2000, No. 1 COPEIA February 1

Copeia, 2000(1), pp. 1–10

Phylogenetic Relationships in the North American Cyprinid Genus *Cyprinella* (Actinopterygii: Cyprinidae) Based on Sequences of the Mitochondrial ND2 and ND4L Genes

RICHARD E. BROUGHTON AND JOHN R. GOLD

Shiners of the cyprinid genus *Cyprinella* are abundant and broadly distributed in eastern and central North America. Thirty species are currently placed in the genus: these include six species restricted to Mexico and three barbeled forms formerly placed in different cyprinid genera (primarily *Hybopsis*). We conducted a molecular phylogenetic analysis of all species of *Cyprinella* found in the United States, using complete nucleotide sequences of the mitochondrial, protein-coding genes ND2 and ND4L. Maximum-parsimony analysis recovered a single most-parsimonious tree for *Cyprinella*. Among historically recognized, nonbarbeled *Cyprinella*, the mitochondrial (mt) DNA tree indicated that basal lineages in *Cyprinella* are comprised largely of species with linear breeding tubercles and that are endemic to Atlantic and/or Gulf slope drainages, whereas derived lineages are comprised of species broadly distributed in the Mississippi basin and the American Southwest. The Alabama Shiner, *C. callistia*, was basal in the mtDNA tree, although a monophyletic *Cyprinella* that included *C. callistia* was not supported in more than 50% of bootstrap replicates. There was strong bootstrap support (89%) for a clade that included all species of nonbarbeled *Cyprinella* (except *C. callistia*) and two barbeled species, *C. labrosa* and *C. zanema*. The third barbeled species, *C. monacha*, fell outside of *Cyprinella* sister to a species of *Hybopsis*. Within *Cyprinella* were a series of well-supported species groups, although in some cases bootstrap support for relationships among groups was below 50%. A derived clade consisting of *C. spiloptera*, *C. whipplei*, *C. venusta*, and the southwestern *C. lutrensis* group was strongly supported. The species *C. lutrensis* and *C. lepida* were not monophyletic, suggesting further study and revision within this group are warranted. In general, the most-parsimonious mtDNA tree was similar in terms of relationships among species to those proposed more than 40 years ago by R. H. Gibbs.

THE largest number of freshwater fish species in North America belong to the family Cyprinidae, and because of their diversity, cyprinids have played an important role in studies of speciation and North American biogeography (Mayden, 1988; Buth et al., 1991; Kristmundsdòttir and Gold, 1996). However, North American cyprinids have received limited attention with respect to phylogenetic relationships among genera and species (Mayden, 1989; Coburn and Cavender, 1992; Simons and Mayden, 1998). The second largest cyprinid genus in North America (after *Notropis*), *Cyprinella* presently comprises of 30 species (Mayden, 1989; Robins et al., 1991) and includes some of the most abundant freshwater fishes in eastern and central North America. Adult *Cyprinella* seldom reach more than about 12 cm (SL), but breeding males often exhibit striking colors and elaborate tuberculation. The native distribution of the genus is streams and rivers from the Atlantic coast to the Great Plains and from southern Canada to the Gulf Coast and northern Mexico.

Girard (1856) originally described the genus *Cyprinella*, with *Leuciscus bubalinus* Baird and Girard (= *C. lutrensis*) as the type species. The next 100 years saw extensive confusion regarding taxonomy and systematics of *Cyprinella*, with member species being variously placed in at least 15 genera or subgenera (Gibbs, 1957). However, many workers noted that species of *Cyprinella* were a natural group, and Bailey and Gibbs (1956) recognized *Cyprinella* as a subgenus of the large shiner genus *Notropis*. Mayden

(1989) identified more than 34 osteological, behavioral, coloration, and tuberculation characters supporting monophyly of *Cyprinella*; consequently, he elevated *Cyprinella* to generic status (along with the related groups *Luxilus* and *Lythrurus*). Species relationships in *Luxilus* (Gilbert, 1964; Buth, 1979; Dowling and Naylor, 1997) and *Lythrurus* (Snelson, 1972; Mayden, 1989; Schmidt et al., 1998) have been studied extensively using both morphological and molecular characters. However, morphologically based assessments of relationships in *Cyprinella* (Gibbs, 1957; Mayden, 1989) yielded conflicting results that have contributed to uncertainty regarding phylogenetic relationships of several species and/or species groups.

Gibbs (1957) proposed relationships of *Cyprinella* (Fig. 1A), based primarily on external morphology and pharyngeal tooth patterns. Although his analysis was essentially phenetic, Gibbs distinguished two groups of species, based largely on the presence of breeding tubercles in rows (considered ancestral) or scattered (considered derived). Basal linages included species endemic to drainages of either the Atlantic (*C. callisema, C. nivea, C. pyrrhomelas*, and *C. xaenura*) or Gulf of Mexico (*C. callistia*, and *C. callitaenia*), or both (*C. leedsi*). The derived group included six species (*C. camura, C. galactura, C. lutrensis, C. spiloptera, C. venusta*, and *C. whipplei*) broadly distributed in Mississippi River drainages along with two from the Atlantic slope (*C. analostana* and *C. chloristia*) and one from the Gulf slope (*C. caerulea*). Species of *Cyprinella* from the southwestern United States and Mexico (the *C. lutrensis* group) were not included beyond the suggestion that this group is sister to *C. spiloptera*. Contreras-Balderas (1975) examined members of the *C. lutrensis* group, noting the existence of additional, undescribed species, but did not propose explicit phylogenetic relationships. Mayden (1989) proposed relationships within *Cyprinella* based on a cladistic analysis of numerous osteological and morphological characters. This study incorporated all species historically recognized as *Cyprinella* including all of the nominal species in the *C. lutrensis* group. Mayden's phylogenetic hypothesis (Fig. 1B) identified two primary clades: a *C. lutrensis* clade, comprised of 10 species primarily inhabiting the American Southwest and Mexico, and a *C. whipplei* clade that included the remaining 17 species. Within Mayden's *C. whipplei* clade, species from the Mississippi basin (corresponding largely to Gibbs' tubercles-scattered group) were basal relative to a clade composed primarily of Atlantic and/or Gulf slope endemics (corresponding largely to Gibbs' tubercles-linear group). An exception was *C. caerulea*, a species with scattered tubercles but that was sister to a clade of *C. gibbsi* and *C. trichroistia* (both with linear tubercles). Both Gibbs (1957) and Mayden (1989) placed *C. analostana* and *C. chloristia* (primarily Atlantic slope endemics with scattered tubercles) as sister to *C. whipplei*, a cosmopolitan, tubercles-scattered species distributed in Mississippi River and Gulf slope drainages. Thus, Gibbs (1957) and Mayden (1989) agreed, in large part, on the distinction between two groups of species, one of Atlantic and/or Gulf slope endemics with linearly arranged tubercles and one of more widespread species primarily

