

Phylogenetic analysis of the mitochondrial cytochrome *c* oxidase subunit 1 gene from 13 sipunculan genera: intra- and interphylum relationships

Joseph L. Staton^a

Department of Biology, University of Michigan, Ann Arbor, MI 48109-1048, USA

Abstract. Sipunculans are a phylum of non-segmented, marine worms. Although they are well characterized morphologically, relationships within the phylum and the relationship of Sipuncula to other spiralian phyla have been strongly debated. I analyzed representatives of 13 of 17 described genera using a 654-bp fragment of the mitochondrial gene, cytochrome *c* oxidase subunit I, to construct the first intraphylum phylogenetic hypothesis for sipunculans based on molecular sequence data. Within the phylum, tree topologies are loosely congruent with a previously published morphological analysis, except that the monotypic genus *Phascolopsis* occurred within the Golfingiaformes as a sister group to, or nested within, the Themistidae. Phylogenetic analyses, including 30 sequences from additional invertebrate taxa, suggest that sipunculans are most closely related to the Annelida (including Echiura). A previously proposed sipunculan-molluscan relationship is not supported. While not universally accepted, this hypothesis is consistent with other recent and past data on phylum-level relationships.

Additional key words: peanut worm, phylogeny, mitochondrial DNA, character analysis, spiralian evolution

Sipuncula (=peanut worms) is a phylum of non-segmented worms that occurs in nearly every marine environment, from the intertidal zone to hadal trenches. Sipunculans are filter- or deposit-feeding burrowers or crevice dwellers that often tunnel through marine sediment or bore into harder substrates ranging from dead coral to a decaying whale's skull (Gibbs 1987). Members of the phylum possess a complete, spirally coiled, U-shaped gut and a simple body plan consisting of two major regions: a fully retractable, anterior introvert and a posterior trunk. The gut coils around a central spindle muscle, and the introvert, which terminates anteriorly in a tentacular crown, retracts entirely into the trunk by means of long retractor muscles.

The systematic relationship of the Sipuncula to other metazoans has been unresolved from the earliest descriptions of these worms. Historically, sipunculans were confused with sea cucumbers (=holothurians) (Lamarck 1816), considered as a subgroup within the segmented worms (=annelids) (Delle Chiaje 1823), and then as a component of the Gephyrea (Quatrefages 1847) along with echiuran worms, sternaspid annelids,

and priapulans. Since their elevation to the level of phylum by Sedgwick (1898), sipunculans have been considered as spiralian phyla either closely allied to the annelids on the basis of biochemical similarities (Florkin 1976, Henry 1987), or as a sister group to molluscs on the basis of developmental similarities ("the molluscan cross" pattern of cleavage; Gerould 1903; Heath 1899; Rice 1975, 1985; Scheltema 1993) and hypothesized homologous structures between larval sipunculans and larval and adult molluscs: a ventral buccal organ homologized with the odontophore of molluscs, the ciliated lip below the mouth of the pelagosphaera homologized with the foot of a pediveliger, and the lip glands homologized with pedal glands of larval chitons and adult neomenioids (Scheltema 1996).

Relationships among the 17 described genera, which comprise 162 species and subspecies (Cutler 1994), have received only limited attention. Subdivision of sipunculans into families was first proposed over 130 years ago (Baird 1868) while they were still a part of the Gephyrea inermia. Not until this century did Pickford (1947) suggest dividing sipunculan genera into four "groups" based on the occurrence and position of ectodermal thickenings or "plates." She considered sipunculan genera to be so closely related that attempts at assigning genera to families were "doomed to fail-

^a Present address:

University of South Carolina, 801 Carteret Street, Beaufort, SC 29902. E-mail: jstaton@sc.edu

