Copeia, 2001(1), pp. 65–81

Phylogenetic Relationships among Populations of Northern Swordtails *(Xiphophorus)* as Inferred from Allozyme Data

MOLLY R. MORRIS, KEVIN DE QUEIROZ, AND DON C. MORIZOT

Twenty-nine populations of *Xiphophorus* fishes representing nine species of northern swordtails, one southern swordtail and a platyfish were assayed electrophoretically for allozyme variation. Phylogenetic relationships were inferred using parsimony and likelihood analysis of gene frequency characters, as well as Fitch-Margoliash, minimum evolution and neighbor-joining analyses of genetic distances. The phylogenetic relationships among species that were well supported in all analyses included (1) monophyly of the northern swordtails, (2) the *pygmaeus* clade of *Xiphophorus nigrensis*, *X. multilineatus*, and *X. pygmaeus*, and (3) the clade of *X. nigrensis* and *X. multilineatus*. Of those species represented by more than one population, all analyses supported monophyly of *X. montezumae* and weakly supported monophyly of *X. nezahualcoyotl* and *X. birchmanni*. Only the distance analyses supported monophyly of *X. cortezi*, and the support was weak. Finally, all analyses supported a clade including *X. nezahualcoyotl* from the Río Tamesí drainage and some populations from the Río Pánuco drainage, that is, nonmonophyly of the set of populations from the Río Pánuco drainage. Previously published trees based on morphology, behavior and randomly amplified DNAs were generally congruent with the optimal trees for the allozyme data and were not rejected by those data; in contrast, trees based on DNA sequences were more incongruent with the optimal trees for the allozyme data and were rejected by those data.

XIPHOPHORUS (Cyprinodontiformes: Poeciliidae) is a clade of small, attractive, freshwater fishes that has been studied intensively in terms of systematics, biogeography, genetics, oncology, and behavior (reviewed by Meffe and Snelson, 1989). Studies on the evolution of sexually selected traits within this group (e.g., Basolo, 1990; 1995; Morris and Ryan, 1996) based on earlier published phylogenies (e.g., Rosen, 1979; Rauchenberger et al., 1990) have prompted a surge of more recent phylogenetic studies (Meyer et al., 1994; Borowsky et al., 1995; Meyer, 1997). This recent interest in the phylogeny of *Xiphophorus* has led to an increase in systematic data, but it has also led to conflicting phylogenetic hypotheses.

We examined relationships among and within species from one of the major clades within *Xiphophorus* using allozyme data. This clade, informally known as the northern swordtails, is composed of nine currently recognized species. All nine species occur in the Río Pánuco drainage in eastern México along the eastern slope of the Sierra Madre Oriental, although a few species have populations in adjacent drainages to the north and south (Rauchenberger et al., 1990). Although a previous study of the relationships among the northern swordtails (Rauchenberger et al., 1990) included some of the same allozyme data analyzed in the present study, we here present a more extensive dataset by including both additional taxa and additional loci and by incorporating information on the frequencies of alleles in natural populations. We analyze these data using several phylogenetic methods and examine congruence among the results, and we evaluate previously published trees for this group using the allozyme data.

MATERIALS AND METHODS

Sampling.—Twenty-seven natural populations representing all currently recognized species of northern swordtails were sampled (Fig. 1). For four of the five species inhabiting more than one river, *X. cortezi, X. montezumae, X. nezahualcoyotl, X. birchmanni* but not *X. malinche* (Rauchenberger et al., 1990), samples from several localities were included. Some genetically similar populations from nearby localities were combined, namely, two samples of *X. montezumae* from Río Ojo Frío (north of Damian Carmona and El Quince, approximately 5 km apart), and three samples of *X. montezumae* from the Río Santa Maria drainage (two from Arroyo Cienega Grande 13 km apart and one from Arroyo La Cienega). None of the combined samples was significantly different from one another at any of the loci based on chi-square tests of their genotype arrays. A representative population of the southern swordtails, *X. helleri* was included in the analysis so that we could test the mono-

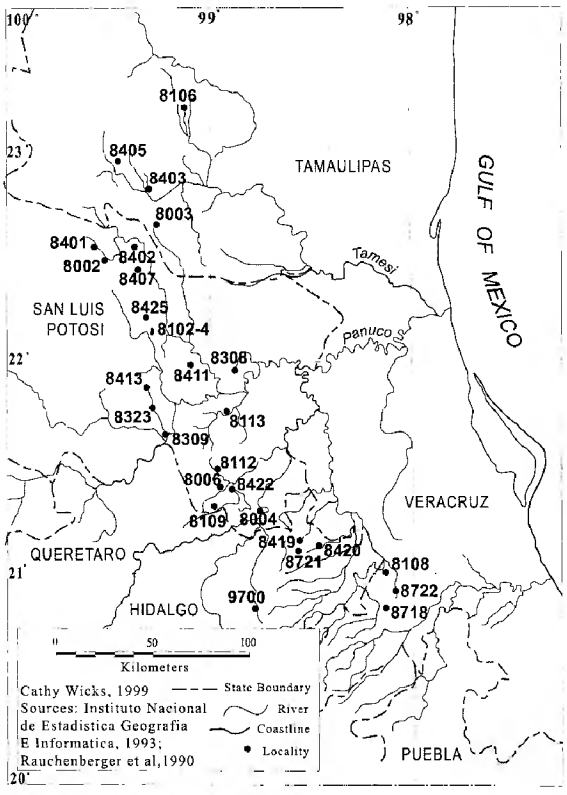

Fig. 1. Map of the Río Pánuco drainage system showing most major drainages. Localities are marked with D. C. Morizot's collection numbers (see Appendix 1).

phyly of the northern swordtails. Data for the *CK-A** locus were missing from the wild caught sample of *X. helleri* and were supplemented with data from two laboratory lines (Stock number 3062, Belize River drainage Guatemala, and 2997, Rio Coatzacoalcos, Mexico; Morizot and Siciliano, 1982). A representative population of the platyfishes, *X. variatus* from from Río Atoyac, Veracruz Mexico, was included to root the trees. Collection numbers, sample sizes, and localities for all populations used in this study are presented in Appendix 1.

Specimen preparation and electrophoresis.—Sample preparation procedures follow Morizot et al. (1977) and Morizot and Siciliano (1979). Tissues were dissected from freshly killed specimens or from specimens frozen in 50 ml centrifuge tubes in liquid N_2 in the field, maintained at −76 C and subjected to vertical starch gel electrophoresis using the methods of Siciliano and Shaw (1976), Morizot et al. (1977) and Morizot and Schmidt (1990). Twenty-eight loci were scored. The locus symbols, numbers (recommended by International Union of Biological Nomenclature, 1984), tissue and electrophoretic conditions used for each protein are listed in Table 1.

Phylogenetic analysis.—Numerous alternative methods exist for inferring phylogenies, each with its own strengths and weaknesses (for review see Felsenstein, 1988; Swofford and Olsen, 1990; Swofford et al., 1996). We analyzed our data using several methods to assess the robustness of the results across methods.

Character analyses.—Recent studies have shown that incorporating frequency information in character based analyses improves the accuracy of phylogenetic inference (Wiens, 1995, 1999; Wiens and Servedio, 1997). Therefore, we used methods that incorporated frequency information by treating loci as characters and allele frequencies as character states. Populations were used as OTUs (terminal taxa), and the data were analyzed using both parsimony and likelihood.

For the parsimony analyses, phylogenetic trees were estimated using the method designated MANOB by Swofford and Berlocher (1987; see also Berlocher and Swofford, 1997) as implemented in PAUP* (vers. 4.0.0 b2-4, D. L. Swofford, Natural History Museum, Smithsonian Institution, Washington, DC., unpubl). In this method, each unique combination of allele frequencies is treated as a different character state, and the cost of transformation between states is the Manhattan distance between them (implemented using a step matrix). A heuristic search was performed, with starting trees obtained by random stepwise addition (25 replicates), tree bisection reconstruction (TBR) branch-swapping, and saving all optimal trees (MULTREES option). A bootstrap analysis (Felsenstein, 1985) was performed with 1000 replicates, using heuristic searches with simple stepwise addition, TBR branch-swapping, and the MULTREES option.