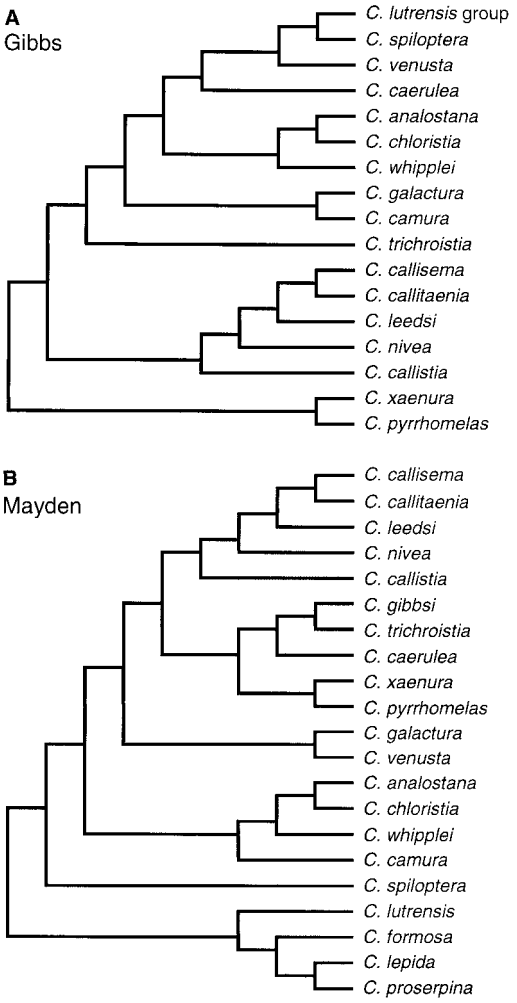

Fig. 1. Phylogenetic hypotheses (redrawn for comparison) for genus *Cyprinella*: (A) Gibbs (1957) and (B) Mayden (1989). Only species included in the present study are shown. Gibbs excluded species in the *C. lutrensis* group and *C. gibbsi* had not been described.

from the Mississippi basin and southwest with tubercles scattered. However, a trenchant difference between their hypotheses is the relative position of the two groups as basal versus derived.

In addition to the 27 species included in *Cyprinella* by Mayden (1989), three or four barbeled species, formerly placed in the genus *Hybopsis* Bailey (1951), possess morphological and behavioral similarities to nonbarbeled *Cyprinella* (Jenkins and Lachner, 1971; Coburn and Cavender, 1992; Dimmick, 1993). These include *C. monacha, C. labrosa, C. zanema*, and the undescribed thinlip chub, presently considered a subspecies of *C. zanema* (Jenkins and Lachner, 1980). Although *C. labrosa* and *C. zanema* share morphological similarities, they differ more from each other than do most sister-species pairs of *Cyprinella*, and neither appears to be closely related to *C. monacha* (Jenkins and Burkhead, 1994). Mayden (1989) placed *C. monacha* in the genus *Erimystax* and *C. labrosa* and *C. zanema* in his modified genus *Hybopsis*, yet there remains uncertainty regarding the phylogenetic and taxonomic status of these barbeled taxa.

In this paper, we report the phylogenetic analysis of nucleotide sequences of the mitochondrially encoded ND2 and ND4L genes from 21 nonbarbeled and three barbeled species of *Cyprinella*. The study included all species of *Cyprinella* occurring in the United States (six species of the *C. lutrensis* group restricted to Mexico were not available). The ND2 gene (1047 nucleotides) has been shown to be informative phylogenetically in studies of numerous vertebrate groups, including fishes (Mindell and Thacker, 1996; Zardoya and Meyer, 1996; Kocher and Carleton 1997). ND4L is smaller, consisting of 297 nucleotides but has been shown to be useful in phylogenetic analysis of closely related taxa (Bielawski and Gold, 1996).

Materials and Methods

Complete ND2 and ND4L gene sequences were obtained from all species and included at least two individuals of each species except for *C. gibbsi, C. formosa, C. labrosa, C. zanema, C. monacha*, and the outgroup species. Where possible, conspecific individuals were from geographically distant localities. Outgroup taxa, *Notropis atherinoides, Lythrurus roseipinnis, Luxilus cornutus*, and *Hybopsis winchelli*, were selected based on hypothesized phylogenetic relationships to *Cyprinella* (Mayden, 1989; Coburn and Cavender, 1992). All sequences were generated in our laboratory (except *L. cornutus*, provided by T. Dowling) and may be found in GenBank (accession numbers AF111205–AF111262). Specimens were collected by seine, placed in liquid nitrogen or dry ice for transport, and stored at −80 C until DNA extraction. Collection localities and voucher information are given in Material Examined.

Total DNAs were obtained by grinding tissues in liquid nitrogen, extracting with phenol and chloroform, and precipitating with ethanol. Amplification of the ND2 gene employed primers (developed by T. Dowling) in the flanking $tRNA^{MET}$ and $tRNA^{TRP}$ genes: ND2B-L: 5′-aagctttcgggcccataccc-3′, and ND2E-H: 5′-ttctacttaaagctttgaaggc-3′. The gene was sequenced in three segments using external primers and an internal primer, developed during the course of the study: ND2G-L: 5′-cacaacaataatccttgccgc-3′. Amplification of the ND4L gene employed primers in $tRNA^{ARG}$ (ArgB-L: 5′-caagacccttgatttcggctca-3′) and the coding region of the adjacent ND4 gene (NAP2: 5′-tggagcttctacgtgrgcttt-3′); ArgB-L was used for sequencing. Fifty μl amplification reactions included 1X *Taq* buffer (Promega Inc.), 200 mM of each dNTP, 2 mM $MgCl_2$, 0.5 μM each primer, approximately 100 ng of template DNA, and 0.25 units of *Taq* polymerase. Reaction conditions were 25 cycles of 95 C for 1 min, 55 C for 1 min, and 72 C for 2 min. Double-stranded reaction products were purified with Prep-A-Gene (Bio-Rad Laboratories). Sequences were generated via dye-terminator reactions read on an ABI 377 Prism sequencer or manually with the fmol system (Promega Inc.), using ^{32}P-labeled primers.

