

## Patterns of Nucleotide Change in Mitochondrial Ribosomal RNA Genes and the Phylogeny of Piranhas

Guillermo Ortí,<sup>1,\*</sup> Paulo Petry,<sup>2</sup> Jorge I.R. Porto,<sup>2</sup> Michel Jégu,<sup>2,3</sup> Axel Meyer<sup>1,4</sup>

<sup>1</sup> Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, NY 11794-5245, USA

<sup>2</sup> Instituto Nacional de Pesquisas da Amazonia (INPA/CPBA), CxP 478, CEP 69. 083-000, Manaus, AM, Brazil

<sup>3</sup> ORSTOM, Département Des Eaux Continentales, 213 Rue La Fayette, 75480 Paris Cedex 10, France

<sup>4</sup> Program in Genetics, State University of New York at Stony Brook, Stony Brook, NY 11794, USA

Received: 12 June 1995 / Accepted: 18 October 1995

**Abstract.** The patterns and rates of nucleotide substitution in mitochondrial ribosomal RNA genes are described and applied in a phylogenetic analysis of fishes of the subfamily Serrasalminae (Teleostei, Characiformes, Characidae). Fragments of 345 bp of the 12S and 535 bp of the 16S genes were sequenced for 37 taxa representing all but three genera in the subfamily. Secondary-structure models based on comparative sequence analysis were derived to characterize the pattern of change among paired and unpaired nucleotides, forming stem and loop regions, respectively. Base compositional biases were in the direction of A-rich loops and G-rich stems. Ninety-five percent of substitutions in stem regions were compensatory mutations, suggesting that selection for maintenance of base pairing is strong and that independence among characters cannot be assumed in phylogenetic analyses of stem characters. The relative rate of nucleotide substitution was similar in both fragments sequenced but higher in loop than in stem regions. In both genes, C-T transitions were the most common type of change, and overall transitions outnumbered transversions by a factor of two in 16S and four in 12S. Phylogenetic analysis of the mitochondrial DNA sequences suggests that a clade formed by the genera *Piaractus*, *Colossoma*, and *Mylossoma* is the sister group to all other serrasalmines and that the genera *Myleus*, *Ser-*

*rasalmus*, and *Pristobrycon* are paraphyletic. A previous hypothesis concerning relationships for the serrasalmines, based on morphological evidence, is not supported by the molecular data. However, phylogenetic analysis of host-specific helminth parasites and cytogenetic data support the phylogeny of the Serrasalminae obtained in this study and provide evidence for coevolution between helminth parasites and their fish hosts.

**Key words:** Mitochondrial rRNA sequences — rRNA secondary structure — Molecular phylogeny — Serrasalminae — Coevolution

### Introduction

The studies of molecular systematics and molecular evolution are mutually informative (Mindell and Honeycutt 1990). Since molecular systematics relies largely on empirical results, increasing knowledge of patterns and rates of nucleotide change greatly improves our capacity for refining methods and assumptions used for phylogenetic inference and, in general, for deciding which molecular markers would be most appropriate for specific evolutionary questions. Analysis of comparative data in molecular evolution, however, takes place within a phylogenetic framework. In particular, secondary-structure models of RNA molecules have been based almost exclusively on comparative sequence analysis (reviewed in Guttel et al. 1985, 1993). Because within-molecule base

\* Present address: Department of Genetics, University of Georgia, Athens, GA 30602, USA

Correspondence to: G. Ortí

pairing determines RNA secondary structure (Noller 1984), evidence for positional covariance (or “compensatory mutations”) has been the primary tool for establishing the configuration of RNA structures. The reality of these structural elements has been confirmed by more direct, experimental approaches like X-ray crystallography for small tRNA molecules (e.g., Schimmel et al. 1979), by chemical protection and crosslinking, and by studying interactions between ribosomes and tRNAs, mainly for the large and small subunit RNAs of *E. coli* (reviewed in Gray and Cedergren 1993, and references therein).

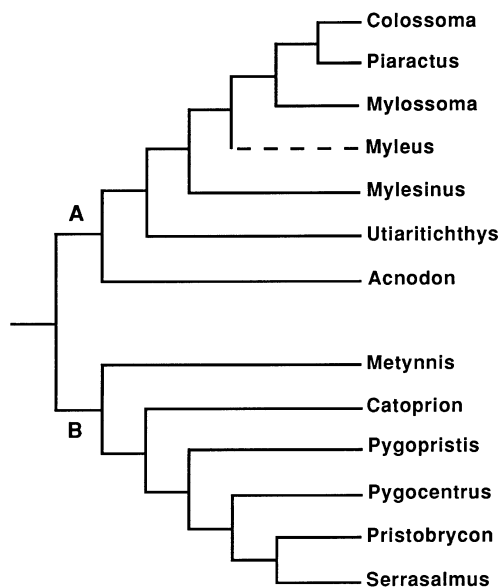
The use of ribosomal RNA sequences for phylogenetic studies has been extensively discussed (e.g., Hixon and Brown 1986; Wheeler and Honeycutt 1988; Larson and Wilson 1989; Mindell and Honeycutt 1990; Dixon and Hillis 1993; Vawter and Brown 1993; Olsen and Woese 1993). A recurring theme, and one of the most crucial issues in molecular systematics, is understanding what makes a molecule a suitable marker to trace organismal history. The evolutionary dynamics of each molecular species must be understood as well as possible, and this understanding provides the basis for phylogenetic analysis. In their review, Mindell and Honeycutt (1990) depicted a hypothetical curve showing the relationship between overall sequence divergence and estimated times of divergence as a heuristic tool for suggesting time intervals within which phylogenetic questions could be addressed using appropriate ribosomal RNA sequences. For the mitochondrial rRNA genes the recommended period of utilization was between 150 and 300 million years before present. Hillis and Dixon (1991), in contrast, suggested that mitochondrial rRNA genes are particularly useful for looking at relationships among groups that diverged no more than 65 million years ago. These estimates were based on a few pairwise comparisons of relatively distantly related taxa, but no detailed information on the evolutionary dynamics of the genes was offered. Relative rates of nucleotide substitution in the different molecules, across structural elements within the same molecule, and among different kinds of base substitutions (e.g., transition/transversion ratios), all constitute important information that should be incorporated into phylogenetic inferences. Vawter and Brown (1993) conducted a detailed analysis of the rate and patterns of base change in the small subunit rRNA gene among metazoans and found base compositional biases among different structural elements and lack of a consistent transition/transversion bias. They suggested how these findings could be applied for the refinement of phylogenetic analyses using rRNA sequences.

The effects of structural constraints of the ribosomal genes on phylogenetic analyses were also discussed in terms of character independence, a basic assumption of most phylogenetic methods (Wheeler and Honeycutt 1988; Swofford and Olsen 1990; Dixon and Hillis 1993). Because paired nucleotides forming stem regions appear

to undergo compensatory mutations that maintain secondary structure, Wheeler and Honeycutt (1988) suggested weighting stem positions by one-half for phylogenetic analysis. Dixon and Hillis (1993), analyzing data from the 28S ribosomal RNA gene of vertebrates, found that stem bases sustain a greater number of compensatory mutations than expected at random, but that this number is significantly smaller than what is expected under a hypothesis of perfect compensation. They therefore suggested a weighting value closer to one than to one-half for phylogenetic analyses of stem bases to remove the effect of nonindependence. The approach described by Dixon and Hillis (1993) could be used to investigate the most appropriate weighting scheme for different data sets.

Following the methods described by Vawter and Brown (1993) and Dixon and Hillis (1993), we present a detailed analysis of partial sequences of the large (16S) and small (12S) subunits of the mitochondrial ribosomal genes, which are frequently used in molecular systematics (e.g., Kocher et al. 1989; reviewed in Meyer 1993, 1994). Because most previous studies compared distantly related taxa, we focus on a relatively closely related group of taxa so that the rate and pattern of nucleotide change can be traced starting at the early stages of divergence. Subsequent studies might incorporate more distantly related taxa. We compare fish taxa thought to have diverged during the Tertiary and range from the intraspecific to the interfamilial level. For the ingroup, the subfamily Serrasalminae (family Characidae, order Characiformes), the earliest fossils known are from mid-to late-Tertiary deposits, some 30–13 million years old (Lundberg et al. 1986; Lundberg 1993).

The subfamily Serrasalminae (sensu Machado-Allison 1983) consists of at least 60 species (in 14 genera), including the well-known predatory “piranha” and the seed-eating “tambaquí” and “pacu.” These fish are endemic to the Neotropics and are widely distributed in all major river systems of South America. Several species are economically important, representing an important fraction of the commercial catch in the Amazon, and are also used for aquaculture (Junk 1984; Marshall 1995). In addition to their wide geographic distribution, serrasalmins occupy diverse habitats, from lowland flood plains, lakes, and flooded forests to rapids and waterfalls in the upper reaches of the rivers. Serrasalmins are characterized by a compressed body, long dorsal fin with more than 16 rays, and the presence of highly modified abdominal scales which give rise to a series of sharp serrae that vary in number from 6 to 9 in *Acnodon* to more than 60 in *Piaractus*. Three general feeding habits can be found among serrasalmins: carnivory, frugivory (fruits), and lepidophagy (scales) (Géry 1977). Carnivorous taxa typically have one row of sharp tricuspid teeth on each jaw, whereas frugivorous taxa have two series of incisor or molariform teeth on the pre-maxilla, one row of teeth on the dentaries, and commonly display one pair



**Fig. 1.** Phylogeny of the Serrasalminae proposed by Machado-Allison (1983).

of symphyseal teeth. Exclusively lepidophagous taxa have specialized tuberculated teeth that are used to remove scales from other fish and that are located on the outer side of the premaxilla. Some species, however, feed on all three types of food, depending on age or food availability (Goulding 1980; Nico and Taphorn 1988; Winemiller 1989; Leite and Jégu 1990).