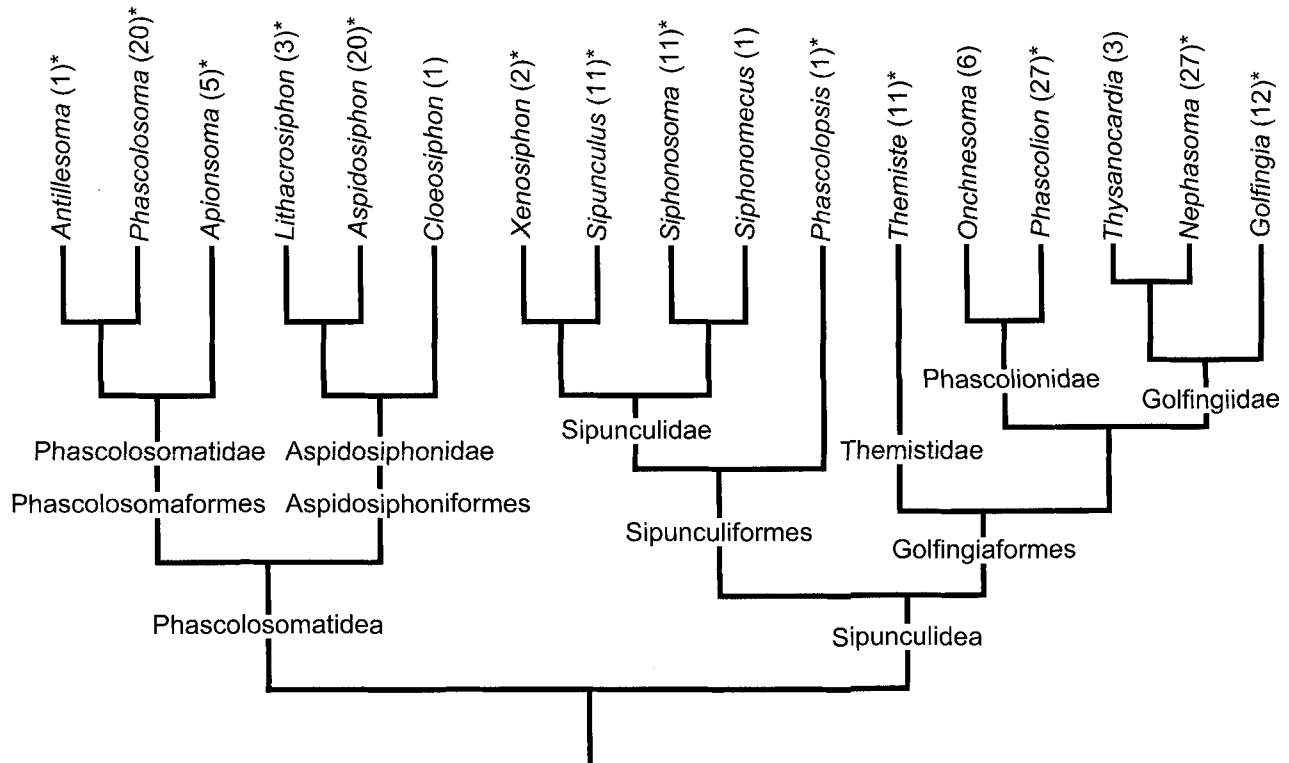


Fig. 1. Phylogenetic relationships of sipunculan genera based on morphologic traits (modified from Cutler & Gibbs 1985). Numbers in parentheses denote the number of reported species and subspecies for each genus in Cutler (1994).

ure" (Pickford 1947: 718A). Åkesson (1958) proposed division of sipunculans into 3 groups, different from those of Pickford, based on characteristics of the nervous system and epidermal organs. Neither author issued formal taxonomic names for their groupings.

Stephen & Edmonds (1972) named 4 families within the phylum Sipuncula, and Cutler & Gibbs (1985; hereafter referred to as C&G) proposed the first phylogeny of sipunculan genera based on phylogenetic analysis of 12 morphological characters (Fig. 1). Gibbs & Cutler (1987) used the resulting phylogeny to erect a classification scheme composed of six families distributed among four orders in two classes. Recently, these classifications were further refined (Cutler 1994).

In the present study, I evaluate the morphology-based phylogenetic hypothesis of C&G by comparison to an analysis of a 654-bp region of the mitochondrial cytochrome *c* oxidase subunit I gene (CO1), which has been used to resolve ambiguous relationships among polychaete taxa (Black et al. 1997). The resulting molecular phylogeny, the first of its kind examining within-phylum relationships for the Sipuncula, serves as an independent assessment of evolution within the Sipuncula and a basis to evaluate the morphological characters analyzed by C&G. I use a variety of taxa from

previously published studies to identify a sister group for the sipunculans.

