CHAPTER 1

Phylogeny and Classification

Guillermo Ortí and Chenhong Li

1.1 INTRODUCTION

Our current knowledge about phylogeny and classification of "fishes" is in a state of flux. Most classification schemes which proposed to organize the vast fish biodiversity (Helfman *et al.* 1997; Nelson 2006) have been based on loosely formulated syntheses of many, largely disconnected phylogenetic studies among some of its components. An explicit cladistic analysis including representatives of all major taxonomic groups across the diversity of fishes has never been accomplished. As a consequence, phylogenetic relationships among the major groups of fishes are still controversial and unresolved, as are many of the proposed higher-level taxa (Greenwood *et al.* 1973; Lauder and Liem 1983; Jamieson 1991; Stiassny *et al.* 1996; Kocher and Stepien 1997; Chen *et al.* 2003; Meyer and Zardoya 2003; Miya *et al.* 2003; Chen *et al.* 2004; Cloutier and Arratia 2004; Stiassny *et al.* 2004). We expect this situation to change soon, as ongoing efforts by morphologists and molecular systematists are seeking to converge on a synthesis. While DNA sequence data are being collected rapidly and cost effectively, and provide a useful way to reconstruct phylogeny, the promise of a data-rich supermatrix approach to explicitly analyze phylogenetic relationships among representatives of *all* major groups of "fishes" still is unaccomplished. These are, therefore, exciting times for molecular systematics in general, and fish phylogenetics in particular. Molecular data sets are proliferating and rapidly transitioning towards phylogenomic proportions (Miya and Nishida 2000; Rokas *et al.* 2003; 2005; McMahon and Sanderson 2006; Comas *et al.* 2007) and a more thorough interpretation of morphological and paleontological material is also underway (Diogo 2007; Mabee *et al.* 2007). For example, a recent analysis of higher-level relationships among the major early-branching lineages of sarcopterygians and actinopterygians (Diogo 2007), based on osteological and

School of Biological Sciences, University of Nebraska, Lincoln, Nebraska 68588, USA.

Change running head throughout books

myological characters, could pave the way for an expanded effort that may include the most diverse euteleostean taxa. We anticipate that in just a few years, efforts along these two fronts will converge to produce a well-supported phylogenetic classification based on genealogical analyses of large numbers of genes and a better understanding of morphological homologies based on detailed analysis of genetic and developmental pathways. In this chapter, we summarize some of the most recent results, with major emphasis on hypotheses for actinopterygian fishes derived from our own molecular studies.

1.2 PHYLOGENETIC ANALYSIS OF MOLECULAR DATA

Much of what we know of the relationships of fishes has been the result of a long history of morphological research (e.g., (Rosen 1973; 1982; Stiassny 1986; Johnson 1992; Johnson and Patterson 1993). But at this time, however, there is no single resource that presents a comprehensive picture or synthesizes our current understanding of higher-level actinopterygian morphology, particularly within the species-rich percomorph crown group. Efforts underway (Diogo 2007; Ed Wiley, pers. com.), as stated above, are setting the stage to solving this shortcoming, especially with the insights that a combination of morphological and molecular data can make available. In this review, we compare hypotheses based largely on morphology with new proposals from molecular systematics. Although the relative merits of the different kinds of data commonly used for phylogenetic analysis remain in dispute—see Scotland *et al*. (2003) and subsequent reaction (Jenner 2004; Wiens 2004; Smith and Turner 2005) for a recent reincarnation of this debate—there is little doubt that molecular data are and will be most commonly used for phylogenetics. Part of the reason is the ease of collection and of establishing primary homology across vast taxonomic ranges (Li *et al.*, 2007). But molecular data, as any other kind of data, are not without problems.

delete bracket

Molecular phylogenies based on DNA sequences of a single locus or a few loci often suffer from low resolution and marginal statistical support due to limited character sampling. Individual gene genealogies also may differ from each other and from the organismal phylogeny under study. This discordance, know as the "gene-tree vs. species-tree" issue (Fitch 1970; Pamilo and Nei 1988) can be caused by several factors. In many cases, systematic biases leading to statistical inconsistency in phylogenetic reconstruction (i.e., base-compositional bias, long-branch attraction, heterotachy) may cause spurious results (Felsenstein 1978; Weisburg *et al.* 1989; Foster and Hickey 1999; Lopez *et al.* 2002). In other cases, discordance may be due to the actual history of gene duplication/extinction events leading to mistaken assumptions about orthology (Fitch 1970). Even though the correct gene tree may be obtained in the analysis, genealogical discordance between the history of the gene and the organsimal phylogeny may persist. Undetected paralogy (the relationship of homology among loci originating from gene duplication events) may result

from sampling genomes of distantly related species using a direct-PCR approach. Ideally, to avoid this problem, only single-copy genes that did not undergo a complex history of duplication and extinction should be used for phylogenetic analysis. This condition may be hard to find among fishes in light of mounting evidence supporting a fish-specific whole-genome duplication event (Amores *et al.* 1998; Meyer and Van de Peer 2005) and the more general observation that gene duplications are a common mechanism of molecular evolution (Ohno 1970; Taylor and Raes 2004).

A phylogenomic approach—using genome-scale data sets to study evolutionary relationship—may provide the best solution to these problems (Eisen and Fraser 2003; Delsuc *et al.* 2005) but it requires compilation of large data sets that include many independent nuclear loci for many species (Bapteste *et al.* 2002; Rokas *et al.* 2003; Driskell *et al.* 2004; Philippe *et al.* 2004). Such data sets are less likely to succumb to sampling and systematic errors (Rokas *et al.* 2003) by offering the possibility to survey characters that are phylogenetically reliable and also to test phylogenetic results with alternative taxonomic samples. Some simple criteria can be used to assess the reliability of molecular markers (e.g. testing homogeneity of base composition, relative rates of evolution, saturation of base substitutions, etc). Taxonomic-rich data sets also allow the possibility of using different subsets of representative species for each group to test for consistency in the results. In spite of rapid success and initial optimism generated by phylogenomic approaches (Gee 2003; Rokas *et al.* 2003), large and complex data sets also exacerbate the unresolved methodological challenges (Li *et al.* 2008). Many long-standing challenges such as sparse taxon-sampling (Soltis *et al.* 2004), base compositional bias (Phillips *et al.* 2004), missing data (Wiens 2003; Waddell 2005) or incomplete lineage sorting (Kubatko and Degnan 2007) also increase in relevance as multi-locus data sets grow in size and complexity. We elaborate below on two major potential obstacles for recovering a comprehensive phylogenetic hypothesis for ray-finned fishes: base compositional bias and undetected paralogy due to gene duplications.

1.2.1 Base Compositional Bias

Stationarity (i.e. that evolutionary processes do not change significantly across or within lineages) and time-reversibility (i.e. that the rate of change from one nucleotide to another is the same in each direction) often are assumed in standard inference models used for phylogenetic analysis, in part to simplify computations and also due to the expectation that base frequencies in DNA sequences remain constant along the evolutionary path. However, highly variable base composition among orthologous DNA sequences sampled from different species is not uncommon (Jukes and Bhushan 1986; Bernardi 1993). This is especially true for the nuclear gene RAG-1 among fishes (Ortí *et al.* 2005), as shown in Figure 1.1. Some fish taxa show extremely high content of G and C at the third codon positions of this gene (e.g., elopomorphs, galaxiids, stomiiforms) while other taxa show extremely low

frequencies (e.g., Ostariophysi). Significant variation in base composition also is evident within some groups (e.g., among acanthomorph fishes, within Clupeiformes and Elopiformes). Base compositional bias is a well-known source for systematic error in phylogenetic inference, usually resulting in groups with similar nucleotide frequencies that do not represent true evolutionary relationships. An example for fishes and several possible solutions that have been proposed (Steel *et al.* 1993; Lockhart *et al.* 1994; Gu and Li 1996; Foster and Hickey 1999; Foster 2004; Collins *et al.* 2005) are discussed below.

One simple way to address the potentially confounding effect of base compositional bias is to carefully choose genes that do not show this pattern (Collins *et al.* 2005). With phylogenomic-sized data sets containing large numbers of genes (>100), this may be a feasible approach, but usually most studies are confronted with limited data and some base compositional bias. The next simple solution in this case would be to recode the data as purines and pyrimidines (RY-coding, where R=G=A and Y=C=T). This approach homogenizes base composition among divergent sequences and removes the GC-bias (Woese *et al.* 1991; Phillips *et al.* 2004). However, the method also leads to loss of phylogenetic information. We applied this approach to assess the effect of base-composition in our study of clupeiform relationships based on DNA sequences of RAG1 and RAG2 genes (Li and Ortí 2007). Most clupeiform fishes have high (> 70%) GC content at the variable positions in these genes, except for *Denticeps clupeoides* (61%) and *Spratelloides delicatulus* (59%), that are closer to the average frequency observed among fishes (65%, Fig.1.1). In contrast, ostariophysans have a relatively low average GC content (55%). This pattern is repeated, albeit to a lesser degree for mitochondrial ribosomal genes (12S and 16S). Analyses of these sequences invariably grouped *Denticeps* with ostariophysans rather than with other clupeiforms (Li and Ortí 2007). Support for the *Denticeps*+ostriophysi clade should decrease significantly when using RY-coded data if this relationship is artificially obtained due to non-stationarity. Indeed, when branch weights (total number of characters supporting each alternative hypothesis) were calculated for RY-coded data, higher support was obtained for the *Denticeps*+Clupeiformes hypothesis, in contrast to the result obtained with non-coded data. Support for both hypothesis, however, was lower under the RY-coding strategy, consistent with the expected loss of phylogenetic information caused by this method (Li and Ortí 2007).

