
Phylogeography and monophyly of the swordtail fish species *Xiphophorus birchmanni* (Cyprinodontiformes, Poeciliidae)

CARLA GUTIÉRREZ-RODRÍGUEZ, A. ELIOT SHEARER, MOLLY R. MORRIS & KEVIN DE QUEIROZ

Submitted: 15 May 2007

Accepted: 4 October 2007

doi:10.1111/j.1463-6409.2007.00311.x

Gutiérrez-Rodríguez, C., Shearer, A. E., Morris, M. R. & de Queiroz, K. (2008). Phylogeography and monophyly of the swordtail fish species *Xiphophorus birchmanni* (Cyprinodontiformes, Poeciliidae). — *Zoologica Scripta*, 37, 129–139.

We used sequences of the mitochondria control region to assess the distribution of genetic variation within and among populations of the poeciliid fish species *Xiphophorus birchmanni*. We collected 122 *X. birchmanni* samples from 11 sites in three drainage systems comprising the distribution of the species. We found low levels of polymorphism among aligned sequences and low levels of genetic variation within populations but high levels of genetic differentiation among populations. Haplotypes are exclusive to three river drainages (Los Hules, Calabozo and San Pedro). Mantel tests revealed correlations between geographical (both straight-line and river distances) and genetic distance, consistent with an isolation by distance scenario, while nested clade analysis suggested allopatric fragmentation between haplotypes from two of the major drainages, and isolation by distance with restricted gene flow within those drainages. Finally, monophyly of *X. birchmanni* is strongly supported while the previous hypothesis of the evolutionary origin of this species from *X. malinche* is not.

Corresponding author: *Carla Gutiérrez-Rodríguez, 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. E-mail: carla.gutierrez@inecol.edu.mx*

A. Eliot Shearer and Molly R. Morris, Department of Biological Sciences, Ohio University, Athens OH 45701, USA. E-mails: as257400@ohio.edu; mormism@ohiou.edu

Kevin de Queiroz, Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA. E-mail: dequeirozk@si.edu

Introduction

The swordtails and platyfishes (*Xiphophorus*) have been the focus of extensive research in diverse areas, including genetics, biogeography, endocrinology, behaviour and oncology (reviewed in Meffe & Snelson 1989). The relationships among three presumed clades within *Xiphophorus* (northern swordtails, southern swordtails and platyfish) as well as the relationships of the species within these groups have also been studied extensively (Rosen 1979; Rauchenberger *et al.* 1990; Meyer *et al.* 1994; Borowsky *et al.* 1995; Meyer 1997; Morris *et al.* 2001), though some of the relationships are still under debate (Meyer *et al.* 1994; Morris *et al.* 2001; Gutiérrez-Rodríguez *et al.* 2007). The most complete sampling and study of the northern swordtail clade was conducted over 15 years ago (Rauchenberger *et al.* 1990), before many of the current population genetic and phylogenetic methods were available. Of the more recent studies that have examined the phylogenetic relationships among the nine species within the northern swordtail clade (Meyer *et al.* 1994; Borowsky *et al.* 1995; Meyer 1997; Morris *et al.* 2001), there is still a lack of

consensus for some of the relationships. As several of the species are widely distributed, examining the genetic variation within those species can provide insights into their evolutionary history and lead to a better understanding of the relationships among species.

Recently, Gutiérrez-Rodríguez *et al.* (2007) investigated genetic variation within and among populations of the swordtail species *X. cortezi* in a phylogenetic and geographical (i.e. a phylogeographic) context. Based on sampling across the complete range of this species, they found low gene and nucleotide diversity within populations and high degrees of genetic differentiation among populations. In addition, phylogenetic analyses suggested that *X. cortezi* is paraphyletic relative to *X. malinche*, raising questions concerning the status of these taxa as separate species. While *X. birchmanni* has been considered to be most closely related to *X. malinche* within the *X. cortezi* clade (Rauchenberger *et al.* 1990; Borowsky *et al.* 1995) the placement of *X. birchmanni* has not been consistent across studies (Meyer *et al.* 1994; Morris *et al.* 2001). The *X. birchmanni* haplotype sampled by Gutiérrez-Rodríguez

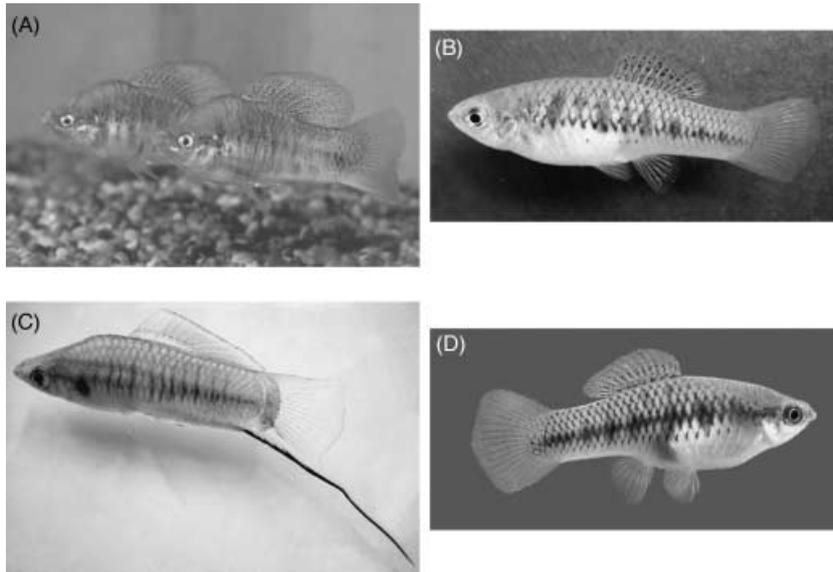


Fig. 1 A–D. Morphological characteristics nuchal hump (A) seen here on *X. birchmanni* male and oval vertical bars (B) seen here on *X. malinche* female are not synapomorphies of *X. birchmanni* and *X. malinche* as nuchal hump is also found on *X. multilineatus* males (C) and oval bars on *X. cortezi* females (D). Photos by K. de Queiroz (A), M. S. Tudor (B) (C) and M. R. Morris (D).

et al. (2007) was much more divergent from the *X. malinche* and *X. cortezi* haplotypes than any of those haplotypes were from each other, suggesting that a more detailed phylogeographical study of *X. birchmanni* species is warranted.

The most distinguishing morphological characteristics of the swordtail *X. birchmanni* are the absence of a caudal fin extension (i.e. sword) and the presence of a nuchal hump, which gives this species its common name of sheephead swordtail (Fig. 1A). Its distribution as described in Rauchenberger *et al.* (1990) is the third largest among the nine species in the northern swordtail clade (after *X. nezabualcoyotl* and *X. cortezi*), stretching over the uplands of the Río Tempoal drainage and the Río Tuxpan drainage in the states of Hidalgo & Veracruz, Mexico. Rauchenberger *et al.* (1990) considered *X. birchmanni* to be the sister species of *X. malinche* and hypothesized that the ancestral form of *X. birchmanni* entered the Río Tempoal drainage from an upper tributary of the Río Moctezuma (Río Claro) via stream capture and then migrated south-east to the Río Tuxpan, and north-west along the foot of the Sierra Madre, where it came into contact with *X. cortezi*.