Sequences of the complete protein-coding region of each gene were aligned by eye and combined into a single data matrix. The possibility of nucleotide compositional bias among taxa, including outgroups, was examined with a chi-square test of nucleotide frequency homogeneity. Potential differences in evolutionary rate among lineages were investigated with Tajima's (1993) 1D test of rate homogeneity. PAUP* vers. 4.01b was used for heuristic parsimony searches with 100 random addition replicates, TBR branch swapping, and ACCTRAN optimization. Equal weighting of all character state changes was used in initial searches. Bootstrapping (1000 replicates) was used to assess relative support for clades on the most-parsimonious tree. Congruence of phylogenetic signal among genes was examined by comparison with random partitions of the data with the incongruence length test (Farris et al., 1995) implemented in PAUP*.

Results and Discussion

Sequence variation.—Only one alignment gap, involving a deletion of a single codon in the ND2

TABLE 1. NUCLEOTIDE VARIATION BY GENE AND CODON POSITION.

	Total sites	Parsimony informative[a]	Number of variable sites[b]			
			All positions	1st	2nd	3rd
ND2	1047	460	560	161 (0.461)	60 (0.172)	339 (0.971)
ND4L	297	113	161	39 (0.394)	29 (0.293)	93 (0.939)
Total	1344	573	721	200	89	432

[a] Sites with alternative nucleotide state in more than one individual.
[b] Parentheses represent proportion of variable sites.

gene in *C. caerulea*, was observed. Mean nucleotide frequencies for all taxa were A = 0.254, C = 0.303, G = 0.167, and T = 0.276. There was no significant difference in nucleotide composition among the taxa [$\chi^2_{(87)} = 69.8$, $P = 0.91$]. The pronounced bias against G nucleotides is similar to that observed in mtDNA coding genes in other vertebrate taxa (e.g., Collins et al., 1994). Uncorrected sequence difference varied among ingroup taxa, from 2.08% (*C. callitaenia* vs *C. zanema*) to 24.26% (*C. proserpina* vs *C. spiloptera* and *C. proserpina* vs *C. formosa*). Among ingroup taxa, 721 of 1344 sites were variable with 573 sites informative phylogenetically (Table 1). Nearly all third-codon positions were variable in both genes. More second-codon positions were variable in the ND4L gene, whereas ND2 exhibited more variation at first codon positions. This is consistent with Bielawski and Gold's (1996) observation that among cyprinids ND4L appears to evolve rapidly at the amino acid level. We examined relative amounts of homoplasy in the data with a modification of the consistency index (ci; Kluge and Farris, 1969) for each of the six classes of nucleotide transformation (Table 2). All changes on the tree (output as a ''change list'' by PAUP) were sorted by codon position and by class of change. For each group, the minimum number of changes (= the number of characters exhibiting a change of that type) was divided by the total number of changes observed. Resulting values provide a useful measure of the amount of change consistent with the tree (or homoplasy, taken as 1 − ci). The ci values were lower for the two transition changes (T ↔ C, A ↔ G) than for transversion changes (all others) and third-codon positions exhibited lower ci values relative to first- and second-codon positions. ND4L exhibited slightly higher ci values than ND2, although ND4L provided substantially fewer useful characters. Despite variable levels of homoplasy across genes, phylogenetic signal of each gene did not differ significantly from random data partitions (incongruence length test, $P = 0.366$).

TABLE 2. CONSISTENCY INDICES FOR CLASSES OF NUCLEOTIDE CHANGE BY GENE AND CODON POSITION.

Change	ND2			ND4L		
	1st	2nd	3rd	1st	2nd	3rd
A ↔ G	0.428	0.313	0.327	0.643	0.813	0.462
C ↔ T	0.404	0.508	0.346	0.379	0.846	0.412
A ↔ C	0.776	0.889	0.568	0.889	0.909	0.759
A ↔ T	0.730	1.000	0.683	0.800	0.750	0.810
C ↔ G	0.680	0.800	0.633	0.833	0.833	0.727
T ↔ G	0.750	0.800	0.741	0.833	0.714	0.750

Phylogenetic analysis.—Maximum-parsimony analysis, employing equal weights for all characters, recovered a single most-parsimonious tree for *Cyprinella* (Fig. 2). However, because of the high proportion of changes at third-codon positions in both genes (Table 1), and low ci values for transitions relative to transversions (Table 2), we considered the possibility that unweighted characters might support spurious phylogenetic hypotheses. Although differential weighting of characters (by codon position) and character-state changes (transitions vs transversions) might overcome potentially misleading phylogenetic signal at rapidly evolving sites (e.g., Simon et al., 1994; Kocher and Carleton, 1997; Martin and Bermingham, 1998), it is not clear when multiple nucleotide changes generate spurious or misleading hypotheses of relationships. Where homoplasy is evident, the goal is to extract true phylogenetic signal from background noise. To this end, it may be useful to identify a priori characters that are misleading (and to what extent), and impose a weighting scheme that penalizes homoplastic changes without losing historically consistent information.

The successive approximations approach of Farris (1969) imposes weights that are based on the relative consistency of characters on a tree. An advantage of this approach is that it provides an objective criterion for determining the relative weight for a character; a disadvantage is that

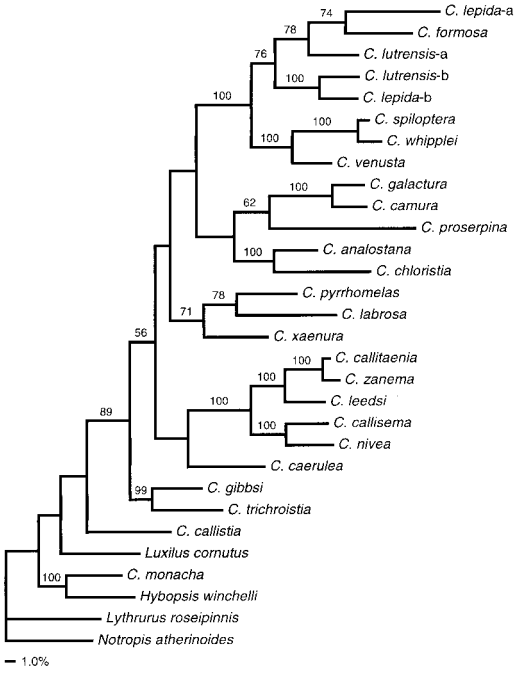

Fig. 2. Single, most-parsimonious tree for ND2 and ND4L nucleotide sequences in the genus *Cyprinella*. Branch lengths reflect number of character state changes. Scale bar corresponds to 1% difference in nucleotide sequence. Numbers indicate proportion of bootstrap replicates supporting a node (branch). Tree length = 3320 steps, CI = 0.353, CI (excluding uninformative) = 0.318, and RI = 0.455. *C. lepida*-a, Frio River, Texas; *C. lepida*-b, Nueces River, Texas; *C. lutrensis*-a, Brazos and Pecos rivers, Texas and New Mexico, respectively; *C. lutrensis*-b, Colorado River, Texas.

it is not clear to what extent final results are dependent on the initial topology used to assess character consistency (Swofford et al., 1996). If the same characters are consistently misleading based on a variety of different starting trees, one can place reasonable confidence in the successive weighting approach. Its efficacy is based on the notion that characters reflecting the true phylogeny will support the same tree (i.e., will be additive), whereas homoplastic characters support numerous, essentially random trees. Thus, homoplastic characters, as a group, should not be consistent on any tree, making it possible to extract phylogenetic signal from noise even when historically consistent characters are in the minority (Farris, 1983).

