



ACADEMIC  
PRESS

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Molecular Phylogenetics and Evolution 26 (2003) 262–288

MOLECULAR  
PHYLOGENETICS  
AND  
EVOLUTION

[www.elsevier.com/locate/ympev](http://www.elsevier.com/locate/ympev)

# Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa

Wei-Jen Chen,<sup>1</sup> Céline Bonillo, and Guillaume Lecointre\*

*Laboratoire d'Ichtyologie générale et appliquée, et service de systématique moléculaire (IFR CNRS 1541), Muséum National d'Histoire Naturelle, 43 rue Cuvier, 75231 Paris cedex 05, France*

Received 17 April 2002; revised 30 August 2002

## Abstract

Although much progress has been made recently in teleostean phylogeny, relationships among the main lineages of the higher teleosts (Acanthomorpha), containing more than 60% of all fish species, remain poorly defined. This study represents the most extensive taxonomic sampling effort to date to collect new molecular characters for phylogenetic analysis of acanthomorph fishes. We compiled and analyzed three independent data sets, including: (i) mitochondrial ribosomal fragments from 12S and 16s (814 bp for 97 taxa); (ii) nuclear ribosomal 28S sequences (847 bp for 74 taxa); and (iii) a nuclear protein-coding gene, rhodopsin (759 bp for 86 taxa). Detailed analyses were conducted on each data set separately and the principle of taxonomic congruence without consensus trees was used to assess confidence in the results as follows. Repeatability of clades from separate analyses was considered the primary criterion to establish reliability, rather than bootstrap proportions from a single combined (total evidence) data matrix. The new and reliable clades emerging from this study of the acanthomorph radiation were: Gadiformes (cods) with Zeioids (dorids); Beloniformes (needlefishes) with Atheriniformes (silversides); blenioids (blennies) with Gobiesocoidei (clingfishes); Channoidei (snakeheads) with Anabantoidei (climbing gouramies); Mastacembeloidei (spiny eels) with Synbranchioidei (swamp-eels); the last two pairs of taxa grouping together, Syngnathoidei (aulostomids, macroramphosids) with Dactylopteridae (flying gurnards); Scombroidei (mackerels) plus Stromatoidei plus Chiasmodontidae; Ammodytidae (sand lances) with Cheimarrichthyidae (torrentfish); Zoarcoidei (eelpouts) with Cottoidei; Percidae (perches) with Notothenioidei (Antarctic fishes); and a clade grouping Carangidae (jacks), Echeneidae (remoras), Sphyaenidae (barracudas), Menidae (moonfish), Polynemidae (threadfins), Centropomidae (snooks), and Pleuronectiformes (flatfishes).

© 2002 Elsevier Science (USA). All rights reserved.

**Keywords:** 12s mtDNA; 16s mtDNA; 28s rDNA; Acanthomorpha; Teleostean phylogeny; Rhodopsin; Separate analyses; Reliability; Robustness; Taxonomic congruence; Teleostei

## 1. Introduction

With advances in the collection of molecular data, phylogenetic results obtained from molecular sources are used to a greater extent to interpret organismal diversity (Moritz and Hillis, 1996). These phylogenetic hypotheses rely increasingly on the information ob-

tained from different genes. The benefit of sampling several independent gene genealogies to infer phylogenetic relationships among taxa is well established (e.g., Cao et al., 1994; Cumming et al., 1995; Russo et al., 1996; Zardoya and Meyer, 1996), since ultimately a better representation of the whole genome is highly desirable. However, the issue of how to analyze multiple sources of data appears to remain unsettled (Lecointre and Deleporte, 2000; Miyamoto and Fitch, 1995). Extreme views emphasize separate analysis (Mickevich, 1978) or simultaneous analysis (e.g., Nixon and Carpenter, 1996), also called the “total evidence” approach by Kluge (1989). Even if the importance of different

\* Corresponding author. Fax: +33-1-40-79-37-71.

E-mail addresses: [wchen@biocomp.unl.edu](mailto:wchen@biocomp.unl.edu) (W.-J. Chen), [lecointre@mnhn.fr](mailto:lecointre@mnhn.fr) (G. Lecointre).

<sup>1</sup> Present address: 315 Manter Hall, School of Biological Sciences, University of Nebraska-Lincoln, NE 68511-0118, USA.

protocols of analyses was discussed by the flurry of recent papers (De Queiroz et al., 1995; Huelsenbeck et al., 1996; Levasseur and Lapointe, 2001; Miyamoto and Fitch, 1995; Nixon and Carpenter, 1996; Lecointre and Deleporte, 2000), the “total evidence approach” is currently the most widely employed. In this paper, we present new molecular data for the Acanthomorpha (Teleostei) to question whether, in terms of reliability, the direct application of the “total evidence” approach is the best solution for a difficult phylogenetic problem.

### 1.1. The paradox of reliability in total evidence approach

One of the central questions in systematics is how phylogenetic hypothesis can be assessed for confidence. As claimed by Hennig (1966), “. . . the reliability of hypothesis increases with number of individual characters that can be fitted into transformation series. . .” Following this claim, supporters of the “total evidence” approach advocate combining all available data in a single matrix in order to globally maximize congruence of the whole set of available relevant characters, the principle of character congruence (Barrett et al., 1991; Eernisse and Kluge, 1993; Kluge, 1989). The basic assumption for this approach is that there are no significant differences *in nature* between partitions, thus implying that any delineation of data partitions is only product of technical and/or historical artifacts. The total evidence approach performs well (securing increasing robustness as more characters are analyzed) when this basic assumption is met and when the distribution of homoplasy (non-historical signal) is randomly distributed among the data partitions. In this case, it is expected that phylogeny will be inferred correctly, if enough data are collected, because historical signal will rise above random homoplasy (Farris, 1983). That is, stochastic errors in the data may lead to the incorrect inference when sample size is small but will disappear with infinite sample size (Swofford et al., 1996). However, molecular systematists have recognized that homoplasy tends to accumulate within genes in ways that are not completely random (Naylor and Brown, 1998). Non-random aspects of molecular homoplasy may be understood by analyzing functional constraints and can be detected without phylogenetic tools, for example by identifying mutational and/or base compositional biases within some positions or regions free to vary. These molecular processes may originate and accumulate non-random homoplasy within a gene and potentially mislead phylogenetic reconstruction. Furthermore, these properties that can be very different from one gene to another and provoke different kinds of deceptive signals. For instance, a set of unrelated taxa sharing the same strong compositional bias in a gene will be erroneously clustered in a tree based on DNA sequences of this gene (Hasegawa and Hashimoto, 1993; Leipe et al., 1993; Chang and Campbell, 2000; Gautier,

2000). It is possible that the contribution of each data matrix to the final topology may be disproportionate, and have unexpected effects in simultaneous analysis. In the worst case, a topology could be completely determined by one of the matrices which contains strong hierarchic but non-historical signal when the others present weak but truly historical signal (Naylor and Adams, 2001; Chen, 2001). In such cases, the preferred strategy to obtain a reliable result would not be a simple total evidence analysis but a careful dissection of noise and signal among the different data partitions. Clearly, reliability of the inference will not necessarily increase with increasing number of characters by just combining heterogeneous sources of data. Warnings against simultaneous analysis have been addressed repeatedly in the recent literature, for instance under the notion of “process partitions” by Bull et al. (1993), who emphasized that verification of congruence or homogeneity between data sets is necessary and critical before combining data and performing simultaneous analysis. Finally, if homoplasy accumulates in a non-random manner within genes while in a heterogeneous manner between genes, data partitions have some degree of naturalness, so acceptance a priori of the null hypothesis of the total evidence approach is not a reasonable practice.

### 1.2. Reliability of clades needs separate analysis

The most common way in systematic studies to assess “reliability” of phylogenetic inferences is the use of indicators of robustness, such as the Bremer (or decay) index (Bremer, 1994) and bootstrap proportions (Felsenstein, 1985). Robustness is attached numerical value to internal branches in trees, calculated from a given (single) data set to measure the strength of support for those branches and corresponding groups. One must keep in mind that these indicators merely assess the *strength* of the signal used to order the data hierarchically (Swofford et al., 1996). That signal can originate either from common ancestry or non-phylogenetic sources like convergent strong selective constraints. Therefore, the numerical value of a robustness indicator does not measure the *reliability of a phylogenetic inference*. Robustness would be considered as reliability only if (1) assumptions about independence of characters and homogeneous distribution of homoplasy were not violated (Kluge and Wolf, 1993; Sanderson, 1989); and (2) all the available knowledge at the time has been taken into account (Carnap, 1950; Lecointre and Deleporte, 2000). However, as stated above, the ideal data set may not be easy to collect and this may be particularly true for molecular data. According the simulation studies, support indices could over- or under-estimate the real expected reliability (Hillis and Bull, 1993). These indicators could be totally misleading due to classical pitfalls of phylogenetic reconstruction provoked by unequal rates of

changes among lineages or base compositional bias and/or by long branch attraction (Felsenstein, 1978; Huelssenbeck, 1997; Philippe and Adoutte, 1998; Philippe and Douzery, 1994; Philippe and Laurent, 1998). One must wonder whether a high bootstrap proportion should be given higher confidence than a lower one. It is often impossible to know *from a single tree* (such as a tree inferred from simultaneous analysis) whether the grouping patterns are due to artifacts of phylogenetic reconstruction or due to common ancestry, whatever the statistical robustness associated. However, separate analysis provides other opportunities for assessing reliability.

Reliability is the quality of being trustworthy given to a statement at a given time. It is never associated with a numerical statistical value drawn from a single data set isolated from other remaining evidence (Carnap, 1950). In other words, when analyzing several data sets separately (which is what the world-wide scientific community does every day), a given bootstrap proportion obtained for a clade from a single data set cannot be a measure of reliability. In science, reliability depends on the repeatability of results through different investigations (Grande, 1994). It is not surprising that experienced molecular systematists converge on a “taxonomic congruence” approach, proposing to analyze data sets separately (Grande, 1994; Mickevich, 1978; Miyamoto and Fitch, 1995; Nelson, 1979), at least as a heuristic step. The congruence of inferences separately drawn from *independent* data is considered as strong indicator of reliability. If we keep in mind the fact that molecular homoplasy may have different effects on tree reconstruction from one gene to another, obtaining the same clade from separate analysis of several genes *despite this fact* renders the clade even more reliable. In other words, obtaining the same tree or even some common clades means that there is a common structure in these data sets that must come from common evolutionary history. Miyamoto and Fitch (1995) suggested that relationships among taxa that are supported by different independent data sets are particularly robust even if the statistical support for each individual result is weak. This is equivalent to obtaining independent verification of an experimental hypothesis from an additional experimental source. This independent type of verification may be lost in combining data sets right from the beginning. Empirically, this point of view implies that two independent genes are not likely to harbor the same positively misleading signals. Even if it is always possible to imagine that two or three genes can exhibit the same positively misleading signals (for instance the same long branches due to a common taxonomic sampling issue), the risk here is by far lower than blindly trusting the bootstrap proportions from the direct simultaneous analysis. The same reasoning can be used to reply to the objection made to separate analyses, that different genes may contribute to resolve different parts of a phylogeny.

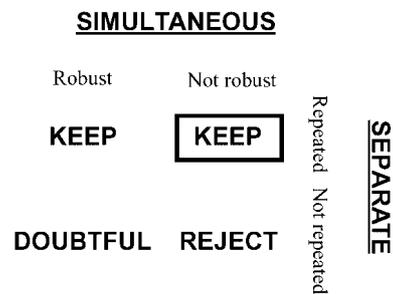


Fig. 1. General protocol for assessing reliability to a clade. A clade can be repeated (or not) across separate analyses, and can be robust (or not) in the simultaneous analysis of all available data. The square shows that repeatability of a clade is a more convincing indicator of its reliability than bootstrap proportions or other indices of robustness.

Finding the same clade repeated *despite* the possibility that different genes may resolve different parts of the phylogeny is using repeatability in a conservative manner, securing reliability.

Thus, the main advantage of separate analysis (without consensus trees) is that it provides a measure of repeatability, but more than a simple majority-rule consensus tree, an additional opportunity to detect tree reconstruction artifacts due to local positively misleading signals. We would be inclined to prefer the same clade that is inferred repeatedly from several data sets with low bootstrap proportions than a highly supported clade inferred from a single data set. We will therefore not use consensus trees, instead we will use repeatability through separate analyses to assess reliability of the clades found in the tree from the simultaneous analysis (Fig. 1). In other words, we use the simultaneous analysis to obtain the complete tree, and separate analyses to determine which clades of that tree are reliable.

### 1.3. *Acanthomorpha* as a case study

The spiny teleost fishes grouped within *Acanthomorpha* (Rosen, 1973) comprise more than 14,736 species (Helfman et al., 1997; Nelson, 1994) and represent one third of the extant vertebrate species of the world. This clade is divided into three large assemblages: the *Paracanthopterygii* (cods, goosefishes), the *Atherinomorpha* (silversides), and the most species-rich group, the *Percomorpha* (perch-like fishes). The earliest acanthomorph fossils known are aipichthyids and polymixiids from the Cenomanian, Upper Cretaceous (Gaudant, 1978; Gayet, 1980a; Otero and Gayet, 1996; Patterson, 1964). Shortly after this period, a vast diversity of acanthomorphs (representing 80 families) suddenly appears in the fossil record, starting in the Early Eocene between 45–55 million years ago (Benton, 1993; Patterson, 1993). This pattern suggests a putative rapid radiation, which resulted in the most diverse vertebrate group of the modern fauna.

Since the pioneering work on systematics of fishes by Greenwood et al. (1966), many studies were published proposing hypotheses of relationships for lower teleosts, but relatively few for the higher teleosts, especially for the Acanthomorpha (Lecointre, 1994; Rosen, 1982, 1985). Consequently, Nelson (1989) concluded his survey of teleostean phylogeny with the following statement “*recent work has resolved the bush at the bottom, but the bush at the top persists,*” a bush already clearly illustrated by Rosen (1982). Our knowledge of high-level acanthomorph phylogeny is very poor considering their sizeable species diversity, especially within the major clade Percomorpha. The vast majority of studies of higher teleosts have focused on relationships at the specific and generic levels or between closely related families. So far, the only three cladograms based on morphological characters depicting interrelationships among acanthomorphs are those of Johnson and Patterson (1993), Lauder and Liem (1983), and Stiassny and Moore (1992). In spite of showing resolution for the basal parts of tree, and in spite of proposing new hypothesis (e.g., Smegmamorpha in Johnson and Patterson, 1993), substantial disagreement persists, especially on the phylogenetic positions of the Zeiformes (dorids), Beryciformes (squirrel fishes), and Synbranchiformes (swamp or spiny eels). Clearly, the phylogenetic relationships reflecting the main acanthomorph radiation are still unclear. Molecular data are only slowly starting to produce results, such as the two recent studies published during the preparation of this paper (Miya et al., 2001; Wiley et al., 2000). These phylogenetic trees are based on a combined matrix (1722 characters) of 12S mitochondrial DNA, 28S nuclear DNA, and morphological data (Johnson and Patterson, 1993) and selected nucleotides sequences (7002 characters) from whole mitochondrial genomes (Miya et al., 2001). If merely increasing the number of characters for analysis and if performing a “total evidence” approach could give reliable results, these two studies should provide a better insight of acanthomorph phylogeny. However, interrelationships between acanthomorph orders or suborders representing major lineages remain poorly resolved in terms of statistical support, with few exceptions. Moreover, as discussed above, robustness does not necessarily mean reliability. Without comparing trees from independent data sets, it is not possible to assess reliability of newly proposed acanthomorph clades. Following this view, the acanthomorph problem still needs to be examined, especially by way of separate analyses.

#### *1.4. More intense taxonomic sampling is required for acanthomorph phylogeny*

Adequate taxonomic sampling to fairly represent the highly complex patterns of diversification of Acanthomorpha is a compulsory issue requiring careful

consideration. In general, major lineages within Acanthomorpha are poorly defined, especially for Percomorpha. In such situations, taxonomic sampling must be extended to neighboring lineages, until the sample is sufficiently inclusive to contain the clade of interest. This is the case for the Percomorpha (Johnson, 1993; Johnson and Patterson, 1993; Rosen, 1973, 1985; Stiassny, 1986) and explains why one must sample the whole acanthomorph diversity when just trying to investigate percomorph phylogeny. A related problem is that some traditionally recognized percomorph subdivisions have been shown to be polyphyletic (e.g., Perciformes, Trachiniodei, Percoidei, Scorpaeniformes; Gill, 1996; Johnson, 1993; Patterson and Rosen, 1989; Stiassny, 1990; Stiassny and Moore, 1992; Travers, 1981) and may not even belong to this group. Since monophyly of such groups is questionable, using reduced sampling from predefined groups is risky. When sampling taxa from paraphyletic or polyphyletic groups, phylogenetic conclusions will depend on the choice of representatives. To address correctly the phylogenetic hypothesis, sampling a large variety of terminals within each of the putatively polyphyletic subdivisions is required. This drastically increases the necessary taxonomic sampling. However, all previous studies sampled very few acanthomorph terminals. One of the best sampling efforts includes merely 32 acanthomorph taxa (Miya et al., 2001), with only a single representative from the large group Perciformes, which is clearly a polyphyletic group (Johnson and Patterson, 1993)!

