SHORT COMMUNICATION

The Position of Arthropods in the Animal Kingdom: Ecdysozoa, Islands, Trees, and the "Parsimony Ratchet"

Gonzalo Giribet and Ward C. Wheeler

Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024

INTRODUCTION

In a recent paper published in *Molecular Phylogenetics and Evolution,* Giribet and Ribera (1998) placed arthropods within the Animal kingdom by analyzing the 18S rDNA locus of 133 metazoan taxa. The goal of that paper was twofold: (1) to investigate the position of the Arthropoda within the protostome animals and (2) to determine arthropod sister-group relationships. Sequence data were analyzed by means of unweighted parsimony (gaps coded as a character state) in NONA 1.5.1 (Goloboff, 1993) using different tree-searching strategies, with the combination of different swappers. Giribet and Ribera found 756 trees of 7032 steps. A strict consensus of these 756 trees is shown in Fig. 1A. The phylogenetic conclusions were summarized as follows:

(1) The Bilateria are supported as monophyletic.

(2) The Mesozoa appear as a monophyletic group and are the sister group to the remaining Bilateria. Other monophyletic groups are Ctenophora, Placozoa, Cnidaria, Urochordata, Craniata, Echinodermata, Enteropneusta, Priapula, Tardigrada, Nematoda, Nematomorpha, Onychophora, Arthropoda, Platyhelminthes, Clitellata, Sipuncula, Ectoprocta, Pogonophora (including Vestimentifera), and Acanthocephala. Nevertheless, Mollusca, Annelida, Phoronida, Brachiopoda, Bryozoa, Nemertea, and Rotifera appear as nonmonophyletic groups.

(3) The Deuterostomia appear as a paraphyletic group.

(4) The Protostomia appear monophyletic and split into two main lines: (1) a clade constituted by animals with typical spiralian development plus some "aschel-

minths" and the "lophophorates" and (2) a clade grouping arthropods with other molting animals (Ecdysozoa *sensu* Aguinaldo *et al.,* 1997).

One of the main findings of the paper was the identification of a "molting clade" or Ecdysozoa. However, Giribet and Ribera (1998: 487) recognized the difficulties of analyzing such a large data set.
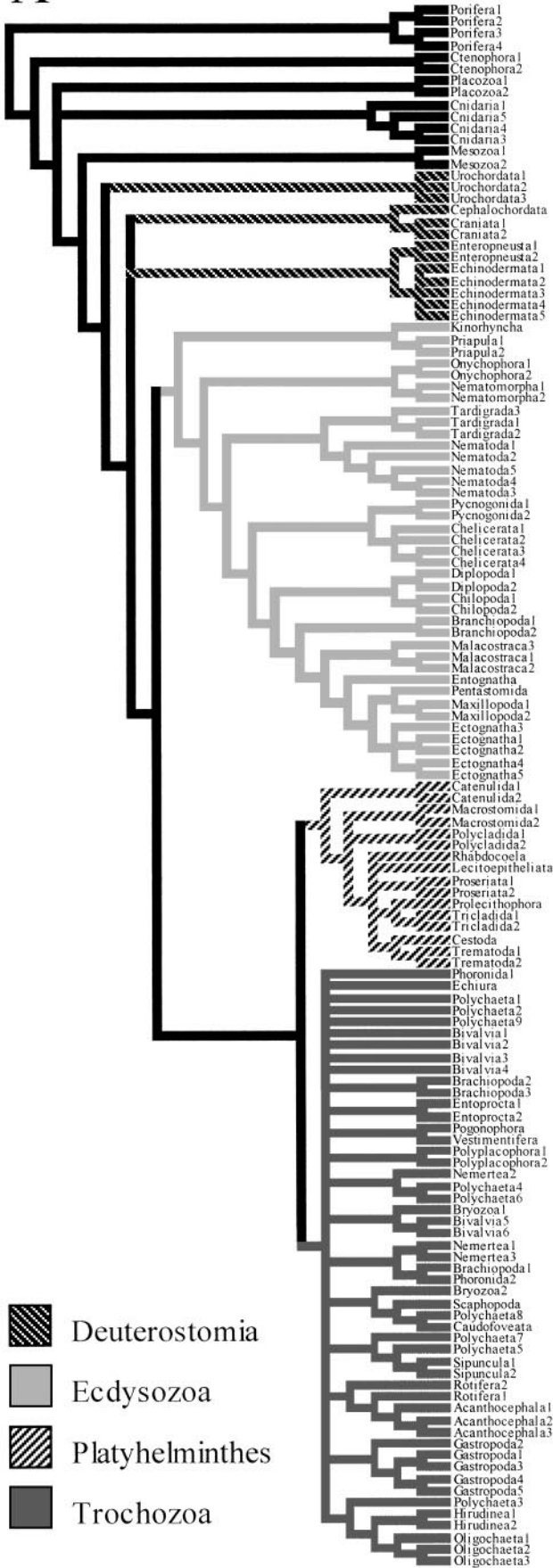
MOST-PARSIMONIOUS TREES AND ISLANDS

