Small Ribosomal Subunit RNA and the Phylogeny of Mollusca

Birgitta Winnepenninckx
Departement Biochemie
Universiteit Antwerpen (U.I.A.)
Universiteitsplein 1,
B–2610 Antwerpen, Belgium

Thierry Backeljau
Afdeling Malacologie
Koninklijk Belgisch Instituut voor
Natuurwetenschappen
Vautierstraat 29
B–1040 Brussel, Belgium

Rupert De Wachter[1]
Departement Biochemie
Universiteit Antwerpen (U.I.A.)
Universiteitsplein 1,
B–2610 Antwerpen, Belgium

ABSTRACT

We determined the complete sequence of the small ribosomal subunit RNA of the pulmonate snail *Onchidella celtica*. This sequence and the one recently determined for the chiton *Acanthopleura japonica* were added to an alignment of 25 18S rRNA sequences of Metazoa, including three other Mollusca. The data set was used to assess certain aspects of molluscan phylogeny by distance matrix and character state methods. The trees obtained were tested for effects of random and systematic errors. The results of our analyses support: (a) molluscan monophyly; (b) gastropod monophyly; (c) bivalve monophyly; (d) a sister group relationship of Gastropoda and Polyplacophora. The position of the phylum among other Metazoa remains uncertain due to a lack of representatives of many invertebrate phyla in our data set. Most of our results are congruent with existing hypotheses.

Key Words: 18S rRNA, phylogeny, Metazoa, Gastropoda, Bivalvia, Polyplacophora, *Onchidella celtica*.

INTRODUCTION

Historical Background

Many aspects of molluscan phylogeny are still uncertain. The huge phenotypic diversity within the phylum obscures the evolutionary relationships between the larger molluscan taxa (*e.g.* von Ihering, 1876; Milburn, 1960; von Salvini-Plawen, 1969, 1972, 1985, 1990a,b; Stasek, 1972; Götting, 1980; Wingstrand, 1985; Scheltema, 1988; Brusca & Brusca, 1990). Nevertheless, it is generally accepted that the "shell- bearing molluscs" (Conchifera, *i.e.* Cephalopoda, Scaphopoda, Bivalvia, Gastropoda and Monoplacophora) are monophyletic with Polyplacophora as sister group (*e.g.* von Salvini-Plawen, 1969, 1985, 1990a; Stasek, 1972; Götting, 1980; Wingstrand, 1985; Scheltema, 1988; Brusca & Brusca, 1990). Indeed, the loss of spicules and the presence of three mantle margin folds, an univalve shell consisting basically of three layers, jaws, a head with cerebrally innervated appendages, a nervous system differentiated in axons and ganglia, a crystalline style and statocysts are considered to be synapomorphies uniting the five conchiferan classes (*e.g.* Götting, 1980; von Salvini-Plawen, 1985; Wingstrand, 1985; Brusca & Brusca, 1990). However, different interpretations exist about the phylogenetic relationships within this subphylum. Milburn (1960) suggested three conchiferan clades: Monoplacophora, Cephalopoda and a Bivalvia-Gastropoda-Scaphopoda clade. The branching pattern of these three groups and the topology of the Bivalvia-Gastropoda-Scaphopoda clade, remains unresolved. Götting (1980) proposed a Bivalvia-Scaphopoda sister relationship and relied on the shell structure and form of the larval shell to conclude that Gastropoda and Monoplacophora are sister groups. Cephalopoda is then a sister group to the other four conchiferan classes. However, Wingstrand (1985), Brusca and Brusca (1990) and von Salvini-Plawen (1985, 1990a) considered "Monoplacophora" (*i.e.* class Tergomya *sensu* Peel, 1991 or Tryblidiida *sensu* Wingstrand, 1985; see Peel, 1991 for a discussion) as a sister group to the four other conchiferan classes, which in turn consist of a Bivalvia-Scaphopoda clade and a Gastropoda-Cephalopoda clade. The former is characterized by the presence of a mantle surrounding the entire body, reduction of the head and a laterally compressed form. The latter is determined by the presence of a well developed head, dorsoventral elongation, dorsal concentration of the viscera and shell coiling (Brusca & Brusca, 1990). Except for the position of the "Monoplacophora", this view agrees well with the division of the Conchifera into the clades Diasoma (classes Bivalvia, Scaphopoda and the fossil Rostroconchia) and Cyrtosoma (classes "Monoplacophora", Gastropoda and Cephalopoda), which is widely accepted among paleontologists (Runnegar & Pojeta, 1974, 1985; Pojeta, 1980; Steiner, 1992). Yet, Peel (1991) recently suggested that Cyrtosoma and Diasoma are both polyphyletic.

The oldest fossil molluscs date from 570 MYA (*e.g.* Runnegar & Pojeta, 1974, 1985; Valentine, 1980), near the Precambrian-Cambrian boundary. This period was marked by an explosive radiation of animals resulting in the appearance of most extant invertebrate phyla (*e.g.*

---

[1] Author for correspondence.

Table 1. List of 17 oligonucleotides complementary to conserved regions in eukaryotic 18S rRNA genes. These were used to determine the sequence of both strands of the 18S rRNA gene of *Onchidella celtica*.

Sequence[1]	Strand[2]	Corresponding position in the 18S rRNA gene of *Onchildella celtica*
CTGGTTGATYCTGCCAGT	R	4–21
GAAACTGCGAATGGCTCATT	R	82–101
AATGAGCCATTCGCAGTTTC	C	101–82
AGGGYTCGAYYCCGGAGA	R	393–410
TCTCCGGRRTCGARCCCT	C	410–393
TCTCAGGCTCCYTCTCCGG	C	422–404
ATTACCGCGGCTGCTGGC	C	605–588
CGCGGTAATTCCAGCTCCA	R	597–615
TTGGYRAATGCTTTCGC	C	990–974
TTRATCAAGAACGAAAGT	R	1002–1019
CCGTCAATTYYTTTRAGTTT	C	1188–1169
AATTTGACTCAACACGGG	R	1221–1238
GGGCATCACAGACCTGTTAT	C	1479–1460
ATAACAGGTCTGTGATGCCC	R	1460–1479
TTTGYACACACCGCCCGTCG	R	1666–1685
GACGGGCGGTGTGTRC	C	1684–1669
CYGCAGGTTCACCTACRG	C	1833–1816

[1] Sequence positions where both purines (A and G) are present are indicated by "R", those where both pyrimidines (C and T) are present by "Y".

[2] Oligonucleotides with a sequence corresponding to that of the RNA-like strand are indicated by a "R", those whose sequence is complementary to it, by a "C".

Bergström, 1991; Erwin, 1991; Valentine, 1991). Several aspects of the metazoan branching pattern still remain confused due to the doubtful homology of the relatively few morphological, anatomical and embryological characters shared by different phyla (*e.g.* Nielsen, 1977; Anderson, 1981; Inglis, 1985; Bergström, 1986; Ax, 1989; Schram, 1991; Backeljau *et al.*, 1993). Nevertheless, Mollusca appear to be a monophyletic group belonging to the Spiralia (*i.e.* Platyhelminthes, Nemertini, Mollusca, Sipuncula, Echiura and Annelida, and probably Gnathostomulida and Entoprocta) (*e.g.* Wingstrand, 1985; Brusca & Brusca, 1990; Willmer, 1990), but no synapomorphies are known linking the Mollusca unambiguously to any other spiralian phylum (*e.g.* Wingstrand, 1985; Erwin, 1991). Some authors (*e.g.* von Salvini-Plawen, 1990a) suggest a sister group relationship to Turbellaria (Platyhelminthes) considering the flat, often ciliated, ventral creeping foot as a synapomorphy relating both phyla. Many others however, include the Mollusca in the protostome clade (*e.g.* Wingstrand, 1985; Brusca & Brusca, 1990; Willmer, 1990; Schram, 1991).

Biochemical and molecular characters have been introduced as an independent source of phylogenetic information. A serological study of molluscs, echinoderms, annelids and arthropods suggested that Mollusca are most closely related to Annelida (Wilhelmi, 1944). Lyddiatt *et al.* (1978) used cytochrome *c* amino acid sequence data to deduce a sister group relationship between molluscs and echinoderms. In studies using 5S ribosomal RNA (rRNA) sequences (Ohama *et al.*, 1984; Hendriks *et al.*, 1986; Hori & Osawa, 1987), the Mollusca (represented by Bivalvia, Gastropoda and Cephalopoda) appeared as a polyphyletic group. From the analysis of Lenaers and Bhaud (1992) on the basis of partial sequences of 28S rRNA, Mollusca (represented by *Mytilus edulis*) appeared to be a sister group to Annelida. Holland *et al.* (1991) used partial small subunit (SSU) rRNA (18S rRNA) sequences and suggested that Mollusca (represented by *Mytilus edulis*) and Arthropoda are sister taxa. On the basis of mitochondrial SSU rRNA sequences the Mollusca, represented by a prosobranch and a chiton, appeared as sister group to the Annelida or as a paraphyletic group including the latter phylum (Ballard *et al.*, 1992) . In all these studies, however, the data sets were too limited to allow reliable conclusions. Field *et al.* (1988) determined partial sequences of SSU rRNA from representatives of ten different metazoan phyla including four Mollusca, viz. an opisthobranch gastropod, two bivalves and a chiton. Yet, different phylogeny inference methods yielded contradictory results. Field *et al.* (1988; see also Raff *et al.*, 1989) used a distance method to conclude that Mollusca form a clade with Annelida, Sipuncula, Brachiopoda and Pogonophora. However, the relationships between the five groups were not resolved. Ghiselin (1988, 1989) reanalyzed this data set with a "signature" approach and concluded that molluscs are a sister group to the Annelida *sensu lato* (*i.e.* Annelida *sensu strictu*, Brachiopoda, Pogonophora and Sipuncula). A maximum parsimony analysis of the same data produced a similar clade containing Sipuncula, Pogonophora, Brachiopoda, Annelida and Mollusca (Patterson, 1989) but with the latter two phyla not being monophyletic. Lake (1989), who applied evolutionary parsimony, also concluded that Mollusca are paraphyletic.