In spite of their economic importance and being so notorious, the taxonomy and systematics of these fishes are relatively poorly known. Morphology and arrangement of the teeth have traditionally been the main characters used to classify serrasalmins. Eigenmann (1915) used dental characters to erect the subfamily Serrasalminae, comprising six genera with a single row of teeth on each jaw, and the subfamily Mylinae (9 genera) with two rows of teeth on the pre-maxillary which also included the lepidophagous monotypic genus *Catoprion*. Similarly, Norman (1929), Gosline (1951), and Géry (1977) based their classifications on dental morphology. Their hypotheses differed mostly on the relative rank of some of the taxa, which changed from generic to subgeneric status. Machado-Allison (1982) was the first to apply cladistic methods to serrasalmin systematics (Fig. 1). He defined the subfamily as a monophyletic unit based on 27 morphological synapomorphies, and subdivided it into two clades. His "lineage A" contained all the pacu (most of the Mylinae of Eigenmann) and "lineage B" contained all piranhas (Serrasalminae of Eigenmann) plus the genera *Catoprion* and *Metynnis* (Fig. 1), but he emphasized the need to further explore generic limits within lineage "A." Although he considered *Utiaritchthys* and *Mylesinus* valid genera, he stated that both taxa might be specialized forms of *Myleus*, which would be a paraphyletic genus as presently constituted. The generic status of taxa within lineage "B" was strongly supported

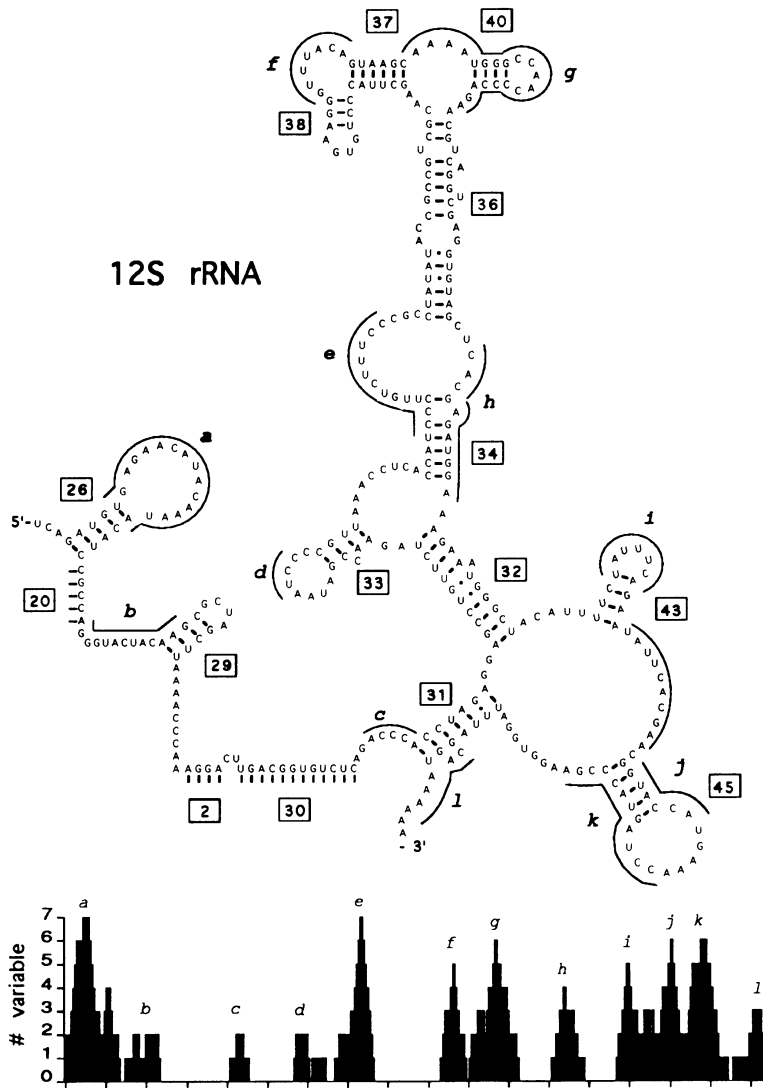
by several synapomorphies. In this paper we provide a test for the phylogeny proposed by Machado-Allison (1982) based on an independent data set.

## Materials and Methods

**Fish Specimens and DNA Sequences.** A total of 37 specimens representing most genera included in the subfamily Serrasalminae (except *Pygopristis*, *Ossubtus*, and *Utiaritchthys*) and three characid outgroups were used for this study. Outgroup designation was based on a preliminary analysis of mitochondrial DNA (mtDNA) sequences of several taxa in the family Characidae and other characiform families (Ortí and Meyer 1996). Each sample was assigned a number (from 1 to 37). Species identifications, localities of origin, museum collection numbers, GenBank accession numbers, and other relevant information are given in Appendix 1. Fish tissues for DNA extraction were preserved in 70% ethanol. Genomic DNA was extracted from muscle tissue by Proteinase K/SDS dissolution and purified by phenol-chloroform extraction and ethanol precipitation (Maniatis et al. 1982; Kocher et al. 1989). The polymerase chain reaction (PCR, Saiki et al. 1988) was used to amplify segments of the small (12S) and large (16S) subunit ribosomal RNA mitochondrial genes. Double-stranded amplifications were performed in 25- $\mu$ l volumes containing 67 mM Tris (pH 8.8), 6.7 mM MgCl<sub>2</sub>, 16.6 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM 2-mercaptoethanol, 0.5 mM of each dNTP, 1  $\mu$ M of each primer, 10–1,000 ng genomic DNA, and 0.5 units of *Taq* Polymerase (Perkin-Elmer-Cetus). The following primers were used: for 12S, L1091 and H1478 (Kocher et al. 1989); for 16S, 16Sar-L and 16Sbr-H (Palumbi et al. 1991). These primers amplify fragments of the 12S and 16S mitochondrial rRNA genes corresponding to positions 1091–1478 and 2510–3059 in the human mitochondrial genome, respectively (Anderson et al. 1981). Gel purification of double-stranded products was followed by generation of single-stranded DNA for direct sequencing from both directions (Gyllensten and Erlich 1988), obtained by asymmetric PCR, using the protocol described in Kocher et al. (1989). Single-stranded DNA was concentrated and desalted in Millipore Ultrafree-MC filters and sequenced by the dideoxy chain-termination method (Sanger et al. 1977) with the limiting primer from the asymmetric PCR amplification, using a commercial kit (Sequenase Version 2, Unites States Biochemical).

**Sequence Alignment and Phylogenetic analyses.** The orthologous DNA sequences obtained were aligned using CLUSTAL W version 1.4 with default settings (Thompson et al. 1994). Computer-generated alignments were compared to secondary-structure models described by Guttel et al. (1985), Dams et al. (1988), and Neefs et al. (1991) for the small subunit rRNA and by Guttel and Fox (1988) and Guttel et al. (1993) for the large subunit rRNA molecule. Based on these models, the partial ribosomal gene sequences obtained for serrasalmins were folded to secondary-structure format by eye. When structural elements similar to those of the published models could not be formed with the fish sequences, the most consistent secondary structure that included compensatory nucleotide changes shared by most sequences was chosen. Some putative stem regions could not be confirmed for our model since the complementary strands of some helices were not sequenced. Alignment gaps placed by CLUSTAL W that disrupted paired regions were moved to contiguous unpaired regions, as defined in our structural model.

Phylogenetic inference by maximum parsimony analyses was performed using PAUP version 3.1.1 (Swofford 1993) and MacCLADE version 3 (Maddison and Maddison 1992); maximum likelihood analyses (Felsenstein 1981) using FASTDNAML version 1.0.8 (Olsen et al. 1994) and NUCML (Adachi and Hasegawa 1994), and neighbor-joining (Saitou and Nei 1987) using MEGA version 1.0 (Kumar et al.



**Fig. 2.** Secondary-structure model for the 12S mitochondrial rRNA fragment of *Pygocentrus nattereri*. Helix numbering system follows Neefs et al. (1991, numbers in boxes). Watson-Crick pairings (C-G and A-U) are identified by *dashes* and G-U pairs by a *point*. The graph at the bottom is a sliding-window analysis of variability among the 33 sequences determined for the present study (window size = 7, offset by one); the vertical axis shows the number of variable sites in the window and the horizontal axis the position along the sequence. Lowercase letters (*a-l*) map the variable regions onto the secondary-structure model.

1993). Parameters and settings used with these programs are reported with the resulting topologies. Bootstrapping (Felsenstein 1985) was used to estimate confidence in the results.

