Phylogeny of Protostome Worms Derived from 18S rRNA Sequences

Birgitta Winnepenninckx, Thierry Backeljau,* and Rupert De Wachter
Departement Biochemie, Universiteit Antwerpen (UIA); and *Koninklijk Belgisch Instituut voor Natuurwetenschappen

The phylogenetic relationships of protostome worms were studied by comparing new complete 18S rRNA sequences of Vestimentifera, Pogonophora, Sipuncula, Echiura, Nemertea, and Annelida with existing 18S rRNA sequences of Mollusca, Arthropoda, Chordata, and Platyhelminthes. Phylogenetic trees were inferred via neighbor-joining and maximum parsimony analyses. These suggest that (1) Sipuncula and Echiura are not sister groups; (2) Nemertea are protostomes; (3) Vestimentifera and Pogonophora are protostomes that have a common ancestor with Echiura; and (4) Vestimentifera and Pogonophora are a monophyletic clade.

Introduction

The phylogenetic and systematic status of several protostome worms, including Pogonophora, Vestimentifera, Annelida, Echiura, Sipuncula, and Nemertea, is still uncertain. Pogonophora and Vestimentifera (both beard worms) are tube-dwelling marine animals with remarkably long, thin bodies (see, e.g., Brusca and Brusca 1990, p. 450). Because of their supposed tripartite body plan, beard worms were classified among the deuterostomes (Ivanov 1955, 1975; Southward 1963; Godeaux 1974). However, once it was shown that beard worms also possess a posterior metameric body region (Webb 1964), they were related to the protostomes, particularly to the Annelida (Nörrøn 1970; Gupta and Little 1975; Van der Land and Nörrøn 1975; Jones 1985; Southward 1988; Brusca and Brusca 1990, p. 882; Conway Morris 1993). On the basis of several ambiguous characters, Cutler (1975) considered beard worms to be intermediate between the Protostomia and Deuterostomia. Jones (1985) separated the Vestimentifera from the Pogonophora at phylum level. However, since Pogonophora and Vestimentifera also have some characters in common, they have often been considered to be closely related, and some authors have assigned Vestimentifera class rank within the Pogonophora (e.g., Southward 1988).

Sipuncula and Echiura are two small phyla of unsegmented worms, which resemble each other in many aspects. Both are characterized by a sausage-shaped body, divided into an anterior region and a trunk. They are undoubtedly protostomes, although their phylogenetic position within that assemblage is not entirely clear (e.g., Brusca and Brusca 1990, p. 456). Larval and embryological features suggested that Sipuncula and Echiura are closely related and that both groups are allied to the Annelida (Clark 1969).

Nemertea are dorsoventrally flattened, cephalized, unsegmented worms. On the basis of anatomical and embryological similarities, they are considered either as acelomates, closely related to Turbellaria (e.g., Norensburg 1985; Bartolomaeus, 1988; Willmer, 1990, p. 207), or as protostomes (see, e.g., Berg 1985; Turbeville and Ruppert 1985; Turbeville 1986).

The objective of this study is to assess the following questions: Should Pogonophora and Vestimentifera be assigned to a single phylum or to separate phyla? Are Sipuncula and Echiura sister groups? What is the phylogenetic position of Annelida, Sipuncula, Echiura, Pogonophora, Vestimentifera, and Nemertea? Therefore, we determined complete 18S rRNA sequences of Vestimentifera, Pogonophora, Echiura, Sipuncula, Nemertea, and Annelida (Oligo- and Polychaeta) and compared them with existing ones of other Metazoa (Van de Peer et al. 1994).

Material and Methods

Specimen Collection

The animals examined are Lineus sp. (Nemertea; Roscoff, France), Phascolosoma granulatum (Sipuncula; São Miguel, Azores), Ochetostoma erythrogrammon
DNA Extraction