In a preliminary attempt, we use complete SSU rRNA sequences to assess molluscan phylogeny. We consider sequences to be complete if (1) the sequence of the entire 18S rRNA molecule is known or (2) if only a total number on the order of 50 nucleotides at the 5′ and 3′ terminal parts are missing because they are used as PCR primer annealing sites (*e.g.* Rice, 1990; Littlewood, 1991). Hitherto, complete 18S rRNA sequences of only three molluscan species viz., the bivalves *Placopecten magellanicus* and *Crassostrea virginica* and the gastropod *Limicolaria kambeul*, have been published respectively by Rice (1990), Littlewood (1991) and Winnepenninckx *et al.* (1992) . In this paper we present the complete 18S rRNA sequence of the gymnomorphan snail *Onchidella celtica* (Cuvier, 1817). A fifth molluscan sequence (Winnepenninckx *et al.* , 1993), that of the chiton *Acanthopleura japonica* (Lischke, 1873) is also included.

Small ribosomal subunit RNA sequences

SSU rRNA sequences combine several features that make them appropriate for phylogenetic studies (Raff *et al.*,

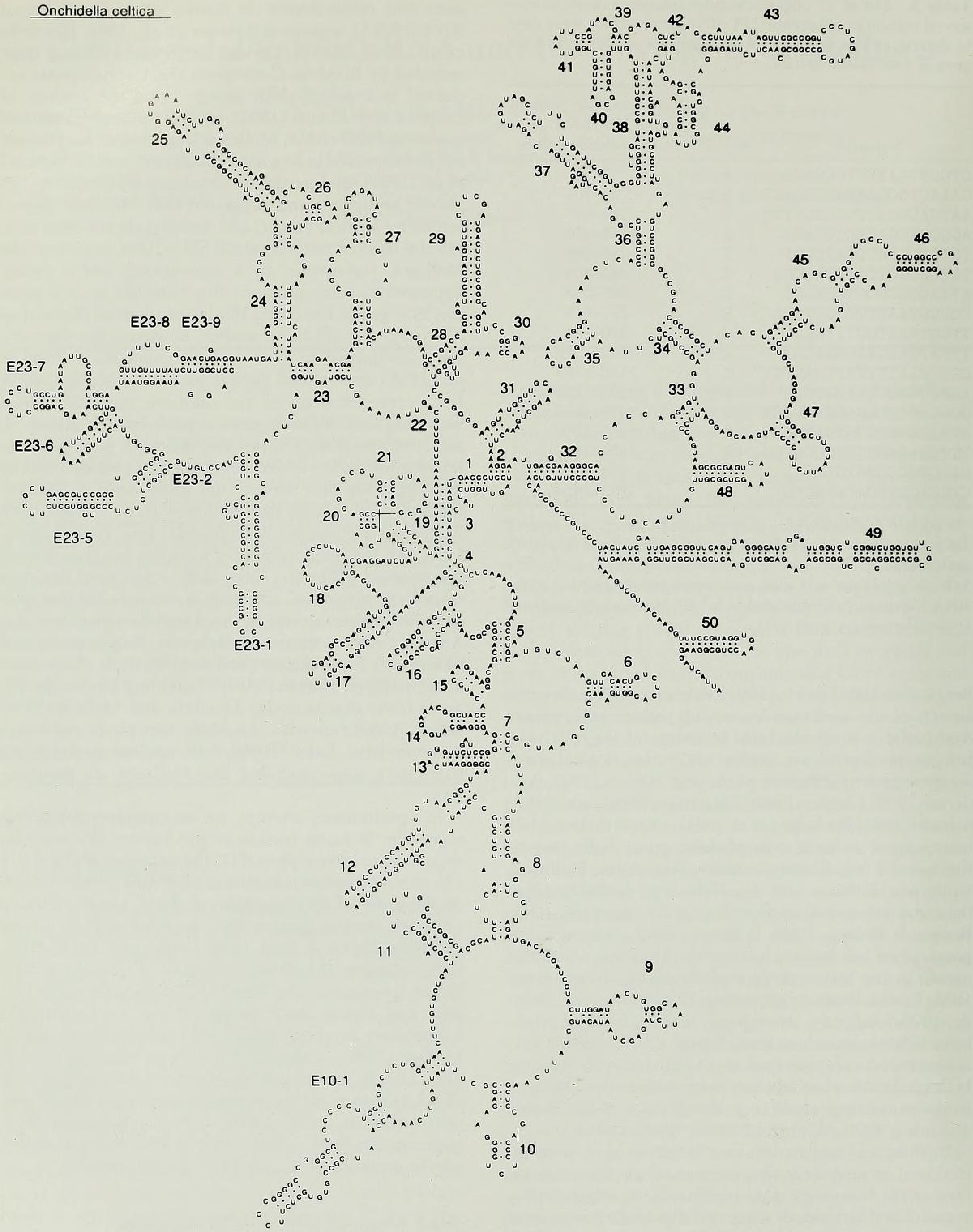

Figure 1. Secondary structure model for the 18S rRNA of *Onchidella celtica*. Helix numbering from helix 19 onward has changed with respect to the numbering used by De Rijk *et al.* (1992), due to the discovery of a tertiary structure interaction in helix 19 (Woese & Gutell, 1989).

Table 2. The 18S rRNA sequence of the gastropod *Onchidella celtica.*

0001	UAUCUGGUUGAUCCUGCCAGUAGUCAUAUGCUUGUCUCAAAGAUUAAGCC
0051	AUGCAUGUCUAAGUUCACACUGUCUCACGGUGAAACCGCGAAUGGCUCAU
0101	UAAAUCAGUCGAGGUUCCUUAGAUGACACGAUCCUACUUGGAUAACUGUG
0151	GCAAUUCUAGAGCUAAUACAUGCUAUUCAAGCUCCGACCCUCUGGGGAAG
0201	AGCGCUUUUAUUAGUUCAAAACCAAUCGCCGUGUGCUCUUCCCGGGGCCG
0251	GGCGUCCCCCUUGGUGACUCUGGAUAACUUUGUGCUGAUCGCAUGGCCUU
0301	UUGCGCCGGCGACGCAUCUUUCAAAUGUCUGCCCUAUUAAAUGCGAUGGU
0351	ACGUGAUAUGCCUACCAUGUUUGUAACGGGUAACGGGGAAUCAGGGUUCG
0401	AUUCCGGAGAGGGAGCAUGAGAAACGGCUACCACAUCCAAGGAAGGCAGC
0451	AGGCGCGCAACUUACCCACUCCCGGCACGGGGAGGUAGUGACGAAAAAUA
0501	ACAAUACGGGACUCUUUCGAGGCCCAGUAAUUGGAAUGAGUACACUUUAA
0551	ACCCUUUAACGAGGAUCUAUUGGAGGGCAAGUCUGGUGCCAGCAGCCGCG
0601	GUAAUUCCAGCUCCAAUAGCGUAUAUUAAAGUUGUUGCAGUUAAAAAGCU
0651	CGUAGUUGGAUCUCAGGCGCAGGCGGGCGGUCCGGCUCGCGCCGCUCACU
0701	GCCCGUUGUCUCCUGCCCUACCUGUUGCCGGCUCUCUCCCGUGGGUGCUC
0751	UUCGCUGAGCGUCCGGGUGGCCGGCGCGUUUACUUUGAAAAAAUUAGAGU
0801	GUUCAAAGCAGGCCUCGCCUGCCUGAAUAAUUGCGCAUGGAAUAAUGGAA
0851	UAGGACCUCGGUUCUAUUUUGUUGGUUUUCGGAACUGGAGGUAAUGAUUA
0901	ACAGGGACAAACGGGGGGAUUCGUAUUGCGGCGUUAGAGGUGAAAUUCUU
0951	GGAUCGCCGCAAGACGAGCUACUGCGAAAGCAUUUGUCAAGAAUGUUUUC
1001	AUUAAUCAAGAACGAAAGUCAGAGGCGAGAAGACGAUCAGAUACCGUCGU
1051	AGUUCUGACCAUAAACGAUGCCGACCAGCGAUCCGCAGGAGUUGCUUCGA
1101	UGACUCUGCGGGCAGCUUCCGGGAAACCAAAGUGUUUGGGUUCCGGGGGA
1151	AGUAUGGUUGCAAAGCUGAAACUUAAAGGAAUUGACGGAAGGGCACCACC
1201	AGGAGUGGAGCCUGCUGCUUAAUUUGACUCAACACGGGAAAACUCACCCG
1251	GUCCGGACACUGUAAGGAUUGACAGAUUGAUAGCUCUUUCUUGAUUCGGU
1301	GGGUGGUGGUGCAUGGCCGUUCUUAGUUGGUGGAGCGAUUUGUCUGGUUA
1351	AUUCCGAUAACGAACGAGACUCUAGCCUAUUAAAUAGUUCGCCGGUCCCU
1401	CGAUGCGCCGGCGCAACUUCUUAGAGGGACGAGUGGCGUUUAGCCAACGA
1451	GAUUGAGCAAUAACAGGUCUGUGAUGCCCUUAGAUGUCCGGGGCCGCACG
1501	CGCGCUACACUGAAGGAAUCAGCGUGGAUGCCUCCCUGGCCCGAAAGGCU
1551	GGGAAACCCGUUGAAUCUCCUUCGUGCUAGGGAUUGGGGCUUGUAAUUCU
1601	UCCCCAUGAACGAGGAAUUCCCAGUAAGCGCGAGUCAUAAGCUCGCGUUG
1651	AUUACGUCCCUGCCCUUUGUACACACCGCCCGUCGCUACUAUCGAUUGAG
1701	CGGUUCAGUGAGGGCAUCGGAUUGGUCUCGGUCUGGUGUUCGCGCACCGG
1751	CACCGCUGGCCGAGAAGACGCUCGAACUCGAUCGCUUGGAGAAAGUAAAA
1801	GUCGUAACAAGGUUUCCGUAGGUGAACCUGCGGAAGGAUCAUUA

