Phylogenetic Relationships of *Danio* Within the Order Cypriniformes: A Framework for Comparative and Evolutionary Studies of a Model Species

RICHARD L. MAYDEN[1]*, KEVIN L. TANG[1], KEVIN W. CONWAY[1], JÖRG FREYHOF[2], SARAH CHAMBERLAIN[1], MIRANDA HASKINS[1], LEAH SCHNEIDER[1], MITCHELL SUDKAMP[1], ROBERT M. WOOD[1], MARY AGNEW[1], ANGELO BUFALINO[1], ZOHRAH SULAIMAN[3], MASAKI MIYA[4], KENJI SAITOH[5], AND SHUNPING HE[6]

[1]*Department of Biology, Saint Louis University, St. Louis, Missouri*
[2]*Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany*
[3]*Department of Biology, Faculty of Science, Universiti Brunei Darussalam, Tungku BE, Brunei Darussalam*
[4]*Department of Zoology, Natural History Museum & Institute, Chiba, Chuo-ku, Chiba, Japan*
[5]*Tohoku National Fisheries Research Institute, Fisheries Research Agency, Shiogama-shi, Miyagi, Japan*
[6]*Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, China*

ABSTRACT The evolutionary relationships of species of *Danio* and the monophyly and phylogenetic placement of the genus within the family Cyprinidae and subfamily Rasborinae provide fundamentally important phyloinformatics necessary for direct evaluations of an array of pertinent questions in modern comparative biology. Although the genus *Danio* is not one of the most diverse within the family, *Danio rerio* is one of the most important model species in biology. Many investigations have used this species or presumed close relatives to address specific questions that have lasting impact on the hypothesis and theory of development in vertebrates. Largely lacking from this approach has been a holistic picture of the exact phylogenetic or evolutionary relationships of this species and its close relatives. One thing that has been learned over the previous century is that many organismal attributes (e.g., developmental pathways, ecologies, behaviors, speciation) are historically constrained and their origins and functions are best explained via a phylogenetic approach. Herein, we provide a molecular evaluation of the phylogenetic placement of the model species *Danio rerio* within the genus *Danio* and among hypothesized closely related species and genera. Our analysis is derived from data using two nuclear genes (RAG1, rhodopsin) and five mitochondrial genes (ND4, ND4L, ND5, COI, cyt *b*) evaluated using parsimony, maximum likelihood, and Bayesian analyses. The family Cyprinidae is resolved as monophyletic but the subfamily Rasborinae (priority over Danioinae) is an unnatural assemblage. *Danio* is identified as a monophyletic group sister to a clade inclusive of the genera *Chela, Microrasbora, Devario*, and *Inlecypris*, not *Devario* nor *Esomus* as hypothesized in previous studies. *Danio rerio* is sister

WILEY InterScience®
DISCOVER SOMETHING GREAT

Danio rerio, commonly referred to as the zebrafish or zebra danio, is a small cyprinid fish native to the streams of South-eastern Himalayan region (Talwar and Jhingran, '91; Menon, '99). Owing to its small body size, short life span, and its ability to reproduce in captivity, the zebrafish has become one of the most important model organisms for vertebrate developmental biology and genetics (Detrich et al., '99). Despite its importance as a model organism, the phylogenetic relationships of the zebrafish, and the composition and relationships of the entire genus *Danio*, remain unclear. This is a major gap in our knowledge because the phylogenetic placement of any model organism is of great importance if it is to be used in comparative biological studies.

Danio is a member of the order Cypriniformes, a large group of freshwater fishes distributed throughout North America, Africa, and Eurasia. Cypriniforms are in turn placed within the series Otophysi (a subgroup of the larger superorder Ostariophysi), a clade of freshwater fishes that also includes the tetras (order Characiformes), South American knifefishes (order Gymnotiformes), and catfishes (order Siluriformes). Members of the Otophysi are characterized by the presence of the Weberian apparatus (Rosen and Greenwood, '70; Fink and Fink, '81). This structure comprises modified perilymphatic and endolymphatic spaces of the inner ear, which are connected to a modified swimbladder through a series of modified anterior vertebral elements, termed the Weberian ossicles (Chranilov, '27). The Otophysi accounts for about 30% of the known fish species and 64% of all freshwater species (Nelson, 2006). The order Cypriniformes comprises six families: Cyprinidae, Catostomidae, Gyrinocheilidae, Psilorhynchidae, Cobitidae, and Balitoridae (Nelson, 2006). *Danio* is placed in the family Cyprinidae. Cyprinidae is the largest clade of freshwater fishes and the second largest vertebrate family with approximately 2,420 currently recognized species (Nelson, 2006). Although cyprinid monophyly has been substantiated (Siebert,

'87; Cavender and Coburn, '92) the intrarelationships of the family remain uncertain (Howes, '91; Nelson, 2006). The family is believed to be composed of between seven and ten nominal subfamily groupings (Howes, '91; Nelson, 2006), one of which, subfamily Rasborinae, includes *Danio*.

The name Rasborinae appears to have been proposed first by Weber and de Beaufort ('16) to include several cyprinid genera (including *Danio*) composed of small-sized to medium-sized species, and species lacking features thought to be characteristic of other groups. Other authors have chosen to use the name 'Danioinae' in the same meaning (e.g., Banarescu, '68; Rainboth, '91). As the name Rasborinae has priority over Danioinae, we chose to use the former, following Gosline ('75). The intrarelationships and interrelationships of the Rasborinae have yet to be thoroughly analyzed and several researchers (Howes, '91; Cavender and Coburn, '92; Nelson, 2006) have suggested that the Rasborinae may represent a non-monophyletic grouping. As with the poorly understood relationships within the Cyprinidae, relationships of Rasborinae to other subfamilies and the relationships of purported members of this subfamily, including *Danio* and *D. rerio*, remain unresolved.

The genus *Danio* was originally described as a subgenus of *Cyprinus* by Hamilton (1822). At present, the genus consists of 66 nominal species, of which about 45 are considered valid (Fang, 2003). Eight junior synonyms of *Danio* are currently recognized (Eschmeyer et al., '98). The name *Brachydanio* was later generally adopted as the valid genus name for the smaller, slender-bodied danios (including *D. rerio*), particularly in the aquarium literature (Axelrod, '85; Riehl and Baensch, '91), with the name *Danio* reserved only for the larger, deeper-bodied *Danio* species. Chu ('81) considered *Brachydanio* as a synonym of *Danio*, based on overlapping meristic (fin ray and scale) counts between species of the two genera, and placed the small, slender-bodied species of

Brachydanio back into *Danio*. Despite this change, the generic name *Brachydanio* was still used often in combination with the specific name *rerio* (e.g., Westerfield, '89). However, the name *Danio rerio* is currently accepted as the correct name for the zebrafish (Eschmeyer et al., '98).

