the *X. cortezi* haplotypes in general, and to those from the southeastern branches of the Tempoal drainage (H7–9) in particular (see previous section). In contrast, the *X. birchmanni* haplotype connects to the phylogenetic trees between a distantly related outgroup (*X. montezumae*) and the entire set of haplotypes from *X. cortezi* and *X. malinche*. Moreover, in both the haplotype trees and networks, the *X. cortezi* and *X. malinche* haplotypes form a tight group whose members are separated by relatively short branches. In contrast, the *X. birchmanni* haplotype connects to the trees via a much longer branch and is so divergent that is not connected to the statistical parsimony network of *X. cortezi* and *X. malinche* haplotypes.

As noted above, there is ambiguity concerning the root of the haplotype trees and the relationships of *X. birchmanni* and *X. montezumae* to the rest of the taxa in our study (see Section 4 of alternative root positions). However, if the tree root position indicated by the outgroup *X. multilineatus* is correct, then the *X. malinche* and *X. birchmanni* haplotypes are most closely related to different *X. cortezi* haplotypes (H7–9 versus H6, respectively). This result is congruent with the results obtained from several different analyses of allozyme data for the northern swordtails, which place *X. malinche* and *X. birchmanni* populations closest to different populations of *X. cortezi* (Morris et al., 2001; esp. Fig. 4). Alternatively, if the root position indicated by the outgroup *X. montezumae* is correct, then the *X. malinche* haplotypes are once again closest to some of the *X. cortezi* haplotypes from the Tempoal drainage (H7–9), but the *X. birchmanni* haplotype is outside of a monophyletic group composed of the *X. malinche* and all the *X. cortezi* haplotypes. This result is congruent with those of a previous analysis of *Xiphophorus* DNA sequence data (Meyer et al., 1994), which placed either *X. birchmanni* (when analyzed by neighbor-joining) or a group composed of *X. birchmanni* and *X. pygmaeus* (when analyzed by parsimony) as the sister group to all the rest of the northern swordtails.

Sequence divergence corroborates the results obtained from the phylogenetic analyses (Table 5). Maximum-likelihood (HKY + I + G) distances between the *X. cortezi* and *X. malinche* haplotypes (0.0012–.0111) are substantially

Table 5
Pairwise total base differences (above diagonal) and maximum-likelihood (HKY + I + G model) distances (below diagonal) for mitochondrial haplotypes in *X. cortezi* (H1–H10) and four outgroup species (H10, *X. birchmanni*; H11, *X. malinche*; H12, *X. montezumae*; H13, *X. multilineatus*)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13
H1	—	1	1	2	1	7	6	7	5	34	4	32	14
H2	0.00115	—	2	3	2	8	7	8	6	35	5	33	15
H3	0.00116	0.00235	—	1	2	6	7	8	6	33	5	30	15
H4	0.00233	0.00356	0.00114	—	1	7	8	9	7	32	6	29	16
H5	0.00114	0.00233	0.00233	0.00116	—	8	7	8	6	33	5	30	15
H6	0.00849	0.00988	0.00715	0.00841	0.00979	—	11	10	10	33	9	30	17
H7	0.00713	0.00847	0.00848	0.00976	0.00839	0.01377	—	1	1	31	2	34	14
H8	0.00838	0.00975	0.00976	0.01106	0.00966	0.01242	0.00114	—	2	32	3	33	13
H9	0.00590	0.00721	0.00721	0.00848	0.00714	0.01242	0.00114	0.00231	—	30	1	33	13
H10	0.05526	0.05818	0.05253	0.05062	0.05338	0.05255	0.04938	0.05121	0.04746	—	30	37	38
H11	0.00469	0.00597	0.00597	0.00721	0.00590	0.01111	0.00231	0.00349	0.00115	0.04753	—	31	12
H12	0.05214	0.05502	0.04681	0.04495	0.04755	0.04755	0.05611	0.05400	0.05410	0.06692	0.04941	—	36
H13	0.01850	0.02022	0.02025	0.02176	0.02000	0.02337	0.01849	0.01701	0.01704	0.06770	0.01563	0.06414	—

Estimates for the number of mutational differences between pairs of haplotypes that correct for multiple substitutions at individual sites can be obtained by multiplying the maximum-likelihood distances (number of substitutions/site) by the number of sites (in this case 881).

smaller than those between both the *X. cortezi* and *X. birchmanni* haplotypes (0.0475–0.0582) and the *X. malinche* and *X. birchmanni* haplotypes (0.0475). In contrast, distances between the *X. cortezi* and *X. malinche* haplotypes and the haplotypes from *X. birchmanni* (0.0475–0.0582) are similar to the distances of the former two taxa from one relatively distantly related outgroup (*X. montezumae* 0.0450–0.0561), and smaller than those from another (*X. multilineatus* 0.0156–0.0234).

Because of limited sampling of other swordtail species and ambiguities concerning the position of the root, our results cannot rule out the possibility that *X. birchmanni* is the sister to a group composed of *X. malinche* and *X. cortezi*. Regardless, of whether these three species form a clade, *X. birchmanni* seems to be much more distantly related to *X. cortezi* and *X. malinche* than these two taxa are to each other.