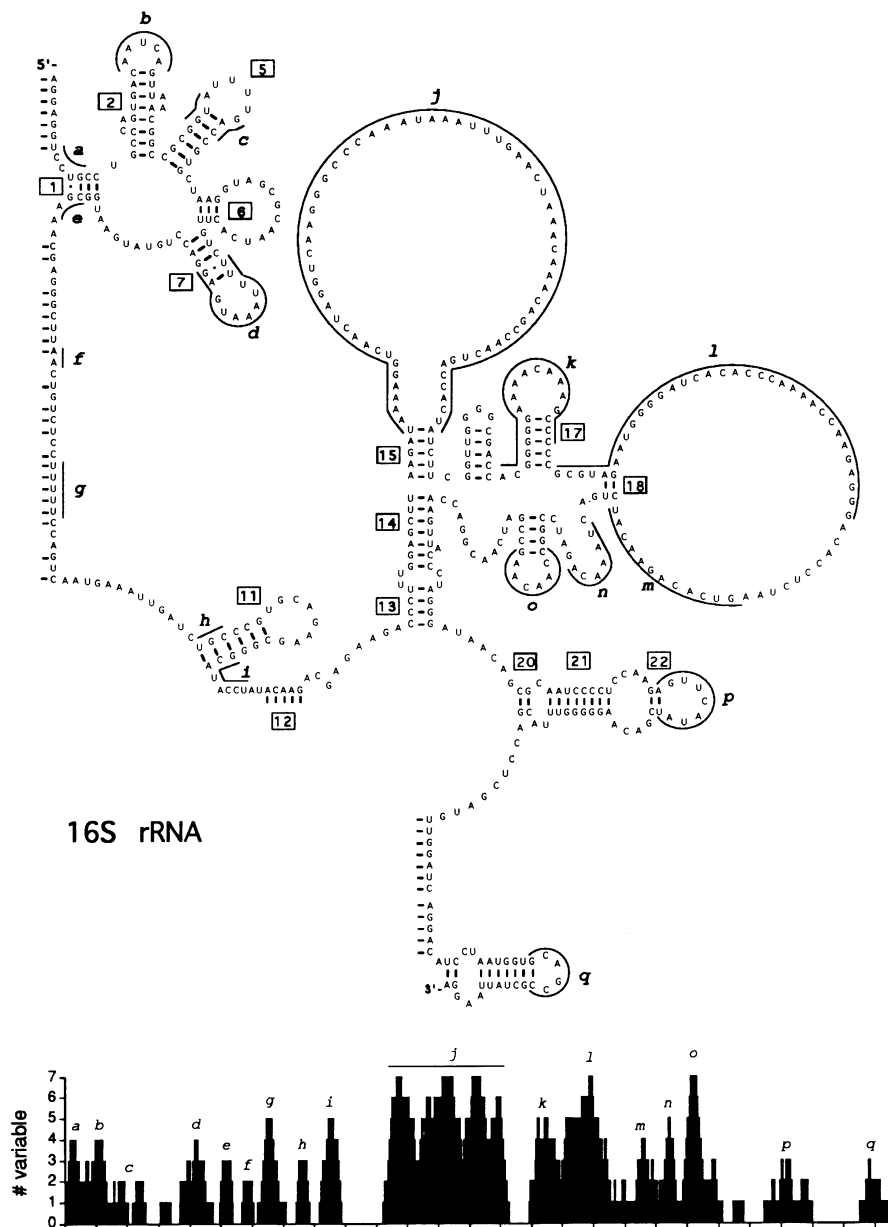
The computation of nucleotide substitutions among sequences was done by tracing them on the best tree obtained by the above methods rather than by pairwise comparisons (Fitch and Markowitz 1970; but see Collins et al. 1994). This approach provides an estimate of the changes that occurred across the phylogeny, under the assumption of maximum parsimony (Swofford and Maddison 1992). Average changes of character state were computed using MacCLADE (Maddison and Maddison 1992). Based on the secondary-structure models, nucleotide composition and substitution rates for stem and loop regions were also calculated separately, following the method outlined in Vawter and Brown (1993). Relative rates of change for each structural category and for each gene fragment were calculated with and without corrections for category size. Relative rates of each kind of nucleotide substitution (e.g., A $\leftrightarrow$ G, A $\leftrightarrow$ C, A $\leftrightarrow$ T, etc.) for each structural category and for each gene fragment were corrected for base composition. Because loops for example, are AC-rich, more changes between A and C than between any other pair of nucleotides are expected to occur there. Correction factors were defined by adding percent composition for the category of both bases involved in the change and dividing the sum by 50%. The observed number of changes was then divided by this factor, so that if there is compositional bias in favor of a pair of

nucleotides, division by a number larger than unity corrects for this overrepresentation (Vawter and Brown 1993).

For stem regions, a tally of the changes in paired nucleotides (single changes and double changes) that maintain and disrupt the pairing was used for comparison with expected values, following the method of Dixon and Hillis (1993).

## Results and Discussion

A fragment of approximately 345 bp from the 3' half of the 12S mitochondrial ribosomal gene was sequenced for all taxa. The proposed secondary model for the piranha (*Pygocentrus nattereri*) 12S fragment is shown in Fig. 2. A fragment of approximately 530 bp for the 16S mitochondrial ribosomal gene was obtained, and the corresponding piranha secondary-structure model is shown in Fig. 3. The computer-generated alignment (with CLUSTAL W) of a single file containing the 12S and the 16S sequences arranged sequentially required a total of 10–16 gaps per sequence (1.1–1.8% of the aligned se-



**Fig. 3.** Secondary-structure model for the 16S mitochondrial rRNA fragment of *Pygocentrus nattereri*. Numbers in boxes identify helix regions discussed in the text. The graph at the bottom is the same as in Fig. 2.

quence length). Most gaps were of length = 1 nucleotide, but a few indels (insertion/deletion events) comprised up to six nucleotides. The aligned sequences are in Ortí (1995). Indels assigned by CLUSTAL W to a stem region of the 12S sequence (Fig. 2, stem labeled 26) were moved to the contiguous hairpin loop (Fig. 2, loop labeled ‘a’).

Sequence divergence among taxa, for the 12S and 16S data combined, is summarized in Table 1. Variation among species of a genus ranged from 0 to 5.8%, reflecting in part the uneven representation of species among genera. Three *Pygocentrus nattereri* samples (numbered 1–3 in Appendix 1) yielded identical sequences for the two ribosomal fragments, as did samples 6 and 7 (both *Serrasalmus* specimens) and samples 11 and 12 (*Catoprion mento*). Samples 21 and 22 (*Mylesi-*

*paraschomburkii*) had identical 12S sequences but differed by a single A to G transition in their 16S sequences. In contrast, sequence difference was greatest within the genera *Pristobrycon* and *Metynnis* (Table 1). Phylogenetic analysis (see below), morphological evidence (Machado-Allison et al. 1989), and parasitology (Van Every and Kritsky 1992) suggest that, as currently defined, *Pristobrycon* may be a paraphyletic unit. No such evidence is available to question the monophyletic status of *Metynnis*, in spite of significant sequence differences (5.4% between samples 13 and 14), which includes a 6-bp deletion within one of the hairpin loops of the 16S sequence (Fig. 3, loop labeled ‘j’). Sequence divergence among serrasalmin genera was below 8.9%, whereas variation between the serrasalmins and the out-group taxa ranged from 9.9 to 17.3%.

**Table 1.** Percent sequence difference (range) among DNA sequences (12S and 16S combined), within and between taxa

Taxa	Within	Between											
		2	3	4	5	6	7	8	9	10	11	12	13
1 <i>Pygocentrus</i>	0.3	0.9–1.7	1.5–5.8	5.1–5.3	5.5–7.4	5.1–6.7	5.9–6.1	5.8–6.0	8.6–9.0	6.9–7.5	6.6–7.0	6.2–6.9	10.9–16.8
2 <i>Serrasalmus</i>	1.7	—	1.4–5.6	5.1–5.6	5.9–7.7	5.3–7.3	5.5–6.5	5.3–6.5	8.2–9.1	6.9–8.3	6.8–7.8	6.5–7.4	10.6–16.8
3 <i>Pristobrycon</i>	0.2–5.8	—	—	5.0–5.7	5.8–7.6	5.0–7.0	5.4–5.8	5.4–5.9	8.0–9.3	6.8–8.1	6.6–7.6	6.4–7.2	11.0–16.6
4 <i>Catoptrion</i>	0.0	—	—	—	5.7–6.7	5.1–6.9	6.0–6.1	5.4	8.5–8.7	7.6–8.1	7.2–7.4	7.0–7.4	11.1–17.3
5 <i>Metynnis</i>	5.4	—	—	—	—	4.9–8.4	5.3–6.9	5.3–7.3	8.1–9.8	6.5–8.7	6.7–8.4	5.8–7.8	10.2–16.2
6 <i>Myleus</i>	0.3–3.2	—	—	—	—	—	1.1–3.9	0.8–2.1	6.6–7.9	6.0–8.2	5.5–7.2	5.4–7.5	9.7–16.3
7 <i>Mylesinus</i>	0.1	—	—	—	—	—	—	1.2–1.4	6.6–7.0	6.0–7.1	5.9–6.2	5.2–5.9	10.1–16.3
8 N. gen. A	—	—	—	—	—	—	—	—	6.7–7.0	6.1–6.6	5.4–5.6	5.6–6.0	9.9–16.3
9 <i>Acnodon</i>	0.3	—	—	—	—	—	—	—	—	8.3–8.9	8.3–8.9	7.3–7.9	11.9–16.9
10 <i>Mylossoma</i>	0.9–3.2	—	—	—	—	—	—	—	—	—	3.5–4.2	3.1–4.4	8.4–16.5
11 <i>Colossoma</i>	0.2	—	—	—	—	—	—	—	—	—	—	2.8–3.5	8.8–16.0
12 <i>Piaractus</i>	0.1–0.9	—	—	—	—	—	—	—	—	—	—	—	8.5–15.5
13 “Outgroups”	10.1–15.5	—	—	—	—	—	—	—	—	—	—	—	—

### Secondary-Structure Models and Selection for Compensatory Mutations

The proposed secondary structure for the piranha 12S fragment (Fig. 2) shows remarkable conservation with those of other mitochondrial 12S RNAs of vertebrates (Guttel et al. 1985; Neefs et al. 1991; Alves-Gomes et al. 1995). Variation among sequences occurs mainly, but not exclusively, in loop regions (Fig. 2). Among stem regions, helix 26 (part of region “a”) is the most variable, with the highest proportion of noncompensated mutations. Length of this helix varies from 4 to 6 bp among taxa. In addition to helix 26, length variation is present in the regions identified as “d,” “f,” “g,” “i,” “j,” and “k” in Fig. 2. Helix 39 varies the most among deuterostomes (Neefs et al. 1991), and as in sea urchins, *Xenopus*, and artiodactyls, it could not be identified in the piranha model (Fig. 2).