In the current study, we completed a survey of most of the known localities of *X. birchmanni*, examining genetic variation within and across populations. The purpose of this study was to assess the distribution of genetic variation within and among populations of *X. birchmanni* to infer relationships among populations, and to test the hypothesis of monophyly for this species, as well as the previously mentioned hypothesis about its evolutionary origin. Finally, the data we present will be useful for further examining the relationships of *X. birchmanni* to *X. cortezi* and *X. malinche*.

Materials and methods

Sample collection

We collected a total of 122 *X. birchmanni* samples from 11 sites located in the states of Hidalgo and Veracruz, Mexico (Fig. 2, Table 1). Samples were collected in December 2004 except for those from Pezmatlán, which were collected in June 1999. The collection sites are located in three different river systems, Los Hules, Calabozo and San Pedro all within the Río Pánuco drainage system (Fig. 2, Table 1). The historical distribution of this species includes localities in the Río Vinazco and Río Beltrán drainages of the Río Tuxpan system (Rauchenberger *et al.* 1990), however, we did not find *X. birchmanni* in several of the rivers in this system that we sampled in 2004 including sites in the Río Vinazco at El Naranjal, a tributary of Río El Cerro near the town of El Terrero and Río El Cerro at the town of Huachuaco. Four specimens of *X. malinche* were collected in the Río Soyatla to use as outgroups in the phylogenetic analyses.

Samples were collected by either electrofishing or seining. After collection, the fish were anaesthetized using tricaine methane sulphonate (MS222), a fin clip was immediately taken and stored in salt-saturated 20% dimethyl sulphoxide solution (Seutin *et al.* 1991) for genetic analyses and the rest of the specimen was preserved in 95% ethanol. In 1999, whole fish were collected and preserved in 95% ethanol, a fin clip was taken immediately prior to DNA extraction in 2005.

mtDNA amplification and sequencing

Total DNA was extracted from the fin clips using the DNeasy tissue kit (Qiagen Inc., Valencia, CA, USA). The entire

Table 1 Description of *X. birchmanni* and outgroup species collection localities and measures of gene diversity.

No.	Locality abbreviation	River	Species	No. sequenced	Elevation (m)	N	W	Haplotypes (n)	h	π	River drainage
1	CAN	Candelaria	<i>X. birchmanni</i>	14	184.7	21.08109	98.40320	B1 (14)	0.000	0.0000	Los Hules
2	HUA	Huazalingo	<i>X. birchmanni</i>	16	163.4	21.03681	98.37305	B1 (11), B6 (5)	0.458	0.0005	
3	ATL	Atlapexco	<i>X. birchmanni</i>	8	164	21.01739	98.33978	B1 (8)	0.000	0.0000	
4	PEZ	Calnali	<i>X. birchmanni</i>	2	700–800	20.5437	98.3234	B1 (2)	0.000	0.0000	
5	GAR	Garces/Calabozo	<i>X. birchmanni</i>	14	195.3	20.94014	98.28164	B2 (13), B3 (1)	0.143	0.0000	Calabozo
6	TLA	Trib. of Encinal/Calabozo	<i>X. birchmanni</i>	13	347.2	20.85682	98.28227	B2 (13)	0.000	0.0000	
7	SAS	Trib. of Calabozo	<i>X. birchmanni</i>	16	277.4	20.93895	98.02093	B2 (1), B3 (9), B5 (6)	0.575	0.0006	
8	CHI	Trib. of Zontecomatlán	<i>X. birchmanni</i>	16	309.1	20.79609	98.18227	B2 (15), B3 (1)	0.125	0.0000	
9	OTL	Trib. of Zontecomatlán	<i>X. birchmanni</i>	7	406.3	20.77811	98.24849	B2 (7)	0.000	0.0000	
10	ZON	Zontecomatlán	<i>X. birchmanni</i>	11	496.2	20.76392	98.34348	B2 (11)	0.000	0.0000	
11	XIL	Xiliatl (headwater)	<i>X. birchmanni</i>	5	469.1	21.0957	98.56269	B4 (5)	0.000	0.0000	San Pedro
12	CLA	Claro	<i>X. malinche</i>	1	655.9	20.8809	98.79833	M1 (1)	—	—	Moctezuma
13	SOY	Soyatla	<i>X. malinche</i>	4	1292.7	20.68265	98.62003	M2 (4)	—	—	Calabozo
14	AMA	Amacuzac	<i>X. cortezi</i>	1		21.17603	98.71567	C1 (1)	—	—	Moctezuma

h, mean gene diversity; π, nucleotide diversity; Trib., tributary; —, calculations were not performed (outgroups).

CAN, Candelaria; HUA, Huazalingo; ATL, Atlapexco; PEZ, Pezmatlán; GAR, Garces; TLA, Tlalatlá; SAS, Sasaltitla; CHI, Chicontepec; OTL, Otlatzintla; ZON, Zontecomatlán; XIL, Xiliatl; CLA, Claro; SOY, Soyatla; AMA, Amacuzac.

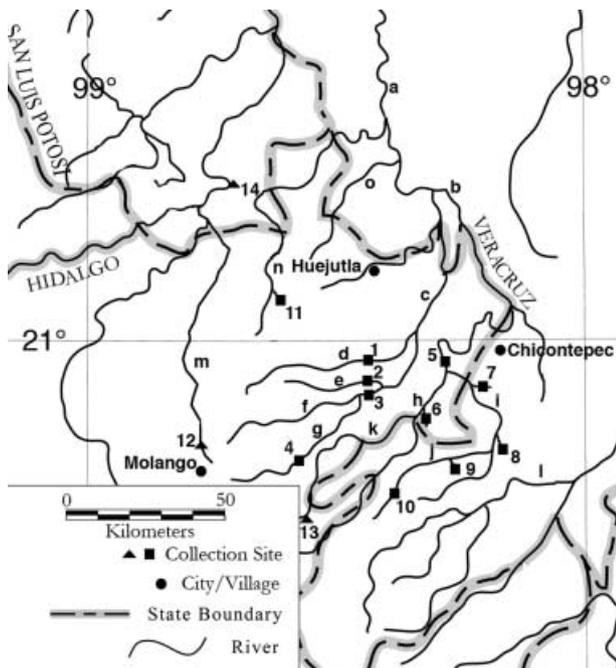


Fig. 2 Map showing collection localities for *X. birchmanni* and for the outgroups (adapted with permission from Morris *et al.* 2001). Cities are labelled with filled circles, rivers and streams by lower case letters, *X. birchmanni* localities are labelled with squares, and outgroup localities with triangles. Numbers correspond with those in Table 1. The rivers and streams are as follows: a, R. Tempoal; b, R. Calabozo; c, R. Los Hules; d, R. La Candelaria; e, R. Huazalingo; f, R. Atlapexco; g, R. Calnali; h, R. Garces; i, R. Calabozo; j, R. Zontecomatlán; R. Otlatzintla; k, R. Encinal; l, R. Vinazco; m, R. Claro; n, R. Xiliatl; o, R. San Pedro.

mitochondrial control region was then amplified using the polymerase chain reaction (PCR). Samples collected in 2004 were amplified using the forward primer K (5'-AGCTCAGCGCCAGAGCGCCGGTCTTGTA-3') and reverse primer G (5'-CGTCGGATCCCATCTTCAGTGTATGCTT-3'), previously described by Lee *et al.* (1995). PCR consisted of 50-μL reactions containing 5 μL of 10× reaction buffer, 4 μL of dNTP mix (10 mM), 2 μL of MgCl₂ (50 mM), 2.5 μL of each primer (5 μM), 0.2 μL of Taq (5 U/μL) polymerase and approximately 10 ng of template DNA. Reactions were performed under the following conditions: one denaturation cycle at 94 °C for 2 min, 30 cycles of 30 s denaturation at 90 °C, 30 s annealing at 53 or 55 °C, and 45 s extension at 72 °C, followed by a final 7 min extension at 72 °C.

