Journal of Fish Biology (2002) **61,** 293–297
doi:10.1006/jfbi.2002.2002, available online at http://www.idealibrary.com on IDEAL®

BRIEF COMMUNICATION

Phylogenetic relationships within the genus *Pimephales* as inferred from ND4 and ND4L nucleotide sequences

J. P. Bielawski*, A. Brault and J. R. Gold

Center for Biosystematics and Biodiversity, Texas A&M University, College Station, TX 77843-2258, U.S.A.

(*Received 24 October 2001, Accepted 13 May 2002*)

In phylogenetic analyses, conducted on the ND4L gene and part of the ND4 gene from species of the genus *Pimephales*, maximum parsimony yielded four trees, with the strict consensus providing no resolution of relationships among species. Maximum likelihood and minimum evolution methods yielded identical tree topologies, which differed from previous hypotheses of relationships for these species. If this topology is correct, it implies independent evolution of morphological characters, possibly associated with convergent trophic specialization.

Key words: *Pimephales*; phylogenetics; mtDNA; convergence; morphology.

The North American cyprinid genus *Pimephales* is comprised of four extant species: *P. notatus* (Rafinesque), *P. promelas* (Rafinesque), *P. tenellus* (Girard) and *P. vigilax* (Baird & Girard). Three of these (*P. notatus*, *P. promelas* and *P. vigilax*) are widespread, occupying a variety of habitats in drainages east of the Rocky Mountains. The fourth species, *P. tenellus*, is restricted to highlands in Missouri, Kansas, Oklahoma and Arkansas, preferring sand and gravel bottomed streams and small rivers. Species of *Pimephales* are small (30–100 mm total length, L_T), round bodied, and typically inhabit demersal freshwater habitats. The females lay a single layer of eggs on the undersides of rocks, and males develop three rows of prominent breeding tubercles on the snout and participate in post-fertilization egg-guarding (Mayden, 1987; Johnston, 1999).

Monophyly of the four species of *Pimephales* is supported by morphological (Mayden, 1987; Coburn & Cavender, 1992) and chromosomal (Li & Gold, 1991) characters; relationships among the four species, however, are not well resolved. Mayden (1987) conducted a small-scale cladistic analysis of morphological characters and suggested that *P. tenellus* was sister to *P. vigilax* and that *P. notatus* was sister to *P. promelas*. Based on analysis of variation in chromosomal nucleolar organizer regions (NORs), Li & Gold (1991) hypothesized that *P. notatus* was sister to *P. vigilax*. Li & Gold (1991) could not infer relationships of *P. promelas* and *P. tenellus*, as both species possessed the hypothesized primitive NOR character state. Schmidt *et al.* (1994) attempted to resolve this conflict by analysing restriction site variation of mitochondrial (mt)DNA. Unfortunately, results of their analysis were sensitive to analytical method; a topology based on Dollo parsimony suggested the same relationships as previously inferred from morphology (Mayden, 1987), whereas a topology based upon a distance method indicated the novel set of relationships (*P. promelas*, (*P. notatus*, (*P. tenellus*, *P. vigilax*))) (Schmidt *et al.*, 1994). As bootstrap support was not strong for either hypothesis of

*Author to whom correspondence should be addressed at present address: Department of Biology, University College London, Darwin Building, Gower Street, London, WC1E 6BT, U.K. Tel.: +44 (0) 20 7679 4388; fax: +44 (0) 20 7679 2887; email: j.bielawski@ucl.ac.uk

0022–1112/02/070293+05 $35.00/0

TABLE I. Collection localities and drainages of species

Species	Collection locality[a]	State	Drainage (river)
Pimephales notatus	Middle Creek	Kansas	Missouri
P. notatus	Saline River	Michigan	Raisin
P. promelas	South Platt River	Nebraska	Platte
P. promelas	Tributary of Rouge River	Michigan	Rouge
P. tenellus	Neosho River	Kansas	Neosho
P. tenellus	Salt Creek	Oklahoma	Arkansas
P. vigilax vigilax	Brazos River	Texas	Brazos
P. v. perspicuus	Embarras River	Illinois	Wabash
Opsopoeodus emiliae	Navasota River	Texas	Brazos

[a]Where required, all collections were made under valid scientific collecting permits.

relationships, Schmidt *et al.* (1994) concluded that none of the hypotheses previously proposed could be disregarded unequivocally.

