J Mol Evol (1986) 24:103–109

Journal of Molecular Evolution

Primary Structures of the 5S Ribosomal RNAs of 11 Arthropods and Applicability of 5S RNA to the Study of Metazoan Evolution*

L. Hendriks, E. Huysmans, A. Vandenberghe, and R. De Wachter

Department of Biochemistry, University of Antwerp (UIA), Universiteitsplein 1, 2610 Wilrijk, Belgium

Summary. 5S Ribosomal RNA sequences have proven to be useful tools in the study of evolutionary relationships among species. However, in reviewing previously published trees constructed from alignments of metazoan 5S RNAs, we noticed several discrepancies with classical evolutionary views. One such discrepancy concerned the phylum Arthropoda, where a crustacean, *Artemia salina,* seemed to be evolutionarily very remote from four insects. The cause of this phenomenon was studied by determining the 5S RNA sequences of additional arthropods, viz. *Limulus polyphemus, Eurypelma californica, Lasiodora erythrocythara, Areneus diadematus, Daphnia magna, Ligia oceanica, Homarus gammarus, Cancer pagurus, Spirobolus* sp., *Locusta migratoria,* and *Tenebrio molitor.* A tree was then constructed from a dissimilarity matrix by a clustering method known as weighted pair grouping. Application of a correction for unequal evolutionary rates improved the apparent evolutionary position of the arthropods and of some other metazoan species. However, neither the uncorrected nor the corrected tree permitted a completely acceptable reconstruction of metazoan evolution. We presume that this phenomenon is due to random deviations in the evolutionary rate of 5S RNA.

Key words: Evolution — 5S Ribosomal RNA sequences — Arthropoda — Secondary structure

Offprint requests to: R. De Wachter

* Presented at the FEBS Symposium on Genome Organization and Evolution, held in Crete, Greece, September 1–5, 1986

Introduction

The 5S ribosomal RNA sequences of about 350 organisms have been published (Erdmann and Wolters 1986). After alignment of these primary structures, one can construct phylogenetic trees. These phenograms have proven useful in the study of evolutionary relationships among species (De Wachter et al. 1985 and references cited therein; Vandenberghe et al. 1985; Willekens et al. 1986a,b). However, in the Metazoa, several discrepancies between the resulting trees and paleontological, morphological, and embryological views have been noticed (Huysmans et al. 1983). The most striking discrepancy manifests itself in the phylum Arthropoda. A crustacean, *Artemia salina,* appears to be evolutionarily very remote from four insects, viz. *Acyrthosiphon magnoliae, Phylosamia cynthia-ricini, Bombyx mori* and *Drosophila melanogaster.* Indeed, *Artemia salina* seems to branch off before the divergence of vertebrates and invertebrates.

This phenomenon can be explained in several ways: Possibly, 5S rRNA is too small a molecule to allow reliable reconstruction of the evolutionary relationships among species, the relative standard deviation of the number of mutations accumulated per unit of time being too large. This could result in a tree showing false relationships among species. Another possibility is that the apparent evolutionary gap within the arthropods is due to an artefact of the clustering procedure. In this case, inclusion of additional arthropod sequences might alter the clustering order and consequently improve the topology of the tree. A third possibility is that the apparent evolutionary gap between *Artemia salina*

and insects reflects a biological phenomenon such as more rapid accumulation of mutations in crustaceans than in insects. Species with fast evolutionary clocks seem to branch off early in evolutionary-tree reconstructions. The apparent evolutionary distance between *Artemia salina* and insects also might be related to a polyphyletic origin of the arthropods. Certain paleontologists advocate a monophyletic origin of arthropods, i.e., that all arthropods have descended from a hypothetical ancestor that itself was an arthropod (Waterlot 1953). Other authors advocate a polyphyletic origin for the arthropods, i.e., that features that are typical of this phylum were acquired independently in different evolutionary lines arising from segmented worms. According to Manton (1969) and Størmer (1977), at least three such groups can be distinguished. The first group consists of the Myriapoda, Onychophora, and Insecta. The second group is represented by the Chelicerata (Merostomata and Arachnida) and the extinct Trilobita. The Crustacea form the third group.