1989; Hillis & Dixon, 1991; Solignac *et al.*, 1991; Woese, 1991): (1) universality; (2) constancy of function; (3) alternation of conserved regions with variable ones, allowing phylogenetic studies at a broad range of taxonomical levels; (4) presence of conservative regions that allow the design of "universal" primers; (5) a conservative secondary structure facilitating the identification of homologous positions in regions with little sequence similarity; (6) apparent absence of lateral gene transfer; (7) a large information content (1800–1900 bp) (8) intraspecific sequence homogeneity among different gene copies (Gerbi, 1985; Dover, 1986).

Gene cloning

Much sequence information on rRNAs has been obtained by direct RNA sequencing using reverse transcriptase (Lane *et al.*, 1985; Solignac *et al.*, 1991) or by direct sequencing of the rRNA genes after PCR amplification (Saiki *et al.*, 1988). Both techniques are very rapid. Yet we prefer to clone and sequence the 18S rRNA genes, for direct RNA sequencing has some disadvantages: (1) RNA is less stable than DNA; (2) subsequent checking of sequences is not possible; (3) sequencing of regions with strong secondary structure is difficult; (4) reverse transcriptase has a rather high error frequency; (5) only one strand is available and thus two-strand verification is not possible. All this results in an overall error rate of about 1% (Lane *et al.*, 1985). Although PCR amplification eliminates a great deal of these problems, it also has some drawbacks (Hillis & Dixon, 1991): (1) *Taq* polymerase has a high error rate, viz. $\approx 2 \times 10^{-4}$ to $< 1 \times 10^{-5}$ according to Eckert and Kunkel (1991) and 2.75 10^{-3} according to Bej *et al.* (1991); (2) the 3′ and 5′ parts of the gene itself have to be used as primer annealing sites, if the sequence of the adjacent regions is unknown; (3) direct sequencing of PCR amplified fragments is difficult (*e.g.* Gyllenstein, 1989); (4) the product is afterwards not available to others for verification. By cloning the PCR product prior to sequencing, the latter two problems can be overcome, but the sequencing of numerous clones is necessary to avoid an enhancement of the error rate

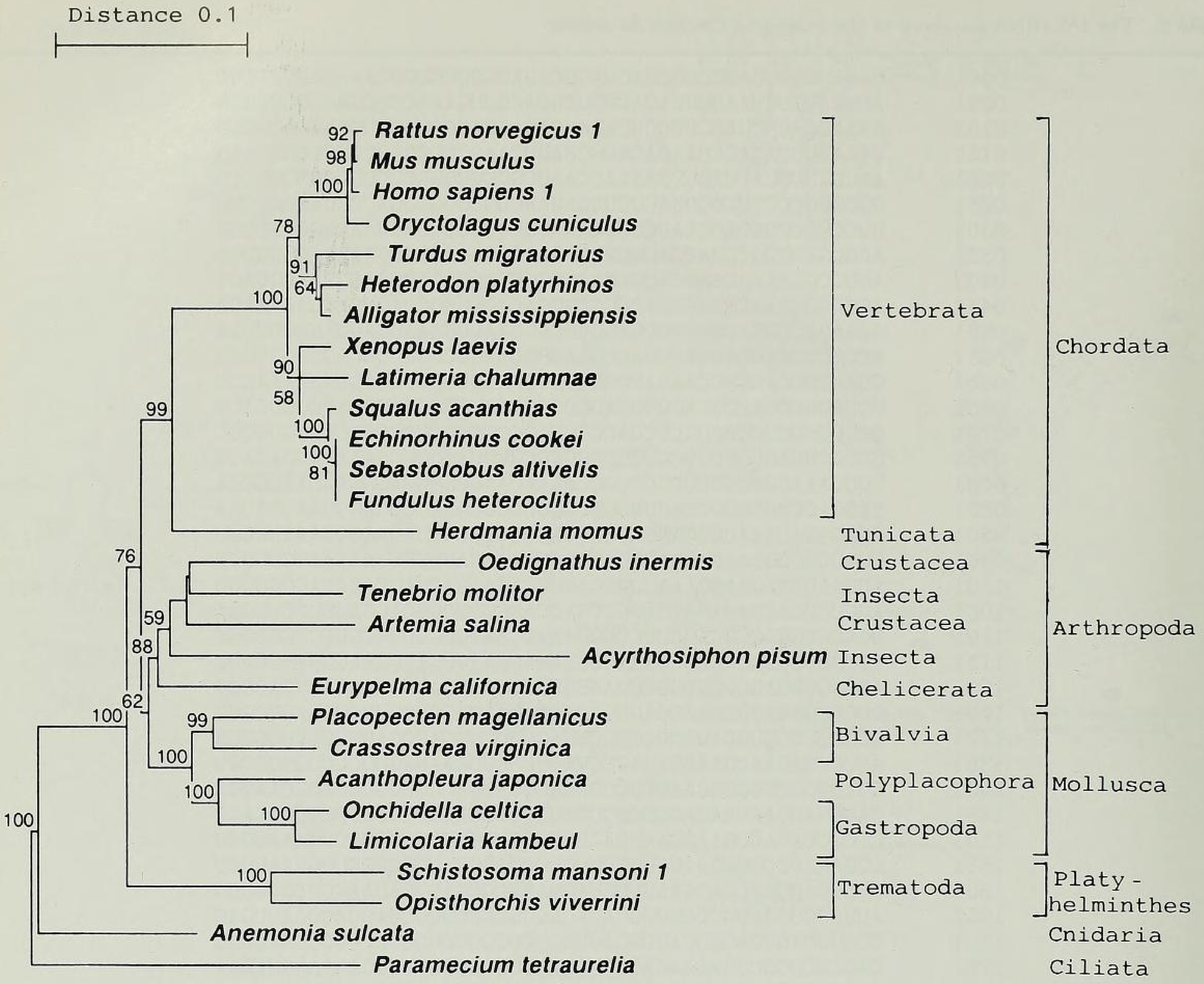

Figure 2. Neighbor-joining tree based on the 18S rRNA sequences from 27 Metazoa. All sequences were complete except for the following (number of sequenced nucleotides between brackets): *Turdus migratorius* (1753), *Alligator mississippiensis* (1691), *Heterodon platyrhinos* (1717) and *Latimeria chalumnae* (1777). *Paramecium tetraurelia* was chosen as an outgroup. Bootstrap values are indicated at the root of each clade, but only if they exceed 50%.

(Bevan *et al.*, 1992). This of course reduces the time advantage of PCR amplification.

MATERIALS AND METHODS

Animals: Specimens of *Onchidella celtica* collected at Vila Franca do Campo (São Miguel, Azores) were frozen alive and preserved at −80°C. Voucher material was deposited in the collections of the "Koninklijk Belgisch Instituut voor Natuurwetenschappen", Brussels (general inventory number, I.G. No. 28053).

DNA extraction: Digestive glands of ten specimens were pooled and homogenized under liquid nitrogen in a prechilled mortar and transferred to 15 ml of preheated (60°C) 2% CTAB buffer (2% (w/v) CTAB; 0.2% (v/v) 2-mercaptoethanol; 1.4 M NaCl; 20 mM EDTA; 100 mM Tris-HCl pH=8; 100 μg/ml proteinase K). After incubation at 60°C for 30 min., further extraction was done

Table 3. Organisms that were used as outgroup in our analyses.