For this study we sampled acanthomorph diversity thoroughly, including representation of 48 suborders and more than 60 families. We present and analyze new data from four genes with different properties in their cellular location, function, and sequence variation. These include two nuclear genes: portions of the 28S ribosomal DNA (domains C1–C2, D3, D6, C12, and D12) and the gene encoding rhodopsin; and two mitochondrial ribosomal genes: 12S and 16S (Table 1). Using both separate analysis and simultaneous analysis, this study aims to discover reliable clades among the main lineages within the acanthomorph radiation, with particular attention to the phylogenetic relationships of the order Zeiformes and the interrelationships of members of the Smegmamorpha (new clade defined by Johnson and Patterson, 1993) and of “Perciformes.” We present a detailed analysis that shows the use of repeatability as the main criterion to postulate the validity of some previously unrecognized clades.

## **2. Materials and methods**

### *2.1. Taxon sampling and DNA extraction*

Taxa were selected to represent a large proportion of acanthomorph diversity, including representatives of 40

Table 1  
Taxa included in this study

Order/suborder	Family	Taxon	GenBank Accession No.			
			28S	12S	16S	Rhodopsin
<b>Outgroups</b>						
Osmeriformes	Bathylagidae	<i>Bathylagus euryops</i>	AY141465–68	AY141325	AY141395	AY141255
Salmoniformes	Salmonidae	<i>Oncorhynchus mykiss</i>	<u>U34341<sup>a</sup></u>	<u>L29771</u>	<u>L29771</u>	
Stomiiformes	Gonostomatidae	<i>Gonostoma atlanticum</i>		<u>D84033</u>	<u>D84049</u>	
		<i>Gonostoma bathyphilum</i>	AY141469–72			AY141256
Aplepisauroidei	Synodontidae	<i>Synodus saurus</i>		<u>AF049723</u>	<u>AF049733</u>	
Chlorophthalmoidei	Ipnopidae	<i>Bathypterois dubius</i>	AY141473–76	AY141326	AY141396	AY141257
Aulopoidei	Aulopidae	<i>Aulopus purpurissatus</i>		<u>AF049722</u>	<u>AF049732</u>	
Myctophiformes	Myctophidae	<i>Electrona antarctica</i>	AY141477–80	AY141327	AY141397	AY141258
		<i>Hygophum hygomii</i>		<u>AF049724</u>	<u>AF049734</u>	
<b>Acanthomorpha</b>						
Lampridiformes	Lampridae	<i>Lampris immaculatus</i>	AY141481–84	AY141328	AY141398	AY141259
	Veliferidae	<i>Metavelifer multiradiatus</i>		<u>AF049725</u>	<u>AF049735</u>	
Polymixiiformes	Polymixiidae	<i>Polymixia japonica</i>		<u>AF049730</u>	<u>AF049740</u>	
Paracanthopterygii						
Gadiformes	Gadidae	<i>Gadus morhua</i>	AY141485–88	AY141329	AY141399	<u>AF137211</u>
		<i>Merlangius merlangus</i>	AY141489–92	AY141330	AY141400	AY141260
Percopsiformes	Percopsidae	<i>Percopsis omiscomaycus</i>		<u>AF049731</u>	<u>AF049741</u>	
Lophiiformes	Ceratiidae	<i>Ceratias holboelli</i>	AY141505–08	AY141334	AY141404	AY141263
Zeiformes						
Zeioidei	Zeidae	<i>Zeus faber</i>	AY141493–96	AY141331	AY141401	<u>Y14484</u>
	Oreosomatidae	<i>Neocyttus helgae</i>	AY141497–00	AY141332	AY141402	AY141261
Caproidei	Caproidae	<i>Capros aper</i>	AY141501–04	AY141333	AY141403	AY141262
Beryciformes						
Trachichthyoidei	Trachichthyidae	<i>Hoplostethus mediterraneus</i>	AY141509–12	AY141335	AY141405	AY141264
Berycoidei	Berycidae	<i>Beryx splendens</i>	AY141513–16	AY141336	AY141406	AY141265
Holocentroidei	Holocentridae	<i>Myripristis botche</i>	AY141517–20	AY141337	AY141407	
		<i>Myripristis violacea</i>				<u>U57539</u>
<b>Percomorpha</b>						
<b>Segmamorpha</b>						
Mugiloidei	Mugilidae	<i>Liza</i> sp.	AY141521–24	AY141338	AY141408	AY141266
<b>Atherinomorpha</b>						
Atherinoidei	Atherinidae	<i>Atherina boyeri</i>				<u>Y18676</u>
Bedotioidei	Bedotiidae	<i>Bedotia geayi</i>	AY141525–28	AY141339	AY141409	AY141267
Belonoidei	Belonidae	<i>Belone belone</i>	AY141529–32	AY141340	AY141410	AY141268
	Hemiramphidae	<i>Dermogenys pusilla</i>		AY141341	AY141411	
Cyprinodontoidei	Poeciliidae	<i>Poecilia reticulata</i>	AY141533–36	AY141342	AY141412	AY141269
<b>Gasterosteriformes</b>						
Gasterosteoidi	Gasterosteidae	<i>Spinachia spinachia</i>	AY141585–88	AY141356	AY141426	AY141281
Syngnathoidei	Aulostomidae	<i>Aulostomus chinensis</i>	AY141577–80	AY141353	AY141423	AY141279
	Fistulariidae	<i>Fistularia petimba</i>		AY141355	AY141425	AY141324
	Macroramphosidae	<i>Macroramphosus scolopax</i>	AY141581–84	AY141354	AY141424	AY141280
<b>Synbranchiformes</b>						
Synbranchoidei	Synbranchidae	<i>Monopterus albus</i>	AY141565–68	AY141350	AY141420	AY141276
Mastacembeloidei	Mastacembelidae	<i>Mastacembelus erythrotaenia</i>	AY141561–64	AY141349	AY141419	AY141275
<b>Unnamed</b>						
Dactylopteriformes	Dactylopteridae	<i>Dactylopterus volitans</i>	AY141589–92	AY141357	AY141427	AY141282
<b>Scorpaeniformes</b>						
Scorpaenoidei	Scorpaenidae	<i>Scorpaena onaria</i>	AY141617–20	AY141364	AY141434	AY141288
	Triglidae	<i>Chelidonicichthys lucerna</i>	AY141609–12	AY141362	AY141432	AY141287
Cottoidei	Cottidae	<i>Taurulus bubalis</i>	AY141613–16	AY141363	AY141433	<u>U97275</u>
<b>Tetraodontiformes</b>						
Tetraodontoidei	Tetraodontidae	<i>Lagocephalus laevigatus</i>	AY141601–04	AY141360	AY141430	AY141285
		<i>Tetraodon nigroviridis</i>				<u>AJ293018</u>
		<i>Takifugu rubripes</i>				<u>AF137214</u>
	Balistidae	<i>Balistes</i> sp.				<u>AF137212</u>
	Ostraciidae	<i>Ostracion</i> sp.				<u>AF137213</u>
	Molidae	<i>Mola mola</i>	AY141605–08	AY141361	AY141431	AY141286

Table 1 (continued)

Order/suborder	Family	Taxon	GenBank Accession No.			
			28S	12S	16S	Rhodopsin
Pleuronectiformes						
Psettoidoidei	Psettodidae	<i>Psettodes</i> sp.				<u>AF148143</u>
Pleuronectoidei	Bothidae	<i>Arnoglossus imperialis</i>	AY141593–96	AY141358	AY141428	AY141283
	Paralichthyidae	<i>Paralichthys olivaceus</i>		<u>AB028664</u>	<u>AB028664</u>	
	Citharidae	<i>Citharus linguatula</i>				AY141323
	Soleidae	<i>Microchirus variegatus</i>	AY141597–00	AY141359	AY141429	AY141284
		<i>Solea vulgaris</i>				<u>Y18672</u>
Perciformes						
Percoidaei	Serranidae	<i>Serranus accraensis</i>	AY141621–24	AY141365	AY141435	AY141289
		<i>Holanthias chrysostictus</i>	AY141625–28	AY141366	AY141436	AY141290
		<i>Epinephelus aeneus</i>	AY141629–32	AY141367	AY141437	AY141291
		<i>Pogonoperca punctata</i>		AY141368	AY141438	AY141292
	Centropomidae	<i>Lates calcarifer 1</i>	AY141641–44	AY141371	AY141441	AY141294
		<i>Lates calcarifer 2</i>				<u>AF148144</u>
	Moronidae	<i>Lateolabrax japonicus</i>	AY141633–36	AY141369	AY141439	AY141293
		<i>Dicentrarchus labrax</i>	AY141637–40	AY141370	AY141440	<u>Y18673</u>
		<i>Morone chrysops</i>		<u>AF055589</u>	<u>AF055610</u>	
	Percidae	<i>Perca fluviatilis</i>	AY141645–48	AY141372	AY141442	AY141295
		<i>Gymnocephalus cernuus</i>	AY141649–52	AY141373	AY141443	AY141296
	Chaetodontidae	<i>Chaetodon striatus</i>		<u>AF055592</u>	<u>AF055613</u>	
	Drepanidae	<i>Drepane punctata</i>		<u>AF055595</u>	<u>AF055616</u>	
		<i>Drepane africana</i>	AY141749–52			AY141321
	Pomacanthidae	<i>Holacanthus ciliaris</i>	AY141753–56	<u>AF055593</u>	<u>AF055614</u>	AY141322
	Sparidae	<i>Sparus aurata</i>				<u>Y18664</u>
	Mullidae	<i>Mullus surmuletus</i>				<u>Y18666</u>
	Menidae	<i>Mene maculata</i>	AY141729–32	AY141390	AY141460	AY141316
	Polynemidae	<i>Pentanemus quinquarius</i>	AY141733–36	AY141391	AY141461	AY141317
	Pomatomidae	<i>Pomatomus saltatrix</i>		<u>AF055591</u>	<u>AF055612</u>	
Carangoidei	Carangidae	<i>Chloroscombrus chrysurus</i>	AY141717–20	AY141387	AY141457	AY141313
		<i>Caranx latus</i>		<u>AF055590</u>	<u>AF055611</u>	
		<i>Trachinotus ovatus</i>	AY141721–24	AY141388	AY141458	AY141314
	Echeneidae	<i>Echeneis naucrates</i>	AY141725–28	AY141389	AY141459	AY141315
Acanthuroidei	Acanthuridae	<i>Ctenochaetus striatus</i>	AY141745–48	AY141394	AY141464	AY141320
		<i>Acanthurus xanthopterus</i>		<u>AF055609</u>	<u>AF055630</u>	
		<i>Zebrasoma scopas</i>		<u>AF055606</u>	<u>AF055627</u>	
		<i>Naso lituratus</i>		<u>AF055603</u>	<u>AF055624</u>	
		<i>Prionurus maculatus</i>		<u>AF055604</u>	<u>AF055625</u>	
	Ephippidae	<i>Platax orbicularis</i>		<u>AF055597</u>	<u>AF055618</u>	
	Luvaridae	<i>Luarus imperialis</i>		<u>AF055601</u>	<u>AF055622</u>	
	Scatophagidae	<i>Scatophagus argus</i>		<u>AF055598</u>	<u>AF055619</u>	
	Siganidae	<i>Siganus canaliculatus</i>		<u>AF055600</u>	<u>AF055621</u>	
	Zanclidae	<i>Zanclus cornutus</i>		<u>AF055602</u>	<u>AF055623</u>	
Labroidei	Labridae	<i>Labrus bergylta</i>	AY141737–40	AY141392	AY141462	AY141318
	Scaridae	<i>Scarus hoefleri</i>	AY141741–44	AY141393	AY141463	AY141319
Zoarcoidei	Zoarcidae	<i>Austrolycus depressiceps</i>	AY141653–56	AY141374	AY141444	AY141297
	Pholidae	<i>Pholis gunnellus</i>	AY141657–60	AY141375	AY141445	AY141298
Notothenioidei	Bovichtidae	<i>Bovichtus variegatus</i>	AY141661–64	<u>Z32702</u>	<u>Z32721</u>	AY141299
		<i>Cottoperca gobio</i>	AY141665–68	AY141376	AY141446	AY141300
		<i>Pseudaphritis urvillii</i>	AY141669–72	AY141377	AY141447	AY141301
	Nototheniidae	<i>Notothenia coriiceps</i>	AY141673–76	<u>Z32712</u>	<u>Z32731</u>	AY141302
		<i>Eleginops maclovinus</i>				AY141303
	Channichthyidae	<i>Chionodraco hamatus</i>		<u>Z32704</u>	<u>Z32723</u>	
		<i>Neopagetopsis ionah</i>	AY141677–80			
Trachioidei	Trachinidae	<i>Trachinus draco</i>	AY141681–84	AY141378	AY141448	AY141304
	Uranoscopidae	<i>Uranoscopus albesca</i>	AY141685–88	AY141379	AY141449	AY141305
	Ammodytidae	<i>Ammodytes tobianus</i>	AY141689–92	AY141380	AY141450	AY141306
	Cheimarrichthyidae	<i>Cheimarrichthys fosteri</i>	AY141693–96	AY141381	AY141451	AY141307
	Chiasmodontidae	<i>Kali macrura</i>	AY141697–00	AY141382	AY141452	AY141308
Blennioidei	Blenniidae	<i>Parablemnus gattorugine</i>	AY141545–48	AY141345	AY141415	AY141271
	Tripterygiidae	<i>Forsterygion lapillum</i>	AY141549–52	AY141346	AY141416	AY141272
Gobiesocoidei	Gobiesocidae	<i>Lepadogaster lepadogaster</i>	AY141553–56	AY141347	AY141417	AY141273
		<i>Apletodon dentatus</i>	AY141557–60	AY141348	AY141418	AY141274

Table 1 (continued)

Order/suborder	Family	Taxon	GenBank Accession No.			
			28S	12S	16S	Rhodopsin
Callionymoidei	Callionymidae	<i>Callionymus lyra</i>	AY141541–44	AY141344	AY141414	AY141270
Gobioidei	Gobiidae	<i>Pomatoschistus</i> sp.	AY141537–40	AY141343	AY141413	
		<i>Pomatoschistus minutus</i>				<u>X62405</u>
Scombroidei	Sphyraenidae	<i>Sphyraena sphyraena</i>	AY141713–16	AY141386	AY141456	AY141312
	Scombridae	<i>Scomber japonicus</i>	AY141709–12	AY141385	AY141455	AY141311
Stromateoidei	Stromateidae	<i>Pampus argenteus</i>	AY141701–04	AY141383	AY141453	AY141309
	Centrolophidae	<i>Psenopsis anomala</i>	AY141705–08	AY141384	AY141454	AY141310
Channoidei	Channidae	<i>Channa striata</i>	AY141569–72	AY141351	AY141421	AY141277
Anabantoidei	Anabantidae	<i>Ctenopoma</i> sp.	AY141573–76	AY141352	AY141422	AY141278

Note 1. Classification following Nelson (1994) and listing order following the cladogram proposed by Johnson and Patterson (1993).

Note 2. Sequences retrieved from GenBank are underlined.

<sup>a</sup> C12D12 sequence retrieved from AF061801.

suborders and more than 60 families, plus outgroup taxa from seven different orders (Table 1). The sampling backbone followed the cladogram proposed by Johnson and Patterson (1993), one of morphological hypotheses we intended to test. All terminal clades are represented except Stephanoberyciformes and Elasmobranchii. For the questionable “Perciformes” clade, 41 species were chosen to represent 14 of the 18 recognized perciform suborders. When an order or suborder was represented by more than one taxon, species were sampled from different families, and if possible, including one from a putatively basal branch. Though this sampling strategy may decrease overall statistical support (Rannala et al., 1998), it is more likely to represent the evolutionary history of the group. From each sample, a small piece of muscle tissues was stored at  $-80^{\circ}\text{C}$  or fixed in 70% ethanol. DNA extraction followed the standard phenol/chloroform method described in Winnpenminck et al. (1993).

## 2.2. DNA amplification and sequencing

DNA amplification was performed by PCR (Mullis and Faloona, 1987; Saiki et al., 1988) in a 50  $\mu\text{l}$  volume containing 20 mM Tris-HCl, pH 8.55, 16 mM  $(\text{NH}_4)_2\text{SO}_4$ , 2.5 mM MgCl<sub>2</sub>, 150  $\mu\text{l}/\text{ml}$  BSA, 5% DMSO, 330  $\mu\text{M}$  of each dNTP, 0.3  $\mu\text{l}$  (1.5 U) of Hi-Taq polymerase (Bioprobe), 50 pmol of each of the two primers, and 0.3–1.2  $\mu\text{g}$  of template DNA. Primers used for amplifying different genes are listed in Table 2. PCR was carried out using a Biometra triblock cyler with denaturation at  $94^{\circ}\text{C}$  for 4 min; annealing temperature (AT) for 2 min; extension at  $72^{\circ}\text{C}$  for 2 min; followed by 29 cycles of ( $94^{\circ}\text{C}$ , 30 s, AT 30 s,  $72^{\circ}\text{C}$ , 30 s); and finally one step of  $72^{\circ}\text{C}$  for 4 min. The annealing temperature varied between 50 and  $60^{\circ}\text{C}$  depending on the species and the region amplified. PCR products were visualized and purified by agarose gel extraction using Qiaex II kit (Quiagen). The Thermo Sequenase Cycle Sequencing Kit (Amersham) was used for direct sequencing of the

purified PCR products using 5'  $\gamma$ -<sup>33</sup>P-labeled primers (the same primers used for PCR). The reacted samples were resolved by acrylamide-urea gel electrophoresis. Some internal primers were also necessary for completing the sequencing when PCR products were longer than 500 bp (See Table 2).