For samples collected in 1999, DNA was amplified using four primers because of DNA degradation: forward primers were L15995n (5'-AACCTCCRCYCTAACTCCCAAAG-3') and L16378n1 (5'-ATGYAGTAAGARACCA-3'); reverse primers were H16498n (5'-GGGTAAYGAGGAGTATG-3') and G (see above). These primers are modifications of L15995 (Meyer *et al.* 1994), L16378 (Faber & Stepien 1997) and H16498 (Meyer *et al.* 1990), respectively, that were designed to anneal more specifically to our samples. The total reaction volume, reagent amounts and concentrations were the same as for the 2004 samples except for the amount of the primers, which was doubled to 5 μL (5 μM). PCR conditions for these samples were as described above except for the number of cycles (40) and the annealing temperature (48 °C). All PCR reactions were conducted on a MJ Research PTC-100 thermocycler (GMI, Inc., Ramsey, MN, USA).

PCR products were purified with the QIAquick PCR Purification Kit (Qiagen, Inc., Valencia, CA, USA) and sequenced using the BigDye terminator cycle sequencing kit and read in a 3730 automated sequencer (Applied Biosystems, Foster City, CA, USA) at the Ohio State University Plant Microbe sequencing facility. Sequences were performed in both directions for 37 of the 122 samples collected in 2004 and no ambiguities were discovered, thus the remaining 85 samples were sequenced using only the forward K primer. The two samples collected in 1999 were sequenced with the two forward and the two reverse primers, as described above.

Forward and reverse sequences were aligned using the computer program SEQMAN II (DNASTar Inc., Madison, WI, USA). Sequences were aligned using the default settings in CLUSTAL X v.1.83.1 (Thompson *et al.* 1997) followed by manual alignment using the computer program SEQUENCE ALIGNMENT EDITOR v2.0a11. All haplotypes were deposited in GenBank (Accession numbers: EF533642–EF533649).

Population genetics analyses

The following population genetic analyses were run in ARLEQUIN 2.0 (Schneider *et al.* 2000). We estimated intrapopulation variation by calculating mean gene diversity (h) and nucleotide diversity (π). Two analyses of molecular variance (AMOVA) (Excoffier *et al.* 1992) were performed using the Tamura-Nei (TrN) model to calculate the genetic distances (Tamura & Nei 1993) with the transition : transversion (ti : tv) ratio = 3.5878 as obtained from MODELTEST (see below) and the gamma shape parameter (α) = 0. HKY + I, which is the model of best fit for our data (see below), is not an option in ARLEQUIN 2.0, and TrN is the closest available approximation to HKY. A total of 16 000 permutations were calculated to assess the significance of each AMOVA. In one of the AMOVAs we used groups that corresponded to the three drainages where the collection sites are located (Los Hules, Calabozo and San Pedro, Table 1, Fig. 2), and in the other one, we did not define groups. Genetic differentiation between pairs of populations was calculated with F -statistics (Weir & Cockerham 1984) and their significance was determined by performing 1000 permutations. In order to test for a relationship between geographical and genetic distance, we performed two Mantel tests, each one with 10 000 permutations using the computer program IBD 1.52 (Bohonak 2002). We used Nei's average number of pairwise differences between populations (Nei & Li 1979) as the genetic distance and river distances (the paths along the water courses connecting two localities) and straight-line distances (the minimum great circle distance between two localities) as geographical distances. The computer software packages ARLEQUIN 2.0 (Schneider *et al.* 2000) and EXPERTGPS 1.3.7 (Topografix 2003) were used to calculate genetic and geographical (river and straight) distances between sites, respectively.

Network estimation and nested clade analysis

A statistical parsimony network (Templeton *et al.* 1992; Crandall *et al.* 1994) connecting the haplotypes was inferred with TCS v.1.21 (Clement *et al.* 2000), treating gaps as a fifth state and using the default 0.95 probability connection limit. We included the outgroup sequences that were used in the phylogenetic analyses (see below). The resulting network was used to construct the nested clade design following the procedure in Templeton *et al.* (1995). Based on the resulting nested clade design, we used the software GEO DIS vs. 2.5 (Posada *et al.* 2000) to test for geographical structuring of genetic variation by evaluating the geographical distribution of the haplotypes. This software uses the geographical coordinates to calculate the clade distance (D_c), which measures how widely distributed are the haplotypes within a given clade, and the nested clade distance (D_n), which determines the geographical distribution of clades relative to other clades within the same nesting category. We performed this test using both river distances and straight-line distances (that we specified) and 10 000 bootstrap resamples to test for significant departures of these two measures from expected values under a null hypothesis of random distribution. Outgroup weights were not included because they disagreed with the root position suggested by outgroups; interior and tip status of haplotypes was based on the root position inferred from outgroups. We used the GEO DIS inference key (Templeton *et al.* 1987 version of 11 November 2005) to infer causal processes (e.g. restricted gene flow, population fragmentation, range expansion).

Phylogenetic analyses

We analysed the phylogenetic relationships among the distinct haplotypes with maximum likelihood (ML) using PAUP* 4.0b10 (Swofford 2002). The model of molecular evolution was selected based on information from hierarchical likelihood ratio tests and the Akaike information criterion (AIC) as implemented in MODELTEST (Posada & Crandall 1998). We used the four sequences of *X. malinche* from the Río Soyatla collected in the present study as well as one sequence of *X. malinche* from the Río Claro and one sequence of *X. cortezi* from the Río Amacuzac previously sequenced for use in other studies (Dries *et al.* in prep.; Gutiérrez-Rodríguez *et al.* 2007) as predefined outgroups.

The ML analyses were performed using the technique of successive approximation (Swofford *et al.* 1996; Sullivan *et al.* 2005), starting with the parameter values obtained on neighbor-joining trees used for model selection and re-estimating the parameter values on the optimal tree obtained from each ML search until the estimated parameter values and tree topology did not change in successive iterations. Exhaustive searches were used in all iterations. Support for monophyletic groups of haplotypes was estimated using a non-parametric bootstrap analysis, which was performed with 1000 replicates,

branch and bound searches, and using the final parameter estimates for the HKY + I model (see below) obtained in the successive approximations search.

Tests of hypotheses corresponding to specific tree topologies were performed by comparing optimal trees constrained to correspond to those hypotheses with optimal trees in the absence of those constraints. Trees were constructed with only those groups relevant to the specific hypothesis being tested (e.g. monophyly of *X. birchmanni* haplotypes) and loaded as topological constraints into PAUP*. A successive approximations analysis was then conducted using the HKY + I model with starting parameter estimates as in the unconstrained analysis, exhaustive searches, and keeping only trees compatible or incompatible with the constraint (depending on the hypothesis being tested). The best tree obtained under the constraint was then compared with the best tree obtained in the absence of the constraint (i.e. the ML tree) using the significance tests proposed by Kishino & Hasegawa (1989) (one-tailed test) and Shimodaira & Hasegawa (1999) using full optimization and 1000 bootstrap replicates.