A more effective approach to avoid artifacts caused by base compositional biases involves accounting for the non-stationarity explicitly in the evolutionary model used for analysis. Several alternative models have been proposed, including the LogDet distance method (Lockhart *et al.* 1994), maximum-likelihood methods assigning local base frequencies to each branch (Yang and Roberts 1995; Galtier and Gouy 1998), and Bayesian methods assigning different base frequencies to predefined number of clades (Foster 2004). In many cases, the relatively simple LogDet distance approach has been

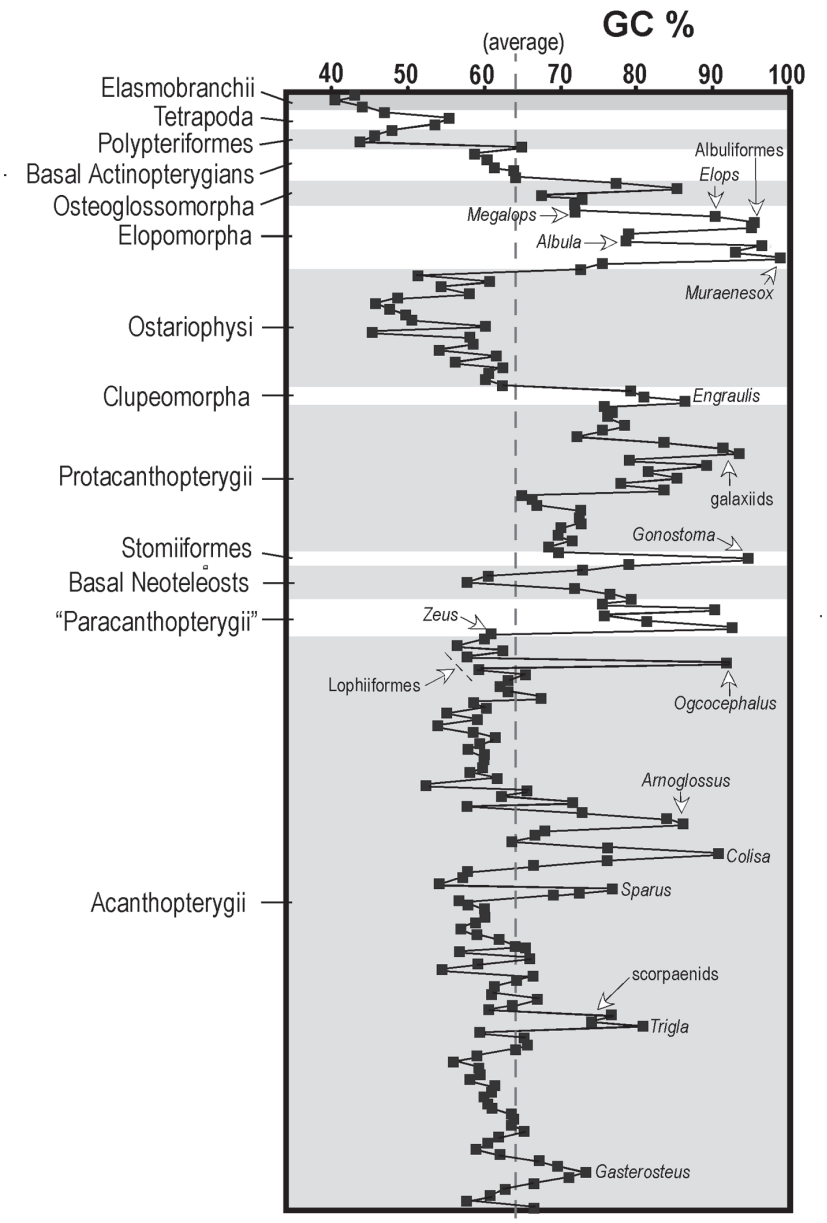

Fig. 1.1 Base composition (% content of G and C) at the third codon positions of the nuclear gene RAG-1 among fishes. Only sequences of exon 3 (*ca.* 1500 bp) of the RAG-1 gene were compared; most of the variation among these sequences is found at third codon positions. Taxonomic groups are indicated on the ~~x-axis~~ vertical axis following the sequential order presented in most current classifications; some representative genera or families are identified in the graph. For a complete list of taxa, please contact the authors. Original.

shown to fail and did not recover the expected tree topology (Foster and Hickey 1999). More realistic, parameter-rich, models that attempt to assign branch-specific base frequencies (Yang and Roberts 1995; Galtier and Gouy 1998) or clade-specific base frequencies (Foster 2004) may be too complex and over-parameterized to produce reliable results. Two recent developments to account for non-stationarity that were also aimed at reducing the high dimensionality of these models may provide promising options to address this problem. Blanquart and Latrillot (2006) proposed a new model that estimates variation among base frequencies across lineages by a stochastic process using a Poisson distribution. Their method is more realistic because it decouples the change of base frequencies from speciation events and also reduces the number of parameters to estimate. A second approach by Gowri-Shankar and Rattray (2007) extended Foster's (2004) methods by introducing a reversible-jump Monte Carlo Markov Chain (MCMC) sampler for efficient Bayesian inference of the model order along with other phylogenetic parameters of interest. The methods of Blanquart and Latrillot (2006) and Goweri-Shankar and Rattray (2007), implemented in the computer programs PhyloBayes and PHASE, respectively, should provide more robust phylogenetic results for large-scale analysis of nuclear genes when base compositional bias may be rampant.

1.2.2 Gene Duplication and Paralogy

Another important issue associated with the use of nuclear protein-coding genes for phylogeny inference is uncertainty about their orthology (Fitch 1970). This uncertainty may lead to the inference of erroneous relationships among species even when the true genealogical histories of specific loci are recovered in the analysis. As stated above, sophisticated phylogenetic methods exist and continue to be developed to identify and circumvent potential analytical artifacts, but confounding biological factors arising from the dynamic nature of the genome remain. Among these, the complex history of gene or genome duplication/extinction events that has been documented for ray-finned fishes (Van de Peer *et al.* 2003) is especially challenging for fish phylogenetics.

Most genes are represented in genomes by more than one copy, usually as members of a gene family. But some genes are unique ("single-copy"), meaning that no other region of the genome contains a sequence with high similarity to them. This definition of "single copy" is somewhat arbitrary and operational (it depends on definition of a threshold of similarity), since fragments of the genome with lower values of similarity to any gene may presumably be found. This suggests that no gene could be truly single copy unless duplicates have been lost from the genome or modified so drastically that they are no longer recognizable as such. Nonetheless, and to simplify interpretation of phylogenetic results, it is better to use single-copy nuclear genes to minimize the chance of sampling paralogous genes among taxa (Li *et al*. 2007). Even in the case that gene duplication events may have occurred

during evolution of the taxa of interest (Van de Peer *et al.* 2003; Meyer and Van de Peer 2005), duplicated copies of a single-copy nuclear gene tend to be lost quickly, possibly due to dosage compensation, a mechanism that balances the phenotypic expression of genes with unequal copy number (Lynch and Conery 2000; Blomme *et al.* 2006). Almost 80% of the duplicated genes can be secondarily lost shortly after a genome-duplication event (Jaillon *et al.* 2004; Woods *et al.* 2005). Therefore, if duplicated copies are lost before the relevant speciation events occur (Fig. 1.2a, b), there will be no discrepancy among the inferred gene trees and the species tree. In contrast, if the unfortunate situation depicted in Fig. 1.2c occurs, paralogous comparisons will result in topological discordances among genes and among some of these and the species tree.

delete full stops

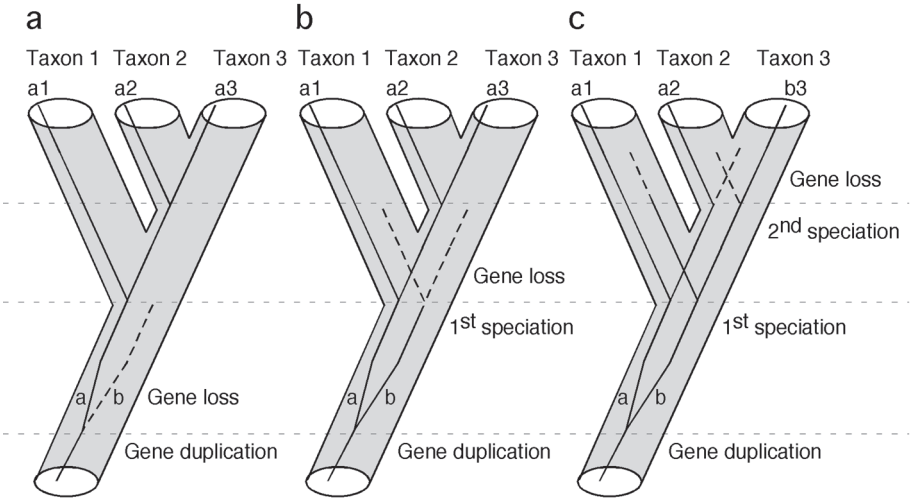

Fig. 1.2 Gene duplication and subsequent loss may not cause incongruence between gene tree and species tree if gene loss occurs before the first speciation event (a), or before the second speciation event (b). The only case that would cause incongruence is when the gene survived both speciation events and is asymmetrically lost in taxon 2 and taxon 3 (c). From Li, C., Ortí, G., Zhang, G., and Lu, G. 2007. A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. BMC Evolutionary Biology 7: 44, Fig. 1.