Species	Position
Zea mays	angiosperms
Neurospora crassa	ascomycetes
Saccharomyces cerevisiae	ascomycetes
Rhodosporidium toruloides	basidiomycetes
Gracilaria lemaneiformis	red algae
Porphyra umbilicalis	red algae
Chlorella ellipsoidea	green algae
Volvox carteri	green algae
Prorocentrum micans	dinoflagellates
Giardia duodenalis	diplomonads
Trypanosoma brucei	kinetoplastids
Paramecium tetraurelia	ciliates
Oxytricha nova	ciliates
Plasmodium berghei	apicomplexa (Sporozoa)
Dictyostelium discoideum	slime molds

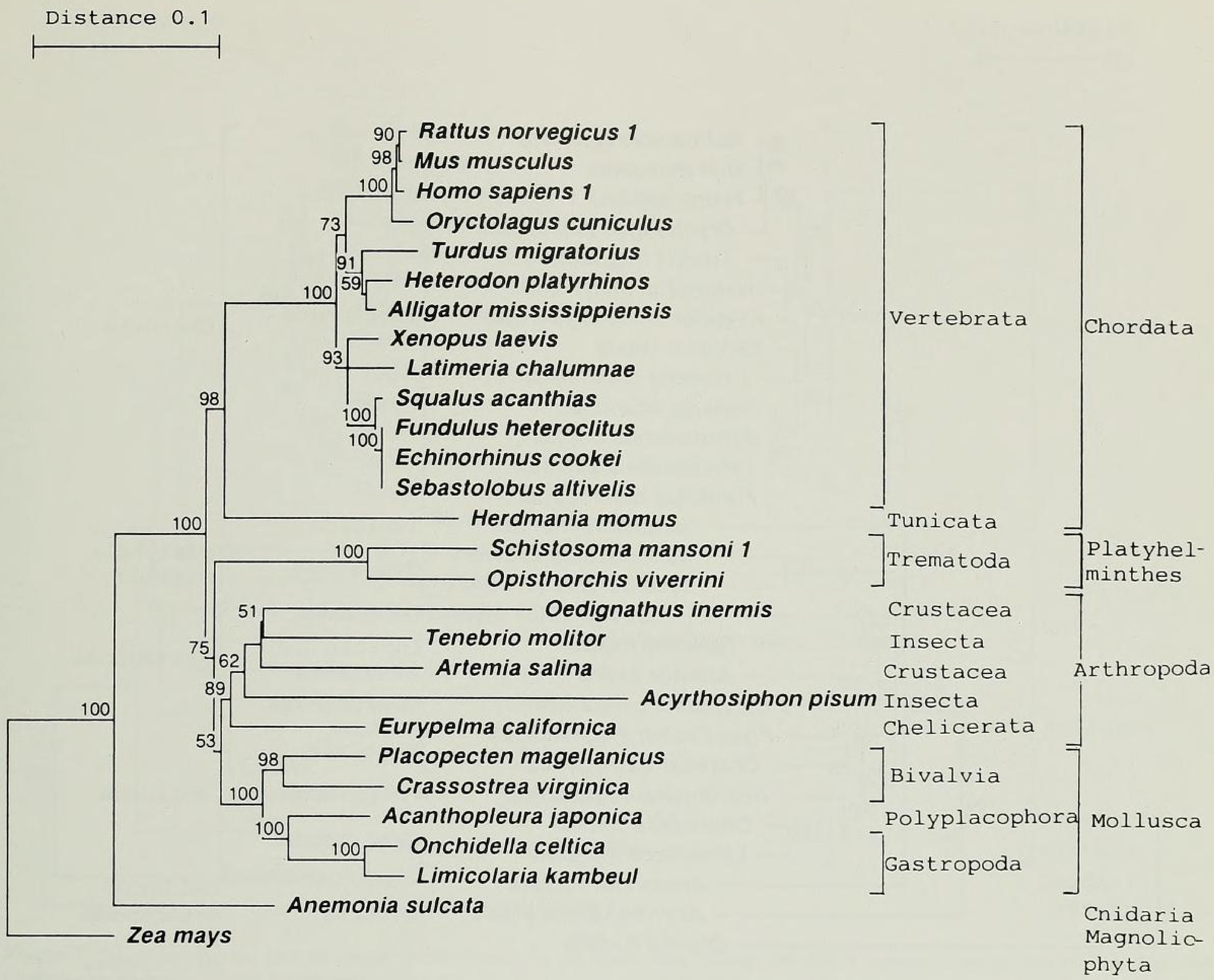

Figure 3. Neighbor-joining tree based on the same set of metazoan 18S rRNA sequences as in Fig. 2, but with *Zea mays* as an outgroup. Bootstrap values are indicated as in Fig. 2.

as described by Winnepenninckx *et al.* (1993a). The DNA yield amounted to 60 μg.

Gene cloning and sequencing: Restriction enzymes suitable for isolation of a DNA fragment containing the 18S rRNA gene were identified as described by Winnepenninckx *et al.* (1992). After digestion of 1.2 μg DNA with *Bam*HI and separation on a 0.8% (w/v) agarose gel, restriction fragments of 4 kb containing the 18S rRNA gene were eluted (Heery *et al.*, 1990). Competent DH5α *E. coli* cells (Gibco BRL Life Technologies; Gaithersburg, USA) were transformed with these DNA restriction fragments ligated into pBluescriptSK$^+$ (Stratagene; La Jolla, California, USA). Colony screening was performed using a PCR fragment of the gastropod *Limicolaria kambeul* (Winnepenninckx *et al.*, 1992), labeled with ^{32}P via nick translation (Rigby *et al.*, 1977). Plasmids were isolated (Birnboim & Doly, 1979) from a single clone and sequencing was performed by the dideoxynucleotide method (Sanger *et al.*, 1977) using Sequenase 2.0 (USB; Cleveland, Ohio, USA). The 18S rRNA primers used are given in Table 1.

Sequence alignment and construction of phylogenetic trees: The *Onchidella celtica* 18S rRNA sequence was aligned with other SSU rRNA sequences present in our database (De Rijk *et al.*, 1992). Alignment was done manually taking into account the secondary structure features of the molecule, as described by De Rijk *et al.* (1992). For tree construction, pairwise distances were calculated using the formula of Jukes and Cantor (1969) modified to take into account gaps (Van de Peer *et al.*, 1990). They served to derive neighbor-joining trees (Saitou & Nei, 1987), whose reliability was tested by bootstrapping (Felsenstein, 1985) over 100 replicates. According to the guidelines of Hillis and Bull (1993), only branching points with bootstrap values higher than 70% were considered to be reliable. Estimated internal branches with bootstrap values above 70% should represent true clades over 95% of the time (Hillis & Bull,

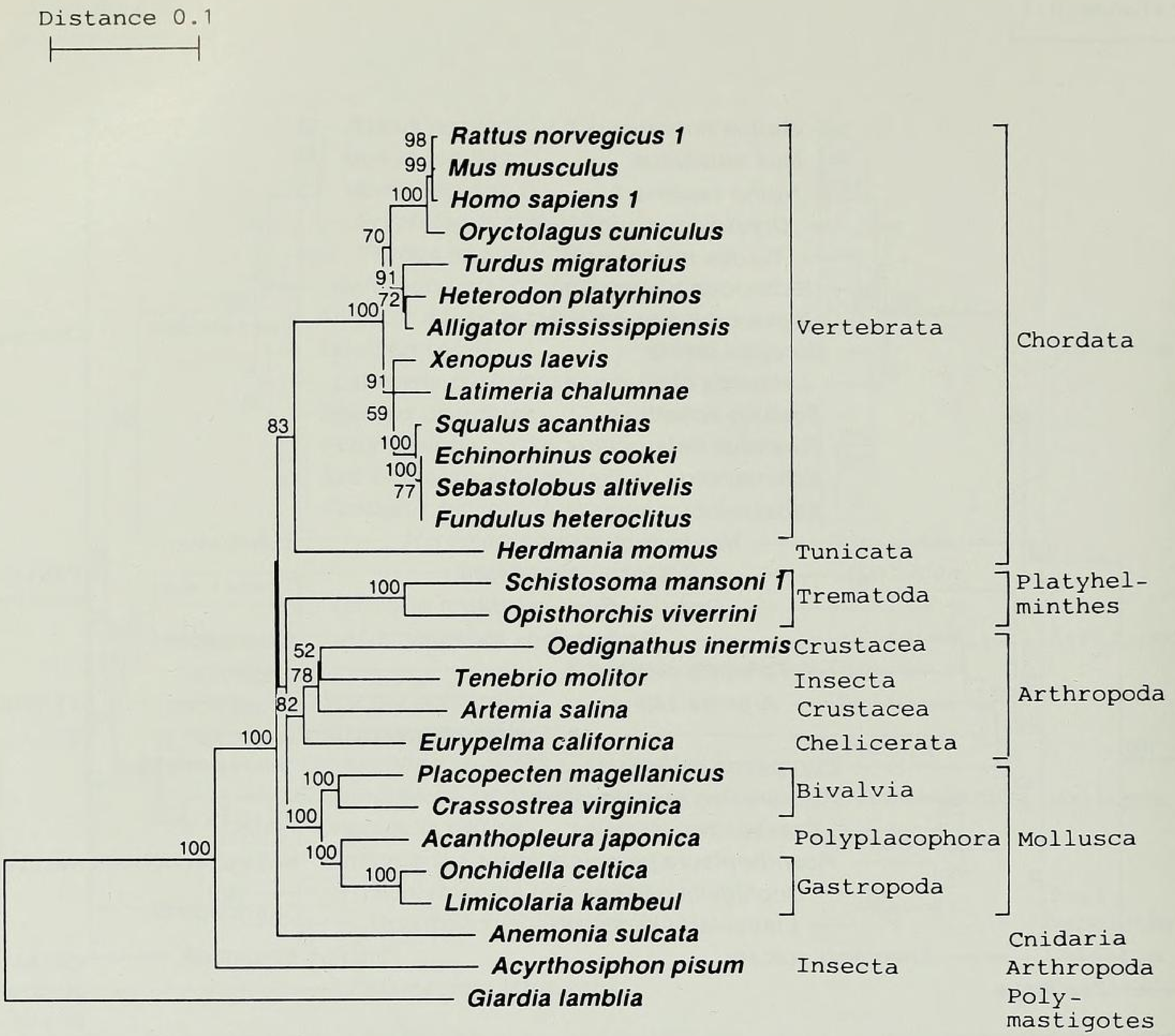

Figure 4. Neighbor-joining tree based on the same set of metazoan 18S rRNA sequences as in Fig. 2, but with *Giardia duodenalis* (often called *Giardia lamblia* or *Giardia intestinalis*) as an outgroup. Bootstrap values are indicated as in Fig. 2.

