Evaluating Hypotheses of Deuterostome Phylogeny and Chordate Evolution with New LSU and SSU Ribosomal DNA Data

Christopher J. Winchell,[1] *Jack Sullivan,*† *Christopher B. Cameron,*‡ *Billie J. Swalla,*§ *and Jon Mallatt**

*School of Biological Sciences, Washington State University; †Department of Biological Sciences, University of Idaho; ‡Marine Science Institute, Department of Marine Science, University of Texas at Austin; and §Department of Zoology, University of Washington

We investigated evolutionary relationships among deuterostome subgroups by obtaining nearly complete large-subunit ribosomal RNA (LSU rRNA)-gene sequences for 14 deuterostomes and 3 protostomes and complete small-subunit (SSU) rRNA-gene sequences for five of these animals. With the addition of previously published sequences, we compared 28 taxa using three different data sets (LSU only, SSU only, and combined LSU + SSU) under minimum evolution (with LogDet distances), maximum likelihood, and maximum parsimony optimality criteria. Additionally, we analyzed the combined LSU + SSU sequences with spectral analysis of LogDet distances, a technique that measures the amount of support and conflict within the data for every possible grouping of taxa. Overall, we found that (1) the LSU genes produced a tree very similar to the SSU gene tree, (2) adding LSU to SSU sequences strengthened the bootstrap support for many groups above the SSU-only values (e.g., hemichordates plus echinoderms as Ambulacraria; lancelets as the sister group to vertebrates), (3) LSU sequences did not support SSU-based hypotheses of pterobranchs evolving from enteropneusts and thaliaceans evolving from ascidians, and (4) the combined LSU + SSU data are ambiguous about the monophyly of chordates. No tree-building algorithm united urochordates conclusively with other chordates, although spectral analysis did so, providing our only evidence for chordate monophyly. With spectral analysis, we also evaluated several major hypotheses of deuterostome phylogeny that were constructed from morphological, embryological, and paleontological evidence. Our rRNA-gene analysis refutes most of these hypotheses and thus advocates a rethinking of chordate and vertebrate origins.

Introduction

The deuterostomes are a major clade of triploblastic metazoans and include three phyla: Echinodermata (sea stars, sea cucumbers, sea urchins, etc.), Hemichordata (acorn worms and pterobranchs), and Chordata (urochordates, lancelets, and vertebrates). Although this classification is widely accepted, some authors recognize three chordate phyla: Urochordata (or Tunicata) for the urochordates, Cephalochordata for the lancelets, and Vertebrata (or Craniata) for the vertebrates (Nielsen 1995; Giribet et al. 2000). It has also been suggested that Urochordata should be a separate phylum from vertebrates and cephalochordates (Cameron, Garey, and Swalla 2000). Apart from the deuterostomes, the other triploblastic metazoan lineage consists of the protostomes (arthropods, annelids, molluscs, etc.). Various differences in early embryology have historically distinguished protostomes from deuterostomes (Brusca and Brusca 1990), but the plasticity of these characters has recently been acknowledged, and their utility for determining evolutionary relationships has been questioned (Nielsen 1994; Valentine 1997). Regardless, molecular data sets have verified the monophyly of both groups (Giribet et al. 2000; Peterson and Eernisse 2001; Mallatt and Winchell 2002). Until recently, two enigmatic groups, the chaetognaths and the lophophorates, were considered to be deuterostomes but are now recognized as protostomes (see Telford and Holland 1993; Wada and Satoh 1994; Halanych 1996 for chaetognaths and Halanych et al. 1995; Mackey et al. 1996; de Rosa et al. 1999; Stechmann and Schlegel 1999 for lophophorates).

For over a century, biologists have debated the interrelationships among major taxa within the deuterostomes, focusing particularly on chordate and vertebrate ancestry (Gee 1996). A large number of studies based on anatomical, embryological, and paleontological data have generated contrasting hypotheses of deuterostome evolution, sustaining the controversy over chordate and vertebrate origins (see for example Gaskell 1890; Garstang 1928; Berrill 1955; Romer 1967; Jollie 1973; Gutmann 1981; Jefferies 1986; Maisey 1986; Schaeffer 1987; Jensen 1988; Swalla 2001).

Molecular studies of deuterostome phylogeny have used small-subunit ribosomal RNA (SSU rRNA), also called 18S rRNA, gene sequences. The results of Turbeville, Schulz, and Raff (1994) and Wada and Satoh (1994) showed that 18S data alone did not provide strong support for interrelationships among deuterostome subgroups and thus could not reveal chordate monophyly. However, utilizing the total-evidence approach (see review by de Queiroz, Donoghue, and Kim 1995), Turbeville, Schulz, and Raff (1994) combined their 18S data matrix with 11 morphological traits and an additional molecular character and found some evidence for chordate monophyly (61% bootstrap support).

In the first thorough analysis of metazoan phylogeny incorporating total evidence, Zrzavy et al. (1998) compared 18S data and 276 morphological characters across nearly all animal phyla. Through a sensitivity

[1] Present address: Department of Organismic Biology, Ecology, and Evolution, University of California, Los Angeles.

Key words: deuterostome, chordate, phylogeny, 28S rRNA, SSU, LSU, spectral analysis.

Address for correspondence and reprints: Jon Mallatt, P.O. Box 644236, School of Biological Sciences, Washington State University, Pullman, Washington 99164-4236. E-mail: jmallatt@mail.wsu.edu.

Mol. Biol. Evol. 19(5):762–776. 2002

analysis of these data, they inferred a monophyletic Deuterostomia, Chordata, and Ambulacraria (echinoderms + hemichordates; Metschnikoff 1881). Giribet et al. (2000) expanded on the study of Zrzavy et al. (1998) by adding many new 18S sequences, using only triploblast animals (to avoid the confounding effects of too distant and long-branched outgroups), and employing unique analyses. They also recovered monophyletic deuterostome, chordate, and Ambulacrarian clades, but within the chordates, their results supported a nontraditional pairing of urochordates and vertebrates.

Cameron, Garey, and Swalla (2000) added many new hemichordate and urochordate taxa to the SSU rRNA-gene database and found strongly supported relationships among the major deuterostome groups. Their analyses utilized species with slow rates of sequence substitution to minimize taxonomic artifacts caused by long-branch attraction. For some analyses, they used a phylogenetic search algorithm, Gambit, designed to be insensitive to unequal rate effects and able to accommodate an evolutionary model accounting for site-to-site variation in substitution rate (Lake 1995). Despite finding ample support for a lancelet + vertebrate clade and an Ambulacrarian clade, a monophyletic Chordata was not supported by most of their analyses. However, by using Gambit and restricting their comparison to the 16 deuterostomes and 1 protostome with the slowest rates of SSU rDNA evolution, Cameron, Garey, and Swalla (2000) did find evidence for a monophyletic Chordata (85% bootstrap support). Additionally, their SSU data suggested that within the hemichordates, the pterobranchs might have evolved from an enteropneust (acorn worm) ancestor, thus supporting the notion of acorn worm paraphyly, a topology first shown by Halanych (1995).

In the current study, we continue the investigation of deuterostome interrelationships by adding different rRNA genes. Mallatt and Sullivan (1998) and Mallatt, Sullivan, and Winchell (2001) have demonstrated the usefulness of chordate large-subunit ribosomal RNA (LSU rRNA)-gene sequences for deep-level phylogenetics, especially when combined with SSU-gene sequences. The large subunit of the metazoan ribosome contains three RNA molecules (5S, 5.8S, 28S), whereas the 18S rRNA molecule is the only one incorporated into the small ribosomal subunit (Alberts et al. 1994, p. 379). Here, we have obtained nearly complete 28S sequences for 14 deuterostomes and 3 protostomes, complete 18S sequences for five of these taxa, and we sequenced all or part of the 5.8S rRNA gene for many of them. By including GenBank sequences from 11 additional animals, we employed the largest rRNA-gene data set to date (consisting of 18S + 5.8S + 28S rRNA-gene sequences, averaging about 4,000 aligned sites per taxon) for estimating the phylogeny of deuterostomes. We constructed these relationships with minimum evolution (ME) (using LogDet distances), maximum parsimony (MP), and maximum likelihood (ML) analyses. Nonparametric bootstrapping (Felsenstein 1985) and spectral analysis (Lento et al. 1995; Penny et al. 1999) provided examination of nodal support. We used our results to evaluate various classical hypotheses of deuterostome evolution.

Materials and Methods

Specimens and Sequences

A table listing the sources of the animals, the number of individuals sequenced, new GenBank accession numbers, and sequences taken from the literature is available online through the journal (http://www.smbe.org). We compared sequences of 10 vertebrates: *Petromyzon marinus* (lamprey), *Eptratretus stouti* (hagfish), *Hydrolagus colliei* (ratfish), *Squalus acanthias* and *Triakis semifasciata* (sharks), *Raja schmidti* (skate), *Acipenser brevirostrum* (sturgeon), *Latimeria chalumnae* (coelacanth), *Xenopus laevis* (frog), *Ambystoma macrodactylum* (salamander); the cephalochordate *Branchiostoma floridae*; four urochordates: *Oikopleura* sp. (larvacean), *Thalia democratica* (thaliacean), *Styela plicata*, and *Ciona intestinalis* (ascidians); six hemichordates: *Ptychodera flava* and *Ptychodera* sp. (ptychoderid acorn worms), *Saccoglossus kowalevskii*, *Saccoglossus* sp. *CC-03-2000*, and *Harrimania* sp. *CC-03-2000* (harrimaniid acorn worms; Cameron 2000), *Cephalodiscus gracilis* (pterobranch); four echinoderms: *Florometra serratissima* (crinoid), *Asterias forbesi* (sea star), *Strogylocentrotus purpuratus* (sea urchin), *Cucumaria salma* (sea cucumber); and three protostomes: *Limulus polyphemus* (Horseshoe crab), *Eisenia fetida* (earthworm), *Proceraea cornuta* (polychaete worm). Except for a few difficult-to-obtain taxa, where only one individual was sequenced, every sequence came from at least two individuals. We obtained 99%-complete 28S sequences (all but the last ~41 nucleotides from the 3′ end) from *Triakis*, *Raja*, *Florometra*, *Asterias*, *Saccoglossus* sp. *CC-03-2000*, *S. kowalevskii*, *P. flava*, *Proceraea*, *Eisenia*, and *Limulus*; and 92%-complete sequences (all but the last ~300 nucleotides at the 3′ end) from *Ambystoma*, *Cucumaria*, *Strongylocentrotus*, *Ptychodera* sp., and *Harrimania* sp. *CC-03-2000*. *Ciona* is missing the first 24 and last ~41 nucleotides (98% complete) and *Cephalodiscus* is missing the last ~70 nucleotides (98% complete). We also obtained complete 5.8S rRNA-gene sequences for *Florometra*, *Cucumaria*, *Asterias*, *Saccoglossus* sp. *CC-03-2000*, *S. kowalevskii*, *Proceraea*, and *Limulus* and determined partial 5.8S sequences (ranging from 45 to 75 nucleotides of the 3′ end of the gene) for *P. flava*, *Harrimania*, and *Cephalodiscus*. Finally, we sequenced complete 18S genes for *P. flava*, *Ptychodera* sp., *R. schmidti*, *T. semifasciata*, and *P. cornuta*.