To test these hypotheses we determined the 5S RNA sequences of 11 more arthropods of various classes. Phylogenetic trees containing these additional sequences were then constructed in several ways.

Materials and Methods

Materials

The horseshoe crab, *Limulus polyphemus* (class Merostomata), and the bird spider *Eurypelma californica* were obtained from the Carolina Biological Supply Company (Burlington, NC, USA). Another bird spider, *Lasiodora erythrocythara,* was kindly donated by Dr. S. Löser (Löbbecke Museum und Aquarium Naturkundliches Heimat Museum, Düsseldorf, FRG). The spider *Areneus diadematus* was collected in the field. The crustacean species *Daphnia magna, Homarus gammarus,* and *Cancer pagurus* were bought in a fish and aquarium shop. *Ligia oceanica* was collected in the tidal zone of the seashore at Wimereux (France). A Diplopoda species, identified tentatively as *Spirobolus* sp., was also supplied by Dr. Löser. The insect species *Tenebrio molitor* and *Locusta migratoria* were obtained from the Antwerp Zoo (Belgium).

Methods

Isolation and Sequencing of 5S RNA. The extraction of nucleic acids from homogenized whole tissue or, in the case of *C. pagurus* and *H. gammarus,* muscle tissue was done according to Bartnik et al. (1981). The purification of 5S RNA was done as described previously (Fang et al. 1982). After 3'-terminal labeling of the molecule with [5'-^{32}P]pCp the sequence was determined by partial chemical degradation (Peattie 1979). The 5'-terminal nucleotide was identified as a nucleoside bisphosphate after total alkaline hydrolysis of unlabeled 5S RNA and high-pressure liquid chromatography (Vandenberghe and De Wachter 1982).

Construction of Phylogenetic Trees. The phylogenetic trees were constructed by a weighted pairwise grouping method using arithmetic averages (WPGMA), starting from a dissimilarity matrix (Sneath and Sokal 1973). The dissimilarity values (D_{ij}) were calculated and corrected for multiple and back mutations as described previously (Willekens et al. 1986b). Figure 3 represents such a tree. Because the evolutionary rate is not constant in different organisms, dissimilarities, corrected for multiple hits and back mutations, can be over- or underestimated. A correction for unequal evolutionary rates can be introduced as follows: First, organisms are divided into two groups, with n_1 and n_2 members. The latter group is a reference group. The mean dissimilarity m_i between organism i of the first group and all n_2 organisms of the reference group is calculated as follows:

$$m_i = \frac{1}{n_2} \sum_{j=1}^{n_2} D_{ij} \tag{1}$$

Next the average m of all m_i values is calculated:

$$m = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} D_{ij} \tag{2}$$

The dissimilarity between organism i and the reference group differs from the mean dissimilarity by

$$d_i = m_i - m \tag{3}$$

The corrected dissimilarity D'_{AB} between species A and B is then calculated as

$$D'_{AB} = D_{AB} - d_A - d_B \tag{4}$$

The reference group can be reduced to a single reference organism if desired. The initial division into two groups can be done by clustering or arbitrarily.

Results and Discussion

Primary Structure

The 5S RNA sequences of 16 arthropods, including the 11 sequences determined in the present work, are aligned in Fig. 1. Length heterogeneity at the 3' terminus was detected for several of the newly sequenced 5S RNAs by polyacrylamide gel electrophoresis of labeled 5S RNA. In such cases the sequence of each component was determined separately.