1993). All calculations were carried out with the TREECON package of Van de Peer and De Wachter (1993). Character state analyses using maximum parsimony were performed using the package HENNIG86 (version 1.5; Farris, 1989) with the heuristic algorithms MHENNIG* and BB* combined. The results were summarized in a strict consensus tree, *i.e.* a tree that contains only those clusters that are common to all competing trees ("nelsen" command of HENNIG86). Nucleotides were treated as non-additive characters and no differential weighting was done.

RESULTS

Sequence Alignment

The 18S rRNA of *Onchidella celtica* (EMBL accession number X70211), of which the nucleotide sequence is shown in Table 2, is 1844 nucleotides long. The 3′ and 5′ termini of the gene were located on the basis of similarity with those of other 18S rRNA sequences. Figure 1 shows a secondary structure model of the molecule in accordance with the one published for *Limicolaria kambeul* (Winnepenninckx *et al.*, 1992). Both models show high similarity to each other and are in accordance with the general model proposed for eukaryotic SSU rRNA (De Rijk *et al.* 1992). Based on our latest insights into the secondary structure of 18S RNA, modifications were made in helices 19, 20, 21 and 38. The new gastropod sequence as well as the one of *Acanthopleura japonica* (Winnepenninckx *et al.*, 1993b) were added to an alignment of other SSU rRNA sequences (De Rijk *et al.*, 1992). This alignment can be obtained on request. Trees were constructed on the basis of a set of 27 metazoan sequences which are either complete or nearly complete.

Distance Matrix Analyses

Figure 2 shows the neighbor joining (NJ) tree obtained with the ciliate *Paramecium tetraurelia* as outgroup. It

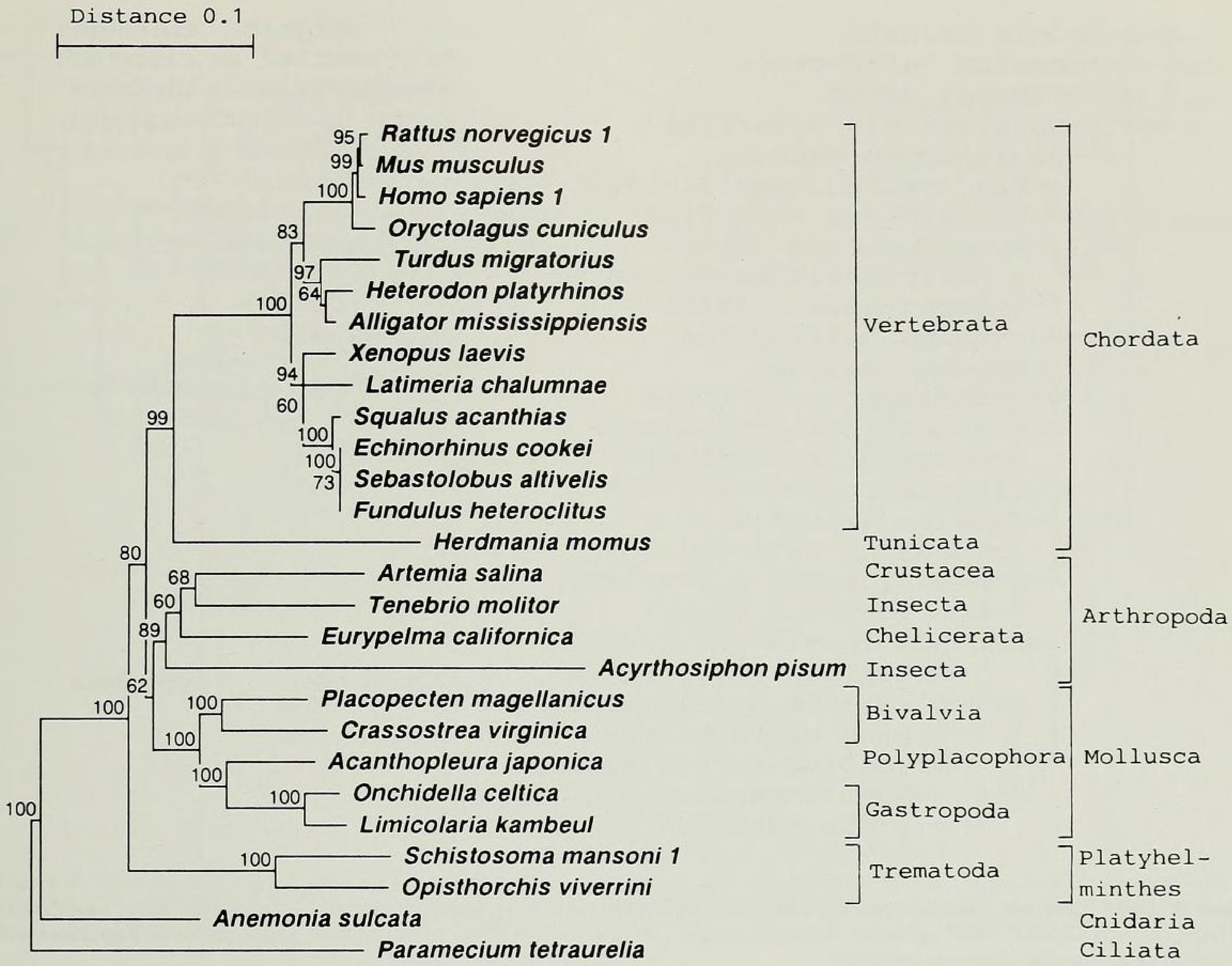

Figure 5. Neighbor-joining tree obtained on the basis of a set containing all the 18S rRNA metazoan sequences of Fig. 2 except *Oedignatus inermis*. *Paramecium tetraurelia* was used as an outgroup. Bootstrap values are indicated as in Fig. 2.

suggests that (bootstrap values in parentheses): (1) Mollusca are a monophyletic group (100/100) within a relatively poorly supported protostome clade (62/100); (2) Gastropoda (100/100) and Bivalvia (99/100) are monophyletic as well; (3) Polyplacophora appears as a sister group to the Gastropoda (100/100). The tree also indicates that : (1) Cnidaria are a sister group to Eubilateria (100/100); (2) Acoelomata, represented by two Trematoda, are a sister group to the Eucoelomata (76/100); (3) Arthropoda are a monophyletic group (88/100); (4) neither Insecta nor Crustacea are monophyletic; (5) Chordata (99/100) and Vertebrata (100/100) are both monophyletic.

We attempted to assess the stability of our tree by testing its sensitivity to the presence of specific taxa. First we studied the influence of the outgroup by successively replacing *Paramecium tetraurelia* by each of the 14 other organisms listed in Table 3. We observed only two topological changes. In nine out of the 15 cases, the topology shown in Figure 3 was obtained, *i.e.* the Platyhelminthes appeared as a sister group to the Arthropoda-Mollusca clade. In one case, the topology shown in Figure 4 was obtained, viz. when the diplomonad *Giardia duodenalis*, was chosen as outgroup. This organism forms a very long branch in previously published trees comprising organisms from different eukaryotic kingdoms (*e.g.* Van de Peer *et al.*, 1993). In this case, the aphid *Acyrthosiphon pisum*, which is also marked by an exceptionally long branch, became a sister group to all other Metazoa. The latter observation is probably due to the fact that errors in distance estimation increase with the amount of divergence. Long branches will provoke an underestimation of the evolutionary distance and will systematically attract each other, causing biased topologies (Felsenstein, 1978; Olsen, 1987; Swofford & Olsen, 1990; Lake, 1991). The changes in the position of the Platyhelminthes, which do not have exceptionally long branches, is probably not due to such a systematic error. The low bootstrapping values on their branching point, suggests uncertainty as to their position. Inclusion of representatives of more invertebrate phyla might be helpful in this case.