In this study, a new dataset was analysed comprised of protein-coding mtDNA sequences. ND4–ND4L sequences were acquired from nine individuals of *Pimephales* and one individual of *Opsopoeodus emiliae* (Hay); the latter served as outgroup in phylogenetic analysis (Cavender & Coburn, 1986; Coburn & Cavender, 1992). Collection localities are given in Table I. For each specimen, a 500-bp DNA fragment that included the complete ND4L gene (297 bp) and 210 base pairs of the adjacent ND4 gene (ND4–ND4L fragment) was sequenced. Sequences of ND4 and ND4L do not sum to 507 because reading frames of these two genes overlap by seven base pairs. Procedures for DNA extraction, PCR amplification and DNA sequencing followed those in Schmidt *et al.* (1998). PCR amplifications and sequencing reactions employed the universal primer NAP2 (Hogan *et al.*, 1997) and primers ARGBL (5′-CAAGACCCTTGATTTC GGCTCA-3′), ND4LB (5′-CAAAACCTTAATCTYCTACAATGCT-3′), and LEUAH (5′-CAAGAGTTTCAGGCTCCTAAGAACA-3′). DNA sequences were determined from a minimum of two independent sequencing-reactions for each primer. GenBank accession numbers for sequences in this paper are AY102287–AY102306.

Phylogenetic analyses were conducted using maximum parsimony (MP), maximum likelihood (ML) and minimum evolution (ME), as implemented in PAUP* ver. 4.0b4a (Swofford, 2000). Tree searches were carried out under MP (unweighted) by using the branch-and-bound algorithm. Tree searches were carried out under ME and ML by using TBR branch swapping. ME analysis was based on LogDet (Lockhart *et al.*, 1994) genetic distances. ML analysis assumed a GTR substitution matrix (Yang, 1994*a*), in combination with a discrete Gamma (dΓ) model of among-sites rate variation (Yang, 1994*b*). This model was selected according to a procedure described in Bielawski & Gold (2001) that is based on the likelihood ratio test. For all substitution matrices tested, the dΓ model of among sites rate variation (ASRV) was optimal. The optimal substitution matrix was sensitive to the assumed model of ASRV, however, under the dΓ model the GTR matrix was optimal. Model selection was conducted with the tree as a free parameter, and the same topology was optimal under all models tested. Relative support for individual nodes of a topology was assessed by non-parametric bootstrapping (Felsenstein, 1985; Penny & Hendy, 1985). Bootstrap proportions (P_B) were based on 1000 (MP and ME) and 100 (ML) pseudoreplications. Under ML, bootstrap analysis was conducted by specifying substitution parameters estimated from the ML topology as fixed-model parameters (Kishino *et al.*, 1990). The molecular clock was tested by using a likelihood ratio test (Yang *et al.*, 1995) and the ML estimate of the topology.

Patterns of sequence variation followed those expected for mtDNA protein-coding sequences. Sequence divergence was greatest at third codon positions (25 ± 7·7%; mean ± S.D.), followed by first and second codon positions (8·8 ± 3·2% and 2·8 ± 1·3%,