We applied character weights implied by the ci or the retention index (ri; see Archie, 1996) on three different starting trees: (1) the equally weighted, mtDNA tree; (2) Gibbs' hypothesis (Fig. 1A); and (3) Mayden's hypothesis (Fig. 1B). Topological constraints consistent with each of the morphological studies (but including additional unresolved taxa in the case of Gibbs' hypothesis) were used to obtain starting trees. After one round of reweighting, each analysis converged on the same topology, that of the equally weighted mtDNA tree. Thus, regardless of the initial tree or index used for weights, the equal weights mtDNA tree was recovered. Furthermore, employing subjective a priori character-state weights (transversions vs transitions: 5:1, 10:1, and 1:0) resulted in tree topologies that differed little from the topology derived from equal weighting of characters. These results support the hypothesis that even when there have been numerous changes per nucleotide site leading to apparent "saturation" for some types of changes, substantial phylogenetic signal may be retained (Fitch and Ye, 1991; Yang, 1998).

Relationships among species.—Bootstrap proportions exceeding 50% are shown on the most-parsimonious mtDNA tree (Fig. 2). Of the species examined only *C. lutrensis* and *C. lepida* (see below) were not monophyletic. The traditionally recognized *Cyprinella* (exclusive of barbeled species) form a monophyletic group relative to outgroup taxa in the shortest mtDNA tree (Fig. 2). With respect to barbeled species, *C. labrosa* and *C. zanema* were deeply embedded within *Cyprinella* as sister species to *C. pyrrhomelas* and *C. callitaenia*, respectively, whereas *C. monacha* fell outside a monophyletic *Cyprinella* as sister to *Hybopsis winchelli*. These results contrast with Dimmick's (1993) hypothesis that *C. monacha* is allied with *Cyprinella*, whereas *C. labrosa* and *C. zanema* are related to *Hybopsis*, and with Jenkins and Lachner's (1971) morphologically based hypothesis that *C. labrosa* and *C. zanema* are sister taxa. A close relationship between *C. zanema* and *C. callitaenia* is consistent with Jenkins and Burkhead's (1994) suggestion that *C. zanema* and *C. labrosa* may be related to certain southeastern species with an inferior mouth, for example, *C. nivea* and *C. leedsi*. However, a close phyletic relationship between *C. labrosa* and *C. pyrrhomelas* was unexpected, because the two species share fewer morphological similarities. Because *C. labrosa* and *C. pyrrhomelas* are sympatric where they were collected [Green River (Santee River drainage), Polk County, North Carolina] hybridization may have generated a *C. labrosa*-like population with mitochondrial genomes derived from *C. pyrrhomelas*. *Cyprinella pyrrhomelas* to *C. labrosa* would be the inferred direction of introgression as the mtDNA sequences of these species were included in a

clade with *C. xaenura* (Fig. 2), a species hypothesized previously (Gibbs, 1957; Mayden, 1989) to be sister to *C. pyrrhomelas*. Hybridization among animal taxa (e.g., Harrison, 1993), and specifically among North American cyprinids, is common (Hubbs, 1955; Smith, 1992) and may result in introgression of mtDNA genomes between species (Dowling et al., 1989). A potential hybridization event between *C. labrosa* and *C. pyrrhomelas* would appear to be of some antiquity, given the 14.4% nucleotide sequence difference between them.

Among historically recognized, nonbarbeled *Cyprinella*, the mtDNA tree (Fig. 2) indicated that basal lineages in *Cyprinella* are composed largely of Gibbs' (1957) tubercles-linear species. These include all the Atlantic and/or Gulf slope endemics except *C. analostana* and *C. chloristia*, both of which have scattered tubercles. Included also in a basal position is *C. caerulea*, a Gulf slope endemic with scattered tubercles. Somewhat surprisingly, *C. callistia* occupied the basal position in the mtDNA tree. Both Gibbs (1957) and Mayden (1989) placed *C. callistia* as sister to a clade that included *C. nivea, C. leedsi, C. callitaenia*, and *C. callisema*. These latter four species (and *C. zanema*) also formed a clade in the mtDNA tree, but in the mtDNA tree, this clade was sister to *C. caerulea*, not *C. callistia*. However, the basal position of *C. callistia* in the mtDNA tree is consistent with Gibbs' (1957) observations that (1) retention of linear tubercles on the head and notal-ridge was the only character (he found) linking *C. callistia* to *Cyprinella*, and (2) *C. callistia* was ''. . . by far the most divergent form in the [sub]genus.'' Relative to bootstrap support, a monophyletic *Cyprinella* that included *C. callistia* was not supported by more than 50% of replicates.

There was strong bootstrap support (89%) for a clade of all species of *Cyprinella* except *C. callistia* (and *C. monacha*). The sister species *C. gibbsi* and *C. trichroistia* are the basal lineage within this clade and this sister relationship is consistent with their prior status as conspecifics (Howell and Williams 1971). However, placement of the two as the base of *Cyprinella* is inconsistent with Gibbs' (1957) and Mayden's (1989) hypotheses in which *C. trichroistia* (and *C. gibbsi* in Mayden's study) was embedded well within *Cyprinella* (Fig. 1).