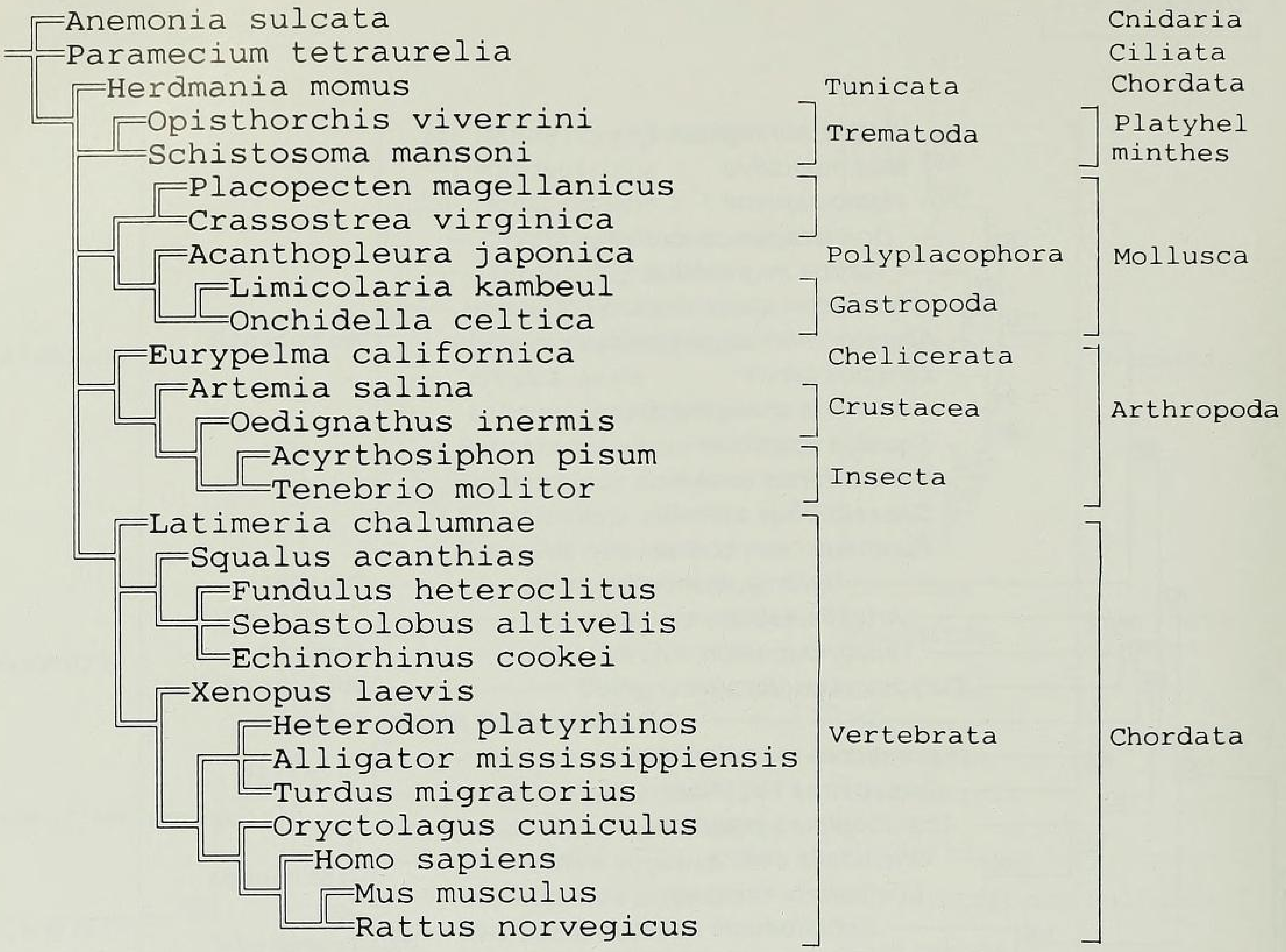

Figure 6. Strict consensus tree constructed from three maximum parsimony trees (length=3114 steps; c.i.= 0.51) obtained by applying the MHENNIG*+BB* option of Hennig86 on the 706 informative positions of the same alignment as in Fig. 2 and with *Paramecium tetraurelia* chosen as outgroup.

Subsequently, we constructed 27 trees with *Paramecium tetraurelia* as outgroup, but each time omitting one species. Only one topological change was observed: when excluding *Oedignathus inermis*, *Acyrtosiphon pisum* branched off first within the arthropod clade (Figure 5). The fact that this change involves the species with the longest branch, again points to the above mentioned "long branch effect" (Felsenstein, 1978; Swofford & Olsen, 1990). Since the placement of the two Platyhelminthes was ambiguous (*cfr.* Figures 2 and 3) and since we suspected *Acyrthosiphon pisum* to be a source of systematic errors, we removed all three species from our data set to assess their impact. However, the topology of the tree we obtained did not differ from the one in Figure 2.

Character State Analyses

The 28 species analysed, with *Paramecium tetraurelia* as outgroup, yielded 706 informative sites. A position is informative if it contains at least two different nucleotides, each of them present in at least two species (Nei, 1987). Ambiguous nucleotides were not used to ascertain the informative character of a position. Three maximum parsimony (MP) trees of 3114 steps and with a consistency index (c.i.) (Kluge & Farris, 1969) of 0.51 were found. The strict consensus tree shown in Figure 6 suggests that (1) Mollusca, Bivalvia and Gastropoda are monophyletic groups; (2) Polyplacophora appear as a sister group to Gastropoda; (3) Arthropoda are a monophyletic clade in which Chelicerata branch off first; (4) Insecta are monophyletic but Crustacea are paraphyletic; (5) Vertebrata are monophyletic. Ten different data input orders did not change this topology. Again we tested the stability of our results. If the placement of a taxon is biased, its removal should cause an increase of the consistency index (Swofford & Olsen, 1990). We checked this by successively removing those species, the position of which appeared unstable in our distance matrix analyses, viz. *Acyrthosiphon pisum*, *Schistosoma mansoni*, *Opisthorchis viverrini*, and a fourth species, *Mus musculus*, which occupied a stable position. Each time we identified the informative positions anew and applied HENNIG86 with *Paramecium tetraurelia* as outgroup. Omitting *Acyrthosiphon pisum* increased the c.i. to 0.53, while removing any of the other species did not change the c.i. This again suggests that the placement of *Acyrthosiphon pisum* is liable to a systematic error. As for the ambiguous position of the Platyhelminthes, this may

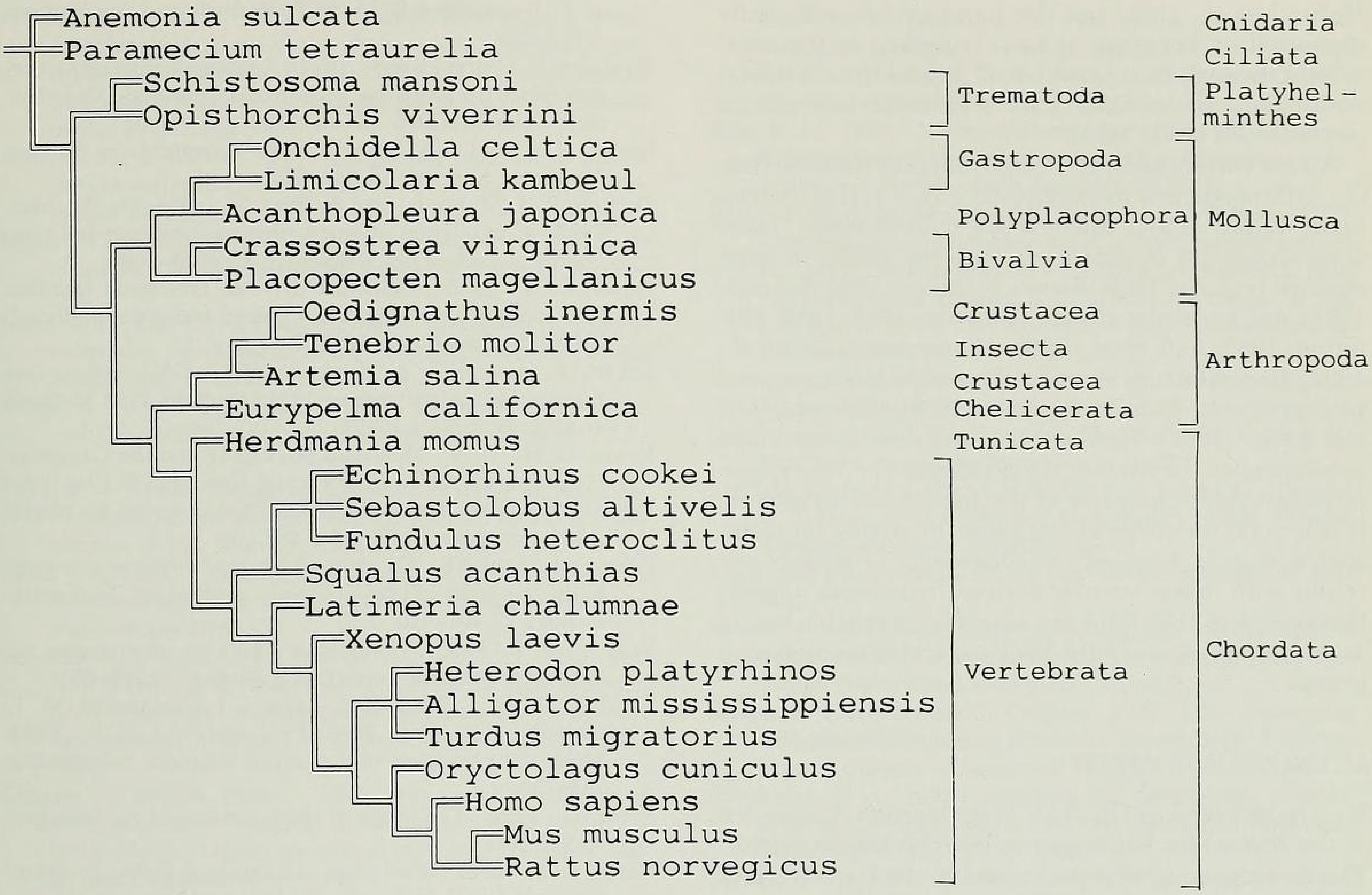

Figure 7. Strict consensus tree constructed from two maximum parsimony trees (length=2741 steps; c.i.=0.53) obtained with the MHENNIG*+BB* option on the 658 informative sites of the same alignment of Fig. 2 from which the insect *Acyrthosiphon pisum* was removed.

be due to the lack of other invertebrate phyla and classes. Figure 7 shows the strict consensus tree of the two MP trees (length=2741; c.i.=0.53) obtained when *Acyrthosiphon pisum* was excluded. All our conclusions based on the tree in Figure 6 remain valid, but in addition the bilaterian pentachotomy of Figure 6 is now resolved. Mollusca appear as a sister group to a clade containing Arthropoda and Chordata. It is also suggested that (1) Bilateria are monophyletic; (2) Acoelomata are a sister group to Eucoelomata; (3) Chordata are monophyletic.