High molecular weight DNA of *O. erythrogrammon* (proboscis), *Lineus sp.* (complete body), *E. fetida* (complete body except intestine and body wall), and *L. concilia* (digestive gland) was extracted as described by Winnepenninckx et al. (1993). To extract DNA of *P. granulatum*, 10 complete specimens, from which body wall and intestine were removed, were ground under liquid nitrogen and incubated for 30 min at 65°C in a solution containing 0.5% SDS (w/v), 10 mM Tris, 5 mM EDTA, and 100 μg/ml Proteinase K. DNA was purified via phenol/chloroform/isoamylalcohol and chloroform/isoamylalcohol extraction followed by isopropanol precipitation (Sambrook et al. 1989). After washing with 70% ethanol, the pellet was dried and suspended in a minimal volume of buffer containing 10 mM Tris-HCl (pH = 7.4) and 1 mM EDTA. Only a tiny tissue sample was available from *S. fioridum*. This sample was cut in pieces and incubated for 15 h at 65°C in buffer containing 100 mM Tris (pH = 7.4), 1 mM EDTA, 0.5% SDS, and 3 mg/ml Proteinase K. DNA was purified by phenol, phenol/chloroform/isoamylalcohol, and chloroform/isoamylalcohol extraction followed by isopropanol precipitation (Sambrook et al. 1989). DNA of *R. piscesae* was provided by D. R. Dixon.

18S rDNA Amplification and Sequencing

An 8 kb HindIII restriction fragment containing the complete 18S rRNA gene of *E. fetida* was cloned as described by Winnepenninckx et al. (1992). The complete 18S rRNA gene of *E. fetida* was sequenced from one clone, as described below. The major part of the sequence of *P. granulatum*, starting from a HindIII site at nucleotide 250, was determined from a cloned restriction fragment (Winnepenninckx et al. 1992). The remaining part was determined on a PCR amplification product using primers 1 and 2 (table 1). The 18S rRNA genes from the other five species were amplified via PCR using primers 1 and 4 (table 1), complementary to the 5' and 3' ends of the gene. The missing 3' part of the gene, where the reverse primer anneals, was PCR-amplified using primer 3 and a primer complementary to the 5' end of the 28S rRNA gene (5'-ACCCGCTGTA-

### Table 1

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Direction</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CTGGTTGATYCTGCCCAGT</td>
<td>F</td>
<td>4–21</td>
</tr>
<tr>
<td>2</td>
<td>ATTACCGGGGCTGCGGC</td>
<td>R</td>
<td>605–588</td>
</tr>
<tr>
<td>3</td>
<td>TTTGYAACACCCGCCGC</td>
<td>F</td>
<td>1666–1685</td>
</tr>
<tr>
<td>4</td>
<td>CVGCAAGTTACCTACCRG</td>
<td>R</td>
<td>1833–1816</td>
</tr>
</tbody>
</table>

*Primer combinations are from Winnepenninckx et al. (1994).*

*Primers with a nucleotide sequence complementary to the RNA encoding strand are indicated by R (reversed). Primers with a sequence corresponding to the one of the RNA strand are indicated by F (forward).*

*Position is corresponding position in the 18S rRNA gene of *Onchidella celtica.*

AYTGAAGCAT-3'). The PCR amplification products, which were purified via agarose gel electrophoresis (Heery et al. 1990), were ligated in a T-tailed vector (Biorad) and DH5α-*E. coli* cells were transformed. The plasmid was isolated from recombinant cells using the alkaline lysis method (Birnboim and Doly 1979).

Both strands of the 18S rRNA genes were sequenced by the dideoxynucleotide method (Sanger et al. 1977) using 17 specific 18S rRNA primers (Winnepenninckx et al. 1994) and the Sequenase 2.0 kit (USB, Cleveland).