It should be noted that in a few cases, the SSU and LSU genes were from different species within a genus: 28S from *A. macrodactylum* and 18S from *A. mexicanum*; 28S from *A. forbesii* and 18S from *A. amurensis*; 28S from *C. salma* and 18S from *C. sykion*. However, rRNA genes evolve slowly enough that there should be only minimal differences between organisms in the same genus (Hillis and Dixon 1991). Also, we combined the 28S rRNA gene of the crinoid *F. serratissima* with the 18S gene of *Antedon serrata*, which is in a different

but closely related genus of the same family (Hyman 1955, pp. 95–97). This is justifiable because the first 371 bases of the 28S gene of *Antedon* are known (GenBank accession AJ225818; Lafay, Smith, and Christen 1995), and they are 96% similar to our *Florometra* sequence, suggesting that the entire 28S genes of these two genera are very similar.

Genomic DNA was extracted from tissue preserved in 70%–95% ethanol, using the cetyltrimethylammonium bromide method (Winnepenninckx, Backeljau, and De Wachter 1993). Tissue sources included: vertebrate tail fin clips; urochordate pharynx and gonad; acorn worm proboscis muscle and skin; entire *Cephalodiscus* animals; echinoderm muscle, gonads, and tube feet; and *Eisenia* cuticle, muscle, and pharynx. For the small polychaete, *Proceraea* (about 1 mm in length), we obtained DNA by grinding a whole animal in 15 μl of TE buffer (pH 8.0) and included 1 μl of this slurry in the PCRs. We used this same approach with *Limulus* book-gill tissue. We performed DNA amplification, purification, sequencing, and fragment assembly, as described in Mallatt and Sullivan (1998). In some animals, we could only amplify 28S genes in smaller segments; occasionally, when sequencing such segments, we found the overlap between them to alternate between matching perfectly and showing minor divergences. Such discordant segments were obtained in *Triakis*, *Asterias*, and *Cucumaria* and must represent parts of alternate forms of the 28S genes or even pseudogenes. Wada (1998) also experienced this phenomenon with the urochordate, *Doliolum nationalis*. Resequencing from larger amplified segments allowed us to recognize and eliminate the pseudogenes in all cases. For amplifying the 18S genes (previously unreported), we used the same procedures but with primers 18e 5′-CTGGTTGATCCTGCCAGT-3′ (Hillis and Dixon 1991) and 18P-C 5′-TAATGATCCTTCCGCAGGTTCACCT-3′ (K. Halanych, personal communication). For sequencing, we used additional 18S primers: 18Q-C 5′-GTTATCGGAATTAACCAGACA-3′ and its complement 18Q-S (K. Halanych, personal communication); 18h 5′-AGGGTTCGATTCCGGAGAGGGAGC-3′ and its complement 18i; 18j 5′-GCCTGCGGCTTAATTTGACTCAACACGGG-3′ and its complement 18k (Hillis and Dixon 1991).

We imported the rRNA-gene sequences into SeqLab, a Macintosh X Window application (see Smith et al. 1994). We aligned concatenated 28S, 5.8S, and 18S rRNA genes entirely by eye, using, as guides, the LSU rRNA secondary structure of *X. laevis* (Schnare et al. 1996) and the SSU secondary structures of *X. laevis* and *S. purpuratus* (Gutell 1994).

Two characteristic types of regions exist within the eukaryote 28S rRNA genes: conserved regions, which have a slow evolutionary rate of nucleotide substitution and are collectively called the conserved core of the molecule; and divergent domains, which evolve at a much higher rate (Hassouna, Michot, and Bachellerie 1984). The 28S genes contain 12 divergent domains, which comprise about one-third of the molecule; see Mallatt, Sullivan, and Winchell (2001) for the locations of and sequences immediately before and after all divergent domains in *X. laevis*, *B. floridae*, and *S. plicata*. Because it is difficult to align the divergent domains of distantly related taxa (such as those compared here), we excluded these variable regions from our analyses and restricted our 28S comparisons to the core regions only, which are easily and unambiguously aligned by eye.

Overall, we compared 1,635, 151, and 2,338 aligned sites in the 18S, 5.8S, and 28S genes, representing about 90%, 90%, and 65% of the total length of these genes, respectively, and a total of about 4,000 bases. Our alignments are deposited in the EMBL Nucleotide Sequence Database (http://www.ebi.ac.uk/embl/index.html) under the accession numbers ALIGN_000057 (LSU genes) and ALIGN_000058 (SSU genes).

Phylogenetic Analyses

Tree Estimation

We used three data sets to conduct our analyses: (1) LSU genes only, (2) SSU genes only, and (3) combined LSU + SSU genes. The issue of whether to analyze multiple genes separately or combined is complex, and no single test can provide utterly reliable guidance (e.g., Sullivan 1996). Although the incongruence length difference test (Farris et al. 1994; Cunningham 1997) suggested some degree of incongruence between the LSU and SSU genes ($P = 0.052$), the phylogenetic signal in these data sets may well be additive. Thus, we analyzed the data sets both separately and combined. The search strategies employed for inferring optimal trees included equally weighted MP, ML (Felsenstein 1981), and ME using LogDet-Paralinear distances (Lake 1994; Lockhart et al. 1994). We implemented each of these optimality criteria with PAUP*, Version 4.0 beta 4a (Swofford 1998). We used ML as in our previous studies (Mallatt and Sullivan 1998; Mallatt, Sullivan, and Winchell 2001): first, we loaded a starting tree (constructed with MP) into PAUP and then calculated the likelihood score of this tree under 16 different models of nucleotide substitution, as well as the parameters specifying each model. The most general of these models, GTR + I + Γ, described our data best in all cases (results not shown). We then followed the iterative search strategy outlined by Swofford et al. (1996) to obtain the optimal ML tree. Basically, using the GTR + I + Γ model parameters of the MP starting tree, we performed an ML search which found a tree with a better likelihood score. We then calculated parameters of this first ML tree under the same model and incorporated these values into a second ML search. For each data set, the first ML tree was the best because it was found again after one iteration.

ML and MP methods assume stationary nucleotide frequencies across all aligned sequences. If unrelated taxa converge in their nucleotide composition, these methods may incorrectly join them on a tree, despite their different ancestry (Lake 1994; Lockhart et al. 1994). Because our sequences did exhibit nonstationary base frequencies (see *Results*), we used the ME method based on the LogDet (Lockhart et al. 1994), or Paralinear (Lake 1994), data transformation. This technique

was specifically designed to produce additive genetic distances from sequences that differ in base composition by assuming a very general evolutionary model that allows each of the 12 nonreversible nucleotide substitution types to occur at a different rate. However, unlike ML, this method does not account for different rates of evolution across nucleotide sites. Nonetheless, a good way of approximating this rate heterogeneity, and accounting for functional constraint in these rRNA molecules, is to use an invariable sites model (e.g., Waddell, Penny, and Moore 1997; Mallatt, Sullivan, and Winchell 2001). To this end, we estimated the proportion of invariable sites (P_{inv}) across taxa (for each of our data sets) via ML using the GTR + I model. However, as mentioned previously, our data violate at least one of the ML assumptions; so we used a range of P_{inv} values from 0.15 to 0.6067 (for our combined-gene data set only) to gauge the sensitivity of our findings to variation in this P_{inv} parameter. The upper limit of this range (0.6067) is the highest possible value for P_{inv} because it represents the actual percentage of sites observed to be constant across all taxa in our combined LSU + SSU data set. Because our taxa so clearly have nonstationary nucleotide frequencies—which violate the assumptions of MP and ML—we emphasize the ME method using LogDet distances in this study.

To measure nodal support on our ME, ML, and MP trees, we performed nonparametric bootstrap analyses (Felsenstein 1985) with 1,000 replicates for all MP and ME searches and 100 replicates for all ML searches.

Spectral Analysis

Because bootstrapping only assesses support for clades and does not directly report any information regarding contradictory signal for a certain grouping (i.e., conflict), we also performed spectral analyses (e.g., Lento et al. 1995; Penny et al. 1999) on the calculated LogDet distances, in order to assess relative amounts of support and conflict for competing hypotheses of deuterostome phylogeny. These analyses were performed on our combined LSU + SSU-gene data set using the Spectrum program (Charleston 1998, http://taxonomy.zoology.gla.ac.uk/~mac/spectrum/spectrum), with the P_{inv} set to the ML estimate of 0.5948.

Spectral analysis is a quantitative method that is well suited for testing alternative phylogenetic hypotheses when relationships among taxa are controversial. With the ability to use either sequence data or distance matrices, the Spectrum program creates a complete array of bipartitions, or splits, in the data, which includes every possible grouping among the taxa used. Therefore, a strength of this method is that it works independently of any one particular tree. We displayed the calculated support and conflict values as a histogram called a Lento plot (Lento et al. 1995), which provides a visual representation of the phylogenetic spectrum of bipartitions (groupings of taxa) and facilitates comparison of competing hypotheses of evolutionary history. Support values were computed using a threshold value of 0.0003, in order to limit the number of groups returned.

In Spectrum, the conflict for a given bipartition is measured as the sum of support values for all bipartitions that contradict the given split. Because there may be many alternatives to a given bipartition, its conflict value may seem so large as to overwhelm its level of support. Therefore, we standardized all conflict values following Lento et al. (1995).

Because the Spectrum program has a limit of 20 taxa, we had to delete 8 of the 28 taxa that were used in the tree-building analyses. We trimmed taxa from the more heavily sampled groups of deuterostomes, retaining taxa that provided essentially the same ME bootstrap support values and tree topology as the 28-taxa data set. Specifically, we excluded five gnathostomes (*Xenopus*, *Latimeria*, *Acipenser*, *Raja*, *Squalus*) and three hemichordates (*Saccoglossus* sp. *CC-03-00*, *Harrimania*, *Ptychodera* sp.) from the spectral analysis.

Results

As seen in table 1, the rRNA genes of some taxa examined have relatively high proportions of C and G nucleotides: the pterobranch *Cephalodiscus* and hagfish *Eptatretus*. Other taxa have high A and T levels: the urochordates (*Ciona*, *Thalia*, *Styela*, and *Oikopleura*—which also have low levels of C) and the protostomes *Eisenia* and *Proceraea*. These differences in nucleotide proportions across taxa resulted in a strongly rejected ($P \ll 0.001$) χ^2 test of stationary nucleotide frequencies. Although this test ignores correlation caused by common ancestry, we assume that such a strong effect reflects a real feature of the data set. By analyzing the SSU and LSU sequences separately, we determined that the nonstationarity of nucleotide frequencies is almost exclusively in the LSU genes; the 18S nucleotide frequencies were stationary (results not shown).