The remaining species of *Cyprinella* fell into a series of fairly well-supported lineages, however, bootstrap support for relationships among some of these lineages was not high. In one group were six Atlantic and/or Gulf slope endemics, including four species (*C. callisema, C. callitaenia, C. leedsi*, and *C. nivea*) with linear tubercles, one (*C. caerulea*) with scattered tubercles, and one barbeled species (*C. zanema*). Bootstrap support in the mtDNA tree for the hierarchical arrangement ((*C. callisema, C. nivea*)(*C. leedsi* (*C. callitaenia, C. zanema*))) was 100% for all branches. Gibbs (1957) and Mayden (1989) concurred on monophyly of the four nonbarbeled species and on the hypothesis *C. nivea* (*C. leedsi* (*C. callisema, C. callitaenia*)). Inclusion of *C. caerulea* in the mtDNA tree as the basal member of this assemblage was not strongly supported. Gibbs (1957) placed *C. caerulea* as basal in a group that contained *C. venusta, C. spiloptera*, and the *C. lutrensis* group, in large part because *C. caerulea* has scattered tubercles and (along with *C. venusta* and *C. spiloptera*) yellow pigmentation on fins of breeding males and reduced dorsal fin pigmentation. Mayden (1989), alternatively, placed *C. caerulea* as sister to the *C. gibbsi/C. trichroistia* species pair. Because of low bootstrap support, *C. caerulea* is likely best considered for now as a single lineage within the large clade that is sister to the *C. gibbsi/C. trichroistia* species pair. However, the strong bootstrap support in the mtDNA tree for clades that include *C. venusta, C. spiloptera*, and the *C. lutrensis* group (see below) argue against Gibbs' hypothesis regarding placement of *C. caerulea*.

A second clade (i.e., sister to the one containing *C. caerulea, C. nivea*, and others) contained a number of mtDNA lineages with fairly strong bootstrap support. Basal in this large clade was an assemblage with *C. xaenura* as sister to the species pair *C. pyrrhomelas* and *C. labrosa*. All three are Atlantic slope endemics: *C. pyrrhomelas* and *C. xaenura* are tubercles-linear species at the base of Gibbs' (1957) tree; whereas *C. labrosa* is one of the three barbeled *Cyprinella*. Mayden (1989) also found *C. pyrrhomelas* and *C. xaenura* to be sister species with this group sister to a clade that included *C. caerulea, C. gibbsi*, and *C. trichroistia*. Two additional assemblages within this second clade were (1) the species pair *C. analostana* and *C. chloristia*, and (2) a clade that included *C. proserpina* as sister to the species pair *C. camura* and *C. galactura*. Both Gibbs (1957) and Mayden (1989) placed *C. analostana* as sister to *C. chloristia*. Based on morphological similarities, Gibbs (1957, 1963) hypothesized *C. whipplei* as sister to this species pair, placing this larger clade within his derived, tubercles-scattered group. Mayden (1989) also found *C. whipplei* to be the sister of the *C. analostana/C. chloristia* species pair. The sister relationship between *C. camura* and *C. galactura* in the mtDNA tree (supported by 100% of bootstrap replicates) is consistent with Gibbs (1957, 1961) but not with Mayden (1989) who placed *C. camura*

as basal to the *C. whipplei* (*C. analostana, C. chloristia*) clade and *C. galactura* as sister to *C. venusta*. The relationship of *C. proserpina* with the *C. camura/C. galactura* species pair in the mtDNA tree was unexpected and is discussed more fully below.

A derived clade in the mtDNA tree included many of the cosmopolitan, tubercles-scattered species distributed primarily in the Mississippi River basin and the American Southwest. Relationships of species within this clade generally had strong support and are similar to Gibbs' (1957) hypothesis where *C. venusta* is sister to a clade of *C. spiloptera* and the *C. lutrensis* group. Alternatively, according to Mayden (1989), the *C. lutrensis* group (including *C. proserpina*) is sister to all other *Cyprinella*, with *C. spiloptera* as the basal lineage in his *C. whipplei* clade. As noted above, both Gibbs (1957) and Mayden (1989) placed *C. whipplei* as sister to the *C. analostana/C. chloristia* species pair. However the sister relationship of *C. whipplei* and *C. spiloptera* in the mtDNA tree is consistent with their earlier consideration as conspecifics (Hubbs and Greene, 1928). Furthermore, analysis of over 700 bp from the mtDNA control region from several allopatric populations did not find these species to be reciprocally monophyletic (Broughton, 1995).

Placement of *C. proserpina* outside of the *C. lutrensis* group is consistent with Gold and Richardson (1999) who estimated 17.2% sequence divergence between *C. proserpina* and the *C. lutrensis* group but 10.5% divergence within the *C. lutrensis* group. Morphological similarity and geographic distribution have been used to support the hypothesis that *C. proserpina* belongs to the *C. lutrensis* group (Hubbs and Miller, 1978), where it was hypothesized by Mayden (1989) to be sister to *C. lepida*.

The terminal branch leading to *C. proserpina* in the mtDNA tree (Fig. 2) is atypically long (177 inferred nucleotide changes), suggesting that mtDNA in *C. proserpina* may have evolved unusually rapidly. We evaluated evolutionary rate variation with Tajima's (1993) relative rate test on 10 species representing major ingroup clades in the mtDNA phylogeny. The included species were *C. callistia, C. trichroistia, C. nivea, C. xaenura, C. analostana, C. proserpina, C. camura, C. venusta, C. whipplei*, and *C. formosa*. All nine pairwise comparisons involving *C. proserpina* differed significantly (Bonferroni corrected for multiple tests) from rate homogeneity. By comparison, only two of the other 36 comparisons (*C. whipplei* vs *C. callistia* and *C. trichroistia*) were significant. Because large differences in the number of changes per branch increase the opportunity for homoplasy to obscure historical phylogenetic signal (Kuhner and Felsenstein, 1994; Lockhart et al., 1994), we view the phylogenetic placement of *C. proserpina* in the mtDNA tree as tentative. However, a maximum likelihood analysis accounting for rate heterogeneity (PAUP*, HKY model, 4-class discrete gamma model of among-site rate variation) recovered a tree (not shown) consistent with the parsimony analysis, including the position of *C. proserpina*. Moreover, a uniquely shared chromosomal nucleolus organizer region (NOR) character state supports the mtDNA hypothesis that *C. proserpina* is related to *C. camura* and *C. galactura* (Amemiya and Gold, 1990; Amemiya et al., 1992).

Cyprinella from the American Southwest (except *C. proserpina*) formed a monophyletic group in the mtDNA tree, however, within this group, neither *C. lepida* nor *C. lutrensis* were monophyletic (Fig. 2). These results are consistent with the complex and confusing taxonomic history of the species group (Chernoff and Miller, 1982; Matthews, 1987). As an example, both *C. formosa* (Contreras-Balderas, 1975; Chernoff and Miller, 1982) and *C. lepida* (Hubbs, 1956, 1972: Matthews 1987) have at times been considered to be distinct species or subspecies of *C. lutrensis*. Previous mtDNA studies (Richardson and Gold, 1995, 1999) indicated that *C. lepida* in the Frio and Nueces Rivers were specifically distinct and that *C. lepida* from the Frio River was more closely related to *C. formosa* than to *C. lepida* from the Nueces River. Hybridization among these taxa may have generated the conflicting mtDNA haplotype relations, but the absence of a clear geographic pattern relating mtDNA haplotypes could suggest multiple (and perhaps undecipherable) hybridization events.

