

Conflict and Resolution Between Phylogenies Inferred From Molecular and Phenotypic Data sets for Hagfish, Lampreys, and Gnathostomes

THOMAS J. NEAR*

Department of Ecology and Evolutionary Biology and Peabody Museum of Natural History, Yale University, New Haven, Connecticut

ABSTRACT One of the most problematic issues in vertebrate phylogenetics is the disagreement between phenotypic and molecular inferences regarding the relationships among hagfishes, lampreys, and gnathostomes. Phenotypic characters support monophyly of lampreys and gnathostomes, whereas nearly all published analyses of molecular data sets support monophyly of hagfishes and lampreys. In this study I present results of phylogenetic analyses of combined phenotypic and molecular data sets that focus on relationships among hagfishes, lampreys, and gnathostomes. Maximum parsimony analyses of 115 phenotypic characters combined with 4,638 rRNA sites and more than 10,000 amino acids each result in monophyly of lampreys and gnathostomes, demonstrating that the addition of relatively few phenotypic characters can alter phylogenetic inferences from large molecular data sets. On the other hand, Bayesian analyses of the combined data sets support monophyly of hagfish and lampreys, indicating that model-based analyses may be prone to data “swamping,” where the phylogenetic signal of the larger molecular data sets overwhelm the signal present in the much smaller phenotypic data set. Nodes that relate hagfish and lampreys were recovered at a low frequency in parametric bootstrapping analyses, indicating that the timing of diversification among hagfishes, lampreys, and gnathostomes has created a difficult phylogenetic problem for molecular data. The fact that addition of relatively few phenotypic characters can alter phylogenetic inferences of cyclostome monophyly obtained from molecular data sets, and the inability of simulated data sets to recover key nodes in the craniate phylogeny provide reasons to view the strong support for cyclostome monophyly inferred from molecular data sets with a measured degree of skepticism. *J. Exp. Zool. (Mol. Dev. Evol.)* 312B:749–761, 2009. © 2009 Wiley-Liss, Inc.

How to cite this article: Near TJ. 2009. Conflict and resolution between phylogenies inferred from molecular and phenotypic data sets for hagfish, lampreys, and gnathostomes. *J. Exp. Zool. (Mol. Dev. Evol.)* 312B:749–761.

The conflict between molecular and phenotypic data sets regarding the relationships of hagfish and lampreys remains one of the most vexing problems in vertebrate phylogenetics (Forey and Janvier, '93; Patterson et al., '93; Forey, '95; Janvier, '96b; Donoghue et al., 2000; Meyer and Zardoya, 2003; Janvier, 2007a,b). Extant craniates are represented by three major clades, the hagfishes (Myxinoidea), lampreys (Petromyzontiformes), and the gnathostomes, or jawed vertebrates. Huxley (1876) and Parker (1883) hypothesized that the similar appearance of hagfish and lampreys was the result of degenerative changes associated with parasitism. Cope (1889) classified

hagfish and lampreys in the class Agnatha, and this classification was used in Hubbs and Potter's ('71) comprehensive and influential review of lamprey species diversity. Noting the absence of jaws, paired fins, and exoskeletons, later investigators hypothesized that hagfish and lampreys

Grant sponsor: National Science Foundation; Grant number: DEB-0716155.

*Correspondence to: Thomas J. Near, Department of Ecology and Evolutionary Biology and Peabody Museum of Natural History, Yale University, New Haven, CT 06520. E-mail: thomas.near@yale.edu

Received 12 January 2009; Revised 21 March 2009; Accepted 1 April 2009

Published online 28 April 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jez.b.21293

shared common ancestry relative to gnathostomes, and were grouped together in Cyclostomata (Dean, 1895; Goodrich, '09, '30; Jollie, '62; Jarvik, '68; Stensiö, '68). This hypothesis was later supported in comparative anatomical studies of the feeding system and branchial arch morphology that is very similar between hagfish and lampreys, and quite different from gnathostomes (Shaeffer and Thomson, '80; Yalden, '85).

With the advent of phylogenetic systematics and increased knowledge of the diversity of extinct jawless vertebrate lineages (e.g. "ostracoderms"), the hypothesis that lampreys and gnathostomes are sister lineages is consistently supported in analyses of morphological, physiological, and biochemical characters (Løvtrup, '77; Hardisty, '79; Janvier and Blicek, '79; Janvier, '81; Hardisty, '82; Maisey, '86; Forey and Janvier, '93; Forey, '95; Donoghue et al., 2000; Gess et al., 2006). Molecular phylogenetic investigations of craniate relationships were initiated in the early 1990s, and essentially all published analyses using DNA or amino acid sequences have resulted in a monophyletic Cyclostomata with high bootstrap and significant Bayesian posterior support. The types of genes used to investigate cyclostome phylogeny have included nuclear-encoded rRNA genes (Stock and Whitt, '92; Lipscomb et al., '98; Mallatt and Sullivan, '98; Mallatt et al., 2001; Mallatt and Winchell, 2007), DNA sequences from single copy nuclear genes (Kuraku et al., '99; Furlong and Holland, 2002), hemoglobin amino acid sequences (Lanfranchi et al., '94), concatenated amino acid sequences from single copy nuclear genes (Hedges, 2001; Takezaki et al., 2003; Blair and Hedges, 2005; Delsuc et al., 2006; Kuraku and Kuratani, 2006; Yu et al., 2008), and whole mtDNA genomes (Delarbre et al., 2002).

Quite surprisingly, there has been no effort to investigate the relationships of lampreys, hagfishes, and gnathostomes using phylogenetic analyses that combine molecular and phenotypic characters. The reasoning for such an exercise is not to satisfy a broad spectrum of investigators that may feel more confident in phylogenetic trees inferred from either molecular or phenotypic data sets, but rather to reveal aspects of character evolution in these data sets that may result in the observed phylogenetic incongruence. True incongruence between data sets should involve more than a simple comparison of tree topologies resulting from separate analyses of multiple data sets (Mickey and Farris, '81; Farris et al., '94; de Queiroz et al., '95; Huelsenbeck et al., '96a;

Baker et al., '98), because the differences may involve the phylogenetic relationships of only a subset of all the sampled lineages, or the incongruence may derive from divergent patterns of character diversification among the contrasted data sets (e.g. long-branch attraction).

Another important reason to explore phylogenetic analyses of combined molecular and phenotypic data sets is the ability to include extinct lineages in analyses that include data, such as molecular characters, that can only be sampled from extant species. This is a particularly promising avenue in the investigation of craniate phylogeny, because it would facilitate the inclusion of extinct clades (e.g. Osteostraci, Arandaspida, Heterostraci, and Anaspida) that along with hagfish and lampreys are hypothesized to form a paraphyletic grade of jawless vertebrates, or agnathans (Janvier, '81; Shu et al., '99; Donoghue et al., 2000; Gess et al., 2006). The inclusion of these fossil lineages are potentially critical to the estimation of craniate phylogeny (Donoghue et al., 2000; Gess et al., 2006), because it is well known that fossil taxa can exhibit dramatic influence on the results of phylogenetic analyses that include only extant lineages (Doyle and Donoghue, '87; Gauthier et al., '88; Donoghue et al., '89).