Early studies of the phylogenetic relationships of *Danio* appeared rapidly when *D. rerio* was identified as a model organism species; however, these studies did not include a very thorough sampling of species that needed to be evaluated for a more comprehensive test of the monophyly of the genus, likely relationships of sister taxa, and the monophyly of the subfamily Rasborinae to which these fishes belong. Meyer et al. ('93) presented the first phylogenetic analysis inclusive of *Danio rerio* using 12S and 16S rDNA. That study, which only included a few species of *Danio* and related genera, was expanded by Meyer et al. ('95), which examined more genes for more species. Though the number of taxa and the number of genes had doubled between studies, the outcome was the same: *D. rerio* (+*D. 'frankei',* a synonym of *D. rerio*) was identified as the sister group to all other small-bodied species, with this 'slender-bodied' clade collectively forming the sister group of the 'deep-bodied' danios. The same result has been recovered frequently using different molecular and morphological data sets (Fig. 1A). From Meyer et al. ('93) to Sanger and McCune (2002), all previous phylogenetic studies of the zebrafish demonstrated monophyly for *Danio*. It was not until Fang (2003) included a wider assortment of small Southeast Asian 'rasborin' cyprinids, including species that had never been analyzed with *Danio* before, that different phylogenetic relationships were inferred. In Fang's (2003) analysis, *Esomus* Swainson, 1839 was found to be the sister group of the 'slender-bodied' species of *Danio*, rendering *Danio* as an unnatural, polyphyletic assemblage (Fig. 1B). Though Fang (2003) did not resolve the relationships of the smaller species of *Danio* (including *D. rerio*), she suggested for the first time that *Danio* was not monophyletic. Fang (2003) reserved the name *Danio*, for the slender-bodied clade that included *D. dangila* (the type species of *Danio*), and employed the name *Devario* for the clade composed of the deeper-bodied species of *Danio*.

Taxon and character sampling are critical components for the development of robust hypotheses of evolutionary relationships (Hillis, '98). Limited taxon sampling and character evaluation can both result in an incomplete interpretation of
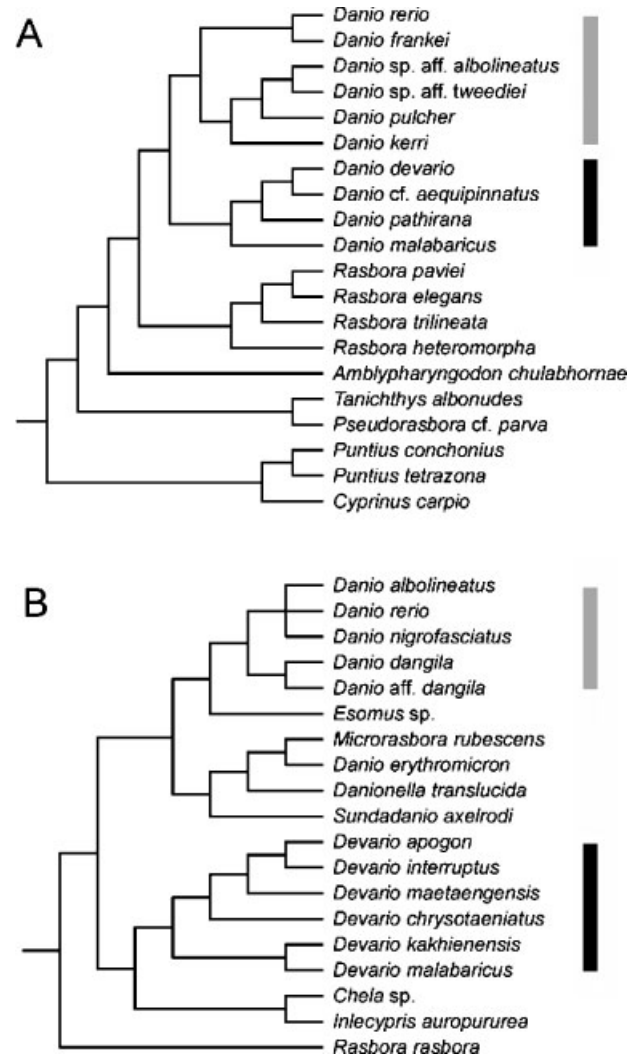

Fig. 1. Previous hypotheses of zebrafish relationships. (**A**) The phylogenetic position of *Danio rerio* and its close relatives based on 16S and 12S mitochondrial ribosomal genes (modified from Meyer et al., '95; Fig. 2). Similar topologies have been obtained by Meyer et al. ('93), Zardoya et al. ('96), Sanger and McCune (2002), Quigley et al. (2004, 2005); (**B**) Cladogram of rasborin taxa representing a strict consensus tree of two equally parsimonious trees of 97 steps (modified from Fang, 2003; Fig. 1). 'Slender-bodied' (gray) and 'deep-bodied' (black) clades are highlighted in each.

sister group relationships, and can consequently mislead the scientific community as to the evolutionary relationships of species. This can have a dramatic impact on the construction of predictive hypotheses for these species in diverse areas of biology when researchers rely on a phylogenetic hypothesis to propose developmental, evolutionary, ecological, physiological, and other hypotheses of genetically mediated attributes.

Using molecular sequence data from a variety of nuclear and mitochondrial genes and phylogenetic analyses, we provide an evaluation of the phylogenetic placement of the zebrafish within the genus *Danio*, evaluate the general relationships

of species within *Danio* and many closely related genera and species. We also provide a much broader survey of taxa that have been traditionally allied to the subfamily Rasborinae to test the monophyly of this group. Using this more comprehensive phylogenetic perspective for *D. rerio*, it is hoped that the evolutionary hypothesis will provide a variety of scientific communities interested in this species and its relatives with a more powerful predictive framework with which to develop and test hypotheses regarding the features discovered in this model species.

MATERIALS AND METHODS

Taxon sampling

Representative taxa from four of the 11 subfamilies of Cyprinidae (Cyprininae, Gobioninae, Leuciscinae, and Rasborinae) belonging to 28 genera were included in this analysis. In an effort to represent the diversity of the Cyprinidae, taxon sampling focused on including members of genera that previously have been hypothesized as the sister group of *Danio* (Sanger and McCune, 2002; Fang, 2003), in addition to a representative sampling of rasborin taxa. Outgroup taxa were drawn from the families Balitoridae, Catostomidae, and Cobitidae, as these families, along with the Cyprinidae, form the bulk of cypriniform diversity. A total of 56 cyprinid species and six non-cyprinid cypriniform species were examined for this study; a full list of species examined is given in Table 1.