Secondary Structure

The boxes superimposed upon the sequences aligned in Fig. 1 enclose double-stranded areas of the secondary-structure models described in recent 5S RNA sequence compilations (Erdmann et al. 1985; Erdmann and Wolters 1986). The 5S RNAs of the Chelicerata (the spiders *A. diadematus, E. californica,* and *L. erythrocythara,* and the Merostomata species *L. polyphemus*) apparently have a structure in helix E different from that found in the other arthropod 5S RNAs. The two structural types are exemplified by models for the 5S RNA of *L. migratoria* (Fig. 2a) and *A. diadematus* (Fig. 2b). As demonstrated

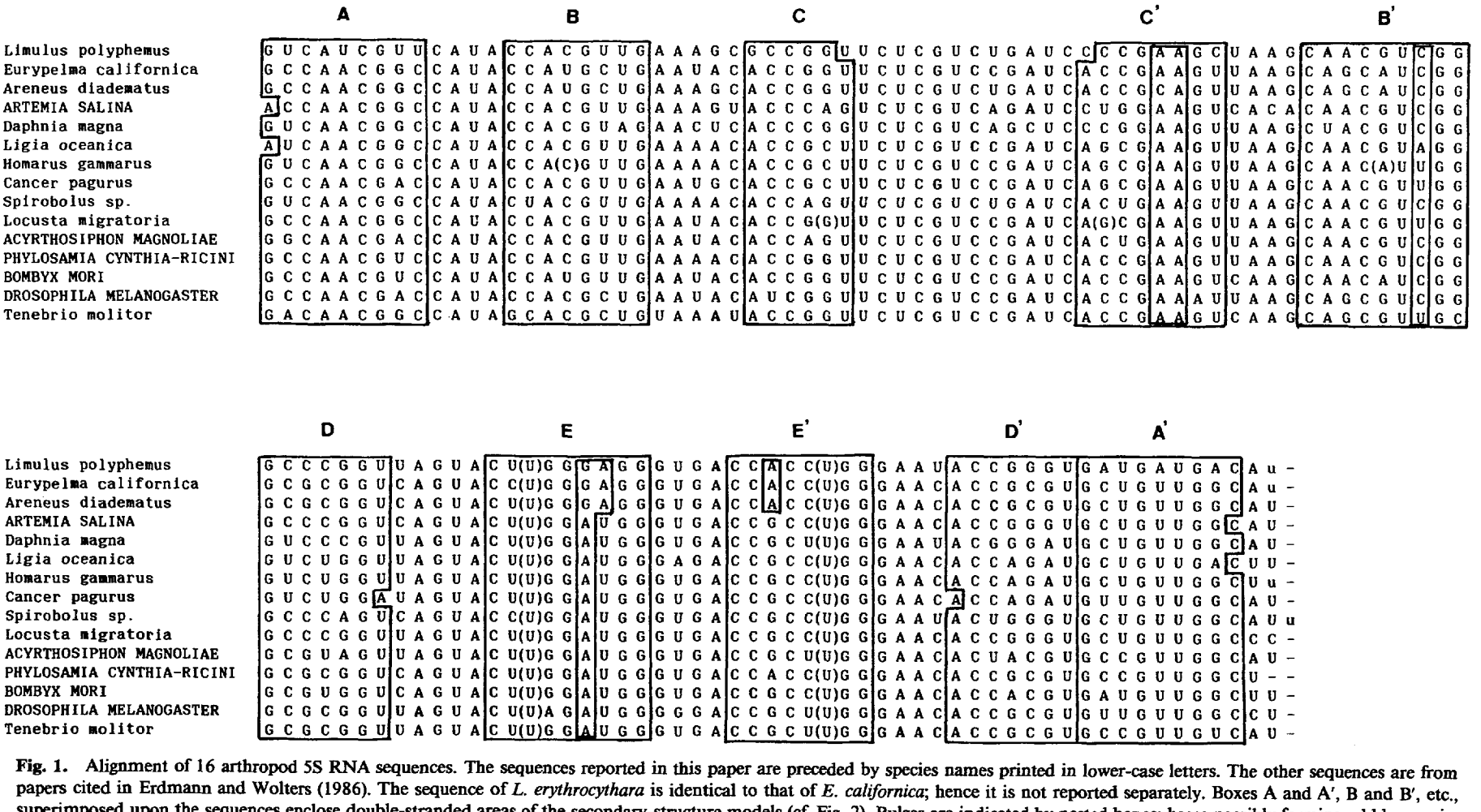

Fig. 1. Alignment of 16 arthropod 5S RNA sequences. The sequences reported in this paper are preceded by species names printed in lower-case letters. The other sequences are from papers cited in Erdmann and Wolters (1986). The sequence of *L. erythrocythara* is identical to that of *E. californica*; hence it is not reported separately. Boxes A and A', B and B', etc., superimposed upon the sequences enclose double-stranded areas of the secondary-structure models (cf. Fig. 2). Bulges are indicated by nested boxes; bases possibly forming odd base pairs (pairs other than GC, AU, and GU) are in parentheses. Length heterogeneity is indicated by terminal residues printed in lower-case letters