The rhodopsin gene used for this study is a member of the opsin gene family that has five main paralogous genes in vertebrates (Chang et al., 1995; Yokoyama, 1997). Organismal phylogeny can be misrepresented if genes used for analysis represent orthologous and paralogous copies. We used the following strategy to guarantee orthology among the rhodopsin sequences collected for this work. First, when designing primers for rhodopsin, we selected priming sites that differ among paralogous genes; divergence among rhodopsin and paralogous opsin genes is far greater than the divergence observed among all vertebrate rhodopsins. Second, other opsin genes have introns, unlike the rhodopsin genes of bony fishes (Fitzgibbon et al., 1995; Venkatesh et al., 1999). Third, the duplication event separating rhodopsin from other opsin genes occurred before the diversification of vertebrates (Yokoyama, 1997; Yokoyama et al., 1999). If we had sequenced by mistake a paralogous opsin gene, the sequence alignment would have shown this extreme divergence. So far, two studies reported two copies of the rhodopsin gene in fishes: Archer et al. (1995) for *Anguilla anguilla* and Lim et al. (1997) for *Cyprinus carpio*. These duplications are very recent events. Similar events among the present sample of fishes would have no impact on our phylogenetic inferences. Finally, the present study focuses on repeated clades obtained from different genes trees. It is very unlikely that an erroneous clade resulting from mistaken orthology would be obtained again in an independent gene tree. This is not a justification for orthology, but stresses that the “repeated clades” approach presented in this paper cannot be challenged by undetected paralogies.

Table 2  
Primers used in this study

Gene name	Sequences (5'–3')	Source
<b>28S</b>		
C1'	ACC CGC TGA ATT TAA GCA T	Lê et al. (1993)
C2	TGA ACT CTC TCT TCA AAG TTC TTT TC	Lê et al. (1993)
C3'p	CCG YGG CGC AAT GAA AGT GA	This study
C6'p	TCA CCT GCC GAA TCA ACT AGC	This study
C7	ACT ACC ACC AAG ATC TGC AC	This study
C12'p	TTA TGA CTG AAC GCC TCT AAG	This study
<i>D12pr</i>	TGA CTT TCA ATA GAT CGC AG	This study
<i>D12r-acan</i>	AGC ACC AGG TTC TCC ACA AAC A	This study
<b>12S</b>		
L1091R	AA ACT GGG ATT AGA TAC CCC ACT AT	Kocher et al. (1989)
<i>H1478</i>	TGA CTG CAG AGG GTG ACG GGC GGT GTG T	Kocher et al. (1989)
<b>16S</b>		
16S INT	GGT CCG CCT GCC CTG TGA C	This study
16S INT bis	CCG CGG TAT TTT GAC CGC G	This study
<i>16S INT bis</i>	GGA TGT CCT GAT CCA AC	This study
<b>Rhodopsin</b>		
Rh193	CNT ATG AAT AYC CTC AGT ACT ACC	This study
Rh545	GCA AGC CCA TCA GCA ACT TCC G	This study
<i>Rh667r</i>	AYG AGC ACU GCA UGC CCU	This study
<i>Rh1039r</i>	TGC TTG TTC ATG CAG ATG TAG A	This study
<i>Rh1073r</i>	CCR CAG CAC ARC GTG GTG ATC ATG	This study

Italics: reverse primers.

### 2.3. Data management and sequence alignment

Sequences were read and entered twice using the MUST package (Philippe, 1993). Data matrices were first prepared using ED of MUST. Sequence files were then exported to Se-Al (Rambaut, 1996), for future data management. The possibility of sequencing errors resulting from sample mix-up or contamination was checked by comparing our sequences to the sequence of a second exemplar or of a putatively closely related taxon, or to sequences from GenBank using BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>). Additional sequences were retrieved from GenBank (Table 1); these were previously described in papers by Archer et al. (1992), Archer and Hirano (unpublished), Bargelloni et al. (1994), Hunt et al. (1997), Miya and Nishida (1996), Ritchie et al. (1997), Saitoh et al. (2000), Tang et al. (1999), Venkatesh et al. (1999) Wiley et al. (1998), and Zardoya and Meyer (1996). We did not include sequences available from GenBank after the year 2000 in our mitochondrial data set (e.g., Miya et al., 2001). Ongoing research of high-order actinopterygian phylogeny based on large molecular data sets, including 12S and 16S mtDNA fragments, is described elsewhere ([www.deepfin.org/](http://www.deepfin.org/)).

For the 28S, 12S, and 16S rDNA fragments, preliminary alignments were achieved using CLUSTAL X (Thompson et al., 1997) with default gap penalties. These were subsequently adjusted by eye on the basis of their secondary structure. Inclusion of secondary struc-

ture information in alignment for phylogenetic studies based on ribosomal genes is strongly recommended (Buckley et al., 2000; Hickson et al., 1996; Hickson et al., 2000; Kjer, 1995; Morrison and Ellis, 1997; Page, 2000; Titus and Frost, 1996). For the 12S and 16S rDNA data sets, we localized stems (base-paired regions) and loops (non-paired regions) following the secondary structure models published for sternoptychids (Miya and Nishida, 1998), *Pygocentrus nattereri* (Ortí and Meyer, 1996), *Fundulus heteroclitus* (Parker and Kornifield, 1996), and *Galaxias brevipinnis* (Waters et al., 2000). For the 28S rDNA data set (domains: C1–C2, D3, D6) we followed the *Xenopus laevis* model of Maidak et al. (1999), and for 28S domain D12, the model of Chen (2001) for acanthomorphs. Stem regions were first aligned following the protocol proposed by Kjer (1995) and Hickson et al. (1996); detailed procedures also were described by Chen (2001, pp. 64–69). The pairing regions were checked by identification of compensatory mutations between stem pairs.

Loop regions were aligned according to sequence similarity or conserved motifs. Major alignment structures given by CLUSTAL X were conserved but adjusted manually to avoid the discontinuity of individual gaps. Instead of deleting some “variable” sequences, which might contain high degree of homoplasy (Wiley et al., 2000) or “ambiguous” alignment sites (Miya and Nishida, 1998; Yamaguchi et al., 2000), we attempted to keep most sequences to extract the maximum amount of information for phylogenetic analysis. Most

‘ambiguous’ regions are alignable with careful observation following secondary structure models. Furthermore, an improved model for alignment can only be obtained by comparing a diversity of taxa. For instance, Waters et al. (2000) showed that the helices G8–G14 (encompassing variable regions from *l* to *n*) in the 16s model of Alves-Gomes et al. (1995) and Ortí (1997) were improperly paired or absent in more divergent taxa, resulting in a large loop. In general, the definition of ambiguous regions in sequence alignments remains subjective and arbitrary and high levels of variability in these regions usually result from a few divergent taxa. In this paper, only three large insertion/deletions segments showing high dissimilarity in sequence length as well as composition were excluded from phylogenetic analysis. These correspond to loop regions in the 28s data set (D3 domain from 342 to 356 and D12 domain from 676 to 686), and in the 16S data set (G10 region of Waters et al. (2000) or stem 40 of Miya and Nishida (1998) from positions 683–713). The alignments are available upon request.

#### 2.4. Phylogenetic analysis

Three data partitions were defined conditioned on putative gene independence in terms of both, functional constraints and selective pressures (Slowinski and Page, 1999): 28s (ribosomal nuclear gene, variable domains C1–C2, D3, D6, and D12 together), MT (partial mitochondrial ribosomal genes 12s and 16s together), and rhodopsin (nuclear visual pigment coding gene, corresponding to transmembrane domains II–VII).

Phylogenetic trees were reconstructed by unweighted maximum parsimony (MP) and model based methods: minimum evolution (ME; Rzhetsky and Nei, 1992) and Maximum Likelihood (ML; Felsenstein, 1981), as im-

plemented in PAUP\* version 4.0b8 (Swofford, 2001). MP trees were obtained by heuristic search with random stepwise addition sequences (MULPARS on) followed by TBR swapping, using 1000 replicates for the MT and rhodopsin data sets and 100 replicates for the 28s data set. Gaps were treated as fifth state, because short length variations in loops are more likely due to discrete insertion/deletion events than to single events involving multiple positions in the sequence. Otherwise, large stretches of indels have been removed from the analysis (see above). To obtain ME and ML optimal trees, a neighbor-joining tree (NJ; Saitou and Nei, 1987) was used as a starting tree for heuristic searches with TBR and NNI branch swapping under the ME and ML criteria, respectively. The distance measure used for ME searches is based on the maximum likelihood model (Waddell and Steel, 1997), as described below.

Likelihood ratio tests (Goldman, 1993; Huelsenbeck and Crandall, 1997), as implemented in MODELTEST 3.06 (Posada and Crandall, 1998), were employed to choose models for ML and ME analyses. The following models were suggested by MODELTEST: TrN + *G + I*, TrN + *G + I*, GTR + *G + I*, GTR + *G + I* for 28S, MT, rhodopsin, and the combined data set, respectively (Gu et al., 1995; Lanave et al., 1984; Rodríguez et al., 1990; Tamura, 1993; Tavaré, 1986). A test of homogeneity of base frequencies across taxa using  $\chi^2$  test was performed using PAUP\* and Puzzle 4.02 (Strimmer and von Haeseler, 1996) (see Table 3). Although the null hypothesis of homogeneity of base composition across taxa was not rejected (*p*-value > 5%) for each data set, the *p*-value for the rhodopsin was very close to the 5% threshold (9%). When the same tests were repeated for each codon position of the rhodopsin gene separately, the third codon position exhibited an extremely high degree of heterogeneity in base composition across taxa

Table 3  
Descriptive statistics for maximum parsimony and maximum likelihood analyses

	28S	MT	Rhodopsin	Combined
No. taxon	74	97	86	72
Length of sequences (bp)	847	814	759	2420
No. parsimony-informative sites	246	490	352	1052
No. variable sites (in %)	405 (48%)	619 (76%)	442 (58%)	1428 (59%)
A	0.1808	0.3077	0.1794	N/A
C	0.2973	0.2568	0.2931	N/A
G	0.3588	0.2197	0.2503	N/A
T	0.1632	0.2158	0.2772	N/A
Base frequencies homogeneity ( <i>p</i> )	1.00	0.99999	0.09115623	N/A
MP tree length	1616	7484	3816	11181
No. MP trees	>4813	8	912	5
No. isl. from MP	5	1	8	1
C.I.	0.4	0.18	0.21	0.27
R.I.	0.46	0.32	0.42	0.33
ML tree score	8417.345	27141.26	17719.14	47147.1444
Alpha shape parameter ( <i>G</i> )	0.27411	0.57636	0.75108	0.66761
Proportion of invariable site ( <i>I</i> )	N/A	0.22712	0.367286	0.3608

( $p$ -value  $< .00001$ ). The deviant taxa detected by the chi-square test in Puzzle 4.02 are indicated as open or full circles in the rhodopsin tree. Therefore, for rhodopsin, the LogDet distance (Lockhart et al., 1994) was also employed for ME analysis. The LogDet distance ME tree was constructed according to the suggestion of Swofford et al. (1996): constant sites were removed in proportion to base frequencies estimated from constant sites only. Model parameters were all estimated via maximum likelihood as implemented in PAUP\* through an iterative process. For each data partition, MP trees were taken as a starting point (Swofford et al., 1996, p. 445) and used for the initial estimation of  $G$  (gamma shape parameter) and  $I$  (proportion of invariable sites, pinvar). These parameters were then fixed and used in heuristic searches (under ME and ML criteria). The new topology obtained was used to re-optimize the parameter values and another search was started with the new parameters. Cycling between parameter estimation and optimal tree searching was continued until the same topology was found in successive iterations (Swofford et al., 1996, p. 445).

Bootstrap analysis was used to assess the robustness of clades (Felsenstein, 1985). Full heuristic searches with TBR branch-swapping were conducted for 100 replicates (MP method) with 20 random addition sequences for each replicate, and for 500 replicates (ME method). The bootstrap procedure could not be applied to the ML method due to computer time limitation.

Clade repeatability was used as a central criterion to assess reliability of our results. Repeated clades were preliminarily determined through comparing separate phylogenetic ME, MP, and ML trees without consensus. Consensus techniques were not directly used because special attention was paid not only to repeated clades but also to: (1) the number of times that particular clades occur over other alternatives, (2) branch lengths, in order to assess potentially misleading groups, (3) to detect nearly repeated clades (e.g., when a single taxon escapes once from a clade in only one tree with most of the other taxa remaining within the clade, as may be the case under rate acceleration in the evolution of the corresponding gene in this taxon);. All this information might be lost in a strict consensus tree. Most importantly, classic consensus techniques cannot be used to summarize trees with unequal number of taxa.

ME trees from the different data partitions were used as a first step to identify repeated clades because: (1) model-based trees are generally more consistent than trees constructed by equal weighted MP method, which are more sensitive to long branch attraction (Huelsenbeck and Hillis, 1993; Sullivan and Swofford, 2001); and (2) ML trees obtained here rely only on less rigorous heuristic searches (using NNI branch swapping). When repeated clades from ME trees were not found in either the MP or ML trees, these particular clades were en-

forced as constraints for new MP and ML searches. We then tested whether the constrained trees were significantly worse than the optimal tree using the Kishino and Hasegawa test (Kishino and Hasegawa, 1989) implemented in PAUP\*. In all cases tested, no significant difference ( $p$ -value  $> 0.005$ ) (Table 4) was detected between constrained trees and optimal trees, suggesting that there were no real conflicts for assigning repeated clades between different methods.

While clade repeatability across data partitions is used here as a criterion for reliability, we need a tree on which comments will be made and the history of characters studied. We acknowledge the need for simultaneous analysis, basing that tree on the whole available data. Maximizing congruence among all available characters and using a larger number of characters should provide the tree that best summarizes the results, the tree on which character evolution must be studied. In other words, for obtaining a topology, the simultaneous analysis is the best approach, and for assessing reliability of the clades the appropriate approach is separate analyses without consensus trees. This point of view is summarized in Fig. 1, indicating that priority is given to the criterion of repeatability over simple bootstrap proportions to assess the reliability of clades found in the tree derived from simultaneous analysis.

To gain further insight on repeated clades and to overcome some shortcomings of taxonomic congruence based only on optimal trees, we develop a protocol called repeated-bootstrap components. Many authors have stressed that the reliance on optimal trees only ignores the fact that all phylogenetic estimates are made with some degree of error and uncertainty (Lanyon, 1993; Miyamoto and Fitch, 1995; Penny and Hendy, 1986). Therefore, taxonomic congruence should consider not only optimal trees but also near-optimal trees (Hillis, 1995; Rodrigo et al., 1993; Swofford, 1991). Because the bootstrap procedure is designed to assess errors due to limited sample size (Page and Holmes, 1998, pp. 218–222), some clades not found in the optimal trees may appear in majority-rule bootstrap consensus trees. To make use of this potential information we mapped bootstrap values for repeated clades from each data partition (repeated-bootstrap components) onto the optimal tree from simultaneous analysis (e.g., Fig. 5, ME tree). Bootstrap repeated components for each data partition are shown in the form of a histogram for each node Fig. 5. If bootstrap values can be regarded as measure of hierarchical signal (Hillis and Bull, 1993), under our criterion of repeatability, the histogram mapped on the tree can be interpreted as the contribution of phylogenetic signal from each data partition to support the corresponding node. To extend this protocol further, we relax our criterion of reliability and also score the repeated clades found in the listing of bootstrap bipartitions produced by PAUP\* for each separate

Table 4

Recurrent clades found in the analysis of taxonomic congruence from ME, MP, and ML trees from three molecular and combined data sets

Clade	Taxa included	28S			MT			Rhodopsin			Combined			Morphological hypothesis
		ME	MP	ML	ME	MP	ML	ME	MP	ML	ME	MP	ML	
A	<i>a1 + a2</i>	<b>X</b>	<b>X</b>	x	x	x	x				<b>X</b>	<b>X</b>	x	Inclusion of zeioids in the Paracanthopterygii (1)
<i>a1</i>	<i>Gadus, Merlangius</i>	<b>X</b>	<b>X</b>	x	<b>X</b>	<b>X</b>	x	<b>X</b>	<b>X</b>	x	<b>X</b>	<b>X</b>	x	Monophyly of the Gadiformes (2)
<i>a2</i>	<i>Zeus, Neocyttus</i>	<b>X</b>	<b>X</b>	.	<b>X</b>	<b>X</b>	x	<b>X</b>	<b>X</b>	x	<b>X</b>	<b>X</b>	x	Monophyly of Zeioidei (2)
B	<i>Hoplostethus, Beryx</i>	x	.	.				x	.	.	<b>X</b>	.	x	Trachichthyoids plus berycoids (2,9)
C	<i>belonoids (a), atheriniforms (b)</i>	x	x	x				x	.	x	x	.	x	
D	<i>d1 + d2</i>	x	x	x				x	x	.	x	x	x	Gobiesocoids plus blennioids (3)
<i>d1</i>	<i>Parablennius, Forsterygion</i>	x	x	x	x	.	.	x	.	.	<b>X</b>	.	x	Monophyly of the Blennioidei (2,4)
<i>d2</i>	<i>Lepadogaster, Apletodon</i>	<b>X</b>	<b>X</b>	x	x	x	x	<b>X</b>	<b>X</b>	x	<b>X</b>	<b>X</b>	x	Monophyly of the Gobiesocoidei (2)
E	<i>Macroramphosus + e</i>	x	.	x	x	x		x	.	x	<b>X</b>	x	x	Inclusion of dactylopterids in syngnathoids (5)
<i>e</i>	<i>Aulostomus, Dactylopterus</i>	x	x	x	x	x	x	x	.	x	<b>X</b>	<b>X</b>	x	
F	<i>f1 + f2</i>					x	x	<b>X</b>	x	x	<b>X</b>	<b>X</b>	x	Relationships of channoids and synbranchiforms (6,10)
<i>f1</i>	<i>Channa, Ctenopoma</i>	x	x	x				<b>X</b>	<b>X</b>	x	<b>X</b>	.	x	Relationships between channoids and anabantoids (12)
<i>f2</i>	<i>Mastacembelus, Monopterus</i>				x	x	x	<b>X</b>	<b>X</b>	x	<b>X</b>	<b>X</b>	x	Monophyly of the Synbranchiformes (2,7,8,9)
G	<i>Ammodytes, Cheimarrichthys</i>	x	x	x	x	x	x				x	x	x	
H	<i>Scomber, Psenopsis, and h1</i>	x	x	.	x	.	x	x	x	x	<b>X</b>	<b>X</b>	x	
<i>h</i>	<i>Pampus, Kali</i>	x	x	x				x	.	x				
I	<i>Taurulus + i</i>				x	x	x		x	x	<b>X</b>	<b>X</b>	x	
<i>i</i>	<i>Austrolycus, Pholis</i>	x	x	x	x	x	.	x	<b>X</b>	x	<b>X</b>	<b>X</b>	x	Monophyly of the Zoarcoidei (2)
K	<i>k1 + k2</i>	x	.	.	x	.	.	x(f)	.	.	x	x	x	
<i>k1</i>	<i>Perca, Gymnocephalus</i>	x	<b>X</b>	x	<b>X</b>	<b>X</b>	x	<b>X</b>	<b>X</b>	x	<b>X</b>	<b>X</b>	x	Monophyly of the Percidae (2)
<i>k2</i>	<i>Notothenioids (c)</i>	x	.	.	x	x	x		x	x	<b>X</b>	x	x	Monophyly of the Notothenioidei (2)
L	<i>Carangids (d), Mene, Ech., Sphy., Pent., Lates, pleuronectiforms (e)</i>				x	x	x	x(g)	.	x(g)	x(g)	x(g)	x(g)	
M	<i>Labrus + Scarus</i>				x	.	.	<b>X</b>	x	x	<b>X</b>	<b>X</b>	x	Relationships of labrids and scarids (12)

Note 1. Crosses in ME, ML, and MP column means that clade is found from ME, ML trees, and the strict consensus of MP trees, respectively. Crosses in bold indicate bootstrap proportions over 80% (bootstrap values did not apply for ML because of huge computation time). Dots means there is no significant contradiction between the MP or ML trees and their equal or near-optimal trees chosen for exhibiting a particular recurrent clade found in ME trees, as confirmed by Kishino–Hasegawa tests (two tailed) performed using PAUP\*.