The MANOB method, which does not allow hypothetical ancestors to have states different than those observed in the terminal taxa, is an approximation of a method called MANAD by Swofford and Berlocher (1987) in which hypothetical ancestors are not so constrained. In three allozyme datasets analyzed by Berlocher and Swofford (1997), the optimal tree obtained under MANOB was identical to at least one of those obtained under MANAD. Preliminary analysis of our data also found a very strong correlation between tree scores for the two methods. Therefore, we used MANOB for the rest of our parsimony analyses. Because some of the programs we used do not allow missing data, or

TABLE 1. PRESUMPTIVE GENETIC LOCI, SYMBOLS, E.C. NUMBERS, TISSUES, AND ELECTROPHORETIC CONDITIONS USED TO STUDY PROTEIN VARIATION IN NORTHERN SWORDTAIL FISHES.

Locus	Symbol[a]	E.C. number[b]	Tissue[c]	Buffer[d]
Aconitate hydratase, cytosolic	*sAH** (*ACO1*)	4.2.1.3	L	TC
Creatine kinase-A (*MP4* of M&S, 1982)	*CK-A**	2.7.3.2	M	TEB (TC)
Glyceraldehyde-3-phosphate dehydrogenase-*1*	*GAPDH-1**	1.2.1.12	B&E, M	TEB
Glyceraldehyde-3-phosphate dehydrogenase-*2*	*GAPDH-2**	1.2.1.12	M (L)	TEB or TC
Glycerate-dehydrogenase	*GLYDH**	1.1.1.29	L	TC
Glucose-6-phosphate isomerase-1	*GPI-1**	5.3.1.9	M (B&E)	TEB
Glucose-6-phosphate isomerase-2	*GPI-2**	5.3.1.9	B&E, M, L	TEB
Guanylate kinase-1	*GUK1**	2.7.4.8	B&E	TEB or TC
Isocitrate dehydrogenase-2, cytosolic	*sIDHP-2** (*IDH1*)	1.1.1.42	L	TEB
Lactate dehydrogenase-C (*LDH1* of M&S, 1982)	*LDH-C**	1.1.1.27	B&E	TC
Lactate dehydrogenase-A (*LDH2* of M&S, 1982)	*LDH-A**	1.1.1.27	M (B&E)	TC
Lactate dehydrogenase-B (*LDH3* of M&S, 1982)	*LDH-B**	1.1.1.27	L (B&E, M)	TC
Malate dehydrogenase-2, cytosolic	*sMDH-2** (*MDH1*)	1.1.1.37	M (B&E)	TC
Malate dehydrogenase-1, cytosolic	*sMDH-1** (*MDH1*)	1.1.1.37	M	TC
Mannose-6-phosphate isomerase	*MPI**	5.3.1.8	M (B&E, L)	TEB (T)
Muscle protein-5 (*MP5* of M&S, 1982)	*PROT5**		M	TEB
Parvalbumin-1 (*MP1* of M&S, 1982)	*PVALB-1**		M	TEB
Parvalbumin-2 (*MP2* of M&S, 1982)	*PVALB-2**		M	TEB
Parvalbumin-3 (*MP3* of M&S, 1982)	*PVALB-3**		M	TEB
Peptidase S (*PEP2* of M&S, 1982)	*PEPS**	3.4.11 or 3.4.1	M (B&E, L)	TEB
Peptidase B (*PEP3* of M&S, 1982)	*PEPB**	3.4.11 or 3.4.12	M (B&E, L)	TEB
6-phosphogluconate dehydrogenase	*PGDH**	1.1.1.44	B&E, M, L	TC
Phosphoglucomutase	*PGM**	5.4.2.2	M (B&E, L)	TEB
Superoxide dimutase, cytosolic	*sSOD** (*SOD1*)	1.15.1.1	M, B&E, L	TEB
Triosephosphate isomerase-1	*TPI-1**	5.3.1.1	B&E (M)	TEB
Triosephosphate isomerase-2	*TPI-2**	5.3.1.1	B&E (M)	TEB
Uridine monophosphate hydrolase (ACP of M&S, 1982)	*UMPH**	3.1.3.2	L	TC
UDP glucose-hexose uridylytransferase	*UGHUT** (*GALT*)	2.7.7.12	L, B&E	TC (TEB)

M&S, 1982 = Morizot and Silicano, 1982.

[a] Gene symbols follow standardized fish gene nomenclature (Schaklee et al., 1990). Symbols of human homologues when significantly different from fish gene nomenclature are given in parentheses (McAlpine et al., 1988).

[b] Numbers recommended by the International Union of Biological Nomenclature (1984).

[c] First entry is tissue of choice; if multiple entries separated by commas, all tissues listed provide adequate activity and resolution. Entries in parentheses are usable but less desirable tissues. Abbreviations: B&E = brain and eye; L = liver; M = skeletal muscle.

[d] Buffer compositions are described in Siciliano and Shaw (1976). Abbreviations: TEB = Tris-EUTA-borate, pH 8.0; TC = Tris-citrate, pH 7.0.

allow only binary trees, we reduced the dataset by eliminating the four taxa with missing data (*birchmanni* 8419a, *birchmanni* 8718, *birchmanni* 8721, *birchmanni* 8722). We also eliminated the most deeply nested populations from monophyletic groups of conspecific populations according to the strict consensus tree from the MANOB analysis of the full dataset (*cortezi* 8420, southern *nezahualcoyotl* 8407, southern *nezahualcoyotl* 8402, northern *nezahualcoyotl* 8403, northern *nezahualcoyotl* 8405). This resulted in a reduced dataset of 20 taxa. To find the optimal trees for this reduced dataset, we ran a MANOB analysis in PAUP*, using a branch-and-bound search and saved all optimal trees (MULTREES option). A bootstrap analysis (Felsenstein, 1985) was performed with 1000 replicates, using heuristic searches with simple stepwise addition, TBR branch-swapping, and the MULTREES option.

The character data were also analyzed under continuous character maximum likelihood (Cavalli-Sforza and Edwards, 1967; Felsenstein, 1981, J. Felsenstein, Univ. of Washington, Seattle, 1993, unpubl.), using the Contml program in PHYLIP (vers. 3.572c, J. Felsenstein, Univ. of Washington, Seattle, 1993, unpubl.). Because this program does not allow missing data, the analysis was carried out on the reduced dataset of 20 taxa, using global rearrangements and 10 random input orders of taxa. Bootstrap analysis was not done because subsampling resulted in taxa with identical transformed gene frequencies, which are not permitted by the Contml program.

Distance analyses.—The reduced dataset of 20 taxa was used to generate Nei (1972) and Cavalli-Sforza and Edwards (1967) distances using PHYLIP. These distances were analyzed using the method of Fitch and Margoliash (1967) as implemented in PHYLIP using global rearrangements and 10 random input orders of taxa. They were also analyzed using the neighbor-joining method (Saitou and Nei, 1987) in PHYLIP and the minimum evolution method (Rzhetsky and Nei, 1992) in PAUP* using a heuristic search with starting trees obtained by random stepwise addition (25 replicates), TBR branch-swapping and the MULTREES option. Bootstrap analyses were performed with 100 replicates for the Fitch-Margoliash and neighbor-joining analyses. Bootstraps were not done for minimum evolution because PAUP* does not calculate Nei and Cavalli-Sforza distances, which were imported from PHYLIP.

Comparison with previous studies.—Previously published hypotheses of phylogenetic relationships within *Xiphophorus* were evaluated using the allozyme data. The trees included Rosen's (1979) morphology-based tree, the morphology and allozyme-based tree of Rauchenberger et al. (1990), Haas's (1992) behavior and morphology-based tree, two DNA sequence trees from Meyer et al. (1994), six trees based on RAPDs from Borowsky et al. (1995:figs. 4–9), and six trees from Marcus and McCune (1999:figs. 3–6) based on various combinations of the data from the previous papers. We also tested monophyly of the northern swordtails relative to the southern swordtails (represented in our analyses by *X. helleri*). Because previous studies did not always include all species sampled in our study, and because most previous studies used species rather than populations as OTUs, previous hypotheses were evaluated using constraint trees. The relationships among our populations were constrained based on the published phylogeny, retaining the between species relationships but allowing relationships among conspecific populations to vary, allowing species to be paraphyletic, and allowing unsampled species to be placed anywhere in the tree. Unrooted trees from Borowsky et al. (1995) were rooted either between the platies (*X. maculatus, X. variatus, X. xiphidium*) and the swordtails or between the northern and southern swordtails if no platies were included (i.e., in a position similar to that found in the current study). The trees were constructed in MacClade (vers. 3.06, W. P. Maddison and D. R. Maddison, Sinauer Assoc., Inc., Sunderland, MA, 1992, unpubl.) and then loaded as backbone constraints in PAUP*, which we used to search for the shortest trees that were consistent with each of the constraint trees under MANOB. Branch-and-bound searches were performed with the MULTREES option. Analyses were performed on the reduced dataset of 20 taxa because missing data in the full dataset resulted in large numbers of equivalent trees.