Figure 1 displays ME trees based on LogDet distances. The two types of rRNA gene used in this study, SSU and LSU, produced broadly similar trees (compare fig. 1*A* and *B*). The LSU data (fig. 1*A*) place the urochordates as the sister group to the remaining deuterostomes and not with other chordates. However, this relationship is poorly supported because the ME bootstrap value for this particular node is only 32% and is thus not displayed on the tree. On the other hand, the SSU genes grouped urochordates with lancelets and vertebrates as a monophyletic Chordata but mainly with the LogDet method (70% ME bootstrap support: fig. 1*B*). LSU genes did not definitively place the lancelet, *Branchiostoma*, with any particular group, but the SSU genes showed consistently strong support for a lancelet + vertebrate clade. Vertebrate monophyly was strongly supported by both genes, as were the following groups within the vertebrates: cyclostomes, gnathostomes, and Chondrichthyes. Resolution within the gnathostomes was weak for both genes, especially LSU, which reversed the expected positions of *Acipenser* and *Ambystoma*. Within the urochordates, both genes placed *Oikopleura* (class Larvacea) as the sister group to all other urochordates. However, relationships among the three other urochordates were inconsistent: the SSU genes in-

Table 1
Proportions of Different Nucleotide Types in the Combined SSU and LSU rRNA Genes of Taxa Used in this Study (includes both stems and loops, but the 28S divergent domains are excluded). The Number of Sites Used (18S + 28S core + partial 5.8S) is Listed at Far Right

Taxon	A	C	G	T	# Sites
Eptatretus	0.2343	0.2503	0.3161	0.1994	4,012
Petromyzon	0.2465	0.2334	0.3028	0.2172	4,121
Xenopus	0.2502	0.2344	0.2995	0.2159	4,117
Ambystoma	0.2419	0.2199	0.2852	0.2122	3,902
Acipenser	0.2553	0.2255	0.2948	0.2244	3,819
Latimeria	0.2535	0.2245	0.2986	0.2235	3,835
Hydrolagus	0.2536	0.2307	0.2961	0.2193	4,113
Triakis	0.2553	0.2296	0.2947	0.2199	4,007
Squalus	0.2583	0.2263	0.2935	0.2216	3,867
Raja	0.2552	0.2290	0.2949	0.2208	4,008
Branchiostoma	0.2519	0.2325	0.2963	0.2185	4,009
Florometra	0.2539	0.2213	0.2861	0.2290	4,030
Cucumaria	0.2527	0.2402	0.2905	0.2156	4,017
Strongylocentrotus	0.2514	0.2317	0.2931	0.2145	3,954
Asterias	0.2428	0.2398	0.3068	0.2107	4,016
Saccoglossus kowalevskii	0.2465	0.2421	0.3019	0.2095	4,081
Saccoglossus sp. *CC-03-2000*	0.2452	0.2426	0.3038	0.2085	4,118
Harrimania sp. *CC-03-2000*	0.2466	0.2428	0.3039	0.2067	3,958
Ptychodera flava	0.2535	0.2326	0.2959	0.2181	4,012
Ptychodera sp.	0.2555	0.2314	0.2930	0.2202	3,946
Cephalodiscus	0.2380	0.2518	0.3148	0.1908	3,984
Styela	0.2588	0.2228	0.2911	0.2272	4,115
Ciona	0.2622	0.2143	0.2920	0.2315	3,901
Thalia	0.2622	0.2177	0.2885	0.2316	4,024
Oikopleura	0.2797	0.1969	0.2721	0.2513	3,947
Proceraea	0.2586	0.2237	0.2863	0.2314	4,118
Eisenia	0.2666	0.2195	0.2827	0.2312	3,923
Limulus	0.2603	0.2270	0.2880	0.2248	4,115
Mean	0.2534	0.2291	0.2948	0.21989	3,999

NOTE.—χ^2 test of homogeneity of base frequencies: $\chi^2 = 187.35$ (degrees of freedom = 81), $P < 0.00000001$.

dicated paraphyly of the class Ascidiacea by placing the ascidian of the suborder Phlebobranchiata (*Ciona*) with *Thalia* (class Thaliacea); the LSU genes, in contrast, found weak support for ascidian monophyly.

Both genes supported monophyly of the echinoderms and the hemichordates and united these two phyla as Ambulacrarians; the LSU data provided the stronger support for this relationship. Within the echinoderms, the LSU genes strongly supported the crinoid, *Florometra*, as sister taxon to the other echinoderm classes, but the SSU genes showed very weak support for relationships at this level, and both genes paired the sea cucumber, *Cucumaria*, with the sea urchin, *Strongylocentrotus*. The two genes are clearly in conflict with respect to relationships within the hemichordates, the difference concerning the placement of *Cephalodiscus*. The LSU data placed this pterobranch as the sister group to the class Enteropneusta (the acorn worms), thus supporting monophyly of the acorn worms. In contrast, the SSU data supported enteropneust paraphyly by placing acorn worms of the family Ptychoderidae as the sister to a group consisting of pterobranchs + acorn worms of the family Harrimaniidae.

The combined LSU and SSU tree is shown in figure 1*C* and can be compared to the tree based on the SSU gene alone (fig. 1*B*). This comparison indicates that adding LSU data increases the bootstrap support for the following groups, mostly to around 100%: lancelets + vertebrates; Ambulacraria; hemichordates; cyclostomes; Chondrichthyes; elasmobranchs; non–larvacean urochordates; non–crinoid echinoderms; and sea cucumber *Cucumaria* + sea urchin *Stongylocentrotus*. On the other hand, addition of LSU data slightly lowered the bootstrap support for the group of *Cephalodiscus* + harrimaniid enteropneusts (compare fig. 1*C* with *B*). Support for most other groups stayed about the same.

Unlike the 70% bootstrap support shown for chordates in the SSU tree, the combined-gene tree failed to support a monophyletic Chordata. Specifically, the optimal ME tree of figure 1*C* placed urochordates as the outgroup to all other deuterostomes, thus displaying an ([echinoderm, hemichordate], [lancelet, vertebrate]) subtree. The ME bootstrap value for this particular node was only 47%, whereas the bootstrap value for the alternate clade showing a monophyletic Chordata is slightly higher, 49.5% (not illustrated). Clearly, the support for either of these groupings is weak and indistinguishable. In other words, these tree-based analyses effectively show a three-group polytomy at the base of the deuterostomes: urochordates; Ambulacraria; and lancelets + vertebrates.

Table 2 shows how changing the proportion of invariable sites (P_{inv}) in the combined data set affects ME bootstrap support for eight splits. Although support for some groups remained high over the entire range of P_{inv} (>75% support: deuterostomes, Ambulacraria, lancelet

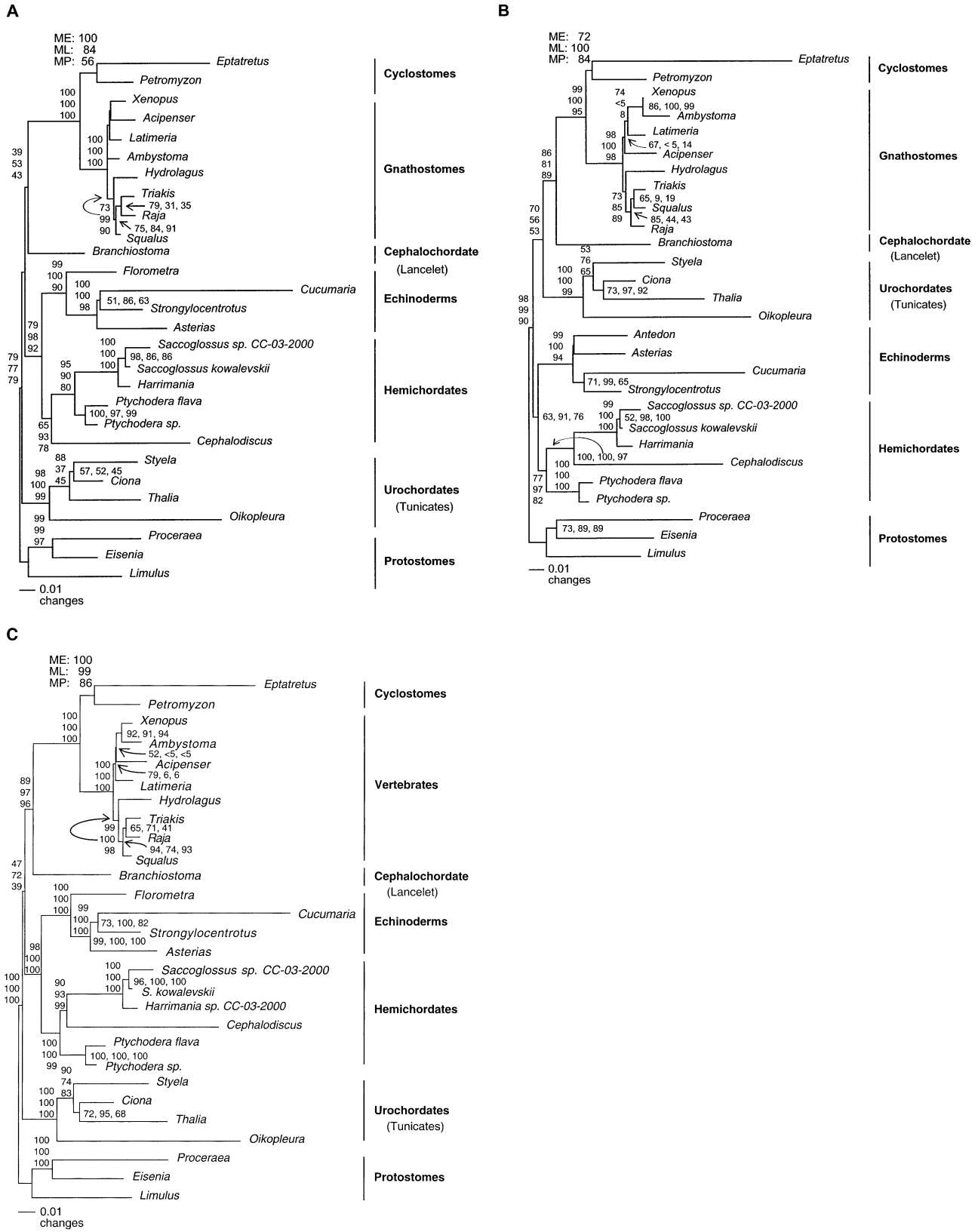

FIG. 1.—Phylogenetic trees of 25 deuterostomes and three protostomes based on rRNA genes. These trees were calculated with the ME optimization method using LogDet distances (with the proportion of invariable sites, P_{inv}, set at likelihood-calculated values). Trees were constructed with: (*A*) LSU (28S and partial 5.8S) rRNA-gene sequences and $P_{inv} = 0.6043$, (*B*) SSU (18S) rRNA-gene sequences and $P_{inv} = 0.5632$, and (*C*) a combined data set of LSU + SSU sequences and $P_{inv} = 0.5883$. The percentage of bootstrap replicates supporting each node is shown and represents, from top to bottom (or left to right), ME-, ML-, and MP-based values. Nodes supported less than 50% of the time by all three methods show no values.