The 16S sequences were also folded into a secondary structure (Fig. 3), following the model proposed by Guttel and Fox (1988) for *Xenopus*. As in 12S, most of the observed variation occurs in loop regions, particularly in the large hairpin-loop labeled “j” in Fig. 3. Some variation in stem regions is also present, most noticeably in helices number 5, 7, 11, 17, and 18 (Fig. 3), but no variation in length by disruption of paired bases was observed. Length variation in the aligned 16S sequences occurs in regions “j,” “l,” “n,” and “o” shown in Fig. 3. For all subsequent analyses, nucleotides were assigned to stem regions only if they were paired in the secondary-structure models (nucleotides in “bulges” are not included), and to loop regions if they were shown forming hairpin-loops, multibranch loops, or internal loops in stem regions (single-stranded unpaired regions are not included). Structural elements are defined as in Vawter and Brown (1993). A synopsis of the data assigned to stems and loops is given in Table 2.

The observed and expected numbers of compensatory mutations in stem regions, following Dixon and Hillis (1993), are shown in Table 3. Most substitutions do not

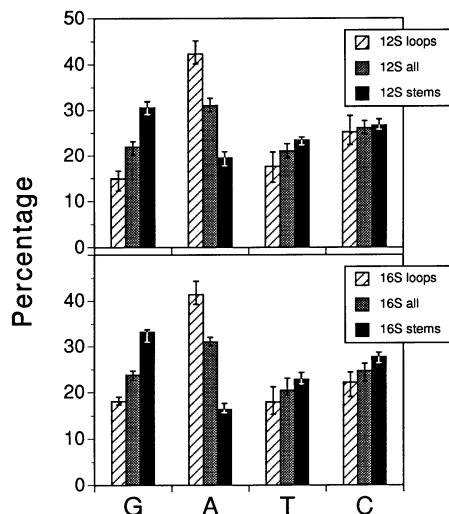
**Table 2.** Summary of sequence data from the 12S and 16S rRNA genes (based on the structural models presented in Figs. 2 and 3)

	Total no. of aligned nucleotides	No. of variable sites	No. of phylogenetically informative sites
12S stems	163	19	15
12S loops	173	62	46
12S all	348	84	63
16S stems	216	41	18
16S loops	245	111	66
16S all	539	177	101
Total (12S + 16S)	887	261	164

**Table 3.** Substitutions observed in stem regions of the 12S and 16S sequence data (based on the secondary-structure models shown in Figs. 2 and 3)

Type of substitution	No. expected	No. observed
Single:		
Base pairing to base pairing	7	38
Base pairing to non-base pairing	49	18
Double:		
Base pairing to base pairing	4.6	17
Base pairing to non-base pairing	13.4	1

disrupt pairings in stem regions, and there are significantly more compensatory mutations than expected by chance ( $\chi^2 = 156.9$  and  $\chi^2 = 44.9$ ,  $df = 1$ ,  $P < 0.001$ , for single and double substitutions, respectively). A total of 92 substitutions were reconstructed for stem regions of the 12S and 16S sequences, of which 72 represented compensatory changes, compared to 16.2 expected by random mutation alone. These results suggest that selection for compensatory substitutions is strong and that a weighting close to one-half should be used for phylogenetic analysis of stem sequences to account for the lack of independence among substitutions. The exact weight-

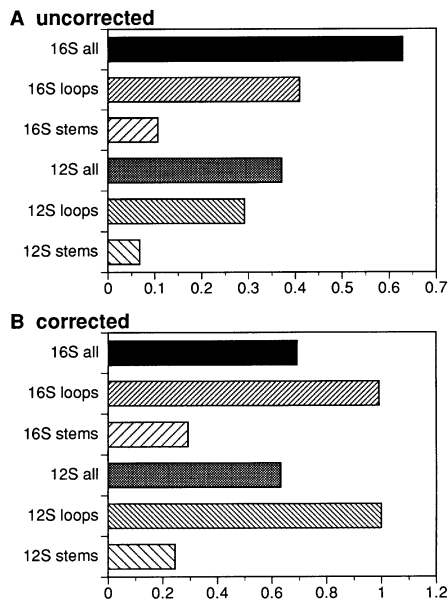


**Fig. 4.** Mean values and ranges of the percent base composition for stems, loops, and the whole fragment of the 12S and 16S sequences.

ing for stem sequences suggested by the method of Dixon and Hillis (1993) should be 0.63. This value is intermediate between what Wheeler and Honeycutt (1988) suggested for 5S and 5.8S rRNA molecules and the weight of 0.8 suggested by Dixon and Hillis (1993) for 28S rRNA. The effects of differential weighting of stem substitutions for the phylogenetic results of the current study are discussed below.

#### Base Composition and Pattern of Nucleotide Change

Base compositions for the different structural categories of the 12S and 16S sequences are shown in Fig. 4. Overall, both gene fragments show very similar base composition, with a slight overrepresentation of A and C, as was previously observed in mitochondrial genomes of fish (e.g., Kocher et al. 1989; Tzeng et al. 1992; Zardoya et al. 1995; reviewed in Meyer 1993). No significant differences in base composition among taxa were observed. However, stem and loop regions differ markedly in their content of A and G, but not of T and C (Fig. 4). Bias in favor of G and C was previously observed in paired regions of RNA (Guttel et al. 1985; Vawter and Brown 1993) and attributed to selection for the more stable conformation of the G-C interaction, compared to the A-U or G-U pairs (reviewed in Turner et al. 1988). Bias in favor of A in unpaired (or loop) regions (Fig. 4) was also previously reported for a vast diversity of RNA molecules (e.g., Guttel et al. 1985). Hydrophobic interaction with proteins was suggested by Guttel et al. (1985) as a possible role for unpaired adenines, since adenine is the least polar of the four bases, and has been observed to occur in protein binding sites. An alternative explanation for the abundance of A and the underrepresentation of G in loop regions (in mitochondrial rRNA) might be found in the higher sequence variability and rate of change in loops (Fig. 5). Because of low structural con-

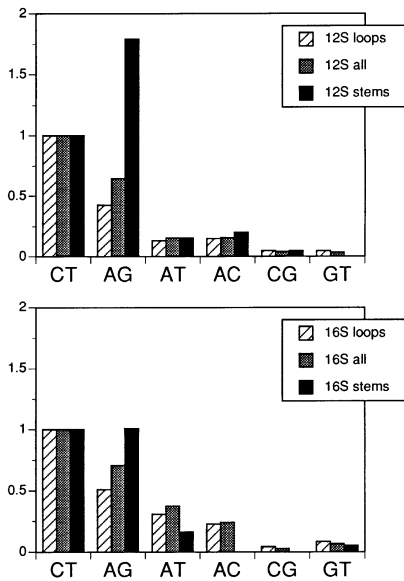


**Fig. 5.** Rates of nucleotide substitution in the 12S and 16S sequences. The number of changes in each category was divided by the total number of changes observed in panel A, without correcting for category size. In panel B, the number of changes observed in each category was divided by category size, showing relative rates of change among categories.

straint in loop sites, underlying mutational biases might be responsible for the observed biased base composition. Silent sites at protein-coding mitochondrial genes are likewise A-rich and G-poor, presumably reflecting a mutational bias in vertebrates (Lee and Kocher 1995).

Unlike the small subunit nuclear rRNA investigated by Vawter and Brown (1993), loop regions evolve at a faster rate than stems (at least four times faster) in both mitochondrial rRNA subunits (Fig. 5). Therefore most of the phylogenetically informative sites are in unpaired regions of both mitochondrial genes (Table 2). Also unlike previous reports (e.g., Lee and Kocher 1995), the 12S and 16S mitochondrial genes seem to be evolving at a similar rate (Fig. 5). Although small sample size in the present study might account for this discrepancy, since Lee and Kocher (1995) compared the entire ribosomal sequences, it may also be possible that similar rates of change between recently diverged sequences may not be detected when more distantly related sequences are compared (e.g., Lee and Kocher compared sequences from all classes of vertebrates).