We determined the symmetric differences (Penny and Hendy, 1985) between the most-parsimonious unconstrained trees and the most-parsimonious trees under each of the constraints using PAUP*. These values reflect how many groups different trees have in common. We also compared the optimal trees in the absence of constraints with the optimal trees under each constraint by calculating the minimum difference in number of steps for each character ($n = 28$) between sets of constrained and unconstrained trees using a Wilcoxon signed ranks test to determine whether the trees corresponding with previously proposed hypothe-

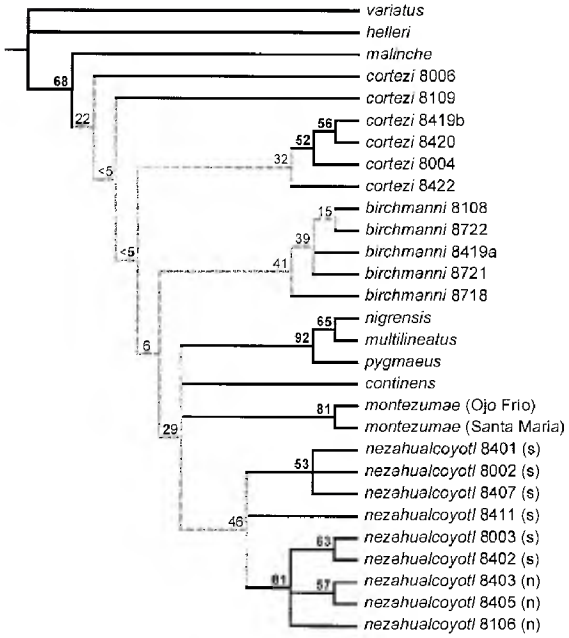

Fig. 2. Strict consensus tree of 60 shortest trees found using MANOB on the full dataset with bootstrap values for the various nodes. Length of optimal trees equals 61.96. Gray branches are supported by bootstrap values of < 50%.

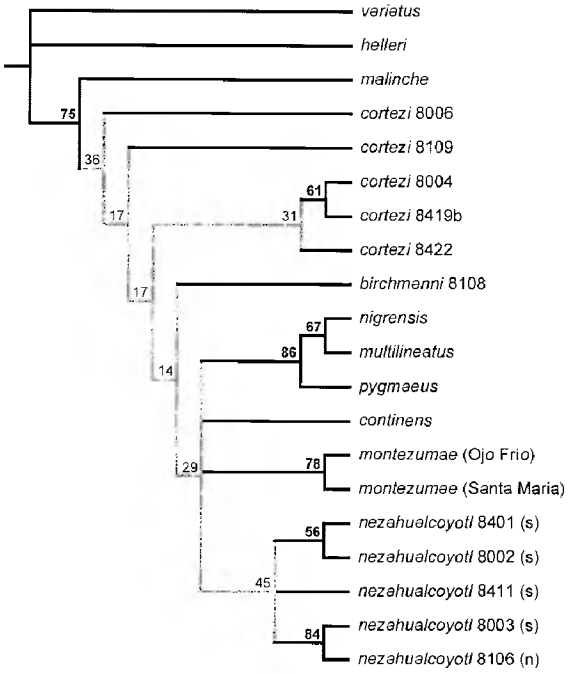

Fig. 3. Strict consensus tree of 10 shortest trees found using MANOB on the reduced dataset with bootstrap values for the various nodes. Length of optimal trees equals 53.625.

ses required more steps for significantly more allozyme characters than the optimal MANOB trees (Templeton, 1983). Similar tests were performed under the likelihood criterion (Kishino and Hasegawa, 1989) by comparing the first binary tree from each set of constrained and unconstrained trees obtained in the MANOB analysis against the maximum-likelihood tree using the User Tree option in the Contml program of PHYLIP.

Results

The alleles at the 28 loci and their frequencies in the 29 populations are summarized in Appendix 2. Three of the characters were monoallelic within the northern swordtails, whereas 25 were variable, 20 of which were informative under parsimony (19 for the 20 taxon dataset). The number of alleles for variable loci ranged from 2 to 6 (2 to 5 for the 20 taxon dataset).

Character analyses.—Sixty shortest trees were found using MANOB for the full dataset, all with a length of 61.96. The strict consensus of these 60 trees with bootstrap values for the various nodes is shown in Figure 2. The differences between the alternative topologies involved relationships among populations of *X. nezahualcoyotyl,* populations of *X. birchmanni,* and among *X. continens, X. montezumae,* and a clade of *X. pygmaeus, X. nigrensis,* and *X. multilineatus* (the *pygmaeus* clade of Rauchenberger et al., 1990). Ten nodes were supported in more than 50% of the bootstrap replicates: monophyly of the northern swordtails (68%), *X. nigrensis* and *X. multilineatus* as sister species (65%), the *pygmaeus* clade (92%), a clade of the two populations of *X. montezumae* (81%), a clade of two populations of *X. cortezi* (8419b and 8420; 56%), the previous clade plus *X. cortezi* 8004 (52%), and four clades within *X. nezahualcoyotl* (southern *nezahualcoyotl* 8401, 8002 and 8407, 53%; southern *nezahualcoyotl* 8003 and 8402, 63%; northern *nezahualcoyotl* 8403 and 8405, 57%; and a clade of the previous two clades plus northern *nezahualcoyotl* 8106, 81%).

The MANOB analysis of the reduced dataset yielded 10 shortest trees, all with a length of 53.625. The strict consensus of these 10 trees with bootstrap values is shown in Figure 3. Seven nodes were supported in more than 50% of the bootstrap replicates, all of which were similarly supported in the full dataset, allowing for deleted taxa: monophyly of the northern swordtails (75%); *X. nigrensis* and *X. multilineatus* as sister species (67%); the *pygmaeus* clade (86%); a clade of the two populations of *X. montezumae* (78%); a clade of two populations of *X. cortezi* (8419b and 8004, 61%); and a clade of two pop-

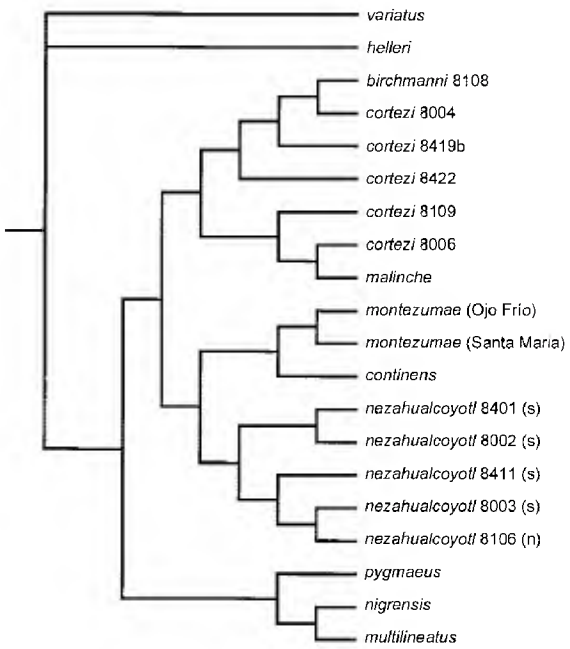

Fig. 4. The single best tree from continuous character maximum likelihood analysis on the reduced dataset. Ln-likelihood equals 1167.69.