My goal in this study is to assess the relationships of hagfish, lampreys, and gnathostomes when molecular and phenotypic characters are combined in maximum parsimony and Bayesian phylogenetic analyses. I also use parametric bootstrapping to investigate the ability of DNA data sets simulated on different craniate time-calibrated phylogenies to resolve the phylogenetic relationships of hagfish, lampreys, and gnathostomes. I had initially predicted that the phylogenetic signal of the phenotypic characters would be overwhelmed, or "swamped," by the signal present in the much larger molecular phylogenetic data sets, and phylogenetic analysis of simulated DNA data sets would converge on relationships that mirrored the tree topology used in the parametric bootstrapping simulations. The results did not support my preliminary predictions. Combination of molecular and phenotypic data sets had dramatic effects on the inferred phylogenies of hagfish and lampreys in maximum parsimony analyses, and parametric bootstrap analyses indicated that simulated data sets had difficulty resolving nodes in the model trees that involved relationships of hagfish, lampreys, and gnathostomes.

MATERIALS AND METHODS

Phylogenetic data sets

Data sets for phylogenetic analyses were gathered from three of the most recently published studies and include aligned nuclear-encoded ribosomal RNA genes (Mallatt and Sullivan, '98), aligned concatenated amino acid sequences sampled from 146 nuclear-encoded genes (Delsuc et al., 2006), and 115 discretely coded phenotypic characters (Gess et al., 2006). All of these studies minimally sampled one hagfish species, one lamprey species, and the molecular data sets included several gnathostome species. The phenotypic data set presented in Gess et al. (2006) sampled several extinct jawless vertebrate lineages that included Anaspida, Arandaspida, Heterostraci, and Osteostraci. The taxon sampling used for the phenotypic data is the one with the reduced taxon sampling used in Gess et al. (2006), because the inclusion of lineages with a large fraction of missing characters in the data set were leading to reduced phylogenetic resolution. These lineages included the Euconodonata, *Mayomyzon*, and Galeaspida. The inclusion or exclusion of these taxa did not result in any change in phylogenetic inferences with regard to the relationships of hagfish and lampreys (Gess et al., 2006). Outgroup sampling in each study consisted of a cephalochordate and a urochordate species. Delsuc et al. (2006) sampled broadly across the major animal phyla, but in this study the amino acid data set was restricted to cephalochordates, urochordates, hagfish, lampreys, and gnathostomes.

One of three rRNA alignments used in Mallatt and Sullivan ('98) was downloaded from the EMBL Nucleotide Sequence Database (DS35049), and the taxon sampling is presented in Table 1. This alignment contained the 18S, 5.8S rRNA, core regions of the 28S gene, and portions of the

28S divergent domains that produced satisfactory alignments. See Mallatt and Sullivan ('98) for a discussion of alignment protocols used for the rRNA data set. The alignment was modified by not including the teleost *Onchorhynchus mykiss*, because there was no data available for the 18S rRNA gene. The first 54 base pairs in the alignment that were missing for *Acipenser brevirostrum* and *Latimeria chalumnae*, and these sites were removed from the alignment and not used in the subsequent phylogenetic analyses.

The amino acid alignment of 146 concatenated nuclear genes used in Delsuc et al. (2006) was obtained directly from Hervé Philippe (Université de Montréal). The taxon sampling for the amino acid data is presented in Table 2. Amino acid sites that were missing data for hagfish or lamprey in the alignment were removed and not used in the subsequent phylogenetic analyses, leaving 14,140 aligned amino acid sites from approximately 95 genes (see Supplementary Table S2 in Delsuc et al., 2006). Hagfish (*Eptatretus burgeri* and *Myxine glutinosa*) and eutherian mammals (five species, but mostly represented by *Homo sapiens*) were sampled as mixed species chimaeras (Delsuc et al., 2006).

The phenotypic data was taken from Gess et al. (2006) and is a matrix of coded character states for 115 characters assembled from previously published phylogenetic analyses of major craniate lineages (Janvier, '96a; Donoghue et al., 2000; Donoghue and Smith, 2001). Characters in this data matrix include those scored from the sensory and nervous system, morphology of the mouth and branchial system, circulatory system, fin morphology, osteology, and physiology. Yalden ('85) presented a list of eleven synapomorphies from the feeding apparatus supporting cyclostome monophyly that were summarized as a single character in subsequent phylogenetic analyses (Donoghue

TABLE 1. Taxon sampling for the ribosomal RNA sequences used in phylogenetic analyses

| Clade | Species | Common name |
|--------------------|-------------------------------|-----------------|
| Urochordata | <i>Herdmania momus</i> | Tunicate |
| Cephalochordata | <i>Branchiostoma floridae</i> | Cephalochordate |
| Myxiniiformes | <i>Eptatretus stouti</i> | Hagfish |
| Petromyzontiformes | <i>Petromyzon marinus</i> | Lamprey |
| Gnathostomata | <i>Hydrolagus colliei</i> | Chimaera |
| Gnathostomata | <i>Squalus acanthias</i> | Shark |
| Gnathostomata | <i>Acipenser brevirostrum</i> | Sturgeon |
| Gnathostomata | <i>Latimeria chalumnae</i> | Coelacanth |
| Gnathostomata | <i>Xenopus laevis</i> | Frog |

Common names are those used in the phylogenetic trees.

TABLE 2. Taxon sampling for the amino acid data set used in phylogenetic analyses

| Clade | Species | Common name |
|--------------------|--|-----------------|
| Urochordata | <i>Ciona intestinalis</i> | Tunicate |
| Cephalochordata | <i>Branchiostoma floridae</i> | Cephalochordate |
| Myxiniiformes | <i>Eptatretus burgei</i> / <i>Myxine glutinosa</i> | Hagfish |
| Petromyzontiformes | <i>Petromyzon marinus</i> | Lamprey |
| Gnathostomata | <i>Tetraodon nigroviridis</i> | Pufferfish |
| Gnathostomata | <i>Danio rerio</i> | Zebrafish |
| Gnathostomata | <i>Homo sapiens</i> | Mammal |
| Gnathostomata | <i>Gallus gallus</i> | Chicken |
| Gnathostomata | <i>Xenopus laevis</i> | Frog |
| Gnathostomata | <i>Ambystoma mexicanum</i> | Salamander |

Common names are those used in the phylogenetic trees.

et al., 2000; Donoghue and Smith, 2001). Gess et al. (2006) reexamined this set of characters through specimen dissections and the single character of the feeding apparatus was expanded conservatively to three characters. The character coding for the phenotypic characters in the phylogenetic analyses of combined phenotypic–molecular data sets was modified slightly to accommodate differences in taxon sampling between the molecular and phenotypic character matrices. The gnathostomes were represented as one taxon in the phenotypic character matrix, so this scoring was used for each of the gnathostome species represented in the molecular data matrices in the combined data analyses.

Phylogenetic analyses

The computer program PAUP* 4.0 (Swofford, 2003) was used to perform separate maximum parsimony analyses on each of the rRNA, amino acid, and phenotypic data sets. In addition, maximum parsimony tree searches were performed on the combined rRNA sequences and phenotypic characters (with and without the inclusion of extinct agnathan lineages), and the amino acid sequences combined with phenotypic characters (with and without the inclusion of extinct agnathan lineages). In all analyses the most parsimonious tree(s) was found using a branch-and-bound search algorithm, and the sampled cephalochordate and urochordate species were used as outgroup taxa in all analyses. Node support in each analysis was assessed using a bootstrap resampling analysis, with a branch-and-bound search algorithm, and 2,000 pseudoreplicates.