DNA amplification and sequencing

Genomic extractions were taken from muscle tissue or fin clips either frozen at $-80°C$ or preserved in ethanol ($>95\%$ concentration), using DNeasy tissue extraction kits (Qiagen, Valencia, CA, USA). Target regions of mitochondrial and nuclear DNA were amplified using polymerase chain reaction (PCR) and the primers listed in Table 2. Target loci included cytochrome *b* (1,141 bp), a 712-bp fragment of cytochrome *c* oxidase subunit I, nicotinamide adenine dinucleotide (reduced form) (NADH) dehydrogenase subunits 4L and 4 (1,748 bp), a 894-bp fragment of

NADH dehydrogenase subunit 5, a 1,518-bp fragment of exon III of recombinase activating gene 1, and a 905-bp fragment of rod opsin (protein component of the rhodopsin photoreceptor). Amplification used the following thermal cycling profiles: 94°C denaturing (30–60 sec), 45–55°C annealing (30–60 sec), and 72°C extension (2 min 30 sec), for 30–40 cycles; an initial heating step at 94°C for 30–60 sec preceded cycling and some profiles included a final extension step at 72°C for 2–5 min after cycling was complete. Amplified products were either directly purified using QIAquick PCR purification kits (Qiagen) or loaded onto agarose gels and electrophoresed, followed by excision of the target DNA from the gel and purification using QIAquick gel extraction kits (Qiagen). Purified PCR samples were then sent to a commercial sequencing facility at Macrogen Inc. for cycle sequencing on an Applied Biosystems 3730xl automated sequencer, using the primers indicated in Table 2. Both strands were sequenced for all target gene regions. A consensus light strand sequence for each gene region for each taxon was assembled from the two complementary sequences; these consensus sequences have been deposited with GenBank (see Table 2 for accession numbers). An initial alignment was made by eye, comparing our sequence data with the previously published sequences of *Danio rerio*; due to the protein-coding nature of these loci, alignment was relatively straightforward. Some loci for some taxa could not be successfully sequenced, these base positions were coded as missing for those taxa. Additional sequence data for the outgroup taxa as well as the non-rasborin cyprinids were obtained from GenBank (Table 2).

Phylogenetic analyses

Three separate types of searching methods were performed on the compiled data matrix: parsimony, maximum likelihood, and Bayesian. Parsimony searches were carried out using PAUP* 4.0b10 (Swofford, 2002) in conjunction with PAUPRat (Sikes and Lewis, 2001), which implements a parsimony ratchet search (Nixon, '99) in PAUP*. PAUPRat was used to perform 1,000 parsimony ratchet replicate searches in PAUP* using 5, 10, 15, 20, and 25% perturbation, with 200 replicates for each percentage value. Gaps were treated as a fifth character state. The best trees recovered via the ratchet searches were then used as the starting trees for a heuristic parsimony search in PAUP*

TABLE 1. Species of Cypriniformes sequenced for nuclear and mitochondrial genes, with GenBank accession numbers

Taxon	COI	Cyt b	ND4	ND5	RAG1	Rhod
Order Cypriniformes						
Family Balitoridae						
Crossostoma lacustre	NC001727	NC001727	NC001727	NC001727	N/A	N/A
Lefua echigonia	NC004696	NC004696	NC004696	NC004696	EF458305	N/A
Family Catostomidae						
Carpiodes carpio	NC005257	NC005257	NC005257	NC005257	N/A	N/A
Myxocyprinus asiaticus	NC006401	NC006401	NC006401	NC006401	N/A	N/A
Family Cobitidae						
Cobitis sinensis	NC007229	NC007229	NC007229	NC007229	N/A	N/A
Cobitis striata	NC004695	NC004695	NC004695	NC004695	EF458303	N/A
Family Cyprinidae						
Subfamily Cyprininae						
Carassius auratus	NC006580	NC006580	NC006580	NC006580	N/A	N/A
Carassius carassius	NC006291	NC006291	NC006291	NC006291	N/A	N/A
Cyprinus carpio	NC001606	NC001606	NC001606	NC001606	EF458304	N/A
Sawbwa resplendens	EF452895	N/A	EF452824	N/A	N/A	N/A
Subfamily Gobioninae						
Sarcocheilichthys variegatus	NC004694	NC004694	NC004694	NC004694	N/A	N/A
Subfamily Leuciscinae						
Campostoma anomalum	EF452850	AF452079	EF452784	EF452751	EF452827	EF452898
Hemitremia flammea	EF452851	AY281054	EF452785	EF452752	EF452828	EF452899
Luxilus chrysocephalus	EF452852	AF117167	EF452786	EF452753	EF452829	EF452900
Nocomis biguttatus	EF452853	AY486057	EF452787	EF452754	EF452830	EF452901
Notemigonus crysoleucas	EF452854	U01318	EF452788	EF452755	EF452831	EF452902
Notropis atherinoides	EF452855	AY281062	EF452789	EF452756	EF452832	EF452903
Opsopoeodus emiliae	EF452856	U17270	EF452790	EF452757	EF452833	EF452904
Phoxinus erythrogaster	EF452857	AY281055	EF452791	EF452758	EF452834	EF452905
Richardsonius balteatus	EF452858	AY096011	N/A	EF452759	EF452835	EF452906
Subfamily Rasborinae						
Boraras brigittae	N/A	N/A	EF452792	EF452760	N/A	N/A
Boraras maculatus	EF452859	N/A	EF452793	EF452761	N/A	N/A
Boraras merah	EF452884	N/A	EF452815	EF452778	EF452838	EF452909
Boraras sp. cf. *micros*	EF452885	N/A	EF452816	EF452779	EF452839	EF452910
Boraras urophthalmoides	EF452886	N/A	EF452817	EF452780	EF452840	EF452911
Chela cachius	EF452891	EF452745	EF452821	N/A	EF452845	EF452914
Chela dadiburjori	EF452892	EF452746	N/A	N/A	N/A	EF452915
Danio choprai	EF452879	EF452740	EF452810	N/A	N/A	N/A
Danio erythromicron	EF452867	EF452737	EF452800	EF452768	N/A	N/A
Danio feegradei	EF452861	EF452732	EF452795	EF452763	N/A	N/A
Danio kyathit	EF452862	EF452733	EF452796	EF452764	N/A	N/A
Danio nigrofasciatus	EF452863	N/A	N/A	N/A	N/A	N/A
Danio sp. ''hikari''	EF452860	EF452731	EF452794	EF452762	N/A	N/A
Danio sp. ''panther''	EF452864	EF452734	EF452797	EF452765	N/A	N/A
Danio rerio	NC002333	NC002333	NC002333	NC002333	NM131389	NM131084
Danio roseus	EF452865	EF452735	EF452798	EF452766	N/A	N/A
Danionella sp.	EF452887	EF452741	EF452818	EF452781	EF452841	N/A
Devario devario	EF452866	EF452736	EF452799	EF452767	N/A	N/A
Esomus sp. cf. *ahli*	EF452888	EF452742	N/A	EF452782	EF452842	N/A
Inlecypris auropurpurea	EF452889	EF452743	EF452819	N/A	EF452843	EF452912
Microrasbora kubotai	EF452868	EF452738	EF452801	EF452769	N/A	N/A
Microrasbora rubescens	EF452890	EF452744	EF452820	EF452783	EF452844	EF452913
Opsaridium sp.	EF452893	EF452747	EF452822	N/A	EF452846	N/A
Opsariichthys uncirostris	EF452894	EF452748	EF452823	N/A	EF452847	EF452916
Rasbora argyrotaenia	EF452880	N/A	EF452811	EF452776	EF452836	EF452907
Rasbora brittani	EF452869	N/A	EF452802	EF452770	N/A	N/A
Rasbora caudimaculata	EF452870	N/A	EF452803	EF452771	N/A	N/A
Rasbora cephalotaenia	EF452881	N/A	EF452812	N/A	N/A	N/A
Rasbora daniconius	EF452872	N/A	EF452805	EF452773	N/A	N/A