The problem of finding the most-parsimonious trees using heuristic searches is one of the major concerns in phylogenetic systematics. Several strategies for finding different islands of trees have been proposed (e.g., Maddison, 1991; Goloboff, 1996) involving randomization of taxon order and branch swapping on different starting trees. However, finding the most-parsimonious trees in an efficient way with current computing technology seemed to be almost intractable when dealing with large data sets. For example, Rice *et al.* (1997) re-ran the data set of Chase *et al.* (1993) and employed 11.6 months of CPU time to find trees five steps shorter than those published in 1993. The strategy followed by Rice *et al.* was swapping to completion the maximum number of possible starting trees.

A problem found in the commonly used programs PAUP 3.1.1 (Swofford, 1993) and PAUP* 4.0b1 (Swofford, 1998) is that if a starting tree is not swapped up to completion, the program stops and does not continue to the next replicate (random addition option).[1] This is especially problematic for large data sets because of the large number of trees that the program must handle, provoking the commonly known "out of memory" problem. In addition to that, swapping a large number of trees to completion results in extremely long analyses (e.g., spending up to 11.6 months in an analysis [Rice *et al.,* 1997] that could be handled in several hours or days with other search stratagies).

A novel method called "parsimony ratchet" that is extremely efficient at finding different islands of trees has recently been described (Nixon, 1998; Nixon, in

---

[1]This can be avoided by specifying not to store more than a few trees of length 1 and do several replicates (F. Anderson, pers. com.). This constraint can be inactivated and the stored trees can be swapped to completion. This strategy avoids unnecessary swapping in replicates that do not hit on one of the islands containing optimal trees.