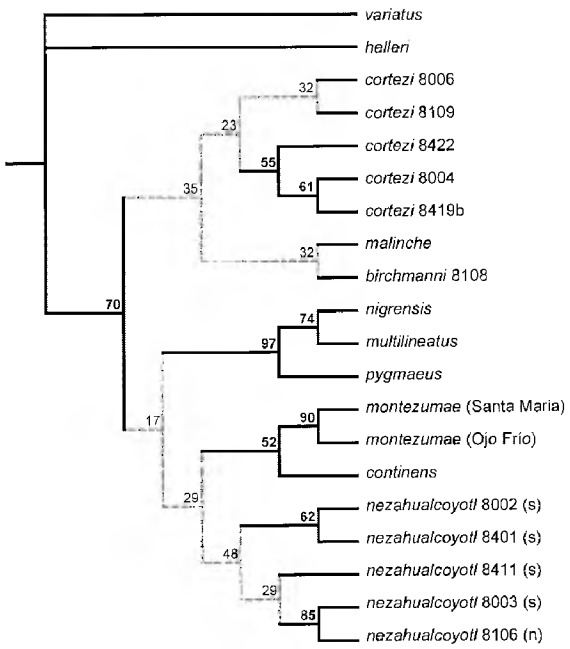

Fig. 5. The single best tree from Fitch-Margoliash analysis of the reduced dataset converted to Nei genetic distances with bootstrap values for the various nodes. Sum of squares equals 3.74291.

ulations of southern *X. nezahualcoyotl,* (8401 and 8002, 56%), and a clade of two populations of *X. nezahualcoyotl* (southern *nezahualcoyotl* 8003 and northern *nezahualcoyotl* 8106, 84%).

Continuous character maximum likelihood analysis on the reduced dataset yielded a single best tree (Fig. 4) with a ln-likelihood score of 1167.69. This tree differs from the parsimony trees in placing the root for the northern swordtails between the *pygmaeus* clade and the remaining northern swordtails rather than within the *cortezi* clade (a putative clade composed of *X. cortezi, X. birchmanni,* and *X. malinche* according to Rauchenberger et al., 1990), in the position of *X. birchmanni* within the *cortezi* clade, and in resolving the positions of southern *X. nezahualcoyotl* 8411 and *X. continens.*

Distance analyses.—Fitch-Margoliash analysis of the data converted to Nei genetic distances produced a single best tree (sum of squares = 3.74291), which is shown with bootstrap values for the various nodes in Figure 5. Nine nodes are supported in more than 50% of the bootstrap replicates and include the seven nodes similarly supported under parsimony analysis of the reduced dataset. The nodes supported in more than 50% of the bootstrap replicates in the Fitch-Margoliash analysis that were not comparably supported by the parsimony analysis include a clade of three populations of *X. cortezi* (8004, 8419b and 8422, 55%), and a clade of *X. montezumae* and *X. continens* (52%). Fitch-Margoliash analysis of the data converted to Cavalli-Sforza distances yielded an optimal tree that differed from the tree based on the Nei distances only in the placement of the three larger clades: with the Cavalli-Sforza distances the *cortezi* clade and Rauchenberger et al.'s (1990) *montezumae* clade (which includes *X. montezumae, X. continens,* and *X. nezahualcoyotl*) are sister groups; with the Nei distances the *pygmaeus* clade and *montezumae* clade are sister groups. Neighbor-joining analyses of the data converted to Nei and Cavalli-Sforza genetic distances produced identical trees (not shown). The two neighbor-joining trees were similar to the Fitch-Margoliash trees based on Cavalli-Sforza distances in placing the *montezumae* and *cortezi* clades together but differed from both Fitch-Margoliash trees in how they resolved relationships within *X. nezahualcoyotl* (specifically, the placement of southern *nezahualcoyotl* 8411).

The optimal tree (S = 1.80852) produced by the minimum evolution analysis of the data converted to Nei distances is presented in Figure 6. The optimal tree (S = 2.74356) produced by minimum evolution analysis of Cavalli-Sforza distances differed from the tree based on Nei distances in placing *X. malinche* and *X. birchmanni* with *X. cortezi,* placing the *pygmaeus* clade as sister group to the rest of the northern swordtails, and in the placement of southern *neza-*

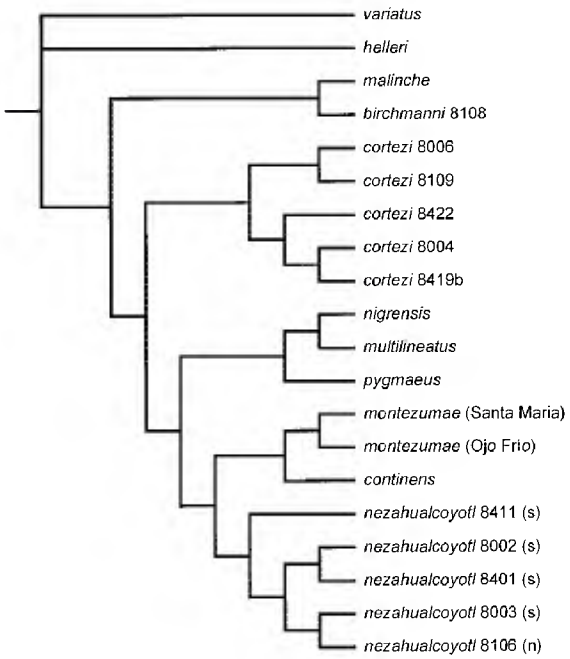

Fig. 6. The single best tree produced by the minimum evolution analysis of the reduced dataset converted to Nei genetic distances. S = 1.80852.

hualcoyotl 8411 within *X. nezahualcoyotl*. Overall, the trees from the various distance analyses were similar, differing only in the relationships among the three larger clades, the monophyly of the *cortezi* clade, and the relationships of southern *nezahualcoyotl* 8411.

Comparison with previous studies.—Comparisons between the most-parsimonious trees found under constraints corresponding with previously published trees and the most-parsimonious unconstrained trees are given in Table 2. In general, the constrained trees that were most different from the unconstrained trees as assessed by symmetric differences also provided the worst fit to the data as assessed by tree length (Spearman Correlation, $Z = 2.50$, $P = 0.01$). Comparison of the optimal trees without constraints with the optimal trees under various constraints using the Wilcoxon signed ranks test revealed that five of the constraint trees corresponding with previously published hypotheses are rejected by the allozyme data (Table 2): the parsimony and neighbor-joining trees presented by Meyer et al. (1994) as well as figures 4, 5, and 6b presented by Marcus and McCune (1999). Results under likelihood were similar, except that only the neighbor-joining tree of Meyer et al. (1994) and figure 4 of Marcus and McCune (1999) were rejected, although the parsimony tree of Meyer et al. (1994) and figures 5 and 6b of Marcus and McCune (1999) were all the next worst trees in terms of their ln-likelihood scores.

Discussion

The allozyme data reported here provide different amounts of resolution and support concerning relationships among species of northern swordtails. Comparison of the optimal trees based on different phylogenetic methods indicates that the trees are congruent for nodes that were reasonably well supported by the individual methods. We considered relationships well supported if they were present in the optimal trees of all the different analyses and were also supported in more than 50% of the bootstrap replicates within the individual analyses where such analyses were performed. Three of these relationships were between species: monophyly of the northern swordtails, monophyly of the *pygmaeus* clade of *X. nigrensis, X. multilineatus,* and *X. pygmaeus*, and *X. nigrensis* as the sister species of *X. multilineatus*. The strongly supported relationships within species included a clade of the two populations of *X. montezumae*, suggesting that this species is monophyletic, and a clade of southern *X. nezahualcoyotl* 8401, 8002 and 8407 (when present). The last two populations (southern *X. nezahualcoyotl* 8401 and 8002) are from adjacent tributaries of the Río El Salto and are separated by only 8 km. In addition, there was strong support to suggest that populations of *X. nezahualcoyotl* do not form monophyletic groups based on the drainage in which they occur. Instead, two of five populations from the Río Pánuco drainage (southern *X. nezahualcoyotl* 8003 and southern *X. nezahualcoyotl* 8402), both collected upstream in the Río El Salto and Arroyo Los Gatos, are more closely related to the northern populations from the Río Tamesí drainage than to other populations from the Río Pánuco drainage.