Data Analysis

The seven new 18S rRNA sequences were fitted into an alignment of small subunit rRNA sequences (Van de Peer et al. 1994) using the computer program DCSE (De Rijk and De Wachter 1993), which considers primary as well as secondary structure similarity. If necessary, manual adjustments were made with the alignment editor provided by the same program. Phylogenetic trees were inferred by the neighbor-joining (NJ) (Saitou and Nei 1987) and maximum parsimony (MP) methods. Neighbor-joined trees were constructed with the program TREECON (Van de Peer and De Wachter 1993), using the distances of Jukes and Cantor (1969), Tajima and Nei (1984), Kimura (1980), and Jin and Nei (1990), all modified to take gaps into account (Van de Peer et al. 1990). All alignment sites were considered. The NJ tree was bootstrapped (Felsenstein 1985), running 1,000 replicates. Bootstrapting is a common method for assessing confidence in phylogenetic analyses (see Sanderson 1989; Hillis and Bull 1993; Doye 1994 for a discussion). Maximum parsimony trees were constructed using the PAUP heuristic search (Swofford 1993) option, treating gaps as missing data. Only phylogenetically informative sites were included. They were identified using TREECON (Van de Peer and De Wachter 1993). A bootstrap analysis running 1,000 replicates was performed. The relative robustness of the MP tree was fur-
ther assessed by generating cladograms of a few steps longer than the MP tree. These cladograms were combined into strict consensus trees. This procedure yields a decay index (Bremer 1988; Donoghue et al. 1992), indicating the number of steps that must be added before each clade in the MP tree is no longer unequivocally supported. Because of computer limitations the MAXTREE option in PAUP was set to 2850.

Sequence Availability

The nucleotide sequence data reported in this article have been submitted to the EMBL Nucleotide Sequence Library under the following accession numbers: E. fetida, X79872; L. conchilega, X79873; P. granulatum, X79874; O. erythropagummon, X79875; S. fiordicum, X79876; R. piscesae, X79877; Lineus sp., X79878.

Results

Alignment

Trees were constructed on the basis of 22 complete or nearly complete metazoan sequences. To limit computer time needed for the analyses, a selection was made among the 37 representatives of the vertebrates and the 13 bivalves, for which complete 18S rRNA sequences are available in the database of Van de Peer et al. (1994). The sequences of three arthropods (Aedes albopictus, Drosophila melanogaster, and Acyrophis piscum) and six nematodes (Caenorhabditis elegans, Haemonchus contortus, Haemonchus placei, Haemonchus similis, Nematodirus battus, and Strongyloides stercoralis) were excluded because their high evolutionary rates tend to induce systematic errors (Swofford and Olsen 1990; Penny et al. 1991). The alignment can be obtained from the authors.

Neighbor-joining Analysis

On the basis of a more extensive NJ analysis, including also unpublished protostome 18S rRNA sequences and published sequences of Cnidaria, plants, fungi, ciliates, and protists, Platyhelminthes appeared to be a sister group to the monophyletic coelomates (unpublished data). Therefore, the platyhelminth Opisthorchis viverrini was used as an outgroup in all further analyses. The NJ tree derived from the Jukes and Cantor (1969) distances is shown in figure 1. The same topology was obtained with the two-parameter equation of Kimura (1980). It suggests the following. First, head worms, consisting of Pogonophora and Vestimentifera, are monophyletic, and this clade is sister to Echiura. Second, the nonarthropod protostomes (= Pogonophora + Vestimentifera + Echiura + Mollusca + Nemertea + Annelida + Sipuncula) are a monophyletic clade. Third, Arthropoda, Chordata, Vertebrata, Gastropoda, and Bivalvia are monophyletic clades. Neighbor-joining trees on the basis of Tajima and Nei (1984) or Jin and Nei (1990) distances yielded nearly the same topology as in figure 1, except for the monophyletic status of annelids. Indeed, in these trees (not shown), the oligochaete Eisenia fetida branches off before the polychaete Lanicie conchilega. The sensitivity of the tree topology of figure 1 to specific taxa was evaluated by leaving out the ingroup taxa one by one. Exclusion of Homo sapiens or Oedipontis inermis caused Annelida to become a paraphyletic group. Without Crassostrea virginica, the branching order of Eisenia fetida, Lanicie conchilega, and the cluster containing Nemertea, Mollusca, Echiura, Vestimentifera, and Pogonophora could not be determined. When Lineus sp. was removed, the branching order of Phascolosoma granulatum, the annelid cluster and the cluster including the rest of the nonarthropod protostomes was an unresolved trichotomy. However, these changes did not affect the general results since the topological changes only involved those parts of the tree where the bootstrap values were lower than the 70% limit of Hillis and Bull (1993).