Results

Sequence variation and description of haplotypes

Mitochondrial control region sequences used for the analyses consisted of 861 aligned characters for all haplotypes including outgroups. A total of 40 (4.65%) variable sites among the *X. birchmanni* and the outgroup species were found and only 6 (0.92%) within *X. birchmanni* sequences were detected. Mutations among *X. birchmanni* sequences include an indel at position 17, three C–T transitions at positions 110, 563 and 699, and two A–T transversions at positions 603 and 856.

Sequence variation among *X. birchmanni* haplotypes was low, as a single mutation separated five of the six haplotypes from their nearest neighbours, and four mutational steps separated farthest neighbours (Fig. 3). Most of the collection localities (7 out of 11) contained just one haplotype; the other four localities, Huazalingo, Garces, Chila and Sasaltitla had either two or three haplotypes (Table 1). All six haplotypes are restricted to particular river drainages (Table 1). Haplotypes B1 and B6 were found only in the populations located in the Los Hules drainage, with B1 occurring in all four of those populations and B6 in only one of them. The most widely distributed haplotype, B2, was present in all the populations from the Calabozo drainage; haplotypes B3 and B5 were detected in three and one of the four populations of the Calabozo drainage, respectively. Haplotype B4 was located at the most northern site, Xiliatl, which is located in the Río San Pedro drainage. All of our samples from the different sites were specimens that were identified as *X. birchmanni* based on morphology, and none had a haplotype more similar to those of *X. malinche*.

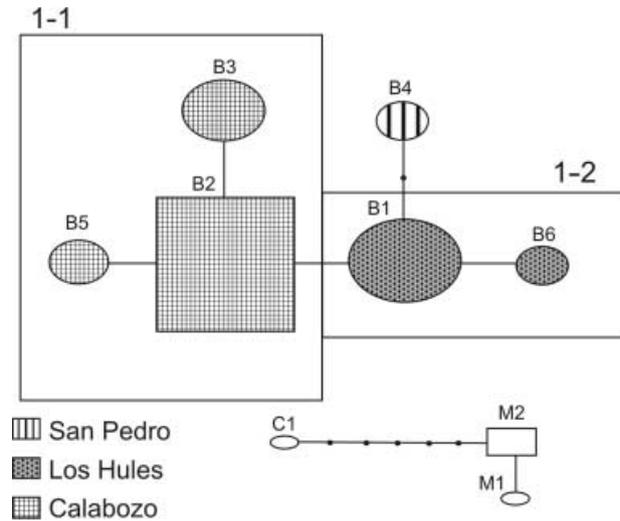


Fig. 3 Statistical parsimony haplotype network created with *trcs* software with gaps treated as a fifth state. Each line corresponds to one mutational change. The size of the ellipse for each haplotype corresponds to the haplotype's frequency. The square represents the ancestral haplotype inferred by *trcs*. Unobserved haplotypes are represented by a small circle. Boxes labelled 1-1 and 1-2 enclose groups of haplotypes used in the Nested Clade Analysis.

Table 2 Hierarchical analysis of molecular variance (AMOVA) of *X. birchmanni*.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
AMOVA with drainage grouping				
Among groups	2	38.75	0.635 (Va)	87.67*
Among populations within groups	8	2.73	0.024 (Vb)	3.37*
Within populations	111	7.21	0.065 (Vc)	8.96*
Total	121	48.68	0.724	
AMOVA without grouping				
Among population	10	41.48	0.375 (Va)	85.23*
Within populations	111	7.21	0.065 (Vb)	14.77
Total	121	48.69	0.440	

*Significant values at $P < 0.05$.

Population genetics

All but four populations (Huazalingo, Garces, Sasaltitla and Chila) had gene diversity values (b) of zero reflecting the presence of a single haplotype in most populations (Table 1). Nucleotide diversity (π) was also zero in most populations, with the exception of Huazalingo and Sasaltitla, indicating little variation among sequences from the same locality. The results of the AMOVA grouping by drainage (Los Hules, Calabozo and San Pedro) revealed significant genetic differentiation at all hierarchical levels (Table 2). While most of the variation (87.67%) was explained by differences among

Table 3 Pairwise F_{ST} comparisons (above the diagonal) and Nei & Li (1979) genetic distances (below the diagonal) among *X. birchmanni* populations. Abbreviations refer to those in Table 1.

		Los Hules				Calabozo						San Pedro
		CAN	HUA	ATL	PEZ	GAR	TLA	SAS	CHI	OTL	ZON	XIL
Los Hules	CAN	—	0.249*	0.000	0.000	1.000*	1.000*	0.808*	1.000*	1.000*	1.000*	1.000*
	HUA	0.313	—	0.179	0.000	0.815*	0.809*	0.717*	0.826*	0.765*	0.797*	0.850*
	ATL	0.000	0.313	—	0.000	1.000*	1.000*	0.765*	1.000*	1.000*	1.000*	1.000*
	PEZ	0.000	0.313	0.000	—	1.000*	1.000*	0.683*	1.000*	1.000*	1.000*	1.000
Calabozo	GAR	3.659	3.972	3.659	3.659	—	0.000	0.314*	0.000	0.000	0.000	1.000*
	TLA	3.588	3.900	3.588	3.588	0.071	—	0.304*	0.000	0.000	0.000	1.000*
	SAS	5.496	5.808	5.496	5.496	1.899	1.908	—	0.333*	0.225	0.282	0.887*
	CHI	3.650	3.963	3.650	3.650	0.125	0.063	1.900	—	0.000	0.000	1.000*
	OTL	3.588	3.900	3.588	3.588	0.071	0.000	1.908	0.063	—	0.000	1.000*
	ZON	3.588	3.900	3.588	3.588	0.071	0.000	1.908	0.063	0.000	—	1.000*
San Pedro	XIL	4.588	4.900	4.588	4.588	8.247	8.176	10.084	8.238	8.176	8.176	—

the drainages, a much smaller proportion was attributed to variation within populations (8.96%) and to variation among populations within groups (3.37%). In the AMOVA performed without groupings, 85.23% of the variation was due to differences among populations and 14.77% was the result of variation within populations (Table 2). A pairwise F_{ST} analysis between populations showed that all populations from different drainages were significantly differentiated and most comparisons between populations from the same drainage were not significantly differentiated (Table 3). Pairwise comparisons of populations within the same drainage that were significant include populations located in Río Los Hules (Candelaria-Huazalingo) and in Río Calabozo (Garces-Sasalitla, Sasalitla-Tlalteatlá and Salsalitla-Chila). Mantel tests comparing genetic distances to river distances and straight-line distances among the populations were significant in each case ($r = 0.823$, $P < 0.01$, and $r = 0.747$, $P < 0.01$, respectively). None of the data points in either analysis fell outside of the 95% confidence intervals (graphs not shown).

Haplotype network and nested clade analysis

The statistical parsimony network (Fig. 3) shows the most parsimonious connections between the haplotypes as well as the frequencies of each haplotype. The six *X. birchmanni* haplotypes are very closely related to each other, with only one or two mutations between nearest neighbours, and only four mutations between farthest neighbours. In contrast, the outgroup (*X. cortezi* and *X. malinche*) haplotypes form a separate (i.e. unconnected) statistical parsimony network, indicating that they are separated by a sufficiently large number of mutational steps that the exact number is likely to be underestimated by parsimony. The two *X. malinche* haplotypes (M1 and M2) were separated by just one mutational step, five and six mutations separated *X. malinche* haplotypes M2 and M1 from the *X. cortezi* haplotype (C1), respectively.