A bias in favor of transitions over transversions has been found in all genes of the mitochondrial genomes so far examined (e.g., Brown et al. 1982; Aquadro and Greenberg 1983; Wolstenholme and Clary 1985; reviewed in Meyer 1993). This bias is more pronounced among least diverged sequences, where transitions may outnumber transversions by a factor of ten, but it drops among distantly related sequences to a factor of about two. Relative rates of substitutions observed among the serrasalmid sequences are as expected (Fig. 6); transi-

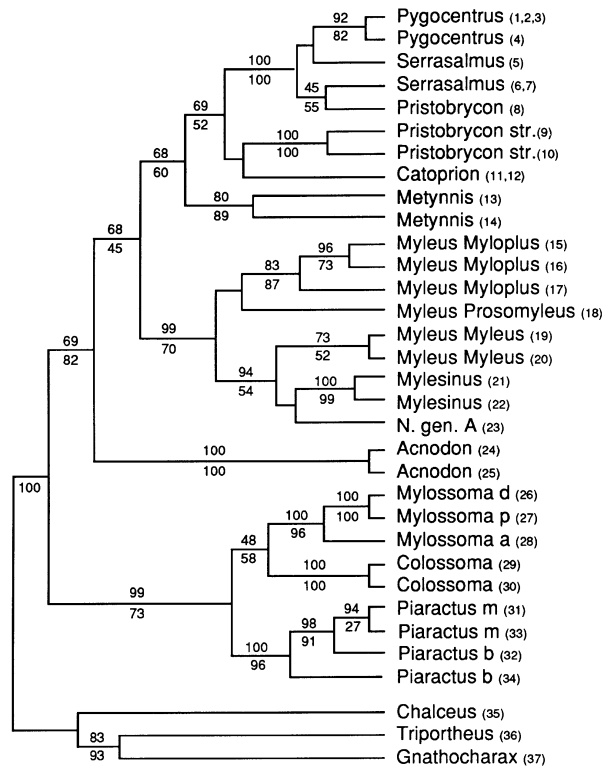


**Fig. 6.** Relative rates of the different types of base change (corrected for base composition) in the 12S and 16S sequences. For each gene and structural category, rates are expressed as a proportion of the number of C $\leftrightarrow$ T changes.

tions are the most frequent kind of substitution in all structural categories. However, the proportion of A $\leftrightarrow$ G and C $\leftrightarrow$ T transition differs among structural classes, especially in the 12S gene. In both genes, C $\leftrightarrow$ T transitions are the most common type of change in loop regions and overall, but A $\leftrightarrow$ G transitions are more abundant among stem substitutions, especially in the 12S sequences. This elevated rate of A $\leftrightarrow$ G transitions in stems has been postulated to be consistent with selection for maintenance of base pairing in stem regions (Vawter and Brown 1993), since in RNA the G-U pair is stable. Unlike the small subunit nuclear rRNA genes analyzed by Vawter and Brown (1993), where the A $\leftrightarrow$ G transition is the least common change in loops, in the mitochondrial 12S and 16S genes in serrasalmians the A $\leftrightarrow$ G transformation outnumber all kinds of transversions in loops (Fig. 6). Overall, transitions in the 12S and 16S sequences outnumber transversions by a factor of two and four, respectively.

#### Phylogenetic Analysis

All three methods of phylogenetic inference resulted in largely congruent topologies, differing only in minor rearrangements. The best tree obtained by maximum likelihood analysis is shown in Fig. 7. It was obtained using the FASTDNAML program (Olsen et al. 1994) with default parameters, global rearrangements, and the jumble options, five times out of five runs. The transition/transversion ratio used was 2.0, but a value of 4.0 gave identical results. All branch lengths were significantly greater than zero ( $P < 0.01$ ), except some terminal branches, and those grouping *Serrasalmus* (6, 7) and *Pristobrycon* (8), and the *Piaractus* taxa (31, 32, 33). A



**Fig. 7.** Phylogenetic tree obtained by maximum likelihood (FASTDNAML). Bootstrap values for neighbor-joining and parsimony analyses are shown above and below each branch, respectively. Branches without numbers received bootstrap values  $< 50$ , or were not supported by neighbor joining or parsimony.

comparison of this tree topology with that shown in Fig. 1 resulted in statistically significant differences, using the method of Kishino and Hasegawa (1989) with the program NUCML (Adachi and Hasegawa 1994). The Ln likelihood of the best tree (Fig. 7) was  $-4,435.1$  and the differences in Ln likelihood ( $\pm$  standard error) between topologies was  $-119 \pm 26$ .

Neighbor-joining analysis was done using Kimura's (1981) genetic distances, which take into account different rates for transitions and transversions, as seemed appropriate for this data set (see above). The result was a tree that differs from the one shown in Fig. 7 only by the placement of *Catoprion* as the sister group of the *Pygocentrus* + *Serrasalmus* + *Pristobrycon* clade, rather than being the sister group of *Pristobrycon striolatus* (9, 10). Bootstrap support for all branches is high (Fig. 7, values above branches), except for those grouping *Serrasalmus* (5) with *Pygocentrus* (1–4) and *Serrasalmus* (6, 7) with *Pristobrycon* (8). Some of the branches within the “Myleus” clade (clade grouping *Myleus*, *Mylesinus*, and N. gen. A) also receive low bootstrap support (Fig. 7).

Parsimony analysis using different weights for transitions and transversions also gave results mostly congruent with the topology shown in Fig. 7. When all characters were uniformly weighted, three equally short trees were found (heuristic search with ten replications using random addition of taxa,  $L = 642$ , consistency index =



0.51, excluding uninformative characters). These trees differed from each other only in the placement of the *Piaractus* species (31, 32, 33), but differed from the one shown in Fig. 7 in: (1) the inclusion of *Acnodon* as the sister group of the “Myleus” clade; (2) relationship within the “Myleus” clade; (3) the position of *Metynnix* and the *Catoprion* + *Pristobrycon striolatus* clade were reversed. However, the 50% majority-rule consensus tree resulting from 100 bootstrap replications (with heuristic search with three replications using random addition of taxa) is congruent with the tree shown in Fig. 7. Resulting bootstrap values are also shown in Fig. 7 (below the branches). A posteriori reweighting (or “successive approximations,” Farris 1969) of the uniformly weighted data set also resulted in three trees that differed in the positions of the *Piaractus* species (with respect to each other), but in which *Acnodon* is placed as shown in Fig. 7. Weighting schemes in which transversions were counted two and four times as much as transitions also gave congruent results and similar bootstrap values to those shown in Fig. 7. The topology shown in Fig. 1 required 705 steps (63 extra steps) when enforced on the molecular data set.

When only stem characters (33 phylogenetically informative sites) were used, a total of 842 equally parsimonious trees were recovered, a strict consensus of which still showed serrasalmin monophyly, and the basal *Piaractus* + *Colossoma* + *Mylossoma* clade. The small number of informative characters in stem regions is clearly not sufficient to solve relationships among most taxa, but it is still able to recover the most basal nodes of the tree shown in Fig. 7. This suggests that the more slowly evolving stem sites did not accumulate enough phylogenetic information to resolve recent divergences. Since most of the resolution is provided by the loop and “other” sites in the rRNA sequences, downweighting stem changes because of nonindependence is irrelevant in the context of the present phylogenetic analysis. Loop sites alone (418 sites), however, result in 18 parsimonious trees, a strict consensus of which has most of the structure shown in Fig. 7. Loop changes alone, however, fail to recover the monophyly of the *Piaractus* + *Colossoma* + *Mylossoma* clade and result in a trichotomous resolution among *Acnodon*, the “Myleus” clade, and the “piranha” clade. A combination of the slowly evolving stem regions and the fast-evolving loop regions results in the complete resolution of relationships, as depicted in Fig. 7.

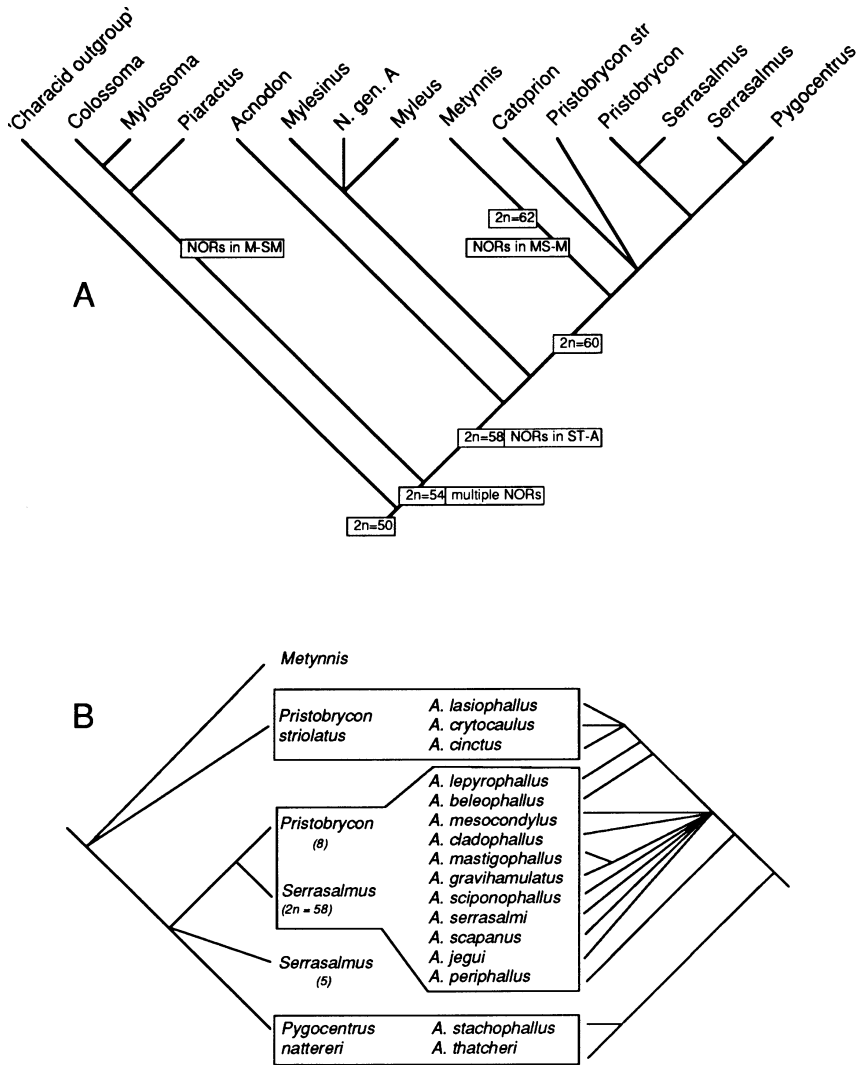
#### *Systematic and Taxonomic Implications*