Methods

Molecular procedures

I isolated total DNA from sipunculans using a modified CTAB method described in Collins et al. (1996). I used the primers of Folmer et al. (1994) to amplify a 710-bp fragment of the CO1 gene for the following 14 species (representing 13 genera) of sipunculan worms: *Antillesoma antillarum* (GRÜBE & OERSTED 1858), *Apionsoma misakianum* (IKEDA 1904), *Aspidosiphon laevis* QUATREFAGES 1865, *Golfingia elongata* (KEFERSTEIN 1862), *Lithacrosiphon cristatus* (SLUITER 1902), *Nephasoma rimicola* (GIBBS 1973), *Phascolion cryptum* HENDRIX 1975, *Phascolopsis gouldii* (PORTALÉS 1851), *Phascolosoma nigrescens* (KEFERSTEIN 1865), *Siphonosoma cumense* (KEFERSTEIN 1865), *Sipunculus nudus* LINNAEUS 1766, *Themiste alutacea* (GRÜBE & OERSTED 1858), *Themiste lageniformis* (BAIRD 1868), and *Xenosiphon branchiatus* (FISCHER 1895). Genera not treated here are *Siphonomecus*, *Thysanocardia*, *Cloeosiphon* (see following paragraph), and *Onchnesoma*.

Polymerase chain reaction (PCR) was performed under the following cycling conditions: 94°C for 30 s, 40–47°C for 30 s, and 72°C for 1 min for a total of 40 cycles. [Mg²⁺] varied, depending on species, from 1.5–3.0 mM. Cycling was performed on either a Perkin-Elmer 9600™ or an MJ-Research DNA Engine™. Manual dideoxy-sequencing was performed on single stranded products using individual PCR primers, modified T7 polymerase (Sequenase™), and standard protocols. Products were labeled with [³⁵S]dATPaS (1000 Ci/mmol; New England Nuclear) and separated on 60-cm, 6% acrylamide gels. Some sequences were analyzed by cycle sequencing double-stranded fragments using PCR primers and ABI PRISM™ dye terminator chemistry according to manufacturers specifications, with subsequent analysis on ABI 373 and 377 automated sequencers (ABI/Perkin-Elmer). In the case of *G. elongata* and *N. rimicola*, reamplifications were carried out on gel-purified products (using an annealing temperature of 52°C) to eliminate smaller secondary bands incurred during the initial amplification procedure. One other species, *Cloeosiphon aspergillus* (QUATREFAGES 1865), never yielded a distinct CO1 product, despite repeated attempts and multiple CTAB extractions. Sequences produced for this work are deposited in GenBank (Accession number AF374337 for *P. gouldii*, as reported in Boore & Staton (2002), and AY161122–AY161134).

Phylogenetic analyses

Nucleic acids were translated into amino acids using the invertebrate mitochondrial genetic code and aligned; the amino-acid alignment was then used to align corresponding triplet nucleotide codons by eye. Another sipunculan sequence previously published for CO1 (*Phascolosoma* sp. GenBank Accession No. U74071; Black et al. 1997) was included in this phylogenetic analysis. Sequences from 30 non-sipunculan species were obtained from GenBank for comparison as possible outgroup species and to address issues of higher-level phylogeny, which fall into annelid, echiuran, molluscan, arthropod, and echinoderm groups. Previously published sequences used in this study are as follows (Species name [Accession No.]): Annelids—*Ridgeia piscesae* [AF022235], *Tevnia jerichonana* [U74075], *Riftia pachyptila* [U74053], *Galathealinum brachiosum* [U74066], *Paralvinella palmiformis* [U74070], *Tubifex tubifex* [U74076], *Hirudo medicinalis* [U74067], *Piscicola geometra* [AF003280], *Desserobdella phalera* [AF003278], *Lumbriculus variegatus* [AF003257], *Pontodrilus bermudensis* [AF003256], *Lumbricus terrestris* [U24570]; Echiuran—*Urechis* sp. [U74077]; Molluscs—*Rangia cuneata* [U47652], *Mercenaria mercenaria* [U47648],

Geukensia demissa [U56844], *Mytilus californianus* [U68776], *Calyptogenia* sp. [AF008297], *Gyraulus albus* [Y14710], *Solemya velum* [U56852], *Lepetodrilus elevatus* [U56846], *Oliva sayana* [U86333], *Arctomelon stearnsii* [U86334], *Katharina tunicata* [U09810]; Arthropods—*Daphnia pulex* [Z15015], *Artemia franciscana* [X69067], *Anopheles gambiae* [L20934], *Drosophila yakuba* [X03240]; and Echinoderms—*Strongylocentrotus purpuratus* [X12631], and *Parastichopus parvimensis* [U32199].