Maximum Parsimony Analysis

For the 22 species analyzed, 609 sites were phylogenetically informative. Using Opisthorchis viverrini as an outgroup, the MP analysis found a single MP tree (length = 2,689; consistency index = 0.428) shown in figure 2. The results were broadly congruent with those of the NJ tree (fig. 1). Differences in topology between the MP and the NJ tree were limited to weakly bootstrap supported branching points. In contrast to the NJ topology of the nonarthropod protostome clade (fig. 1), the MP tree (fig. 2) shows the following. First, annelids are paraphyletic, including the sipunculid Phascolosoma granulatum. Second, the nemertine Lineus sp. is a sister taxon to the bivalves. Third, the chiton Acanthopleura japonica branches off after the annelid-sipunculid clade. Furthermore, also in contrast to the NJ tree, Deuterostomia is a sister group to the nonarthropod protostome clade, and Herdmania momus and Stichopus japonicus form a monophyletic group. Ten different data input orders had no effect on the MP results (e.g., Blair Hedges et al. 1991; Templeton 1991; Stoneking et al. 1992).

Table 2 gives the number of cladograms that are one to 13 steps longer than the minimum-length tree of figure 2. The number of extra steps necessary for a monophyletic clade not to be unequivocally supported in the strict consensus tree (Bremer 1988; Donoghue et al. 1992), are indicated in figure 2. Because of computer limitations (MAXTREE = 2,850), it was not possible to pursue this analysis to the point at which the strict consensus tree was totally collapsed. The 2,850 trees found that are 13 steps longer (L = 2,702) than the min-
imum length tree still had several monophyletic clades in common: the gastropods, the bivalves, and the beard worms.

The ingroup taxa of figure 2 were removed one by one and the MP analysis was repeated on the informative sites of each new data set. The consistency index ranged from 0.429 to 0.446 and either one, two, three, or four trees were found (not shown). They differed from each other and from the tree of figure 2 within the deuterostome, the arthropod, and the nonarthropod protostome clades and in the branching order of these three clades. When Herdmania momus was removed, Stichopus japonicus became either a sister group to the Chordata or formed a clade with Branchiostoma floridae. Without Stichopus japonicus, the branching order of Oedignathus inermis, Artemia salina, and Tenebrio molitor remained unresolved. After excluding Tenebrio molitor or Lanice conchilega from the data set, the branching order of arthropods, nonarthropod protostomes, and deuterostomes was an unresolved trichotomy. The removal of any of the nonarthropod protostomes, except Placopecten magellanicus, Limicolaria kamebul, or Siboglinum fiordicum, caused changes within the nonarthropod protostome cluster. However, all the changes caused by removing an ingroup taxon involved branching points with little bootstrap support (less than 70%; Hillis and Bull 1993).

Discussion

Our results suggest that (1) Vestimentifera and Pogonophora are a monophyletic protostome group, having a common ancestor with the Echiura but not with the Annelida, (2) Echiura and Sipuncula are not sister groups, and (3) Nemertea are protostomous coelomates. Measures of reliability firmly support a monophyletic cluster comprising the protostome phyla Echiura, Si-
puncta, Annelida, Mollusca, Vestimentifera, and Pogonophora. Arthropoda do not belong to this clade. The existence of an annelid-arthropod sister group relationship was already questioned by Eernisse et al. (1992) on anatomical and embryological grounds. Our analyses do not even consistently support a sister-group relationship between the Arthropoda and the other protostomes. This suggests that Protostomia may be not monophyletic. However, the exact position of the Arthropoda is still uncertain and requires more data. The eventual polyphyly of Protostomia would be consistent with the result of Valentine (1992), who combined morphological, paleontological, and partial 18S rRNA data to conclude that Arthropoda are a sister group to the Deuterostomia.

The protostome status of beard worms agrees with several cladistic analyses of anatomical and larval features (e.g., Brusca and Brusca 1990, p. 882; Mehlisch and Schram 1991, p. 596; Backeljau et al. 1993; Eernisse et al. 1993). The presence of a posterior metameric body region (opisthosoma) with serially arranged septae, paired setae, and a coelomic cavity in each segment (Webb 1964), which is a typical protostome character, also supports this hypothesis. The protostome status of beard worms suggests that the interpretation of a ventral localization of the nerve cord (Nörrevang 1970; Jones and Gardiner 1988), a forward-running blood flow (Nörrevang 1970), and spiral egg cleavage (Bakke 1980) is correct. The protostome affinities of Vestimentifera were previously suggested by other molecular studies.
Table 2  
Number of Cladograms Which Are One to 13 Steps Longer than the Minimum-Length Tree of Figure 2