Note 2. Taxon abbreviations: *Ech.*, *Echeneis*; *Sphy.*, *Sphyaena*; *Pent.*, *Pentanemus*.

Note 3. Recurrent clades within the Notothenioidei not shown.

Note 4. For 28S model-based trees, one parameter-less model (TrN + G) is used here: see results for explanation. *a*: Taxa included are *Belone* (present in all data sets) and *Dermogenys* (present only in MT). *b*: Taxa included in all data set is *Bedotia*; additional taxon included in Rhodopsin data set is *Atherina*. *c*: Taxa included are *Bovichtus*, *Cottoperca*, *Pseudaphritis*, *Notothenia*, *Eleginops* (only in Rhodopsin), *Chionodraco* (only in MT), *Neopagetopsis* (only in 28S). *d*: Taxa included are *Chloroscombrus*, *Trachinotus*, *Caranx* (only in MT). *e*: Taxa included in all data sets are *Arnoglossus* and *Microchirus*; additional taxon included in MT is *Paralichthys*; additional taxa included in Rhodopsin data set are *Psettodes*, *Citharus*, and *Solea*. *f*: One notothenioid taxon, *Pseudaphritis*, escapes from the clade K. *g*: One pleuronectiform taxon, *Arnoglossus*, escapes from the clade L.

(1) Gayet (1980b, 1980c); (2) Nelson (1994); (3) Rosen and Patterson (1990); (4) Spriger (1993); (5) Pietsch (1978); (6) Lauder and Liem (1983); (7) Gosline (1983); (8) Travers (1984a, 1984b); (9) Johnson and Patterson (1993); (10) Roe (1991); (12) Gosline (1971).

data partition (repeated-bootstrap components). Non-repeated clades are not considered reliable, even if bootstrap support for a clade is high for a single data partition (lower left case in Fig. 1). For example, assume that three data partitions are analyzed separately. If clade X receives a bootstrap value of 40% from the first data partition, of 60% from the second, and of 20% from the third, the composite bootstrap value will be calculated as  $(50 + 60 + 20)/300 = 43\%$ . If clade Y has a bootstrap value of 20% for the first data partition, 80% from the second and was not found in the third, the final value will be  $(20 + 80)/300 = 33\%$ . If clade Z receives a bootstrap value of 95% in a single data partition but is not found in the other two, it will not be considered reliable because it is not repeated. To define the lower bootstrap value for discarding clades, we generated 500 random trees and found that the frequency of clades appearing by chance was maximally 2.2% for 72 taxa data matrices. We discarded clades with bootstrap values  $<6\%$ . The computer program of repeated-bootstrap components is available upon request.

Although this protocol may avoid a strong misleading signal, it may also annihilate a single strong “truly phylogenetic signal” when this signal is not found in the other two (or more) data partitions. This protocol is based on two assumptions: (1) if partitions are really independent, misleading signals (e.g., caused by base compositional bias) are likely to be restricted to a single data partition rather than reproduced in all partitions. (2) Such artifacts are more commonly found than the situation where a “real” and strong phylogenetic signal is clearly contained in a single data set but not in others. In other words, it is assumed that we have higher chances of detect a “true” signal through repeatability than through a single high bootstrap proportion.

### 3. Results

#### 3.1. Characterization of nucleotide substitution patterns

Sequences were successfully obtained using the primers listed in Table 2 for all species except *Pogonoperca punctata* and *Fistularia petimba* (for the 28S domains D3, D6, and D12). The rhodopsin sequence of *Lampris immaculatus* could only be amplified successfully using primers rh545 and rh1073. The 5' end portion (321 bp) of the rhodopsin gene for this taxon was replaced by question marks. If failure of amplification of rhodopsin in *L. immaculatus* is not related to the presence of an intron, our results seem to support the intronless hypothesis for this gene in ray-finned fishes except bichirs (Fitzgibbon et al., 1995; Venkatesh et al., 1999). All rhodopsin sequences obtained contain a single open reading frame. For the three data partitions we were able to collect sequences for a common set of 72 taxa of

the same species, or at least the same genus. The length of the aligned sequences (after removal of a few ambiguous alignment regions in 28S as described above), the total number of taxa, and other descriptive statistics for each data set are summarized in Table 3. Sequences have been deposited in GenBank with accession numbers listed in Table 1.

Base composition among sequences differs among genes but not among taxa within genes (Table 3). The second codon positions of rhodopsin exhibit an excess in T, whereas the third codon positions are high in C. In contrast to the mitochondrial cytochrome *b* gene of all fishes that typically exhibit an anti-G bias at third codon positions (Lydeard and Roe, 1997), the third codon positions of rhodopsin show a relatively low frequency of A. This is similar to other two nuclear genes characterized for fishes: mixed-lineage leukemia-like (Mll) and ependymin (Chen, 2001; Ortí and Meyer, 1996). However, the high frequency of T at the second codon positions of rhodopsin is similar to cytochrome *b*, reflecting a strong functional constraint. Both genes code for transmembrane proteins that are rich in hydrophobic amino acid residues: Phe (TTY), Leu (TTR or CTR), and Ile (ATY) (Naylor et al., 1995).

The MP trees for each data partition were used to estimate the frequencies of nucleotide changes using MacClade version 3.07 (Maddison and Maddison, 1992). As expected, inferred transitions occur more frequently than inferred transversions and the MT sequences showed a remarkable low frequency of G–C and G–T interchanges. Bubble diagrams of inferred nucleotide changes for mitochondrial and nuclear data show heterogeneity among substitution types, in agreement with the choice of parameter-rich models selected by MODELTEST (see above). The absolute saturation test (Hassanin et al., 1998; Lavoué et al., 2000; Philippe et al., 1994; Philippe and Forterre, 1999) was performed on transitions and transversions for each gene and codon position taken separately (saturation plots available upon request, Chen, 2001). Only the transitions in 12s and 16s genes exhibit a clear plateau suggesting a high frequency of multiple substitutions. According to these observations, in addition to the MP method, the alternative model-based approaches are justified for phylogenetic reconstruction (Swofford et al., 1996).

#### 3.2. Phylogenetic analyses

The ME trees obtained from separate analyses are shown in Figs. 2–4. The shapes of the trees inferred differ among data partitions. The 28S tree (Fig. 2) may be divided into two sections, a basal section with relatively long internal branches and a terminal component that looks like a radiation (with very short internal branch lengths). In contrast, the MT tree (Fig. 3) is somewhat homogeneous in terms of extreme terminal

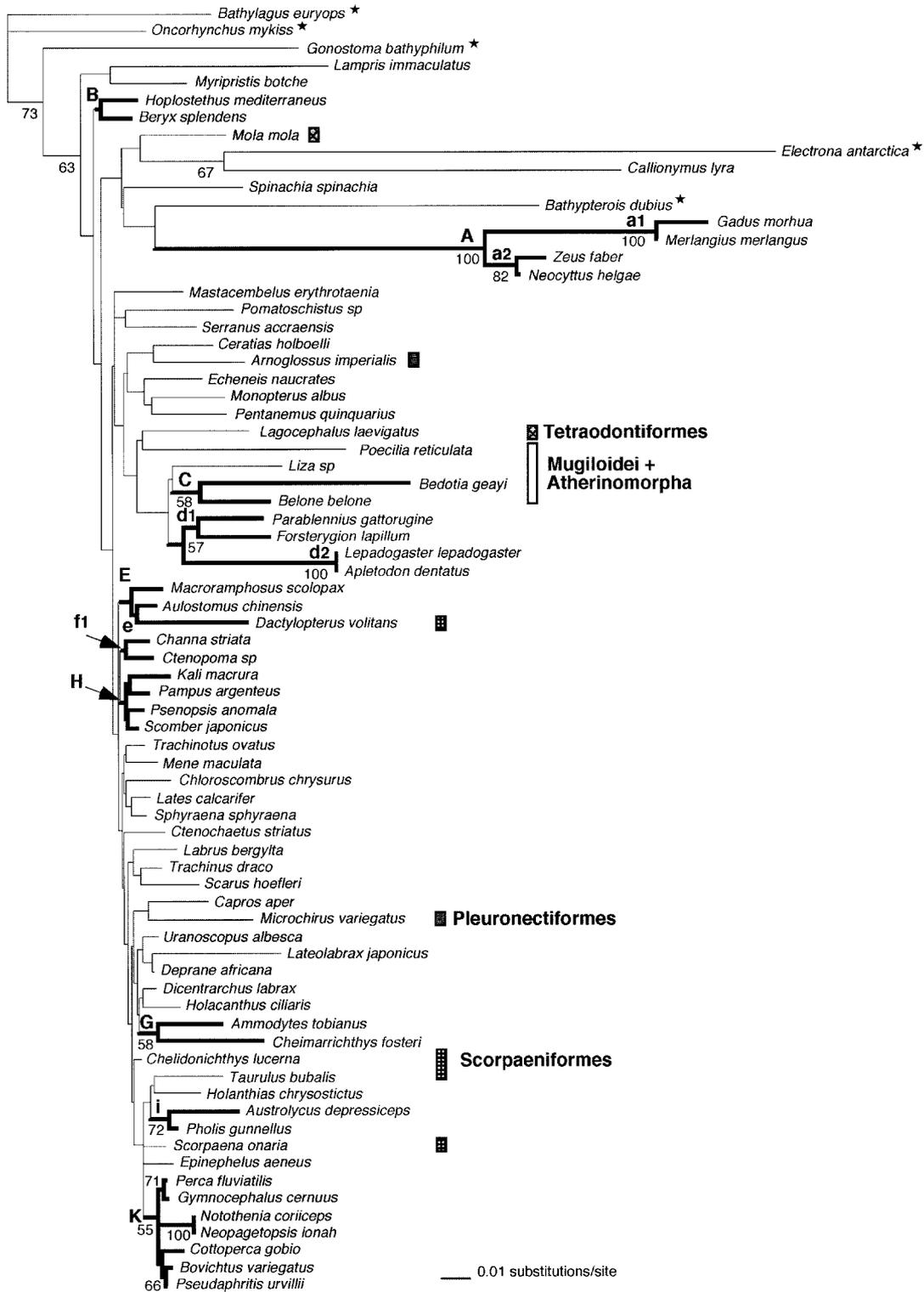


Fig. 2. Minimum evolution (ME) tree from partial 28s sequences (domains: C1–C2, D3, and D12) using TrN + G transformation maximum likelihood distance (instead of TrN + G + I model as suggested by MODELTEST, one parameter-less model is used here: see results for explanation). ME score is 2.41272. The branch lengths are proportional to inferred distances. Outgroups are marked with a star. Clades indicated by letters are recurrently found at least twice out of three of our data sets. They are shown with bold lines. Numbers represent bootstrap proportions from 500 replicates.

branch lengths and short deep internal branches, reflecting the mutational saturation detected above. The rhodopsin tree (Fig. 4) has longer deep branches, but

has an asymmetrical base and a symmetrical crown. Such a tree shape must be interpreted with caution since the long deep branches may suffer long-branch attrac-



Fig. 3. Minimum evolution (ME) tree from mitochondrial sequences (MT) of part 12S and 16S genes using TrN + G + I transformation maximum likelihood distance. ME score is 13.93740. The branch lengths are proportional to inferred distances. Outgroups are marked with a star. Clades indicated by letters are recurrently found at least twice out of three of our data sets. They are shown with bold lines. Numbers represent bootstrap proportions from 500 replicates.



Fig. 4. Minimum evolution (ME) tree from partial rhodopsin sequences using GTR + G + I transformation maximum likelihood distance. ME score is 6.28146. The branch lengths are proportional to inferred distances. Outgroups are marked with a star. Clades indicated by letters are recurrently found at least twice out of three of our data sets. They are shown with bold lines. Numbers represent bootstrap proportions from 500 replicates. Taxa with significant higher GC contents at the third codon position, as detected by  $\chi^2$  tests, are indicated with open circles. Taxa with significant lower GC contents at the third codon position, as detected by  $\chi^2$  tests, are indicated with full circles.

tion towards divergent outgroup sequences (Philippe and Adoutte, 1998; Philippe and Laurent, 1998; Philippe et al., 2000).

Although the branching order is not always the same across the three ME trees, the most basal acanthomorphs are the same: Lampridiformes (*Lampris*, *Metavelifer*), Polymixiiformes (*Polymixia*), Paracanthopterygii (*Percopsis*, *Gadus*, and *Merlanginus*), Beryciformes (*Myripristis*, *Beryx*, and *Hoplostethus*), Zeioids (*Zeus*, *Neocyttus*). Interestingly, gadids (clade a1) and zeioids (clade a2) are sister-groups in the 28S and MT trees, a finding already obtained by Wiley et al. (2000) by analyzing another portion of the 28S gene and by Miya et al. (2001) from whole mitogenomic data. The Lophiiformes (*Ceratias*), a member of Paracanthopterygii, are not present in the basal group but appear among the more derived percomorph lineages, as already reported by Lê et al. (1993). Compared to the other two trees, the base of the rhodopsin tree (Fig. 4) is rather surprising because the basal acanthomorphs indicated above are separated by several percomorph lineages, most noteworthy components *a1* + *a2*, from clade B containing beryciforms and lampridiforms. A close analysis of base compositional bias using the  $\chi^2$  homogeneity test suggests that this bias could play a role in determining the deep branching of these percomorph lineages. In fact, heterogeneity of base composition among taxa in the rhodopsin data set is very high at third codon positions. The GC3 (GC contents at third codon positions) ranged from 58% (*Aulostomus chinensis*) to 92% (*Callionymus lyra*). Indeed, most of the deviant taxa with high GC3 (marked with open circles in Fig. 4) are concentrated at the base of tree, perhaps attracted to each other by similar base composition. Some basal acanthomorphs, including *Lampris*, *Hoplostethus*, and *Beryx*, and one outgroup taxon (*Bathypterois*), have base composition close to the average value and are placed in the tree closer to the central percomorph crown. In general, the degree of congruence between ME trees is low and none of the three data partitions alone recovers traditionally recognized monophyletic groups, such as Acanthomorpha, Paracanthopterygii, Scorpaeniformes, Pleuronectiformes, Smegmarphorpha (Johnson and Patterson, 1993), and Percomorpha. However, tree comparisons allow the identification of repeated clades (bold lines in Figs. 2–4), found in at least two of the three ME trees (Table 4).

Results from MP and ML analyses support the identification of repeated clades found by the ME approach, with the single exception of the 28S data partition. ML and ME analysis of 28S based on the TrN + I + G model selected by MODELTEST failed to identify about half of the repeated clades, and supported a rather atypical tree. Reanalysis of 28S under ML and ME with a simpler TrN + G model (one parameter-less than TrN + I + G) produced results more in line with the

other analyses (including the analyses with varied simpler models) and identified all the repeated clades. The fact that use of particular models of nucleotide substitution may change the results of model-based analysis has been previously recognized (e.g., Cunningham et al., 1998; Kelsey et al., 1999; Leitner et al., 1997; Sullivan and Swofford, 1997). Accuracy and consistency of phylogenetic results generally depend on assuming the “right” model of evolution, and several statistical procedures to identify the best-fit model have been proposed (e.g., Huelsenbeck and Crandall, 1997). However, the parameter-rich model selected by MODELTEST in this study seems to be less accurate or consistent than a simpler (“wrong”) model (TrN + G) that was rejected by this test. A few unusual cases of phylogenetic bias, by which “incorrect” models can give “correct” answers, have been identified in both simulation (e.g., Saitou and Nei, 1987; Takahashi and Nei, 2000; Tateno et al., 1994; Yang, 1997) and empirical studies (Posada and Crandall, 2001). The causes for this bias may be complex and perhaps related to problematic alignment among highly divergent sets of sequences (Posada and Crandall, 2001).