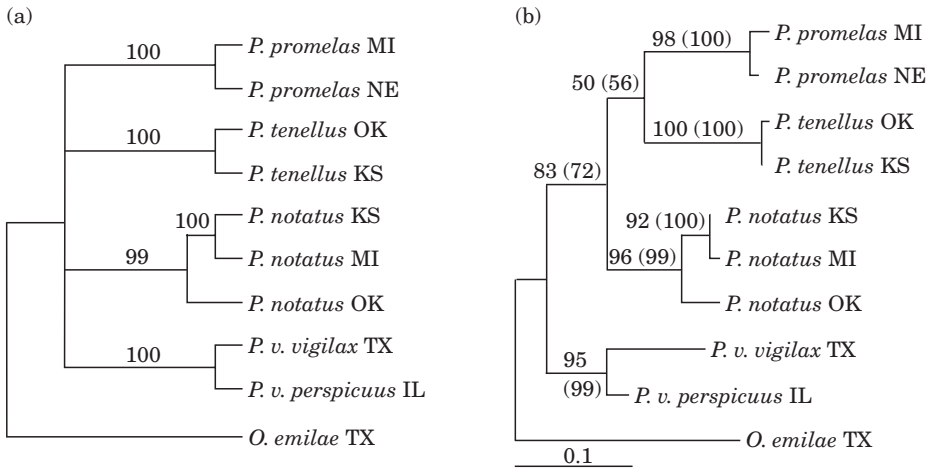

FIG. 1. Estimates of the *Pimephales* phylogeny derived from analysis of mitochondrially-encoded ND4-ND4L sequences. (a) Strict consensus of four equally parsimonious (MP) trees (length = 244 steps). Numbers at nodes are bootstrap proportions. (b) Tree topology derived from both maximum-likelihood (ML) analysis ($\ell = -1728{\cdot}04$) under the GTR substitution matrix (Yang, 1994*a*) in combination with a discrete Gamma (dΓ) model of among sites rate variation (Yang, 1994*b*), and minimum-evolution (ME) analysis (tree score = 0.4936) of LogDet distances (Lockhart *et al.*, 1994). Branch lengths are proportional to mean number of nucleotide substitutions per site, as inferred under the GTR + dΓ model. Numbers at nodes are bootstrap proportions for ML and, in parentheses, ME. IL, Illinois; KS, Kansas; MI, Michigan; NE, Nebraska; OK, Oklahoma; TX, Texas.

respectively). Nucleotide sequence variation within *Pimephales* was predominantly (70%) at synonymous sites. Average amino acid and nucleotide sequence divergence among species of *Pimephales* was $7{\cdot}5 \pm 2{\cdot}7\%$ and $11{\cdot}3 \pm 0{\cdot}4\%$, respectively. Although nucleotide composition exhibited bias typical for mitochondrial genes (unpubl. data), the bias was homogenous among all sampled lineages: first positions, $\chi^2 = 1{\cdot}85$, d.f. = 27, $P > 0{\cdot}05$; second positions, $\chi^2 = 1{\cdot}15$, d.f. = 27, $P > 0{\cdot}05$; and third positions, $\chi^2 = 15{\cdot}24$, d.f. = 27, $P > 0{\cdot}05$. A likelihood ratio test of the molecular clock indicated homogeneity of rates among lineages ($2\delta = 11{\cdot}82$, d.f. = 8, $P > 0{\cdot}05$).

There are two subspecies of *P. vigilax* (Hubbs & Black, 1947): *P. v. vigilax*, distributed from the Trinity River in Texas to drainages of the Rio Grande system in Mexico, and *P. v. perspicuus*, distributed widely in the Mississippi River basin and in Gulf Slope drainages of Georgia and Alabama. Interestingly, ND4–ND4L sequence divergence between these subspecies (7%) was similar to, or greater than, that for several other pairs of cyprinids recognized as full species (Schmidt *et al.*, 1998; Broughton & Gold, 2000; Bielawski & Gold, 2001). This finding is consistent with patterns of sequence divergence inferred from mtDNA restriction site variation (Schmidt *et al.*, 1994), and reinforces the suggestion (Schmidt *et al.*, 1994) that morphological and ecological variation within *P. vigilax* warrants additional examination, as separate species might exist.

Maximum-parsimony analysis generated four equally parsimonious trees (length = 244, CI = 0·73, RI = 0·72) [Fig. 1(a)]. Although monophyly of each species was strongly supported, there was no resolution of relationships among species. ML and ME analyses, alternatively, recovered the same resolved topology, where P_B values were $\geq 50\%$ for all internodes [Fig. 1(b)]. Similar to MP analysis, values of P_B under ML and ME indicated strong support for monophyly of each species. The ML–ME topology, however, differed from previous topologies in suggesting a sister relationship between *P. promelas* and *P. tenellus*, a sister relationship between *P. notatus* and the *P. promelas*–*P. tenellus* clade, and placement of *P. vigilax* as the basal lineage within *Pimephales*. Strongest support (P_B) was for the basal placement of *P. vigilax*. The ML–ME topology

differs from that of Mayden (1987) and Schmidt *et al.* (1994), who placed *P. vigilax* as sister to *P. tenellus* on the basis of morphology and mtDNA restriction-site variation, respectively; and from that of Li & Gold (1991), who placed *P. vigilax* as sister to *P. notatus* on the basis of a chromosomal NOR character state. Placement of *P. promelas* as sister to *P. tenellus* was not strongly supported (50–56%) by bootstrap analysis of ND4–ND4L sequences [Fig. 1(b)], and differed from Mayden (1987), who placed *P. promelas* as sister to *P. notatus*, and from Schmidt *et al.* (1994), who placed *P. promelas* as the basal lineage within *Pimephales*. The low bootstrap support means that alternative hypotheses of relationships cannot be disqualified. It seems likely that a lack of a consensus opinion of relationships for species of *Pimephales* will remain without collection and anlysis of additional data.