<table>
<thead>
<tr>
<th>Length</th>
<th>Number of Cladograms Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,690</td>
<td>4</td>
</tr>
<tr>
<td>2,691</td>
<td>4</td>
</tr>
<tr>
<td>2,692</td>
<td>10</td>
</tr>
<tr>
<td>2,693</td>
<td>27</td>
</tr>
<tr>
<td>2,694</td>
<td>51</td>
</tr>
<tr>
<td>2,695</td>
<td>113</td>
</tr>
<tr>
<td>2,696</td>
<td>189</td>
</tr>
<tr>
<td>2,697</td>
<td>334</td>
</tr>
<tr>
<td>2,698</td>
<td>517</td>
</tr>
<tr>
<td>2,699</td>
<td>933</td>
</tr>
<tr>
<td>2,700</td>
<td>1448</td>
</tr>
<tr>
<td>2,701</td>
<td>2260</td>
</tr>
<tr>
<td>2,702</td>
<td>&gt;2850</td>
</tr>
</tbody>
</table>

Note.—For 13 steps longer, \( L = 2,690-2,702 \). Minimum-tree length \( L = 2,689 \).

but the relationships with other protostomes remained uncertain. The 5S rRNA data (e.g., Lane et al. 1985) suggested a close relationship between Mollusca and Vestimentifera. Analyses of extracellular hemoglobins (Terwilliger et al. 1980, 1985; Suzuki et al. 1988, 1989) and elongation factor-1α (Kojima et al. 1993) pointed to a close affinity between Vestimentifera and Annelida. Partial 18S rRNA data (Field et al. 1988) were insufficient to identify the sister group of the Vestimentifera. Our 18S rRNA trees contradict a specific relationship between beard worms and either Oligochaeta or Polychaeta, as has often been suggested (e.g., Liwanow and Porfirjewa 1967; Nörrevang 1970; George and Southward 1973; Gupta and Little 1975; Van der Land and Nörrevang 1975; Jones 1985; Jones and Gardiner 1988; Southward 1988; Brusca and Brusca 1990, p. 457; Conway Morris 1993), but support a sister group relationship between beard worms and Echiura. This suggests that, for example, the typical “annelid” setae evolved more than once.

The 18S rRNA analyses confirm that Pogonophora and Vestimentifera form a monophyletic clade (e.g., Southward 1988; Brusca and Brusca 1990, p. 882; Conway Morris 1993). This is in accordance with the presence of several shared anatomical traits such as the absence of a digestive system in adults and the presence of a metameric opisthosoma. Therefore, this study supports the unification of Pogonophora and Vestimentifera in a single phylum (e.g., Southward 1988; Brusca and Brusca 1990; Meglitsch and Schram 1991, p. 300).

Our 18S rRNA analyses confirm the protostome status of Echiura and Sipuncula, as already suggested by embryological features (e.g., Rice 1985) and by several cladistic anatomical studies (e.g., Brusca and Brusca 1990, p. 882; Eernisse et al. 1992; Backeljau et al. 1993). Previous analyses of partial 18S rRNA sequences (Field et al. 1988; Ghiselin 1988; Patterson 1989) also placed Sipuncula in a protostome cluster but could not resolve the relationships of Sipuncula to other protostomes. A reanalysis of these data by Lake (1990) positioned Sipuncula closest to the Mollusca and Brachiopoda, with Annelida and Vestimentifera as sister groups. This evolutionary relationship between Mollusca and Sipuncula was in accordance with some embryological and larval features (Inglis 1985; Scheltema 1993). However, on the grounds of other embryological and larval similarities like the pattern and size of the prototroch cells, the formation of the apical plate, and the ventral nerve cord, the derivation of the larval cuticle from the egg membrane, and the inverted larval photoreceptors, Clark (1969) suggested a close affinity between Sipuncula and Annelida. The present analyses based on complete 18S rRNA sequences also fail to resolve the relationships of Sipuncula to other protostomes. The NJ tree suggests, although it is not bootstrap-supported, that Sipuncula branch off first within the nonarthropod protostome clade. However, the MP analyses show a monophyletic Sipuncula-Polychaeta group. This would confirm the affinity between Sipuncula and Annelida as suggested by Clark (1969). The absence of segmentation in Sipuncula would then be a secondary loss, probably associated with the exploitation of a sedentary, burrowing lifestyle. However, many Annelida have retained their metamersim despite burrowing lifestyles (Brusca and Brusca 1990, p. 456). More data are necessary to resolve this issue.