Phylogenomic approaches allow the comparison of such potential discordances for a large number of genes. Therefore, a gene-by-gene analysis of the topological distribution of the observed discordance may be used to reconstruct (reconcile) putative duplication/extinction events and avoid the pitfall of mistaken paralogy (Page and Cotton 2002). For example, gene duplication events that occurred before the inferred origin of the ingroup of interest that were followed by differential losses of the duplicates among ingroup taxa, may lead to the inclusion of paralogous sets of genes ("out-paralogs" sensu Remm *et al.*, (2001). In this case, an a posteriori examination

italics

of unrooted tree topologies and associated branch lengths may help detect the putative out-paralogs because they will form two highly divergent clades. If this pattern is detected, it would be safer to infer phylogenies only for the reduced taxonomic sets represented in each of the orthologous datasets. In contrast, if duplication/differential loss events occurred within the ingroup of interest ("in-paralogs") these will not be as easily detected by inspection of topology and branch lengths because these duplicates will be equidistant to other ingroup lineages that are not descendants of the ancestor in which the duplication took place. Genes that have this history must meet two conditions to remain undetected and have an effect on phylogenetic conclusions: (1) none of the taxa affected by the duplication maintain both copies of the gene or any existing duplicates remain undetected by PCR assays; and (2) the same taxonomic distribution of duplication and loss is repeated across multiple genes. Although possible, it seems unlikely that both of these conditions will be met.

In summary, although molecular characters are not free of many potential problems that usually confuse phylogenetic results, careful analysis of large numbers of single-copy nuclear genes (the phylogenomic approach) may provide a realistic means towards inferring the tree of life of "fishes" in the near future.

1.3 THE TREE OF LIFE OF "FISHES"

In this section, we outline some currently accepted hypotheses of relationships among the major groups of "fishes" relevant to subsequent chapters of this book. Although used freely in the literature, the term "fishes" does not refer to a natural group (a monophyletic lineage). The term is used to describe a heterogeneous collection of distantly related vertebrates such as hagfish, dogfish, knifefish, killifish, cowfish, and lungfish. The term could be restricted to a monophyletic group if it were applied only to the largest and most diverse clade of fishes (Actinopterygii). Because the tetrapods are always excluded, "fishes" form a paraphyletic group and classification schemes do not give this term taxonomic rank (Nelson 2006). Figure 1.3 is a summary of the most likely hypothesis upheld by phylogenetic analyses of morphology and molecular data. In the following subsections, we discuss evidence supporting or contradicting this hypothesis and provide more detailed phylogenetic relationships among some relevant groups of ray-finned fishes. For a discussion on sarcopterygian relationships, see Jamieson, **Chapter 16**.

1.3.1 Jawless Fishes (Agnathans)

The living jawless fishes (hagfishes and lampreys, Jamieson, ~~**Part A,**~~ **Chapter 6**) represent early-branching lineages at the base of the vertebrate tree of life. Their relationship to other long-extinct jawless fishes, to each other, and to the jawed vertebrates remains controversial. Most morphological and paleontological analyses (Hardisty 1982; Mallat 1984; Janvier 1996; Mallat 1997; Donoghue *et*

delete

al. 2000; Ota and Kuratani 2007) but see (Mallat 1997; Ota and Kuratani 2007) support the view that agnathans form a paraphyletic group, with lampreys more closely related to the gnathostomes than to hagfishes. Molecular evidence, in contrast, keeps mounting to overwhelmingly support a sister-group relationship between hagfishes and lampreys (Cyclostomata) as shown in Figure 1.3.

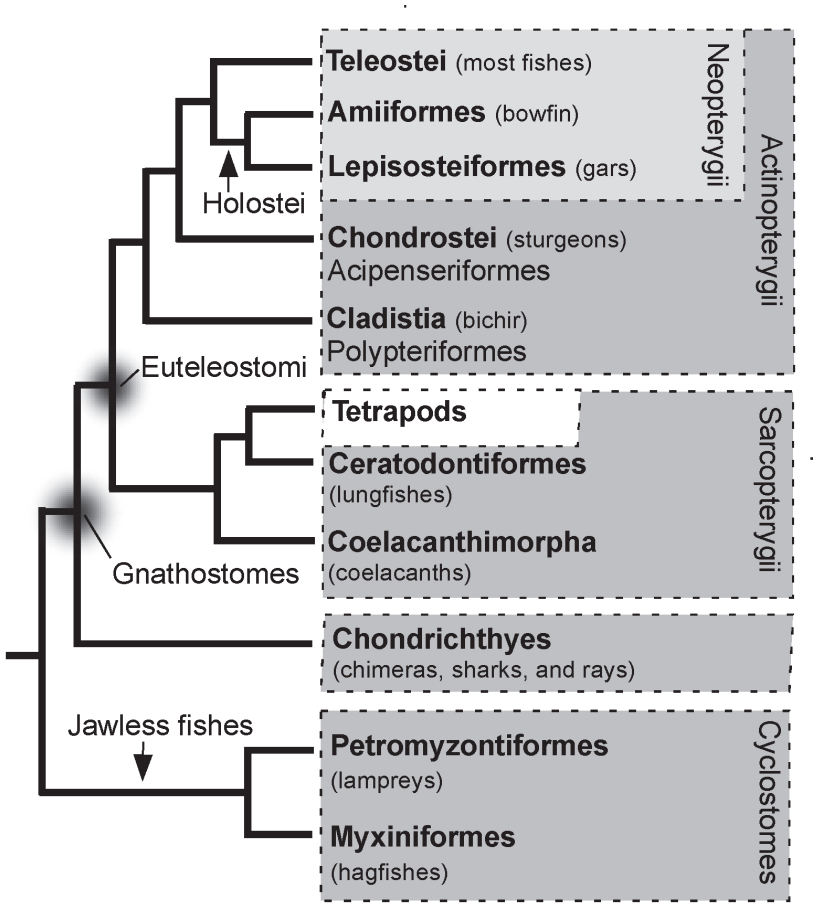

Fig. 1.3 Phylogenetic relationships among major groups of living "fishes" (in gray boxes). Common names are given in parentheses. ~~Evidence for and against this hypothesis is discussed in the text.~~ Original.

boxes), based on several morphological and molecular anlayses discussed in the text where evidence for and against this hypothesis is discussed.

delete

The first molecular study to address this issue was based on comparison of nucleotide sequences of 18S ribosomal RNA (Stock and Whitt 1992). Subsequently, Mallat and Sullivan (1998) added sequences of the 28S rRNA gene and also recovered Cyclostomata. These early studies, however, included only a single representative of hagfish (*Epatretus*) and lamprey (*Petromyzon*),

risking potential analytical artifacts due to long-branch attraction. To address this issue, Mallat and Winchell (2007) added distantly related hagfish and lamprey taxa (*Myxine* and *Geotria*) and their analysis upheld the previous result with even higher support, rejecting the alternative "lampreys plus gnathostomes" hypothesis with confidence. Furlong and Holland (2002) reanalyzed 18S sequences using a Bayesian approach and also included in their analysis protein-coding mitochondrial DNA (mtDNA) and two protein-coding nuclear genes, triose phosphate isomerase (TPI) and superoxide dismutase (SOD). Their study provided strong support for Cyclostomata, again corroborating previous results with increasing confidence. To overcome potential biases in mtDNA sequences and alignment issues with rRNA genes, Takezaki *et al*. (2003) assembled an impressive data set with 35 loci to test cyclostome monophyly. The genes analyzed by Takezaki and collaborators represented a diverse group of nuclear protein-coding genes, including housekeeping and regulatory genes, with about half of them encoding ribosomal proteins which are known to evolve slowly. This study provided definitive evidence that molecular genetic data support the cyclostome hypothesis (Fig.1.3). Additional molecular studies in favor of this view include papers by Kuraru *et al*. (1999), Cotton and Page (2002), Delarbre *et al*. (2002), Blair and Hedges (2005), and Delsuc *et al.* (2006).

Agnathans and gnathostomes exhibit striking differences in their immune system. Neither hagfishes or lampreys possess the essential components that gnathostomes use for adaptive immunity, namely immunoglobulins (Ig), T cell receptors, recombination activating genes RAG1 and -2, and MHC class I and II molecules, but they share a fundamentally similar immune mechanism of generating variable lymphocyte receptors, VLRs (Pancer *et al*. 2004; Pancer *et al.* 2005). Although the VLR-based immune system could represent the plesiomorphic condition predating evolution of the vertebrate Ig-response, these two systems could have evolved simultaneously as early vertebrates experienced intense selective pressures to develop an anticipatory molecular recognition response. Whether the VLR system can be considered a synapomorphy supporting the monophyly of cyclostomes depends on the (largely unknown) condition observed in deuterostome outgroups (tunicates, hemichordates, echinoderms). A recent review of the immune system of the sea urchin based on comparative genomics shows that echinoderms exhibit immune signalling mediators and much of the gene regulatory toolkit for immunity known previously only for vertebrates, including a homologous Rag1/2 functional gene cluster (Rast *et al.* 2006). This finding suggests that the VLR system of hagfishes and lampreys could indeed be interpreted as a cyclostome synapomorhy.