Acknowledgments

We are grateful to the Mexican government for permission to collect the fishes. We would like to thank Oscar Rios-Cardenas and Abby Darrah for their help in the field. This research was supported by funding from the National Science Foundation (IBN 9983561) and Ohio University (Research Incentive) to MRM. Collecting methods comply with the Animal Care Guidelines of Ohio University (Animal Care and Use approval No. L01-01).

Appendix A. Results of phylogenetic analyses

A.1. Parsimony analyses

Of 881 unordered characters, 814 were constant, 67 were variable, and 25 were informative (42 uninformative) under the parsimony criterion with gaps treated as a fifth state. The branch and bound search found 12 optimal trees of length 86, the strict consensus of which had the following topology:

(H13,H6,(H4,H5,H3),H1,H2,(((H7,H8),H9),H11),(H10,H12)).

A.2. Likelihood analyses

Of 881 characters, 49 distinct character patterns were observed. The successive approximation heuristic search required three iterations, and although it identified two trees as maximum-likelihood estimates of the phylogeny under fixed parameter values, only one of these trees (Fig. 2) was identified as optimal when the parameters were estimated for that tree ($-\ln L = 1632.18538$). The parameter estimates from the final iteration of the successive approximation analysis on the optimal tree were ti/tv = 3.663520, $i = 0.783835$, and $\alpha = 0.792371$, frequency of $A = 0.308990$, $C = 0.227496$, $G = 0.138244$, $T = 0.325269$. The strict consensus of the two trees identified as optimal under fixed parameter estimates had the following topology:

(H13,(((H6,(H10,H12)),H4,H3),H5,H1,H2),(((H7,H8),H9),H11)).

A.3. Distance analyses

Using HKY85 + G + I distances (with the same parameter estimates used in the first iteration of the likelihood analysis) and an unweighted least squares optimality criterion, the heuristic search found 12 optimal trees with a score of 0.00026. The strict consensus of the 12 trees had the following topology:

(H13,((H6,(H10,H12)),(H4,H5,H1,H2,H3)),((H7,H8),H9),H11). In the individual optimal trees, the haplotype from *Xiphophorus malinche* (H11) always grouped with one group (H1–6) or the other (H7–9) of the *X. cortezi* haplotypes, resulting in an unresolved position in the strict consensus tree.

References

Aboim, M.A., Menezes, G.M., Schlitt, T., Rogers, A.D., 2005. Genetic structure and history of populations of the deep-sea fish *Helicolenus dactylopterus* (Delaroche, 1809) inferred form mtDNA sequence analysis. Mol. Ecol. 14, 1343–1354.

Alves, J., Coelho, M.M., 1994. Genetic variation and population subdivision of the endangered Iberian cyprinid *Chondrostoma lusitanicum*. J. Fish Biol. 44, 627–636.

Bohonak, A.J., 2002. IBD (Isolation by Distance): a program for population genetic analyses of isolation by distance. J. Hered. 93, 153–154.

Borowski, R.L., McClelland, M., Cheng, R., Welsh, J., 1995. Arbitrarily primed DNA fingerprinting for phylogenetic reconstruction in vertebrates: the *Xiphophorus* model. Mol. Biol. Evol. 12, 1022–1032.

Carvalho, G.R., Shaw, P.W., Magurran, A.E., Sehers, B.H., 1991. Marked genetic divergence revealed by allozymes among populations of the guppy *Poecilia reticulata* (Poeciliidae), in Trinidad. Biol. J. Linn. Soc. 57, 219–234.

Carvalho, G.R., 1993. Evolutionary aspects of fish distribution: genetic variability and adaptation. J. Fish Biol. 43, 53–73.

Carvalho, G.R., Shaw, P.W., Hauser, L., 1996. Artificial introductions, evolutionary change and population differentiation in Trinidadian guppies (*Poecilia reticulata*: Poeciliidae). Biol. J. Linnean Soc. 57, 219–234.

Castelloe, J., Templeton, A.R., 1994. Root probabilities for intraspecific gene trees under neutral coalescent theory. Mol. Phylogenet. Evol. 3, 102–113.

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

Coelho, M.M., Alves, J., Rodrigues, E., 1997. Patterns of genetic divergence in *Chondrostoma lusitanicum* Collares-Pereira, in intermittent Portuguese rivers. Fish Manage. Ecol. 4, 223–232.

Excoffier, L., Smouse, P., Quattro, J., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479–491.

Fajen, A., Breden, F., 1992. Mitochondrial DNA sequence variation among natural populations of the Trinidad guppy, *Poecilia reticulata*. Evolution 46, 1457–1465.

Hänfling, B., Brandl, R., 1998a. Genetic and morphological variation in a common European cryprinid, *Leuciscus cepahlus* within and across Central European drainages. J. Fish Biol. 52, 706–715.

Hänfling, B., Brandl, R., 1998b. Genetic variability, population size and isolation of distinct populations in the freshwater fish *Cottus gobio* L. Mol. Ecol. 7, 1625–1632.