A

B

Tree A taxon labels (top to bottom):
Porifera1, Porifera2, Porifera3, Porifera4, Ctenophora1, Ctenophora2, Placozoa1, Placozoa2, Cnidaria1, Cnidaria5, Cnidaria4, Cnidaria3, Mesozoa1, Mesozoa2, Urochordata1, Urochordata2, Urochordata3, Cephalochordata, Craniata1, Craniata2, Enteropneusta1, Enteropneusta2, Echinodermata1, Echinodermata2, Echinodermata3, Echinodermata4, Echinodermata5, Kinorhyncha, Priapula1, Priapula2, Onychophora1, Onychophora2, Nematomorpha1, Nematomorpha2, Tardigrada3, Tardigrada1, Tardigrada2, Nematoda1, Nematoda2, Nematoda5, Nematoda4, Nematoda3, Pycnogonida1, Pycnogonida2, Chelicerata1, Chelicerata2, Chelicerata3, Chelicerata4, Diplopoda1, Diplopoda2, Chilopoda1, Chilopoda2, Branchiopoda1, Branchiopoda2, Malacostraca3, Malacostraca1, Malacostraca2, Entognatha, Pentastomida, Maxillopoda1, Maxillopoda2, Ectognatha3, Ectognatha1, Ectognatha2, Ectognatha4, Ectognatha5, Catenulida1, Catenulida2, Macrostomida1, Macrostomida2, Polycladida1, Polycladida2, Rhabdocoela, Lecitoepitheliata, Proseriata1, Proseriata2, Prolecithophora, Tricladida1, Tricladida2, Cestoda, Trematoda1, Trematoda2, Phoronida1, Echiura, Polychaeta1, Polychaeta2, Polychaeta9, Bivalvia1, Bivalvia2, Bivalvia3, Bivalvia4, Brachiopoda3, Entoprocta1, Entoprocta2, Pogonophora, Vestimentifera, Polyplacophora1, Polyplacophora2, Nemertea2, Polychaeta4, Polychaeta6, Bryozoa1, Bivalvia5, Bivalvia6, Nemertea1, Nemertea3, Brachiopoda1, Phoronida2, Bryozoa2, Scaphopoda, Polychaeta8, Caudofoveata, Polychaeta7, Polychaeta5, Sipuncula1, Sipuncula2, Rotifera2, Rotifera1, Acanthocephala1, Acanthocephala2, Acanthocephala3, Gastropoda2, Gastropoda1, Gastropoda3, Gastropoda4, Gastropoda5, Polychaeta3, Hirudinea1, Hirudinea2, Oligochaeta1, Oligochaeta2, Oligochaeta3

Tree B taxon labels (top to bottom):
Porifera1, Porifera2, Porifera3, Porifera4, Ctenophora1, Ctenophora2, Cnidaria1, Cnidaria5, Cnidaria4, Cnidaria3, Placozoa1, Placozoa2, Mesozoa1, Mesozoa2, Scaphopoda, Polychaeta8, Caudofoveata, Rotifera2, Rotifera1, Acanthocephala1, Acanthocephala2, Acanthocephala3, Brachiopoda2, Brachiopoda3, Echiura, Polyplacophora1, Polyplacophora2, Nemertea2, Polychaeta4, Polychaeta6, Polychaeta5, Sipuncula1, Sipuncula2, Phoronida1, Bryozoa1, Bivalvia5, Bivalvia6, Bivalvia1, Bivalvia4, Bivalvia3, Bivalvia2, Gastropoda2, Gastropoda1, Gastropoda3, Gastropoda4, Gastropoda5, Polychaeta9, Entoprocta1, Entoprocta2, Nemertea1, Nemertea3, Phoronida2, Polychaeta1, Polychaeta2, Pogonophora, Vestimentifera, Bryozoa2, Polychaeta3, Hirudinea1, Hirudinea2, Oligochaeta1, Oligochaeta2, Oligochaeta3, Catenulida1, Catenulida2, Macrostomida1, Macrostomida2, Polycladida1, Polycladida2, Rhabdocoela, Lecitoepitheliata, Proseriata1, Proseriata2, Prolecithophora, Tricladida1, Tricladida2, Cestoda, Trematoda1, Trematoda2, Urochordata1, Urochordata2, Urochordata3, Cephalochordata, Craniata1, Craniata2, Enteropneusta1, Enteropneusta2, Echinodermata1, Echinodermata2, Echinodermata3, Echinodermata4, Echinodermata5, Kinorhyncha, Priapula1, Priapula2, Onychophora1, Onychophora2, Nematomorpha1, Nematomorpha2, Tardigrada3, Tardigrada1, Tardigrada2, Nematoda1, Nematoda2, Nematoda5, Nematoda4, Nematoda3, Pycnogonida1, Pycnogonida2, Chelicerata1, Chelicerata2, Chelicerata3, Chelicerata4, Diplopoda1, Diplopoda2, Chilopoda1, Chilopoda2, Branchiopoda1, Branchiopoda2, Malacostraca3, Malacostraca1, Malacostraca2, Entognatha, Pentastomida, Maxillopoda1, Maxillopoda2, Ectognatha3, Ectognatha1, Ectognatha2, Ectognatha4, Ectognatha5