Bayesian phylogenetic analyses were performed on the rRNA, phenotypic, and combined rRNA–phenotypic data sets using partitioned analyses

(Ronquist and Huelsenbeck, 2003). Metropolis-coupled Markov chain Monte Carlo (MC3) was used to estimate posterior probabilities (Larget and Simon, '99; Huelsenbeck et al., 2001), as executed in the computer program MrBayes 3.1 (Ronquist and Huelsenbeck, 2003). The optimal maximum likelihood model for the rRNA data set was determined using the Akaike Information Criterion (AIC) as executed in the computer program Modeltest 3.6 (Posada and Crandall, '98). Phenotypic characters were treated as “standard” in MrBayes and used the M_{kv} model of character evolution (Lewis, 2001). This model allows for there to be k number of discrete character states and the rate of character change to be equal or to vary. Bayes factors were calculated from the harmonic mean of the maximum likelihood scores from Bayesian analyses and were used to determine whether a gamma distribution of among character rate variation was a more optimal model of evolution for the phenotypic characters (Nylander et al., 2004; Brandley et al., 2005; Müller and Reisz, 2006). The Bayes factor was calculated by multiplying by two the difference of the harmonic mean of the maximum likelihood scores (Brandley et al., 2005). A Bayes factor score greater than 10 was considered very strong support for the alternative model as following Kass and Raftery ('95).

Five sets of Bayesian analyses were run, one for the rRNA data set, two for the phenotypic data sets (with and without the inclusion of extinct agnathan lineages), and two for the combined rRNA and phenotypic data sets (with and without the inclusion of extinct agnathan lineages). The amino acid data set was not subjected to Bayesian analyses, owing to the computational intensity (i.e. estimated 1,450 hr on a Macintosh G5 computer). In each analysis MrBayes 3.1 was run for

6.0×10^6 generations to ensure convergence of the MC3 algorithm and the number of generations discarded as the burn-in was determined by plotting the maximum likelihood score vs. the number of generations. The posterior probabilities of nodes in the phylogeny were calculated from the set of post burn-in trees. All phylogenetic trees resulting from the Bayesian analyses were rooted with the cephalochordate and urochordate out-group species.

Parametric bootstrapping

The ability of simulated data sets to recover nodes on model phylogenies depicting the timing of diversification among the major craniate clades was explored using parametric bootstrapping (Felsenstein, '88; Goldman, '93; Huelsenbeck et al., '96b). The computer program Seq-Gen was used to generate two sets of 1,000 simulated data sets that were based on the number of nucleotides and optimal molecular evolutionary model observed for the combined 18S, 5.8S, and 28S rRNA data set used in the phylogenetic analyses (Rambaut and Grassly, '97). The two sets of simulated nucleotide rRNA data were generated on two different model phylogenies that reflected the two divergent hypotheses of craniate relationships (Fig. 1). The branch lengths of the model trees were based on paleontological age estimates that time calibrate phylogenies using stratigraphic ranges and inferred ghost lineages (Donoghue et al., 2000, 2004; Janvier, 2007b). Each of the two ultrametric trees was scaled by 0.2 to reflect the observed mean path length from the tips to the deepest node in the phylogeny in a maximum likelihood analysis.

The best maximum likelihood tree was estimated from each 1,000 simulated data sets generated from each of the two model trees using a heuristic tree search in PAUP* 4.0. The optimal nucleotide substitution model and model parameter values used in the data simulations were used in the maximum likelihood tree searches. The frequency of each clade present in the optimal maximum likelihood trees was recorded. For example, if 889 of the phylogenies estimated from the simulated data sets contained a clade that included hagfish and lamprey, this clade was present in 88.9% of the simulated data phylogenies.

RESULTS

Maximum parsimony analyses

With regard to cyclostome relationships, maximum parsimony analysis of each data set resulted in phylogenies that were consistent with previous results (Mallatt and Sullivan, '98; Mallatt et al., 2001; Delsuc et al., 2006; Gess et al., 2006; Mallatt and Winchell, 2007). Details of the maximum parsimony analyses are presented in Table 3. The most parsimonious trees found using the rRNA and amino acid data sets included a monophyletic Cyclostomata, but bootstrap support for this node was low in the rRNA analysis (Fig. 2A and B). Relationships among the sampled gnathostome lineages were inconsistent with generally accepted phylogenetic hypotheses, specifically only the gnathostome node and the chimera-shark node are consistent with hypotheses based on substantial phenotypic and molecular evidence (e.g. Meyer and Zardoya, 2003; Kikugawa et al., 2004; Stiasny et al., 2004). Lampreys were the sister lineage of

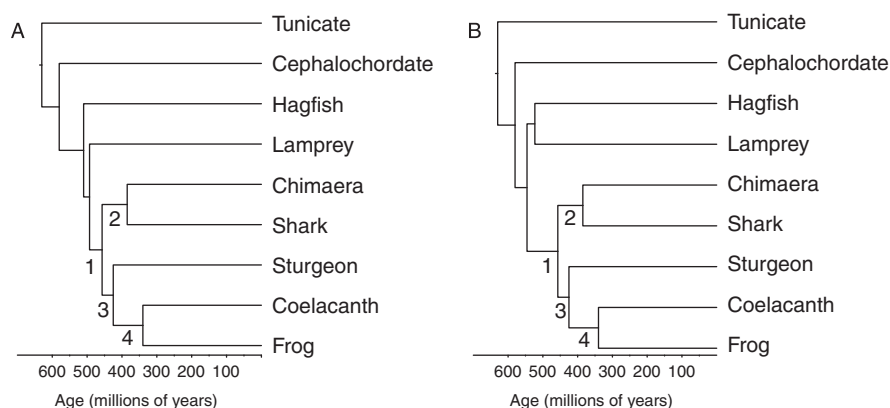


Fig. 1. Phylogenies that depict temporal hypotheses of diversification among major craniate lineages. Major vertebrate clades are identified with numbers, 1. Gnathostomata, 2. Chondrichthyes, 3. Osteichthyes, and 4. Sarcopterygii. (A) Hypothesis depicting monophyly of lampreys and gnathostomes. (B) Hypothesis depicting monophyly of hagfishes and lampreys.

TABLE 3. Summary of maximum parsimony analyses (branch-and-bound tree searches) using ribosomal RNA (rRNA), amino acid, phenotypic, and combined data sets

| Data set | MPT | Consistency index ^a | Tree length |
|---|-----|--------------------------------|-------------|
| rRNA (Fig. 2A) | 1 | 0.630 | 1,994 |
| Amino acids (Fig. 2B) | 1 | 0.733 | 12,044 |
| Phenotypic (Fig. 2C) | 1 | 0.793 | 130 |
| Phenotypic, extant only (Fig. 2D) | 1 | 0.707 | 181 |
| rRNA and phenotypic (Fig. 3A) | 1 | 0.642 | 2,177 |
| rRNA and phenotypic, extant only (Fig. 3B) | 1 | 0.652 | 2,126 |
| Amino acids and phenotypic (Fig. 4A) | 1 | 0.731 | 12,244 |
| Amino acids and phenotypic, extant only (Fig. 4B) | 1 | 0.733 | 12,193 |

The number of most parsimonious trees (MPT), consistency index, and tree length are reported for each analysis.

^aExcluding phylogenetically uninformative characters.