TABLE 1. Continued

Taxon	COI	Cyt *b*	ND4	ND5	RAG1	Rhod
Rasbora dorsiocellata	EF452873	N/A	EF452806	EF452774	N/A	N/A
Rasbora kottelati	N/A	N/A	N/A	EF452775	N/A	N/A
Rasbora rubrodorsalis	EF452874	N/A	EF452807	N/A	N/A	N/A
Rasbora sp. cf. *bankanensis*	EF452871	N/A	EF452804	EF452772	N/A	N/A
Rasbora sumatrana	EF452882	N/A	EF452813	EF452777	EF452837	EF452908
Rasbora trilineata	EF452883	N/A	EF452814	N/A	N/A	N/A
Rasbora vaterifloris	EF452876	N/A	N/A	N/A	N/A	N/A
Rasbora vulcanus	EF452875	N/A	N/A	N/A	N/A	N/A
Sundadanio axelrodi	N/A	EF452739	EF452808	N/A	N/A	N/A
Trigonostigma espei	EF452877	N/A	N/A	N/A	N/A	N/A
Trigonostigma hengeli	EF452878	N/A	EF452809	N/A	N/A	N/A
Zacco platypus	EF452896	EF452749	EF452825	N/A	EF452848	EF452917
Zacco temminckii	EF452897	EF452750	EF452826	N/A	EF452849	EF452918

Abbreviations: COI, cytochrome *c* oxidase subunit I; Cyt *b*, cytochrome *b*; ND4, NADH dehydrogenase subunit 4; ND5, NADH dehydrogenase subunit 5; RAG1, recombination activating gene 1; Rhod = opsin.

TABLE 2. List of primers and primer sequences used in reconstructing relationships of Cypriniformes species

Primer	Sequence (5′–3′)	Source
Cytochrome *b*		
LA-danio	GACTYGAARAACCACYGTTG	This study
HA-danio	CTCCGATCTTCGGATTACAAG	This study
Cytochrome *c* oxidase subunit I		
LCO1490	GGTCAACAAATCATAAAGATATTGG	Folmer et al. ('94)
HCO2198	TAAACTTCAGGGTGACCAAAAAATCA	Folmer et al. ('94)
NADH dehydrogenase subunit 4 & 5		
L10444	AAGACCTCTGATTTCGGCTCA	This study
L11427-ND4-C	CCWAAGGCSCATGTWGARGC	Miya et al. (2006)
L12328-Leu-C	AACTCTTGGTGCAAMTCCAAG	Miya et al. (2006)
H11618a	TGRCTKACSGAKGAGTAGGC	This study
H12296-Leu-C	CAAGAGTTTTTGGTTCCTAAG	Miya et al. (2006)
H13393-ND5-C	CCTATTTTKCGGATGTCTTGYTC	Miya et al. (2006)
Opsin		
Rh193	CNTATGAATAYCCTCAGTACTACC	Chen et al. (2003)
Rh1073r	CCRCAGCACARCGTGGTGATCATG	Chen et al. (2003)
RAG1		
2533F	CTGAGCTGCAGTCAGTACCATAAGATGT	Lopez et al. (2004)
4090R	CTGAGTCCTTGTGAGCTTCCATRAAYTT	Lopez et al. (2004)

with TBR branch swapping. The most-parsimonious trees recovered were evaluated using summary values reported by PAUP* (e.g., tree length, consistency index). Branch support was evaluated by calculating decay index values (Bremer, '88) for each clade using TreeRot v. 2 (Sorenson, '99). Bootstrap values were calculated in PAUP* using a simple heuristic search and 10,000 bootstrap replicates (1,000 heuristic searches, each with ten random addition replicates).

Maximum likelihood searches were conducted with GARLIv0.941 (Zwickl, unpublished data; http://www.bio.utexas.edu/grad/zwickl/web/garli.html). Twenty individual runs were performed using the default search settings (5,000,000 generations) and termination criteria (genthreshfortopoterm = 20,000; scorethreshforterm = 0.05) with random starting topologies. Hierarchical likelihood ratio tests (hLRTs) performed with MrModelTest v.2.2 (Nylander, 2004) and PAUP* found that the best-fit model for a majority of loci was GTR+I+Γ see Results). As this is the default model implemented by GARLI the GTR+I+Γ model was used for the likelihood searches and gaps were treated as missing. Resulting topologies were optimized in PAUP* to calculate log likelihood scores, using the rates matrix, base composition frequencies, proportion of invariable sites, and gamma distribu-

tion estimated by the individual GARLI searches. The lnL scores estimated by GARLI versus PAUP* differed by less than 0.00001 when they differed at all. The topology with the best likelihood score was retained. Bootstrap values were calculated in GARLI using 100 bootstrap replicates and the termination criteria were relaxed slightly (genthreshfortopoterm = 10,000).

Bayesian analyses were conducted with the MPI version (Altekar et al., 2004) of MrBayes v.3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) on a parallel computing cluster with 12 Apple Xserve dual processor nodes running UNIX. *Cobitis striata* was designated as the single outgroup taxon. Before the analysis, the sequence data were partitioned by codon position and loci, resulting in 15 separate data partitions (one for each codon position for each of the five loci). MrModelTest v.2.2 (Nylander, 2004) and PAUP* were used to perform hLRTs on each partition to determine the best-fit model of nucleotide substitution. These models were then applied to the appropriate partitions during the subsequent MrBayes analyses. Two independent Bayesian searches were conducted using the same search parameters: 10,000,000 generations and 24 chains (with one cold chain and 23 heated chains), with sampling every 10,000 generations. The distribution of log likelihood scores was examined to determine burn-in time for each of the analyses. Trees recovered after stationarity had been achieved were retained. Branch support for each clade was based on clade credibility values, indicated by the frequency of occurrence of each clade among the trees retained after the initial burn-in topologies were discarded. The trees retained from the two independent analyses were combined into a single pool for the purposes of calculating clade credibility values, which was accomplished by constructing a 50% majority-rule consensus of these trees in PAUP*.

RESULTS

A total of 6,921 aligned base positions were obtained for 62 taxa, 56 cyprinid species and six species representing other cypriniform families. Our searches utilizing different optimality criteria converged on very similar hypotheses of relationships. The parsimony analysis recovered 18 most-parsimonious trees with a total length of 21,861 steps (CI = 0.270, RI = 0.431, RC = 0.116); the strict consensus of these trees is shown in

Figure 2. The family Cyprinidae is recovered as monophyletic, as are the subfamilies Cyprininae and Leuciscinae. The subfamily Rasborinae, however, is rendered polyphyletic by the placement of *Opsariichthys* and *Zacco* in a clade with species of the subfamilies Gobioninae and Leuciscinae. The genus *Danio* is found to be monophyletic within a clade containing the remaining Rasborinae representatives, with *Danio rerio* occupying an apical position within the group and sister to *D. kyathit*. *Danio erythromicron* is clearly placed within *Danio* and not *Microrasbora*, as has been previously hypothesized for this species (Kottelat and Witte, '99). The sister taxon of *Danio* is a clade composed of *Chela*, *Inlecypris*, *Microrasbora*, and most notably *Devario*, a previously hypothesized sister group of *Danio* (sensu Sanger and McCune, 2002). The genus *Esomus* is recovered as the basal sister group to the larger clade containing *Danio*, and is not the sister group of the genus as previously hypothesized by Fang (2003). A monophyletic group containing *Sundadanio* sister to *Danionella* plus *Opsaridium* forms the sister group to the remaining members of this clade, excluding *Esomus*. This large clade containing *Danio* and relatives (*Chela*, *Danionella*, *Devario*, *Esomus*, *Inlecypris*, *Microrasbora*, *Opsaridium*, and *Sundadanio*) is recovered as the sister group to a clade that includes *Rasbora*, *Rasboroides*, *Boraras*, and *Trigonostigma*. The genus *Rasbora* is polyphyletic relative to *Rasboroides vaterifloris*, a monophyletic *Boraras*, and species of *Trigonostigma*, with some *Rasbora* more closely related to members of the other three genera than to other *Rasbora*. The largely unresolved relationships among some species of *Rasbora* in the *Rasbora-Rasboroides-Trigonostigma* clade is a result of limited sequence variability in these species for the genes examined, not conflicting relationships.