1.3.2 Actinopterygii (Ray-finned Fishes)

This group contains nearly 27,000 described species, currently classified into three subclasses, 44 orders and 453 families (Nelson 2006). We review here some outstanding controversies regarding relationships among higher-level

taxa that have become classic debates and discuss the incidence of new data. We begin by characterizing the early-branching lineages at the base of the tree of ray-finned fishes and progress towards the derived euteleostean crown groups, where the highest diversity among living actinopterygians can be found.

Despite previous hypotheses linking polypterids to sarcopterygians (see Jamieson and Mattei, **Chapter 7**), it is now quite well established that the extant sister group to all other ray-finned fishes is the lineage leading to the bichir (*Polypterus*) and its living relatives, some 11 species of African freshwater fishes (family Polypteridae). This view (Figs. 1.3 and 1.4) has been recently supported by molecular evidence (Venkatesh *et al.* 2001; Inoue *et al.* 2003; Kikugawa *et al.* 2004; Li *et al.* 2008) as well as morphological analysis (Diogo 2007). The classic concept of "Chondrostei" that grouped *Polypterus* and living sturgeons and paddlefishes (Acipenseriformes) and their fossil relatives (Schaeffer 1973), received some support in a recent analysis of 10 nuclear genes (Ortí and Li 2007; Li *et al.* 2008), albeit with low bootstrap (65%) and posterior probability (0.74) values (Fig. 1.4). Most evidence from both morphological (Grande and Bemis 1996; Gardiner *et al.* 2005; Grande 2007) and molecular data (see above) suggests that "Chondrostei" is actually a paraphyletic group. Therefore, the current consensus is that polypterids are the sister taxon to all other living actinoterygians, and Acipenseriformes (or Chondrostei *sensu stricto;* Jamieson, **Chapter 8**) are considered as the sister group to neopterygians (Nelson 2006; Jamieson and Mattei, **Chapter 7**; Figs. 1.3 and 1.4).

Relationships among the non-teleost actinopterygians have been somewhat controversial, but a consensus seems to be emerging (at least for the extant taxa). While most morphological (Regan 1923; Patterson 1973) and molecular evidence (Lê *et al.* 1993; Kikugawa *et al.* 2004; Hurley *et al.* 2007; Li *et al.* 2008) supports the monophyly of Neopterygii, a group represented by gars (Lepisosteiformes), *Amia*, and teleosts (Fig. 1.3), relationships among these three lineages are hotly debated; see Arratia (2001) for a review of alternative schemes of relationships. Historically, *Lepisosteus* and *Amia* were grouped in a single clade (Holostei), placed as the sister-group to Teleostei (Nelson 1969; Jessen 1972). Subsequently, Holostei was dissolved in favor of alternative hypotheses suggesting that either Amiiformes (Patterson 1973; Olsen 1984; Grande and Bemis 1996; Diogo 2007) or Lepisosteiformes (Olsen 1984) alone represent the sister-group of teleosts. Yet another hypothesis, derived from analysis of mitogenomic data or indel patterns in the nuclear gene RAG2, supports a monophyletic "ancient fish" group composed by Acipenseriformes, Lepisosteidae and *Amia* (Venkatesh *et al.* 2001; Inoue *et al.* 2003). This group was placed as the sister-group to Teleostei. Most recently, however, both molecular data (Kikugawa *et al.* 2004; Hurley *et al.* 2007; Ortí and Li 2007; Li *et al.* 2008) and a reassessment of morphology (Grande 2007) advocate the "resurrection" of Holostei as the sister group of Teleostei. This is our currently preferred hypothesis, presented in Figures 1.3 and 1.4 (see also Jamieson, **Chapter 9**).

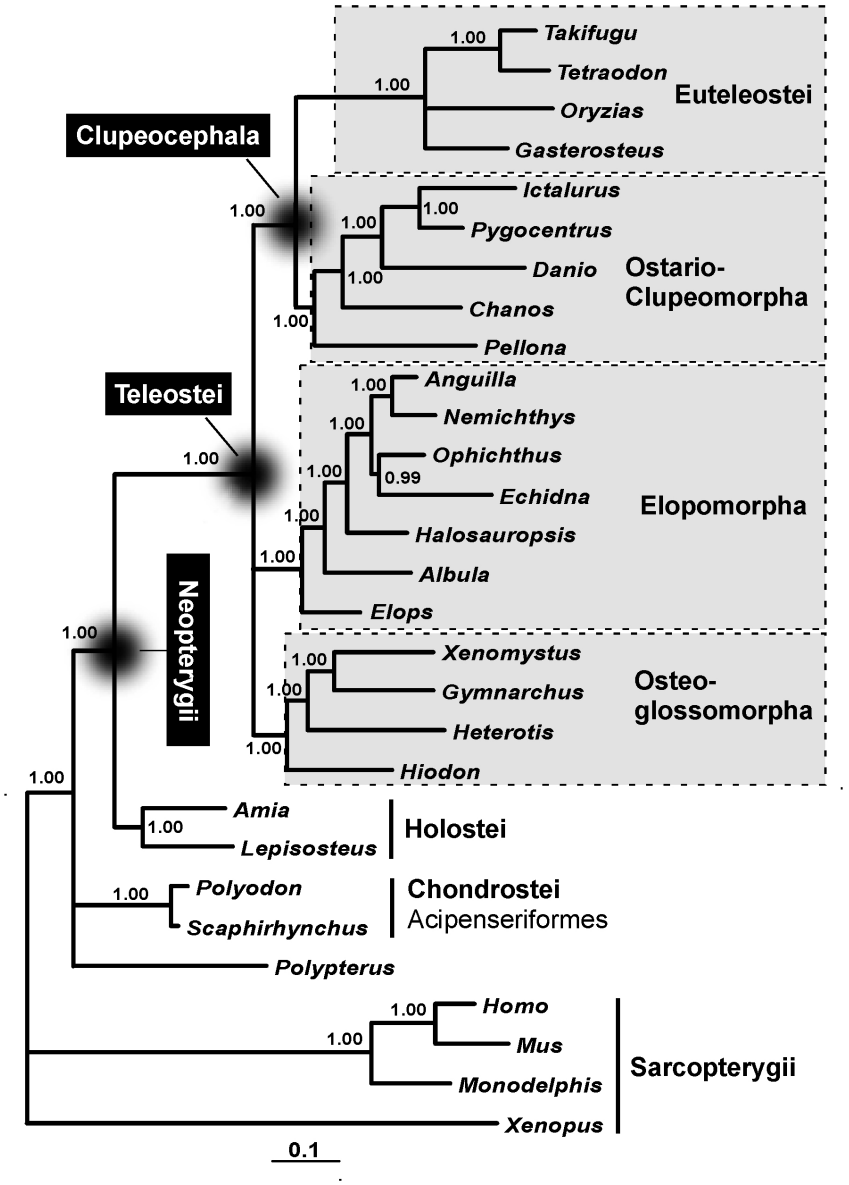

Fig. 1.4 Phylogeny of lower Actinopterygians and the early-branching teleost lineages. The tree is a consensus obtained by Bayesian analysis of nuclear gene DNA sequences (8 genes, 11,766 bp) for 29 representative taxa. Posterior probabilities (only values >0.95) are shown for well-supported groups and branches with low support were collapsed (e.g., a branch uniting *Polypterus* and Acipenseriforms and a branch uniting elopomorphs and osteoglossomorphs). Original.

1.3.2.1 Teleostei

The monophyly of Teleostei is supported by many morphological characters (de Pinna 1996; Arratia 2000). Four major teleostean lineages are currently recognized: Elopomorpha, Osteoglossomorpha, Ostarioclupeomorpha (or Otocephala = Clupeiformes plus Ostariophysi), and Euteleostei (Nelson 2006). Ostarioclupeomorphs are generally placed as the sister-group of euteleosts (Lê *et al.* 1993; Arratia 1997; Inoue *et al.* 2001) a grouping named Clupeocephala, that excludes elopomorphs and osteoglossomorphs (fig. 1.4). Interrelationships among elopomorphs, osteoglossomorphs, and clupeocephalans are still controversial. Both morphological (Patterson and Rosen 1977) and molecular (Inoue *et al.* 2001) studies support the position of osteoglossomorphs at the base of the teleosts, but this view was challenged by the alternative placing of elopomorphs as the living sister-group of all other teleosts (Arratia 1991; Shen 1996; Arratia 1997, 2000). A third alternative was suggested by Lê *et al.* (1993) based on relatively weak evidence from 28S ribosomal gene sequences, with osteoglossomorphs and elopmorphs forming a clade that is sister to clupeocephalans. We obtained the same result, albeit with low support (posterior probablility = 0.72) based on Bayesian analysis of 8 nuclear genes (Fig. 1.4). Therefore, at this time there is no unequivocal evidence to resolve with confidence the basal teleost trichotomy.

Support is strong for the Ostarioclupeomorpha hypothesis (Otocephala), placing the Clupeiformes as a sister group to Ostariophysi (Figs. 1.4 and 1.5). A recent result based on mitogenomic data (Saitoh *et al.* 2003), however, suggests that gonorynchiforms are more closely related to Clupeiformes, but this result could be due to poor taxonomic sampling or an analytical artifact. In most molecular studies (Dimmick and Larson 1996; Ortí and Meyer 1996; Saitoh *et al.* 2003), relationships within Ostariophysi are consistent with the traditional view (Fink and Fink 1981) placing Cypriniformes as a sister group to the rest, but relationships among Characiformes, Siluriformes, and Gymnotiformes cannot be resolved with confidence (see discussion in Saitoh *et al.* 2003). The close relationship shown in Fig. 1.5. between *Ictalurus* (Siluriformes) and *Pygocentrus* (Characiformes), to the exclusion of *Apteronotus* (Gymnotiformes) should, therefore, be taken with caution.