Haplotype B2 is inferred to be ancestral according to outgroup weights (Castelloe & Templeton 1994), which are based on both haplotype frequencies and positions (interior vs. tip) in the graph.

The results of the nested clade analysis (NCA) are summarized in Table 4. In general, haplotypes and groups of haplotypes were non-randomly distributed geographically (Chi-squared contingency tests). More specifically, haplotype B2 within group 1-1 had a significantly larger geographical spread, haplotypes B3 and B5 within group 1-1 and haplotype B6 within group 1-2 had significantly smaller geographical spreads, and groups 1-1 and 1-2 within the total haplotype tree had significantly smaller geographical spreads than expected by chance (D_c probabilities). Interpretation of the patterns of both clade and nested clade distances (D_c and D_n) in the context of the NCA inference key resulted in inferences of restricted gene flow with isolation by distance within groups 1-1 and 1-2 and of allopatric fragmentation between the two haplotype groups.

Phylogenetic analyses

Hierarchical likelihood ratio tests identified HKY as the most appropriate model for our data. In contrast, the AIC identified GTR + I as the most appropriate model. Use of GTR for our data set is questionable because of the small number (35 with gaps treated as missing data) of variable sites, which means that some of the rate matrix parameters are unlikely to be estimated accurately (particularly C–G and G–T transversions, of which there were no observed instances). Moreover, under the AIC, all models within 2.0 AIC units of the highest scoring model are to be considered candidate models (Burnham & Anderson 2004), and these included both HKY + I (delta = 0.7705) and HKY + Γ (delta = 1.2520), both of which have only two rate parameters compared to six for GTR + I. Finally, the hierarchical likelihood

Table 4 Results of the Nested Clade Analysis using river distances. Group designations (1-1 and 1-2) correspond to those on the statistical parsimony network (Fig. 3). Abbreviations are as follows: D_o , observed distance; D_e , expected distance (based on permuted data); D_c , within clade distance; D_n , nested clade distance. Inferences about the processes underlying the observed patterns are as follows (parenthetical series of numbers indicate steps in the Nested Clade Analysis inference key): Group 1-1 = restricted gene flow with isolation by distance (1-2-3-4); Group 1-2 = restricted gene flow with isolation by distance (1-2-3-4); Total haplotype tree = allopatric fragmentation (1-19).

Group 1-1			
Chi-squared contingency test			
$\chi^2 = 61.6521$			$P = 0.0000$
Geographic distance analysis:			
Haplotype B2 (Interior)			
Type of distance	Distance (D_o)	$P(D_o < D_e)$	$P(D_o > D_e)$
D_c	33.5403	0.9962	0.0038
D_n	32.9841	1.0000	0.0000
Haplotype B3 (Tip)			
Type of distance	Distance (D_o)	$P(D_o < D_e)$	$P(D_o > D_e)$
D_c	12.8243	0.0003	0.9997
D_n	27.0289	0.0087	0.9913
Haplotype B5 (Tip)			
Type of distance	Distance (D_o)	$P(D_o < D_e)$	$P(D_o > D_e)$
D_c	0.0000	0.0001	1.0000
D_n	26.6876	0.0353	0.9647
Interior vs. Tip clades:			
Type of distance	Distance (D_o)	$P(D_o < D_e)$	$P(D_o > D_e)$
D_c	25.2422	1.0000	0.0000
D_n	6.0757	1.0000	0.0000
Group 1-2			
Chi-squared contingency test:			
$\chi^2 = 8.5714$			$P = 0.0366$
Geographic distance analysis:			
Haplotype B1 (Interior)			
Type of distance	Distance (D_o)	$P(D_o < D_e)$	$P(D_o > D_e)$
D_c	22.0280	0.9211	0.0870
D_n	21.0407	0.9211	0.0870
Haplotype B6 (Tip)			
Type of distance	Distance (D_o)	$P(D_o < D_e)$	$P(D_o > D_e)$
D_c	0.0000	0.0101	1.0000
D_n	14.4865	0.0635	0.9446
Interior vs. Tip clades:			
Type of distance	Distance (D_o)	$P(D_o < D_e)$	$P(D_o > D_e)$
D_c	22.0280	1.0000	0.0081
D_n	6.5542	0.9429	0.0652
Total haplotype tree			
Chi-squared contingency test:			
$\chi^2 = 117.0000$			$P = 0.0000$
Geographic distance analysis:			
Clade 1-1 (Tip)			
Type of distance	Distance (D_o)	$P(D_o < D_e)$	$P(D_o > D_e)$
D_c	32.1752	0.0000	1.0000
D_n	49.3010	0.0000	1.0000
Clade 1-2 (Interior)			
Type of distance	Distance (D_o)	$P(D_o < D_e)$	$P(D_o > D_e)$
D_c	20.9988	0.0000	1.0000
D_n	61.6999	1.0000	0.0000
Interior vs. Tip clades:			
Type of distance	Distance (D_o)	$P(D_o < D_e)$	$P(D_o > D_e)$
D_c	-11.1764	0.0084	0.9916
D_n	12.3989	1.0000	0.0000

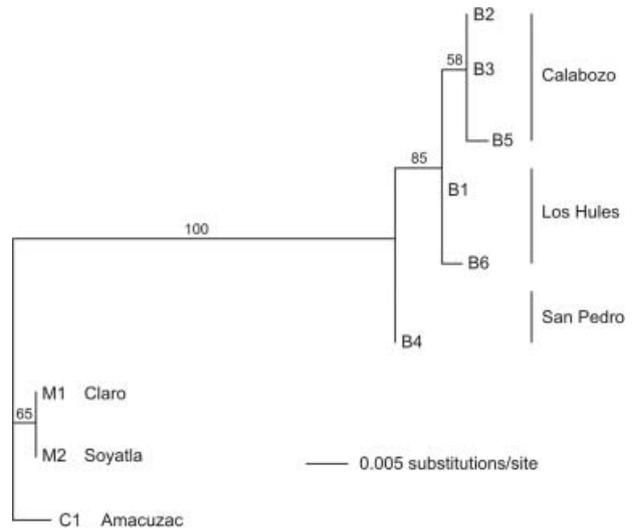


Fig. 4 Maximum likelihood phylogram illustrating haplotype relationships and their geographical distributions (river drainages). Bootstrap proportions greater than 50% are shown above branches.

ratio tests indicated that addition of a parameter for either invariant sites (I) or rate variation among sites (Γ) yielded a significant improvement to the HKY model at $\alpha = 0.05$ but not at $\alpha = 0.01$ (the significance level adopted in MODELTEST), with a greater improvement for I ($2[\ln L1 - \ln L0] = 4.4883$; $P = 0.017064$) than for Γ ($2[\ln L1 - \ln L0] = 4.0068$; $P = 0.022658$). Therefore, we selected HKY + I as the most appropriate model for our data (though the ML trees under this model had the same topology and exhibited only minor differences in branch lengths from those under HKY and GTR + I). The successive approximations ML analyses under the HKY + I model yielded a single tree ($-\ln L = 1368.37718$, Fig. 4) and the following parameter estimates: base frequencies, A = 0.308863, C = 0.223824, G = 0.143183, T = 0.324130; TiTv ratio = 3.587839 ($\kappa = 7.610642$); I = 0.845168. In this tree, the *X. birchmanni* haplotypes are separated by short branches and exhibit the same relationships as in the statistical parsimony network. Outgroups connect to the *X. birchmanni* haplotypes through haplotype B4, suggesting that this is the ancestral *X. birchmanni* haplotype. The branch length separating the *X. birchmanni* and outgroup haplotypes is estimated to be approximately 34 substitutions (SE = 8).