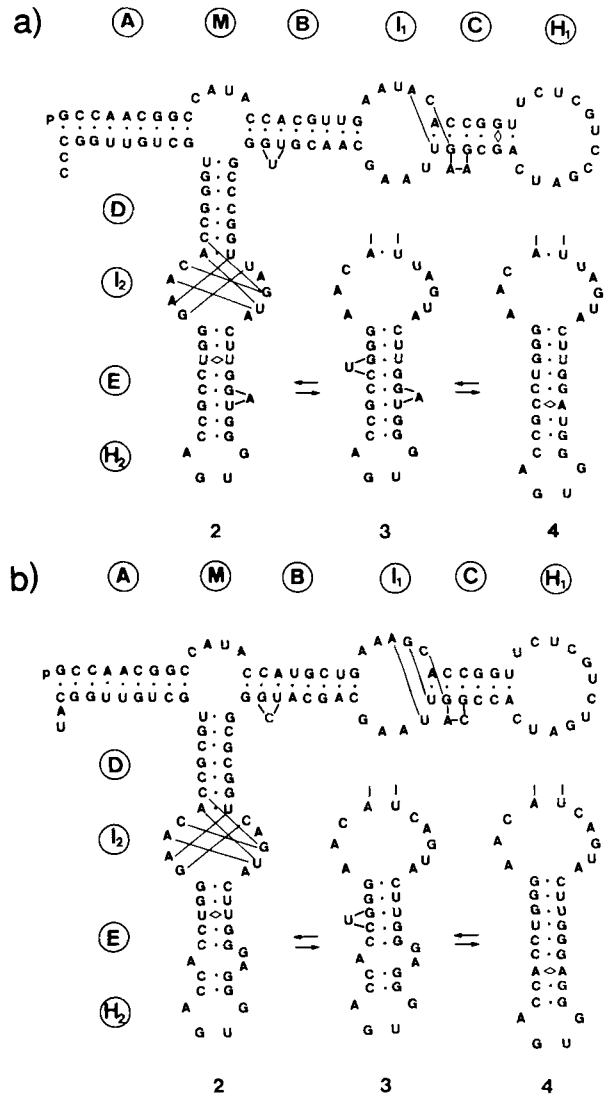

Fig. 2a,b. Secondary structures of 5S rRNAs of the insect *Locusta migratoria* (**a**) and the spider *Areneus diadematus* (**b**). Helices are labeled A–E; loops are labeled M (multibranched), I_1 and I_2 (internal), and H_1 and H_2 (hairpin). Odd base pairs are indicated by a lozenge. As in other metazoan 5S RNAs, three different folding schemes for helix E, labeled 2, 3, and 4 in accordance with De Wachter et al. (1984), can be considered. Formation of base pairs according to the lines drawn in loops I_1 and I_2 would result in alternative secondary structures for these areas also. The possibility of secondary-structure switches in 5S RNAs has been discussed by De Wachter et al. (1984)

previously (De Wachter et al. 1984) it is actually possible to consider three structural forms for helix E of metazoan 5S RNA. The differences between the Chelicerata 5S RNAs and other arthropod 5S RNAs is most apparent if helix E is assumed to adopt form 2 or 3. If helix E is fitted into form 4, the difference is restricted to the position of the odd base pair (AC or AA) within the helix.