The scores and corresponding statistics of MP and ML trees are summarized in Table 3 (trees are not shown but are available upon request). Repeated clades found by all three analytical approaches are summarized in Table 4. Although, the topologies are not always identical between the trees constructed by different methods, choice of method generally has a very weak impact on repeatability (except for the case of 28S, as discussed above). There are more differences between data partitions than between methods applied to the same data. Moreover, there is no significant contradiction between the MP or ML trees and their equal or near-optimal trees chosen for exhibiting a particular recurrent clade found in ME trees (dots in Table 4), as confirmed by Kishino–Hasegawa tests.

### 3.3. Simultaneous analysis

Analyses of all data combined were performed under three different phylogenetic methods—MP, ML, and ME. A summary of the results is presented in Table 3, and the ME tree is shown in Fig. 5. Topological differences among trees obtained from the different methods are smaller than the differences among trees from different data sets. All recurrent clades identified by the separate analyses also are recovered in the tree from simultaneous analysis (bold lines in Fig. 5, Table 4) and the bootstrap support for these clades increased dramatically, indicating additive phylogenetic signal from each data partition to confirm the same clades. Some patterns previously suspected to be influenced by base compositional biases (high GC content in rhodopsin) still persist in the tree obtained from simultaneous analysis (e.g., a group containing *Spinachia*, *Taurulus*

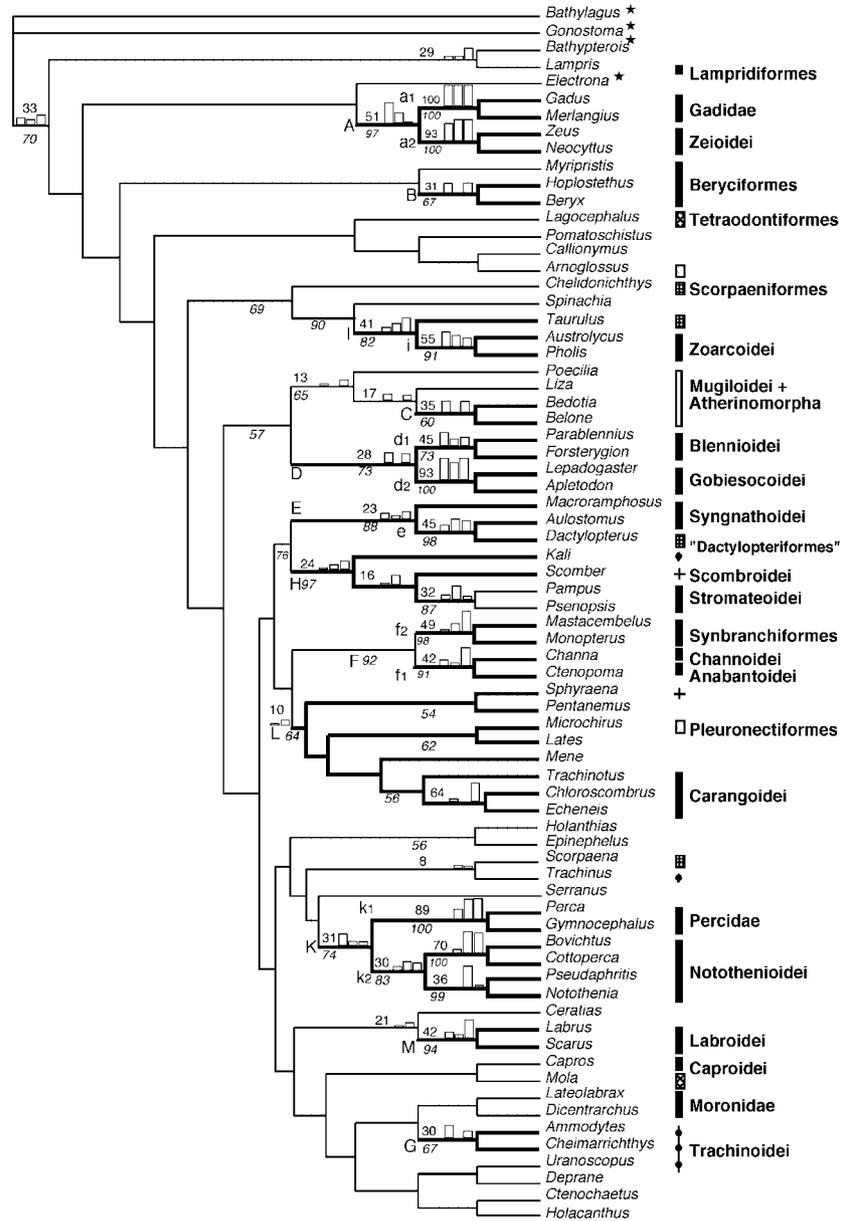


Fig. 5. Minimum evolution (ME) tree from simultaneous analysis using GTR + G + I transformation maximum likelihood distance. ME score is 6.21173. Bold lines: clades which are recurrent across separated analyses (Figs. 2–4 and Table 4). Two kinds of bootstrap proportions (BP) are shown. Values above branches are for repeated-bootstrap components (only those ones congruent with this ME tree are shown). Classical BPs from simultaneous analysis are values below branches (BPs below 50% not shown). Bootstrap resamplings are performed with 500 replicates. Small histograms over branches are the BPs from repeated-bootstrap components. They are BPs taken from each of the three listings of repeated bipartitions obtained from separate bootstrap analyses using the same taxa as simultaneous analysis, displayed for 28s, MT, and rhodopsin, respectively.

and two zoarcoids). Though basal acanthomorphs and outgroup taxa are not clearly separated into two blocks as seen in the rhodopsin tree, some deviant taxa from derived groups with high GC content (e.g., *Arnoglossus*) are still placed among the basal groups.

Repeated clades, also recovered in the topology obtained from simultaneous analysis suggest that the following groups should be considered reliable: Gadidae (cods, clade a1); Zeioidei (dorries, clade a2); Zeioidei + Gadidae (clade A); Trachichthyoidei + Berycoidei

(clade B); Beloniformes + Atheriniformes (needlefishes and silversides, clade C); blenioids (blennies, clade d1); Gobiesocoidi (clingfishes, clade d2); Gobiesocoidi + Blennioidei (clade D); Dactylopteriformes + Syngnathoidi (flying gurnards, trumpetfishes, and snipefishes, clade E); Channoidei + Anabantoidei (snakeheads and climbing gouramies, clade f1); Mastacembeloidei + Synbranchioidei (spiny eels and swamp-eels, clade f2); Ammodytidae + Cheimarrichthyidae (sand lances and torrentfish, clade G); Stromatoidi-

dei + Scombridae (mackerels) + Chiasmodontidae (clade H); Zoarcoidei (eelpouts, clade *i1*); Percidae (perches, clade *k1*); Notothenioidei (Antarctic fishes, clade *k2*); Percidae + Notothenioidei (clade K), and a bigger clade L grouping Carangidae (jacks), Sphyrnidae (barracudas), Echeineidae (remoras), Polynemidae (threadfins), Menidae (moonfish), Centropomidae (snooks), and Pleuronectiformes (flatfishes). There are two special clades: Cottoidei + Zoarcoidei (sculpins and eelpouts, clade I), and Clade F (*f1* + *f2*), which are recurrently found only in trees obtained using MP and ML methods and found as well in all trees from simultaneous analysis regardless of method.

Further investigation of repeated clades using the protocol of repeated-bootstrap components identified clades that were previously not found in any of the optimal trees based on separate analyses, possibly due to sampling error. For example, clade I, not found in the 28S and rhodopsin ME trees is of interest because it is found in the bootstrap listings of all three data partitions. The following repeated-bootstrap clades have very weak partition bootstrap support: a clade grouping Mugiloidei and Atherinomorpha; the clade (*Scorpaena* + *Trachius*); the clade (*Ceratias* + *labroids*), which require further investigation.

The histograms at each node in the simultaneous tree (Fig. 5) show bootstrap values for each partition. In general, phylogenetic signal—as judged by bootstrap support—is not homogeneously distributed across different data partitions nor throughout the tree. It appears that the contribution of the MT data is rather weak, while the rhodopsin data set contains more phylogenetic signal for derived clades. This is also indicated by the consistency (CI) and retention (RI) indices (Table 3), which are relatively low for the MT data, reflecting more homoplasy content in mitochondrial than in other data sets. Regarding particular parts of tree, the 28S data seem to perform well for resolving the interrelationships within clades A and D, while performing poorly within the “perciform” crown (as indicated as well by short internodes in the 28S tree, Fig. 2). The rhodopsin data perform well for resolving inter-relationships within clade F. This information may be useful for future studies of particular acanthomorph groups, by focusing on signal-rich genes when the target taxonomic samples become available.

## 4. Discussion

### 4.1. Phylogenetic trees based on rhodopsin and base compositional bias

Although the rhodopsin tree contains the highest number of well-supported clades, base compositional bias across taxa at third codon positions may be af-

fecting the accuracy of phylogenetic inference. When base composition varies significantly among taxa, all classical methods (MP, ML, and ME) tend to group sequences of similar nucleotide composition together, regardless of evolutionary history (Lockhart et al., 1994). The LogDet transformation, designed to correct this problem (Lockhart et al., 1994), has also been applied to the rhodopsin data (under ME). Although the LogDet tree shows more symmetric topology in the basal part than the GTR + G + I tree, the basal acanthomorphs are still separated by high GC3 percomorph taxa. Regarding the recurrent clades defined above, the LogDet tree recovers one more group (clade I) and shows monophyly of notothenioids (clade K). The notothenioid clade did not appear in the GTR + G + I tree because *Pseudaphritis* was excluded. The high-GC pleuronectiform taxon *Arnoglossus* failed to group with other pleuronectiform taxa in clade L in both LogDet and GTR + G + I trees. These observations indicate that: (1) using LogDet transformation distance did not alter identification of repeated clades; (2) the LogDet method might not correct the apparent bias introduced by base composition similarity. Given that the topology obtained by total evidence is somehow similar to the topology obtained by rhodopsin alone, the phylogenetic bias introduced by skewed base composition may be eclipsing the simultaneous analysis. We illustrate that the shape of the tree based on the whole set of genes can be determined by a single gene.

### 4.2. Implications for morphological hypotheses

*Basal acanthomorphs and paracanthopterygians.* As previous morphological studies (Nelson, 1989), this molecular study cannot elucidate phylogenetic interrelationships of the main basal acanthomorph lineages such as Lampridiformes, Polymixiiformes, Paracanthopterygii, Beryciformes, and Zeiformes. Nevertheless, application of the criterion of repeatability focuses attention on two clades: A (zeioids and gadids) and B (*Hoplostethus* and *Beryx*). The order Zeiformes has already been suspected to be paraphyletic (Johnson and Patterson, 1993; Rosen, 1984; Stiassny and Moore, 1992). Contradicting the sister-group relationship of Zeiformes and Beryciformes proposed by Johnson and Patterson (1993), Lauder and Liem (1983), placed Zeiformes (excluding Caproidei) in a very basal position among acanthomorphs, and as the sister-group of Beryciformes plus Percomorpha (as defined by them, see Fig. 6). Results of this molecular study also suggests Zeiformes paraphyly, however by placing zeioids as the sister-group to gadiforms, and also excluding caproids that remain in an ambiguous position. Support for a sister-group relationship between gadids and Zeioides comes from taxonomic congruence (28s and MT data sets) and from a unique pattern of sequence variation

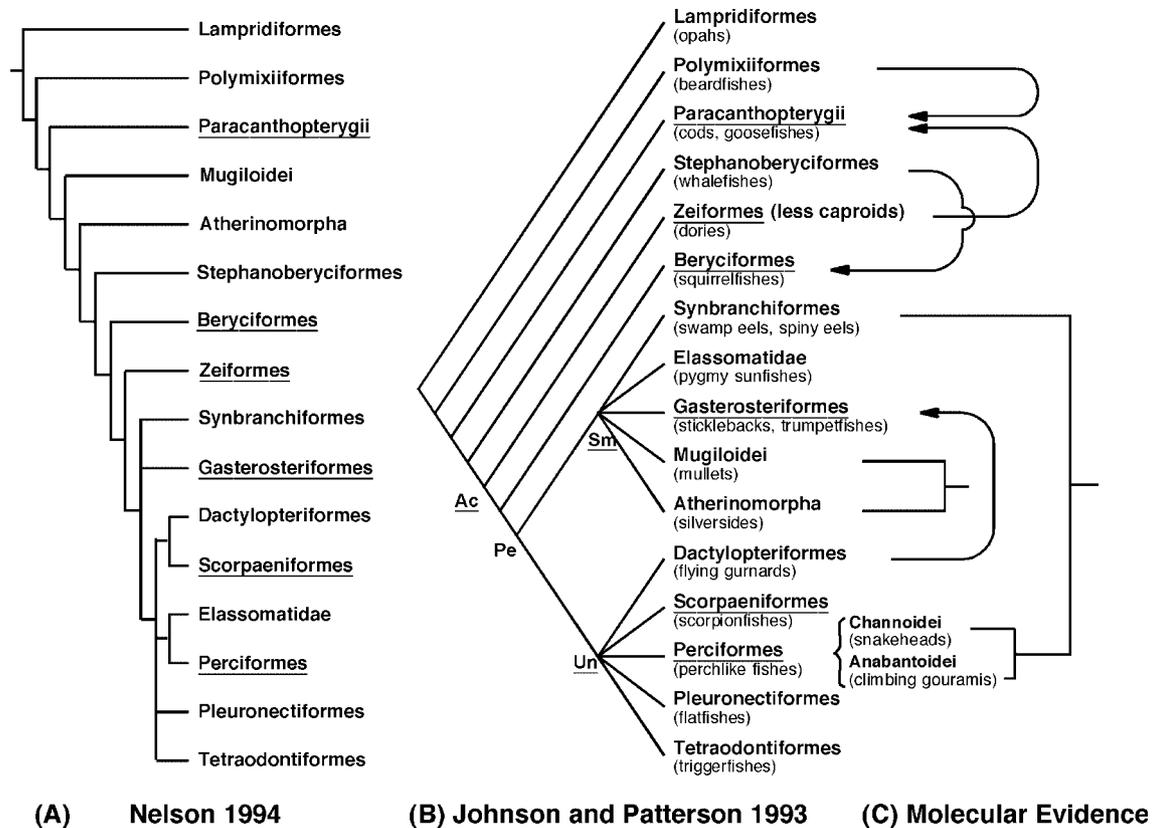


Fig. 6. (A) Phylogenetic summary depicting the relationship of the Acanthomorpha followed by Nelson (1994). (B) Cladogram depicting the relationship of the Acanthomorpha as presented by Johnson and Patterson (1993). (C) Summary of the phylogenetic revisions from molecular data proposed by this study and corroborated by Miya et al. (2001). The changes indicated concern the main acanthomorph lineages, most of the conclusions within perciforms are not shown. Ac, Acanthopterygii; Pe, Percormorpha; Sm, Smegmamorpha; Un, Unnamed clade. Taxonomic groups underlined means that the groups are shown to be paraphyletic or polyphyletic from our study, also corroborated in Miya et al. (2001). An arrow means the inclusion of the clade from which it starts within the clade to which it goes.

across the C1–C2, D3, and C12–D12 domains of the 28S rDNA. This evidence is strong because such a pattern of uniquely derived nucleotides is not likely the result of convergence. Non-overlapping 28S sequence data published previously also support this clade (Wiley et al., 2000, their Fig. 8c), however the grouping is not supported by morphological data (Wiley et al., 2000, their Fig. 8a: the three trees seem to have been mislabeled in the original publication, 8a should be for morphology, 8b for 12S, and 8c for 28S). More recently, a study based on whole mitogenomic data (Miya et al., 2001) confirmed that clade with high bootstrap support (94%). The corroboration from different studies, in which different protocols of analyses are employed, is another strong sign of reliability for this clade. The gadid + Zeioidei clade has never been directly proposed by any explicit phylogenetic study using morphological data. Gayet (1980b, 1980c) analyzed fossil data and included two zeoid families (Zeidae and Oreosomatidae) within the “Paracanthopterygii,” not based on putative synapomorphies but on global similarities with aipichthyids (a group of Cretaceous acanthomorph fossils). This hypothesis actually depends on how the aipichthyids are

placed in the phylogeny of Acanthomorpha. They were later included among the paracanthopterygians (Patterson and Rosen, 1989). The concept of Paracanthopterygii was proposed by Patterson and Rosen (1989), who listed four synapomorphies for the group. The taxonomic content of the group has changed according to different authors (Tables 1 and 2 in Patterson and Rosen, 1989), contributing to the uncertainty of its monophyly. In fact, the four synapomorphies proposed by Patterson and Rosen (1989) are questionable (see Gill, 1996 and Johnson and Patterson, 1993) because they appear also in the other basal acanthomorphs and even among perciform lineages. The present molecular study rejects Paracanthopterygian monophyly, especially by the sister-group of Zeioidei, the sister-group of Gobiesocoidei, and the evidence summarized by the MT tree (Fig. 3) showing that *Percopsis* (paracanthopterygian) and *Polymixia* (non-paracanthopterygian) form a clade with robust support (87%). This clade was not repeated because the taxa were not sequenced for the other two genes.