Non-parametric bootstrap resampling (Fig. 4) yielded 100% support for monophyly of the *X. birchmanni* haplotypes. Confidence is also reasonably high (85%) for placing B4 outside of a group formed by the remaining *X. birchmanni* haplotypes, and there is weaker support (58%) for a group formed by haplotypes B2, B3 and B5. In order to further evaluate the support for monophyly of the *X. birchmanni* haplotypes, we tested the null hypothesis that those haplotypes do

not form a monophyletic group using a constraint tree in which the *X. birchmanni* haplotypes formed a monophyletic group and searching for the best tree that was incompatible with that constraint. The null hypothesis was rejected by both a one-tailed KH test ($P < 0.001$) and the SH test ($P < 0.001$), providing strong support for the monophyly of the *X. birchmanni* haplotypes. Tests of the hypothesis that haplotypes B1 and B6 were sister to the other *X. birchmanni* haplotypes were not significant (KH $P = 0.131$; SH $P = 0.131$) as were tests of the hypothesis that haplotype B2 was sister to the other *X. birchmanni* haplotypes (KH $P = 0.144$; SH $P = 0.144$).

Discussion

Genetic variation

Mitochondrial control region sequences showed low levels of genetic variability as only 0.92% of the sites were variable among *X. birchmanni* sequences from the different populations. Such levels of genetic variation are even lower than those found in *X. cortezi*, in which 1.8% of the sites of the aligned sequences of the same mtDNA fragment were variable (Gutiérrez-Rodríguez et al. 2007). In contrast to the low levels of sequence variation found in *X. cortezi* and *X. birchmanni*, higher levels of genetic variation have been reported in other teleostean fishes (Fajen & Breden 1992; Lee et al. 1995; Salzburger et al. 2003; Stefanni & Thorley 2003; Aboim et al. 2005).

The low levels of genetic variation observed in this study could be the result of recent divergence and/or other population level processes, such as isolation followed by genetic drift in small populations (e.g. Templeton et al. 1990). Studies using other mitochondrial genes as well as nuclear markers on this and other *Xiphophorus* taxa are necessary to determine which processes are responsible for the low levels of polymorphism observed in this study.

Although sample sizes were reasonable in most populations (mean = 11.1), indices of genetic variation within populations (nucleotide and gene diversity) were zero for all but four populations (Huazalingo, Garces, Sasaltitla and Chila). This finding is concordant with the low levels of genetic variation within populations found in *X. cortezi* (Gutiérrez-Rodríguez et al. 2007) and in other poeciliids (Fajen & Breden 1992; Carvalho et al. 1996; Paul Shaw et al. unpubl. data). Low genetic diversity within populations may reflect founder effects (Shaw et al. 1992; Carvalho et al. 1996) and/or population bottlenecks.

While some of the haplotypes were found in more than one population, other haplotypes were unique to single populations (e.g. B2 is only present in Huazalingo, B5 is only present in Sasaltitla and B4 is only present in Xiliatl), and all haplotypes were unique to one of three river drainages, suggesting genetic differentiation among populations within

and among those drainages. In addition, haplotypes tend to be most closely related (in terms of both mutational steps and common ancestry) to other haplotypes found in the same drainage, as can be seen in both the haplotype parsimony network (Fig. 3) and the ML tree (Fig. 4). The AMOVA without groupings revealed that most of the variation (85.23%) was due to differences among rather than within populations, and the AMOVA that grouped the populations by drainage revealed that most of the variation among populations (87.67%) was due to differences among drainages. F_{ST} values between populations located in different drainages (and between a few populations within the same drainage) were high and significant suggesting low levels of gene flow between these populations. NCA (Table 4) based on haplotype relationships revealed significant geographical associations between the two one-step *X. birchmanni* haplotype groups (1-1 and 1-2) and an inference of allopatric fragmentation between Río Calabozo (B2, B3 and B5) and Río Los Hules (B1 and B6) drainages.

High levels of genetic differentiation among populations from different drainages were also found in the swordtail *X. cortezi* (Gutiérrez-Rodríguez et al. 2007), and among populations of other freshwater fishes (Carvalho et al. 1991; Shaw et al. 1991, 1994; Fajen & Breden 1992; Alves & Coelho 1994; Coelho et al. 1997; Hänfling & Brandl 1998a,b,c; Mesquita et al. 2001). The combination of low genetic diversity within populations and the high genetic differentiation among populations could be the result of bottlenecks in the individual populations of *X. birchmanni* after they last shared a common ancestor. Bottlenecks themselves could have resulted from habitat disturbances as a consequence of human activities, resulting in inbreeding and genetic drift (Nei et al. 1975; Wishard et al. 1984).

The significant correlation found with both Mantel tests supports an isolation-by-distance scenario, in which genetic difference increases with geographical distances measured along both river courses and great circles. These results are corroborated by the results of the NCA with haplotype groups 1-1 and 1-2, which resulted in an inference of restricted gene flow with isolation by distance within the Calabozo and Los Hules drainages. Patterns of isolation by distance are relatively common in other riverine species and have been found in the Atlantic salmon *Salmo salar* (Primmer et al. 2005), in brook charr *Salvelinus fontinalis* (Poissant et al. 2005), in the guppy *Poecilia reticulata* (Crispo et al. 2006), in the swordtail *X. cortezi* (Gutiérrez-Rodríguez et al. 2007) and in the bullfrog *Rana catesbeiana* (Austin et al. 2004).

Phylogenetic relationships

According to the hypothesis that *X. birchmanni* and *X. malinche* are closest relatives (Rauchenberger et al. 1990), paraphyly of the *X. birchmanni* haplotypes relative to those of *X. malinche* would be a possibility. Such a relationship, however, is not

supported by our data. Instead, monophyly of the *X. birchmanni* haplotypes (Fig. 4) is strongly supported (bootstrap = 100%; KH and SH tests highly significant), and *X. malinche* appears to be more closely related to *X. cortezi* than to *X. birchmanni* (Gutiérrez-Rodríguez et al. 2007).

Previous studies have suggested that *X. birchmanni* is nested within the *cortezi* clade as the sister group of *X. malinche* (Rauchenberger et al. 1990; Borowsky et al. 1995). Although our limited sample of other *Xiphophorus* species (only *X. cortezi* and *X. malinche*) does not allow us to address this hypothesized relationship explicitly (i.e. by outgroup rooting of the three species tree), our results are consistent with the findings of Meyer et al. (1994) and Gutiérrez-Rodríguez et al. (2007) in suggesting that *X. birchmanni* is not particularly closely related to *X. malinche*, or more specifically, that *X. malinche* is more closely related to *X. cortezi* than to *X. birchmanni*. As seen in the parsimony network (Fig. 3) and ML phylogram (Fig. 4), the *X. birchmanni* haplotypes are all very closely related to one another (one to four mutations) and are highly divergent (34 ± 8 mutations) from those of *X. malinche* and *X. cortezi*, which are themselves very closely related (one to seven mutations).