Interestingly, the ML–ME tree implies that the single morphological character (shape of the pharyngeal pad of the basioccipital) supporting a sister relationship between *P. promelas* and *P. notatus* (Mayden, 1987) either evolved independently in each lineage or evolved once at the base of the *P. notatus–P. tenellus–P promelas* clade and then was lost in *P. tenellus*. This suggestion is not unwarranted, as this character exhibits homoplasy in other cyprinids (Mayden, 1989; Coburn & Cavender, 1992; Schmidt, 1994). It is important to note that this finding for *Pimephales* depends on resolution of a branch with bootstrap support of only 50–56% (ML–ME). More remarkable is that the ML–ME topology also implies that three characters (broad and notched anterior wing of the hyomandibular, high and broad posterior process on the metapterygoid and flattened dentary) evolved independently in *P. tenellus* and *P. vigilax*. These three skeletal elements, however, are not independent, as the posterior ascending process of the metapterygiod articulates with the hyomandibular, and all three elements (hyomandibular, metapterygiod and dentary) are associated with the adductor mandibulae, a large muscle coupled with mandibular movement. Again, the pattern homoplasy for this suite of three characters is dependent on the branch supporting a sister relationship between *P. promelas* and *P. tenellus* (P_B=50–56%).

The broad anterior wing of the hyomandibula is associated with a broadened insertion of muscles required for the lateral movement of the opercula (Coburn, 1982), and lateral movement of the opercula is an important specialization required by species with suction-feeding capabilities (Sibbing, 1991; Gerking, 1994). If the ML–ME tree is correct, replicated patterns of morphological evolution could reflect similar functional requirements in both *P. tenellus* and *P. vigilax* for oral food acquisition and food processing. Such a situation is not unusual among teleosts, as replicated patterns of trophic adaptation have led to conflict between morphological an molecular hypotheses for North American cyprinids (Bielawski & Gold, 2001; J. P. Bielawski & J. R. Gold, unpubl. data) and other teleost lineages (Martin & Bermingham, 1998; Rüber *et al.*, 1999; Hanel & Sturmbauer, 2000). Additional nucleotide sequences will be required to better resolve relationships among species of *Pimephales* and to determine the extent of evolutionary plasticity of trophic characters in this genus.

We thank R. Cashner, T. Dowling, A. Echelle, W. Matthews and L. Page for assistance in obtaining samples used in this study, and M. Coburn for his helpful comments on this manuscript. We thank two anonymous reviewers for helpful comments. Research was supported in part by a National Science Foundation Doctoral Dissertation Improvement Grant (DEB-9700717), in part by a Thomas Slick Research Fellowship, and in part by the Texas Agricultural Experimental Station under Project H-6703. This paper represents Contribution Number 108 of the Center for Biosystematics and Biodiversity at Texas A&M University.

References

Bielawski, J. P. & Gold, J. R. (2001). Phylogenetic relationships of cyprinid fishes in subgenus *Notropis* inferred from nucleotide sequences of the mitochondrially encoded cytochrome *b* gene. *Copeia* **2001,** 656–667.

Broughton, R. E. & Gold, J. R. (2000). Phylogenetic relationships in the North American cyprinid genus *Cyprinella* (Actinopterygii: Cyprinidae) based on sequences of the mitochondrial ND2 and ND4L genes. *Copeia* **2000,** 1–10.

Cavender, T. M. & Coburn, M. M. (1986). Cladistic analysis of eastern North American Cyprinidae. *Ohio Journal of Science* **86,** 1.

Coburn, M. M. (1982). Anatomy and relationships of *Notropis atherinoides*. Unpubl. PhD dissertation, Ohio State University, Columbus, Ohio.

Coburn, M. M. & Cavendar, T. M. (1992). Interrelationships of North American fishes. In *Systematics, Historical Ecology and North American Freshwater Fishes* (Mayden, R. L., ed.), pp. 328–373. Stanford, CA: Stanford University Press.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39,** 783–791.