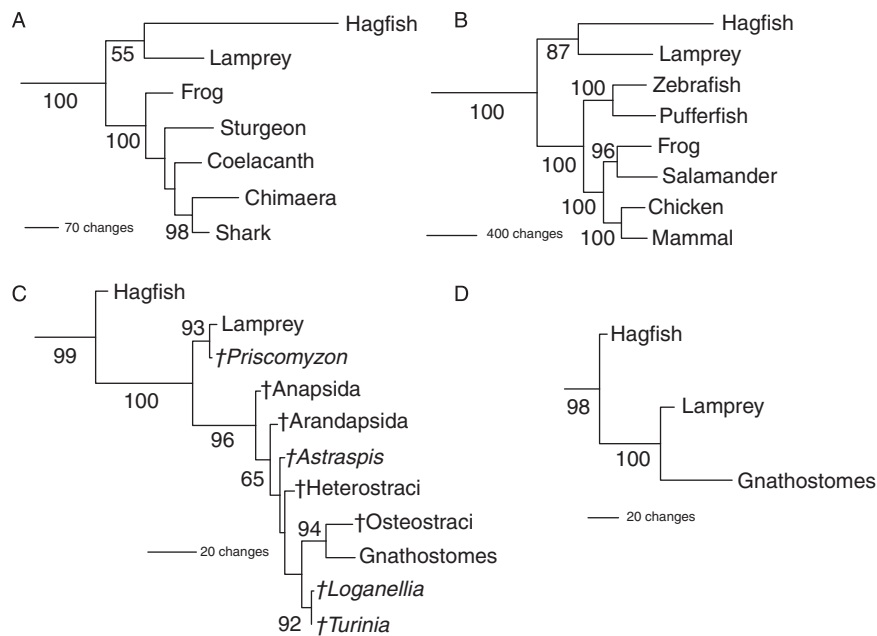


Fig. 2. Phylogenies resulting from maximum parsimony analyses (see Table 3). Numbers at nodes report bootstrap pseudoreplicate scores. Extinct vertebrate lineages are marked with a dagger. (A) Most parsimonious tree resulting from analysis of the rRNA data set. (B) Most parsimonious tree resulting from analysis of the amino acid data set. (C) Most parsimonious tree resulting from phenotypic data set that includes extinct vertebrate lineages. (D) Most parsimonious tree resulting from phenotypic data set that includes only extant craniate lineages.

gnathostomes in the parsimony analysis of phenotypic characters (Fig. 2C and D). Bootstrap support for the lamprey–gnathostome node was high and inferred relationships and levels of bootstrap support were consistent between analyses that included and excluded the extinct fossil agnathan lineages.

The maximum parsimony analyses of the concatenated molecular and phenotypic data sets resulted in cyclostome paraphyly and monophyly of lampreys and gnathostomes (Figs. 3A and B, 4A and B). The most parsimonious trees resulting from analyses of the concatenated rRNA and

phenotypic data sets showed no difference in tree topology and bootstrap support between the inclusion and exclusion of fossil agnathan lineages. In both analyses, there was high bootstrap support for the monophyly of lampreys and gnathostomes (Fig. 3A and B). Similar to the rRNA inferred phylogeny (Fig. 2A), analysis of the concatenated rRNA and phenotypic data set resulted in relationships among the sampled gnathostome lineages that are massively contradicted by other phenotypic and molecular evidence. The most parsimonious trees resulting from analyses of the concatenated amino acid

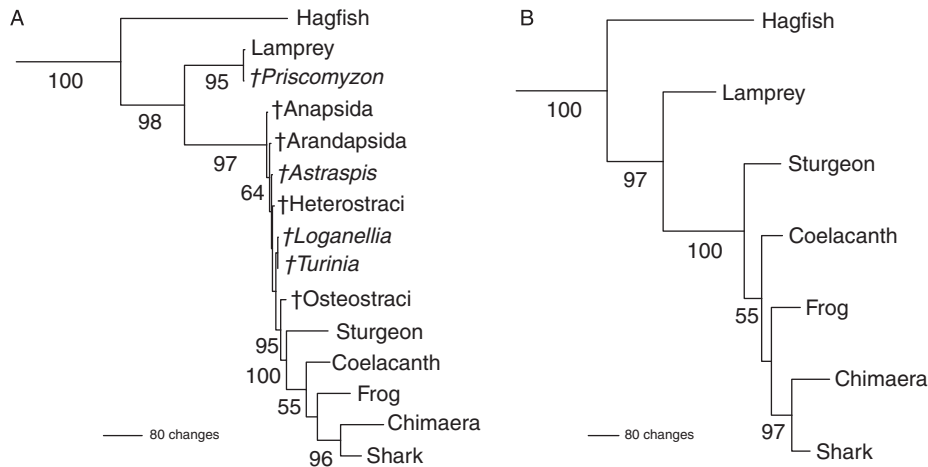


Fig. 3. Phylogenies resulting from maximum parsimony analyses (see Table 3). Numbers at nodes report bootstrap pseudoreplicate scores. Extinct vertebrate lineages are marked with a dagger. (A) Most parsimonious tree resulting from analysis of the combined phenotypic and rRNA data set that includes extinct vertebrate lineages. (B) Most parsimonious tree resulting from analysis of the combined phenotypic and rRNA data set that includes only extant craniate lineages.

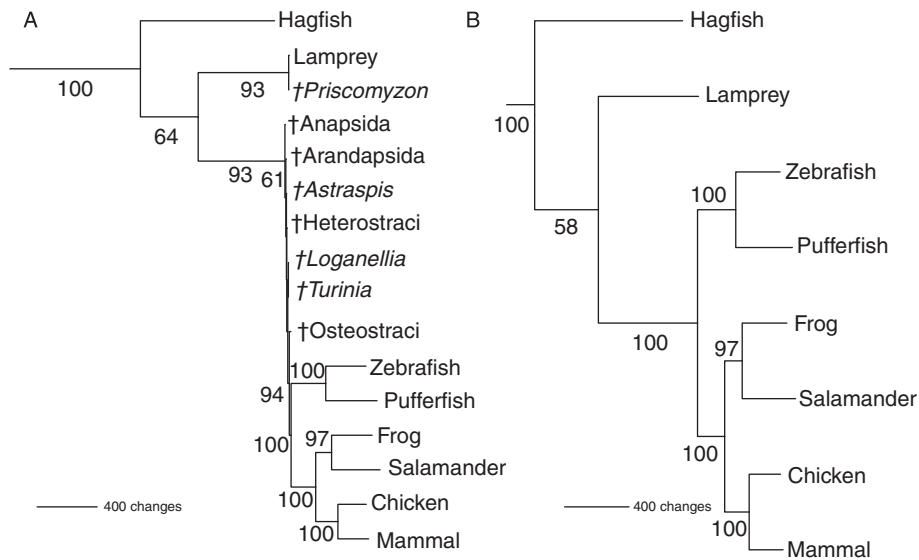


Fig. 4. Phylogenies resulting from maximum parsimony analyses (see Table 3). Numbers at nodes report bootstrap pseudoreplicate scores. Extinct vertebrate lineages are marked with a dagger. (A) Most parsimonious tree resulting from analysis of the combined phenotypic and amino acid data set that includes extinct vertebrate lineages. (B) Most parsimonious tree resulting from analysis of the combined phenotypic and amino acid data set that includes only extant craniate lineages.

and phenotypic data sets supported the monophyly of lampreys and gnathostomes, but this relationship was not strongly supported with bootstrap pseudoreplicate scores in the analyses that included and excluded the extinct fossil agnathan lineages (Fig. 4A and B).