In general, the most-parsimonious mtDNA tree (Fig. 2) appears to be more similar to the hypothesis of Gibbs (1957; Fig. 1A) than to the hypothesis of Mayden (1989; Fig. 1B). Much of the apparent similarity derives from placement of the tubercles-linear, Atlantic and/or Gulf slope endemics as basal in the *Cyprinella* tree. We evaluated the degree of similarity among phylogenetic hypotheses by considering only the 17 species common to all three studies and estimating the number of steps required for the most-parsimonious, unrooted network for mtDNA (1843 steps), and for Gibbs' hypothesis (2098 steps) and Mayden's hypothesis (2225 steps) when topological constraints corresponding to the two morphological studies were enforced on the mtDNA data. Gibbs' and Mayden's hypotheses were 13.8% and 20.7% longer, respectively, than the shortest mtDNA tree. A

nonparametric Wilcoxon signed-ranks test (Templeton, 1983) indicated that both of the morphologically based topologies were significantly longer than the most-parsimonious mtDNA tree ($P < 0.0001$, each). However, the greater similarity of Gibbs' hypothesis to the mtDNA tree appears to be based on the polarity of the tree (southeastern endemics basal and Mississippi basin/Southwestern forms derived). Differences among the three phylogenetic hypotheses may be related, in part, to interpretation of morphological characters (exemplified by differences in the two morphological studies) and by possible differences in evolutionary rates among morphological and mtDNA characters. Ultimately, more nucleotide characters from additional genes may aid in recovering more robust support for all clades in an hypothesis of phylogenetic relationships of *Cyprinella*.

Material Examined

Voucher specimens are deposited in the Texas Cooperative Wildlife Collection (TCWC) at Texas A&M University. Collection localities (drainage) of material examined are as follows: *Cyprinella spiloptera*, Big River, Washington Co., Missouri (Meramec River), Stony Creek, Vermillion Co., Illinois (Wabash River), Susquehanna River, Tioga Co., New York; *C. whipplei*, Big Darby Creek, Pickaway Co., Ohio (Scioto River), Hocking River, Athens Co., Ohio (Ohio River), Ouachita River, Montgomery Co., Arkansas; *C. galactura*, Choats Creek, Giles Co., Tennessee (Tennessee River), Holston River, Scott Co., Virginia (Tennessee River); *C. venusta*, Cahaba River, Bibb Co., Alabama, Llano River, Kimble Co., Texas (Colorado River), TCWC: 7283.01; *C. camura*, Thompson Creek, West Feliciana Par., Louisiana (Mississippi River), Chickaskia River, Kingman Co., Kansas (Arkansas River); *C. callistia*, Cahaba River, Bibb Co., Alabama, Conasauga River, Bradley Co., Tennessee (Coosa River); *C. leedsi*, Alpaha River, Echolz Co., Georgia (Suwanee River); *C. pyrrhomelas*, Green River, Polk Co. North Carolina (Santee River), TCWC: 7982.02; *C. analostana*, Eno River, Durham Co., North Carolina (Nuese River), White Oak Creek, Johnston Co., North Carolina (Nuese River), TCWC: 7985.01; *C. gibbsi*, Tallapoosa River, Randolph Co., Alabama; *C. chloristia*, Broad River, Rutherford Co., North Carolina (Santee River), TCWC: 7984.01; *C. nivea*, Broad River, Maddison/Elbert Co., Georgia (Savannah River), TCWC: 8292.01, Deep River, Lee/Chatham Co., North Carolina (Cape Fear River), TCWC: 7981.01; *C. xaenura*, Middle Oconee River, Oconee/Clarke Co., Georgia (Altamaha River), TCWC: 8294.01; *C. callisema*, Middle Oconee River, Oconee/Clarke Co., Georgia (Altamaha River), TCWC: 8294.02; *C. callitaenia*, Flint River, Meriwether Co., Georgia (Apalachicola River), TCWC: 7986.02; *C. trichroistia*, Little Cahaba River, Bibb Co., Alabama (Cahaba River), Conasaugua River, Bradley Co., Tennessee (Coosa River); *C. caerulea*, Conasaugua River, Bradley Co., Tennessee (Coosa River), Little River, Cherokee Co., Alabama (Coosa River); *C. lutrensis*, Brazos River, Milam Co., Texas, Colorado River, Concho Co., Texas, Pecos River, De Baca Co., New Mexico; *C. proserpina*, Devil's River, Val Verde Co., Texas, TCWC: 6988.01; *C. lepida*, Frio River, Bandera Co., Texas, TCWC: 7264.01, Nueces River, Edwards/Real Co., Texas, TCWC: 7267.01; *C. formosa*, Dexter National Fish Hatchery, TCWC: 6622.04; *C. monacha*, Buffalo River, Lewis Co., Tennessee (Tennessee River); *C. zanema*, Second Broad River, Rutherford Co., North Carolina (Santee River), TCWC: 7983.01; *C. labrosa*, Green River, Polk Co., North Carolina (Santee River), TCWC: 7982.01; *Hybopsis winchelli*, Cahaba River, Bibb Co., Alabama; *Luxilus cornutus*, Kalamazoo River, Calhoun Co., Michigan; *Lythrurus roseipinnis*, Chickashay River, Clarke Co., Mississippi (Pascagoula River), TCWC: 7658.01; *Notropis atherinoides*, Brazos River, Haskell Co., Texas.