Other structural variations are observed in the 5S RNAs of *A. salina* and *L. oceanica,* where helix A is shortened by 1 bp; in *C. pagurus,* where helix D loses one base pair at the expense of internal loop I_2; and in *L. polyphemus,* where helix C is shortened by 1 bp at the expense of hairpin loop H_1.

Molecular Evolution

Figure 3 shows a phylogenetic tree constructed without correction for unequal evolutionary rates from 77 metazoan 5S RNA sequences as described in Methods. The tree, which comprises 15 arthropod sequences, shows hardly any improvement in topology over a previously published tree (Huysmans et al. 1983) that contained only five arthropod sequences. Branchiopoda (*D. magna* and *A. salina*) still seem to diverge very early and seem to be more related to *Halocynthia roretzi* (phylum Chordata) than to other invertebrates. Also, *L. polyphemus* seems evolutionarily very remote from other arthropods, branching off before their divergence from most other invertebrates. Classical evolutionary data (Størmer 1977) pointing to a common origin for all crustaceans are not reflected in the tree of Fig. 3. Although Branchiopoda, Malacostraca (*L. oceanica, C. pagurus,* and *H. gammarus*) and Arachnida each form subclusters, the different arthropod classes are interwoven within the insect cluster, which itself does not show any apparent order. The early divergence of crustaceans seen in the previously published tree based on only five arthropod sequences (Huysmans et al. 1983) cannot be ascribed to a polyphyletic origin of arthropods. Had such an origin occurred, the species belonging to the three evolutionary lines proposed by Størmer (1977) and Manton (1969), viz. Chelicerata, Crustacea, and Myriapoda plus Insecta, should have formed three homogeneous clusters in the tree of Fig. 3, which result is not observed.

To investigate whether the apparent distortions in the tree of Fig. 3 are due to differences in evolutionary rate among species, a tree with a correction for unequal evolutionary rates was constructed (Fig. 4). The correction was based on using the mesozoan *Dicyema misakiense* as the external reference organism. In this tree, arthropods form a rather homogeneous cluster. Relatively close associations appear between members of the Chelicerata (Merostomata and Arachnida) and between members of the Crustacea (Branchiopoda and Malacostraca), as is expected on the basis of classical evolutionary views. The Insecta, however, do not appear to be monophyletic, and the arthropod cluster contains some species belonging to other phyla, namely the Nemertini and Mollusca. As for the other phyla, only the evolutionary position of Porifera species is improved with respect to the uncorrected tree of Fig. 3. The remaining invertebrates included in the tree are scattered without any apparent clustering