Bayesian analyses

The optimal molecular evolutionary model identified for the rRNA data set using AIC was

GTR+I+G. The harmonic mean of the maximum likelihood score for the phenotypic data set (excluding fossil agnathan lineages) with equal rates for characters was -317.33 , and for a model with a gamma distribution of among character rate variation was -310.12 , giving a Bayes factor score of 14.42 that indicates strong support for the M_{kv} +gamma model for the phenotypic data set.

Bayesian phylogenetic analysis of the rRNA data set resulted in cyclostome monophyly with significant posterior support (Fig. 5A). Similar to the

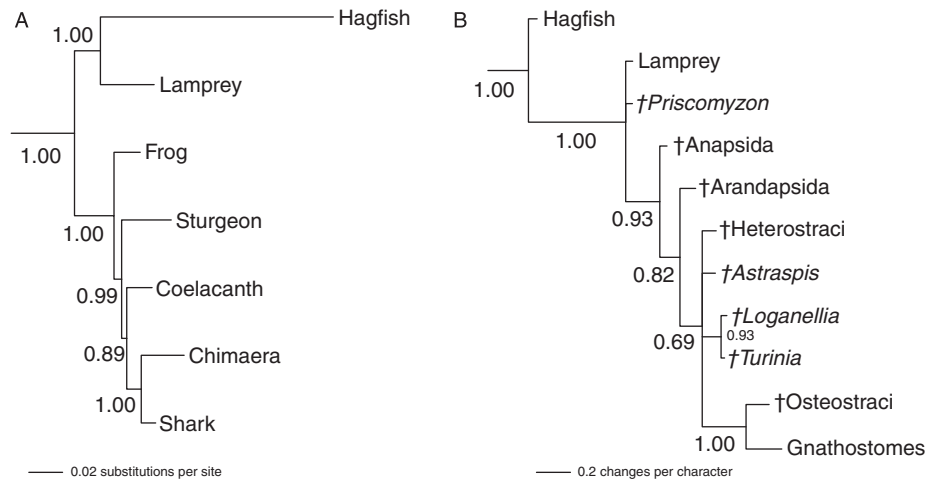


Fig. 5. Phylogenies resulting from Bayesian analyses. Numbers at nodes report Bayesian posterior probabilities. Extinct vertebrate lineages are marked with a dagger. (A) Tree resulting from analysis of the rRNA data set. (B) Tree resulting from analysis of the phenotypic data set that includes extinct vertebrate lineages.

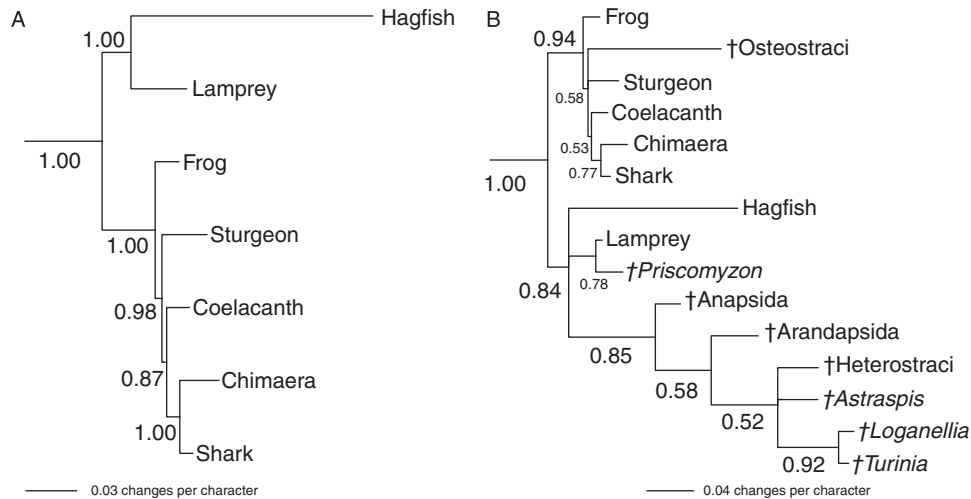


Fig. 6. Phylogenies resulting from Bayesian analyses. Numbers at nodes report Bayesian posterior probabilities. Extinct vertebrate lineages are marked with a dagger. (A) Tree resulting from analysis of the combined phenotypic and rRNA data sets that only includes extant craniate lineages. (B) Tree resulting from analysis of the combined phenotypic and rRNA data sets that includes extinct vertebrate lineages.

parsimony analyses, the Bayesian analysis resulted in a phylogeny that was inconsistent with regards to the relationships among sampled gnathostomes. Bayesian analysis of the phenotypic data set resulted in monophyly of lampreys and gnathostomes with significant posterior support (Fig. 5B). There were no differences in posterior tree topology and node support in the analyses of the phenotypic data set that included and excluded fossil agnathan lineages (tree not shown).

Bayesian analyses of the concatenated rRNA and phenotypic data sets all resulted in cyclostome monophyly (Fig. 6); however, relationships among sampled gnathostomes were again inconsistent

with broadly accepted phylogenetic hypotheses. The analysis that excluded all of the extinct agnathan lineages was identical in posterior tree topology and posterior node support to the analysis of the rRNA data set (Fig. 6A). Inclusion of the fossil agnathan lineages in the Bayesian analysis resulted in cyclostome monophyly, but this node was not supported with a significant Bayesian posterior probability (Fig. 6B).

Parametric bootstrapping

Maximum likelihood phylogenetic analyses of the 1,000 data sets simulated on each of the two

TABLE 4. Results of parametric bootstrapping analyses, showing frequency occurrence of nodes in model phylogenies (Fig. 1) recovered in maximum likelihood analyses of simulated rRNA data sets

| Node | Lamprey–gnathostome monophyly (Fig. 1A) | Hagfish–lamprey monophyly (Fig. 1B) |
|---------------------|---|-------------------------------------|
| Lamprey–gnathostome | 49.2 | NA |
| Hagfish–lamprey | NA | 53.5 |
| 1. Gnathostomata | 100 | 98.9 |
| 2. Chondrichthyes | 100 | 100 |
| 3. Osteichthyes | 99.3 | 99.3 |
| 4. Sarcopterygii | 100 | 100 |

phylogenetic hypotheses did not result in a high frequency of recovery for the nodes in these model trees that involve hagfish and lampreys (Fig. 1; Table 4). On the other hand, all four nodes in the two model phylogenies that involved gnathostome relationships were recovered in 99–100% of the simulations (Table 4). Data sets simulated on the model phylogeny, where lampreys and gnathostomes formed a clade, recovered this node in 53.5% of the simulated data sets (Fig. 1A; Table 4). Similarly, the data sets simulated on the phylogeny that depicted cyclostome monophyly recovered this node in 49.2% of the simulations (Fig. 1B; Table 4).

DISCUSSION

A fundamental problem in phylogeny of vertebrates is the conflict between molecular and phenotypic data sets with regard to the phylogenetic relationships of hagfishes, lampreys, and gnathostomes (Forey and Janvier, '93; Meyer and Zardoya, 2003; Janvier, 2007b). This is the first investigation to explore phylogenetic analyses of combined molecular–phenotypic character data sets, and I have discovered that a modest number of phenotypic characters have an appreciable impact on phylogenetic inferences in combined data analyses. Despite the presence of many more molecular characters relative to phenotypic characters in the concatenated data sets, the phylogenetic signal of the molecular data does not always appear to overwhelm, or “swamp,” the phylogenetic signal present in the phenotypic characters. Maximum parsimony analysis of the combined molecular–phenotypic data sets resulted in monophyly of lampreys and gnathostomes, but Bayesian analyses of the concatenated rRNA–phenotypic data sets support cyclostome monophyly. Perhaps the model-based

Bayesian analysis is more prone to the much larger molecular data partition overwhelming the phylogenetic signal of the smaller phenotypic data set. Such a phenomenon has never been investigated in model-based analyses of combined phenotypic and molecular data sets.