The optimal topology recovered by the maximum likelihood searches has a lnL score of $-92,351.0953$ (Fig. 3). The results of the maximum likelihood analysis are largely congruent with the results of the parsimony analysis. The genus *Danio* is found to be monophyletic and a member of a clade which contains most of the putative members of the Rasborinae, which is again not monophyletic because *Opsariichthys* and *Zacco* are recovered in a clade with the gobionins and leuciscins. *Danio rerio* is again recovered in a derived position in the genus, with *D. kyathit* as its sister species. A clade comprised *Chela*, *Devario*, *Inlecypris*, and *Microrasbora* appears again as the sister taxon of *Danio*. In these analyses the genus
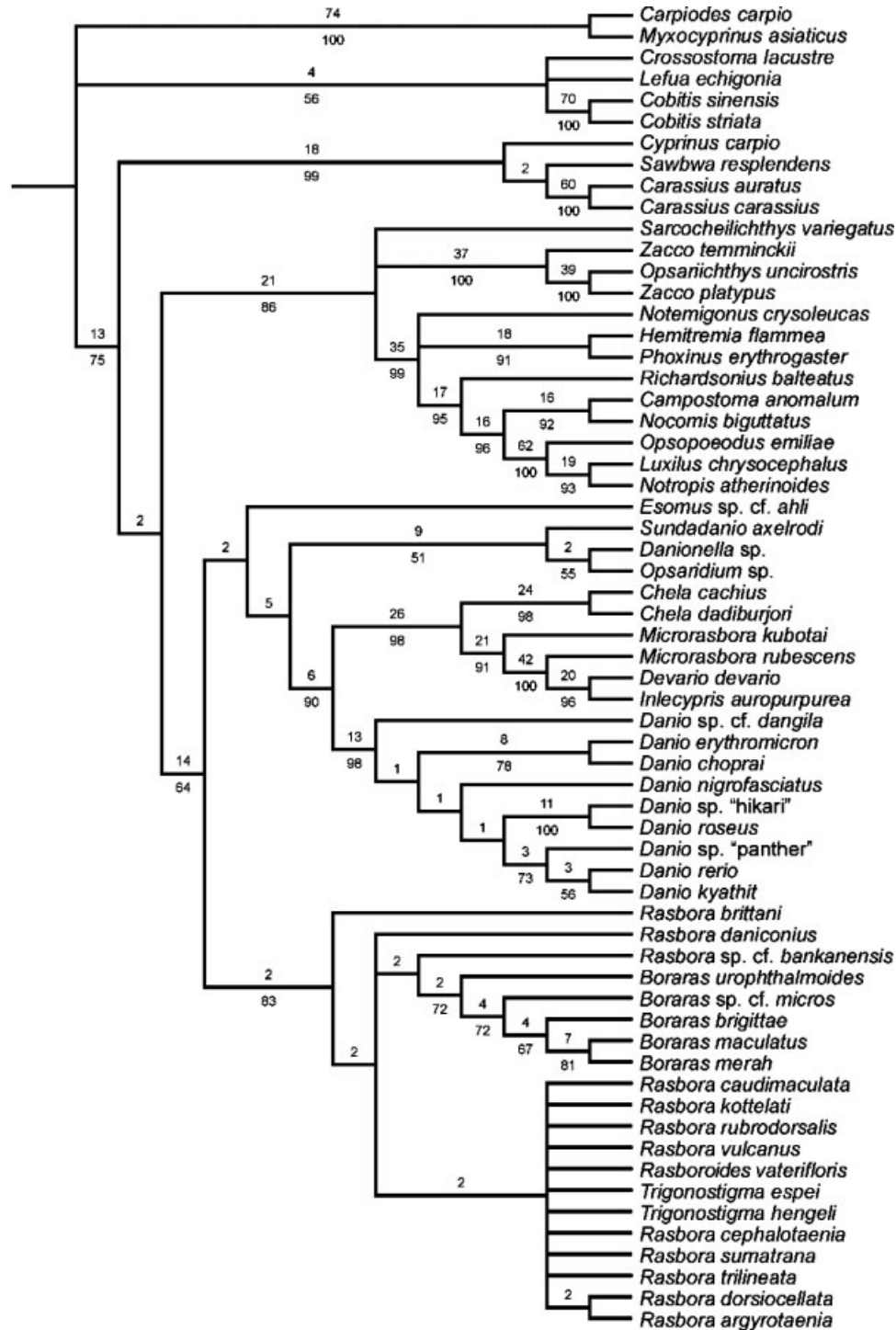
Fig. 2. Strict consensus of 18 equally most-parsimonious trees resulting from a parsimony analysis with parsimony ratcheting; TL = 21,861 steps, CI = 0.270, RI = 0.431, RC = 0.116. Decay index support (above) and bootstrap values for 10,000 replicates (below) are reported at each node.

Rasbora is resolved as polyphyletic relative to monophyletic *Boraras* and *Trigonostigma*. In this analysis, however, *Esomus* forms the basal sister group to the *Rasbora-Rasboroides-Boraras-Trigo-* *nostigma* clade, not the clade containing *Danio* and relatives as described above (Fig. 2). Some resolution of species relationships within *Rasbora* is elucidated in this analysis, but sup-