We considered relationships weakly supported if they were supported by 50% or more of the bootstrap replicates (where applicable) and were not contradicted by the optimal trees in any of the analyses, or they were present in the optimal trees of all analyses, but not necessarily in greater than 50% of the bootstrap replicates. The weakly supported relationships included the interspecific relationship between *X. montezumae* and *X. continens*, and the monophyly of the species *X. nezahualcoyotl*. In addition, the *montezumae* clade was present in the optimal tree of some of the analyses and not contradicted by any of them. The relationships for which there was disagreement between analyses were those for which the data were least infor-

TABLE 2. COMPATIBILITY OF PREVIOUSLY PUBLISHED HYPOTHESES WITH ALLOZYME DATA.

Study/tree	No. constrained nodes	No. conflicting nodes[a]	No. optimal trees	Tree length (MP = 53.625)	Symmetric difference[b]	*Z*-score[c]	*P*-value[c]
X. helleri placed within N. swordtails	2	1 (1)	108	55.260	12	−0.32–−0.53	0.84–0.65
Rosen (1979)	4	1 (1)	6	55.635	10	−0.32	0.84
Rauchenberger et al. (1990)	9	0 (1)	6	54.260	13	−0.68	0.62
Haas (1992)	4	0 (0)	12	54.115	8	−1.07	0.50
Meyer et al. (1994) Parsimony	10	5 (5)	2	59.025	14	−2.37	0.02*
Meyer et al. (1994) Neighbor-Joining	10	4 (5)	6	59.850	15	−2.38	0.02*
Borowsky et al. (1995) fig. 4	3	0 (0)	144	54.260	8	−0.34–−0.74	0.80–0.62
Borowsky et al. (1995) fig. 5	7	0 (2)	6	55.260	13	−0.95	0.44
Borowsky et al. (1995) fig. 6	7	0 (0)	6	54.260	13	−0.68	0.62
Borowsky et al. (1995) fig. 7	9	1 (2)	6	54.380	13	−0.85	0.46
Borowsky et al. (1995) fig. 8	9	1 (1)	6	55.160	13	−0.98	0.38
Borowsky et al. (1995) fig. 9	9	1 (1)	6	55.160	15	−0.98	0.38
Marcus and McCune (1999) fig. 3	8	0 (1)	6	54.260	13	−0.68	0.62
Marcus and McCune (1999) fig. 4	9	5 (6)	2	60.255	14	−2.38	0.02*
Marcus and McCune (1999) fig. 5	9	4 (4)	2	59.025	14	−2.37	0.02*
Marcus and McCune (1999) fig. 6a	8	2 (3)	6	57.280	15	−1.82	0.07
Marcus and McCune (1999) fig. 6b	9	4 (4)	2	59.025	14	−2.37	0.02*
Marcus and McCune (1999) fig. 6c	7	2 (2)	36	57.280	10	−1.82–−1.86	0.08

[a] Number of nodes in the constraint tree that conflict with those in the strict (semistrict) consensus tree for the trees based on parsimony, likelihood, Fitch-Margoliash, neighbor-joining, and minimum evolution analysis of the allozyme data.

[b] Between the strict consensus of the most-parsimonious trees in the absence of constraints and the strict consensus of those constrained to correspond with the trees of previous authors.

[c] For comparison between one of the optimal trees (the one identified as "best" by PAUP) and all of the best trees consistent with a given constraint. Values obtained when polytomies resolved arbitrarily. Limiting comparisons to binary trees gave similar results. Critical values of Wilcoxon rank sum obtained from table for $n < 25$.

* Significant at alpha = 0.05.

mative and included the placement of *X. birchmanni, X. malinche,* and the root for the northern swordtails.

All of the phylogenetic analyses of the allozyme data presented in this study supported monophyly of the northern swordtails. Rosen (1960) recognized the northern swordtails plus *X. milleri* as a group (his *montezumae* species group) but later rejected it as paraphyletic (Rosen, 1979). Rauchenberger et al. (1990) identified several characters that allegedly support monophyly of the northern swordtails, including two pigment characters, two sword characters, and five allozyme characters, but they did not test for monophyly in their analysis (our results suggest that three—SMDH-2, PVALB-1, SSOD—support monophyly). In addition to the current study, monophyly of the northern swordtails has been supported by DNA sequences (Meyer et al., 1994), RAPDs (Borowsky et al., 1995) and a combined phenotypic dataset (Marcus and McCune, 1999). In the constraint tree analysis, the allozyme data alone were unable to reject the placement of *X. helleri* within the northern swordtails; nevertheless, monophyly of the northern swordtails is favored by both the allozyme and other data.

Two other relationships between species were supported in all of the analyses of the allozyme data: the *pygmaeus* clade, and within that clade, *X. nigrensis* as the sister species of *X. multilineatus.* Rauchenberger et al. (1990) described the *pygmaeus* clade at the same time they recognized populations formerly referred to as *X. nigrensis* as two species, *X. nigrensis* from the Río Choy and *X. multilineatus* from the Río Coy. These two species are very similar but nevertheless are distinguishable both morphologically and allozymically (Rauchenberger et al., 1990). DNA sequence data of Meyer et al. (1994) supported the sister species relationship between *X. nigrensis* and *X. multilineatus* in both parsimony and neighbor-joining analyses (95% and 96% bootstrap values respectively). These analyses did not support the *pygmaeus* clade, placing *X. pygmaeus* with *X. birchmanni* in the maximum-parsimony tree and with *X. montezumae* in the neighbor-joining tree, though both relationships were weakly supported (see below). In addition, the *pygmaeus* clade was not supported by the various analyses of the Meyer et al. (1994) sequence data by Marcus and McCune (1999), and again placement of *X. pygmaeus* with other taxa was weakly supported (see below). In contrast, the *pygmaeus* clade is present in all of trees presented in Borowsky et al. (1995), but in the two of their trees where all three species of this clade were included (Borowsky et al., 1995:figs. 8–9), *X. multilineatus* is the sister species of *X. pygmaeus* rather than of *X. nigrensis* (bootstrap values or any other measures of support were not provided). This clade is also present in the combined data analysis of Marcus and McCune (1999:fig. 3). Finally, analysis of behavioral data by Haas (1992) supported a clade of *X. pygmaeus* and *X. nigrensis* (*X. multilineatus* was not included in that study) and so did the combined data analysis of Marcus and McCune (1999:fig. 3). In summary, the *pygmaeus* clade is supported by allozymes, morphology, RAPD's and behavior, and only weakly contradicted by DNA sequences, whereas the sister species relationship between *X. nigrensis* and *X. multilineatus* is supported by allozymes, morphology and DNA sequences and only contradicted by RAPDs.

Of the intraspecific relationships supported in all the analyses of the allozyme data, the close relationship between certain northern populations of *X. nezahualcoyotl* from the Río Pánuco drainage and the populations from the Río Tamesí drainage bears on the hypothesis presented by Rauchenberger et al. (1990) concerning colonization of the Río Tamesí drainage by this species. Those authors suggested the streams in the Ocampo area, which are currently part of the Río Tamesí system, formerly had a southern connection to the Río Pánuco system through the Río Los Gatos or the Río El Salto. Our finding that the population from the Río Los Gatos (southern *X. nezahualcoyotl* 8003) is closely related to the populations from the Río Tamesí drainage supports their proposed connection. The situation is complicated, however, by an apparent close relationship between only one of the four populations (southern *X. nezahualcoyotl* 8402) from the Río El Salto drainage and the population from the Río Los Gatos. Sampling of additional localities in these rivers might clarify the phylogeography of this species.