Table 2
Bootstrap Values for Eight Key Groups Computed Using LogDet Distances of LSU + SSU Data Over a Range of Invariable Sites (P_{inv})

P_{inv}	Deutero-stomes (%)	Urochordates + Protostomes (%)	Chordates (%)	Echinoderms + Hemichordates (%)	Lancelet + Vertebrates (%)	Cyclo-stomes (%)	Pterobranch + Harrimaniid Acorn Worms (%)	Sea Cucumber + Sea Urchin (%)
0.15	100	82	16	100	96	96	75	<5
0.25	100	80	18	100	96	96	76	6
0.35	100	78	20	99	96	97	77	10
0.45	100	71	28	99	95	99	80	20
0.55	100	57	41	99	92	99	86	53
0.5883[a]	100	47	49.5	98	89	100	90	73
0.6067[b]	99	41	56	98	83	100	91	85

NOTE.—The higher P_{inv} values (0.5883, 0.6067) are most likely to be correct (see text).
[a] P_{inv} estimated with the GTR + I maximum likelihood model.
[b] Actual proportion of constant sites.

+ vertebrates, cyclostomes, pterobranch + harrimaniids), the bootstrap support for other groups changed markedly with P_{inv} (urochordates + protostomes, chordates, sea cucumber + sea urchin). Especially noteworthy is the increase in support for monophyletic chordates and the decrease in support for urochordates + protostomes, at the higher P_{inv} values.

The 20 taxa chosen for spectral analysis, and the ME tree calculated for this subset of taxa, are shown in figure 2*A*. The results of the spectral analysis are shown in figure 2*B*, and the support and conflict values from this analysis are listed in table 3. In the Lento plot (fig. 2*B*), splits are analogous to nodes on a tree that join two or more taxa and partition the entire tree into two groups: the group joined by that node and the rest of the tree. Of the 20 most highly supported splits in the spectrum (splits A–T in fig. 2*B*), 15 were also present in both the trimmed and untrimmed ME trees (the black bars in fig. 2*B*) and are therefore unambiguous. The other five splits—M, N, O, Q, and R—bear further consideration. Split N (lamprey + gnathostome) is not favored because it had less support and more conflict than its alternative, split J (monophyletic cyclostomes). Split R, in contrast, is favored and is especially noteworthy: it is the first evidence for chordate monophyly in the present study, with much more support than the nearest contradictory split, which is labeled Urochordates + Protostomes in figure 2*B* (both splits, however, have roughly the same amount of conflict). Split Q weakly favors ascidian monophyly over the alternative of *Thalia* + *Ciona* (split T), and split M seems to favor enteropneust monophyly over enteropneust paraphyly (split O: *Cephalodiscus* + *Saccoglossus*)—both in contrast to the tree-based findings of figure 1*C*.

Of these results, the evidence for enteropneust monophyly (split M in fig. 2*B*) might be a trimming artifact because our taxon-trimmed tree (fig. 2*A*) joined the two enteropneust families, whereas the original 28-taxa tree did not (fig. 1*C*). To test for the possibility of such an artifact, we retrimmed the taxa in multiple ways that always retained all six hemichordate species. However, no matter how the taxa were trimmed, the hemichordate sequences always contained considerable support for both hypotheses. This means that spectral analysis, unlike the tree-based analyses, does not unite the pterobranchs with harrimaniid enteropneusts.

Figure 3 shows the optimal tree that was inferred from the results of our spectral analysis under the manhattan distance criterion (Charleston 1998). The nodes are labeled with letters corresponding to splits in figure 2*B*. As a method of measuring confidence for the clades on the manhattan tree (fig. 3), we compared their spectral analysis support levels to those of alternative clades resulting from nearest-neighbor interchanges (fig. 4). That is, every internal edge on the manhattan tree was rearranged, and spectral analysis support values for the three possible topologies were compared (see fig. 5 for an example). As shown in figure 4, the arrangement of groups congruent with the manhattan tree always had the highest support—usually overwhelmingly. However, splits J (cyclostomes), M (acorn worms), and Q (ascidian urochordates) have alternative pairings with significant support.

Discussion

Significance

This is one of the first studies to use nearly the entire rRNA gene family for deep-level animal phylogeny. We assembled a large data set (about 4,000 bases of sequence data from 28 taxa were compared) to explore the evolutionary relationships among deuterostomes. We added new LSU sequences to the existing SSU database (our own SSU tree in fig. 1*B* is essentially identical to the most recent SSU trees in the literature; see Wada 1998; Winnepenninckx, Van De Peer, and Backeljau 1998; Cameron, Garey, and Swalla 2000; Swalla et al. 2000). We obtained LSU sequences from at least one representative of each major subgroup of each phylum, so taxonomic sampling should be sufficient to indicate relationships between phyla; however, the relationships within phyla are not as certain and will benefit from the addition of more taxa in the future. Given this caveat, the new findings of this study are:

1. The LSU genes produced a tree very similar to the SSU tree (compare fig. 1*A* and *B*).

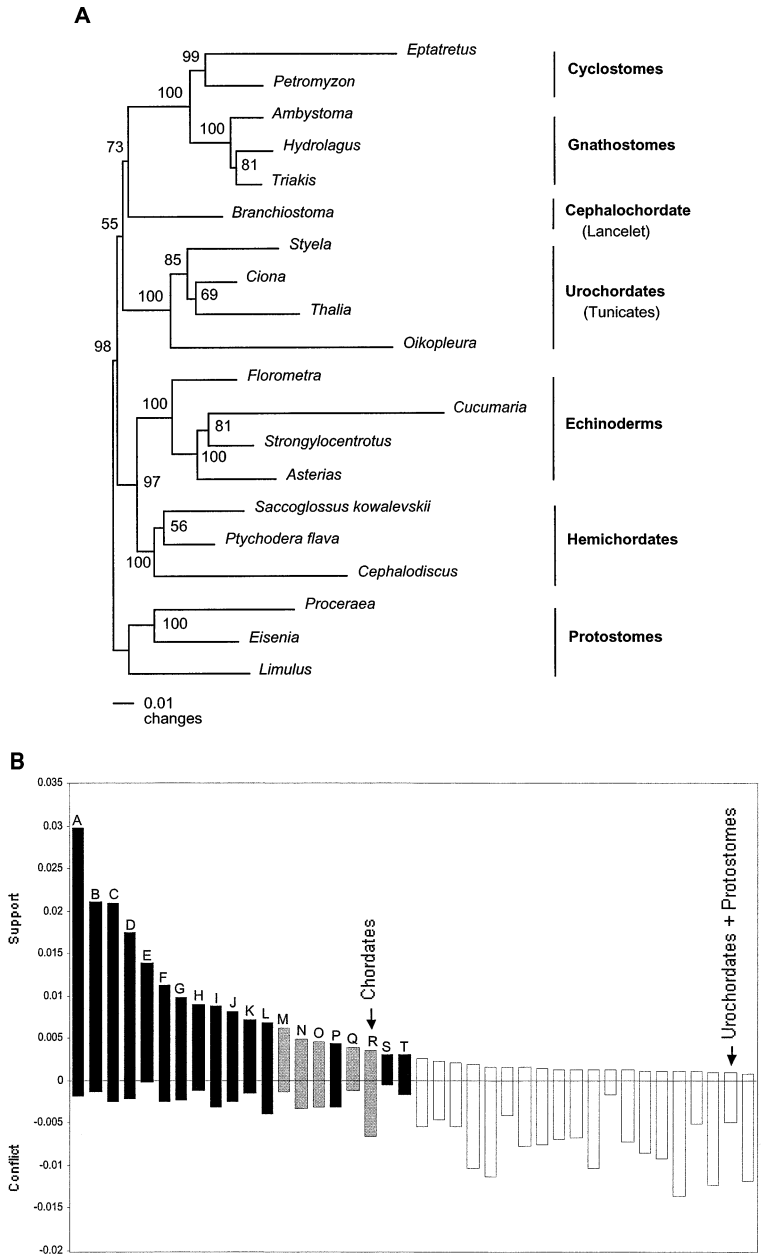

FIG. 2.—Spectral analysis. (*A*) Optimal ME tree for the 20 taxa chosen for spectral analysis. This tree was generated from combined LSU + SSU sequences, P_{inv} = 0.5948. Numbers at nodes are bootstrap support values (1,000 replicates; compare this with fig. 1*C*). (*B*) Lento plot displaying support and conflict values for the 40 most supported splits in the phylogenetic spectrum computed from LSU + SSU LogDet distances among the 20 taxa in part 2*A*, same P_{inv}. The units of the support and conflict values for given splits are numbers of nucleotide changes per site. The conflict values (i.e., the values below zero) for a given split were calculated by summing the support values of splits contradictory to the given split. Black bars designate splits also present in both the untrimmed and trimmed optimal ME trees (figs. 1*C* and 2*A*, respectively). Gray bars denote controversial groupings of two types: (1) those splits occurring in either the untrimmed or trimmed optimal ME trees, but not both; and (2) splits that are absent from both of these trees. White bars represent splits that neither occur in optimal ME trees nor are favored by spectral analyses of the combined data set; however, many of the competing hypotheses of deuterostome phylogeny are found in this category of splits. Note that split R represents monophyletic chordates and is to be compared with the split labeled Urochordates + Protostomes at far right, which was the next best alternative grouping for the urochordates. Code explaining the lettering of the splits: A = vertebrates, B = gnathostomes, C = urochordates, D = echinoderms, E = *Proceraea* + *Eisenia*, F = non–crinoid echinoderms, G = hemichordates, H = non–larvacean urochordates, I = deuterostomes, J = cyclostomes, K = *Cucumaria* + *Strongylocentrotus*, L = hemichordates + echinoderms, M = enteropneusts, N = lamprey + gnathostomes, O = *Saccoglossus* + *Cephalodiscus*, P = vertebrates + lancelet, Q = ascidian urochordates, R = chordates, S = chondrichthyans, T = *Ciona* + *Thalia*.