Our knowledge of the identity and relationships among major euteleostean lineages range from well corroborated to poorly understood. Johnson and Patterson (1996) and Lecointre and Nelson (1996) provide synapomorphies supporting euteleost monophyly, but mitogenomic data suggest an alternative definition (Ishiguro *et al.* 2003). Several early-branching euteleost lineages have been placed in the Protacanthopterygii (Greenwood *et al.* 1966), a supraordinal taxon that has undergone major re-definitions since its creation (Johnson and Patterson 1996). Recent molecular studies of basal euteleosts by Ishiguro *et al.* (2003) based on mitogenomic data and by Lopez *et al.* (2004) based on12S and 16S mitochondrial rRNA genes (815 bp) and exon 3 of the RAG-1 gene (1444 bp) examined protacanthopterygian taxa. Both corroborated the position of Esociformes (pikes, pickerels, and

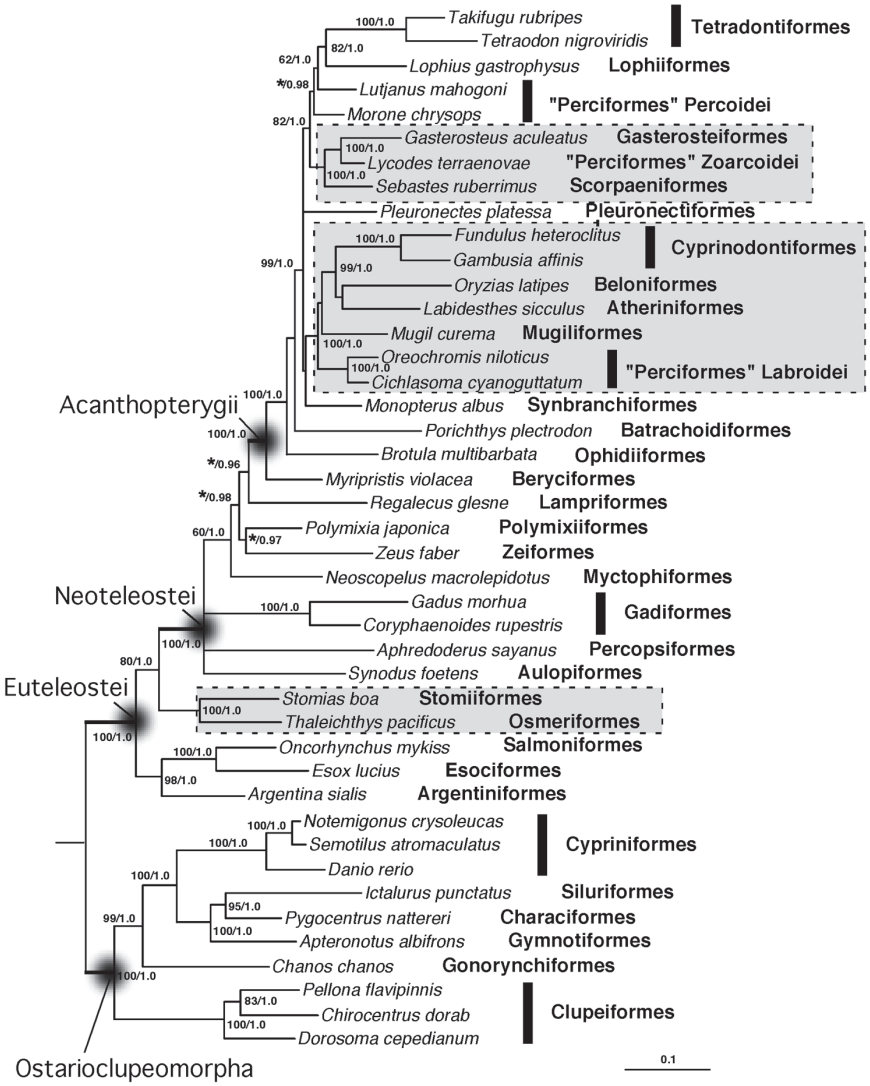

Fig. 1.5 Phylogeny of representative clupeocephalans (euteleosts plus otocephalans) based on analysis of 10 nuclear genes (7995 bp). The numbers on branches are maximum likelihood bootstrap values and Bayesian posterior probabilities. Asterisks indicate a bootstrap value < 50%. The names of representative species, orders, and supraordinal taxa are indicated. In grey boxes we highlight non-traditional groupings discussed in the text. Modified from Li, C., Lu, G., and Ortí, G. 2008. Systematic Biology ~~57: in press~~, Fig. 4. 57(4): 519-539

mudminnows) as the sister group of Salmoniformes, a clade also supported by analysis of 10 nuclear genes (Li *et al.* 2008). This clade is either sister to or in a polytomy with the marine smelts (Argentiniformes), freshwater smelts

(Osmeriformes), and Neoteleostei (Fig 1.5). The identity, composition and relationships of Prothacanthopterygii still await analysis of taxon-rich basal euteleost data sets, but an interesting result of these studies is the suggestion that Osmeriformes and Stomiiformes are closely related, thus removing Stomiiformes from the Neoteleostei (Fig. 1.5, see also Jamieson, **Chapter 12**).

The monophyly and relationships of the basal neoteleost clades (Stomiiformes, Aulopiformes and Myctophiformes) to the crown group Acanthomorpha have been relatively well established based on morphology and mitogenomic data (Stiassny 1986; Johnson 1992; Patterson and Johnson 1995; Miya *et al.* 2003; Springer and Johnson 2004). Nuclear gene data necessary to fully assess these relationships are still unavailable, but our limited taxonomic sampling for 10 genes (Li *et al.* 2008) suggests that traditional hypotheses may not be supported. We discuss next the main implications of our study.

Rosen (1973) proposed Acanthomorpha (including more than 15,000 species) to comprise two subgroups, Acanthopterygii and Paracanthopterygii. Acanthomorph monophyly has since been corroborated by both morphology and DNA data (Stiassny 1986; Johnson and Patterson 1993; Smith and Wheeler 2006), but the monophyly of the two subgroups has been refuted. In addition to the placement of Stomiiformes outside of the Neoteleostei, our analysis rejects the notion of Paracanthopterygii, a classical grouping of neoteleosts that has been extensively debated in the literature (Greenwood *et al.* 1966; Patterson and Rosen 1989; Miya *et al.* 2003, 2005). Taxa traditionally included in this group are placed (Fig. 1.5) either among the early-branching lineages of Neoteleosts close to Aulopiformes and Myctophiformes (Gadiformes, Percopsiformes), closer to the base of Acanthopterygii (Ophidiformes, and Batrachoidiformes), or in a derived position close to Tetraodontiformes (Lophiiformes). The paracanthopterygian hypothesis proposed by mitogenomic analyses (Miya *et al.* 2003, 2005) also was not supported in our study since *Polymixia* and *Zeus* did not form a monophyletic group with gadiforms and percopsiforms.

comma

comma

A problematic group missing in our analysis is the jellynoses (Ateleopodidae), which has been placed at the base of the Acanthomorpha (Miya *et al.* 2003). Lanternfishes (Lampriformes) are currently considered sister to the Acanthomorpha (Johnson 1992; Smith and Wheeler 2006), a position consistent with the placement obtained with the nuclear gene data (Fig. 1.5).

Acanthopterygii is strongly supported in our study as a monophyletic group only if ophidiiforms, batrachoidiforms and lophiiforms are included. Representative taxa for Beryciformes, Ophidiiformes and Batrachoidifromes branch off sequentially from the base of the acanthopterygians. The rest of the taxa included in our study formed a monophyletic group of crown acanthopterygians with a 99% bootstrap support (Fig. 1.5). Most notoriously without resolution is the crown of the teleost tree, represented by just a few species in our analysis. Relationships within the dominant acanthomorph group, Percomorpha, are essentially unknown. Percomorpha comprises more than 13,000 species of fishes (Nelson, 2006). The majority of economically

delete comma

important fishes are percomorphs, yet their monophyly and relationships remain virtually unknown. In our study, taxa traditionally assigned to the order Perciformes (*Lutjanus, Morone, Lycodes, Oreochromis, and Cichlasoma*) do not form a monophyletic group, in agreement with several results suggesting the polyphyletic nature of this group (Lauder and Liem 1983; Johnson and Patterson 1993; Miya *et al.* 2003; Nelson 2006). Although the very limited sampling of perciform taxa in our study of 10 nuclear genes precludes general results, two noteworthy groupings emerge. First, some elements of the suborder Labroidei, such as cichlids, are placed in the same clade with atherinomorphs (Atheriniforms, Beloniformes, and Cyprinodontiformes) and Mugiliformes to the exclusion of other perciforms (Fig. 1.5). A close relationship among rice fish (*Oryzias*) and tilapia (*Oreochromis*) was first suggested by a phylogenomic study (Chen *et al.* 2004). Second, the three-spined stickleback (Gasterosteidae) is grouped with a zoarcid perciform (*Lycoides*), corroborating previous results suggesting this relationship by analysis of mitogenomic data (Miya *et al.* 2003). A recent, expanded study of 11 families of "Gasterosteiformes" based on the same type of data (Kawahara *et al.* 2008) clearly refutes the monophyly of this order, establishing that gasterosteiform fishes form indeed three separate lineages: Syngnathoidei, Gasterosteoidei (minus Indostomidae), and Indostomidae.