Gerking, S. D. (1994). *Feeding Ecology of Fishes*. San Diego, CA: Academic Press.

Hanel, R. & Sturmbauer, C. (2000). Multiple recurrent evolution of trophic types in northeastern Atlantic and Mediterranean seabreams (Sparidae, Percoidei). *Journal of Molecular Evolution* **50,** 276–283.

Hogan, K. M., Davis, S. K. & Greenbaum, I. F. (1997). Mitochondrial DNA analysis of the systematic relationships within the *Peromyscus maniculatus* species group. *Journal of Mammalogy* **78,** 733–743.

Hubbs, C. L. & Black, J. D. (1947). Revision of *Ceratichthys*, a genus of American cyprinid fishes. *Miscellaneous Publication of the Museum of Zoology, University of Michigan* **66,** 1–56.

Johnston, C. E. (1999). The relationship of spawning mode to conservation of North American minnows (Cyprinidae). *Environmental Biology of Fishes* **55,** 21–30.

Kishino, H., Miyata, T. & Hasegawa, M. (1990). Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *Journal of Molecular Evolution* **31,** 151–160.

Li, Y. C. & Gold, J. R. (1991). Cytogenetic studies in North American minnows (Cyprinidae). XXII. Chromosomal NORs in the genus *Pimephales*. *Canadian Journal of Zoology* **69,** 2826–2830.

Lockhart, J. P., Steele, M. A., Hendy, M. D. & Penny, D. (1994). Recovering evolutionary trees under a more realistic model of sequence evolution. *Molecular Biology and Evolution* **11,** 605–612.

Martin, A. P. & Bermingham, E. (1998). Systematics and evolution of lower Central American cichlids inferred from analysis of cytochrome *b* gene sequences. *Molecular Phylogenetics and Evolution* **9,** 192–203.

Mayden, R. L. (1987). Historical ecology and North American highland fishes: a research program in community ecology. In *Community and Evolutionary Ecology of North American Stream Fishes* (Matthews, W. J. & Heins, D. C., eds), pp. 210–221. Norman: University of Oklahoma Press.

Mayden, R. L. (1989). Phylogenetic studies of North American minnows, with emphasis on the genus *Cyprinella* (Teleostei: Cypriniformes). *University of Kansas Museum of Natural History, Miscellaneous Publications* **80,** 1–189.

Penny, D. & Hendy, M. (1985). Testing methods of evolutionary tree construction. *Cladistics* **1,** 266–272.

Rüber, L., Verheyen, E. & Meyer, A. (1999). Replicated evolution of trophic specializations in an endemic cichlid fish lineage from Lake Tanganyika. *Proceedings of National Academy of Science, USA* **96,** 10230–10235.

Schmidt, T. R. (1994). Phylogenetic relationships of *Hybognathus*. (Teleostei: Cyprinidae). *Copeia* **1994,** 622–630.

Schmidt, T. R., Dowling, T. E. & Gold, J. R. (1994). Molecular systematics of the genus *Pimephales* (Teleostei: Cyprinidae). *Southwestern Naturalist* **38,** 241–248.

Schmidt, T. R., Bielawski, J. P. & Gold, J. R. (1998). Molecular phylogenetics and evolution of the cytochrome *b* gene in the cyprinid genus *Lythrurus* (Actinopterygii: Cypriniformes). *Copeia* **1998,** 14–22.

Sibbing, F. A. (1991). Food capture and oral processing. In *Cyprinid Fishes: Systematics, Biology, and Exploitation* (Winfield, I. J. & Nelson, J. S., eds), pp. 377–412. New York: Chapman and Hall.

Swofford, D. L. (2000). *PAUP*. Phylogenetic Analysis using Parsimony (*and other methods), version 4*. Sunderland: Sinauer Associates.

Yang, Z. (1994*a*). Estimating the patterns of nucleotide substitution. *Journal of Molecular Evolution* **10,** 1396–1401.

Yang, Z. (1994*b*). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution* **39,** 306–314.

Yang, Z., Goldman, N. & Friday, A. E. (1995). Maximum likelihood trees from DNA sequences: a peculiar statistical estimation problem. *Systematic Biology* **44,** 384–399.