Clade B (*Beryx* + *Hoplostethus*) is another significant clade identified among basal Acanthomorpha (Fig. 5).

This clade includes two taxa classified in the beryciform suborders Berycoidei and Trachichthyoidei, respectively. This relationship has also been proposed, based on morphological characters, by Johnson and Patterson (1993). However, the monophyly of the beryciforms is still problematic because: (1) none of our results, except the ME and ML trees obtained from simultaneous analysis, showed the Holocentridae (here represented by *Myripristis*) closely related to other beryciforms; (2) no Stephanoberyciform fish was sampled here following Moore (1993) who proposed their inclusion within trachichthyoids. Moreover, beryciform monophyly was challenged by Moore (1993) and Stiassny and Moore (1992), who suggested that the Holocentridae may be more closely related to higher perciforms than to other beryciforms. More recent results from molecular studies (Colgan et al., 2000; Miya et al., 2001), with intensive taxonomic samplings for these groups, did not support the paraphyly (Moore, 1993; Stiassny and Moore, 1992) nor the monophyly (Johnson and Patterson, 1993; Nelson, 1994) of beryciforms suggested by morphological hypotheses. In fact, in Miya et al.'s (2001) study stephanoberyciforms and berycoids are closely related.

*Smegmamorpha*. The monophyly of the Smegmamorpha (Johnson and Patterson, 1993, their Fig. 6B) is not supported by the present study nor any previous molecular phylogeny (Miya et al., 2001; Wiley et al., 2000). Furthermore, the criterion of repeatability developed here provides evidence against this clade. Our results (Fig. 6) suggest that there are at least three other taxa, *Channa*, *Ctenopoma*, and *Dactylopterus* (belonging to the Channoidei, Anabantoidei, and Scorpaeniformes, respectively, following Nelson, 1994), that could be added to the Smegmamorpha because they are closely related to some smegmamorph components. The single synapomorphy, epineural on parapophysis, proposed for the Smegmamorpha (Johnson and Patterson, 1993) is therefore questionable. Among putatively smegmamorph taxa, 28S and mtDNA sequences show that the Mugilomorpha (represented by *Liza*) is the sister-group of some Atherinomorpha. Only in the rhodopsin trees and in results from simultaneous analysis, we found a clade containing *Liza* and all other atherinomorphs, as suggested by Miya et al. (2001) and Stiassny (1990, 1993). However, only beloniform taxa were used to represent Atherinomorpha by Miya et al. (2001) and relationships within Atherinomorpha are still uncertain. For example, the phylogenetic position of Cyprinodontiformes changes across data sets. The trees in Figs 2, 4, and 5 support clade C, grouping beloniforms and atheriniforms, contradicting the propositions of Dyer and Chernoff (1996) and Stiassny (1990). A more complete sampling is required to measure the strength of this conflict and to elucidate the interrelationships of the main atherinomorph lineages.

Our data corroborate the proposition of Gosline (1983), Johnson and Patterson (1993) and Miya et al. (2001), and Travers (1984a, 1984b) to add mastacembeloids to the order Synbranchiformes (clade *f*<sub>2</sub>). Before 1983, mastacembeloids were considered perciforms. Mitochondrial and rhodopsin data sets group *Monopterus* and *Mastacembelus* together (*f*<sub>2</sub>). Moreover, clade *f*<sub>2</sub> representing Synbranchiformes, seems to be the sister-group of clade *f*<sub>1</sub> that groups *Channa* and *Ctenomopa* (from Channoidei and Anabantoidei, respectively). Clade *f*<sub>1</sub> is supported by the 28S and rhodopsin data and clade F (*f*<sub>1</sub> + *f*<sub>2</sub>) also is supported by rhodopsin and MP analysis of MT data. Although clade F is not repeated across ME trees in this study it has been found in a previous study based on Mill genes (Fig. 30 in Chen, 2001) and also has been recognized since the time of Cuvier (1828–49). Taxa included in clade F represent unique labyrinthic freshwater fishes that use suprabranchial labyrinthic chambers for aerial respiration (Helfman et al., 1997, pp. 55–56). However, no other study has provided further evidence to align the components of the group (Rosen and Patterson, 1990), except Lauder and Liem (1983) and Roe (1991). Analyzing brain anatomy, these authors showed that the Channoidei were closely related to the Synbranchiiformes, but rejecting their sister-group relationship with Anabantoidei (e.g., Gosline, 1971) recovered by our results. Although the detailed anatomical background is not always the same for each lineage (Lauder and Liem, 1983), suggesting convergence from similar life habitats and selective pressures to survive in anoxic water, the trend seems worthy of attention for further anatomical studies.

The other taxon closely related to members of the “Smegmamorpha” is *Dactylopterus*, from the scorpaeniform suborder Dactylopteroidei, according to Nelson (1994, see Fig. 6A). Johnson and Patterson (1993) considered dactylopteroidei as an independent lineage within their unnamed polytomic clade grouping Scorpaeniformes, Perciformes, Tetraodontiformes, and Pleuronectiformes (Fig. 6B). In contrast to these propositions, but in agreement with Pietsch (1978) on the basis of jaw anatomy, our results suggest that *Dactylopterus* is closely allied to syngnathoids (clade E). The monophyly of Gasterosteiformes of Johnson and Patterson (1993) (syngnathoids plus gasterosteoids) is not recovered by this study because *Spinachia* (Gasterosteioidei) does not group with clade E.

*The unnamed clade of Johnson and Patterson (1993)*. This large assemblage contains most of the putatively advanced percomorph taxa. For most of these taxa, our sequence data failed to establish unambiguous phylogenetic relationships but at least showed that Scorpaeniformes and Perciformes are polyphyletic (see examples above and below). This lack of resolution might correspond to the rapid percomorph radiation, as

suggested by the fossil record (Benton, 1993). The great majority of percomorph families are known from the Lower Eocene, between 55 and 45 million years ago, a short period during which more than 60 families appeared. So many cladogenetic events in such a short time span might have left insufficient time for the accumulation of molecular synapomorphies, as suggested by the extremely short internal branch lengths in our trees. Such short internal branches might be also due to mutational saturation (i.e., in the mitochondrial data set). However, using non-saturated data (28S), the lack of resolution persists. Only some clades emerged from the radiation. These are clades D, G, H, I, K, L, and M (Table 4). Except for clade D grouping the Blenioidei and the Gobiesocoidei (Rosen and Patterson, 1990) and clade M, grouping two taxa of higher labroid families together (Gosline, 1971), all of these clades are new (i.e., without any previous morphological support). Alternative hypotheses for the sister-group of the Gobiesocoidei have been debated for a long time (Gill, 1996). Our results confirm the observation of Rosen and Patterson (1990, pp. 40–44) and suggests a synapomorphy for gobiesocids and “true” blennioids (as defined by Springer, 1993). Both of them lack pharyngobranchials (PB) 1, 2, and 4 but PB3 persists (Fig. 33A, B and Fig. 35 in Rosen and Patterson, 1990). This characteristic appears to be convergent in some cottids (Fig. 34 in Rosen and Patterson, 1990). In fact, cottids (classified in the Scorpaeniformes by Nelson, 1994) are found closely related to zoarcoids (Perciformes) in our study (clade I).

We sampled representatives from five distinct trachinoid families (Chiasmodontidae, Cheimarrichthyidae, Ammodytidae, Trachinidae, and Uranoscopidae) to check for the paraphyly of the suborder (Mooi and Johnson, 1997), and its possible relationships with the Antarctic notothenioids (Hastings, 1993; Pietsch, 1989). The monophyly and intrarelationships of the Trachinoidei were established by Pietsch (1989) and Pietsch and Zabetian (1990) using morphological characters. We show here that they are not monophyletic, as already suggested by Johnson (1993) and Mooi and Johnson (1997). First, our data show a close relationship between Cheimarrichthyidae and Ammodytidae (clade G) excluding the other families, in contrast to the previous morphological study of Pietsch and Zabetian (1990), which showed the Ammodytidae as the sister-group to the clade Trachinidae plus Uranoscopidae. But the last two families were not found as sister-groups in our study. Second, the Chiasmodontidae (represented by *Kali*) was nested three times within clade H with the Scombroidei and the Stromatoidei.

Regarding the Scombroidei, the most primitive family among scombroids, Sphyracidae (Johnson, 1986), does not group here with *Scomber* but within the clade L, which contains diverse percomorph taxa (percoids): carangids, echeneids, menids, polynemids, centropo-

mids, and pleuronectiforms. Gosline (1968, 1971) recognized a close relationship between mugiloids, atherinoids, sphyracidae, and polynemids. Their common character is the lack of attachment of the pelvic girdle to the cleithra. According to our trees, this character might be homoplastic: clade L included taxa without this character and excluded taxa having it (mugiloidei and atherinoids). Interestingly, clade L contains all or almost all flatfishes (Pleuronectiforms). Although our results fail to show monophyly of flatfishes, we cannot provide strong evidence against its monophyly. Nonetheless, from our clade L it seems that the origin of flatfishes is close to the origin of either centropomid or carangoid fishes. The former case seems to confirm the “percoid origin” of pleuronectiforms (Chapleau, 1993), although anatomical evidence used to reach that result is a combination of generalized percoid characters, and the Percoidei, which contains about 2860 species, is most likely polyphyletic (Johnson and Patterson, 1993; Nelson, 1994).

The sister-group of Antarctic perciform fishes, the Notothenioidei, has not yet been identified (Lecointre et al., 1997). The candidates proposed are zoarcoids (Anderson, 1984, 1990) or trachinoids (Hastings, 1993; Pietsch, 1989). Zoarcoids are recurrently found with cottoids (clade I). By sampling most of perciform suborders, we surprisingly found support for percids as the sister-group of notothenioids (clade K) from the three data partitions. Although the rhodopsin sequence data excluded the GC-aberrant notothenioid taxon *Pseudaphritis* from this clade, clades K and *k2* (monophyly of the Notothenioidei) are recovered by simultaneous analysis. Clade K is challenging, because percids live in the freshwaters of the Northern hemisphere and notothenioids are mostly marine Southern hemisphere fishes, most of them living in the Antarctic Ocean. If this hypothesis is correct, and considering the absence of known notothenioid fossil, there would be an important gap in the history of the clade K: some fossil relatives or extinct species would not have yet been sampled to connect the northern lineage to the southern one. However, according to Eastman (1993), if there were fossils related to notothenioids discovered one day, it would be difficult to recognize them as notothenioids because there is no unique osteological, or any other known character for that matter, that clearly distinguishes this suborder (Eakin, 1981). In fact, anatomical characters investigated in notothenioids so far can also be found among perciforms (Eakin, 1981; Iwami, 1985; Voskoboinikova, 1993). The definition of notothenioids given by Lecointre et al. (1997) excluded *Cottoperca* and *Bovichtus*. Interestingly, the present study recovers monophyletic notothenioids in their classical sense (i.e., including *Cottoperca* and *Bovichtus*). The molecular tree of Lecointre et al. (1997) failed to identify *Perca* as the sister-group of the notothenioids. This was due to the

fact that the 28S sequences used (domains D2 and D8) were far more variable than the present ones, leading to homoplasy obscuring the deepest outgroup interrelationships. In their study, homoplasy was indeed low within notothenioids, but much higher when other suborders were considered.

#### 4.3. Congruence or conflicts between acanthomorph studies

In general, our results are more in line with previous molecular phylogenies (Miya et al., 2001; Wiley et al., 2000) than with other morphological studies (Johnson and Patterson, 1993; Lauder and Liem, 1983; Stiassny and Moore, 1992) and traditional classifications (Nelson, 1994). Fig. 6 presents a comparison of the phylogenetic synthesis offered by Nelson (1994), the cladogram proposed by Johnson and Patterson (1993), a summary of results from our study, corroborated by the study of Miya et al. (2001). Although Wiley et al. (2000) concluded that their results, based on a “total evidence” approach, are *largely* congruent with the morphological hypothesis articulated by Johnson and Patterson (1993), this statement seemed somewhat confusing. We re-analyzed the molecular data (12S + 28S) presented by Wiley et al. (2000), but included characters which were excluded by these authors due to alignment or saturation problems. Interestingly, we found two new terminal clades with high bootstrap support. One of them is the clade grouping *Atherinomorus* and *Strongylura* (Atheriniformes and Beloniformes, respectively). The second is the clade grouping *Scopeloberyx* and *Beryx* (Stephanoberyciformes and Berycoidei, respectively). The former clade is congruent with our results (clade C) and the latter has been proposed by Miya et al. (2001). Surprisingly, the phylogenetic position of these taxa derived from the total evidence approach by Wiley et al. (2000) are identical to those proposed by morphological hypotheses (Dyer and Chernoff, 1996; Johnson and Patterson, 1993; Nelson, 1994). There are several possible explanations for this discrepancy. First, excluding characters (non-conserved loops or “saturated” regions) to avoid potential homoplasy might result in the exclusion of useful phylogenetic information, especially for the derived clades. Several studies have already reported that removing homoplasy also removes phylogenetic structure (Källersjö et al., 1999; Philippe et al., 1996; Sennblad and Bremer, 2000; Wenzel and Siddall, 1999). Second, the impact of sampling errors due to small number of characters (Nei et al., 1998; Takahashi and Nei, 2000) might be more important than the impact of putative homoplasy. Finally, the topology resulting from the total evidence analysis could have been dominated by the morphological data matrix, especially because some informative molecular characters were excluded. Since Wiley et al. (2000) used the morpho-

logical matrix of Johnson and Patterson (1993) in their total evidence analysis, it is not surprising that their result is congruent with morphology-based hypothesis.

According to our results, conflict between molecular and morphological hypotheses seems to be significant, but this is not a rare situation in the phylogenetic literature (e.g., Hillis and Wiens, 2000; Patterson et al., 1993). However, apparent lack of congruence may have several explanations. Incongruence may just be apparent when there is lack of phylogenetic resolution coupled to incomplete samplings, as may be the case here resulting from the fast radiation of acanthomorpha. As stated by Lauder and Liem (1983), the tremendous radiation of these fishes has resulted in extensive variation not only in morphology but also in behavior and ecology. This might explain some of the disagreement found in the phylogenetic interpretation of morphological characters, probably prone to huge plasticity and homoplasy. Better phylogenetic resolution could be accomplished by increasing the number of characters (Poe and Wiens, 2000). However, in spite of more characters being used in molecular studies, the lack of global resolution persists. Indeed, most of the topological disagreement between molecular and morphological studies resides in areas of the phylogeny with the weakest support (e.g., with non-repeated clades in this study). For instance, failure of this study to recover some of the traditional monophyletic groups such as Pleuronectiformes (flatfishes) and Tetraodontiformes (puffers and allies) should not be taken as *prima facie* evidence of conflict with these hypothesis, but most likely as a lack of phylogenetic signal in the data matrices. Thus the apparent incongruence could be only spurious (Hillis and Wiens, 2000).

Other potential source of conflict may be due to undersampling of taxa (Hillis and Wiens, 2000). Consider for example, the hypothesis of sister-group relationship between notothenioids and percids obtained by this study. It could merely mean that we failed to sample relevant intermediate samples, such as perch-like fishes from the Southern Hemisphere. Even with our best effort for taxonomic sampling of acanthomorpha, more exhaustive taxonomic sampling will require a better understanding of the phylogenetic relationships and of the putative conflicts between morphological and molecular hypotheses.

Nonetheless, some conflicts are likely to result from potential problems in morphological studies, such as lack of explicitness of the characters chosen, character coding, and application of non-phylogenetic methodology (Poe and Wiens, 2000). The latter may be one of the most serious problems for Acanthomorph studies. Although, ichthyology was second only to entomology in welcoming phylogenetic systematics in its earliest days of expansion (Lecointre, 1994; Rosen, 1982, 1985), there are still few high-order morphological studies based on

rigorous phylogenetic analyses. Even for the most “famous” groups, monophyly has never been really tested by rigorous phylogenetic analysis. In this light, it is not surprising to see how frequently the definition of acanthomorph subdivisions changed through time (Johnson, 1993).

## 5. Conclusion

In this study, separate analysis of multiple data sets has taken precedence over the total evidence approach for the assessment of phylogenetic reliability. Several main messages emerge: (1) This approach is especially useful when phylogenetic signal in the data is relatively low due to putative radiation and when one of the data partitions may be influenced by strong misleading signal. (2) Blindly trusting the results from simultaneous analysis, even associated with high bootstrap supports, is risky. (3) The present criterion of reliability allowed to reliably hypothesize new clades among acanthomorph fishes (by comparison with previous studies) and the demonstration of paraphyly or polyphyly for some previously recognized acanthomorph “lineages.” However, it is somewhat discouraging to see how little resolution was obtained at the deeper nodes of the acanthomorph radiation, even when high numbers of representative taxa are used. (4) Our results challenge currently accepted points of views based on morpho-anatomic characters (e.g., Johnson and Patterson, 1993; Nelson, 1994; see Fig. 6). Interestingly, some of the “new” clades found here were directly or indirectly suggested by morphological studies published in the 1970s and 1980s, and even much earlier (e.g., Cuvier, 1828–49) (see Table 4). But many of these earlier morphological hypotheses were based on overall similarity, implying that close inspection of the morphology may still be required. The present paper will help in defining new paths to the future of systematic Ichthyology, and probably in resolving acanthomorph relationships.