2. Adding LSU to SSU sequences (fig. 1*C*) strengthens the bootstrap support for many groups above the SSU-only values: e.g., lancelets + vertebrates; Ambulacraria; cyclostomes; Chondrichthyes; elasmobranchs, non–crinoid echinoderms, and sea urchin + sea cucumber. The validity of all these groups is independently supported by classical morphological or independent molecular evidence (lancelets + verte-

Table 3
Support (S) and Conflict (C) Values for the Preferred Splits and Their Best Alternatives Found by Spectral Analysis of LogDet Distances of the Combined LSU + SSU Data Set (the 20 taxa of fig. 2 were used)

Preferred Split (See fig. 2*B*)	Best Alternative Hypothesis by Nearest Neighbor Interchange (See fig. 4)
1. Vertebrates (A)	1. Cyclostomes + lancelet (A1)
S = 0.0298	S = 0.00027[a]
C = −0.0018	
2. Gnathostomes (B)	2. Cyclostomes + Chondrichthyes (B1)
S = 0.0211	S = 0.0012
C = −0.0014	C = −0.0072
3. Urochordates (C)	3. Non–larvacean urochordates + (lancelet + vertebrates) (C2)
S = 0.0209	S = 0.00002[a]
C = −0.0026	
4. Echinoderms (D)	4. Crinoid + hemichordates (D2)
S = 0.0176	S = 0.0003
C = −0.0022	C = −0.0091
5. *Proceraea* + *Eisenia* (E)	5. *Proceraea* + *Limulus* (E1)
S = 0.0139	S = 0.0003
C = −0.00007	C = −0.0060
6. Non–crinoid echinoderms (F)	6. *Florometra* + *Asterias* (F1)
S = 0.0112	S = 0.0016
C = −0.0024	C = −0.0041
7. Hemichordates (G)	7. Echinoderms + enteropneusts (G2)
S = 0.0099	S = 0.0013
C = −0.0023	C = −0.0069
8. Non–larvacean urochordates (H)	8. *Thalia* + *Oikopleura* (H1)
S = 0.0090	S = 0.0022
C = −0.0012	C = −0.0047
9. Deuterostomes versus Protostomes (I)	9. Ambulacraria + *Limulus* (I1)
S = 0.0087	S = 0.0016
C = −0.0032	C = −0.0077
10. Cyclostomes (J)	10. *Petromyzon* + gnathostomes (J1 = N)
S = 0.0081	S = 0.0048
C = −0.0025	C = −0.0032
11. *Cucumaria* + *Strongylocentrotus* (K)	11. *Strongylocentrotus* + *Asterias* (K1)
S = 0.0072	S = 0.0002[a]
C = −0.0015	
12. Echinoderms + hemichordates (L)	12. Hemichordates + protostomes (L2)
S = 0.0069	S = 0.0004
C = −0.0040	C = −0.0083
13. Enteropneusts (M)	13. *Saccoglossus* + *Cephalodiscus* (M2 = O)
S = 0.0062	S = 0.0046
C = −0.0013	C = −0.0032
14. Lancelet + vertebrates (P)	14. Vertebrates + urochordates (P2)
S = 0.0044	S = 0.0008
C = −0.0031	C = −0.0074
15. Ascidian urochordates (Q)	15. *Ciona* + *Thalia* (Q1 = T)
S = 0.0039	S = 0.0030
C = −0.0011	C = −0.0017
16. Chordates (R)	16. Urochordates + protostomes (R2)
S = 0.0036	S = 0.0009
C = −0.0065	C = −0.0050
17. Chondrichthyes (S)	17. *Triakis* + *Ambystoma* (S1)
S = 0.0031	S = 0.0012
C = −0.0006	C = −0.0017

NOTE.—Symbols (A)–(T) refer to the same groups as in figures 2*B* and 4.

[a] Conflict values were not calculated for these groupings because their levels of support were below the 0.0003 threshold; i.e., they had no support.

brates: Ahlberg 2001, pp. 1–32; Ambulacraria: Castresana et al. 1998; Bromham and Degnan 1999; Ahlberg 2001, pp. 5–11, 44–45; cyclostomes: Kuraku et al. 1999; chondrichthyan groups: Stiassny, Parenti, and Johnson 1996, p. 35; echinoderm groups: Brusca and Brusca 1990, p. 801).

3. Our LSU sequences do not support two SSU-based hypotheses (Cameron, Garey, and Swalla 2000; Swalla et al. 2000) that within hemichordates, pterobranchs evolved from an enteropneust and within urochordates, thaliaceans evolved from an ascidian. That is, although our combined LSU + SSU tree did support these hypotheses (fig. 1*C*), our LSU-only tree and our spectral analysis of the LSU + SSU data supported alternate hypotheses: monophyletic enteropneusts and ascidians (figs. 1*A* and 3). Further data

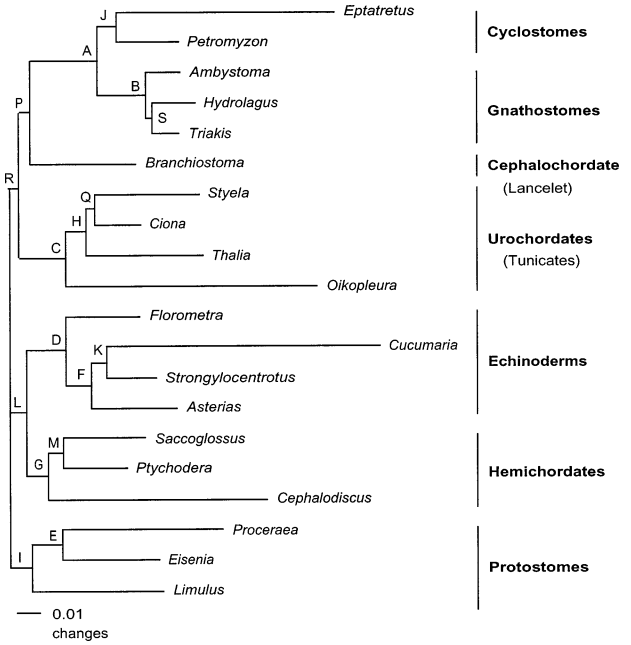

FIG. 3.—Manhattan tree computed by Spectrum (based on the 20 taxa from fig. 2). The expected phylogenetic spectrum of this tree is the closest manhattan (or city-block) distance from the spectrum obtained by the LSU + SSU rRNA gene data. Letters at nodes represent splits identified in figure 2*B*.

are necessary to clarify relationships within these groups.

4. All our new evidence strengthened the SSU-based hypothesis (Wada 1998; Swalla et al. 2000) that larvaceans are an outgroup to tunicates (figs. 1*A*, 1*C*, and 3). However, it has been shown that larvacean sequences have long branches (Cameron, Garey, and Swalla 2000), which can give high support to incorrect phylogenetic relationships. Also, this hypothesis has not yet received independent support, so it is premature to conclude that ancestral urochordates were larvacean-like.
5. The combined LSU + SSU sequences contain ambiguous evidence, both for and against the monophyly of the chordates (see *Chordates*).

Chordates

The animals classified within the phylum Chordata—urochordates, cephalochordates, and vertebrates—share the following anatomic synapomorphies: a notochord, a dorsal hollow nerve cord, a tail extending behind the visceral cavity, a thyroid gland (or endostyle), and other features (Brusca and Brusca 1990, p. 873). Despite such strong morphological evidence for this clade, molecular systematists have not been able to obtain strong evidence for chordate monophyly using the SSU gene (Turbeville, Schulz, and Raff 1994; Wada and Satoh 1994). The only evidence for a chordate clade from molecular (SSU) data was obtained when compar-

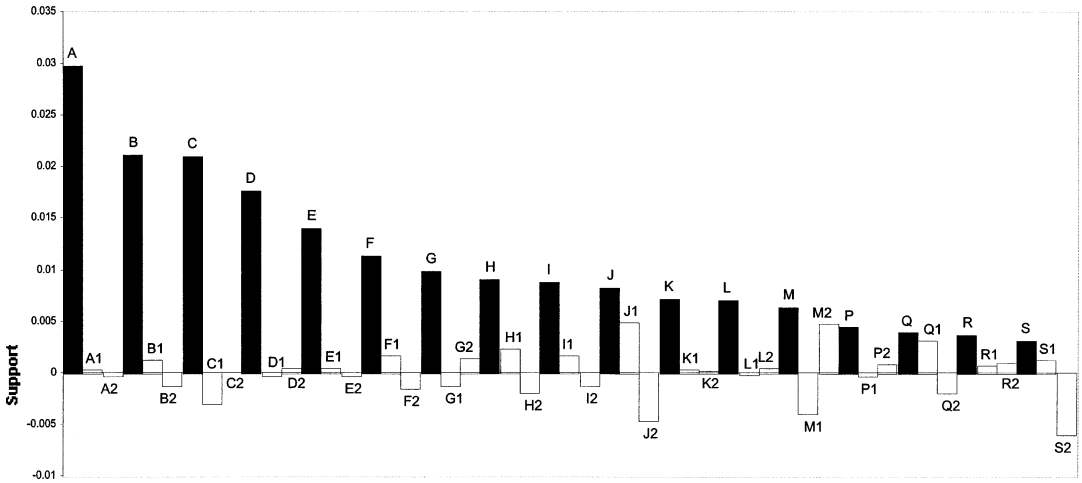

FIG. 4.—Nearest-neighbor interchange plot showing relative amounts of support (measured as the number of changes per site) for groups present in the manhattan tree (fig. 3) and for alternate groups resulting from local rearrangements around every internal edge of this tree (see text and fig. 5 for further explanation). Bars labeled with only letters represent groups occurring in the manhattan tree (fig. 3), and bars labeled with letters plus numbers depict alternate groups resulting from rearrangements of this tree. A = vertebrates, A1 = cyclostomes + lancelet, A2 = gnathostomes + lancelet; B = gnathostomes, B1 = cyclostomes + chondrichthyans, B2 = cyclostomes + tetrapod; C = urochordates, C1 = (lancelet + vertebrates) + *Oikopleura*, C2 = (lancelet + vertebrates) + non–larvacean urochordates; D = echinoderms, D1 = non–crinoid echinoderms + hemichordates, D2 = crinoid + hemichordates; E = *Proceraea* + *Eisenia*, E1 = *Proceraea* + *Limulus*, E2 = *Eisenia* + *Limulus*; F = non–crinoid echinoderms, F1 = *Florometra* + *Asterias*, F2 = non–asteroid echinoderms; G = hemichordates, G1 = echinoderms + pterobranch, G2 = echinoderms + acorn worms; H = non–larvacean urochordates, H1 = *Thalia* + *Oikopleura*, H2 = non–thaliacean urochordates; I = deuterostomes, I1 = Ambulacraria + *Limulus*, I2 = Ambulacraria + annelids; J = cyclostomes, J1 = gnathostomes + *Petromyzon*, J2 = gnathostomes + *Eptatretus*; K = *Cucumaria* + *Strongylocentrotus*, K1 = *Strongylocentrotus* + *Asterias*, K2 = *Cucumaria* + *Asterias*; L = hemichordates + echinoderms, L1 = echinoderms + protostomes, L2 = hemichordates + protostomes; M = enteropneusts, M1 = *Ptychodera* + *Cephalodiscus*, M2 = *Saccoglossus* + *Cephalodiscus*; P = vertebrates + lancelet, P1 = lancelet + urochordates, P2 = vertebrates + urochordates; Q = ascidian urochordates, Q1 = *Ciona* + *Thalia*, Q2 = *Styela* + *Thalia*; R = chordates, R1 = urochordates + (hemichordates + echinoderms), R2 = urochordates + protostomes; S = chondrichthyans, S1 = *Triakis* + *Ambystoma*, S2 = *Hydrolagus* + *Ambystoma*.

