Letters to the Editor

5S Ribosomal RNA Sequences Inappropriate for Phylogenetic Reconstruction[1]

Kenneth M. Halanych
Department of Zoology, University of Texas

Phylogenies of several different metazoan clades were generated from 5S ribosomal sequence data obtained from Erdmann et al. (1985) and Hendriks et al. (1986). The present study utilizes maximum-parsimony methods rather than the phenetic methods which are employed by Hendriks et al. I examined clades at the subkingdom, phyla, and class levels by using bootstrap analysis with a variety of weighting schemes (e.g., equal weighting, transversions weighted heavier than transitions, and stem-region nucleotide bases weighted heavier than loop-region nucleotide bases). In all the clades studied, there was not sufficient phylogenetic signal to reconstruct meaningful phylogenies.

The first clade that I examined consisted of 65 sequences representing a wide range of metazoan taxa (see table 1). With the general heuristic search algorithm of PAUP (phylogenetic analysis using parsimony), version 3.0 (Swofford 1990), I analyzed possible phylogenetic relationships by using the two *Euglena gracilis* sequences as outgroups. Of the 123 nucleotide positions in the metazoan 5S ribosomal gene, only 89 are informative. The search yielded 93 trees of 536 steps, and there were >500 trees of 537 steps. The 50% majority-rule consensus tree is shown in figure 1. The consensus tree includes 76 homoplastic characters which account for 353 homoplastic steps.

Since the placement of several taxa is inconsistent with phylogenies based on morphological and developmental data, I conducted a series of 10 bootstraps with 100 iterations each using representative taxa. A bootstrap using all 65 sequences is prohibitive because of the amount of computation time required. All 10 bootstraps, which used the general heuristic search mode of PAUP, are combinations of seven or eight of the taxa designated by an asterisk in figure 1. In five of the bootstraps the poriferan *Halichondria japonica* is the outgroup, and in the other five the hydrozoan *Spirocodon saltatrix* acts as the outgroup.

The bootstraps did not support any node in any of the 10 combinations at a level $\geq 80\%$ ($P = 0.20$).

To determine whether 5S ribosomal sequence data are useful for phylogenetic reconstruction, I examined several different hierarchical levels, a few of which are discussed below.

I analyzed a deuterostome subset which included two echinoids (*Lytechinus variegatus* and *Hemicentrotus pulcherrimus*), two asteroids (*Asterias vulgaris* and *Asterina pectinifera*), a holothuroid (*Stichopus oshimae*), a urochordate (*Halocynthia roretzi*), a hemichordate (*Saccoglossus kowalevskii*), a vertebrate (*Tinca tinca*), and a brachiopod (*Lingula anatina*) as the outgroup. All possible trees were examined by the exhaustive search mode of PAUP. Two most-parsimonious trees with a length of 86 steps were found.

A bootstrap analysis of the deuterostome clade with 1,000 iterations supported only one node at a significant level; the two echinoids were consistently grouped together

1. Key words: 5S RNA, bootstrap, metazoan phylogeny.

Address for correspondence and reprints: Kenneth M. Halanych, Department of Zoology, University of Texas, Austin, Texas 78712.

Mol. Biol. Evol. 8(2):249–253. 1991.

0737-4030/91/0802-0007$02.00

Table 1
Species Used in All-Metazoa Clade

Species	Taxomonic Designation
Euglena gracilis one	Euglenophyta 1a
Euglena gracilis two	Euglenophyta 1b
Crithidia fasciculata	Flagellata
Paramecium tetraurelia	Ciliata
Dicyema misakiense	Mesozoa
Halichondria japonica	Porifera 1
Haliclona oculata	Porifera 2
Halichondria panicea	Porifera 3
Anthopleura japonica	Anthozoa
Spirocodon saltatrix	Hydrozoa 1
Nemopsis dofleini	Hydrozoa 2
Aurelia aurita one	Scyphozoa 1a
Aurelia aurita two	Scyphozoa 1b
Chrysaora quinquecirrha	Scyphozoa 2
Dugesia japonica Furuyu	Turbellaria 1a
Dugesia japonica Sanage	Turbellaria 1b
Planocera reticulata	Turbellaria 2
Lineus geniculatus	Nemertini 1
Emplectonema gracile one	Nemertini 2a
Emplectonema gracile two	Nemertini 2b
Caenorhabditis elegans	Nematoda 2
Rhabditis tokai	Nematoda 3
Brachionus plicatilis	Rotifera
Perinereis brevicirrus	Polychaeta 1
Sabellastarte japonica	Polychaeta 2
Artemia salina	Branchiopoda 1
Daphnia magna	Branchiopoda 2
Cancer pagurus	Malacostraca 1
Homarus gammarus	Malacostraca 2
Ligia oceanica	Malacostraca 3
Limulus polyphemus	Merostomata
Eurypelma californica	Arachnida 1
Areneus diadematus	Arachnida 3
Spirobolus sp.	Diplopoda
Acyrthosiphon magnoliae	Insecta 1
Bombyx mori	Insecta 2
Calliphora erythrocephala	Insecta 3
Drosophila melanogaster one	Insecta 4a
Drosophila melanogaster two	Insecta 4b
Drosophila melanogaster three	Insecta 4c
Locusta migratoria	Insecta 7
Philosania cynthia-ricini	Insecta 8
Tenebrio molitor	Insecta 9
Urechis unicinctus	Echiura
Calyptogena magnifica	Bivalvia 1
Mytilus edulis	Bivalvia 2
Solemya velum	Bivalvia 3
Octopus vulgaris	Cephalopoda 1
Illex illecebrosus	Cephalopoda 2
Helix pomatia	Gastropoda 1
Arion rufus	Gastropoda 2
Riftia pachyptila	Pogonophora
Phascolopsis gouldii	Sipuncula