## Acknowledgments

During this 10-year-long project, numerous people have provided fish samples. We thank Nicolas Bailly, Philippe Bouchet, François Catzeflis, Romain Causse, Pascal Deynat, Catherine Chombard, Guido Dingerkus, Marie-Henriette Dubuit, Guy Duhamel, Yves Fermon, Jin-Chywan Gwo, Michel Hignette, Jean-Claude Hurreau, Sébastien Lavoué, Yves Le Gal, Chen-Hsiang Liu, François Meunier, Pierre Noël, Catherine Ozouf-Co-staz, Eva Pisano, Stuart Poss, Jean-Claude Quérou, François Renaud, Peter Ritchie, Thibaud Roman, Melanie Stiassny, Denis Terver, Annie Tillier, Marino Vacchi, and Dick Williams. We thank Tsui-Yu Chang

and Hansen Chen for help in development of computer program (repeated-bootstrap components), and Cécile Fischer for rhodopsin sequence of *Tetraodon nigroviridis*. Special thanks are given to Dr. Guillermo Ortí for helpful comments and for revising the manuscript. This work was supported by the grant Action spécifique du Muséum: acides nucléiques et Evolution, No. UC1358 (1990), by the Direction de la Recherche et des Etudes Doctorales du Ministère de l'Éducation Nationale: “Evolution: approches interdisciplinaires et développements méthodologiques” (1990–1991); by the Réseau National de Biosystématique: “Exploitation phylogénétique et développements méthodologiques de la congruence de classes de caractères de différentes natures” (1996–1997), by funds from the GDR CNRS 1005 (1997–1998), by funds from IFR CNRS 1541 (1998–1999), and by a doctoral grant from the “ministère des affaires étrangères” in France for scientific cooperation on marine Biotechnologies between FRANCE and TAIWAN (1997–2000). The senior author received appreciated help from the C.R.O.U.S. of Paris.

## References

- Alves-Gomes, J.A., 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, M.E., 1984. On the anatomy and phylogeny of the Zoarcidae (Teleostei: Perciformes). Ph.D. Dissertation, College of William and Mary, Williamsburg, VA.
- Anderson, M.E., 1990. The origin and evolution of the Antarctic ichthyofauna. In: Gon, O., Heemstra, P.C. (Eds.), *Fishes of the Southern Ocean*. JBL Smith Institute of Ichthyology, Grahamstown, South Africa, pp. 28–33.
- Archer, S.N., Hirano, J., unpublished. Comparative analysis of opsins in Mediterranean coastal fish.
- Archer, S.H., Hope, A.J., Partridge, J.C., 1995. The molecular basis for the blue-green sensitivity in the rod visual pigments of the European eel. *Proc. R. Soc. London B* 262, 289–295.
- Archer, S.N., Lythgoe, J.N., Hall, L., 1992. Rhodopsin cDNA sequence from the sand goby (*Pomatoschistus minutus*) compared with those of other vertebrates. *Proc. R. Soc. London B* 248, 19–25.
- Bargelloni, L., Ritchie, P.A., Patarnello, T., Battaglia, B., Lambert, D.M., Meyer, A., 1994. Molecular evolution at subzero temperatures: mitochondrial and nuclear phylogenies of fishes from Antarctica (suborder Notothenioidei), and the evolution of anti-freeze glycopetide. *Mol. Biol. Evol.* 11, 854–886.
- Barrett, M., Donoghue, M.J., Sober, E., 1991. Against consensus. *Syst. Zool.* 40, 486–493.
- Benton, M.J., 1993. *The fossil record II*. Chapman and Hall, London.
- Bremer, K., 1994. Branch support and tree stability. *Cladistics* 10, 295–304.
- Buckley, T.R., Simon, C., Flook, P.K., Misof, B., 2000. Secondary structure and conserved motifs of the frequently sequenced domains IV and V of the insect mitochondrial large subunit rRNA gene. *Insect Mol. Biol.* 9, 565–580.
- Bull, J.J., Huelsenbeck, J.P., Cunningham, C.W., Swofford, D.L., Waddell, P.J., 1993. Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* 42, 384–397.

- Cao, Y., Adachi, J., Jank, A., Paabo, S., Hasegawa, M., 1994. Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: instability of a tree based on a single gene. *J. Mol. Evol.* 39, 519–527.
- Carnap, R., 1950. *Logical Foundations of Probability*. University of Chicago Press, Chicago.
- Chang, B.S.W., Campbell, D.L., 2000. Bias in phylogenetic reconstruction of vertebrate rhodopsin sequences. *Mol. Biol. Evol.* 17, 1220–1231.
- Chang, B.S.W., Crandall, K.A., Carulli, J.P., Hartl, D.L., 1995. Opsin phylogeny and evolution: a model for blue shifts in wavelength regulation. *Mol. Phylogenet. Evol.* 4, 31–43.
- Chapleau, F., 1993. Pleuronectiform relationships: a cladistic reassessment. *Bull. Mar. Sci.* 52, 516–540.
- Chen, W.-J., 2001. La répétitivité des clades comme critère de fiabilité: application à la phylogénie de Acanthomorpha (Teleostei) et des Notothenioidei (acanthomorphes antarctiques). Ph.D. Thesis, University of Paris VI.
- Colgan, D.J., Zhang, C.-G., Paxton, J.R., 2000. Phylogenetic investigation of the Stephanoberyciformes and Beryciformes, particularly whalefishes (Euteleostei: Cetomimidae), based on partial 12S rDNA and 16S rDNA sequences. *Mol. Phylogenet. Evol.* 17, 15–25.
- Cumming, M.P., Otto, S.P., Wakeley, J., 1995. Sampling properties of DNA sequence data in phylogenetic analysis. *Mol. Biol. Evol.* 12, 814–822.
- Cunningham, C.W., Zhu, H., Hillis, D.M., 1998. Best-fit maximum-likelihood models for phylogenetic inference: empirical tests with known phylogenies. *Evolution* 52, 978–987.
- Cuvier, G., Valenciennes, 1828–49. *Histoire naturelle des poissons*, Paris.
- De Queiroz, A., Donoghue, M.J., Kim, J., 1995. Separate versus combined analysis of phylogenetic evidence. *Ann. Rev. Ecol. Syst.* 26, 657–681.
- Dyer, B.S., Chernoff, B., 1996. Phylogenetic relationships among atheriniform fishes (Teleostei: Atherinomorpha). *Zool. J. Linnean Soc. London B* 117, 1–69.
- Eakin, R.R., 1981. Osteology and relationships of the fishes of the antarctic family Harpagiferidae (Pisces, Notothenioidei). In: Kornicker, L.S. (Ed.), *Biology of the Antarctic Seas IX*, Washington, pp. 81–147.
- Eastman, J.T., 1993. *Antarctic Fish Biology*. Academic Press, San Diego, CA.
- Eernisse, D.J., Kluge, A., 1993. Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. *Mol. Biol. Evol.* 10, 1170–1195.
- Farris, J.S., 1983. The logical basis of phylogenetic analysis. In: Platnick, N.I., Funk, V.A. (Eds.), *Advances in Cladistics*, vol. II. Columbia Press, New York, pp. 7–36.
- Felsenstein, J., 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27, 401–410.
- Felsenstein, J., 1981. 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.
- Fitzgibbon, J., Hope, A.J., Slobodyanyuk, S.J., Bellingam, S.J., Bowmaker, J.K., Hunt, D.M., 1995. The rhodopsin-encoding gene of the bony fish lack introns. *Gene* 164, 273–277.
- Gaudant, M., 1978. Contribution à l'étude anatomique et systématique de l'Ichtyo-faune cénomaniennne du Portugal. Première partie: les Acanthopterygii. *Com. Serv. Geol. Portugal* 63, 105–149.
- Gautier, C., 2000. Compositional bias in DNA. *Curr. Opin. Genet. Dev.* 10, 656–661.
- Gayet, M., 1980a. Recherches sur de l'Ichtyo-faune cénomaniennne des Monts de Judée: Les acanthoptérygiens. *Ann. Paléotol. Vertébrés* 66, 75–128.
- Gayet, M., 1980b. Sur la découverte dans le Crétacé de Hadjula (Liban) du plus ancien Caproidae connu. *C. R. Hebdo. Séances Acad. Sci., Paris* 290D, 447–448.
- Gayet, M., 1980c. Découverte dans le Crétacé de Hadjula (Liban) du plus ancien Caproidae connu. étude anatomique et phylogénétique. *Bull. Mus. Natl. His. Naturelle, Paris, Sér. 4* 2C, 259–269.
- Gill, A.C., 1996. Comments on an intercalary path for the glossopharyngeal (Cranial IX) nerve as a synapomorphy of the Paracanthopterygii and on the phylogenetic position of the Gobiesocidae (Teleostei: Acanthomorpha). *Copeia* 1996, 1022–1029.
- Goldman, N., 1993. Statistical tests of models of DNA substitution. *J. Mol. Evol.* 36, 182–198.
- Gosline, W.A., 1968. The suborders of Perciform fishes. *Proc. U. S. Natl. Mus.* 124, 1–77.
- Gosline, W.A., 1971. *Functional Morphology and Classification of Teleostean Fishes*. University Press Hawaii, Honolulu.
- Gosline, W.A., 1983. The relationships of the mastacembelid and synbranchid fishes. *Jpn. J. Ichthyol.* 29, 323–328.
- Grande, L., 1994. Repeating patterns in nature, predictability, and “impact” in science. In: Grande, L., Rieppel, O. (Eds.), *Interpreting the Hierarchy of Nature*. Academic Press, New York, pp. 61–84.
- Greenwood, P.H., Rosen, D.E., Weitzman, S.H., Mayers, G.S., 1966. Phyletic studies of teleostean fishes, with a provisional classification of living forms. *Bull. Amer. Mus. Nat. Hist.* 131, 339–455.
- Gu, X., Fu, Y.-X., Li, W.-H., 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Mol. Biol. Evol.* 12, 546–557.
- Hasegawa, M., Hashimoto, T., 1993. Ribosomal RNA trees misleading? *Nature* 361, 23.
- Hassanin, A., Lecointre, G., Tillier, S., 1998. The “evolutionary signal” of homoplasy in protein-coding gene sequences and its consequences for a priori weighting in phylogeny. *C. R. Acad. Sci., Sér. III* 321, 611–620.
- Hastings, P.A., 1993. Relationships of the fishes of the perciform suborder Notothenioidei. In: Miller, R.G. (Ed.), *A History and Atlas of the Fishes of the Antarctic Ocean*. Foresta Institute for Ocean and Mountain Studies, Carson City, Nevada, pp. 99–107.
- Helfman, G.S., Collette, B.B., Facey, D.E., 1997. *The Diversity of Fishes*. Blackwell Science, Massachusetts.
- Hennig, W., 1966. *Phylogenetic Systematics*. University of Illinois Press, Urbana, IL.
- Hickson, R.E., Simon, C., Copper, A., Spicer, G.S., Sullivan, J., Penny, D., 1996. Conserved sequence motifs, alignment, and secondary structure for the third domain of animal 12s rRNA. *Mol. Biol. Evol.* 13, 150–169.
- Hickson, R.E., Simon, C., Perrey, S.W., 2000. The performance of several multiple-sequence alignment programs in relation to secondary-structure features for an rRNA sequence. *Mol. Biol. Evol.* 17, 530–539.
- Hillis, D.M., 1995. Approaches for assessing phylogenetic accuracy. *Syst. Biol.* 44, 3–16.
- Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* 42, 182–192.
- Hillis, D.M., Wiens, J.J., 2000. Molecular versus morphology in systematics: conflicts, artifacts, and misconceptions. In: Wiens, J.J. (Ed.), *Phylogenetic Analysis of Morphological Data*. Smithsonian Institution Press, Washington and London, pp. 1–19.
- Huelsenbeck, J.P., 1997. Is the Felsenstein zone a fly trap? *Syst. Biol.* 46, 69–74.
- Huelsenbeck, J.P., Bull, J.J., Cunningham, C.W., 1996. Combining data in phylogenetic analysis. *Trends Ecol. Evol.* 11, 152–157.
- Huelsenbeck, J.P., Crandall, K., 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. *Annu. Rev. Ecol. Syst.* 28, 437–466.

- Huelsenbeck, J.P., Hillis, D.M., 1993. Success of phylogenetic methods in the four-taxon case. *Syst. Biol.* 42, 247–264.
- Hunt, D.M., Fitzgibbon, J., Slobodyanyuk, S.J., Bowmaker, J.K., Dulai, K.S., 1997. Molecular evolution of the cottoid fish endemic to Lake Baikal deduced from nuclear DNA evidence. *Mol. Phylogenet. Evol.* 8, 415–422.
- Iwami, T., 1985. Osteology and relationships of the Family Channichthyidae. *Mem. Natl. Inst. Polar Res. Ser.*, 1–69.
- Johnson, G.D., 1986. Scombroid phylogeny: an alternative hypothesis. *Bull. Mar. Sci.* 39, 1–41.
- Johnson, G.D., 1993. Percomorph phylogeny: progress and problems. *Bull. Mar. Sci.* 52, 3–28.
- Johnson, G.D., Patterson, C., 1993. Percomorph phylogeny: a survey of acanthomorphs and a new proposal. *Bull. Mar. Sci.* 52, 554–626.
- Källersjö, M., Albert, V.A., Farris, J.S., 1999. Homoplasy increases phylogenetic structure. *Cladistics* 15, 91–93.
- Kelsey, C.R., Crandall, K.A., Voevodin, A.F., 1999. Different models, different trees: the geographic origin of PTLV-I. *Mol. Phylogenet. Evol.* 13, 336–347.
- Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order of Hominoidea. *J. Mol. Evol.* 29, 170–179.
- Kjer, K.M., 1995. Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: a example of alignment and data presentation from the frogs. *Mol. Phylogenet. Evol.* 4, 314–330.
- Kluge, A.G., 1989. A concern for evidence and a phylogenetic hypothesis of relationships among Epicrates (Boidae, Serpentes). *Syst. Zool.* 38, 7–25.
- Kluge, A.G., Wolf, A.J., 1993. Cladistics: what's in a word? *Cladistics* 9, 183–199.
- Lanave, C., Preparata, G., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. *J. Mol. Evol.* 20, 86–93.
- Lanyon, S.M., 1993. Phylogenetic frameworks: towards a firmer foundation for the comparative approach. *Biol. J. Linn. Soc.* 49, 45–61.
- Lauder, G.V., Liem, K.F., 1983. The evolution and interrelationships of the actinopterygian fishes. *Bull. Mus. Comp. Zool. Cambridge (Mass)* 150, 95–197.
- Lavoué, S., Bigorne, R., Lecointre, G., Agnèse, J.F., 2000. Phylogenetic relationships of mormyrid electric fishes (Mormyridae; Teleostei) inferred from cytochrome b sequences. *Mol. Phylogenet. Evol.* 14, 1–10.
- Lê, H.L.V., Lecointre, G., Perasso, R., 1993. A 28S rRNA based phylogeny of the Gnathostomes: first steps in the analysis of conflict and congruence with morphologically based cladograms. *Mol. Phylogenet. Evol.* 2, 31–51.
- Lecointre, G., 1994. Aspects historiques et heuristiques de l'Ichtyologie systématique. *Cybiurn* 18, 339–430.
- Lecointre, G., Bonillo, C., Ozouf-Costaz, C., Hureau, J.-C., 1997. Molecular evidence for the origins of Antarctic fishes: paraphyly of the Bovichtidae and no indication for the monophyly of the Notothenioidei (Teleostei). *Polar Biol.* 18, 193–208.
- Lecointre, G., Deleporte, P., 2000. Le Principe du "total evidence" requiert l'exclusion de données trompeuses. In: V. Barriel, T. Bourgoin (Eds.), *Biosystema 18: Caractères*. Published by the Société Française de Systématique, Paris, pp. 129–151.
- Leipe, D.D., Gunderson, J.H., Nerad, T.A., Sogin, M.L., 1993. Small subunit ribosomal RNA of *Hexamita inflata* and the quest for the first branch in the eukaryotic tree. *Mol. Biochem. Parasitol.* 59, 41–48.
- Leitner, T., Kumar, S., Albert, J., 1997. Tempo and mode of nucleotide substitutions in *gag* and *env* gene fragments in human immunodeficiency virus type 1 populations with known transmission history. *J. Virol.* 71, 4761–4770.
- Levasseur, C., Lapointe, F.J., 2001. War and peace in phylogenetics: a rejoinder on total evidence and consensus. *Syst. Biol.* 50 (6), 881–891.
- Lim, J., Chang, J.-L., Tsai, H.-J., 1997. A second type of rod opsin cDNA from the common carp (*Cyprinus caprio*). *Biochem. Biophys. Acta* 1352, 8–12.
- Lockhart, P.J., Steel, M.A., Hendy, M.D., Penny, D., 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* 11, 605–612.
- Lydeard, C., Roe, K.J., 1997. The phylogenetic utility of the mitochondrial cytochrome b gene for inferring relationships among actinopterygian fishes. In: Kocher, T.D., Stepien, C.A. (Eds.), *Molecular Systematics of Fishes*. Academic Press, San Diego, CA, pp. 285–303.
- Maddison, W.P., Maddison, D.R., 1992. *Mac Clade: Analysis of Phylogeny and Character Evolution*. Version 3.01. Sinauer, Sunderland, MA.
- Maidak, B.L., Cole, J.R., Parker, C.T., et al. (14 co-authors), 1999. A new version of the RDP (Ribosomal Database Project). *Nucleic Acids Res.* 27, 171–173.
- Mickevich, M.F., 1978. Taxonomic congruence. *Syst. Zool.* 27, 143–158.
- Miya, M., Kawaguchi, A., Nishida, M., 2001. Mitogenomic exploration of higher teleostean phylogenies: a case study for moderate-scale evolutionary genomics with 38 newly determined complete mitochondrial DNA sequences. *Mol. Biol. Evol.* 18, 1993–2009.
- Miya, M., Nishida, M., 1998. Molecular phylogeny and evolution of deep-sea fish genes *Sternoptyx*. *Mol. Phylogenet. Evol.* 10, 11–22.
- Miya, M., Nishida, M., 1996. Molecular phylogenetic perspective on the evolution of the deep-sea fish genus *Cyclothone* (Stomiiformes: Gonostomatidae). *J. Ichthyol. Res.* 43, 375–398.
- Miyamoto, M.M., Fitch, W.M., 1995. Testing species phylogenies and phylogenetic methods with congruence. *Syst. Biol.* 44, 64–76.
- Mooi, R.D., Johnson, G.D., 1997. Dismantling the Trachinoidei: evidence of a scorpaenoid relationship for the Champsodontidae. *J. Ichthyol. Res.* 44, 143–176.
- Moore, J.A., 1993. Phylogeny of the Trachichthyiformes (Teleostei: Percomorpha). *Bull. Mar. Sci.* 52, 114–136.
- Moritz, G., Hillis, D.M., 1996. Molecular systematics: context and controversies. In: Hillis, M., Moritz, C., Mable, B.K. (Eds.), *Molecular Systematics*. Sinauer, Sunderland, MA, pp. 1–13.
- Morrison, D.A., Ellis, J.T., 1997. Effects of nucleotide sequence alignment on phylogeny estimation: a case study of 18S rDNAs of apicomplexa. *Mol. Biol. Evol.* 14, 428–441.
- Mullis, K.B., Faloona, F.A., 1987. Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. *Methods Enzymol.* 155, 335–350.
- Naylor, G.J.P., Adams, D.C., 2001. Are the fossil data really at odds with the molecular data? Morphological evidence for Cetartioactyla phylogeny reexamined. *Syst. Biol.* 50, 444–453.
- Naylor, G.J.P., Brown, W.M., 1998. Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences. *Syst. Biol.* 47, 61–76.
- Naylor, G.J.P., Collins, T.M., Brown, W.M., 1995. Hydrophobicity and phylogeny. *Nature* 373, 565–566.
- Nei, M., Kumar, S., Takahashi, K., 1998. The optimization principle in phylogenetic analysis tends to give incorrect topologies when the number of nucleotides or amino acids used is small. *Proc. Natl. Acad. Sci. USA* 95, 12390–12397.
- Nelson, G.J., 1979. Cladistic analysis and synthesis: principles and definitions, with a historical note on Adanson's *Familles des Plantes* 1763–1764. *Syst. Zool.* 28, 1–21.
- Nelson, G.J., 1989. Phylogeny of major fish groups. In: Fernholm, B., Bremer, K., Jornavall, H. (Eds.), *The hierarchy of life: molecules and morphology in Phylogenetic analysis*. Elsevier Science Publishers B.V., Amsterdam, The Netherlands, pp. 325–336.
- Nelson, J.S., 1994. *Fishes of the World*, third ed. Wiley, New York.