Hänfling, B., Brandl, R., 1998c. Genetic differentiation of the bullhead *Cottus gobio* L. across watersheds in Central Europe: evidence for two taxa. Heredity 80, 110–117.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29, 170–179.

Lee, W-J., Conroy, J., Howell, W.H., Kocher, T.D., 1995. Structure and evolution of teleost micochondrial control regions. J. Mol. Evol. 41, 54–66.

Meffe, G.K., Snelson, F.F., 1989. An ecological overview of poeciliid fishes. In: Meffe, G.K., Snelson, F.F. (Eds.), Ecology and Evolution of Livebearing Fishes (Poeciliidae). Prentice Hall, Englewood Cliffs, pp. 13–31.

Mesquita, M., Carvalho, G., Shaw, P., Crespo, E., Coelho, M.M., 2001. River basin-related genetic structuring in an endangered fish species, *Chondrostoma lusitanicum*, based on mtDNA sequencing and RFLP analysis. Heredity 86, 253–264.

Meyer, A., Morrissey, J., Shartl, M., 1994. Recurrent evolution of a sexually selected trait in *Xiphophorus* fishes inferred from a molecular phylogeny. Nature 368, 539–542.

Meyer, A., 1997. The evolution of sexually selected traits in male swordtail fishes (*Xiphophorus*: Poeciliidae). Heredity 79, 329–337.

Moritz, C., 2002. Strategies to protect biological diversity and the evolutionary process that sustain it. Syst. Biol. 51, 238–254.

Morris, M.R., de Queiroz, K., Morizot, D.C., 2001. Phylogenetic relationships among populations of northern swordtails (*Xiphophorous*) as inferred form allozyme data. Copeia 1, 65–81.

Nei, M., Maruyama, T., Chakraborty, R., 1975. The bottleneck effect and genetic variability in populations. Evolution 29, 1–10.

Nei, M., Li, W.H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76, 5269–5273.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Rauchenberger, M., Kallman, K.D., Morizot, D.C., 1990. Monophyly and geography of the Río Pánuco basin swordtails (Genus *Xiphophorus*) with descriptions of four new species. Am. Mus. Nov. 2975, 1–41.

Rosen, D.E., 1979. Fishes from the uplands and intermontane of Guatemala: reversionary studies and comparative geography. Bull. Am. Mus. Nat. Hist. 162, 268–375.

Ryan, M.J., Rosenthal, G.G., 2001. Variation and selection in swordtails. In: Dugatkin, L.A. (Ed.), Model Systems in Behavioral Ecology. Princeton University Press, New Jersey, pp. 133–148.

Salzburger, W., Brandästter, A., Gilles, A., et al., 2003. Phylogeography of the vairone (*Leuciscus souffia*, Risso 1826) in Central Europe. Mol. Ecol. 12, 2371–2386.

Schneider, S., Roessli, D., Excoffier, L., 2000. ARLEQUIN, Version 2.00: a Software for Population Genetics Data Analysis. University of Geneva, Geneva.

Seutin, G., White, B.N., Boag, P.T., 1991. Preservation of avian blood and tissue samples for DNA analyses. Can. J. Zool. 69, 82–92.

Shaw, P.W., Carvalho, G.R., Magurran, A.E., Seghers, B.H., 1991. Population differentiation in Trinidad guppies (*Poecilia reticulata*): patterns and problems. J. Fish Biol. 39 (suppl. A), 203–209.

Shaw, P.W., Carvalho, G.R., Seghers, B.H., Magurran, A.E., 1992. Genetic consequences of an artificial introduction of guppies (*Poecilia reticulata*) in N. Trinidad. Proc. R. Soc. B. 248, 111–116.

Shaw, P.W., Carvalho, G.R., Magurran, A.E., Seghers, B.H., 1994. Factors affecting the distriution of genetic variability in the guppy, *Poecilia reticulata*. J. Fish Biol. 45, 875–888.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Stefanni, S., Thorley, J., 2003. Mitochondrial DNA phylogeography reveals the existence of an evlutionary significant unit of the sand goby *Pomatoschistus minutus* in the Adriatic (Eastern Mediterranean). Mol. Phylogenet. Evol. 28, 601–609.

Sullivan, J., Abdo, Z., Joyce, P., Swofford, D.L., 2005. Evaluating the performance of a successive approximations approach to parameter optimization in maximum-likelihood phylogeny estimation. Mol. Biol. Evol. 22, 1386–1392.

Swofford, D.L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4b10. Sinauer Associates, Sunderland, MA.

Swofford, D.L., Olsen, G.J., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic inference. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer Associates, MA, pp. 407–514.

Tajima, F., 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105, 437–460.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512–526.

Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132, 619–633.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignments aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

Topografix 2003. ExpertGPS, Copyright 2000–2006, TopoGrafix, Stow MA.

Weir, B.S., Cockerham, C., 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38, 1358–1370.

Wishard, L.N., Seeb, J.E., Utter, F.M., Stefan, D., 1984. A genetic investigation of suspected redband trout populations. Copeia 1984, 120–132.