Legend:
Deuterostomia
Ecdysozoa
Platyhelminthes
Trochozoa

prep.).[2] This method resulted in search times from 20 to 200 times faster than those using the more standard methods of random addition sequence, Wagner tree, SPR, or TBR (Nixon, pers. commun.). The method therefore extends the size of data sets that can reasonably be analyzed with existing computer hardware/software combinations.

The parsimony ratchet mechanics are summarized as follows (Nixon, in prep.). (1) Generate a starting tree. (2) Reweight a randomly selected subset of the characters. (3) Search on the current tree, holding only one tree, with the new weights. Any kind of swapping strategy might be used. (4) Reweight the characters to the original weights, and swap on the tree found in step 3. (5) Return to step 2 to begin another iteration. After reweighting in step 2, swapping is performed on the tree found in step 4. Continue this cycle for 20, 50, 100, or more iterations.

The parsimony ratchet is an iterative method, the goal of which is to find shorter trees more quickly by avoiding the time spent searching on new starting trees that are much less optimal than the last tree swapped. The reweighted characters favor trees that are potentially not in the same "island" as the current tree.

This novel method of tree searching has already been implemented in DADA (Nixon, 1998) under the name of "island hopper" and in NONA v. 1.9 (Goloboff, 1998) under the command "nixwts."

THE REANALYSIS

We reanalyzed the metazoan data set of Giribet and Ribera (1998) with the "nixwts" command from NONA v. 1.9 (see a full description of the options in the manual [Goloboff, 1998]) and found trees 4 steps shorter than those published (7028 versus 7032 steps). After applying the "jump" command and swapping on the resultant trees using "max*" (tbr branch swapping) to complete the most-parsimonious trees of each one of the islands, we obtained 432 trees of 7028 steps (CI = 0.24; RI = 0.50). A strict consensus of these trees is shown in Fig. 1B.

The phylogenetic conclusions are not definitive for two reasons. First, the alignments used by Giribet and Ribera (1998) may be far from optimal (manual alignments were used). Second, a single parameter value was analyzed (gap cost = transversions = transitions), while parameter sensitivity (Wheeler, 1995) is crucial

for understanding the robustness of phylogenetic hypotheses. However, the main goal of the present paper is to emphasize the analysis of large data sets using the parsimony ratchet and not to discuss metazoan phylogeny. Better analyses of metazoan phylogeny (work in progress by G.G. and W.W.) may be obtained (1) by avoiding alignment ambiguities by using the direct character optimization of Wheeler (1996), (2) by exploring a large number of parameters that would let us derive stricter conclusions, (3) by including a broader taxonomic sampling, and (4) by including morphological data.

The current results are important from a methodological point of view. Clearly, the new method developed by Nixon outperforms any other tree-searching strategy used in preexisting implementations, such as random addition sequences, branch swapping (i.e., SPR and TBR in PAUP or NONA), and branch swapping in suboptimal trees (such as the "jump" command in NONA). The implications of the ratchet are crucial for the analysis of large data sets, since the method allows analyses in a reasonable amount of time. We stress the necessity of more thorough, intense, and aggressive analyses for large data sets to avoid suboptimal solutions (as in Rice *et al.*, 1997; Giribet and Ribera, 1998). The ratchet seems to be one of the most effective methods in that respect.