In a previous phylogenetic analysis of the northern swordtails, Rauchenberger et al. (1990) used 11 of the 28 electrophoretic characters used in our analysis and two that we did not include (CA1, carbonic anhydrase-1 and ADA, adenesine deaminase) in addition to morphological characters. CA1 was not included in this study because of missing data: the stain for CA1 was developed after the beginning of the allozyme study, and therefore several of the early populations were not examined for this locus. ADA was not included because it has numerous alleles, making it difficult to compare all pairs of alleles directly to determine their relative mobilities. Rauchenberger et al. (1990) coded individual alleles as characters and their presence/absence as character states; this method

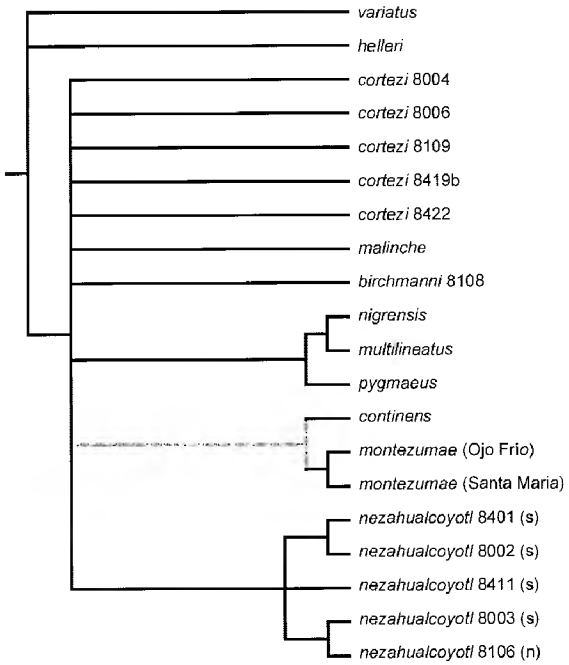

Fig. 7. Strict and semistrict consensus trees of the optimal trees for the allozyme data based on parsimony, likelihood, Fitch-Margoliash, neighbor-joining and minimum evolution methods. The tree as illustrated is the semistrict consensus tree; the branch in gray is missing from the strict consensus tree.

results in unrealistic reconstructions of ancestral allelic conditions and is generally no longer recommended for allozyme data (Buth, 1984; Mickevich and Mitter, 1981, 1982). Nevertheless, the results of our analyses of the allozyme data were incongruent with only one of the relationships they presented (*X. montezumae* as sister species of *X. nezahualcoytl* rather than *X. continens*).

Comparison of previously published and optimal trees using Templeton's (1983) test reveals that 12 of the previously published trees (those of Rosen, 1979; Rauchenberger et al. 1990; Haas, 1992; Borowsky et al., 1995; and three trees of Marcus and McCune, 1999) do not differ significantly from the optimal trees in terms of the number of allozyme characters that favor alternative trees (Table 2). These 12 trees also explain the allozyme data reasonably well in terms of tree length (none are more than 7% longer than the optimal trees) as well as likelihood (all are within 22 ln-likelihood units of the optimal tree). Six of the 12 trees (those of Rauchenberger et al. 1990; Haas, 1992; Borowsky et al., 1995:figs. 4–6; Marcus and McCune, 1999:fig. 3) exhibit no conflicts with the strict consensus of the optimal trees for the allozyme data based on the various analytical methods used in this study (Fig. 7). Four others (those of Rosen, 1979; and Borowsky et al., 1995:figs. 7–9) exhibit only a single such conflict. In some cases, the previously published trees are highly resolved and exhibit only minor differences from the optimal trees for the allozyme data. For example, the tree of Rauchenberger et al. (1990) differs from the optimal trees for the allozyme data only in the relationships among the three species of the *montezumae* clade, the tree from Figure 3 (Marcus and McCune, 1999) based on the combined dataset differs only in the relationships among the *cortezi* clade, whereas two of the trees of Borowsky et al. (1995:figs. 7–8) differ only in the relationships among the three species of the *pygmaeus* clade. In other cases, the previously published trees have relatively few resolved nodes and thus little potential for conflict. For example, the tree of Haas (1992) does not include *X. continens* and *X. malinche* and has only two resolved nodes within the northern swordtails. Similarly, one of the trees of Borowsky et al. (1995:fig. 4) includes only three of the nine northern swordtail species.

The only previously published trees that provide a significantly worse fit to the allozyme data, as assessed by Templeton's (1983) test, are the two trees of Meyer et al. (1994) and the three trees from Marcus and McCune (1999), all of which are based on datasets that include at least some of the DNA sequence data of Meyer et al. (1994). These trees also provide the worst fits to the allozyme data as assessed by both tree length (all are at least 10% longer than optimal trees) and likelihood (none are within 26 ln-likelihood units of the optimal tree), and they are the most different from the optimal trees as assessed by the number of conflicting nodes (Table 2). The two trees from Meyer et al. (1994), based on parsimony and neighbor-joining analyses of DNA sequence data, differ in several ways both from each other and from the optimal trees for the allozyme data. Notable, however, is the placement of *X. pygmaeus*, which is the putative sister group of *X. birchmanni* with the two species as the putative sister group of all other northern swordtails in the parsimony tree [the tree of Marcus and McCune (1999) based on the same data (their fig. 4) is nearly identical], and is the putative sister group of *X. montezumae* in the neighbor-joining tree. Because these relationships contradict one another, the placement of *X. pygmaeus* is unresolved in the strict (and semistrict) consensus of the two trees (not shown) and the alternative relationships are only weakly supported, with bootstrap values less than 50% (the actual values are 31% for *X. pygmaeus* and *X. birchmanni* in the parsimony analysis, and 47% for

X. pygmaeus and *X. montezumae* in the neighbor-joining analysis). The trees in figures 5 and 6b (Marcus and McCune, 1999) place *X. birchmanni* and *X. pygmaeus* outside of a clade formed by the rest of the northern swordtails but not as sister groups (bootstraps were 42% or less). Additional analyses using Templeton's (1983) test with constraint trees and the allozyme data nearly reject a sister-group relationship between *X. pygmaeus* and *X. birchmanni* ($Z = -1.99$, $n = 6$, $P = 0.06$), as well as one between *X. pygmaeus* and *X. montezumae* ($Z = -1.57$–2.02, $n = 6$–5, $P = 0.12$–0.06). Similar analyses with trees constrained to have *X. birchmanni* and *X. pygmaeus* outside of a clade formed by the rest of the northern swordtails rejected that relationship ($Z = -2.20$, $n = 7$, $P = 0.03$).

In summary, the results of phylogenetic analyses of allozyme data are highly congruent with those of similar analyses of datasets derived from morphology, behavior, and randomly amplified DNAs. They are also largely congruent with the results of phylogenetic analyses of DNA sequence data, differing primarily in relationships that are weakly supported by the sequence data. Although the allozyme data are only able to resolve about half of the 17 possible nodes within the northern swordtail clade (bootstraps $> 50\%$), they are nonetheless able to reject several previously proposed hypotheses of relationships within this clade.

Acknowledgments

We thank D. L. Swofford for access to, and permission to publish results from, test versions of PAUP*, K. Kallman., A. Basolo, R. Borowsky, M. Ryan for assistance in collecting samples in the field, and the government of the Republic of Mexico for numerous permits over the years (to K. Kallman, D. Morizot, R. Borowsky, and M. Ryan) including the most recent to M. Morris (permit DAN-00162). We thank J. Felsenstein for help with PHYLIP. Finally, we extend a special thanks to K. Kallman. Without his extensive knowledge of the biogeography of these fishes the collections for this study could not have been made. In addition, we thank him for his critical comments on the manuscript.

Literature Cited

Basolo, A. L. 1990. Female preference predates the evolution of the sword in swordtail fishes. Science 250:808–810.

———. 1995. A further examination of a pre-existing bias favouring a sword in the genus *Xiphophorus*. Anim. Behav. 50:365–375.

Berlocher, S. H., and D. L. Swofford. 1997. Searching for phylogenetic trees under the frequency parsimony criterion: an approximation using generalized parsimony. Syst. Biol. 46:211–215.

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

Buth, D. G. 1984. The application of electrophoretic data in systematic studies. Annu. Rev. Ecol. Syst. 15: 501–522.

Cavalli-Sforza, L. L, and A. W. F. Edwards 1967. Phylogenetic analysis: models for estimation procedures. Evolution 3:550–570.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17:368–376.

———. 1985. Confidence limits on phylogenies: an approach using bootstrap. Evolution 39:783–791.

———. 1988. Phylogenies from molecular sequences: inference and reliability. Annu. Rev. Genet. 22: 521–565.

Fitch, W. M., and E. Margoliash. 1967. Construction of phylogenetic trees. Science 155:279–284.

Haas, V. 1992. *Xiphophorus* phylogeny, reviewed on the basis of courtship behaviour, p 279–288. *In:* Trends in ichthyology, an international perspective. J. H. Schroder, J. Bauer, and M. Schartl (eds.). Blackwell Scientific, Oxford.