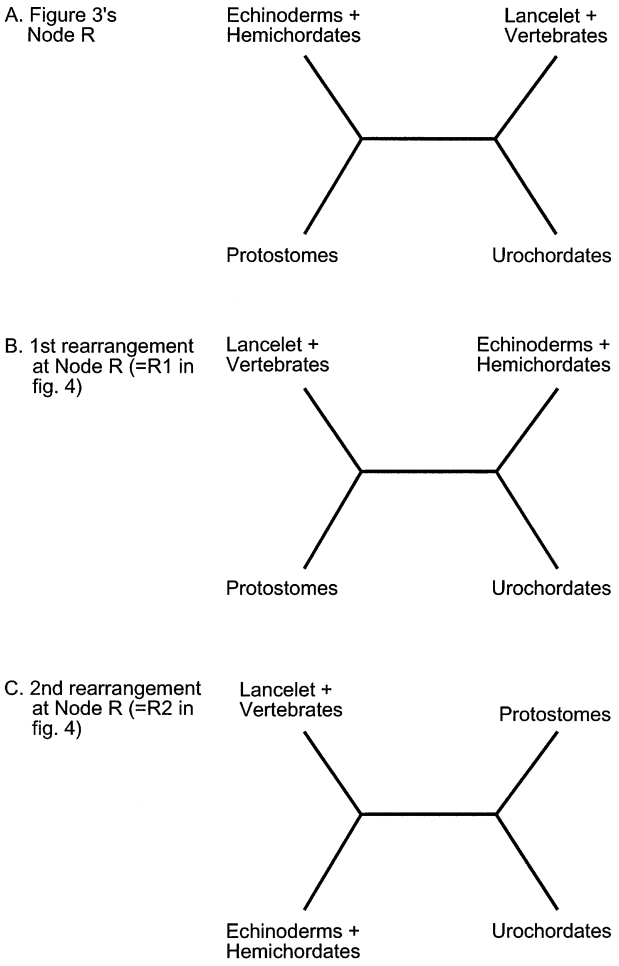

FIG. 5.—Example explaining the nearest-neighbor interchange procedure that was used to find support for every local rearrangement about the nodes in figure 3. These support values are recorded in figure 4. *A,* The configuration of subtrees grouped by node R in figure 3. *B,* Upon one rearrangement of this node, groups have been swapped across the tree, creating new relationships. *C,* A final rearrangement at R results in the last possible grouping of these four branches or subtrees.

ing a select sample of deuterostome and outgroup taxa, i.e., those with the shortest branch lengths (Cameron, Garey, and Swalla 2000).

We experienced the same complications because the urochordate LSU + SSU sequences did not always group with those of other chordates. Although spectral analysis, an efficient method for identifying both support and conflict in the data, showed chordate monophyly (fig. 4 and table 3), none of the tree-based phylogenetic methods placed urochordates with the other chordates (fig. 1*C* and table 2, but see fig. 1*B* for moderate 18S support). Because the internal branches resolving relationships among the three main groups of deuterostomes shown in figure 1*C* (urochordates; echinoderms + hemichordates; and cephalochordates + vertebrates), are so short, it is possible that the combined data contain signal for urochordates as chordates, but it is too weak to overcome noise with only ~4,000 bases. To test this, we conducted bootstrap analyses with increasingly large pseudoreplicate data sets (2×, 3×, 4×, 5×, and 10× the size of the original matrix) using PAUP*. Because this analysis preserves the imbalances in base frequencies in the preliminary data set, we conducted ME analyses on LogDet distances with P_{inv} = 0.588 (the ML estimate under a GTR + I model) for each data set. In none of these analyses did either placement of urochordates—urochordates nested within a monophyletic Chordata or urochordates as the basal deuterostome lineage—receive more than 55% or less than 43% bootstrap support (table 4). Thus, tree-based analyses of rRNA genes cannot discriminate between these hypotheses. Interestingly, no other placement of urochordates received substantial support in any tree-based analysis.

Table 4
Attempt to Amplify Phylogenetic Signal for Chordate Relationships in the Combined LSU + SSU Data

Data Set Size	Chordate Monophyly (%)	Urochordates Basal (%)
1×	53	43
2×	49	51
3×	45	55
4×	51	48
5×	49	51
10×	51	49

NOTE.—Bootstrap support values are listed for monophyly of chordata versus a basal position of urochordates with increasingly large pseudoreplicate data sets. No amount of data of the nature of the rRNA genes examined here will resolve the position of urochordates. Conflicting signal in the data confounds tree-based analysis.

Assuming Chordata is a real group, why was there no clear signal for chordate monophyly in our LSU + SSU data? It may be because of our limited taxonomic sampling of urochordates—but SSU sequences from many more urochordates still do not readily unite them with other chordates (Swalla et al. 2000). If these rRNA-based findings can be taken literally, they might mean that urochordates diverged from the base of the chordate tree much earlier than other chordate groups and that their ancestral genetic signal was overwhelmed with noise from long-term change. Alternatively, it might mean that urochordate genes have evolved rapidly and in an unusual way that obscures their relation to chordates. In support of this latter interpretation, the branch leading to urochordates in the rRNA gene tree (fig. 1*C*) is longer than that leading to any other phylum, and the single mitochondrial genome that has been sequenced from a urochordate, the ascidian *Halocynthia roretzi*, differs strikingly in base composition, gene arrangement, and gene length (mitochondrial rRNA genes) from all other metazoans examined (Yokobori et al. 1999). Whether the entire genome of urochordates is so divergent should be investigated by future studies.

Testing Non–Molecular Hypotheses of Deuterostome Interrelationships

Deuterostome phylogeny has been a topic of interest among biologists of many subdisciplines and is discussed in most biology and comparative vertebrate anatomy textbooks (Pough, Janis, and Heiser 1999, p. 33;

Kardong 2002, p. 47). Many morphological hypotheses of the relationships among the groups of deuterostomes have been championed and will be assessed here. A number of radical hypotheses have been proposed, such as those of Gaskell (1890) and Patten (1890), who claimed vertebrates evolved from an arthropod ancestor similar to the horseshoe crab, *Limulus*, and that of Jensen and others (see Jensen 1988), who claimed a nemertean-worm ancestry for the vertebrates. We will exclude these hypotheses from our discussion because both are firmly refuted by rRNA genes and other evidence indicating arthropods and nemerteans are protostomes (fig. 1*C;* Sundberg, Turbeville, and Härlin 1998; Zrzavy et al. 1998; Giribet et al. 2000; Mallatt and Winchell 2002). For other hypotheses, we will summarize each in the form of a cladogram (fig. 6) and assess its validity using spectral analysis and bootstrapping of our combined-gene data set (table 5). Because spectral analysis is only capable of measuring support and conflict for bipartitions within a data set, we cannot evaluate an overall topology incorporating every subgroup at the same time, but we can dissect relationships essential to each hypothesis and evaluate its components.

The classical hypothesis of deuterostome phylogeny (fig. 6*A*) is the simplest and most often cited (Young 1962, p. 78; Maisey 1986; Schaeffer 1987). It places hemichordates with chordates, not with echinoderms. Molecular evidence, presented here and in previous studies (see previously), that echinoderms and hemichordates form a natural group conflicts with this hypothesis, as does the utter lack of spectral analysis support for a hemichordate + chordate grouping (see table 5, part A).

Garstang's (1928) hypothesis (see Jefferies 1986, p. 347), dominant a generation ago and still cited (Berrill 1955; Romer 1967; Pough, Janis, and Heiser 1999; Kardong 2002), was based on the proposal that the ancestral forms of every deuterostome group were sessile suspension feeders and that a neotenic transformation of a lineage with ascidian-like larvae generated the free-swimming cephalochordate + vertebrate line. Garstang's phylogeny (fig. 6*B*) resembles the classical hypothesis, except that it proposes a paraphyletic Hemichordata with an enteropneust + chordate clade arising from a pterobranch-like ancestor. In contrast, our data strongly favor a monophyletic hemichordate clade over an enteropneust + chordate grouping (see table 5, part B). Garstang's hypothesis also proposed a sister relationship between urochordates of the classes Larvacea and Thaliacea, but our rRNA-gene data strongly united all non–larvacean urochordates (see table 5, part B).

Berrill's (1955) hypothesis (see Gee 1996, p. 122) is much like Garstang's, except that it explicitly states the neotenic ancestor of larvaceans and thaliaceans gave rise to the cephalochordate + vertebrate clade (fig. 6*C*). That is, Berrill proposed only one neotenic event giving rise to a common ancestor of larvaceans, thaliaceans, cephalochordates, and vertebrates, whereas Garstang had proposed two parallel neotenic events: one giving rise to larvaceans and the other to the ancestor of cephalochordates and vertebrates. Berrill's phylogeny (fig. 6*C*) suggests paraphyly of the urochordates (lancelet + vertebrates as the sister group to thaliaceans + larvaceans), but this received much less support from our data than the alternative hypothesis of a monophyletic urochordata (see table 5, part C).