— 0.05 substitutions/site
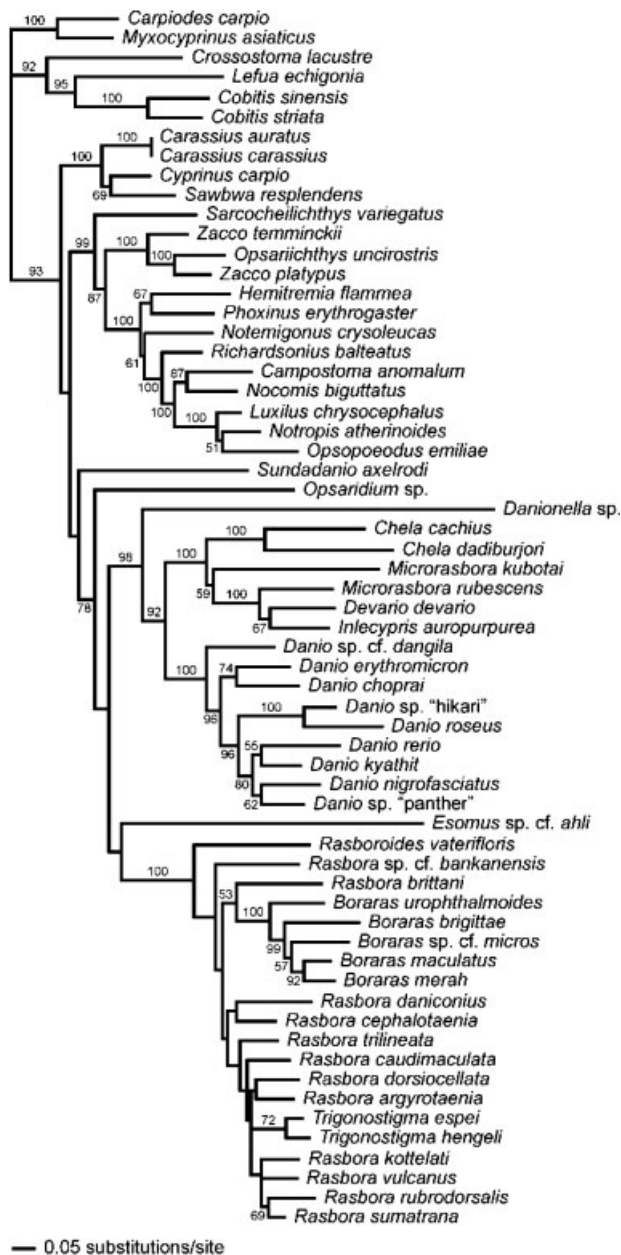
Fig. 3. Tree topology recovered with the best log likelihood score of 20 independent maximum likelihood searches; lnL = −92,351.0953. Bootstrap values for 100 replicates are reported at each node (bootstrap values were less than 50% where no value is given).

port for these relationships is poor and branch lengths are very short, again illustrating the general lack of variation among these taxa for the genes examined. As in the parsimony analyses, the genus *Microrasbora* is not a monophyletic group and does not include *Danio erythromicron*.

For the Bayesian analyses, the hLRTs performed with MrModelTest and PAUP* found that

the best-fit model for a majority of loci was GTR+I +Γ. The following data partitions had GTR+Γ: cyt *b* third positions, RAG1 third positions, and rhodopsin third positions. HKY+I+Γ was found as the best-fit model for second positions of RAG1; F81+I+Γ was the best-fit model for first positions in rhodopsin; and F81+Γ was the best fit for rhodopsin second positions. In both Bayesian searches, it appeared that stationarity had been reached after approximately 50,000 generations. The first 100,000 generations (11 trees) were discarded to ensure all burn-in trees were excluded from the results, leaving 990 of the original 1,001 saved trees (from 10,000,000 generations) from each of the independent searches. This yielded a total of 1,980 trees from a combined 20,000,000 generations that were included in the 50% majority-rule consensus tree (Fig. 4).

The relationships found in the Bayesian tree mostly correspond with those seen in the parsimony and likelihood trees. *Danio rerio* and *D. kyathit* are recovered as sister species in a monophyletic *Danio*, with the clade of *Chela*, *Devario*, *Inlecypris*, and *Microrasbora* as the sister group of *Danio*. One of the major differences between the Bayesian tree and the other two trees is that *Sundadanio axelrodi* is recovered as the sister group of the subfamily Cyprininae, not as part of the ''rasborin'' clade (putative rasborins minus *Opsariichthys*, and *Zacco*), as seen in the other two searches. This highlights one of the only major areas of disagreement between the three different search methods: *Sundadanio*, along with *Danionella*, *Esomus*, and *Opsaridium*, are found in different positions in all three trees. This instability is a result of insufficient sequence data for appropriate genes for these four taxa, as is also observed in species of *Rasbora*. One other notable difference in the Bayesian tree is that the sister group of the ''rasborin'' clade is the Cyprininae (and *Sundadanio axelrodi*) not the leuciscin+ gobionin+*Opsariichthys*+*Zacco* clade seen in the parsimony and likelihood trees.

DISCUSSION

Although the order Cypriniformes contains the world's most diverse group of freshwater fishes, including important species for biological and evolutionary studies as well as many species of great cultural and economic importance, the diversity and relationships of these species are far from being well understood. To some degree, the paucity of information regarding these species
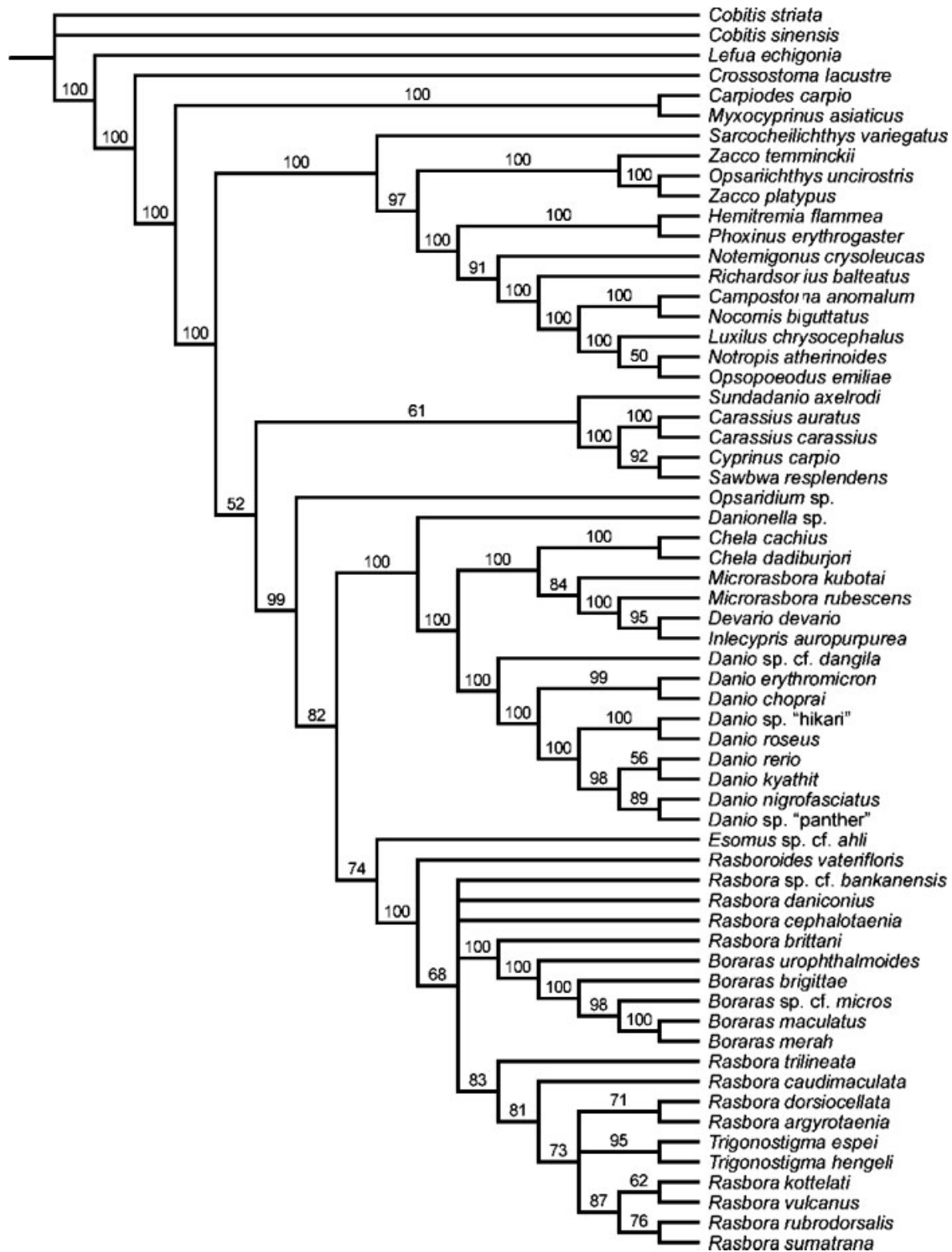
Fig. 4. The 50% majority-rule consensus of 1,980 trees recovered by Bayesian inference, with clade credibility values reported at each node.