Rauchenberger et al. (1990) considered *X. birchmanni* to be the sister species of *X. malinche* based on three shared morphological characters (nuchal hump, mid-dorsal spots and broad vertical bars on females, see Fig. 1), and a shared allele at the PEP-S locus. According to the more extensive electrophoretic survey of Morris et al. (2001), two PEP-S alleles are found in *X. birchmanni* (d and e). The more common allele (d) is not found in *X. malinche* though it is found in some populations of *X. cortezi* as well as several other swordtail species, while the less common allele (e) is the more common allele in *X. malinche* and *X. cortezi*. Although these patterns of shared alleles are ambiguous concerning the relationships among the three species, they are more suggestive of a close relationship between *X. cortezi* and *X. malinche* than between *X. birchmanni* and *X. malinche*.

Of the three putative morphological synapomorphies for *X. birchmanni* and *X. malinche*, the nuchal hump is characteristic of mature *X. birchmanni* males, but has rarely been observed in *X. malinche*, and has also been observed infrequently in at least one other swordtail species (*X. multilineatus*, Fig. 1C). The broad oval vertical bars that were supposed to be distinctive of *X. birchmanni* and *X. malinche* females have been detected in *X. cortezi* females (Molly R. Morris, unpubl. data). Therefore, the only character that remains as a possible synapomorphy of *X. malinche* and *X. birchmanni* is the presence of mid-dorsal spots, though *X. birchmanni* characteristically has one to two rows and *X. malinche* three to four rows of these spots. According to the phylogeny of Rauchenberger et al. (1990), *X. birchmanni* and *X. malinche* are the only sister species within the Northern Swordtail clade that are found in

sympatry. However, if *X. malinche* is the sister species of *X. cortezi* rather than of *X. birchmanni* (see above), then there are no exceptions to the allopatric distribution of sister species in this group. This finding suggests that speciation in the Northern Swordtails is primarily allopatric, and that a hybrid zone between *X. birchmanni* and *X. malinche* (Rosenthal et al. 2003) is the result of secondary contact.

Hybridization between *X. birchmanni* and *X. malinche* (Rosenthal et al. 2003) raises the possibility that the characters cited by Rauchenberger et al. (1990) as evidence that these species are sister groups resulted from examining hybrid specimens. While most of the *X. malinche* specimens they examined were from the Río Claro, where *X. birchmanni* does not occur, they also examined a few specimens from the Río Calnali, where hybridization with *X. birchmanni* is known to occur, and from the Río Contzintla, which is upstream from one of the localities (Atlaxco) where *X. birchmanni* was collected for the present study. The inclusion of hybrids may also have influenced the results of Borowsky et al. (1995), who reported relationships for this group similar to those in Rauchenberger et al. (1990) using RAPD DNA markers, as they included both *X. malinche* and *X. birchmanni* from the hybrid site (i.e. Río Calnali).

Outgroup rooting of the haplotype tree (Fig. 4) suggests that haplotype B4 is ancestral among the *X. birchmanni* haplotypes. Although the analysis performed using the *tcs* software identified haplotype B2 as the ancestral haplotype, this inference ignores outgroup information and is based instead on haplotype frequency and interior vs. tip location in the haplotype tree. High frequency is a questionable basis for inferring ancestral status because ancestral haplotypes should commonly experience decreases in frequency (and even extinction) as the result of both natural selection and stochastic processes. Moreover, the *tcs* analysis treated haplotype B4 as a tip haplotype (thus decreasing its outgroup weight), but the tip vs. interior status of that haplotype is ambiguous. The outgroup haplotypes are sufficiently divergent from the *X. birchmanni* haplotypes that the numbers of mutational steps separating the two groups cannot be estimated accurately by parsimony, which results in the two sets of haplotypes being unconnected in the *tcs* analysis. The ML phylogenetic analysis, which connects the two sets of haplotypes, suggests that B4 is an interior haplotype. Although B4 is identified as the ancestral *X. birchmanni* haplotype in that analysis, the number of mutational steps separating all of the *X. birchmanni* haplotypes is small, and therefore the alternative hypothesis that haplotype B2 is ancestral (or the sister group of the remaining *X. birchmanni* haplotypes) cannot be rejected ($P = 0.144$ for both KH and SH tests). Nonetheless, the best-supported inference is that B4 is ancestral.

Inferences concerning the ancestral haplotype bear on the biogeographical origin and subsequent history of the species.

Rauchenberger *et al.* (1990) hypothesized that the ancestral form of *X. birchmanni* entered the Río Tempoal drainage from the upstream ranges (Río Claro) of the Río Moctezuma drainage via stream capture. Once isolated from *X. malinche*, it then expanded south-east to the Río Tuxpan and north-west along the foot of the Sierra Madre until it came into contact with *X. cortezi*. If haplotype B1 (or B6) were ancestral, this finding would lend support to the Rauchenberger *et al.* (1990) hypothesis, as this haplotype is found in the part of the Tempoal system (Río Los Hules drainage) closest to the Río Claro. However, our results tentatively suggest that haplotype B4, found in the north-western part of the Tempoal system (Río San Pedro drainage), is the ancestral haplotype. Therefore, we hypothesize that *X. birchmanni* arose in the north-west (Río San Pedro drainage) and spread south-east through the Tempoal system (Río Los Hules and Río Calabozo drainages). Unfortunately, we did not detect *X. birchmanni* in our sampling of the Río Vinazco (Tuxpan) drainage. Our hypothesis concerning the phylogeographical history of *X. birchmanni* predicts that the haplotypes from this drainage will be B2, B3 and B5 or derivatives of these haplotypes.

Acknowledgements

We are grateful to the Mexican government for permission to collect fishes, and we would like to thank Oscar Rios-Cardenas, Daniel Wiegmann and M. Scarlett Tudor for their assistance in the field. This research was supported by funding from the National Science Foundation IBN-0316687 to M. R. M. and the Ohio University Provost's Undergraduate Research Fund to A. E. S. Collecting methods comply with the Animal Care Guidelines of Ohio University (Animal Care and Use approval no. L01-01).