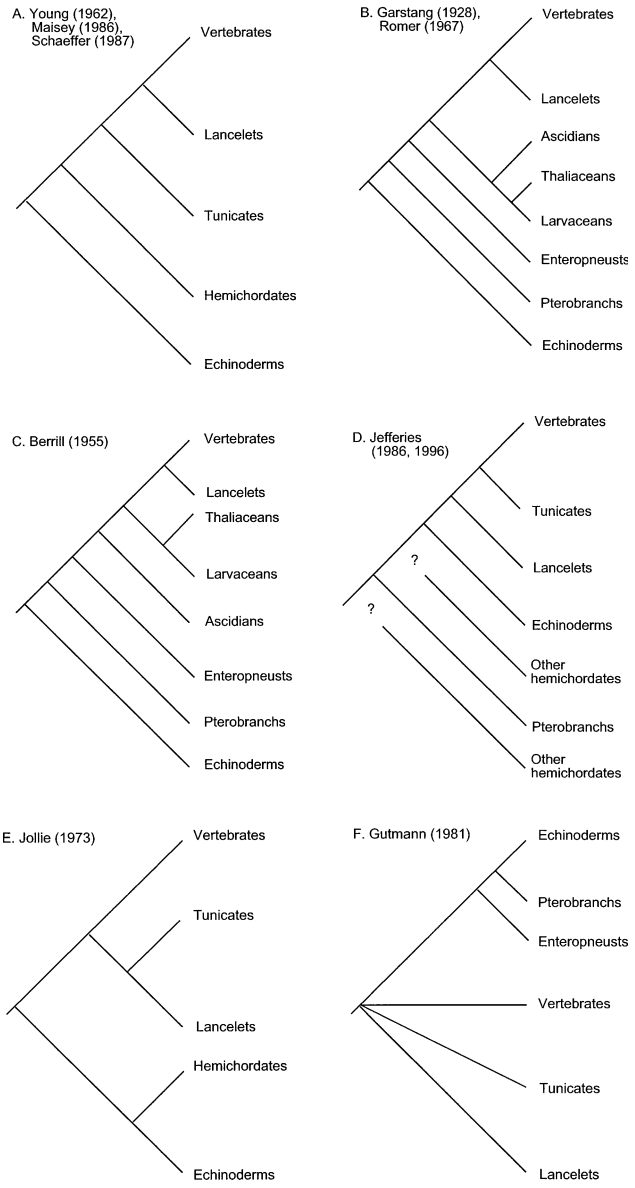

FIG. 6.—Anatomy-based hypotheses of deuterostome phylogeny, as proposed by previous authors.

The strange Paleozoic fossils known as carpoids, which bear large calcium carbonate plates, are considered by most paleontologists to be extinct echinoderms (see Peterson 1995). However, the calcichordate hypothesis (Jefferies 1986, 1996) posits that various carpoids are stem-group echinoderms, urochordates, cephalochordates, and vertebrates. Jefferies' analysis of these fossils led him to conclude that the sister group to vertebrates is urochordates (not cephalochordates) and that echinoderms are not the sister group of hemichordates but instead are the sister group to the chordates (fig. 6*D*). Our results from spectral analysis of rRNA genes do not

Table 5
Evaluation of Anatomy Based Hypotheses of Deuterostome Phylogeny with Spectral Analysis of LSU + SSU rRNA LogDet Distances (among the 20 taxa used in fig. 2)

Hypothesis	Distinguishing Feature	Support	Conflict	Best Alternative to This Feature (See figs. 2*B* and 4)
A. Young (1962), Maisey (1986), and Schaeffer (1987) . .	1) Hemichordates + Chordates	—[a]	—[a]	Split L = hemichordates + echinoderms; (S = 0.0069 and C = −0.0040)
B. Garstang (1928)	1) Enteropneusts + Chordates	0.0008	−0.0126	Split G = monophyletic hemichordates; (S = 0.0099 and C = −0.0023)
	2) Thaliaceans + Larvaceans	0.0023	−0.0047	Split H = thaliaceans + ascidians; (S = 0.0090 and C = −0.0012)
C. Berrill (1955)	1) Thaliaceans + Larvaceans + Vertebrates + Cephalochordates	—[a]	—[a]	Split C = monophyletic urochordates; (S = 0.0209 and C = −0.0026)
D. Jefferies (1996)	1) Urochordates + Vertebrates	0.0008	−0.0074	Split P = lancelet + vertebrates; (S = 0.0044 and C = −0.0031)
	2) Echinoderms + Chordates	0.0004	−0.0083	Split L = echinoderms + hemichordates; (S = 0.0069 and C = −0.0040)
E. Jollie (1973)	1) Echinoderms + Hemichordates	0.0069	−0.0040	This grouping is upheld by our data.
	2) Urochordates + Cephalochordates	—[a]	—[a]	Split P = lancelet + vertebrates; (S = 0.0044 and C = −0.0031)
F. Gutmann (1981)	1) Pterobranchs + Echinoderms	—[a]	—[a]	Split G = monophyletic hemichordates; (S = 0.0099 and C = −0.0023)

[a] Our data show no support for these groupings; i.e., support was below the 0.0003 threshold.

support these claims. As shown in table 5 (see table 5, part D), the support and conflict values strongly favor a cephalochordate + vertebrate pairing over a urochordate + vertebrate pairing and favor a hemichordate + echinoderm pairing over an echinoderm + chordate pairing. Incidentally, we interpret carpoids as resembling stem echinoderms only, which had many pleisiomorphic deuterostome traits; Conway Morris (2000) and Gee (2001) drew the same conclusion.

While this paper was in review, Jefferies accepted the ambulacrarian concept uniting hemichordates and echinoderms (Jefferies 2001) but without changing any key aspect of his calcichordate hypothesis. Nonetheless, our evidence against a urochordate + vertebrate group still contradicts the carpoid theory of vertebrate origins.

Jollie (1973) compared anatomic details of the deuterostome subgroups and proposed: (1) monophyly of the hemichordates + echinoderms based on similarities in larval morphology, and (2) monophyly of the urochordates + lancelets based on details of their embryology (fig. 6*E*). Although the first grouping is supported by our data, the second receives no support and is strongly refuted by our evidence for a lancelet + vertebrate group (see table 5, part E).

Finally, Gutmann (1981) proposed a phylogeny, based on a functional-mechanical analysis of the coelom and body musculature, that concluded deuterostomes evolved from a cephalochordate-like ancestor (fig. 6*F*). His hypothesis implies that cephalochordates are not particularly closely related to vertebrates, that chordates do not form a natural group, and that hemichordates are paraphyletic (that is, pterobranchs gave rise to echinoderms). Our data refute this; we found the cephalochordate + vertebrate pairing to be robust (split P), our spectral analysis supports a monophyletic Chordata (split R), and we do not support his pterobranch + echinoderm clade (instead, spectral analysis confirms a monophyletic Hemichordata; Split S, see table 5, part F).

In summary, our molecular data refute many aspects of previously proposed hypotheses. However, we emphasize that our data are from just one gene family—a minute portion of any genome—and to test these hypotheses rigorously much more genetic data, particularly protein-coding genes, should also be considered. Although they refute all of the above hypotheses, our main results based on rRNA genes (figs. 1*C* and 2*B*) are not radical but uphold many widely accepted and traditional aspects of deuterostome phylogeny. In particular, monophyly of the lancelets + vertebrates has long been a favored view; we do not refute a monophyletic Chordata; and evidence for a hemichordate + echinoderm clade is now gaining widespread acceptance.

Acknowledgments

Many thanks to Mary Fauci, Albrecht Fischer, Julie Fronzuto, Jim Garey, Tom Humphreys, and Andrew Martin for providing animals or other biological samples. The authors appreciate the kind help and advice of Joe Brunelli, Derek Pouchnick, and Gary Thorgaard. Some of the calculations were performed using the Wisconsin Package Version 10.0, Genetics Computer Group, provided by the VADMS Center, a regional biocomputing resource at Washington State University. VADMS is supported by WSU's Division of Sciences and the School of Molecular Biosciences. B.J.S. thanks NSF grant IBN 0096266 and Penn State Start-Up Funds for funding some DNA extractions and sequencing in her laboratory.

LITERATURE CITED

AHLBERG, P. E., ed. 2001. Major events in early vertebrate evolution: palaeontology, phylogeny, genetics, and development. Taylor and Francis, London.

ALBERTS, B., D. BRAY, J. LEWIS, M. RAFF, K. ROBERTS, and J. D. WATSON. 1994. Molecular biology of the cell. 3rd edition. Garland Publishing, New York.

BERRILL, N. J. 1955. The origin of the vertebrates. Oxford University Press, Oxford.

BROMHAM, L. D., and B. M. DEGNAN. 1999. Hemichordates and deuterostome evolution: robust molecular phylogenetic support for a hemichordate + echinoderm clade. Evol. Dev. **1**:166–171.

BRUSCA, R. C., and G. J. BRUSCA. 1990. Invertebrates. Sinauer, Sunderland, Mass.

CAMERON, C. B. 2000. The phylogeny of the Hemichordata and ecology of two new enteropneust species from Barkley Sound. Doctoral dissertation, University of Alberta.

CAMERON, C. B., J. R. GAREY, and B. J. SWALLA. 2000. Evolution of the chordate body plan: new insights from phylogenetic analyses of deuterostome phyla. Proc. Natl. Acad. Sci. USA **97**:4469–4474.

CASTRESANA, J., G. FELDMAIER-FUCHS, S. YOKOBORI, N. SATOH, and S. PÄÄBO. 1998. The mitochondrial genome of the hemichordate *Balanoglossus carnosus* and the evolution of deuterostome mitochondria. Genetics **150**:1115–1123.

CHARLESTON, M. 1998. Spectrum Version 2.3.0. Available from http://taxonomy.zoology.gla.ac.uk/~mac/spectrum/spectrum.

CONWAY MORRIS, S. 2000. The Cambrian ''explosion'': slow-fuse or megatonnage? Proc. Natl. Acad. Sci. USA **97**:4426–4429.

CUNNINGHAM, C. W. 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol. **14**: 733–740.

DE ROSA, R., J. K. GRENIER, T. ANDREEVA, C. E. COOK, A. ADOUTTE, M. AKAM, S. B. CARROL, and G. BALAVOINE. 1999. Hox genes in brachiopods and priapulans and protostome evolution. Nature **399**:772–776.

DE QUEIROZ, A., M. J. DONOGHUE, and J. KIM. 1995. Separate versus combined analysis of phylogenetic evidence. Ann. Rev. Ecol. Syst. **26**:657–681.

FARRIS, J. S., M. KALLERSJO, A. G. KLUGE, and C. BULT. 1994. Testing significance of incongruence. Cladistics **10**:315–319.

FELSENSTEIN, J. 1981. Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. J. Mol. Evol. **17**:368–376.

———. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **17**:368–376.

GARSTANG, W. 1928. The morphology of the tunicata and its bearing on the phylogeny of the chordata. Q. J. Microsc. Sci. **72**:51–187.

GASKELL, W. H. 1890. On the origin of vertebrates from a crustacean-like ancestor. Q. J. Microsc. Soc. **31**:379–444.

GEE, H. 1996. Before the backbone. Chapman and Hall, London.

———. 2001. Deuterostome phylogeny: the context for the origin and evolution of the vertebrates. Pp. 1–14 *in* P. E. AHLBERG, ed. Major events in early vertebrate evolution: palaeontology, phylogeny, genetics, and development. Taylor and Francis, London.