is a result of the overwhelming diversity within the group, as scientists traditionally have had difficulty identifying these species. The Cypriniformes

Tree of Life Initiative (www.cypriniformes.org) is a worldwide consortium of researchers that have recently focused, in a collaborative partner-

ship, to jointly resolve the phylogenetic relationships of these important fishes and better understand their diversity, biogeography, anatomy, ecology, and evolution. This collaborative effort has also involved an important opportunity for systematic ichthyologists to interact with scientists involved in primary research on other aspects of the biology of *Danio rerio*, facilitating a better understanding of the many different biological properties of this species and its relatives. With the current and impending threats to most of our planet's biodiversity and their ecosystems, more focused and collaborative international efforts involving scientists from many different countries are needed to study a group as diverse as the Cypriniformes. Although several researchers have worked toward improving our understanding of the evolutionary relationships of *Danio rerio*, only through this collaborative international effort have we been able to obtain and evaluate a larger number of species of Cypriniformes, providing better insight into the phylogenetic position and biology of this important model species.

Most developmental and genetic studies involving *Danio rerio* make comparisons with sticklebacks or fugu, or even human or mouse. These comparisons are important for understanding evolution, but such comparisons can also aid our understanding of evolutionary developmental biology if they are cast in a phylogenetic framework. A synthetic approach to the evolutionary origins of diversity and/or control mechanisms of morphology can be the focus of research in purely theoretical fields, as well as in more applied fields like medicine, agriculture, and aquaculture. Goldschmidt ('40, pp 205–206) noted that: ''The change from species to species is not a change involving more and more additional atomistic changes, but a complete change of the primary pattern or reaction system into a new one, which afterward may again produce intraspecific variation by micromutation.'' Goldschmidt's hypothesis was universally rejected and widely ridiculed within the biological community of his day, which favored the Neo-Darwinian explanations of R.A. Fisher, J.B.S. Haldane, and Sewall Wright. The Neo-Darwinian paradigm, however, has yet to provide a convincing mechanism for the origin of species that has withstood testing within a phylogenetic framework, largely because adherents of this paradigm have lacked the basic data or testable hypotheses which would allow one to study causal relationships between descent, genetics, and morphology. Model organisms provide clear opportunities

for the scientific community to investigate many evolutionary and developmental questions. However, a well-corroborated phylogeny of the model organism and its closest relatives is required before the evaluation of pertinent questions can begin.

In all of our analyses, the genus *Danio* is recovered as a monophyletic group and *D. rerio* is sister to *D. kyathit*. With more sampling of species of *Danio* in the future, this sister group relationship may change and species not yet sampled may be identified as the closest relative. However, predictive comparative analyses of different aspects of the biology, anatomy, physiology, genetics, and development of *D. rerio* can begin to focus on those species identified as close relatives by this study, which include *D. kyathit, D. nigrofasciatus, D.* sp. ''panther,'' *D. roseus, D.* sp. ''hikari,'' *D. choprai,* and *D. erythromicron* (Quigley et al., 2004, 2005). Our phylogenetic analyses indicate that comparisons of *Danio* with species of *Devario* and/or *Esomus* may be misleading because they are not as closely related as once thought.

Danio erythromicron, a species originally described in *Microrasbora*, is found to be more closely related to *D. rerio* than the putative species of the non-monophyletic *Microrasbora* resolved herein. Although the resulting placement of this species in *Danio* may be somewhat surprising to some, this species was recently referred to *Danio* by Kottelat and Witte ('99). These authors reassigned *D. erythromicron* on the basis of a suite of morphological characters, but also indicated that this change should be further verified with additional data; herein their conclusion is strongly supported by our molecular sequence data and analyses.

The non-monophyly of the subfamily Rasborinae is not too surprising as this grouping has traditionally been used as a ''catch-all'' for taxa not easily assignable to other subfamilies (Howes, '91). Further evaluation of additional cypriniform taxa will provide further tests of the monophyly of this subfamily and other subfamilies. Formal reclassification of species from Rasborinae or other subfamilies should await more comprehensive analyses following from the Cypriniformes Tree of Life initiative or other investigations (Saitoh et al., 2006).

Other important findings include the close relationship between *Devario* and *Inlecypris*, two clades with similar general coloration and morphology, and the observation that the genus *Esomus* is not closely related to the ''slender''

group of *Danio* as previously hypothesized. We have demonstrated the non-monophyly of the genus *Rasbora* for the first time. Although this conclusion is also one that is not too surprising given the tremendous anatomical, morphological, and ecological diversity in this genus, it clearly indicates that *Rasbora*, a highly speciose group of southeast Asian cyprinids, is in need of significant systematic attention. The monophyly of the genus *Boraras* as proposed by Conway (2005), based on morphological characters in a phylogenetic analysis, is corroborated with molecular sequence data herein. The monophyly of the genus *Trigonostigma* is supported in both likelihood and Bayesian analyses, whereas the relationships of species in this genus are unresolved in the parsimony analysis. Regardless of the specifics of relationships within those two genera, the placement of *Boraras* and *Trigonostigma* (and *Rasboroides*, in the parsimony tree) within *Rasbora* reveals *Rasbora* as a unnatural assemblage in need of additional phylogenetic study.

The collaborative efforts initiated recently with the Cypriniformes Tree of Life (www.cypriniformes.org), and between members of the Tree of Life project and molecular geneticists and developmental biologists using *D. rerio* as a model system, provide far-reaching opportunities to advance many areas of comparative and evolutionary biology. Knowledge of the phylogenetic relationships of *D. rerio* provide a fundamental framework in which scientists can begin to better understand the evolution of genetic mechanisms central to morphological and developmental changes in evolutionary time, morphological adaptations to different environments, and the nature of mutants, all major areas of inquiry that are of interest to a broad audience of comparative and evolutionary biologists.

ACKNOWLEDGMENTS

LITERATURE CITED

Altekar G, Dwarkadas S, Huelsenbeck JP, Ronquist F. 2004. Parallel Metropolis-coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics 20: 407–415.

Axelrod HR. 1985. Dr. Axelrod's Atlas of Freshwater Aquarium Fishes. Neptune City, NJ: T.F.H. Publications.