Much remains to be learned about the identity and relationships of many important groups of ray-finned fishes. An exhaustive review of the literature for all groups is beyond the scope of this chapter, so many taxa remain without mention. We are confident that within the next few years important discoveries will be made with insights from a concerted effort underway to combine molecular and morphological data. Resolution of the tree of life of fishes still is far away, but the stage is set for rapid progress to establish the branching pattern that explains their amazing diversity.

1.4 CHAPTER SUMMARY

Our knowledge of relationships among "fishes" ranges from well corroborated to very poorly understood. Fishes constitute a large, heterogeneous and paraphyletic assemblage of distantly related jawless and jawed vertebrates, with most of their living diversity found in the crown group of actinopterygians (the ray-finned fishes). As a group, they have fundamental relevance to understanding the evolution of vertebrate animals and their features, and they also carry great commercial importance. Yet, their phylogenetic relationships remain largely unknown. Currently, no explicit and comprehensive analyses featuring all groups are available to support a sound phylogenetic classification of fishes. This situation is likely to improve relatively soon in light of ongoing efforts to compile and analyze large phylogenomic data sets that span the diversity of fishes in concert with integrated studies of their morphology and development. In this chapter, we summarize some recent advances in these fields and discuss some of the challenges that lie ahead. We emphasize the molecular aspects and illustrate this with some recent and ongoing studies.

Two major challenges for molecular systematics of fishes involve ancient gene and whole genome duplication events and systematic biases such as base compostition heterogeneity among DNA sequences sampled from distantly related taxa. Whereas the latter may lead to spurious phylogenetic results, the former may impede proper interpretation of gene genealogies as indicators of species phylogeny. Some possible solutions to these problems are presented. We also summarize some results from recent studies concerning especially the relationships among jawless fishes, the early branching lineages of ray-finned fishes, and some unexpected relationships among the crown euteleost groups. We find compelling evidence from molecular studies that support the cyclostome hypothesis that groups lampreys and hagfish. Among basal actinopterygians, recent re-interpretation of morphology and new evidence from nuclear gene sequences support the Holostei (grouping amiids and lepisosteids). Relationships among the early-branching teleosts groups remain unresolved with elopomorphs, osteoglossomorphs, and clupeocephalans (ostarioclupeomorphs plus euteleosts) forming a polytomy at the base of the teleost tree. We discuss briefly our current understanding of supra-ordinal groups of euteleosts such as Protacanthopterygii, Paracanthopterygii, and Acanthopterygii. Relationships among the percomorph crown group are virtually unknown.

Much remains to be learned about the identity and relationships among the many groups of fishes, but we anticipate that the next few years will witness significant advances to establish the branching pattern of the tree of life of all fishes.

1.5 ACKNOWLEDGMENTS

We thank Wei-Jei Chen for help with the generation of sequence data and analyses of RAG-1, especially to compile Figure 1.1. This work was funded by grants DEB-9985045 and DEB-0732838 from the National Science Foundation (USA) to G.O.

1.6 LITERATURE CITED

Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Langeland, J., Prince, V., Wang, Y. L., Westerfield, M., Ekker, M., and Postlethwait, J. H. 1998. Zebrafish hox clusters and vertebrate genome evolution. Science 282: 1711-1714.

Arratia, G. 1991. The caudal skeleton of Jurassic teleosts; a phylogenetic analysis. Pp. 249-340. In M.-M. Chang, H. Liu, and G.-R. Zhang (ed.), *Early vertebrates and related problems in evolutionary biology*. Science Press, Beijing.

Arratia, G. 1997. Basal teleosts and teleostean phylogeny. Palaeontologia Ichthyologica. 7: 5-168.

Arratia, G. 2000. Phylogenetic relationships of teleostei: past and present. Estudios Oceanológicos 19: 19-51.

Arratia, G. 2001. The sister group of Teleostei: consensus and disagreements. Journal of Vertebrate Paleontology 21: 767-773.

Bapteste, E., Brinkmann, H., Lee, J. A., Moore, D. V., Sensen, C. W., Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M., and Philippe, H. 2002. The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. Proceedings of the National Academy of Sciences USA 99: 1414-1419.

Bernardi, G. 1993. The vertebrate genome: isochores and evolution. Molecular Biology and Evolution 10: 186-204.

Blair, J. E., and Hedges, S. B. 2005. Molecular Phylogeny and Divergence Times of Deuterostome Animals. Molecular Biology and Evolution 22: 2275–2284.

Blanquart, S., and Lartillot, N. 2006. A Bayesian Compound Stochastic Process for Modeling Nonstationary and Nonhomogeneous Sequence Evolution. . Molecular Biology and Evolution 23: 2058-2071.

Blomme, T., Vandepoele, K., De Bodt, S., Simillion, C., Maere, S., and Van de Peer, Y. 2006. The gain and loss of genes during 600 million years of vertebrate evolution. Genome Biology 7: R43.

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

Chen, W. J., Ortí, G., and Meyer, A. 2004. Novel evolutionary relationship among four fish model systems. Trends in Genetics 20: 424-431.

Cloutier, R., and Arratia, G. 2004. Early diversification of actinopterygians. Pp. 217-270. In G. Arratia, M. V. H. Wilson, and R. Cloutier (eds), *Recent advances in the origin and early radiation of vertebrates*. Verlag Dr Friedrich Pfeil, Munich.

Collins, T. M., Fedrigo, O., and Naylor, G. J. P. 2005. Choosing the Best Genes for the Job: The Case for Stationary Genes in Genome-Scale Phylogenetics. Systematic Biology 54: 493-500.

Comas, I., Moya, A., and Gonzales-Candelas, F. 2007. From phylogenetics to phylogenomics: the evolutionary relationships of insect endosymbiotic gamma-Proteobacteria as a test case. Systematic Biology 56: 1-16.

Cotton, J. A., and Page, R. D. 2002. Going nuclear: gene family evolution and vertebrate phylogeny reconciled. Proceedings of the Royal Society of London B 269: 1555-1561.

de Pinna, M. C. C. 1996. Teleostean monophyly. Pp. 193-207. In M. L. J. Stiassny, L. R. Parenti, and G. D. Johnson (eds), *Interrelationships of fishes*. Academic Press, San Diego.

Delarbre, C., Gallut, C., Barriel, V., Janvier, P., and Gachelin, G. 2002. Complete mitochondrial DNA of the hagfish, *Eptatretus burgeri*: the comparative analysis of mitochondrial DNA sequences strongly supports the cyclostome monophyly. Molecular Phylogenetics and Evolution 22: 184–192.

Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature 439: 965-968.

Delsuc, F., Brinkmann, H., and Philippe, H. 2005. Phylogenomics and the reconstruction of the tree of life. Nat Rev Genet 6: 361-375.

Dimmick, W. W., and Larson, A. 1996. A molecular and morphological perspective on the phylogenetic relationships of the otophysan fishes. Molecular Phylogenetics and Evolution 6: 120-133.

Diogo, R. 2007. The Origin of Higher Clades:Osteology, Myology, Phylogeny and Evolution of Bony Fishes and the Rise of Tetrapods. Science Publishers, Enfield, NH.

Donoghue, P. C., Forey, P. L., and Aldridge, R. J. 2000. Conodont affinity and chordate phylogeny. Biological Reviews (Cambridge) 75: 191-251.

Driskell, A. C., Ane, C., Burleigh, J. G., McMahon, M. M., O'Meara, B. C., and M.J., S. 2004. Prospects for building the tree of life from large sequence databases. Science 306: 1172-1174.

Eisen, J. A., and Fraser, C. M. 2003. Phylogenomics: intersection of evolution and genomics. Science 300: 1706-1707.

Felsenstein, J. 1978. Cases in which parsimony or capatibility methods will be positively misleading. Systematic Zoology 27: 401-410.

Fink, S. V., and Fink, W. L. 1981. Interrelationships of the ostariophysan fishes (Teleostei). Zoological Journal of the Linnean Society 72: 297-353.

Fitch, W. M. 1970. Distinguishing homologous from analogous proteins. Systematic Zoology 19: 99-113.

Foster, P. G. 2004. Modeling compositional heterogeneity. Systematic Biology 53: 485-495.

Foster, P. G., and Hickey, D. A. 1999. Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. Journal of Molecular Evolution 48: 284-290.

Furlong, R. F., and Holland, P. W. 2002. Bayesian phylogenetic analysis supports monophyly of ambulacraria and of cyclostomes. Zoological Science 19: 593-599.

Galtier, N., and Gouy, M. 1998. Inferring pattern and process: maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. Molecular Biology and Evolution 15: 871-879.

Gardiner, B. G., Schaeffert, B., and Masserie, J. A. 2005. A review of the lower actinopterygian phylogeny. Zoological Journal of the Linnean Society 144: 511-525.

Gee, H. 2003. Evolution: ending incongruence. Nature 425:782.

Gowri-Shankar, V., and Rattray, M. 2007. A reversible jump method for Bayesian phylogenetic inference with a nonhomogeneous substitution model. Molecular Biology and Evolution 24: 1286-99.

Grande, L. 2007. Morphology Based Phylogenetic Study of Gars, Basal Neopterygian Interrelationships, and the Resurrection of Holostei. American Society of Ichthyologists and Herpetologists, Abstracts of Annual Meetings. ASIH, St. Louis, Missouri, USA.