IMPLICATIONS FOR METAZOAN PHYLOGENY

The present tree (Fig. 1B) differs considerably from the tree published by Giribet and Ribera (Fig. 1A). The monophyly of Ecdysozoa is still found in the new reanalysis but other conclusions for bilaterian evolution of the previous paper by Giribet and Ribera (1998) [(paraphyletic Deuterostomia (Ecdysozoa (Platyhelminthes + Trochozoa)))] are falsified in the present analysis. It still seems that within the Bilateria there are at least three main clades: one of acoelomate platyhelminthes, one of deuterostomes, and one of Ecdysozoa including arthropods and the remaining molting phyla. Whether the "spiralian-coelomates" (=Trochozoa) constitute a grade (in the present analysis) or a fourth real clade (Eernisse, 1998; Littlewood *et al.*, 1998) needs further analysis and probably a broader sampling.

At this point, it seems that the 18S rDNA gene is very informative, distinguishing between diploblastic and triploblastic (=Bilateria) metazoans (see the length of the branch separating Placozoans from the Mesozoans in Fig. 2) and certain groups of Bilateria (Ecdysozoa, Platyhelminthes, Trochozoa, Deuterostomia). However, 18S does not definitively resolve the internal relation-

FIG. 1. (A) Strict consensus of 756 trees of 7032 steps (CI = 0.24; RI = 0.50) as presented by Giribet and Ribera (1998). (B) Strict consensus of 432 trees of 7028 steps (CI = 0.24; RI = 0.50) after the re-analysis of the data of Giribet and Ribera (1998). See Giribet and Ribera (1998) for taxon names, taxonomy, and GenBank accession numbers.
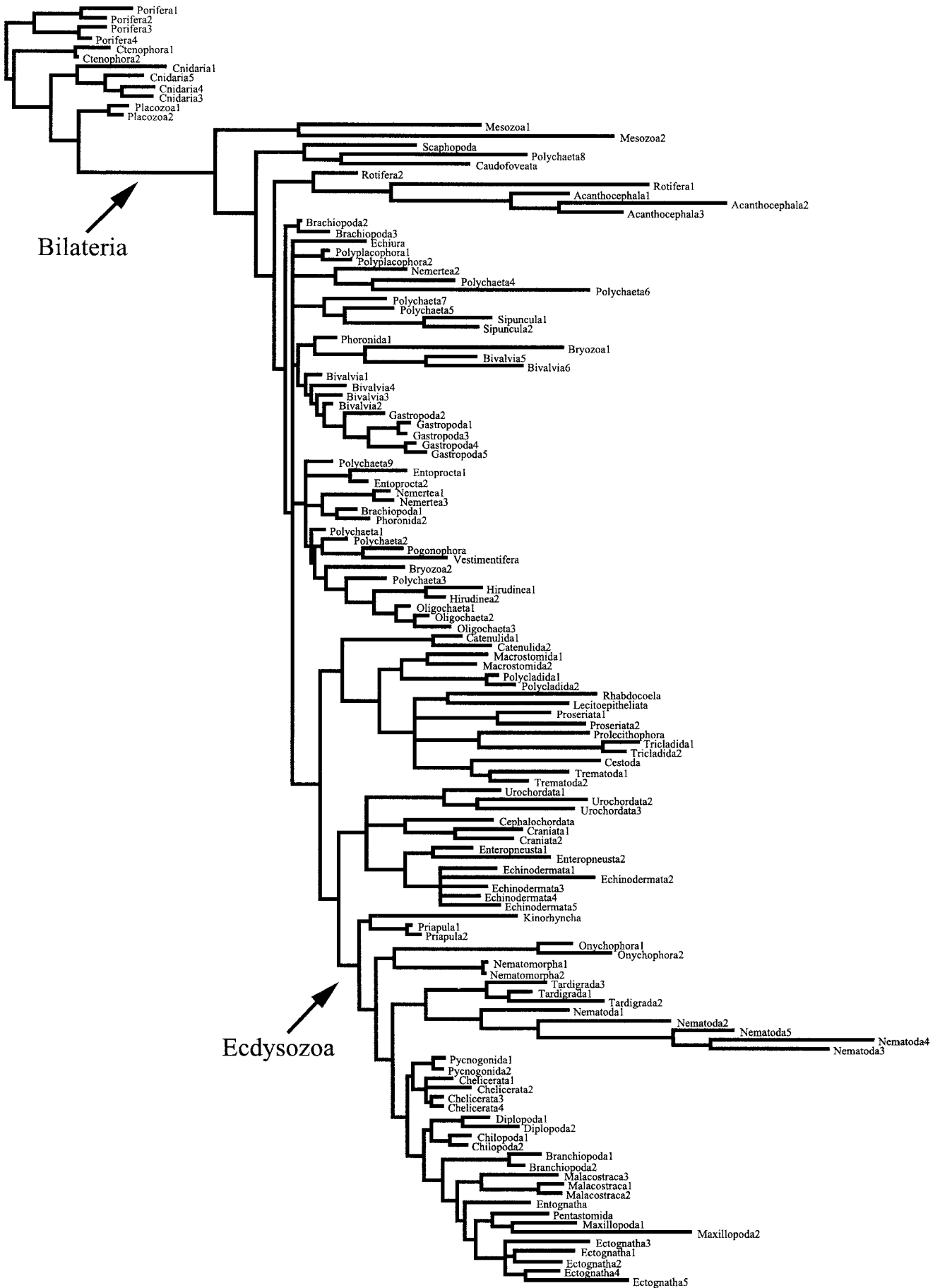
Porifera1
Porifera2
Porifera3
Porifera4
Ctenophora1
Ctenophora2
Cnidaria1
Cnidaria5
Cnidaria4
Cnidaria3
Placozoa1
Placozoa2