GIRIBET, G., D. L. DISTEL, M. POLZ, W. STERRER, and W. C. WHEELER. 2000. Triploblastic relationships with emphasis on the acoelomates and the position of Gnathostomulida, Cycliophora, Plathelminthes, and Chaetognatha: a combined approach of the 18S rDNA sequences and morphology. Syst. Biol. **49**:539–562.

GUTELL, R. R. 1994. Collection of small subunit (16S- and 16S-like) ribosomal RNA structures. Nucleic Acids Res. **22**: 3502–3507.

GUTMANN, W. F. 1981. Relationships between invertebrate phyla based on functional-mechanical analysis of the hydrostatic skeleton. Am. Zool. **21**:63–81.

HALANYCH, K. M. 1995. The phylogenetic position of the pterobranch hemichordates based on 18S rDNA sequence data. Mol. Phylogenet. Evol. **4**:72–76.

———. 1996. Testing hypotheses of chaetognath origins; long branches revealed by 18S rDNA. Syst. Biol. **45**:223–246.

HALANYCH, K. M., J. BACHELOR, A. M. A. AGUINALDO, S. LIVA, D. M. HILLIS, and J. A. LAKE. 1995. 18S rDNA evidence that the lophophorates are protostome animals. Science **267**:1641–1643.

HASSOUNA, N., B. MICHOT, and J.-P. BACHELLERIE. 1984. The complete nucleotide sequence of mouse 28S rRNA gene: implications for the process of size increase of the large subunit rRNA in higher eukaryotes. Nucleic Acids Res. **12**: 3563–3583.

HILLIS, D. M., and M. T. DIXON. 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. Q. Rev. Biol. **66**:411–453.

HYMAN, L. H. 1955. The invertebrates: echinodermata. The coelomate bilateria, Vol. IV. McGraw-Hill, New York.

JEFFERIES, R. P. S. 1986. The ancestry of the vertebrates. British Museum (Natural History), London.

———. 1996. The early phylogeny of chordates and echinoderms and the origin of chordate left-right asymmetry and bilateral symmetry. Acta Zool. Stockh. **77**:101–122.

———. 2001. The origin and early fossil history of the chordate acoustico-lateralis system, with remarks on the reality of the echinoderm-hemichordate clade. Pp. 40–66 *in* P. E. AHLBERG, ed. Major events in early vertebrate evolution: palaeontology, phylogeny, and development. Taylor and Francis, London.

JENSEN, D. D. 1988. Hubrecht, Macfarlane, Jensen, and Willmer: on the nature and testability of four versions of the nemertean theory of vertebrate origins. Hydrobiologia **156**: 99–104.

JOLLIE, M. J. 1973. The origin of the chordates. Acta Zool. **54**: 81–100.

KARDONG, K. V. 2002. Vertebrates: comparative anatomy, function, evolution. 3rd edition. WCW/McGraw Hill, Boston.

KURAKU, S., D. HOSHIYAMA, K. KATOH, H. SUGA, and T. MIYATA. 1999. Monophyly of lampreys and hagfishes supported by nuclear DNA-coded genes. J. Mol. Evol. **49**:729–735.

LAFAY, B., A. B. SMITH, and R. CHRISTEN. 1995. A combined morphological and molecular approach to the phylogeny of asteroids (Asteroidea: Echinodermata). Syst. Biol. **44**:190–208.

LAKE, J. A. 1994. Reconstructing evolutionary trees from DNA and protein sequences: paralinear distances. Proc. Natl. Acad. Sci. USA **91**:1455–1459.

———. 1995. Calculating the probability of multitaxon evolutionary trees: Boostrappers Gambit. Proc. Natl. Acad. Sci. USA **92**:9662–9666.

LENTO, G. M., R. E. HICKSON, G. K. CHAMBERS, and D. PENNY. 1995. Use of spectral analysis to test hypotheses on the origin of pinnipeds. Mol. Biol. Evol. **12**:28–52.

LOCKHART, P. J., M. A. STEEL, M. D. HENDY, and D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. **11**:605–612.

MACKEY, L. Y., B. WINNEPENNINCKX, R. DE WACHTER, T. BACKELJAU, P. EMSCHERMANN, and J. R. GAREY. 1996. 18S rRNA suggests that entoprocta are protostomes, unrelated to ectoprocta. J. Mol. Evol. **42**:552–559.

MAISEY, J. G. 1986. Heads and tails: a chordate phylogeny. Cladistics **2**:201–256.

MALLATT, J., and J. SULLIVAN. 1998. 28S and 18S rDNA sequences support the monophyly of lampreys and hagfishes. Mol. Biol. Evol. **15**:1706–1718.

MALLATT, J., J. SULLIVAN, and C. J. WINCHELL. 2001. The relationships of lampreys to hagfishes: a spectral analysis of ribosomal DNA sequences. Pp. 106–118 *in* P. E. AHLBERG, ed. Major events in early vertebrate evolution: palaeontology, phylogeny, and development. Taylor and Francis, London.

MALLATT, J., and C. J. WINCHELL. 2002. Testing the new animal phylogeny: first use of combined large-subunit and small-subunit rRNA gene sequences to classify the protostomes. Mol. Biol. Evol. **19**:289–301.

METSCHNIKOFF, V. E. 1881. Über die systematische Stellung von *Balanoglossus*. Zool. Anz. **4**:139–157.

NIELSEN, C. 1994. Larval and adult characters in animal phylogeny. Am. Zool. **34**:492–501.

———. 1995. Animal evolution: interrelationships of the living phyla. Oxford University Press, New York.

PATTEN, W. 1890. On the origin of vertebrates from arachnids. Q. J. Microsc. Sci. **31**:317–378.

PENNY, D., M. HASEGAWA, P. J. WADDELL, and M. D. HENDY. 1999. Mammalian evolution: timing and implications from using the LogDeterminant transform for proteins of differing amino acid composition. Syst. Biol. **48**:76–93.

PETERSON, K. J. 1995. A phylogenetic test of the calcichordate scenario. Lethaia **28**:25–38.

PETERSON, K. J., and D. J. EERNISSE. 2001. Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences. Evol. Dev. **3**:170–205.

POUGH, F. H., C. M. JANIS, and J. B. HEISER. 1999. Vertebrate life. 5th edition. Prentice Hall, Upper Saddle River, New Jersey.

ROMER, A. S. 1967. Major steps in vertebrate evolution. Science **158**:1629–1637.

SCHAEFFER, B. 1987. Deuterostome monophyly and phylogeny. Evol. Biol. **21**:179–235.

SCHNARE, M. N., S. H. DAMBERGER, M. W. GRAY, and R. GUTELL. 1996. Comprehensive comparison of structural characteristics in eukaryotic cytoplasmic large subunit (23S-like) ribosomal RNA. J. Mol. Biol. **256**:701–719.

SMITH, S. W., R. OVERBEEK, C. R. WOESE, W. GILBERT, and P. M. GILLEVET. 1994. The genetic data environment: an expandable GUI for multiple sequence analysis. CABIOS **10**:671–675.

STECHMANN, A., and M. SCHLEGEL. 1999. Analysis of the complete mitochondrial DNA sequence of the brachiopod *Terebratulina retusa* places Brachiopoda within the protostomes. Proc. R. Soc. Lond. B **266**:2043–2052.

STIASSNY, M. L. J., L. R. PARENTI, and G. D. JOHNSON, eds. 1996. Interrelationships of fishes. Academic Press, San Diego.

SULLIVAN, J. 1996. Combining data with different distributions of among-site rate variation. Syst. Biol. **45**:375–380.

SUNDBERG, P., J. M. TURBEVILLE, and M. S. HÄRLIN. 1998. There is no support for Jensen's hypothesis of nemerteans as ancestors to the vertebrates. Hydrobiologia **365**:47–54.

SWALLA, B. J. 2001. Phylogeny of the urochordates: implications for chordate evolution. Pp. 219–224 *in* H. SAWADA and C. LAMBERT, eds. Biology of ascidians. Proceedings of the First International Symposium on the Biology of Ascidians. Springer Verlag, Tokyo.

SWALLA, B. J., C. B. CAMERON, L. S. CORLEY, and J. R. GAREY. 2000. Urochordates are monophyletic within the deuterostomes. Syst. Biol. **49**:122–134.

SWOFFORD, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (* and other methods). Version 4.0 beta 4a. Sinauer Associates, Sunderland, Mass.

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, and D. M. HILLIS. 1996. Phylogenetic inference. Pp. 407–514 *in* D. M. HILLIS, C. MORITZ,, and B. K. MABLE, eds. Molecular systematics. 2nd edition. Sinauer, Sunderland, Mass.

TELFORD, M. J., and P. W. H. HOLLAND. 1993. The phylogenetic affinities of the chaetognaths: a molecular analysis. Mol. Biol. Evol. **10**:660–676.

TURBEVILLE, J. M., J. R. SCHULZ, and R. A. RAFF. 1994. Deuterostome phylogeny and the sister group of the chordates: evidence for molecules and morphology. Mol. Biol. Evol. **11**:648–655.

VALENTINE, J. W. 1997. Cleavage patterns and the topology of the metazoan tree of life. Proc. Natl. Acad. Sci. USA **94**: 8001–8005.

WADA, H. 1998. Evolutionary history of free-swimming and sessile lifestyles in urochordates as deduced from 18S rDNA molecular phylogeny. Mol. Biol. Evol. **15**:1189–1194.

WADA, H., and N. SATOH. 1994. Details of the evolutionary history from the invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. Proc. Natl. Acad. Sci. USA **91**:1801–1804.

WADDELL, P. J., D. PENNY, and T. MOORE. 1997. Hadamard conjugations and modeling sequence evolution with unequal rates across sites. Mol. Phylogenet. Evol. **8**:398–414.

WINNEPENNINCKX, B., T. BACKELJAU, and R. DE WACHTER. 1993. Extraction of high molecular weight DNA from molluscs. Trends Genet. **9**:407.

WINNEPENNINCKX, B., Y. VAN DE PEER, and T. BACKELJAU. 1998. Metazoan relationships on the basis of 18S rRNA sequences: a few years later. . . Am. Zool. **38**:888–906.

YOKOBORI, S., T. UEDA, G. FELDMAIER-FUCHS, S. PÄÄBO, R. UESHIMA, A. KONDOW, K. NISHIKAWA, and K. WATANABE. 1999. Complete DNA sequence of the mitochondrial genome of the ascidian *Halocynthia roretzi* (Chordata, Urochordata). Genetics **153**:1851–1862.

YOUNG, J. Z. 1962. The life of vertebrates. 2nd edition. Oxford University Press, Oxford.

ZRZAVY, J., S. MIHULKA, P. KEPKA, A. BEZDEK, and D. TIETZ. 1998. Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. Cladistics **14**:249–285.

STEPHEN PALUMBI, reviewing editor

Accepted January 21, 2002