Grande, L., and Bemis, W. E. 1996. Interrelationships of Acipenseriformes, with comments on "Chondrostei". Pp. 85-115. In M. L. J. Stiassny, L. R. Parenti, and G. D. Johnson (eds), *Interrelationships of fishes*. Academic Press, San Diego.

Greenwood, P. H., Miles, R. S., and Patterson, C. (eds) 1973. Interrelationships of fishes. Academic Press, London.

Greenwood, P. H., Rosen, D. E., Weitzman, S. H., and Meyers, G. S. 1966. Phyletic studies of teleostean fishes, with a provisional classification of living forms. Bulletin of the American Museum National History 131: 339-456.

Gu, X., and Li, W.-H. 1996. Bias-Corrected Paralinear and LogDet Distances and Tests of Molecular Clocks and Phylogenies Under Nonstationary Nucleotide Frequencies. Molecular Biology and Evolution 13: 1375-1383.

Hardisty, M. W. 1982. Lampreys and hagfishes: analysis of cyclosome relationships. Pp. 165–258. In M. W. Hardisy, and I. C. Potter (eds), *The biology of Lampreys*. Academic Press, London.

Helfman, G. S., Collette, B. B., and Facey, D. E. 1997. The Diversity of Fishes. Blackwell Science, Malden, MA.

Hurley, I. A., Mueller, R. L., Dunn, K. A., Schmidt, E. J., Friedman, M., Ho, R. K., Prince, V. E., Yang, Z., Thomas, M. G., and Coates, M. I. 2007. A new time-scale for ray-finned fish evolution. Proceedings of the Royal Society of London B 274: 489-498.

Inoue, J. G., Miya, M., Tsukamoto, K., and Nishida, M. 2001. A mitogenomic perspective on the basal teleostean phylogeny: resolving higher-level relationships with longer DNA sequences. Molecular Phylogenetics and Evolution 20: 275-285.

Inoue, J. G., Miya, M., Tsukamoto, K., and Nishida, M. 2003. Basal actinopterygian relationships: a mitogenomic perspective on the phylogeny of the "ancient fish". Molecular Phylogenetics and Evolution 26: 110-120.

Ishiguro, N. B., Miya, M., and Nishida, M. 2003. Basal euteleostean relationships: a mitogenomic perspective on the phylogenetic reality of the "Protacanthopterygii". Molecular Phylogenetics and Evolution 27: 476-488.

Jaillon, O., Aury, J. M., Brunet, F., Petit, J. L., Stange-Thomann, N. ~~and at, e.~~ 2004. et al. [italics]
Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 431: 946-957.

Jamieson, B. G. M. 1991. Fish Evolution and Systematics: Evidence from Spermatozoa. Cambridge University Press, Cambridge.

Janvier, P. 1996. Early vertebrates. Clarendon Press, Oxford.

Jenner, R. A. 2004. Accepting Partnership by Submission? Morphological Phylogenetics in a Molecular Millennium. Systematic Biology 53: 333 - 342.

Jessen, H. 1972. Schultergürtel und Pectoralflosse bei Actinopterygiern. Fossils and Strata 1: 1–101.

Johnson, G. D. 1992. Monophyly of the euteleostean clades: Neoteleostei, Eurypterygii, and Ctenosquamata. Copeia: 8-25.

Johnson, G. D., and Patterson, C. 1993. Percomorph phylogeny: a survey of acanthomorphs and a new proposal. Bulletin of Marine Sciences 52: 554-626.

Johnson, G. D., and Patterson, C. 1996. Relationships of lower euteleostean fishes. Pp. 251-332. In M. L. J. Stiassny, L. R. Parenti, and G. D. Johnson (ed.), *Interrelationships of fishes*. Academic Press, San Diego.

Jukes, T. H., and Bhushan, V. 1986. Silent nucleotide substitutions and G + C content of some mitochondrial and bacterial genes. Journal of Molecular Evolution 24: 39-44.

Kawahara, R., Miya, M., Mabuchi, K., Lavoué, S., Inoue, J. G., Satoh, T. P., Kawaguchi, A., and Nishida, M. 2008. Interrelationships of the 11 gasterosteiform families (sticklebacks, pipefishes, and their relatives): A new perspective based on whole mitogenome sequences from 75 higher teleosts. Molecular Phylogenetics and Evolution 46: 224-236.

Kikugawa, K., Katoh, K., Shigehiro, K., Sakurai, H., Ishida, O., Iwabe, N., and Miyata, T. 2004. Basal jawed vertebrate phylogeny inferred from multiple nuclear DNA-coded genes. BMC Biology 2: 1-11.

Kocher, T. D., and Stepien, C. A. 1997. Molecular systematics of fishes. Academic Press, San Diego.

Kubatko, L. S., and Degnan, J. H. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Systematic Biology 56: 17-24.

Kuraku, S., Hoshiyama, D., Katoh, K., Suga, H., and Miyata, T. 1999. Monophyly of lampreys and hagfishes supported by nuclear DNA-coded genes. Journal of Molecular Evolution 49: 729-735.

Lauder, G. V., and Liem, K. F. 1983. The evolution and interrelationships of the Actinopterygian fishes. Bulletin of the Museum of Comparative Zoology 150: 95-197.

Lê, H. L., Lecointre, G., and Perasso, R. 1993. A 28S rRNA-based phylogeny of the gnathostomes: first steps in the analysis of conflict and congruence with morphologically based cladograms. Molecular Phylogenetics and Evolution 2: 31-51.

Lecointre, G., and Nelson, G. 1996. Clupeomorpha, sister group of Ostrariophysi. Pp. 193-207. In M. L. J. Stiassny, L. R. Parenti, and G. D. Johnson (ed.), *Interrelationships of fishes*. Academic Press, San Diego.

Li, C., Lu, G., and Ortí, G. 2008. Optimal Data Partitioning and a Test Case for Ray-finned Fishes (Actinopterygii) Based on Ten Nuclear Loci. Systematic Biology ~~57: in press.~~ 57(4): 519-539.

Li, C., and Ortí, G. 2007. Molecular phylogeny of Clupeiformes (Actinopterygii) inferred from nuclear and mitochondrial DNA sequences. Molecular Phylogenetics and Evolution 44: 386-398.

Li, C., Ortí, G., Zhang, G., and Lu, G. 2007. A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. BMC Evolutionary Biology 7: 44.

Lockhart, P. J., Steel, M. A., Hendy, M. D., and Penny, D. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Molecular Biology and Evolution 11: 605-612.

Lopez, A. J., Chen, W. J., and Ortí, G. 2004. Esociform phylogeny. Copeia 2004: 449-464.

Lopez, P., Casane, D., and Philippe, H. 2002. Heterotachy, an important process of protein evolution. Molecular Biology and Evolution 19: 1-7.

Lynch, M., and Conery, J. S. 2000. The evolutionary fate and consequences of duplicate genes. Science 290: 1151-1155.

Mabee, P. M., Ashburner, M., Cronk, Q., Gkoutos, G. V., Haendel, M., Segerdell, E., Mungall, C., and Westerfield, M. 2007. Phenotype ontologies: the bridge between genomics and evolution. Trends in Ecology and Evolution 22: 345-350.

Mallat, J. 1984. Early vertebrate evolution: Pharyngeal structure and the origin of gnathostomes. Journal of Zoology 204: 169–183.

Mallat, J. 1997. Hagfish do not resemble ancestral vertebrates. J. Morphol. 232: 293.

Mallat, J., and Winchell, J. 2007. Ribosomal RNA genes and deuterostome phylogeny revisited: More cyclostomes, elasmobranchs, reptiles, and a brittle star. Molecular Phylogenetics and Evolution 43: 1005-1022.

Mallatt, J., and Sullivan, J. 1998. 28S and 18S rDNA sequences support the monophyly of lampreys and hagfishes. Molecular Biology and Evolution 15: 1706-1718.

McMahon, M. M., and Sanderson, M. J. 2006. Phylogenetic supermatrix analysis of GenBank sequences from 2228 papilionoid legumes. Systematic Biology 55: 818-836.

Meyer, A., and Van de Peer, Y. 2005. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). Bioessays 27: 937-945.

Meyer, A., and Zardoya, R. 2003. Recent advances in the (molecular) phylogeny of vertebrates. Annual Review of Ecology, Evolution, and Systematics 34: 311-318.

Miya, M., and Nishida, M. 2000. Use of mitogenomic information in teleostean molecular phylogenetics: a tree-based exploration under the maximum-parsimony optimality criterion. Molecular Phylogenetics and Evolution 17: 437-455.

Miya, M., Satoh, T. P., and Nishida, M. 2005. The phylogenetic position of toadfishes (order Batrachoidiformes) in the higher ray-finned fish as inferred from partitioned Bayesian analysis of 102 whole mitochondrial genome sequences. Biological Jourbnal of the Linnean Society of London 85: 289-306.

Miya, M., Takeshima, H., Endo, H., Ishiguro, N. B., Inoue, J. G., Mukai, T., Satoh, T. P., Yamaguchi, M., Kawaguchi, A., Mabuchi, K., Shirai, S. M., and Nishida, M. 2003. Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. Molecular Phylogenetics and Evolution 26: 121-138.

Nelson, G. J. 1969. Gill arches and the phylogeny of fishes, with notes on the classification of vertebrates. Bulletin of the American Museum National History 141: 475-552.

Nelson, J. S. 2006. Fishes of the world, 4th edition. John Wiley and Sons, Inc., New York.

Ohno, S. 1970. Evolution by Gene Duplication. Springer-Verlag.