The timing of diversification among the major craniate lineages demands critical consideration when discussing their phylogenetic relationships. Hagfish, lampreys, and the remaining craniates (conodonts, “ostracoderms,” and gnathostomes) are a clade that diversified rapidly in the vastly distant past. The paleontological estimate for the age of the common ancestor of hagfish, lampreys, and gnathostomes is approximately 510 million years ago (mya), and the time period of diversification among the three extant vertebrate clades was less than 15 million years (Donoghue et al., 2000, 2004). This hypothesized timing of diversification is based on the estimated ghost range of lampreys, which is the estimated time between the earliest occurrence of the lineage in the fossil record and the minimal time the lineage must have existed because its sister lineage is present as fossils in older strata. Phylogenetic analyses of phenotypic characters have resulted in trees where lampreys are the sister lineage of the Stem Gnathostomata that includes the conodonts (Donoghue et al., 2000; Donoghue and Sansom, 2002; Gess et al., 2006). The oldest lamprey is from the Devonian and dates to approximately 360 mya (Gess et al., 2006; Janvier, 2007b, 2008), but the ghost range of lampreys can be extended 140 million years because the oldest conodont fossils date to approximately 500 mya and the most inclusive clade containing conodonts and gnathostomes is the sister lineage of lampreys (Donoghue et al., 2000, 2004; Donoghue and Sansom, 2002; Janvier, 2007b). Therefore, the time period of diversification of extant vertebrates is less than 3% of the total clade age. In such a scenario, molecular data may be at a disadvantage (Donoghue et al., '89), because it would have to evolve rapidly enough to generate molecular synapomorphies during the short time of diversification among the extant craniate clades, and remain conserved for nearly 500 million years to preserve the molecular synapomorphies among hagfish, lampreys, and gnathostomes. However, if the rate of molecular evolution remained similar to that which facilitated the generation of the molecular synapomorphies, then subsequent evolution over the course of hundreds of millions of years would result in divergent terminal branches. The limited number

of character states in molecular data and the divergent nature of the extant terminal lineages would result in an erosion of phylogenetic signal through multiple changes at specific nucleotide sites (Felsenstein, '78).

The results of the parametric bootstrap analyses reflect the difficulty in using DNA sequences to reconstruct phylogenetic relationships of two nodes separated by very short branching times, relative to the total clade age of approximately 510 mya. The key nodes that related hagfish and lampreys in the model phylogenies were never recovered at a high frequency in the phylogenetic analyses of the simulated data sets. However, younger nodes in the gnathostomes with larger time intervals separating them were recovered at a very high frequency (Fig. 1; Table 4). I suspect that very long terminal branches that are contrasted with very short internal branches that are temporally spaced close to the common ancestor will continue to hamper the investigation of craniate phylogenies using molecular data. This long-branch bias observed in the vertebrate phylogeny could result in the incorrect grouping of lineages owing to systematic error, which is the deviation of the estimated phylogeny from the true phylogeny owing to incorrect assumptions in the inference method (Felsenstein, '78; Huelsenbeck, '98; Anderson and Swofford, 2004). This is different from classic models of long-branch attraction that involves heterogeneous rates of character evolution among lineages; however, previous studies have demonstrated a similar long-branch effect that leads to inaccurate phylogenetic inference when rates of character evolution are uniform among lineages (Hendy and Penny, '89; Zharkikh and Li, '93; Felsenstein, 2004, p 119). Short internal branches relative to terminal lineages will result in a persistence of error that can potentially increase as the sampled data increases, leading to strong support for nodes in the incorrectly estimated phylogeny (Felsenstein, '78; Swofford et al., 2001). The fact that model-based analyses of molecular data sets support cyclostome monophyly (Fig. 5A, Furlong and Holland, 2002; Delsuc et al., 2006; Mallatt and Winchell, 2007), and cyclostome monophyly is supported with high bootstrap pseudoreplicate scores or significant Bayesian posterior probabilities in these molecular phylogenies is not evidence that these data sets are immune to systematic error. The suspicions of a long-branch bias affecting the analysis of craniate relationships using the available molecular data extend beyond

the obvious long terminal branches in the phylogeny (Fig. 1), but include the fact that in parsimony analyses the inclusion of few phenotypic characters results in very different trees (Figs. 3 and 4), and analysis of individual gene sequences do not always result in cyclostome monophyly (Takezaki et al., 2003; Yu et al., 2008).

When considering the phenotypic characters, it is commonly accepted that rates of morphological evolution are not uniform over long stretches of evolutionary time (Simpson, '44, '53; Eldredge and Gould, '72), and there is no observed correlation of rates of evolutionary change between molecular and phenotypic phylogenetic data sets (Bromham et al., 2002). Unlike rapidly evolving molecular characters, a morphological character state that originates along an early branch in a phylogeny may be retained in the descendants of that lineage, and not erased by subsequent evolutionary change (Donoghue et al., '89). This is a particularly compelling argument when the evolutionary concept of "burden," as introduced by Riedl ('78), is considered, as it provides an explanation for marked variability in character evolution over long periods of evolutionary time. Burden is the function of a trait relative to other characters and is quantified by the number of other characters that are dependent on its presence. As articulated by Donoghue ('89), the magnitude of burden on a particular character will relate to the likelihood of character modification, with changes at heavily burdened characters being less likely. Many of character states in the phenotypic data set that support the monophyly of lampreys and gnathostomes are characters conceivably under a heavy burden. Character states shared between lampreys and gnathostomes include, but are not limited to the presence of lateral line neuromasts, hyperosmoregulation, dorsal fins with fin rays and radial muscles, a braincase with lateral walls of cartilage, extrinsic eye muscles, a closed pericardium, and a heart innervated by a vagus that allows nervous regulation of the heart (Hardisty, '82; Gess et al., 2006). Changes in any of these characters would clearly have effects on a broad collection of associated organismal traits. The explanation for the presence of these traits shared by lampreys and gnathostomes from the perspective of cyclostome monophyly is one of simple character loss in hagfishes, or remarkable convergent evolution in lampreys and gnathostomes (Stock and Whitt, '92; Mallatt et al., 2001; Furlong and Holland, 2002; Takezaki et al., 2003; Mallatt and Winchell, 2007), without any consideration of

evolutionary burden of these traits and the likelihood of such dramatic character evolution across such a diverse array of anatomical and physiological traits.