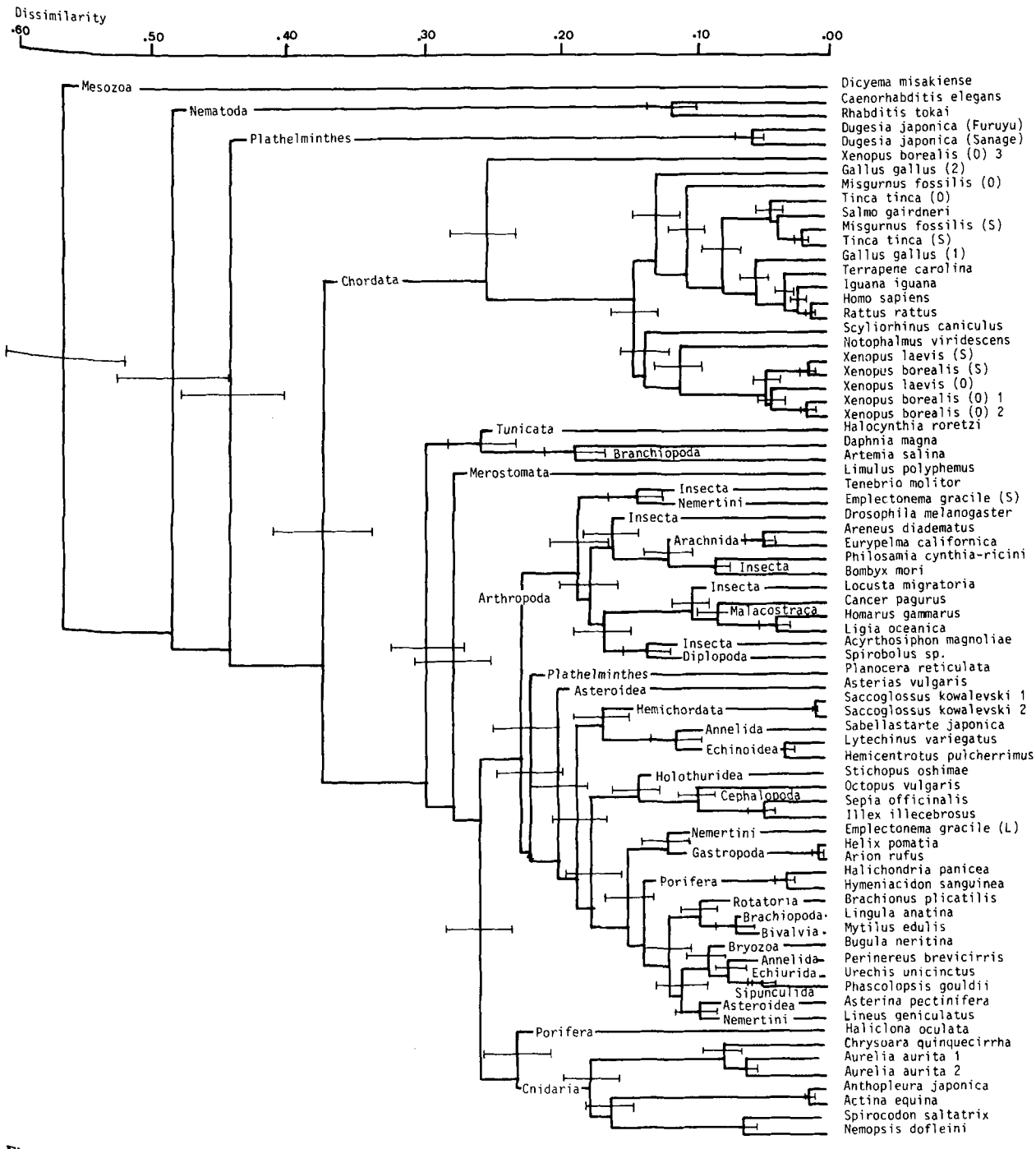

Fig. 3. Evolutionary tree of metazoan species, constructed by a weighted pairwise grouping method as described in Methods. Error bars indicate the standard error of the dissimilarity value as defined by Willekens et al. (1986b). Symbols following certain species names are defined as follows: (O), oocyte; (S), somatic; numbers indicate the presence of different 5S RNAs in an organism, or the fact that different authors reported different sequences

into the 15 phyla to which they belong, namely Mesozoa, Porifera, Cnidaria, Nemathelminthes (Nematoda and Rotatoria), Lophophorata (Bryozoa and Brachiopoda), Pogonophora, Sipunculida, Platyhelminthes, Nemertini, Mollusca (Cephalopoda, Bivalvia, and Gastropoda), Echiurida, Annelida, Echinodermata (Asteroidea, Echinoidea, and Holothuridea), Hemichordata, and Chordata.