DISCUSSION

The monophyletic character of the Mollusca, the Bivalvia and the Gastropoda, which is supported by all our trees, is generally accepted (*e.g.* Brusca & Brusca, 1990; Willmer, 1990; von Salvini-Plawen, 1985, 1990a; Götting, 1980). Using globin amino acid sequences, Goodman *et al.* (1988) agreed with these views. The 5S rRNA based analyses of Ohama *et al.* (1984), Hendriks *et al.* (1986) and Hori and Osawa (1987) also confirmed gastropod monophyly. Ghiselin (1988, 1989) supported molluscan monophyly. But Patterson (1989) and Lake (1989) did not corroborate these well established views, while the question was not resolved by Field *et al.* (1988; see also Raff *et al.*, 1989).

In both the distance and MP trees, we find the chiton included within the conchiferan clade as a sister group to the Gastropoda. This result is in contrast with the results of anatomical (*e.g.* Milburn, 1960; Stasek, 1972; Götting, 1980; Scheltema, 1988; Brusca & Brusca, 1990; von Salvini-Plawen, 1990a) and paleontological (*e.g.* Runnegar & Pojeta, 1974; Pojeta, 1980; Peel, 1991) studies. Neither Field *et al.* (1988; see also Raff *et al.*, 1989), nor Ghiselin (1988, 1989) or Lake (1989) were able to resolve the position of the Polyplacophora, while Patterson (1989) suggested that Polyplacophora and Brachiopoda are sister taxa. Using mitochondrial SSU rRNA sequences (Ballard *et al.*, 1992), the class either appeared as a sister group to the Gastropoda-Annelida clade or formed together with the Gastropoda a sister group to the Annelida. Addition of more molluscan representatives to our data set is necessary to investigate the conflicting position of the Polyplacophora.

Our current data set is also not sufficiently representative to draw conclusions on the position of the Mollusca among other Metazoa. From our NJ analyses, the phylum appears as a sister group to the Arthropoda (see also

Holland *et al.*, 1991) but this topology is insufficiently supported by bootstrap values. According to the character state analysis it branches off before the Chordata-Arthropoda clade. Data from additional invertebrate phyla should be included.

All our current analyses strongly support the view that: (1) Arthropoda is a monophyletic group and (2) Vertebrata, Chordata and Bilateria are monophyletic. These observations are in agreement with the results of some classical (*e.g.* Ax, 1989; Brusca & Brusca, 1990; Schram, 1991) and molecular studies (Ghiselin, 1988, 1989; Patterson, 1989; Raff *et al.*, 1989; Winnepenninckx *et al.*, 1992). Contradictory views on these aspects of metazoan phylogeny were given by *e.g.* Lake (1989), Willmer (1990) and Fryer (1992). However our analyses do not allow conclusions on the status of the Acoelomata and the mono- or paraphyletic character of the Insecta and Crustacea. It is beyond the scope of this paper to expand on metazoan evolution, however the congruence of most of our results with independently derived hypotheses suggests that complete 18S rRNA sequences are a reliable tool to assess the phylogeny of the Mollusca and other metazoan groups.

ACKNOWLEDGMENTS

We are indebted to Dr. A.M. Frias Martins (University of the Azores) for helping to collect *Onchidella celtica*. The investigations were performed in the framework of the Institute for the Study of Biological Evolution of the University of Antwerp. This work was supported by FKFO Grants 2.0023.94 and 2.0003.93. B. Winnepenninckx holds an I.W.O.N.L. scholarship.

LITERATURE CITED

Anderson, D. T. 1981. Origins and relationships among animal phyla. Proceedings of the Linnean Society of New South Wales 106:151–166.

Ax, P. 1989. Basic phylogenetic systematization of the Metazoa. *in*: Fernholm, B., K. Bremer and H. Jörnvall (eds). The Hierarchy of life. Elsevier, Amsterdam, p. 229–245.

Backeljau, T., B. Winnepenninckx and L. De Bruyn. 1993. Cladistic analysis of metazoan relationships: areappraisal. Cladistics 9:167–181.

Ballard, J. W. O., G. J. Olsen, D. P. Faith, W. A. Odgers, D. M. Rowell and P. W. Atkinson. 1992. Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods. Science 258:1345–1348.

Bej, A. K., M. H. Mahbubani and R. M. Atlas. 1991. Amplification of nucleic acids by polymerase chain reaction (PCR) and other methods and their applications. Critical reviews in Biochemistry and Molecular Biology 26:301–334.

Bergström, J. 1986. Metazoan evolution - a new model. Zoologica Scripta 15:189–200.

Bergström, J. 1991. Metazoan evolution around the Precambrian-Cambrian transition. *In*: Simonetta, A.M. and S.C. Morris (eds). The early evolution of Metazoa and the significance of problematic taxa. Cambridge University Press, p. 25–34.

Bevan, I. S., R. Rapley and M. R. Walker. 1992. Sequencing of PCR-amplified DNA. PCR Methods and Applications 1:222–228.

Birnboim, H. C. and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Research 7:1513–1522.

Brusca, R. C. and G. J. Brusca. 1990. Invertebrates. Sinauer Associates, Sunderland, MA 922 pp.

De Rijk, P., J.-M. Neefs, Van de Peer, Y. and R. De Wachter. 1992. Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Research 20:2075–2089.

Dover, G. A. 1986. Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. Trends in Genetics 2:159–165.

Eckert, K. A. and T. A. Kunkel. 1991. DNA polymerase fidelity and the polymerase chain reaction. PCR Methods and Applications 1:17–24.

Erwin, D. H. 1991. Metazoan phylogeny and the Cambrian radiation. Trends in Ecology and Evolution 6: 131–134.

Farris, J. S. 1989. Hennig86: a PC-DOS program for phylogenetic analysis. Cladistics 5:163.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Systematic Zoology 27:401–410.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791.

Field, K. G., G. J. Olsen, D. J. Lane, S. J. Giovannoni, M. T. Ghiselin, E. C. Raff, N. R. Pace and R. A. Raff. 1988. Molecular phylogeny of the animal kingdom. Science 239: 748–753.

Fryer, G. 1992. The origin of the Crustacea. Acta Zoologica 73:273–286.

Gerbi, S. A. 1985. Evolution of ribosomal DNA. *In*: MacIntyre, R.J. (ed). Molecular Evolutionary Genetics. Plenum Press, New York and London, p. 419–517.

Ghiselin, M. T. 1988. The origin of molluscs in the light of molecular evidence. *In*: Harvey, P. H. and L. Partridge (eds). Oxford Surveys in Evolutionary Biology. Volume 5. p. 66–95.

Ghiselin, M. T. 1989. Summary of our present knowledge of metazoan phylogeny. *In*: Fernholm, B., K. Bremer and H. Jörnvall (eds). The Hierarchy of life. Elsevier, Amsterdam, p. 261–272.

Goodman, M., J. Pedwaydon, J. Czelusniak, T. Suzuki, T. Gotoh, L. Moens, F. Shishikura, D. Walz and S. Vinogradov. 1988. An evolutionary tree for invertebrate globin sequences. Journal of Molecular Evolution 27:236–249.

Götting, K.-J. 1980. Origin and relationships of the Mollusca. Zeitschrift für zoologische Systematik und Evolutionsforschung 18:24–27.

Gyllenstein, U. B. 1989. PCR and DNA sequencing. Biotechniques 7:700–708.

Heery, D. M., F. Gannon and R. Powell. 1990. A simple method for subcloning DNA fragments from gel slices. Trends in Genetics 6:173.

Hendriks, L., E. Huysmans, A. Vandenberghe and R. De Wachter. 1986. Primary structures of the 5S ribosomal RNAs of 11 arthropods and applicability of 5S RNA to the study of metazoan evolution. Journal of Molecular Evolution. 24:103–109.

Hillis, D. M. and M. T. Dixon. 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. The Quarterly Review of Biology. 66:411–453.

Hillis, D. M. and J. J. Bull. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Systematic Biology. 42:182–192.

Holland, P. W. H., A. M. Hacker and N. A. Williams. 1991. A molecular analysis of the phylogenetic affinities of *Saccoglossus cambrensis* Brambell & Cole (Hemichordata). Philosophical Transactions of the Royal Society of London (Series B) 332:185–189.

Hori, H. and S. Osawa. 1987. Origin and evolution of organisms as deduced from 5S ribosomal RNA sequences. Molecular Biology and Evolution 4:445–472.

Inglis, W. G. 1985. Evolutionary waves: patterns in the origins of animal phyla. Australian Journal of Zoology 33:153–178.

Jukes T. H. and C. R. Cantor. 1969. Evolution of protein molecules. *In*: Munro, H. N. (ed). Mammalian Protein Metabolism. Academic Press, New York, p. 21–132.