In a review of congruence between phylogenetic inferences from molecular and phenotypic data sets Patterson et al. ('93) stated that the monophyly of cyclostomes inferred from molecular data sets "...seems to be another instance of "not proven" rather than outright incongruence." Addressing the same problem from a much different perspective, a recent study claimed, "The monophyly of cyclostomes can be regarded as one of the most clear-cut examples in which molecular phylogenetics has succeeded in updating phylogenetic relationships based on nonmolecular [*sic*] traits." (Kuraku, 2008). Far from delivering a resolution of the problem of craniate phylogeny, my investigation favors the former of these two assessments and depicts a fragile portrait of the apparently resolute monophyly of cyclostomes resulting from molecular phylogenetic studies. In particular, my study demonstrates that phylogenetic analyses that combine a relatively low number of phenotypic characters with large molecular data sets can dramatically alter the phylogenetic inference of cyclostome monophyly resulting from analyses of molecular data sets alone. Also, the short interval between the diversification events among hagfishes, lampreys, and gnathostomes, relative to the total age of the clade, presents a very hard phylogenetic problem for molecular data, leading to the erosion of phylogenetic signal and the potential for long-branch attraction. The results of the simulation studies are consistent with this scenario (Fig. 1; Table 4). Given the importance to basic organismal function of many of the phenotypic characters that support the monophyly of lampreys and gnathostomes, models of trait conservation consistent with the hypothesis of lamprey and gnathostome monophyly are more compelling than those of simple character loss offered by the hypothesis of cyclostome monophyly. The analyses and observations presented in this study indicate that the monophyly of the cyclostomes inferred from molecular phylogenetic analyses cannot be considered unequivocal.

ACKNOWLEDGMENTS

This study was motivated by undergraduate students in my ichthyology course, offered through Yale College, who were not satisfied with

my explanations for the apparent conflict between phenotypic and molecular data sets in the phylogeny of craniates. They had challenged me to seek a more satisfactory answer to the problem. In particular my efforts were aided by discussions with Matthew Brandley, Michael Donoghue, Alexis Jackson, Evan McCartney-Melstead, Brian Moore, and Robert Quigley. Matthew Brandley provided comments on an earlier version of this manuscript.

LITERATURE CITED

- Anderson FE, Swofford DL. 2004. Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rDNA. *Mol Phylogenet Evol* 33: 440–451.
- Baker RH, Yu X, DeSalle R. 1998. Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. *Mol Phylogenet Evol* 9:427–436.
- Blair JE, Hedges SB. 2005. Molecular phylogeny and divergence times of deuterostome animals. *Mol Biol Evol* 22:2275–2284.
- Brandley MC, Schmitz A, Reeder TW. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. *Syst Biol* 54:373–390.
- Bromham L, Woolfit M, Lee MSY, Rambaut A. 2002. Testing the relationship between morphological and molecular rates of change along phylogenies. *Evolution* 56:1921–1930.
- Cope ED. 1889. Synopsis of the families of Vertebrata. *Am Nat* 23:849–877.
- Dean B. 1895. Fishes, living and fossil. New York: Columbia University Press. 300p.
- Delarbre C, Gallut C, Barriol V, Janvier P, Gachelin G. 2002. Complete mitochondrial DNA of the hagfish, *Eptatretus burgeri*: the comparative analysis of mitochondrial DNA sequences strongly supports the cyclostome monophyly. *Mol Phylogenet Evol* 22:184–192.
- Delsuc F, Brinkmann H, Chourrout D, Philippe H. 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439:965–968.
- de Queiroz A, Donoghue MJ, Kim J. 1995. Separate versus combined analysis of phylogenetic evidence. *Annu Rev Ecol Syst* 26.
- Donoghue MJ. 1989. Phylogenies and the analysis of evolutionary sequences, with examples from seed plants. *Evolution* 43:1137–1156.
- Donoghue MJ, Doyle JA, Gauthier J, Kluge AG, Rowe T. 1989. The importance of fossils in phylogeny reconstruction. *Annu Rev Ecol Syst* 20:431–460.
- Donoghue PCJ, Sansom IJ. 2002. Origin and early evolution of vertebrate skeletonization. *Microsc Res Tech* 59:352–372.
- Donoghue PCJ, Smith MP. 2001. The anatomy of *Turinia pagei* (Powrie), and the phylogenetic status of the Thelodonti. *Trans R Soc Edin-Earth* 92:15–37.
- Donoghue PCJ, Forey PL, Aldridge RJ. 2000. Conodont affinity and chordate phylogeny. *Biol Rev* 75:191–251.
- Donoghue PCJ, Smith MP, Sansom IJ. 2004. The origin and early evolution of chordates: molecular clocks and the fossil record. In: Donoghue PCJ, Smith MP, editors. *Telling the evolutionary time*. London: Taylor & Francis. p 190–223.