Acknowledgments

We thank J. Bielawski, D. Billings, T. Dowling, D. Etnier, B. Freeman, T. Frizzell, B. Jensen, M. Howell, L. Richardson, M. Sabaj, and W. Starnes for providing, or assistance collecting, specimens. Helpful discussions and/or comments on the manuscript were provided by M. Coburn, D. Etnier, M. Howell, R. Jenkins, W. Matthews, L. Richardson, and W. Starnes. We thank R. Ford and L. Stewart for performing some of the sequencing reactions. All collections were made under valid scientific collecting permits issued by the states where permits were required. Specimens of *Cyprinella monacha* (federally threatened) were provided by the Illinois Natural History Survey and collected under Tennessee Wildlife Resources Agency Scientific Study Permit No. 809 issued to L. Page. Specimens of *Cyprinella caerulea* were collected in 1985 and 1988, prior to being listed (federally threatened) in 1992. Specimens of *Cyprinella formosa* were provided by the Dexter National Fish Hatchery, Dexter, New Mexico. Funding was provided in part by the National Science Foundation (BSR-90-20217), and in part by the Texas Agricultural Experiment Station (H-6703). Much of the work was carried out in the Center for Biosystematics

and Biodiversity at Texas A&M University, a facility funded, in part, by the National Science Foundation (DIR-89-07006). The manuscript was prepared while REB was a NSF/Sloan Postdoctoral Research Fellow in Molecular Evolution in the laboratory of R. Harrison at Cornell University. This paper is contribution number 93 of the Center for Biosystematics and Biodiversity at Texas A&M University.

Literature Cited

Amemiya, C. T., and J. R. Gold. 1990. Cytogenetic studies in North American minnows (Cyprinidae). XVII. Chromosomal NOR phenotypes of 12 species, with comments on cytosystematic relationships among 50 species. Hereditas 112:231–247.

———, P. K. Powers, and J. R. Gold. 1992. Chromosomal evolution in the North American cyprinids, p. 534–550. *In:* Systematics, historical ecology, and North American freshwater fishes. R. L. Mayden (ed.). Stanford Univ. Press, Stanford, CA.

Archie, J. W. 1996. Measures of homoplasy, p. 153–188. *In:* Homoplasy. M. Sanderson and L. Hufford (eds.). Academic Press, New York.

Bailey, R. M. 1951. A check-list of the fishes of Iowa, with keys for identification, p. 327–377. *In:* Iowa fish and fishing. J. Harlan and E. Speaker (eds.). Iowa Conservation Commission, Des Moines, IA.

———, and R. H. Gibbs Jr. 1956. *Notropis callistius*, a new cyprinid fish from Alabama, Florida, and Georgia. Occ. Pap. Mus. Zool., Univ. Mich. 576:1–14.

Bielawski, J. P., and J. R. Gold. 1996. Unequal synonymous substitution rates within and between two protein-coding mitochondrial genes. Mol. Biol. Evol. 13:889–92.

Broughton, R. E. 1995. Evolution of duplicated sequences in mitochondrial DNA. Unpubl. Ph.D. diss., Arizona State Univ., Tempe.

Buth, D. G. 1979. Biochemical systematics of the cyprinid genus Notropis. I. The subgenus *Luxilus*. Biochem. Syst. Ecol. 7:69–79.

———, T. E. Dowling, and J. R. Gold. 1991. Molecular and cytological investigations, p. 83–126. *In:* The biology of cyprinid fishes. I. Winfield, and J. Nelson (eds.). Chapman and Hall, London.

Chernoff, B., and R. R. Miller. 1982. *Notropis bocagrande*, a new cyprinid fish from Chihuahua, Mexico, with comments on *Notropis formosus*. Copeia 1982:514–522.

Coburn, M. M., and T. M. Cavender. 1992. Interrelationships of North American cyprinid fishes, p. 328–373. *In:* Systematics, historical ecology, and North American freshwater fishes. R. Mayden (ed.). Stanford Univ. Press, Stanford, CA.

Collins, T. M., P. H. Wimberger, and G. J. P. Naylor. 1994. Compositional bias, character-state bias, and character-state reconstruction using parsimony. Syst. Biol. 43:482–496.

Contreras-Balderas, S. 1975. Zoogeography and evolution of *Notropis lutrensis* and "*Notropis*" *ornatus* in the Rio Grande basin and range, Mexico and United States (Pisces: Cyprinidae). Unpbl. Ph.D. diss., Tulane Univ., New Orleans, LA.

Dimmick, W. W. 1993. A molecular perspective on the phylogenetic relationships of the barbeled minnows, historically assigned to the genus *Hybopsis* (Cyprinidae: Cypriniformes). Mol. Phylogenet. Evol. 2:173–184.

Dowling, T. E., and G. J. P. Naylor. 1997. Evolutionary relationships of minnows in the genus *Luxilus* (Teleostei: Cyprinidae) as determined from cytochrome *b* sequences. Copeia 1997:758–765.

———, G. R. Smith, and W. M. Brown. 1989. Reproductive isolation and introgression between *Notropis cornutus* and *Notropis chrysocephalus* (family Cyprinidae): comparison of morphology, allozymes, and mitochondrial DNA. Evolution 43:620–634.

Farris, J. S. 1969. A successive approximations approach to character weighting. Syst. Zool. 18:374–385.

———. 1983. The logical basis of phylogenetic analysis, p. 7–36. *In:* Advances in cladistics. Vol. 2. N. Platnick and V. Funk (eds.). Columbia Univ. Press, New York.

———, M. Kallersjo, A. G. Kluge, and C. Bult. 1995. Testing significance of incongruence. Cladistics 10:315–319.

Fitch, W. M., and J. Ye. 1991. Weighted parsimony: does it work? p. 147–154. *In:* Phylogenetic analysis of DNA sequences. M. Miyamoto and J. Cracraft (eds.). Oxford Univ. Press, New York.

Gibbs, R. H. 1957. Cyprinid fishes of the subgenus *Cyprinella* of *Notropis* I. Systematic status of the subgenus *Cyprinella*, with a key to the species exclusive of the *lutrensis-ornatus* complex. Copeia 1957:185–195.

———. 1961. Cyprinid fishes of the subgenus *Cyprinella* of *Notropis*. IV. The *Notropis galacturus-camurus* complex. Am. Midl. Nat. 66:337–354.

———. 1963. Cyprinid fishes of the subgenus *Cyprinella* of *Notropis*. The *Notropis whipplei-analostanus* complex. Copeia 1963:511–528.

Gilbert, C. R. 1964. The American cyprinid fishes of the subgenus *Luxilus* (genus *Notropis*). Bull. Fla. State Mus., Biol. Sci. 8:95–194.

Girard, C. 1856. Researches upon the cyprinoid fishes inhabiting the freshwaters of the United States of America, west of the Mississippi Valley, from specimens in the museum of the Smithsonian Institution. Proc. Acad. Nat. Sci., Phila. 8:165–213.

Harrison, R. G. 1993. Hybrids and hybrid zones: historical perspective, p. 3–12. *In:* Hybrid zones and the evolutionary process. R. G. Harrison (ed.). Oxford Univ. Press, New York.