The 12S and 16S molecular data set herein analyzed, using three different methods of phylogenetic inference, results in a robust estimate of serrasalmin relationships (Fig. 7) that does not support a previous hypothesis based on morphological data (Fig. 1, Machado-Allison 1982,

1983). Division of the subfamily Serrasalminae into two groups (lineages A and B, Fig. 1) is not supported by mitochondrial DNA sequences. Instead, three major groups are defined. The earliest divergence within the serrasalmins gave rise to the well-defined group of plant (leaves, fruits, and seeds)-eating fishes which lack a predorsal spine (*Piaractus*, *Colossoma*, and *Mylossoma*), a condition known to be plesiomorphic for characid fishes. According to the hypothesis in Fig. 1, absence of the predorsal spine in this group is interpreted as a secondary loss (Machado-Allison 1983). Another character with similar distribution is an opening in the anterolateral musculature of the body wall, forming a “humeral hiatus.” This hiatus, present in all serrasalmins except these three genera, was considered a synapomorphy for the subfamily and a reversal for the three genera by Machado-Allison (1983). Our hypothesis suggests that the absence of the hiatus in these three genera is best interpreted as a plesiomorphic condition for serrasalmins. Although *Mylossoma* and *Colossoma* are shown as sister groups to the exclusion of *Piaractus*, this relationship does not receive significant bootstrap support (Fig. 7). *Piaractus* is considered a subgenus of *Colossoma* in several previous classifications (e.g., Norman 1929; Gosline 1951; Géry 1972, 1977), but not in others (e.g., Eigenmann 1915; Nelson 1961; Machado-Allison 1982). The basal position of this group among the Serrasalminae was also advocated by Gosline (1951), based on dental, fin, and osteological characters, and by Nelson (1961), based on swim bladder morphology and other internal anatomical features.

The genus *Acnodon* forms the sister group to the other two clades. The genera *Myleus*, *Mylesinus*, and N. gen. A form a group which receives good support from the molecular data. According to Jégu et al. (1992), the specimens assigned by Machado-Allison (1982, 1983) to *Utariitchthys* (Fig. 1) do not belong to that genus. The taxon N. gen. A includes specimens that could not be assigned to *Utariitchthys* or to any other serrasalmin genus (Jégu et al. 1992), and its proper definition is contingent on a complete revision of the genus *Myleus* (Jégu, unpublished data). The only difference between *Utariitchthys* and *Myleus* is the extreme elongation of the body in the former, and Jégu et al. (1992) suggested a sister-group relationship of these two genera. The genus *Myleus*, as currently defined, seems to be paraphyletic (Fig. 7), since the taxa in the subgenus *Myleus* seem to be more closely related to *Mylesinus* and N. gen. A than to other subgenera of *Myleus* (i.e., *M. Prosomeyus* and *M. Myloplus*). An exhaustive revisionary study of this clade seems warranted.

The third group contains the “true piranhas,” predominantly carnivorous fish with one row of sharp tricuspid teeth on each jaw (*Serrasalmus*, *Pygocentrus*, *Pristobrycon*, *Catoprion*, and—missing in this study—*Pygopristis*), and the plant-eating *Metynnix*. Although different relationships among the taxa in this group were



**Fig. 8. A** Chromosomal characters optimized onto the phylogeny produced by molecular data. *NOR*: nucleolar organizer region;  $2n$ : diploid number; *M*: metacentric chromosome; *SM*: submetacentric; *ST*: subtelo-centric; *A*: acrocentric. **B** The topology shown on the left is our hypothesis for piranha relationships. In the same box with the piranha species are the *Acanthurus* species known to be their parasites, and to the left is the parasite cladogram from Van Every and Kristsky (1992).

proposed by Machado-Allison (1983), his lineage B (Fig. 1) has the same taxonomic composition. The sister-group relationship of *Metynnus* to the “true piranhas” seems well supported by the molecular data and by the morphological evidence as well. The monophyly of *Serrasalmus* and *Pristobrycon* is not supported by our data. Most noteworthy, *Pristobrycon striolatus* (9, 10) is very different from the other *Pristobrycon* (8) analyzed in the present study. Sequence divergence between these two taxa is 5.8%, as large as differences between genera from the “Myleus” and the “piranha” clades (Table 1). *Pristobrycon* (8) groups with *Serrasalmus* and *Pygocentrus* species with very high bootstrap support (Fig. 7), to the exclusion of *P. striolatus* (9, 10). This result agrees with observations by Machado-Allison et al. (1989), who have also suggested the paraphyly of *Pristobrycon*, since *Pristobrycon striolatus* lacks a synapomorphic character (the preanal spine) present in all other *Pristobrycon* species and in *Serrasalmus* and *Pygocentrus*. The four specimens of *Pygocentrus nattereri* included in this study form a well-supported monophyletic group, in agreement

with a revision by Fink (1993). However, relationships among the piranha genera in this clade remain somewhat unresolved because relatively few informative characters are available. A molecular marker with higher rate of evolution might be more appropriate than the mitochondrial rRNA genes to investigate this problem further.

#### *Congruence with Chromosome Number and Coevolution of Parasites*

Evidence from other data sets seems to corroborate our phylogenetic findings. Cytological information can be readily mapped onto the hypothesis of relationships among serrasalmids proposed here. Most characid fishes have a chromosome number of  $2n = 24$  to  $2n = 52$ , with a mode of  $2n = 50$  (Oliveira et al. 1988; Porto et al. 1992), and a few or a single nucleolar organizer region (NOR). The presence of multiple NORs is a synapomorphic character for the subfamily (Porto et al. 1992), but the position of these NORs varies among taxa. Serra-

salmins have a higher chromosome number, ranging from  $2n = 54$  exclusively bi-armed chromosomes in the *Colossoma* + *Piaractus* + *Mylossoma* clade to  $2n = 62$  in *Metynnys* (Porto et al. 1991; Cestari and Galetti 1992). The distribution of chromosome numbers and the position of NORs for the different taxa, given in Porto et al. (1989, 1991) and Cestari and Galetti (1992), were mapped onto the phylogenetic tree obtained in this study and shown in Fig. 8A. There is a clear evolutionary trend for chromosome numbers to increase during the evolution of serrasalmins, according to the phylogeny proposed in this study. The genus *Metynnys* exhibits a unique karyotypic pattern, with  $2n = 62$ , 1–2 subtelocentric pairs, and 4 NORs in bi-armed chromosomes. Not shown in Fig. 8 is the autapomorphic reduction in chromosome number of one of the species of *Serrasalmus* ( $2n = 58$ ).

Studying morphological characters of the helminth parasites (*Anacanthorus*, Monogenoidea) of piranhas, Van Every and Kritsky (1992) proposed a cladogram for the parasites that is congruent with the relationships among the piranhas as depicted in Fig. 7 for the genera *Serrasalmus*, *Pristobrycon*, and *Pygocentrus*. A partial representation of the *Anacanthorus* cladogram is shown in Fig. 8B, showing parasite taxa next to piranha species known to be their hosts. A remarkable similarity between the host and parasite cladograms supports the paraphyly of *Serrasalmus* and *Pristobrycon*, strongly suggesting a remarkable case of coevolution (cospeciation) between these fish and their parasites. But the implied polarity of transformations in the parasite cladograms is reversed. Helminth taxa in the most basal branches of the parasite cladogram are associated with piranha species shown to be in the most derived position on the molecular fish cladogram. This reversed pattern of relationships would be hard to explain in terms of parasite–host coevolutionary scenarios. We therefore conducted a reanalysis of Van Every and Kritsky's (1992) data set and found that rerooting their cladogram, such that it matches our fish phylogeny, only required a single extra step. When comparing their tree (shown in Fig. 8B) and the rerooted tree, only two characters favored their hypothesis and one character the alternative, rerooted tree. Apparently more data need to be collected in order to confidently solve relationships among *Anacanthorus* species. However, extensive coevolution between serrasalmins and their parasites is indeed suggested by a more recent (but still unpublished) study of *Anacanthorus* parasites. Parasites of *Colossoma*, *Piaractus*, and *Mylossoma* were found to occupy a basal position in the helminth phylogeny (W.A. Boeger and D. Kritsky, pers. comm.), in agreement with our results based on DNA sequences of their serrasalmin hosts.

**Acknowledgments.** This work was supported by a Doctoral Dissertation Improvement Grant to G. Ortí (BSR9112367) and by grants to A.