- Nixon, K.C., Carpenter, J.M., 1996. On simultaneous analysis. *Cladistics* 12, 221–241.
- Ortí, G., 1997. Radiation of Characiform fishes: evidence from mitochondrial and nuclear DNA sequences. In: Kocher, T.D., Stepien, C.A. (Eds.), *Molecular Systematics of Fishes*. Academic Press, San Diego, CA, pp. 219–243.
- Ortí, G., Meyer, A., 1996. Molecular evolution of Ependymin and the phylogenetic resolution of early divergences among euteleost fishes. *Mol. Biol. Evol.* 13, 556–573.
- Otero, O., Gayet, M., 1996. Anatomy and phylogeny of the Aipichthyoidea nov. of the Cenomanian Tethys and their place in the Acanthomorpha (Teleostei). *N. Jb. Geol. Paläont. Abh.* 202, 313–344.
- Page, R.D.M., Holmes, E.C., 1998. *Molecular Evolution: A Phylogenetic Approach*. Blackwell Science, Abingdon, UK.
- Page, R.D.M., 2000. Extracting species trees from complex gene trees: reconciled trees and vertebrate phylogeny. *Mol. Phylogenet. Evol.* 14, 89–102.
- Parker, A., Kornifield, A., 1996. An improved amplification and sequencing strategy for phylogenetic studies using the mitochondrial large subunit rRNA gene. *Genome* 39, 793–797.
- Patterson, C., 1964. A review of Mesozoic acanthopterygian fishes, with special reference to those of the English Chalk. *Philos. Trans. R. Soc. Lond.* 247, 213–482.
- Patterson, C., 1993. An overview of the early fossil record of acanthomorphs. *Bull. Mar. Sci.* 52, 29–59.
- Patterson, C., Rosen, D.E., 1989. The Paracanthopterygii revisited: order and disorder. In: *Papers on the systematics of gadiform fishes*. Cohen, D.M. (Ed.), *Natl. Hist. Mus. Ser.* 32, Los Angeles City, pp. 5–36.
- Patterson, C., Williams, D.M., Humphries, C.J., 1993. Congruence between molecular and morphological phylogenies. *Annu. Rev. Ecol. Syst.* 24, 153–188.
- Penny, D., Hendy, M.D., 1986. Estimating the reliability of evolutionary trees. *Mol. Biol. Evol.* 3, 403–417.
- Philippe, H., 1993. MUST: a computer package of Management Utilities for Sequences and Trees. *Nucleic Acids Res.* 21, 5264–5272.
- Philippe, H., Adoutte, A., 1998. The molecular phylogeny of Eukaryota: solid facts and uncertainties. In: Coombs, G.H., Vickerman, K., Sleigh, M.A., Warren, A. (Eds.), *Evolutionary Relationships Among Protozoa*. Chapman & Hall, London, pp. 25–56.
- Philippe, H., Forterre, P., 1999. The rooting of the universal tree of life is not reliable. *J. Mol. Evol.* 49, 509–523.
- Philippe, H., Douzery, E., 1994. The pitfalls of molecular phylogeny based on four species, as illustrated by the Cetacea/Artiodactyla relationships. *J. Mam. Evol.* 2, 133–152.
- Philippe, H., Laurent, J., 1998. How good deep phylogenetic trees? *Curr. Opin. Gen. Develop.* 8, 616–623.
- Philippe, H., Lecointre, G., Lê, H.L.V., Le Guyader, H., 1996. A critical study of homoplasy in molecular data with the use of a morphologically based cladogram, and its consequences for character weighting. *Mol. Biol. Evol.* 13, 1174–1186.
- Philippe, H., Lopez, P., Brinkmann, H., Budin, K., Germot, A., Laurent, J., Moreira, D., Müller, M., Le Guyader, H., 2000. Early branching of fast evolving eukaryotes? An answer based on slowly evolving positions. *Proc. R. Soc. Lond. B* 267, 1213–1221.
- Philippe, H., Sorhannus, U., Baroin, A., Perasso, R., Gasse, F., Adoutte, A., 1994. Comparison of molecular and paleontological data in diatoms suggests a major gap in the fossil record. *J. Evol. Biol.* 7, 247–265.
- Pietsch, T.W., 1978. Evolutionary relationships of the sea moths (Teleostei: Pegasidae) with a classification of gasterosteiform families. *Copeia* 1978, 517–529.
- Pietsch, T.W., 1989. Phylogenetic relationships of trachinoid fishes of the family Uranoscopidae. *Copeia* 1989, 253–303.
- Pietsch, T.W., Zabetian, C.P., 1990. Osteology and interrelationships of sand lances (Teleostei: Ammodytidae). *Copeia* 1990, 78–100.
- Poe, S., Wiens, J.J., 2000. Character selection and the methodology of morphological phylogenetics. In: Wiens, J.J. (Ed.), *Phylogenetic Analysis of Morphological Data*. Smithsonian Institution press, Washington and London, pp. 20–36.
- Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Posada, D., Crandall, K.A., 2001. Simple (Wrong) models for complex trees: a case from Retroviridae. *Mol. Biol. Evol.* 18, 271–275.
- Rambaut, A., 1996. Sequence alignment editor version 1.0  $\alpha$ 1, The package and information is available from the WWW site: <http://evolve.zoo.ox.ac.uk/Se-AL/Se-AL.html>.
- Rannala, B., Huelsenbeck, J.P., Yang, Z., Nielsen, R., 1998. Taxon sampling and the accuracy of large phylogenies. *Syst. Biol.* 47, 702–710.
- Ritchie, P.A., Lavoué, S., Lecointre, G., 1997. Molecular phylogenetics and the evolution of Antarctic notothenioid fishes. *Comp. Biochem. Physiol. A* 118, 1009–1025.
- Rodrigo, A.G., Kelly-Borges, M., Bergquist, P.R., Bergquist, P.L., 1993. A randomisation test of the null hypothesis that two cladograms are sample estimates of a parametric phylogenetic tree. *N. Z. J. Bot.* 31, 257–268.
- Rodríguez, F., Oliver, J.L., Marín, A., Medina, J.R., 1990. The general stochastic model of nucleotide substitution. *J. Theor. Biol.* 142, 485–501.
- Roe, L.J., 1991. Phylogenetic and ecological significance of Channidae (Osteichthyes, Teleostei) from the early Eocene kuldana formation of Kohat, Pakistan. *Contrib. Mus. Paleont. Univ. Mich.* 28, 93–100.
- Rosen, D.E., 1973. Interrelationships of higher teleostean fishes. In: Greenwood, R., Miles, S., Patterson, C. (Eds.), *Interrelationships of Fishes*. J. Linn. Soc. (Zool.). Academic Press, New York, pp. 397–513.
- Rosen, D.E., 1982. Teleostean interrelationships, morphological function, and evolutionary inference. *Am. Zool.* 22, 261–273.
- Rosen, D.E., 1984. Zeiformes as primitive plectognath fishes. *Am. Mus. Novit.* 2782, 1–45.
- Rosen, D.E., 1985. An essay on euteleostean classification. *Am. Mus. Novit.* 2827, 1–57.
- Rosen, D.E., Patterson, C., 1990. On Müller's and Cuvier's concepts of pharyngognath and labyrinth fishes and the classification of percomorph fishes, with an atlas of percomorph dorsal gill arches. *Am. Mus. Novit.* 2983, 1–57.
- Russo, C.A.M., Takahashi, K., Nei, M., 1996. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. *Mol. Biol. Evol.* 13, 525–536.
- Rzhetsky, A., Nei, M., 1992. A simple method for estimating and testing minimum-evolution trees. *Mol. Biol. Evol.* 9, 945–967.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S., Higuchi, R., Horn, R., Mullis, K.B., Erlich, H.A., 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA-polymerase. *Science* 239, 487–491.
- Saitoh, K., Hayashizaki, K., Yokoyama, Y., Asahida, T., Toyohara, H., Yamashita, Y., 2000. Complete nucleotide sequence of Japanese flounder (*Paralichthys olivaceus*) mitochondrial genome: structural properties and cue for resolving teleostean relationships. *J. Hered.* 91, 271–278.
- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Sanderson, M.J., 1989. Confidence limits on the phylogenies: the bootstrap revisited. *Cladistics* 5, 113–129.
- Sennblad, B., Bremer, B., 2000. Is there a justification for differential a priori weighting in coding sequences? A case study from rbcL and Apocynaceae s.l. *Syst. Biol.* 49, 101–113.
- Slowinski, J.B., Page, R.D.M., 1999. How should species phylogenies be inferred from sequence data? *Syst. Biol.* 48, 814–825.

- Springer, V.G., 1993. Definition of the suborder Blennioidei and its included families (Pisces: Perciformes). *Bull. Mar. Sci.* 52, 472–495.
- Stiassny, M.L.J., 1986. The limits and relationships of the acanthomorph teleosts. *J. Zool. Lond. B* 1, 411–460.
- Stiassny, M.L.J., 1990. Notes on the anatomy and relationships of the bedotioid fishes of Madagascar, with a taxonomic revision of the genus *Rheocles* (Atherinomorpha: Bedotiidae). *Amer. Mus. Novit.* 2979, 1–33.
- Stiassny, M.L.J., 1993. What are grey mullets? *Bull. Mar. Sci.* 52, 197–219.
- Stiassny, M.L.J., Moore, J.A., 1992. A review of the pelvic girdle of acanthomorph fishes, with comments on hypotheses of acanthomorph intrarelationships. *Zool. J. Linn. Soc.* 104, 209–242.
- Strimmer, K., von Haeseler, A., 1996. Quartet puzzling: A quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* 13, 964–969.
- Sullivan, J., Swofford, D.L., 1997. Are guinea pigs rodents? The importance of adequate models in molecular phylogenies. *J. Mamm. Evol.* 4, 77–86.
- Sullivan, J., Swofford, D.L., 2001. Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution pattern are violated? *Syst. Biol.* 50, 723–729.
- Swofford, D.L., 1991. When are phylogeny estimates from molecular and morphological data incongruent? In: Miyamoto, M.M., Cracraft, J. (Eds.), *Phylogenetic Analysis of DNA Sequences*. Oxford University Press, New York, pp. 295–333.
- Swofford, D.L., 2001. PAUP\*. *Phylogenetic Analysis Using Parsimony (\*and Other Methods)*. Version 4. Sinauer Associates, Sunderland, MA.
- Swofford, D.L., Olsen, G.J., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic inference. In: Hillis, M., Moritz, C., Mable, B.K. (Eds.), *Molecular Systematics*, second ed. Sinauer, Sunderland, MA, pp. 407–514.
- Takahashi, K., Nei, M., 2000. Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution, and maximum likelihood when a large number of sequences are used. *Mol. Biol. Evol.* 17, 1251–1258.
- Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512–526.
- Tang, K.L., Berendzen, P.B., Wiley, E.O., Morrissey, J.F., Winterbottom, R., Johnson, G.D., 1999. The phylogenetic relationships of the suborder Acanthuroidei (Teleostei: Perciformes) based on molecular and morphological evidence. *Mol. Phylogenet. Evol.* 11, 415–425.
- Tateno, Y., Takezaki, N., Nei, M., 1994. Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum-parsimony methods when substitution rate varies with site. *Mol. Biol. Evol.* 11, 261–277.
- Tavaré, S., 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. *Lec. Math. Life Sci.* 17, 57–86.
- Thompson, J.D., Gibson, T.J., Pevniak, F., Jeanmougin, F., Higgins, D.G., 1997. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882.
- Titus, T.A., Frost, D.A., 1996. Molecular homology assessment and phylogeny in the lizard family Opluridae (Squamata: Iguania). *Mol. Phylogenet. Evol.* 6, 49–62.
- Travers, R.A., 1981. The interarcual cartilage; a review of its development, distribution, and value as an indicator of phylogenetic relationships in euteleostean fishes. *J. Nat. Hist.* 15, 853–871.
- Travers, R.A., 1984a. A review of Mastacembeloidei, a suborder of synbranchiform teleost fishes. Part I: anatomical descriptions. *Bull. Brit. Mus. Nat. Hist. (Zool.)* 46, 1–133.
- Travers, R.A., 1984b. A review of Mastacembeloidei, a suborder of synbranchiform teleost fishes. Part II: Phylogenetic analysis. *Bull. Brit. Mus. Nat. Hist. (Zool.)* 47, 83–150.
- Venkatesh, B., Ning, Y., Brenner, S., 1999. Late changes in spliceosomal introns define clades in vertebrate evolution. *Proc. Natl. Acad. Sci. USA* 96, 10267–10271.
- Voskoboinikova, O.S., 1993. Evolution of the visceral skeleton and phylogeny of Nototheniidae. *J. Ichthyol.* 22, 105–111.
- Waddell, P.J., Steel, M.A., 1997. General time-reversible distances with unequal rates across sites: mixing gamma and inverse Gaussian distributions with invariant sites. *Mol. Phylogenet. Evol.* 8, 398–414.
- Waters, J.M., López, J.A., Wallis, G.P., 2000. Molecular phylogenetics and biogeography of Galaxioid fishes (Osteichthyes:Galaxiidae): dispersal, vicariance, and the position of Lepidogalaxias salamandroides. *Syst. Biol.* 49, 777–795.
- Wenzel, J.W., Siddall, M.E., 1999. Noise. *Cladistics* 15, 51–64.
- Wiley, E.O., Johnson, G.D., Dimmick, W.W., 1998. The phylogenetic relationships of lampridiform fishes (Teleostei: Acanthomorpha), based on a total-evidence analysis of morphological and molecular data. *Mol. Phylogenet. Evol.* 10, 417–425.
- Wiley, E.O., Johnson, G.D., Dimmick, W.W., 2000. The interrelationships of Acanthomorph fishes: a total evidence approach using molecular and morphological data. *Biochem. Syst. Ecol.* 28, 319–350.
- Winnepminck, B., Backeljau, T., De Wachter, R., 1993. Extraction of high molecular weight DNA from molluscs. *T.I.G.* 9, 407.
- Yamaguchi, M., Miya, M., Okiyama, M., Nishida, M., 2000. Molecular phylogeny and larval morphological diversity of the lanternfish genus *Hygophum* (Teleostei: Myctophidae). *Mol. Phylogenet. Evol.* 15, 103–114.
- Yang, Z., 1997. How often do wrong models produce better phylogenies? *Mol. Biol. Evol.* 14, 105–108.
- Yokoyama, S., 1997. Molecular genetic basis of adaptive selection: examples from color vision in the vertebrates. *Annu. Rev. Genet.* 31, 315–336.
- Yokoyama, S., Zhang, H., Radlwimmer, B.F., Blow, N.S., 1999. Adaptive evolution of color vision of the comoran coelacanth (*Latimeria chalumnae*). *Proc. Natl. Acad. Sci. USA* 96, 6279–6284.
- Zardoya, R., Meyer, A., 1996. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. *Mol. Biol. Evol.* 13, 933–942.