Table 1
(*Continued*)

Species	Taxomonic Designation
Lingula anatina	Brachiopoda
Bugula neritina	Bryozoa
Lytechinus variegatus	Echinoidea 1
Hemicentrotus pulcherrimus	Echinoidea 2
Asterias vulgaris	Asteroidea 1
Asterina pectinifera	Asteroidea 2
Stichopus oshimae	Holothuroidea
Saccoglossus kowalevskii one	Hemichordata 1a
Saccoglossus kowalevskii two	Hemichordata 1b
Halocynthia roretzi	Tunicata
Tinca tinca s.	Vertebrata 1a
Tinca tinca o.	Vertebrata 1b

in 97% of the iterations ($P = 0.03$). No other node was supported by bootstrap analysis at a level >50% ($P = 0.50$). These low bootstrap values are due to the lack of informative characters. There are only 29 informative characters in the deuterostome subset when all characters are given equal weighting. Only four characters are not homoplastic, and there are 47 homoplastic steps. Since the number of informative nucleotide positions is small, reconstruction is prone to gross inaccuracy because of the lack of phylogenetic signal (Felsenstein 1988).

To determine whether phylogenetic signal is present, I analyzed the skewness of the tree-frequency distribution for the deuterostome subset, as advocated by Hillis (accepted; also see Fitch 1979; Hillis and Dixon 1989). Analysis of the deuterostome subset yields a tree-frequency distribution that is more skewed than would be expected on the basis of random chance. Thus, some signal appears to be present. However, in this case, that signal either appears to be too weak to be useful or represents a phylogenetic constraint due to convergence and thus is an inaccurate signal.

I also examined the 5S data at the class level by using a mollusk clade. This clade consists of three bivalves (*Calyptogena magnifica, Mytilus edulis,* and *Solemya velum*), two gastropods (*Helix pomatia* and *Arion rufus*), and two cephalopods (*Octopus vulgaris* and *Illex illecebrosus*). A polychaete, *Perineries brevicirrus,* is used as an outgroup. An exhausive search yields one most-parsimonious tree with a length of 41 steps (eight characters and nine steps are homoplastic). Five trees are a step longer. Once again the number of informative characters (21) is small, and the attempted reconstructions are very sensitive to which sequences are used. For example, when both the polychaete *Perineries brevicirrus* and the turbellarian *Dugesia japonica* Furuyu are designated as outgroups, many of the most-parsimonious trees could not maintain a monophyletic ingroup. Presumably, if the phylogenetic signal were stronger, the reconstructions would not be as sensitive to this problem.

This issue of sensitivity was addressed by running a bootstrap (1,000 iterations). Only two groupings, the gastropod clade and the cephalopod clade, are supported at a significance level of $P = 0.08$ (≥92% of the iterations). No other branch is supported in >55% ($P = 0.45$) of the iterations.

The fundamental problem is that the 5S subunit is too small to contain phylogenetic signal sufficient to allow accurate reconstruction of evolutionary relationships. In other words, there are too few informative positions in the 5S subunit. The "deuterostome" clade has the highest character-to-taxa ratio, 3.63:1. The "all-metazoa" clade, which is the most problematic, has a ratio of 1.39:1. As Felsenstein (1988)