Bilateria

Mesozoa1
Mesozoa2
Scaphopoda
Polychaeta8
Caudofoveata
Rotifera2
Rotifera1
Acanthocephala1
Acanthocephala2
Acanthocephala3
Brachiopoda2
Brachiopoda3
Echiura
Polyplacophora1
Polyplacophora2
Nemertea2
Polychaeta4
Polychaeta6
Polychaeta7
Polychaeta5
Sipuncula1
Sipuncula2
Phoronida1
Bryozoa1
Bivalvia5
Bivalvia6
Bivalvia1
Bivalvia4
Bivalvia3
Bivalvia2
Gastropoda2
Gastropoda1
Gastropoda3
Gastropoda4
Gastropoda5
Polychaeta9
Entoprocta1
Entoprocta2
Nemertea1
Nemertea3
Brachiopoda1
Phoronida2
Polychaeta1
Polychaeta2
Pogonophora
Vestimentifera
Bryozoa2
Polychaeta3
Hirudinea1
Hirudinea2
Oligochaeta1
Oligochaeta2
Oligochaeta3
Catenulida1
Catenulida2
Macrostomida1
Macrostomida2
Polycladida1
Polycladida2
Rhabdocoela
Lecitoepitheliata
Proseriata1
Proseriata2
Prolecithophora
Tricladida1
Tricladida2
Cestoda
Trematoda1
Trematoda2
Urochordata1
Urochordata2
Urochordata3
Cephalochordata
Craniata1
Craniata2
Enteropneusta1
Enteropneusta2
Echinodermata1
Echinodermata2
Echinodermata3
Echinodermata4
Echinodermata5
Kinorhyncha
Priapula1
Priapula2
Onychophora1
Onychophora2
Nematomorpha1
Nematomorpha2
Tardigrada3
Tardigrada1
Tardigrada2
Nematoda1
Nematoda2
Nematoda5
Nematoda4
Nematoda3

Ecdysozoa

Pycnogonida1
Pycnogonida2
Chelicerata1
Chelicerata2
Chelicerata3
Chelicerata4
Diplopoda1
Diplopoda2
Chilopoda1
Chilopoda2
Branchiopoda1
Branchiopoda2
Malacostraca3
Malacostraca1
Malacostraca2
Entognatha
Pentastomida
Maxillopoda1
Maxillopoda2
Ectognatha3
Ectognatha1
Ectognatha2
Ectognatha4
Ectognatha5