Howell, W. M., and J. D. Williams. 1971. *Notropis gibbsi*, a new cyprinid fish from the Tallapoosa River system in Alabama and Georgia. Copeia 1971:55–64.

Hubbs, C. 1956. Relative variability of hybrids between the minnows, *Notropis lepidus* and *N. proserpinus*. Tex. J. Sci. 8:463–469.

———. 1972. A checklist of Texas freshwater fishes. Texas Parks Wildlife Dept., Tech. Ser. 11:1–11.

Hubbs, C. L. 1955. Hybridization between fish species in nature. Syst. Zool. 4:1–20.

———, and C. W. Greene. 1928. Further notes on

the fishes of the Great Lakes and tributary waters. Pap. Mich. Acad. Sci. Arts Lett. 8:371–392.

———, AND R. R. MILLER. 1978. *Notropis panarcys*, n. sp., and *N. proserpinus*, cyprinid fishes of subgenus *Cyprinella*, each inhabiting a discrete section of the Rio Grande complex. Copeia 1978:582–592.

JENKINS, R. E., AND N. M. BURKHEAD. 1994. Freshwater fishes of Virginia. American Fisheries Society, Bethesda, MD.

———, AND E. A. LACHNER. 1971. Criteria for analysis and interpretation of the American fish genera *Nocomis* Girard and *Hybopsis* Agassiz. Smiths. Contrib. Zool. 90:1–15.

———, AND ———. 1980. Account of *Hybopsis zanema*, p. 197. *In*: Atlas of North American freshwater fishes. D. S. Lee, C. R. Gilbert, C. H. Hocutt, R. E. Jenkins, D. E. McAllister, J. R. Stauffer Jr. (eds.). North Carolina State Musuem of Natural History, Raleigh.

KLUGE, A. G., AND J. S. FARRIS. 1969. Quantitative phyletics and the evolution of anurans. Syst. Zool. 18: 1–32.

KOCHER, T. D., AND K. L. CARLETON. 1997. Base substitution in fish mitochondrial DNA: patterns and rates, p. 13–24. *In:* Molecular systematics of fishes. T. Kocher and C. Stepien (eds.). Academic Press, New York.

KRISTMUNDSDÓTTIR, Á. Ý., AND J. R. GOLD. 1996. Systematics of blacktail shiners (*Cyprinella venusta*) inferred from analysis of mitochondrial DNA. Copeia 1996:773–783.

KUHNER, M. K., AND J. FELSENSTEIN. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Mol. Biol. Evol. 11:459–68.

LOCKHART, J. P. , M. A. STEELE , M. D. HENDY, AND D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Ibid.* 11:605–612.

MARTIN, A. P., AND E. BERMINGHAM. 1998. Systematics and evolution of lower Central American cichlids inferred from analysis of cytochrome *b* gene sequences. Mol. Phylogenet. Evol. 9:192–203.

MATTHEWS, W. J. 1987. Geographic variation in *Cyprinella lutrensis* (Pisces: Cyprinidae) in the United States, with notes on *Cyprinella lepida*. Copeia 1987: 616–637.

MAYDEN, R. L. 1988. Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. Syst. Zool. 37:329–355.

———. 1989. Phylogenetic studies of North American minnows, with emphasis on the genus *Cyprinella* (Teleostei: Cypriniformes). Univ. Kans., Mus. Nat. Hist., Misc. Publ. 80:1–189.

MINDELL, D. P., AND C. E. THACKER. 1996. Rates of molecular evolution: phylogenetic issues and applications. Annu. Rev. Ecol. Syst. 27:279–303.

ROBINS, C. R., R. M. BAILEY, C. E. BOND, J. R. BROOKER, E. A. LACHNER, R. N. LEA, AND W. B. SCOTT (EDS.). 1991. Common and scientific names of fishes from the United States and Canada. American Fisheries Society, Spec. Publ. 20, Bethesda, MD.

RICHARDSON, L. R., AND J. R. GOLD. 1995. Evolution of the *Cyprinella lutrensis* species-complex. II. Systematics and biogeography of the Edwards Plateau shiner, *Cyprinella lepida*. Copeia 1995:28–37.

———, AND ———. 1999. Systematics of *Cyprinella lutrensis* species-group species (Cyprinidae) from the southwestern United States, as inferred from restriction-site variation of mitochondrial DNA. Southwest. Nat. 44:49–56.

SCHMIDT, T. R., J. P. BIELAWSKI, AND J. R. GOLD. 1998. Molecular phylogenetics and evolution of the cytochrome *b* gene in the cyprinid genus *Lythrurus* (Actinopterygii: Cypriniformes). Copeia 1998:14–22.

SIMON, C., F. FRATI, A. BECKENBACH, B. CRESPI, H. LIU, AND P. FLOOK. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87:651–701.

SIMONS, A. M., AND R. L. MAYDEN. 1998. Phylogenetic relationships of the western North American phoxinins (Actinopterygii: Cyprinidae) as inferred from mitochondrial 12S and 16S ribosomal RNA sequences. Mol. Phylogenet. Evol. 9:308–329.

SMITH, G. R. 1992. Introgression in fishes: significance for paleontology, cladistics, and evolutionary rates. Syst. Biol. 41:41–57.

SNELSON, F. F. 1972. Systematics of the subgenus *Lythrurus*, genus *Notropis* (Pisces: Cyprinidae). Bull. Fla. State Mus., Biol. Sci. 17:1–92.

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, AND D. M. HILLIS. 1996. Phylogenetic inference, p. 407–514. *In:* Molecular systematics. D. M. Hillis, C. Moritz, and B. Mable (eds.). Sinauer Associates, Sunderland, MA.

TAJIMA, F. 1993. Simple methods for testing the molecular evolutionary clock hypothesis. Genetics 135: 599–607.

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.

YANG, Z. 1998. On the best evolutionary rate for phylogenetic analysis. Syst. Biol. 47:125–133.

ZARDOYA, R., AND A. MEYER. 1996. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. Mol. Biol. Evol. 13:933–942.

CENTER FOR BIOSYSTEMATICS AND BIODIVERSITY, TEXAS A&M UNIVERSITY, COLLEGE STATION, TEXAS 77843-2258. PRESENT ADDRESS: (REB) OKLAHOMA BIOLOGICAL SURVEY AND DEPARTMENT OF ZOOLOGY, UNIVERSITY OF OKLAHOMA, NORMAN, OKLAHOMA 73019-0235. E-mail: (REB) rbroughton@ou.edu. Send reprint requests to REB. Submitted: 24 Feb. 1999. Accepted: 9 July 1999. Section editor: M. E. Douglas.