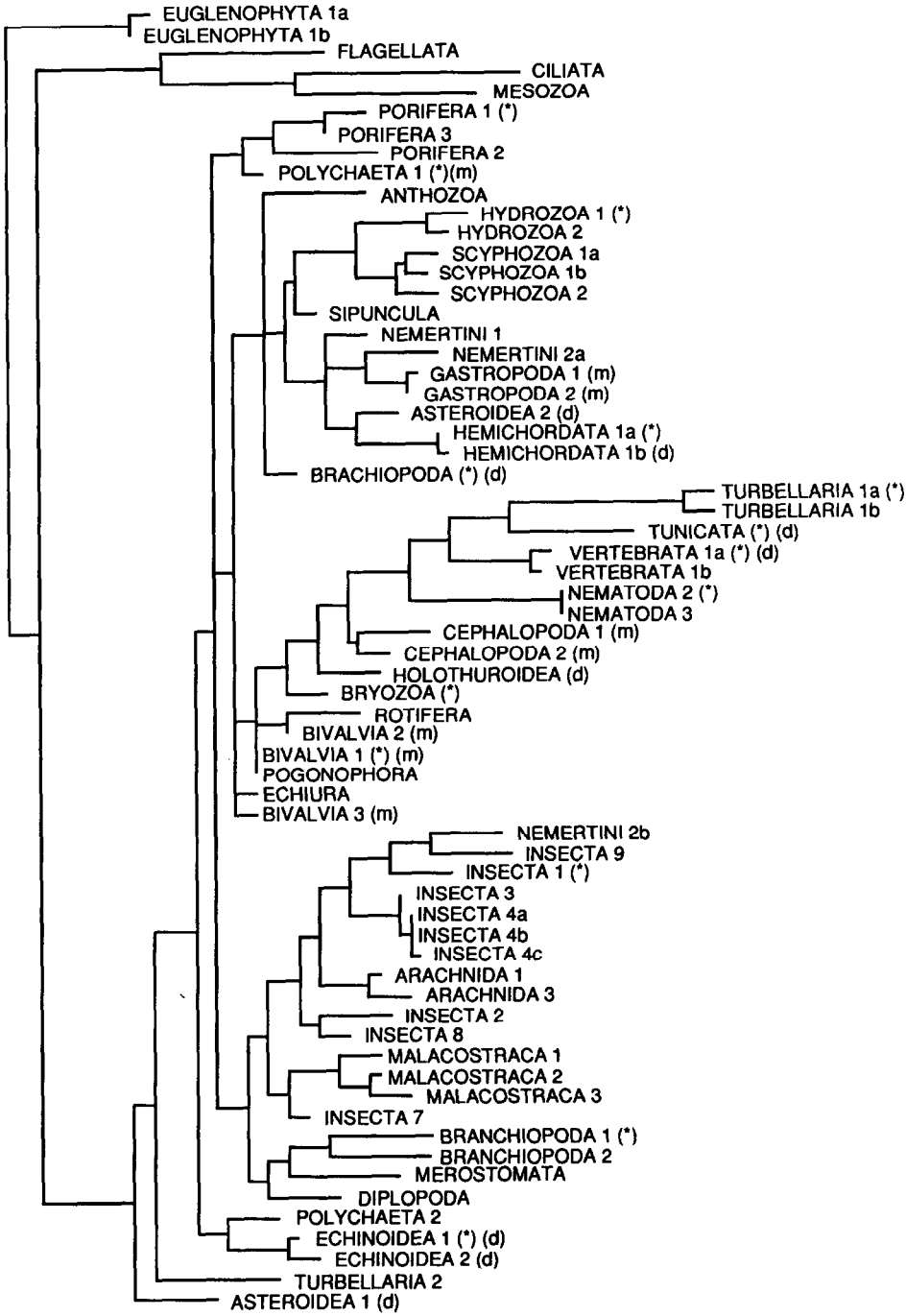

FIG. 1.—Fifty-percent majority-rule consensus tree of 65 metazoan 5S rRNA sequences produced by the heuristic search mode of PAUP. The consensus is a result of 93 trees of 536 steps. The all-metazoa, all-deuterostome, and all-mollusca bootstrap data sets are indicated by an asterisk (*), d, and m, respectively.

argues, these low ratios are not acceptable when one is trying to reconstruct the evolutionary history of a group by using sequence data. Because of its weak phylogenetic signal, the 5S ribosomal subunit alone is not an acceptable source of information for reconstructing phylogenies.

Acknowledgments

I would like to thank C. Pease, T. Engels, and C. Mojica for the constructive comments they contributed to the manuscript. I am especially grateful to B. Crother for his comments and help with the analysis.

LITERATURE CITED

ERDMANN, V. A., J. WOLTERS, E. HUYSMANS, and R. DE WACHTER. 1985. Collection of published 5S, 5.8S, and 4.5S ribosomal RNA sequences. Nucleic Acids Res. **13**:r105–r153.

FELSENSTEIN, J. 1988. Phylogenies from molecular sequences: Inference and reliability. Annu. Rev. Genet. **22**:521–565.

FITCH, W. M. 1979. Cautionary remarks on using gene expression events in parsimony procedures. Syst. Zool. **28**:375–379.

HENDRIKS, L., E. HUYSMANS, A. VANDENBERGHE, and R. DE WACHTER. 1986. Primary structure of the 5S ribosomal RNAs of 11 arthropods and applicability of 5S RNA to the study of metazoan evolution. J. Mol. Evol. **24**:103–109.

HILLIS, D. M. Discriminating between phylogenetic signal and random noise in DNA sequences. *In* M. M. MIYAMOTO and J. CRACRAFT, eds. Phylogenetic analysis of DNA sequences. Oxford University Press, New York (accepted).

HILLIS, D. M., and M. T. DIXON. 1989. Vertebrate phylogeny: evidence from 28S ribosomal DNA sequences. Pp. 355–367 *in* B. FERNHOLM, K. BREMER, and H. JORNVALL, eds. The hierarchy of life. Elsevier Science B.V. (Biomedical Division), New York.

SWOFFORD, D. L. 1990. PAUP, phylogenetic analysis using parsimony, version 3.0. Illinois Natural History Survey, Champaign.

WALTER M. FITCH, reviewing editor

Received August 4, 1990; revision received August 17, 1990

Accepted October 2, 1990