Anatomical data (e.g., Rice 1985) suggested a common premetameric ancestor for Sipuncula and Echiura. Similarly, 5S rRNA data (Ohama et al. 1984; Hendriks et al. 1986; Hori et al. 1988) predicted a sister-group relationship of Echiura and Sipuncula. However, our study points to a distinct origin of Sipuncula and Echiura and suggests that the anatomical and embryological resemblance of both groups is the result of convergence rather than common ancestry. A common ancestor for Sipuncula and Echiura was also questioned, for example, by von Salvini-Plawen (1980) and Inglis (1985) on the basis of embryological features, while the combined morphological, paleontological, and morphological study of Conway Morris (1993) yielded no evidence for a sister-group relationship between both groups.

Concurrent with the results of a study using 946 18S rRNA positions (Turbeville et al. 1992), complete 18S rRNA sequences suggest that Nemertea are protostomal coelomates. This result supports the hypothesis that the rynchocoel and the circulatory system are coelom homologues (e.g., Turbeville 1986).
The complete 18S rRNA sequences cast doubt on two widespread viewpoints, the monophyly of Annelida and Mollusca. Most anatomical studies consider Annelida to be a monophyletic group (e.g., Brusca and Brusca 1990, p. 432), mainly on the basis of the typical annelid head and the presence of epidermal paired setae. However, this monophyly could not be corroborated by partial 18S rRNA sequences (Field et al. 1988; Patterson 1989; Lake 1990; Turbeville et al. 1992). Our 18S rRNA analyses also yielded contradictory results on the monophyly or paraphyly of the Annelida and are not sufficient to resolve this question.

Different mitochondrial gene sequences (Ballard et al. 1992; Lecanidou et al. 1994), partial 18S rRNA sequences (Field et al. 1988; Patterson 1989; Lake 1990), and SS rRNA sequences (e.g., Ohama et al. 1984; Hendriks et al. 1986; Hori and Osawa 1987; Hori et al. 1988) have also cast doubt on the monophyly of the Mollusca. Although not supported by the different robustness tests, the current complete 18S rRNA sequences do not provide evidence for the molluscan monophyly either. Yet, molluscan monophyly is generally accepted on the basis of anatomical data (e.g., Milburn 1960; Götting 1980; von Salvini-Plawen 1990). The resolution of the discrepancies between anatomical and molecular data, with respect to the monophyly of Annelida and Mollusca, will have to await the availability of further anatomical and molecular data.

Finally, the unstable results concerning the relationships within the protostome clade may be due to (1) the supposed fast radiation of the protostome clade during a very short time span in the Cambrian (e.g., Erwin 1991; Valentine 1992) and/or (2) the limited number of taxa presently available for analysis (see Lecointre et al. 1993). Obviously, the latter point may be overcome as soon as more 18S rRNA sequences become available.

Acknowledgments

We are indebted to A. Frías Martins (University of the Azores, Portugal), F. Gentil (Biological Station, Roscoff, France), B. Morton (Swire Marine Laboratory, University of Hong Kong), W. Ponder (Australasian Museum, Sydney, Australia), and T. Rodriguez (University of Santiago de Compostela, Spain) for collecting part of the material. We also thank D. R. Dixon (Plymouth Marine Laboratory, U.K.) for providing us with DNA from R. piscesae and P. W. H Holland (University of Oxford) for tissue samples of S. fiordicum. These investigations were performed in the framework of the Institute for the Study of Biological Evolution of the University of Antwerp and were supported by the Fund for Joint Basic Research (Belgium) grants 2.0023.94 and 2.0030.93. B.W. holds a scholarship at the Institute of Science and Techniques (I.W.T.).

LITERATURE CITED


RODNEY L. HONEYCUTT, reviewing editor

Received August 1, 1994

Accepted February 24, 1995