- Doyle JA, Donoghue MJ. 1987. The importance of fossils in elucidating seed plant phylogeny and macroevolution. *Rev Palaeobot Palynol* 50:63–95.
- Eldredge N, Gould SJ. 1972. Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf TJM, editor. *Models in paleobiology*. San Francisco: Freeman, Cooper and Co. p 82–115.
- Farris JS, Källersjö M, Kluge AG, Bult C. 1994. Testing significance of incongruence. *Cladistics* 10:315–319.
- Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool* 27:401–410.
- Felsenstein J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annu Rev Genet* 22:521–565.
- Felsenstein J. 2004. *Inferring phylogenies*. Sunderland, MA: Sinauer. 664p.
- Forey P, Janvier P. 1993. Agnathans and the origin of jawed vertebrates. *Nature* 361:129–134.
- Forey PL. 1995. Agnathans recent and fossil, and the origin of jawed vertebrates. *Rev Fish Biol Fisheries* 5:267–303.
- Furlong RF, Holland PWH. 2002. Bayesian phylogenetic analysis supports monophyly of ambulacraria and of cyclostomes. *Zool Sci* 19:593–599.
- Gauthier J, Kluge AG, Rowe T. 1988. Amniote phylogeny and the importance of fossils. *Cladistics* 4:105–209.
- Gess RW, Coates MI, Rubidge BS. 2006. A lamprey from the Devonian period of South Africa. *Nature* 443:981–984.
- Goldman N. 1993. Statistical tests of models of DNA substitution. *J Mol Evol* 36:182–198.
- Goodrich ES. 1909. *Vertebrata Craniata*. First fascicle: cyclostomes and fishes. In: Lankester RE, editor. *A treatise on zoology*, Vol. 9. London: A. & C. Black. p 1–518.
- Goodrich ES. 1930. *Studies on the structure and development of vertebrates*. London: Macmillan. 837p.
- Hardisty MW. 1979. *The biology of cyclostomes*. London: Chapman and Hall. 428p.
- Hardisty MW. 1982. Lampreys and hagfishes: analysis of cyclostome relationships. In: Hardisty MW, Potter IC, editors. *The biology of lampreys*. London: Academic Press. p 165–260.
- Hedges SB. 2001. Molecular evidence for the early history of living vertebrates. In: Ahlberg PE, editor. *Major events in early vertebrate evolution: paleontology, phylogeny, genetics and development*. London: Taylor & Francis. p 119–134.
- Hendy MD, Penny D. 1989. A framework for the quantitative study of evolutionary trees. *Syst Zool* 38:297–309.
- Hubbs CL, Potter IC. 1971. Distribution, phylogeny and taxonomy. In: Hardisty MW, Potter IC, editors. *The biology of lampreys*. London: Academic Press. p 1–65.
- Huelsenbeck JP. 1998. Systematic bias in phylogenetic analysis: is the Strepsiptera problem solved? *Syst Biol* 47:519–537.
- Huelsenbeck JP, Bull JJ, Cunningham CW. 1996a. Combining data in phylogenetic analysis. *Trends Ecol Evol* 11:152–158.
- Huelsenbeck JP, Hillis DM, Jones R. 1996b. Parametric bootstrapping in molecular phylogenetics: applications and performance. In: Ferraris JD, Palumbi SR, editors. *Molecular Zoology*. New York: Wiley-Liss. p 19–45.
- Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294:2310–2314.
- Huxley TH. 1876. On the nature of the cranio-facial apparatus of *Petromyzon*. *J Anat Physiol* 10:412–429.
- Janvier P. 1981. The phylogeny of Craniata, with particular reference to the significance of fossil “agnathans”. *J Vert Paleontol* 1:121–159.
- Janvier P. 1996a. The dawn of the vertebrates: characters versus common ascent in the rise of current vertebrate phylogenies. *Palaeontology* 39:259–287.
- Janvier P. 1996b. *Early vertebrates*. New York: Oxford University Press. 393p.
- Janvier P. 2007a. Born-again hagfishes. *Nature* 446:622–623.
- Janvier P. 2007b. Living primitive fishes and fishes from deep time. In: McKenzie DJ, Farrell AP, Brauner CJ, editors. *Primitive fishes*. San Diego: Elsevier. p 1–51.
- Janvier P. 2008. Early jawless vertebrates and cyclostome origins. *Zool Sci* 25:1045–1056.
- Janvier P, Blicek A. 1979. New data on the internal anatomy of Heterostraci (Agnatha) with general remarks on the phylogeny of Craniota. *Zool Scr* 8:287–296.
- Jarvik E. 1968. Aspects of vertebrate phylogeny. In: Orvig T, editor. *Current problems of lower vertebrate phylogeny*. New York: Interscience Publishers. p 497–527.
- Jollie M. 1962. *Chordate morphology*. New York: Reinhold Publishing Corporation. 478p.
- Kass RE, Raftery AE. 1995. Bayes factors. *J Am Stat Assoc* 90:773–795.
- Kikugawa K, Katoh K, Kuraku S, Sakurai H, Ishida O, Iwabe N, Miyata T. 2004. Basal jawed vertebrate phylogeny inferred from multiple nuclear DNA-coded genes. *BMC Biology* 2:1–11.
- Kuraku S. 2008. Insights into cyclostome phylogenomics: pre-2R or post-2R? *Zool Sci* 25:960–968.
- Kuraku S, Kuratani S. 2006. Time scale for cyclostome evolution inferred with a phylogenetic diagnosis of hagfish and lamprey cDNA sequences. *Zool Sci* 23:1053–1064.
- Kuraku S, Hoshiyama D, Katoh K, Suga H, Miyata T. 1999. Monophyly of lampreys and hagfishes supported by nuclear DNA-coded genes. *J Mol Evol* 49:729–735.
- Lanfranchi G, Pallavicini A, Laveder P, Valle G. 1994. Ancestral hemoglobin switching in lampreys. *Dev Biol* 164:402–408.
- Larget B, Simon DL. 1999. Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. *Mol Biol Evol* 16:750–759.
- Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst Biol* 50:913–925.
- Lipscomb DL, Farris JS, Källersjö M, Tehler A. 1998. Support, ribosomal sequences and the phylogeny of the eukaryotes. *Cladistics* 14:303–338.
- Løvtrup S. 1977. *The phylogeny of vertebrata*. New York: John Wiley & Sons. 330p.
- Maisey JG. 1986. Heads and tails: a chordate phylogeny. *Cladistics* 2:201–256.
- Mallatt J, Sullivan J. 1998. 28S and 18S rDNA sequences support the monophyly of lampreys and hagfishes. *Mol Biol Evol* 15:1706–1718.
- Mallatt J, Winchell CJ. 2007. Ribosomal RNA genes and deuterostome phylogeny revisited: more cyclostomes, elasmobranchs, reptiles, and a brittle star. *Mol Phylogenet Evol* 43:1005–1022.
- Mallatt J, Sullivan J, Winchell CJ. 2001. The relationship of lampreys to hagfishes: a spectral analysis of ribosomal DNA sequences. In: Ahlberg P, editor. *Major events in early vertebrate evolution*. London: Taylor & Francis. p 106–118.
- Meyer A, Zardoya R. 2003. Recent advances in the (molecular) phylogeny of vertebrates. *Annu Rev Ecol Evol Syst* 34:311–338.

- Mickevich MF, Farris JS. 1981. The implications of congruence in *Menidia*. *Syst Zool* 30:351–370.
- Müller J, Reisz RR. 2006. The Phylogeny of early eurentiles: comparing parsimony and Bayesian approaches in the investigation of a basal fossil clade. *Syst Biol* 55:503–511.
- Nylander JAA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey JL. 2004. Bayesian phylogenetic analysis of combined data. *Syst Biol* 53:47–67.
- Parker WK. 1883. On the skeleton of Marsipobranch fishes. *Philos Trans R Soc Lond B Biol Sci* 174:411–458.
- Patterson C, Williams DM, Humphries CJ. 1993. Congruence between molecular and morphological phylogenies. *Annu Rev Ecol Syst* 24:153–188.
- Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- Rambaut A, Grassly NC. 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comput Appl Bio Sci* 13:235–238.
- Riedl R. 1978. *Order in living organisms*. New York: Wiley.
- Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
- Shaeffer B, Thomson KS. 1980. Reflections on agnathan–gnathostome relationships. In: Jacobs LL, editor. *Aspects of vertebrate history*. Flagstaff: Museum of Northern Arizona Press. p 19–33.
- Shu D-G, Luo H-L, Morris SC, Zhang X-L, Hu S-X, Chen L, Han J, Zhu M, Li Y, Chen L-Z. 1999. Lower Cambrian vertebrates from South China. *Nature* 402:42–46.
- Simpson GG. 1944. *Tempo and mode in evolution*. New York: Columbia University Press. 237p.
- Simpson GG. 1953. *The major features of evolution*. New York: Columbia University Press. 434p.
- Stensiö E. 1968. The cyclostomes with special reference to the diphyletic origin of the Petromyzontida and Myxinoidea. In: Ørving T, editor. *Current problems of lower vertebrate phylogeny*. New York: Interscience Publishers. p 13–71.
- Stiassny MLJ, Wiley EO, Johnson GD, Carvalho MRd. 2004. Gnathostome fishes. In: Cracraft J, Donoghue MJ, editors. *Assembling the tree of life*. New York: Oxford University Press. p 410–429.
- Stock DW, Whitt GS. 1992. Evidence from 18S ribosomal RNA sequence that lampreys and hagfishes form a natural group. *Science* 257:787–789.
- Swofford DL. 2003. PAUP*. *Phylogenetic analysis using parsimony (*and other methods)*. Sunderland, MA: Sinauer Associates.
- Swofford DL, Waddell PJ, Huelsenbeck JP, Foster PG, Lewis PO, Rogers JS. 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Syst Biol* 50:525–539.
- Takezaki N, Figueroa F, Zaleska-Rutczynska Z, Klein J. 2003. Molecular phylogeny of early vertebrates: monophyly of the Agnathans as revealed by sequences of 35 genes. *Mol Biol Evol* 20:287–292.
- Yalden DW. 1985. Feeding mechanisms as evidence for cyclostome monophyly. *Zool J Linn Soc* 84:291–300.
- Yu SY, Zhang WW, Li L, Huang HF, Ma F, Li QW. 2008. Phylogenetic analysis of 48 gene families revealing relationships between hagfishes, lampreys, and Gnathostomata. *J Genet Genomics* 35:285–290.
- Zharkikh A, Li W-H. 1993. Inconsistency of the maximum-parsimony method: the case of five taxa with a molecular clock. *Syst Biol* 42:113–125.