FIG. 2. Same tree as in Fig. 1B showing branch lengths. Note the length of the branch uniting diploblastic and triploblastic animals.

ships of the four main clades of the Bilateria. It has been suggested that there might be a rooting problem for the Bilateria because the branch separating them from the diploblastic animals is too long and thus may have accumulated too many changes (Eernisse, pers. commun.). Wheeler (1990) has already pointed out that distant outgroups may lead to spurious relationships based on random similarity, and this could be what it is happening within the phylogenetic reconstruction of the Bilateria based solely on 18S rDNA sequences.

ACKNOWLEDGMENTS

REFERENCES

Aguinaldo, A. M. A., Turbeville, J. M., Lindford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A., and Lake, J. A. (1997). Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387:** 489–493.

Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., Duvall, M. R., Price, R. A., Hills, H. G., Qiu, Y.-L., Kron, K. A., Rettig, J. H., Conti, E., Palmer, J. D., Manhart, J. R., Sytsma, K. J., Michaels, H. J., Kress, W. J., Karol, K. G., Clark, W. D., Hedrén, M., Gaut, B. S., Jansen, R. K., Kim, K.-J., Wimpee, C. F., Smith, J. F., Furnier, G. R., Strauss, S. H., Xiang, Q.-Y.,

Plunkett, G. M., Soltis, P. S., Swensen, S. M., Williams, S. E., Gadek, P. A., Quinn, C. J., Eguiarte, L. E., Golenberg, E., Learn, G. H. J., Graham, S. W., Barrett, S. C. H., Dayanandan, S., and Albert, V. A. (1993). Phylogenetics of seed plants: An analysis of nucleic sequences from the plastid gene *rbc*L. *Ann. Mo. Bot. Gard.* **80:** 528–580.

Eernisse, D. J. (1998). Arthropod and annelid relationships re-examined. *In.* "Arthropod Relationships" (R. A. Fortey and R. H. Thomas, Eds.), pp. 43–56. Chapman & Hall, London.

Giribet, G., and Ribera, C. (1998). The position of arthropods in the animal kingdom: A search for a reliable outgroup for internal arthropod phylogeny. *Mol. Phylogenet. Evol.* **9:** 481–488.

Goloboff, P. A. (1993). Nona ver. 1.5.1, American Museum of Natural History, New York.

Goloboff, P. A. (1996). Pee-Wee ver. 2.5.1, American Museum of Natural History, New York.

Goloboff, P. A. (1998). Nona ver. 1.9, American Museum of Natural History, New York.

Littlewood, D. T., Telford, M. J., Clough, K. A., and Rohde, K. (1998). Gnathostomulida—An enigmatic metazoan phylum from both morphological and molecular perspectives. *Mol. Phylogenet. Evol.* **9:** 72–79.

Maddison, D. R. (1991). The discovery and importance of multiple islands of most-parsimonious trees. *Syst. Zool.* **40:** 315–328.

Nixon, K. C. (1998). DADA version 1.4.1. Cornell Univ., Ithaca, NY.

Rice, K. A., Donoghue, M. J., and Olmstead, R. G. (1997). Analyzing large data sets: *rbc*L 500 revisited. *Syst. Biol.* **46:** 554–563.

Swofford, D. L. (1993). PAUP: Phylogenetic Analysis Using Parsimony version 1.3.1, Illinois Natural History Survey, Illinois.

Swofford, D. L. (1998). PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods) version 4, Sinauer, Sunderland, MA.

Wheeler, W. C. (1990). Nucleic acid sequence phylogeny and random outgroups. *Cladistics* **6:** 363–368.

Wheeler, W. C. (1995). Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. *Syst. Biol.* **44:** 321–331.

Wheeler, W. C. (1996). Optimization alignment: The end of multiple sequence alignment in phylogenetics? *Cladistics* **12:** 1–9.