Banarescu P. 1968. Remarks on the genus *Chela* Hamilton-Buchanan (Pisces, Cyprinidae) with description of a new subgenus. Ann Mus Civ Stor Nat 'Giacomo Doria' 77:53–64.

Bremer K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42: 795–803.

Cavender TM, Coburn MM. 1992. Phylogenetic relationships of North American Cyprinidae. In: Mayden RL, editor. Systematics, Historical Ecology, and North American Freshwater Fishes. Stanford, CA: Stanford University Press. p 293–328.

Chen W-J, Bonillo C, Lecointre G. 2003. Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. Mol Phylogenet Evol 26:262–288.

Chranilov NS. 1927. Beiträge zur Kenntnis des Weber'schen Apparates der Ostariophysi. 1. vergleichend-anatomische Übersicht der Knochenelemente des Weber'schen Apparates bei Cypriniformes. Zool Jb Anat 49:501–597.

Chu XL. 1981. A preliminary revision of fishes of the genus *Danio* from China. Zool Res 2:145–156.

Conway KW. 2005. Monophyly of the genus *Boraras* (Teleostei: Cyprinidae). Ichthyol Explor Freshw 16:249–264.

Detrich WH, Westerfield M, Zon L. 1999. The zebrafish, genetics and genomics, Vol. 60. San Diego, CA: Academic Press.

Eschmeyer WN, Ferraris CJ, Hoang MD, Long DJ. 1998. Species of fishes. In: Eschmeyer WN, editor. Catalog of fishes. San Francisco, CA: California Academy of Sciences. p 25–1820.

Fang F. 2003. Phylogenetic analysis of the Asian cyprinid genus *Danio* (Teleostei, Cyprinidae). Copeia 2003: 714–728.

Fink SV, Fink WL. 1981. Interrelationships of the ostariophysan fishes (Teleostei). Zool J Linn Soc 72:297–353.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome *c*

oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299.

Goldschmidt RB. 1940. The material basis of evolution. New Haven, CT: Yale University Press.

Gosline W. 1975. The cyprinid dermosphenotic and the subfamily Rasborinae. Occas Pap Mus Zool Univ Mich 673: 1–13.

Hamilton F. 1822. An account of the fishes found in the River Ganges and its branches. Edinburgh: Archibald Constable and Co.

Hillis DM. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst Biol 47:3–8.

Howes GJ. 1991. Systematics and biogeography: an overview. In: Winfield IJ, Nelson JS, editors. Cyprinid fishes: systematics, biology and exploitation. London: Chapman and Hall. p 1–33.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754–755.

Kottelat M, Witte KE. 1999. Two new species of *Microrasbora* from Thailand and Myanmar, with two new generic names for small Southeast Asian cyprinid fishes (Teleostei: Cyprinidae). J South Asian Nat Hist 4:49–56.

López JA, Chen W-J, Ortí G. 2004. Esociform phylogeny. Copeia 2004:449–464.

Menon AGK. 1999. Check list—fresh water fishes of India. Rec Zool Surv India Misc Publ Occas Pap 175:1–366.

Meyer A, Biermann CH, Ortí G. 1993. The phylogenetic position of the zebrafish (*Danio rerio*), a model system in developmental biology: an invitation to the comparative method. Proc R Soc Lond B 252:231–236.

Meyer A, Ritchie PA, Witte KE. 1995. Predicting developmental processes from evolutionary patterns: a molecular phylogeny of the zebrafish (*Danio rerio*) and its relatives. Philos Trans R Soc Lond B 349:103–111.

Miya M, Saitoh K, Wood RM, Nishida M, Mayden RL. 2006. New primers for amplifying and sequencing the mitochondrial ND4/ND5 gene region of the Cypriniformes (Actinopterygii: Ostariophysi). Ichthyol Res 53:75–81.

Nelson JS. 2006. Fishes of the World. New York, NY: John Wiley and Sons.

Nixon KC. 1999. The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics 15:407–414.

Nylander JAA. 2004. MrModeltest v2.2. Uppsala: Evolutionary Biology Centre, Uppsala University.

Quigley IK, Turner JM, Nuckels RN, Manuel JL, Budi E, MacDonald EI, Parichy DM. 2004. Pigment pattern evolution by differential deployment of neural crest and postembryonic melanophore lineages in *Danio* fishes. Development 131:6053–6069.

Quigley IK, Turner JM, Nuckels RN, Manuel JL, Budi E, MacDonald EI, Parichy DM. 2005. Evolutionary diversification of pigment pattern in *Danio* fishes: differential *fms* dependence and stripe loss in *D. albolineatus*. Development 132:89–104.

Rainboth WJ. 1991. Cyprinids of South East Asia. In: Winfield IJ, Nelson JS, editors. Cyprinid fishes: systematics, biology and exploitation. London: Chapman and Hall. p 156–202.

Riehl R, Baensch HA. 1991. Aquarien Atlas. Band. 1. Melle: Mergus, Verlag für Natur- und Heimtierkunde.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Rosen DE, Greenwood PH. 1970. Origin of the Weberian apparatus and the relationships of the ostariophysan and gonorynchiform fishes. Am Mus Novit 2428:1–25.

Saitoh K, Sado T, Mayden RL, Hanzawa N, Nakamura K, Nishida M, Miya M. 2006. Mitogenomic evolution and interrelationships of the Cypriniformes (Actinopterygii: Ostariophysi): The first evidence towards resolution of higher-level relationships of the world's largest freshwater-fish clade based on 59 whole mitogenome sequences. J Mol Evol 63:826–841.

Sanger TJ, McCune AR. 2002. Comparative osteology of the *Danio* (Cyprinidae: Ostariophysi) axial skeleton with comments on *Danio* relationships based on molecules and morphology. Zool J Linn Soc 135:529–546.

Siebert DH. 1987. Interrelationships among families of the order Cypriniformes (Teleostei). Unpubl. Ph.D. dissertation. New York, NY: City University of New York.

Sikes DS, Lewis PO. 2001. PAUPRat: PAUP* implementation of the parsimony ratchet. Storrs, CT: Department of Ecology and Evolutionary Biology, University of Connecticut.

Sorenson MD. 1999. TreeRot, version 2. Boston, MA: Boston University.

Swofford DL. 2002. PAUP*. Phylogenetic analysis using parsimony (*and Other Methods). Version 4. Sunderland, MA: Sinauer Associates.

Talwar PK, Jhingran AG. 1991. Inland fishes of India and adjacent countries, vol 1. Rotterdam: A. A. Balkema.

Weber M, de Beaufort LF. 1916. The fishes of the Indo-Australian archipelago. III. Ostariophysi: II Cyprinoidea, Apodes, Synbranchi. Leiden: E. J. Brill.

Westerfield M. 1989. The zebrafish book: a guide for the laboratory use of zebrafish (*Brachydanio rerio*). Eugene, OR: University of Oregon Press.

Zardoya R, Abouheif E, Meyer A. 1996. Evolutionary analyses of hedgehog and *Hoxd*-10 genes in fish species closely related to the zebrafish. Proc Natl Acad Sci USA 93:13036–13041.