Olsen, P. E. 1984. The skull and pectoral girdle of the parasemionotid fish *Watsonulus eugnathoides* from the Early Triassic Sakamena Group of Madagascar, with comments on the relationships of the holostean fishes. Journal of Vertebrate Paleontology 4: 481-499.

Ortí, G., Chen, W. J., Mahon, A., Holcroft Benson, N., Unmack, P., and others. 2005. Phylogeny of the ray-finned fish (Actinopterygii): Analysis of more than 500 RAG-1 exon 3 sequences. American Society of Ichthyologists and Herpetologists, Abstracts of Annual Meeting. ASIH, Tampa, Florida, USA.

Ortí, G., and Li, C. 2007. Chondrostei, Holostei, Teleostei? Nuclear DNA Sequence Data. American Society of Ichthyologists and Herpetologists, Abstracts of Annual Meeting. ASIH, St. Louis, Missouri, USA.

Ortí, G., and Meyer, A. 1996. Molecular Evolution of Ependymin and the Phylogenetic Resolution of Early Divergences Among Euteleost Fishes. Molecular Biology and Evolution 13: 556-573.

Ota, K. G., and Kuratani, S. 2007. Cyclostome embryology and early evolutionary history of vertebrates. Integrative and Comparative Biology 47: 329-337.

Page, R. D., and Cotton, J. A. 2002. Vertebrate phylogenomics: reconciled trees and gene duplications. Pacific Symposium of Biocomputing 2002: 536-547.

Pamilo, P., and Nei, M. 1988. Relationships Between Gene Trees and Species Trees. Molecular Biology and Evolution 5: 568-583.

Pancer, Z., Amemiya, C. T., Ehrhardt, G. R., Ceitlin, J., Gartland, G. L., and Cooper, M. D. 2004. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature 430.

Pancer, Z., Saha, N. R., Kasamatsu, J., Suzuki, T., Amemiya, C. T., Kasahara, M., and Cooper, M. D. 2005. Variable lymphocyte receptors in hagfish. Proceedings of the National Academy of Sciences USA 102: 9224-9229.

Patterson, C. 1973. Interrelationships of holosteans. Pp. 207-226. In P. H. Greenwood, R. S. Miles, and C. Patterson (eds), *Interrelationships of fishes*. Academic Press, London.

Patterson, C., and Johnson, G. D. 1995. The intermuscular bones and ligaments of teleostean fishes. Smithsonian Contributions of Zoology 559: 1-85.

Patterson, C., and Rosen, D. E. 1977. Review of ichthyodectiform and other Mesozoic teleost fishes and the theory and practice of classifying fossils. Bulletin of the American Museum National History 158: 81-172.

Patterson, C., and Rosen, D. E. 1989. The Paracanthopterygii revisited: order and disorder. Pp. 5-36. In D. M. Cohen (ed.), *Papers on the systematics of gadiform fishes*. Natural History Museum of Los Angeles County, Los Angeles, California.

Philippe, H., Snell, E. A., Bapteste, E., Lopez, P., Holland, P. W., and Casane, D. 2004. Phylogenomics of eukaryotes: impact of missing data on large alignments. Molecular Biology and Evolution 21: 1740-1752.

Phillips, M. J., Delsuc, F., and Penny, D. 2004. Genome-scale phylogeny and the detection of systematic biases. Molecular Biology and Evolution 21: 1455-1458.

Rast, J. P., Smith, L. C., Loza-Coll, M., Hibino, T., and Litman, G. W. 2006. Genomic Insights into the Immune System of the Sea Urchin. Science 314: 952-956.

Regan, C. T. 1923. The skeleton of *Lepidosteus*, with remarks on the origin and evolution of the lower neopterygian fishes. Proceedings of the Zoological Society: 445-461.

Remm, M., Storm, C. E., and Sonnhammer, E. L. 2001. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. Journal of Molecular Biology 314: 1041-1052.

Rokas, A., Kruger, D., and Carroll, S. B. 2005. Animal evolution and the molecular signature of radiations compressed in time. Science 310: 1933–1938.

Rokas, A., Williams, B. L., King, N., and Carroll, S. B. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425: 798-804.

Rosen, D. E. 1973. Interrelationships of higher euteleostean fishes. . Pp. 397-513. In P. H. Greenwood, R. S. Miles, and C. Patterson (eds), *Interrelationships of Fishes*. Academic Press, London.

Rosen, D. E. 1982. Teleostean interrelationships, morphological function and evolutionary inference. American Zoologist 22: 261-273.

Saitoh, K., Miya, M., Inoue, J. G., Ishiguro, N. B., and Nishida, M. 2003. Mitochondrial genomics of ostariophysan fishes: perspectives on phylogeny and biogeography. Journal of Molecular Evolution 56: 464-472.

Schaeffer, B. 1973. Interrelationships of chondrosteans. Pp. 207-226. In P. H. Greenwood, R. S. Miles, and C. Patterson (eds), *Interrelationships of fishes*. Academic Press, London.

Scotland, R. W., Olmstead, R. G., and Bennett, J. R. 2003. Phylogeny Reconstruction: The Role of Morphology. Systematic Biology 52: 539 - 548.

Shen, M. 1996. Fossil "osteoglossomorphs" in East Asia and their implications in teleostean phylogeny. Pp. 261-272. In G. Arratia, and G. Viohl (eds), *Mesozoic fishes: systematics and plaleoecology*. Verlag Dr. F. Pfeil, München, Germany.

Smith, N. D., and Turner, A. H. 2005. Morphology's Role in Phylogeny Reconstruction: Perspectives from Paleontology. Systematic Biology 54: 166 - 173.

Smith, W. L., and Wheeler, W. C. 2006. Venom evolution widespread in fishes: a phylogenetic road map for the bioprospecting of piscine venoms. Journal of Heredity 97: 206-217.

Soltis, D. E., Albert, V. A., Savolainen, V., Hilu, K., Qiu, Y. L., Chase, M. W., Farris, J. S., Stefanovic, S., Rice, D. W., Palmer, J. D., and Soltis, P. S. 2004. Genome-scale data, angiosperm relationships, and "ending incongruence": a cautionary tale in phylogenetics. Trends in Plant Science 9: 477-483.

Springer, V. G., and Johnson, G. D. 2004. Study of the Dorsal Gill-arch Musculature of Teleostome Fishes, with Special Reference to the Actinopterygii. Bulletin of the Biological Society of Washington 11: 260.

Steel, M. A., Lockhart, P. J., and Penny, D. 1993. Confidence in evolutionary trees from biological sequence data. Nature 364: 440-442.

Stiassny, M. L. J. 1986. The limits and relationships of acanthomorph teleosts. Journal of Zoology B 1: 411-460.

Stiassny, M. L. J., Parenti, L. R., and Johnson, G. D. (eds) 1996. *Interrelationships of fishes*. Academic Press, San Diego.

Stiassny, M. L. J., Wiley, E. O., Johnson, G. D., and de Carvalho, M. R. 2004. Gnathostome fishes. Pp. 410-429. In J. Cracraft, and M. J. Donoghue (eds), *Assembling The Tree of Life*. Oxford University Press, New York.

Stock, D. W., and Whitt, G. S. 1992. Evidence from 18S ribosomal RNA sequences that lampreys and hagfishes form a natural group. Science 257: 787–789.

Takezaki, N., Figueroa, F., Zaleska-Rutczynska, Z., and Klein, J. 2003. Molecular phylogeny of early vertebrates: monophyly of the agnathans as revealed by sequences of 35 genes. Molecular Biology and Evolution 20: 287-292.

Taylor, J. S., and Raes, J. 2004. Duplication and Divergence: The Evolution of New Genes and Old Ideas. Annual Review of Genetics 9: 615-643.

Van de Peer, Y., Taylor, J. S., and Meyer., A. 2003. Are all fishes ancient polyploids? Journal of Structural and Functional Genomics 3: 65-73.

Venkatesh, B., Erdmann, M. V., and Brenner, S. 2001. Molecular synapomorphies resolve evolutionary relationships of extant jawed vertebrates. Proceedings of the National Academy of Sciences USA 98: 11382-11387.

Waddell, P. J. 2005. Measuring the fit of sequence data to phylogenetic model: allowing for missing data. Molecular Biology and Evolution 22: 395-401.

Weisburg, W. G., Giovannoni, S. J., and Woese, C. R. 1989. The Deinococcus-Thermus phylum and the effect of rRNA composition on phylogenetic tree construction. Systematic and Applied Microbiology 11: 128-134.

Wiens, J. J. 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Systematic Biology 52: 528-538.

Wiens, J. J. 2004. The Role of Morphological Data in Phylogeny Reconstruction. Systematic Biology 53: 653 - 661.

Woese, C. R., Achenbach, L., Rouviere, P., and Mandelco, L. 1991. Archaeal phylogeny: reexamination of the phylogenetic position of Archaeoglobus fulgidus in light of certain composition-induced artifacts. Systematic and Applied Microbiology 14: 364-371.

Woods, I. G., Wilson, C., Friedlander, B., Chang, P., Reyes, D. K., Nix, R., Kelly, P. D., Chu, F., Postlethwait, J. H., and Talbot, W. S. 2005. The zebrafish gene map defines ancestral vertebrate chromosomes. Genome Res 15: 1307-1314.

Yang, Z., and Roberts, D. 1995. On the use of nucleic acid sequences to infer early branchings in the tree of life. Molecular Biology and Evolution 12: 451-458.