Kluge, A. G. and J. S. Farris. 1969. Quantitative phyletics and the evolution of anurans. Systematic Zoology 18:1–32.

Lake, J. A. 1989. Origin of the multicellular animals. *In*: Fernholm, B., K. Bremer and H. Jörnvall (eds). The Hierarchy of life. Elsevier, Amsterdam, p. 273–278.

Lake, J. A. 1991. Tracing origins with molecular sequences: metazoan and eukaryotic beginnings. Trends in Biochemical Sciences 16:46–50.

Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin and N. R. Pace. 1985. Rapid determination of 16S-ribosomal RNA sequences for phylogenetic analyses. Proceedings of the National Academy of Science of the USA 82:6955–6959.

Lenaers, G. and M. Bhaud. 1992. Molecular phylogeny of some polychaete annelids: an initial approach to the Atlantic-Mediterranean speciation problem. Journal of Molecular Evolution 35:429–435.

Littlewood, D. T. J., S. E. Ford and D. Fong. 1991. Small subunit rRNA gene sequence of *Crassostrea virginica* (Gmelin) and a comparison with similar sequences from other bivalve molluscs. Nucleic Acids Research 19:6048.

Lyddiatt, A., D. Peacock and D. Boulter. 1978. Evolutionary change in invertebrate cytochrome C. Journal of Molecular Evolution 11:35–45.

Milburn, P. W. 1960. Further remarks on the interpretation of the Mollusca. The Veliger 3:43–48.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, p. 315–319.

Nielsen, C. 1977. The relationships of Entoprocta, Ectoprocta and Phoronida. American Zoologist 17:149–150.

Ohama, T., T. Kumazaki, H. Hori and S. Osawa. 1984. Evolution of multicellular animals as deduced from 5S rRNA sequences: a possible early emergence of the Mesozoa. Nucleic Acids Research 12:5101–5108.

Olsen, G. J. 1987. Earliest phylogenetic branchings: comparing rRNA-based evolutionary trees inferred with various techniques. Cold Spring Harbor Symposia on Quantitative Biology LII:825–837.

Patterson, C. 1989. Phylogenetic relations of major groups: conclusions and prospects. *In*: Fernholm, B., K. Bremer, and H. Jörnvall (eds). The Hierarchy of life. Elsevier, Amsterdam, p. 471–487.

Peel, J. S. 1991. Functional morphology of the class Helcionelloida nov., and the early evolution of the Mollusca. *In*: Simonetta, A.M. and S.C. Morris (eds). The early evolution of Metazoa and the significance of problematic taxa. Cambridge University Press, p. 157–177.

Pojeta, J. Jr. 1980. Molluscan phylogeny. Tulane Studies in Geology and Paleontology 16:55–80.

Raff, R. A., K. G. Field, G. J. Olsen, S. J. Giovannoni, D. J. Lane, M. T. Ghiselin, N. R. Pace and E. C. Raff. 1989. Metazoan phylogeny based on analysis of 18S ribosomal RNA. *In*: Fernholm, B., K. Bremer and H. Jörnvall (eds). The Hierarchy of life. Elsevier, Amsterdam, p. 247–260.

Rice, E. L. 1990. Nucleotide sequence of the 18S ribosomal RNA gene from the Atlantic sea scallop *Placopecten magellanicus*. Nucleic Acids Research 18:5551.

Rigby, P. W. J., M. Dieckmann, C. Rhodes and P. Berg. 1977. Labeling deoxyribonucleic acids to high specific activity in vitro by nick translation with DNA polymerase I. Journal of Molecular Biology 113: 237–251. Runnegar, B. and J. Pojeta, Jr. 1974. Molluscan phylogeny: the paleontological viewpoint. Science 186:311–317.

Runnegar, B. and J. Pojeta, Jr. 1985. Origin and diversification of the Mollusca. *In*: Wilbur, K. M., E. R.

Trueman and M. R. Clarke (eds). The Mollusca. Volume 10. Evolution. Academic Press, Inc. p. 1–57.

Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis and H. A. Erlich. 1988. Primer-directed enzymatic amplification of

DNA with a thermostable DNA polymerase. Science 239:487–491.

Saitou, N. and M. Nei. 1987. The Neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406–425.

Sanger, F., S. Nicklen and R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Science of the USA 74:5463–5467.

Scheltema, A. H. 1988. Ancestors and descendents: relationships of the Aplacophora and Polyplacophora. American Malacological Bulletin 6:57–68.

Schram, F. R. 1991. Cladistic analysis of metazoan phyla and the placement of fossil problematica. *In*: Simonetta, A. M. and S. C. Morris (eds). The early evolution of Metazoa and the significance of problematic taxa. Cambridge University Press, p. 35–46.

Solignac, M., M. Pélandakis, F. Rousset and A. Chenuil. 1991. Ribosomal RNA phylogenies. *In*: Hewitt G.M., A. W. B. Johnston and J. P. W. Young (eds). NATO ASI Series. Volume 57. Molecular Techniques in Taxonomy. Volume H57. Springer-Verlag, Berlin, Heidelberg, p. 73–85.

Stasek, C. R. 1972. The molluscan framework. *In*: Florkin, M. and B. T. Scheer (eds). Chemical Zoology. Volume VII. Mollusca. Academic Press, New York and London, p. 1–44.

Steiner, G. 1992. Phylogeny and classification of Scaphopoda. Journal of Molluscan Studies 58:385–400.

Swofford, D. L. and G. J. Olsen. 1990. Phylogeny reconstruction. *In*: Hillis, D. M. and C. Moritz (eds). Molecular Systematics. Sinauer Associates, Inc., Sunderland, MA, p. 411–501.

Valentine, J. W. 1980. L'origine des grands groupes d'animaux. La Recherche 112:666–674.

Valentine, J. W. 1991. Major factors in the rapidity and extant of the metazoan radiation during the Proterozoic-Phanerozoic transition. *In*: Simonetta, A.M. and S.C. Morris (eds). The early evolution of Metazoa and the significance of problematic taxa. Cambridge University Press, p. 11–13.

Van de Peer, Y., J.-M. Neefs and R. De Wachter. 1990. Small ribosomal subunit RNA sequences, evolutionary relationships among different life forms, and mitochondrial origins. Journal of Molecular Evolution 30:463–476.

Van de Peer, Y., J.-M. Neefs, P. De Rijk and R. De Wachter. 1993. Evolution of eukaryotes as deduced from small ribosomal subunit RNA sequences. Biochemical Systematics and Ecology 21:43–56.

Van de Peer, Y. and R. De Wachter. 1993. TREECON: A software package for the construction and drawing of evolutionary trees. Computer Applications in the Biosciences 9: 177–182.

von Ihering, H. 1876. Versuch eines natürlichen Systemes der Mollusken. Jahrbuch der Deutschen Malakozoologischen Gesellschaft 3: 97–148.

von Salvini-Plawen, L. 1969. Solenogastres und Caudofoveata (Mollusca, Aculifera) Organisation und phylogenetische Bedeutung. Malacologia 9:191–216.

von Salvini-Plawen, L. 1972. Zur Morphologie und Phylogenie der Mollusken; die Beziehungen der Caudofoveata und der Solenogastres als Aculifera, als Mollusca und als Spiralia. Zeitschrift zur Wissenschaftlichen Zoologie 184: 205–394.

von Salvini-Plawen, L. 1985. Early evolution and the primitive groups. *In*: Wilbur, K. M., E. R. Trueman and M. R. Clarke (eds). The Mollusca. Volume 10. Evolution. Academic Press, Inc. p. 59–150.

von Salvini-Plawen, L. 1990a. Origin, phylogeny and classification of the phylum Mollusca. Iberus 9:1–33.

von Salvini-Plawen, L. 1990b. The status of the Caudofoveata and the Solenogastres in the mediterranean sea. Lavori della Società Italiana di Malacologia 23: 5–30.

Wilhelmi, R. W. 1944. Serological relationships between the Mollusca and other invertebrates. Biological Bulletin 87: 96–105.

Willmer, P. 1990. Invertebrate relationships—Patterns in animal evolution. Cambridge University Press, Cambridge.

Wingstrand, K. G. 1985. On the anatomy and relationships of recent Monoplacophora. Galathea Reports 16:7–94.

Winnepenninckx, B., T. Backeljau, Y. Van de Peer and R. De Wachter. 1992. Structure of the small ribosomal subunit RNA of the pulmonate snail *Limicolaria kambeul*, and phylogenetic analysis of the Metazoa. FEBS Letters 309: 123–126.

Winnepenninckx, B., T. Backeljau and R. De Wachter. 1993a. Extraction of high molecular weight DNA from molluscs. Trends in Genetics 9:407.

Winnepenninckx, B., T. Backeljau and R. De Wachter. 1993b. Complete small ribosomal subunit RNA sequence of the chiton *Acanthopleura japonica* (Lischke, 1873) (Mollusca, Polyplacophora). Nucleic Acids Research 21:1670.

Woese, C. R. 1991. The use of ribosomal RNA in reconstructing evolutionary relationships among Bacteria. *In*: Selander, R. K., A. G. Clark and T. S. Whittam (eds). Evolution at the molecular level. Sinauer Associates, Inc., Sunderland, MA p. 1–25.

Woese, C. R. and R. R. Gutell. 1989. Evidence for several higher order structural elements in ribosomal RNA. Proceedings of the National Academy of Science of the USA 86:3119–3122.