Comparison of the uncorrected tree (Fig. 3) with the tree corrected for unequal evolutionary rates (Fig. 4) shows some local improvements of topology in the latter. As an example, the Branchiopoda and *L. polyphemus* occupy positions distant from other arthropods in Fig. 3. This is probably due to high evolutionary rates in these taxa. Hence a correction for evolutionary-rate differences shifts these taxa into

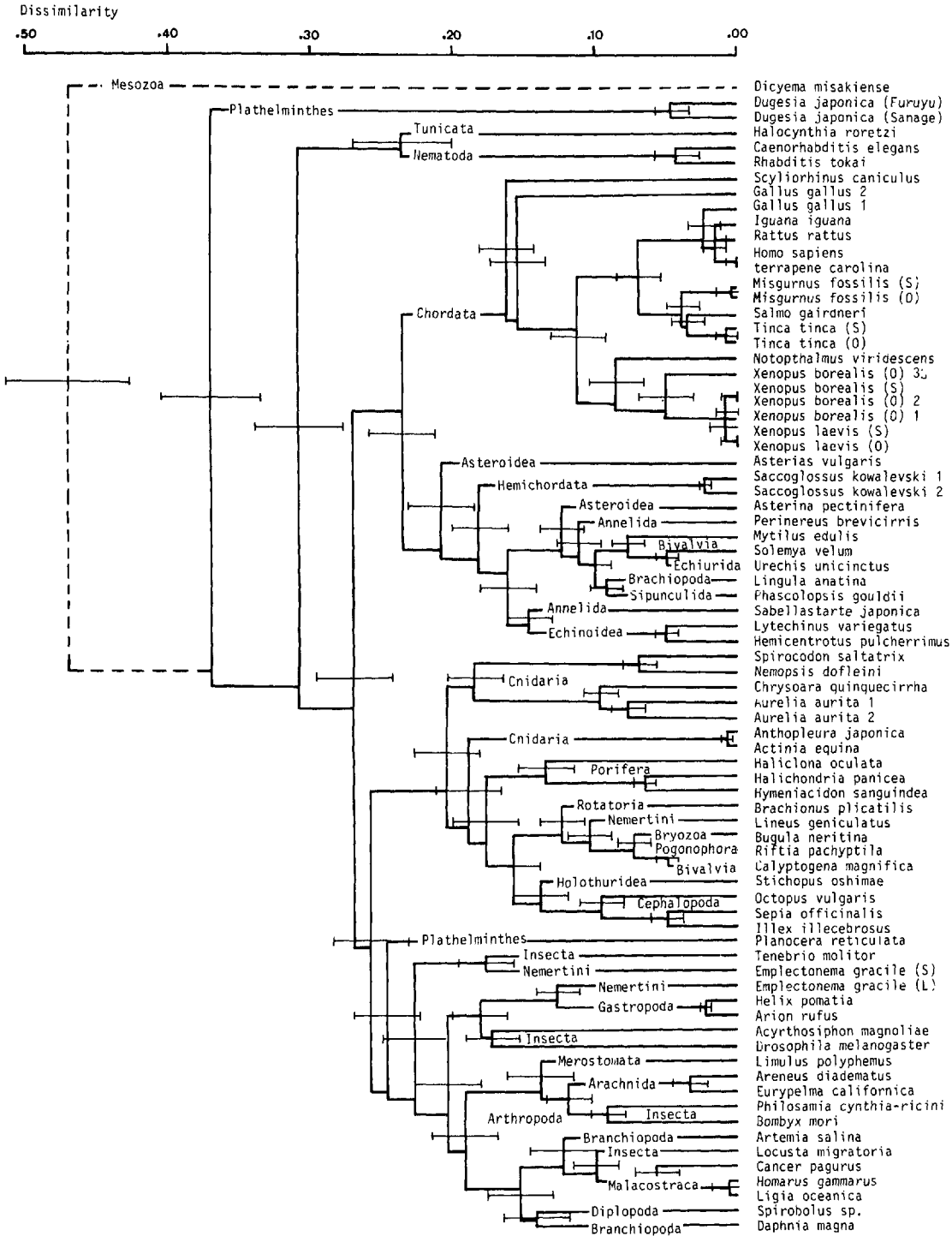

Fig. 4. Evolutionary tree of metazoan species, constructed by a weighted pairwise grouping method involving a correction for unequal evolutionary rates as described in Methods. *D. misakiense* represents the outgroup reference, as indicated by a dashed line. Symbols are as defined in Fig. 3

the arthropod cluster in Fig. 4. However, the improvements thus obtained remain limited and do not bring about a clustering of Metazoa species that is acceptable in the light of classical evolutionary data.