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 from mtDNA sequence analysis. *Molecular Ecology*, *14*, 1343–1354.
- Alves, J. & Coelho, M. M. (1994). Genetic variation and population subdivision of the endangered Iberian cyprinid *Cbondrostoma lusitanicum*. *Journal of Fish Biology*, *44*, 627–636.
- Austin, J. D., Loughheed, S. C. & Boag, P. T. (2004). Controlling for the effects of history and nonequilibrium conditions in northern bullfrog (*Rana catesbeiana*) populations. *Genetics*, *168*, 1491–1506.
- Bohonak, A. J. (2002). IBD (Isolation by Distance): a program for population genetic analyses of isolation by distance. *Journal of Heredity*, *93*, 153–154.
- Borowsky, R. L., McClelland, M., Cheng, R. & Welsh, J. (1995). Arbitrarily primed DNA fingerprinting for phylogenetic reconstruction in vertebrates: the *Xiphophorus* model. *Molecular Biology and Evolution*, *12*, 1022–1032.
- Burnham, K. P. & Anderson, D. R. (2004). Multimodel inference. Understanding AIC and BIC in model selection. *Sociology Methods Research*, *33*, 261–304.
- 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. *Biological Journal of the Linnean Society*, *42*, 389–405.
- Carvalho, G. R., Shaw, P. W. & Hauser, L. (1996). Artificial introductions, evolutionary change and population differentiation in Trinidadian guppies (*Poecilia reticulata*: Poeciliidae). *Biological Journal of the Linnean Society*, *57*, 219–234.
- Castelloe, J. & Templeton, A. R. (1994). Root probabilities for intraspecific gene trees under neutral coalescent theory. *Molecular Phylogenetics Evolution*, *3*, 102–113.
- Clement, M., Posada, D. & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, *9*, 1657–1659.
- Coelho, M. M., Alves, J. & Rodrigues, E. (1997). Patterns of genetic divergence in *Cbondrostoma lusitanicum* Collares-Pereira, in intermittent Portuguese rivers. *Fisheries Management and Ecology*, *4*, 223–232.
- Crandall, K., Templeton, A. R. & Sing, C. F. (1994). Intraspecific cladogram estimation: problems and solutions. In R. W. Scotland & D. J. Siebert & D. M. Williams (Eds) *Models in Phylogeny Reconstruction* (pp. 273–297). Oxford: Clarendon Press.
- Crispo, E., Bentzen, P., Reznick, D. N., Kinnison, M. T. & Hendry, A. P. (2006). The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology*, *15*, 49–62.
- 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.
- Faber, J. E. & Stepien, C. A. (1997). The utility of mitochondrial DNA control region sequences for analyzing phylogenetic relationships among populations, species, and genera of the Percidae. In T. D. Kocher & C. A. Stepien (Eds) *Molecular Systematics of Fishes* (pp. 129–144). San Diego: Academic Press.
- Fajen, A. & Breden, F. (1992). Mitochondrial DNA sequence variation among natural populations of the Trinidad guppy, *Poecilia reticulata*. *Evolution*, *46*, 1457–1465.
- Gutiérrez-Rodríguez, C., Morris, M. R., Dubois, N. S. & De Queiroz, K. (2007). Genetic variation and phylogeography of the swordtail fish *Xiphophorus cortezi* (Cyprinodontiformes, Poeciliidae). *Molecular Phylogenetics and Evolution*, *43*, 111–123.
- Hänfling, B. & Brandl, R. (1998a). Genetic and morphological variation in a common European cyprinid, *Leuciscus cephalus* within and across Central European drainages. *Journal of Fish Biology*, *52*, 706–715.
- Hänfling, B. & Brandl, R. (1998b). Genetic variability, population size and isolation of distinct populations in the freshwater fish *Cottus gobio* L. *Molecular Ecology*, *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. *Journal of Molecular Evolution*, *29*, 170–179.
- Lee, W.-J., Conroy, J., Howell, W. H. & Kocher, T. D. (1995). Structure and evolution of teleost mitochondrial control regions. *Journal of Molecular Evolution*, *41*, 54–66.
- Meffe, G. K. & Snelson, F. F. (1989). An ecological overview of poeciliid fishes. In G. K. Meffe & F. F. Snelson (Eds) *Ecology and Evolution of Livebearing Fishes (Poeciliidae)* (pp. 13–31). Englewood Cliffs: Prentice Hall.

- 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. (1997). The evolution of sexually selected traits in male swordtail fishes (*Xiphophorus*: Poeciliidae). *Heredity*, 79, 329–337.
- Meyer, A., Kocher, T. D., Basasibwaki, P. & Wilson, A. C. (1990). Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature*, 347, 550–553.
- 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.
- Morris, M. R., De Queiroz & Morizot, D. C. (2001). Phylogenetic relationships among populations of northern swordtails (*Xiphophorus*) as inferred from allozyme data. *Copeia*, 1, 65–81.
- Nei, M. & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceeding of the National Academy of Sciences of the United States of America*, 76, 5269–5273.
- Nei, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, 29, 1–10.
- Poissant, J., Knowlton, T. W. & Ferguson, M. M. (2005). Non-equilibrium conditions following landscape rearrangement: the relative contribution of past and current hydrological landscape on the genetic structure of a stream-dwelling fish. *Molecular Ecology*, 14, 1321–1331.
- Posada, D. & Crandall, K. A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics*, 14, 817–818.
- Posada, D., Crandall, K. A. & Templeton, A. R. (2000). GEODIS: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, 9, 487–488.
- Primmer, C. R., Veselov, A. J., Zubchenko, V. A., Poututkin, A. & Bakhmet, I. (2005). Isolation by distance within a river system: genetic population structuring of Atlantic salmon, *Salmo salar*, in tributaries of the Varzuga River in northwest Russia. *Molecular Ecology*, 15, 653–666.
- 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. *American Museum Novitates*, 2975, 1–41.
- Rosen, D. E. (1979). Fishes from the uplands and intermontane of Guatemala: revisionary studies and comparative geography. *Bulletin of the American Museum of Natural History*, 162, 268–375.
- Rosenthal, G. G., de la Rosa Reyna, X. F., Kazianis, S., Stephens, M. J., Morizot, D. C., Ryan, M. J. & García de León, F. J. (2003). Dissolution of sexual signal complexes in a hybrid zone between the swordtails *Xiphophorus birchmanni* and *Xiphophorus malinche* (Poeciliidae). *Copeia*, 2003, 299–307.
- Salzburger, W., Brandstätter, A., Gilles, A., Parson, W., Hempel, M., Sturmhuber, C. & Meyer, A. (2003). Phylogeography of the vairone (*Leuciscus souffia*, Risso 1826) in Central Europe. *Molecular Ecology*, 12, 2371–2386.
- Schneider, S., Roessli, D. & Excoffier, L. (2000). *ARLEQUIN, Version 2.00: a Software for Population Genetics Data Analysis*. Geneva: University of Geneva.
- Seutin, G., White, B. N. & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, 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. *Journal of Fish Biology*, 39 (Suppl. A), 203–209.
- Shaw, P. W., Carvalho, G. R., Magurran, A. E. & Seghers, B. H. (1994). Factors affecting the distribution of genetic variability in the guppy, *Poecilia reticulata*. *Journal of Fish Biology*, 45, 875–888.
- 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. *Proceeding of the Royal Society of London*, 248, 111–116.
- Shimodaira, H. & Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*, 16, 1114–1116.
- Stefanni, S. & Thorley, J. (2003). Mitochondrial DNA phylogeography reveals the existence of an evolutionary significant unit of the sand goby *Pomatoschistus minutus* in the Adriatic (Eastern Mediterranean). *Molecular Phylogenetics and Evolution*, 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. *Molecular Biology and Evolution*, 22, 1386–1392.
- Swofford, D. L. (2002). PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4.0b10. Sunderland, Massachusetts: Sinauer Associates.
- Swofford, D. L., Olsen, G. J., Waddell, P. J. & Hillis, D. M. (1996). Phylogenetic inference. In D. M. Hillis & C. Moritz & B. K. Mable (Eds) *Molecular Systematics* (pp. 407–514). Sunderland, Massachusetts: Sinauer Associates.
- Tamura, K. & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10, 512–526.
- Templeton, A. R., Boerwinkle, E. & Sing, C. F. (1987). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*, 117, 343–351.
- 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.
- Templeton, A. R., Routman, E. & Phillips, C. A. (1995). Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, 140, 767–782.
- Templeton, A. R., Shaw, K., Routman, E. & Davis, S. K. (1990). The genetic consequences of habitat fragmentation. *Annals of the Missouri Botanical Garden*, 77, 13–27.
- 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 Research*, 24, 4876–4882.
- Topografix (2003). *ExpertGPS*. Stow, Massachusetts: TopoGrafix.
- 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.