International Union of Biological Nomenclature. 1984. Enzyme nomenclature, 1984. Academic Press, Orlando, FL.

Kishino, H., and M. Hasegawa 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.

Marcus, J. M., and A. R. McCune. 1999. Ontogeny and phylogeny in the northern swordtail clade of *Xiphophorus*. Syst. Biol. 48:491–521.

McAlpine, P. J., C. Boucheix, A. J. Pakstis, L. C. Stranc, T. G. Berent, and T. B. Shows. 1988. The 1988 catalog of mapped genes and report of the nomenclature committee. Cytogenet. Cell. Genet. 49:4–45.

Meffe, G. K,. and F. F. Snelson Jr. 1989. Ecology and evolution of livebearing fishes. Prentice Hall, Englewood Cliffs, NJ.

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

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

Mickevich, M. F., and C. Mitter. 1981. Treating polymorphic characters in systematics: a phylogenetic treatment of electorphoretic data, p. 46–58. *In:* Advances in cladistics. Vol. 1. V. A. Funk and D. R. Brooks (eds.). New York Botanical Gardens, Bronx.

———, and ———. 1982. Evolutionary patterns in allozyme data: a systematic approach, p. 118–176.

In: Advances in cladistics. Vol. 2. Columbia Univ. Press, New York.

MORIZOT, D. C., AND M. E. SCHMIDT. 1990. Starch gel electrophoresis and histochemical visualization of proteins, p. 23–80. *In:* Applications of electrophoresis and isoelectric focusing in fisheries management. D. H. Whitmore (ed.). CRC Press, Boca Raton, FL.

———, AND M. J. SICILIANO. 1979. Polymorphisms, linkage and mapping of four enzyme loci in the fish genus *Xiphophorus* (Poeciliidae). Genetics 93:947–960.

———, AND ———. 1982. Protein polymorphisms, segregation in genetic crosses and genetic distances among fishes of the genus *Xiphophorus* (Poeciliidae) *Ibid.* 102:539–556.

———, D. A. WRIGHT, AND M. J. SICILIANO. 1977. A three point linkage of enzyme loci in fishes: implications in the evolution of vertebrate chromosomes. *Ibid.* 86:645–656.

MORRIS, M. R., AND M. J. RYAN. 1996. Sexual difference in signal-receiver coevolution. Anim. Behav. 52:1017–1024.

NEI, M. 1972. Genetic distance between populations. Am. Nat. 106:283–292.

PENNY, D., AND M. D. HENDY 1985. The use of tree comparison metrics. Syst. Zool. 34:75–82.

RAUCHENBERGER, M., K. D. KALLMAN, AND D. C. MORIZOT. 1990. Monophyly and geography of the Pánuco Basin swordtails (genus *Xiphophorus*) with description of four new species. Am. Mus. Novit. 2975: 1–41.

ROSEN, D. E. 1960. Middle American poeciliid fishes of the genus *Xiphophorus.* Bull. Fla. State. Mus. Biol. Sci. 5:57–242.

———. 1979. Fishes from the uplands and intermontane of Guatemala: revisionary studies and comparative geography. Bull. Am. Mus. Nat. Hist. 162:268–375.

RZHETSKY, A., AND M. NEI. 1992. A simple method for estimating and testing minimum-evolution trees. Mol. Biol. Evol. 9:945–967.

SAITOU, N., AND M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Ibid.* 4:406–425.

SHAKLEE, J. B., F. W. ALLENDORF, D. C. MORIZOT, AND G. S. WHITT. 1990. Gene nomenclature system for protein-coding loci in fish. Trans. Am. Fish. Soc. 119: 2–15.

SICILIANO, M. J., AND C. R. SHAW. 1976. Separation and visualization of enzymes on gels, p. 185–209. *In:* Chromatographic and electrophoretic techniques. I. Smith (ed.). Wm. Heineman, London.

SWOFFORD, D. L., AND S. H. BERLOCHER. 1987. Inferring evolutionary trees from gene frequency data under the principle of maximum parsimony. Syst. Zool. 36:293–325.

———, AND G. J. OLSEN. 1990. Phylogeny reconstruction, p. 411–501. *In:* Molecular systematics, D. M. Hillis and C. Moritz (eds.). Sinauer Associates, Sunderland, MA.

———, ———, P. J. WADDELL, AND D. M. HILLIS. 1996. Phylogenetic inference, p. 407–425. *In:* Molecular systematics. 2d ed. D. M. Hillis, M. C. Moritz, and B. K. Mable (eds.). Sinauer Associates, Sunderland, MA.

TEMPLETON, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221–244.

WIENS, J. J. 1995. Polymorphic characters in phylogenetic systematics. Syst. Biol. 44:482–500.

———. 1999. Polymorphisms in systematics and comparative biology. Annu. Rev. Ecol. Syst. 30:327–362.

———, AND M. SERVEDIO. 1997. Accuracy of phylogenetic analysis including and excluding polymorphic characters. Syst. Biol. 46:332–345.

(MRM) DEPARTMENT OF BIOLOGICAL SCIENCES, OHIO UNIVERSITY, ATHENS, OHIO 45701; (KQ) DEPARTMENT OF VERTEBRATE ZOOLOGY, NATIONAL MUSEUM OF NATURAL HISTORY, SMITHSONIAN INSTITUTION, WASHINGTON, DC 20560; AND (DCM) UNIVERSITY OF TEXAS M. D. ANDERSON CANCER CENTER, SCIENCE PARK, RESEARCH DIVISION, P.O. BOX 389, SMITHVILLE, TEXAS 78957. E-mail: (MRM) morrism@oak.cats.ohiou.edu. Send reprint requests to MRM. Submitted: 28 Oct. 1998. Accepted: 29 June 2000. Section editor: J. D. McEachran.

APPENDIX 1. TAXA AND LOCALITIES SAMPLED. D. C. Morizot's collection numbers with sample sizes followed by AMNH collection numbers for those collections with specimens deposited in the American Museum of Natural History. Localities (arroyo or river, drainage, state) also given. na = no collection deposited. State abbreviations: VER = Veracruz; SLP = San Luis Potosí; HID = Hidalgo; TAM = Tamaulipas.

Species	Collection numbers		Locality
	DCM# (N)	AMNH#	
X. pygmaeus	8112 (40)	na	Río Huichihuayan at Chimalaco (La Y Griega Vieja), Río Axtla, SLP
X. multilineatus	8113 (28)	na	Río Coy at HWY 85, Río Tampaón, SLP
X. nigrensis	8308 (22)	na	Nacimiento Río Choy, E of Cd. Valles, Río Tampaón, SLP
X. cortezi	8004 (15)	na	Arroyo Lagarto, E of Tamazunchale, Río Moctezuma, SLP
	8006 (28)	na	Arroyo La Conchita at Xilitla, Río Axtla, SLP
	8109 (10)	45300	Río Tancuilin, W of Esclanar off Hwy 85, Río Axtla, SLP
	8419b (6)	75920	Arroyo Santa Cruz, 10 km E of Orizatlan, Río Tempoal, HID
	8420 (2)	75925	Río Tecoluco at Macustepetla, W of Huejutla, Río los Hules, HID
	8422 (23)	75930	Arroyo Seco, 4 km W of Comola, Río Axtla, SLP
X. continens	8102 (38)	45306	Río Ojo Frío, 7 km N of Damian Carmona, Río Gallinas, SLP
X. montezumae	8103 (3)	45307	Río Ojo Frío, 7 km N of Damian Carmona, Río Gallinas, SLP
	8104 (11)	na	Río Ojo Frío, 7 km N of Damian Carmona, Río Gallinas, SLP
	8309 (6)	45344	Arroyo La Cienega at Ojo Caliente, Río Santa Maria, SLP
	8323 (6)	78570	Arroyo La Cienega Grande at Viente de Noviembre, Río Santa Maria, SLP
	8413 (6)	75912	Northernmost source, Arroyo La Cienega Grande, Río Santa Maria, SLP
	8425 (25)	75943	Source of Río Ojo Frío, N of Rascon, Río Santa Maria, SLP
X. nezahualcoyotl	8002 (23)	45260	Arroyo La Barranca at Santa Barbarita, W of Naranjo, Río El Salto, SLP,
(southern/Pánuco)	8003 (10)	45262	Source of Río los Gatos, 8 km N of Nueve Morelos, Río Valles, TAM
	8401 (27)	76790	Arroyo Hondo at Francia Chica, Río El Salto, SLP
	8402 (25)	na	Arroyo El Sabinito, 5 km W of El Naranjo, Río El Salto, SLP
	8407 (30)	75895	Río El Caballete at Minas Viejas, Río El Salto, SLP
	8411 (20)	na	Source of Río Tanchachin, W of Cd. Valles on HWY 70, Río Tampaón, SLP
X. nezahualcoyotl	8403 (25)	75886	Arroyo La Villa, 3 km SE of Ocampo at La Muralla, Río Ocampo, TAM
(northern/Tamesí)	8405 (12)	75891	Río St. Maria de Guadalupe, 18 km NW of Ocampo, Río Ocampo, TAM
	8106 (21)	na	Source of Arroyo El Zarco, NW of El Encino, Río Sabinas, TAM
X. birchmanni	8108 (24)	45314	Arroyo Aclatle, 4 km NE of Chicontepec, Río Calabozo, VER
	8419a (9)	75920	Arroyo Santa Cruz, 10 km E of Orizatlan, Río Tempoal, HID
	8718 (10)	77994	Arroyo Pilpuerta, 11 km SE of Benito Juarez, Río Tuxpan, VER
	8721 (7)	78007	Río Santa Maria, near Santa Maria, Río San Pedro, HID
	8722 (24)	78011	Río Camaitlan, 12 km SE of Chicontepec, Río Calabozo, VER
X. malinche	9700 (20)	na	Río Claro, 13 km NW of Molango, Río Moctezumae, HID
X. variatus	8001 (20)	na	Arroyo El Zarco, NW of El Encino, Río Sabinas, TAM
X. helleri	8116 (11)	na	Nacimiento at Finca Santa Anita, 5 km NW of Protero Nuevo, Río Atoyac, VER