The most obvious conclusion from our study is that random deviations in the evolutionary rate of a molecule the size of 5S RNA are too large to permit accurate reconstitution of the evolutionary record, at least on the time scale of metazoan evolution. Hence it may be necessary to resort to using larger molecules, such as small-ribosomal-subunit RNAs, as molecular clocks for the study of this type of problem.

Acknowledgments. L. Hendriks holds an IWONL scholarship. E. Huysmans was a FKFO research assistant. We thank Dr. Löser for gifts of bird spiders and millipedes. Our research was supported in part by a FKFO grant.

References

Bartnik E, Strugula K, Stepien PP (1981) Cloning and analysis of recombinant plasmids containing genes for *Anacystis nidulans* 5S ribosomal RNA. Curr Genet 4:173–176

De Wachter R, Chen M-W, Vandenberghe A (1984) Equilibria in 5S ribosomal RNA secondary structure. Eur J Biochem 143:175–182

De Wachter R, Huysmans E, Vandenberghe A (1985) 5S ribosomal RNA as a tool for studying evolution. In: Schleifer KH, Stackebrandt E (eds) Evolution of prokaryotes. Academic Press, London, pp 115–141

Erdmann VA, Wolters J (1986) Collection of published 5S, 5.8S and 4.5S ribosomal RNA sequences. Nucleic Acids Res 14: r1–r59

Erdmann VA, Wolters J, Huysmans E, De Wachter R (1985) Collection of published 5S, 5.8S and 4.5S ribosomal RNA sequences. Nucleic Acids Res 13:r105–r153

Fang B-L, De Baere R, Vandenberghe A, De Wachter R (1982) Sequences of three molluscan 5S ribosomal RNAs confirm the validity of a dynamic secondary structure model. Nucleic Acids Res 10:4679–4685

Huysmans E, Dams E, Vandenberghe A, De Wachter R (1983) The nucleotide sequences of the 5S ribosomal RNAs of four mushrooms and their use in studying the phylogenetic position of basidiomycetes among the eukaryotes. Nucleic Acids Res 11:2871–2879

Manton SM (1969) Introduction to classification of arthropods. In: Moore RC (ed) Treatise on invertebrate paleontology, part R. Geological Society of America and University of Kansas Press, pp 3–15

Peattie DA (1979) Direct chemical method for sequencing RNA. Proc Natl Acad Sci USA 76:1760–1764

Sneath PHA, Sokal RR (1973) Numerical taxonomy. WH Freeman, San Francisco

Størmer L (1977) Arthropoda—general features. In: Moore RC (ed) Treatise on invertebrate paleontology, part O. Geological Society of America and University of Kansas Press, pp 3–16

Vandenberghe A, De Wachter R (1982) High pressure liquid chromatography determination of the 5'-terminal residue of small RNA molecules. J Liq Chromatogr 5:2079–2084

Vandenberghe A, Wassink A, Raeymaekers P, De Baere R, Huysmans E, De Wachter R (1985) Nucleotide sequence, secondary structure and evolution of the 5S ribosomal RNA from five bacterial species. Eur J Biochem 149:537–542

Waterlot G (1953) Généralités sur les arthropodes. In: Piveteau J (ed) Traité de paléontologie, tome III. Masson et Cie, Paris, pp 11–24

Willekens P, Huysmans E, Vandenberghe A, De Wachter R (1986a) Archaebacterial 5S ribosomal RNA: nucleotide sequences in two methanogen species, secondary structure models and molecular evolution. System Appl Microbiol 7:151–159

Willekens P, Stetter KO, Vandenberghe A, Huysmans E, De Wachter R (1986b) The structure of 5S ribosomal RNA in the methanogenic archaebacteria *Methanolobus tindarius* and *Methanococcus thermolithotrophicus*. FEBS Lett 204:273–278