Meyer (BSR9119867, BSR9107838) and M.A. Bell (INT9117104), from the U.S. National Science Foundation. P. Petry was supported by CAPES scholarship 1055/90-7. We wish to thank P.R. Souza, A. Fortuny, R. Delfino, and S. Sverlij for specimens, M.A. Bell, D. Futuyama, and R. Vari for comments on earlier versions of the manuscript, and D. Kritsky and W.A. Boeger for sharing preliminary results on serrasalmin parasite studies. This paper was prepared in partial fulfillment of requirements for the Ph.D. in Ecology and Evolution by G. Ortí. This is contribution 956 from the Graduate Program in Ecology and Evolution at SUNY at Stony Brook.

## References

- Adachi J, Hasegawa M (1994) MOLPHY: a program package for molecular phylogenetics, version 2.2. The Institute of Statistical Mathematics, Tokyo
- Alves-Gomes JA, Ortí G, Haygood M, Heiligenberg W, Meyer A (1995) Phylogenetic analysis of the South American electric fishes (Order Gymnotiformes) and the evolution of their electrogenic system: a synthesis based on morphology, electrophysiology, and mitochondrial sequence data. *Mol Biol Evol* 12:298–318
- Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465
- Aquadro CF, Greenberg BD (1983) Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* 103:287–312
- Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J Mol Evol* 18:225–239
- Cestari MM, Galetti Jr PM (1992) Chromosome evolution in the genus *Serrasalmus* and cytotoxic considerations about Serrasalminae (Characidae, Pisces). *Rev Brasil Genet* 15:555–567
- Collins TM, Wimberger PH, Naylor GJP (1994) Compositional bias, character-state bias, and character-state reconstruction using parsimony. *Syst Biol* 43:482–496
- Dams E, Hendriks L, Van de Peer Y, Neefs JM, Smits G, Vandembemt I, De Wachter R (1988) Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Res* 16:r87–r173
- Dixon MT, Hillis DM (1993) Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic analysis. *Mol Biol Evol* 10:256–267
- Eigenmann C (1915) The Serrasalminae and Mylinae. *Ann Carnegie Museum*, Pittsburgh 9:266–272
- Farris JS (1969) A successive approximations approach to character weighting. *Syst Zool* 18:374–385
- Felsenstein J (1969) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17:368–376
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791
- Fink WL (1993) Revision of the piranha genus *Pygocentrus* (Teleostei, Characiformes). *Copeia* 1993:665–687
- Fitch WM, Markowitz E (1970) An improved method for determining codon variability in a gene and its application to the rate of fixation of mutations in evolution. *Biochem Genet* 4:579–593
- Géry J (1972) Poissons characoides des Guyanes II. Famille des Serrasalminae. *Zoologische Verhandelingen Leiden* 122:134–248
- Géry J (1977) Characoids of the World. Tropical Fish Hobbyist Publications, Neptune City, NJ, pp 672
- Gosline W (1951) Notes on the characoid fishes of the Subfamily Serrasalminae. *Proc Cal Acad Sci*, ser 4 27:17–64
- Goulding M (1980) The fishes and the forest. Explorations in Amazonian natural history. University of California Press, Berkeley, p 280
- Gray MW, Cedergren R (1993) The new age of RNA. *FASEB J* 7:4–6

- Gutell RR, Fox GE (1988) A compilation of large subunit rRNA sequences presented in a structural format. *Nucleic Acids Res* 16s:r175–r269
- Gutell RR, Gray MW, Schare MN (1993) A compilation of large subunit 23S and 23S-like ribosomal RNA structures: 1993. *Nucleic Acids Res* 21:3055–3074
- Gutell RR, Weiser B, Woese CR, Noller HF (1985) Comparative anatomy of 16-S-like ribosomal RNA. *Prog Nucleic Acid Res Mol Biol* 32:155–216
- Gyllenstein UB, Erlich HA (1988) Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc Natl Acad Sci USA* 85:7652–7655
- Hillis DM, Dixon MT (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Q Rev Biol* 66:411–453
- Hixon JE, Brown WM (1986) A comparison of the small ribosomal RNA genes from the mitochondrial DNA of the great apes and humans: sequence, structure, evolution and phylogenetic implications. *Mol Biol Evol* 3:1–18
- Jégu M, Tito de Morais L, Mendes dos Santos G (1992) Redescription des types d' *Utiaritchthys sennaebregai* Miranda Ribeiro, 1937 et description d'une nouvelle espèce du bassin Amazonien, (Characiformes, Serrasalminidae). *Cybiurn* 16:105–120
- Junk WJ (1984) Ecology, fisheries and fish culture in Amazonia. In: Sioli H (ed) *The Amazon, limnology and landscape ecology of a mighty tropical river and its basin*. Dr. W. Junk Publishers, Dordrecht, pp 443–476
- Kimura M (1981) Estimation of evolutionary distances between homologous nucleotide sequences. *Proc Natl Acad Sci USA* 78:454–458
- Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J Mol Evol* 29:170–179
- Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals. *Proc Natl Acad Sci USA* 86:6196–6200
- Kumar S, Tamura K, Nei M (1993) MEGA: molecular evolutionary genetics analysis, version 1.0. The Pennsylvania State University, University Park, PA
- Larson A, Wilson AC (1989) Patterns of ribosomal RNA evolution in salamanders. *Mol Biol Evol* 6:131–154
- Lee W, Kocher TD (1995) Complete sequence of a sea lamprey (*Petromyzon marinus*) mitochondrial genome: early establishment of the vertebrate genome organization. *Genetics* 139:873–887
- Leite RG, Jégu M (1990) Food habits of two species of *Acnodon* (Characiformes, Serrasalminidae) and scale-eating habits of *Acnodon normani*. *Cybiurn* 14:353–360
- Lundberg JG (1993) African-South American freshwater fish clades and continental drift: problems with a paradigm. In: Goldblatt P (ed) *Biological relationships between African and South America*. Yale University Press, New Haven, pp 156–199
- Lundberg JG, Machado-Allison A, Kay RF (1986) Miocene characid fishes from Colombia: evolutionary stasis and extirpation. *Science* 234:208–209
- Machado-Allison A (1982) Studies on the systematics of the subfamily Serrasalminae (Pisces-Characidae). Unpublished PhD thesis, The George Washington University, Washington DC
- Machado-Allison A (1983) Estudios sobre la sistemática de la subfamilia Serrasalminae (Teleostei, Characidae), parte II. Discusión sobre la condición monofilética de la subfamilia. *Acta Biol Venez* 11:145–195
- Machado-Allison A, Fink WL, Antonio ME (1989) Revisión del género *Serrasalmus* Lacepede, 1803 y géneros relacionados en Venezuela: I. Notas sobre la morfología y sistemática de *Pristobrycon striolatus* (Steindachner, 1908). *Acta Biol Venez* 12:140–171
- Maddison WP, Maddison DR (1992) MacCLADE: analysis of phylogeny and character evolution, version 3.0. Sinauer, Sunderland, MA
- Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Publications, Cold Spring Harbor, NY
- Marshall E (1995) Homely fish draws attention to Amazon deforestation. *Science* 267:814
- Meyer A (1993) Evolution of mitochondrial DNA of fishes. In: Hochachka PW, Mommsen P (eds) *The biochemistry and molecular biology of fishes*, vol 2. Elsevier Press, Amsterdam, pp 1–38
- Meyer A (1994) DNA technology and phylogeny of fish. In: Beaumont AR (ed) *Genetics and evolution of aquatic organisms*. Chapman and Hall, London pp. 219–249
- Mindell DP, Honeycutt RL (1990) Ribosomal RNA in vertebrates: evolution and phylogenetic implications. *Annu Rev Ecol Syst* 21:541–566
- Neefs JM, Van de Peer Y, De Rijk P, Goris A, De Wachter R (1991) Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Res* 19s:1987–2015
- Nelson EM (1961) The swim bladder in the Serrasalminae, with notes on additional morphological features. *Fieldiana Zool* 39:603–624
- Nico L, Taphorn DC (1988) Food habits of piranhas in the low llanos of Venezuela. *Biotropica* 20:311–321
- Noller HF (1984) Structure of ribosomal RNA. *Annu Rev Biochem* 53:119–162
- Norman JR (1929) The South American Characid fishes of the subfamily Serrasalmoninae with a revision of the genus *Serrasalmus* Lacepede. *Proc Zool Soc London* 52:661–1044
- Oliveira C, Almeida-Toledo LF, Foresti F, Britski H, Toledo-Filho SA (1988) Chromosome formulae of Neotropical freshwater fishes. *Rev Bras Genet* 11:577–624
- Olsen GJ, Matsuda H, Hagstrom R, Overbeek R (1994) fast DNAm1: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *Comput Appl Biosci* 10:41–48
- Olsen GJ, Woese CR (1993) Ribosomal RNA: a key to phylogeny. *FASEB J* 7:113–123
- Ortí G (1995) Molecular systematics of characiform fishes. Unpublished PhD thesis, State University of New York at Stony Brook
- Ortí G, Meyer A (1996) The radiation of characiform fishes and the limits of resolution of mitochondrial ribosomal DNA sequences. *Syst Biol* 45(in press)
- Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowski G (1991) *The simple fool's guide to PCR*. Dept of Zoology and Kewalo Marine Laboratory, University of Hawaii, Honolulu
- Porto JIR, Feldberg E, Nakayama C, Falcao JN (1992) A checklist of chromosome numbers and karyotypes of Amazonian freshwater fishes. *Rev Hydrobiol Trop* 25:287–299
- Porto JIR, Feldberg E, Nakayama CM, Jégu M (1989) Análise cariotípica na família Serrasalminidae (Ostariophysi, Characiformes): aspectos evolutivos. *Ciencia é Cultura (Suppl)* 41:714
- Porto JIR, Feldberg E, Nakayama CM, Maia RO, Jégu M (1991) Cytotaxonomic analysis in the Serrasalminidae (Ostariophysi, Characiformes). VII Congress of Ichthyology. *Bull Zool Mus Univ, Amsterdam, The Hague*
- Saiki RK, Gelfand DH, Stoffel S, Scharf S, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487–491
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–525
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminator inhibitors. *Proc Natl Acad Sci USA* 74:5436–5437
- Schimmel PR, Soll D, Abelson JN (1979) Transfer RNA: structure, properties and recognition. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Swofford DL (1993) PAUP: Phylogenetic Analysis Using Parsimony, ver 3.1.1. Smithsonian Institution, Washington, DC
- Swofford DL, Maddison WP (1992) Parsimony, character-state reconstructions, and evolutionary inferences. In: Mayden RL (ed) *Sys-*