Appendix 2. Allele Frequencies of 28 Presumptive Loci among Populations of Northern Swordtails and 2 Outgroups. 1 = *cortezi*8004; 2 = *cortezi*8006; 3 = *cortezi*8109; 4 = *cortezi*84196b; 5 = *cortezi*8422; 6 = *cortezi*8420; 7 = *malinche*; 8 = *birchmanni*8108; 9 = *birchmanni*8419a; 10 = *birchmanni*8718; 11 = *birchmanni*8721; 12 = *birchmanni*8722; 13 = *nigrensis*; 14 = *pygmaeus*; 15 = *multilineatus*; 16 = *continens*; 17 = *montezumae*(Ojo Frío); 18 = *montezumae*(Santa Maria); 19 = southern *nezahualcoyotl*8401; 20 = southern *nezahualcoyotl*8002; 21 = southern *nezahualcoyotl*8407; 22 = southern *nezahualcoyotl*8411; 23 = southern *nezahualcoyotl*8003; 24 = southern *nezahualcoyotl*8402; 25 = northern *nezahualcoyotl*8403; 26 = northern *nezahualcoyotl*8405; 27 = northern *nezahualcoyotl*8106; 28 = *helleri*; 29 = *variatus*.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
*sAH**																													
a																													100
b													92	97	100							100							
c	100	100	100	100	100	100	100	100	100	100	100	100	8	3		100	93	100	100	100	100		100		100	100	100		
d																	7							100					
e																												100	
*CK-A**																													
a							100																						100
b	100	100	100	100	100	100		100	100				100	100	100	100			100	100	100	100	100	100	100	100	100		
c																	100	100										50	
d																												50	
*GAPDH-1**																													
a																													100
b																												100	
c	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100		
*GAPDH-2**																													
a	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100		100	100	100	98	100	100	100	100	100	100	100	100
b																	100												
c																					2								
*GLYDH**																													
a	100	100	100	100	100	100	100	100					100	100	100		100	100	100	100	100	100			100	100	100	100	
b																							100	100					
c																100													100
*GPI-1**																													
a			5													100													
b	100	100	95	100	76	100	100	100	100	100	100	100	100	100	100		100	100	100	100	100	100	100	100	100	100	100	100	100
d					24																								
*GPI-2**																													
a	100	100	100	100	100	100	100	100	100	100	100	100	95	100	36	100	100	100	100	100	100	100	100	100	100	100	100	100	100
b													5		64														

Appendix 2. Continued.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
*GUK1**																													
a	100	100	100	92	100	75	100	6		100	5					100	100	97	100	100	82	100	100	100	100	100	100	100	
b								63			45	2																	
c				8		25		31			50	98																	
d													100	100	100														57
e																		3											43
f																					18								
*sIDHP-2**																													
a	97	100	100	92	100	100	100	100	50		100	100					100	100	100	100	100	100	100	100	13		100		
b	3																								87	100		100	100
c				8					50					1															
d													100	99	100	100													
*LDH-C**																													
a	100	100	75	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
b			25																										
*LDH-A**																													
a	100	100	100	100	100	50	100	77	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
b								23																					
c						50																							
*LDH-B**																													
a	100	100	100	100	100	100	100	100	100	100	100	100	39	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
b													61																
*sMDH-1**																													
a	100	100	100	100	93	100	100	100	100	100	100	100	52	96	100													100	100
b																100	99	100	100	100	100	100	100	100	100	100	100		
c					7												1												
d														4															
e													48																
*sMDH-2**																													
a																												100	100
b																					2		100	100	100	100	100		
c	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	71	100	100	100	100	98	100							
d																29													

Appendix 2. Continued.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
*PROT5**																													
a	100	100	100	100	100	100		29	61				100	100	100	100	100	100	100	100	72	100						100	
b							100	71	39			100									28		100	100	100	100	100		
c																													100
*MPI**																													
a																	100	100											
b	97	100	100	100	100	100	100	44		50		23	100	100	98				100	100	100	100	100	100	100	100	100	100	100
c															2														
d																100													
e	3							56	100	50		77																	
*PEPS**																													
b																			100	91	92	100	100	100	100	100	100	100	
c													100	95	13														
d	47		5	33	50			100	100	100	100	27		5	87	100	100	100		9	8							100	
e	53	100	95	67	50	100	100					73																	100
*PEPB**																													
a	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	94	100	100	100	100	100	100	100	100	100	100	80
b																		6											20
*PGDH**																													
b	100	100	100	100	100	100	100	100	100	100	100	100	100	100	98	100	100	100	100	100	100	100	100	100	100	100	100	100	100
c															2														
*PGM**																													
b	100	100	100	100	100	100	100	100	100	100	100	100				90	99	100	100	100	100	100	100	100	100	100	100	100	100
c													100		100														
d														100															
e																10													
f																	1												
*PVALB-1**																													
a																												100	100
b	100	100	100	100	100	100	100	100	100	100	100		100		100	100	100	100	100	100	100	100	100	100	100	100	100		
c														100															
*PVALB-2**																													
a	100	100	100	100	100	100	100	100		100			100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	
b																													100

APPENDIX 2. CONTINUED.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
*PVALB-3**																													
a	100	100	100	100	100	100	100	100		100			100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
*sSOD**																													
a													30															100	100
b	100	100	100	100	100	100	100	100		100	100	100	70	100	100	100	100	100	100	100	100	100	100	100	100	100	100		
*TPI-2**																													
a	100	100	100	100	95	100	100	100	100	100	100	100	100	100	100	100	97	100	100	100	100	92		100	100	100	100	100	
b																						8	100						100
c					5																								
d																	3												
*TPI-1**																													
a																								70					
b	100	100	100	100	100	75	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	30	100	100	100		100
c						25																						100	
*UMPH**																													
a	75	9	25	83	45	100					7			1						2			6						
b	25	36	10		38		100	100	100	100	86	100																	100
c		55	65	8	18						7		100	99	100	100	100	100	100	98	100	100	94	100	100	100	100	100	
d				8																									
*UGHUT**																													
a					5								89		100				91	63	43								
b	100	100	100	100	93	100	100	92	100	100	100	69		100		100	100	100	9	37	35	100	100	72	100	100	100	100	100
c								8				31												28					
d					2																22								
e													11																