- tematics, historical ecology, and North American freshwater fishes. Stanford University Press, Stanford, pp 186–223
- Swofford DL, Olsen GJ (1990) Phylogeny reconstruction. In: Hillis DM, Moritz C (eds) Molecular systematics. Sinauer, Sunderland MA, pp 411–501
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Turner DH, Sugimoto N, Freier SM (1988) RNA structure prediction. *Annu Rev Biophys Chem* 17:167–192
- Tzeng CS, Hui CF, Shen SC, Huang PC (1992) The complete nucleotide sequence of the *Crossostoma lacustre* mitochondrial genome: conservation and variations among vertebrates. *Nucleic Acids Res* 20:4853–4858
- Van Every LR, Kritsky DC (1992) Neotropical Monogenoidea. 18. *Anacanthorus Mizelle* and Price, 1965 (Dactylogyridae, Anacanthorinae) of piranha (Characoidea, Serrasalminidae) from the central Amazon, their phylogeny, and aspects of host-parasite coevolution. *J Helminthol Soc Wash* 59:52–75
- Vawter L, Brown WM (1993) Rates and patterns of base change in the small subunit ribosomal RNA gene. *Genetics* 134:597–608
- Wheeler WC, Honeycutt RL (1988) Paired sequence difference in ribosomal RNAs: evolutionary and phylogenetic implications. *Mol Biol Evol* 5:90–96
- Winemiller KO (1989) Ontogenetic diet shifts and resource partitioning among piscivorous fishes in the Venezuelan Llanos. *Environ Biol Fishes* 26:177–199
- Wolstenholme DR, Clary DO (1985) Sequence evolution of *Drosophila* mitochondrial DNA. *Genetics* 109:725–744
- Zardoya R, Garrido-Pertierra A, Bautista JM (1995) The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow trout, *Oncorhynchus mykiss*. *J Mol Evol* 41:(in press)

## Appendix 1: Specimens

All specimens used in this study have been numbered from 1 to 37 and are referred to by these numbers in the text. Numbers in parenthesis following the species name are G. Orti's collection numbers. When voucher specimens were deposited in museum collections, their collection numbers are preceded by INPA for the specimens deposited at the Instituto Nacional de Pesquisa da Amazonia, Manaus, Brazil, and by USNM for those at the US National Museum of Natural History (Washington, DC). Finally GenBank accession number (GB) are given for the 12S and 16S fragments, respectively.

Ingroup taxa:

Family Characidae, subfamily Serrasalminae:

gen *Pygocentrus*:

1—*P. nattereri* (#78), commercial source, locality unknown. GB: U33558, U33590

2—*P. nattereri* (#215), commercial source, locality unknown. GB: U33558, U33590

3—*P. nattereri* (#212), R. Solimoes, Ilha da Marchantaria, AM, Brazil (P. Petry collector). INPA10143. GB: U33558, U33590

4—*P. nattereri* (#155), R. Uruguay, Salto Grande, Argentina (R. Delfino collector, USNM325686). GB: U33559, U33591

gen *Serrasalmus*:

5—*S. spilopleura* (#139), R. Uruguay, Salto Grande, Argentina (S. Sverlij collector, USNM325683). GB: U33560, U33592

6—*S. n.sp.* (2n = 58) (#220), R. Negro-Solimoes, AM, Brazil (J. Porto collector). GB: U33561, U33593

7—*S. compressus (altuvei?)* (2n = 60) (#221), R. Solimoes, Ilha da Marchantaria, AM, Brazil (J. Porto collector). GB: U33562, U33594

gen *Pristobrycon*:

8—*P. sp.* (#224), R. Solimoes, Ilha da Marchantaria, AM, Brazil (J. Porto collector). GB: U33563, U33595

9—*P. striolatus* (#225), R. Pitinga, UHE do Pitinga, AM, Brazil (M. Jégu collector). GB: U33597, U33596

10—*P. striolatus* (#226), R. Pitinga, UHE do Pitinga, AM, Brazil (M. Jégu collector). GB: U33564, U33598

gen *Catoprion*:

11—*C. mento* (#80), commercial source, locality unknown. GB: U33565, U33599

12—*C. mento* (#247), commercial source, locality unknown. INPA10145. GB: U33565, U33599

gen *Metynnis*:

13—*M. sp.* (#81), commercial source, locality unknown. GB: U33566, U33600

14—*M. cf. mola* (#202), R. Miranda, Pantanal Matogrossense, Campo Grande, MS, Brazil (P.R. Souza collector). INPA10146. GB: U33567, U33601

gen *Myleus*:

15—*M. Myloplus rubripinnis* (#79), commercial source, locality unknown. GB: U33568, U33602

16—*M. Myloplus asterias* (#235), R. Pitinga, UHE do Pitinga, AM, Brazil (M. Jégu collector). GB: U33569, U33603

17—*M. Myloplus tiete* (#210), R. Miranda, Pantanal Matogrossense, Campo Grande, MS, Brazil (P.R. Souza collector). INPA10147. GB: U33570, U33604

18—*M. Prosomyleus schomburgkii* (#233), R. Pitinga, Cachoeira 40 Ilhas, AM, Brazil (M. Jégu collector). GB: U33571, U33605

19—*M. Myleus pacu* (#238), R. Pitinga, Cachoeira 40 Ilhas, AM, Brazil (M. Jégu collector). GB: U33572, U33606

20—*M. Myleus pacu* (#239'), R. Xingu, Cachoeira do Kaituka, Pará, Brazil (M. Jégu collector). GB: U33573, U33607

gen *Mylesinus*:

21—*M. paraschomburgkii* (#227), R. Pitinga, Cachoeira 40 Ilhas, AM, Brazil (M. Jégu collector). GB: U33574, U33608

22—*M. paraschomburgkii* (#228), R. Pitinga, Cachoeira 40 Ilhas, AM, Brazil (M. Jégu collector). GB: U33574, U33609

gen 'N. gen. A'

23—N. gen. A n.sp. (#246), R. Xingu, Cachoeira do Kaituka, Pará, Brazil (M. Jégu collector). This specimen could not be assigned to any valid genus of the Serrasalminae, but is similar in many respects to *Utiaritchthys* and *Myleus* (Jégu et al. 1992; Jégu unpublished data). GB: U33575, U33610

gen *Acnodon*:

24—*A. normani* (#244), R. Xingu, Cachoeira do Kaituka, Pará, Brazil (M. Jégu collector). GB: U33576, U33611

25—*A. normani* (#245), R. Xingu, Cachoeira do Kaituka, Pará, Brazil (M. Jégu collector). GB: U33577, U33612

gen *Mylossoma*:

26—*M. duriventri* (#203), R. Solimoes, Ilha da Marchantaria, AM, Brazil (P. Petry collector). INPA10154. GB: U33578, U33613

27—*M. paraguayensis* (#214), R. Miranda, Pantanal Matogrossense, Campo Grande, MS, Brazil (P.R. Souza collector). INPA10152. GB: U33579, U33614

28—*M. aureum* (#204), R. Solimoes, Ilha da Marchantaria, AM, Brazil (P. Petry collector). INPA10153. GB: U33580, U33615

Gen *Colossoma*:

29—*C. macropomum* (#216), R. Solimoes, Ilha da Marchantaria, AM, Brazil (P. Petry collector). INPA10149. GB: U33581, U33616

30—*C. macropomum* (#201), R. Solimoes, Ilha da Marchantaria, AM, Brazil (P. Petry collector). INPA10150. GB: U33582, U33617

gen *Piaractus*:

31—*P. mesopotamicus* (#205), R. Miranda, Pantanal Matogrossense, Campo Grande, MS, Brazil (P.R. Souza collector). INPA10151. GB: U33583, U33618

32—*P. brachipomus* (#200), R. Solimoes, Ilha da Marchantaria, AM, Brazil (P. Petry collector). INPA10148. GB: U33584, U33619

33—*P. mesopotamicus* (#143), A. Fortuny collector, locality unknown. GB: U33585, U33620

34—*P. brachipomus* (#45), commercial source, locality unknown. GB: U33586, U33621

Outgroup taxa:

Family Characidae, subfamily Bryconinae

gen *Chalceus*

35—*C. macrolepidotus* (#40), commercial source, locality unknown. GB: U33587, U33622

gen *Triportheus*

36—*T. paranensis* (#109), R. Paraná, Bella Vista, Corrientes, Argentina (A. Fortuny collector). GB: U33588, U33623

Family Characidae, subfamily Characinae

gen *Gnathocharax*

37—*G. steindachneri* (#123), commercial source, locality unknown. GB: U33589, U33624