The phylogenetic analysis was carried out using maximum parsimony (MP) for amino acids and neighbor-joining (NJ; Saitou & Nei 1987) with Log-determinant/paralinear distances (LD/P; Steel 1994, Lake 1994) on first and second nucleotide positions using PAUP* ver. 4.0b4a (Swofford 2000). Searches for maximum likelihood (ML) trees of amino acid data were conducted using PAML (Yang 1997) with branch lengths estimated from a Jones-Taylor-Thorton (1992) substitution matrix for amino acids. The procedure uses a heuristic tree search by nearest neighbor interchange perturbations to optimize the log-likelihood score to find the tree most consistent with the model of evolution. Analytical package limited analysis to a total of thirty taxa.

Tests of phylum-level relationships were investigated using the Kishino-Hasegawa procedure (Kishino & Hasegawa 1989) and a likelihood-mapping procedure (Strimmer & von Haeseler 1997). A Kishino-Hasegawa (1989) procedure compares the shortest trees (100 heuristic searches using random addition) produced under the constrained searches of ((annelid, sipunculan) remaining taxa) and ((mollusc, sipunculan) remaining taxa) under maximum likelihood criteria in PAUP*. The model of evolution applied to these data was derived by MODELTEST (ver. 1.05; Posada & Crandall 1998). Results of that test (both AIC and likelihood ratio tests) described the best-fit model of evolution to be a general time reversible model (reviewed in Yang 1994) with rate heterogeneity (GTR + Γ ; P < 0.000001) without implementing a molecular clock constraint. Although recent research has demonstrated that the application of the Kishino-Hasegawa procedure recursively to likelihood-derived trees is flawed (Goldman et al. 2000), the trees compared here are the best-constrained trees derived from MP methods. A likelihood mapping procedure was implemented in PUZZLE (Strimmer & von Haeseler 1996, 1997; Strimmer et al. 1997). This procedure randomly samples one taxon from each of four constrained groups (Sipuncula, Annelida, Mollusca, and Arthropoda) for a maximum-likelihood analysis of each sampled quartet, in this case 10,000 random quartets. Level of support for each of the three possible branch patterns is noted by the percent of random quartets favoring each arrangement.

Results

Of the 219 inferred amino acids, 168 were variable and 132 were assessed as parsimony informative for the 45 CO1 sequences analyzed. Within the Sipuncula, 58 amino-acid sites were parsimony informative. An MP heuristic search (100 replicate searches with random additions) using inferred amino acid data yielded 300 most parsimonious trees (952 steps, CI=0.5431, CI excluding uninformative characters=0.5101), for which a majority consensus tree is reported in Fig. 2. In all trees, 14 of 15 sipunculan sequences form a sister group to either the *Urechis* sequence, alone, or to a clade of *Urechis* and oligochaete sequences. In these analyses, I also note consistently the resolution of monophyletic polychaete, arthropod, and leech clades (Figs. 2, 3, & 4).

There was a strong base compositional bias found when analyzing all 654 aligned nucleotides ($\chi^2=671.44$ for 132 d.f.; $p \ll 0.001$), which can produce misleading phylogenetic analyses (Collins et al. 1994). By excluding third positions, this bias was lessened ($\chi^2=129.68$ for 132 d.f.; $p=0.541$). Tree topologies based on MP analysis of first and second nucleotides produced 12 trees (208 parsimony informative characters; 1412 steps; consistency index [CI]=0.3548, CI excluding uninformative characters=0.3114); the consensus topology (not shown) was similar to that of ML analysis on amino-acid data within sipunculans and for the major branching patterns where resolved (see Fig. 3). Analysis using all DNA data produced trees with only 12 of 15 sipunculan genera clustered (tree not shown), and therefore analyses with third positions were not considered further.

A single sipunculan species, *Phascolion cryptum*, clusters with several bivalve species: *Rangia cuneata*, *Mercenaria mercenaria*, *Geukensia demissa*, *Mytilus californianus*, and *Calyptogena* sp. (Fig. 2). The bivalves grouped with *P. cryptum* are, in turn, separate from the other molluscan sequences as is a basal, isolated pulmonate snail, *Gyraulus albus* (Fig. 2). In the case of the bivalve/sipunculan group, *G. demissa*, *M. californianus*, and *Calyptogena* sp. are inferred to share an amino acid insertion, based on published sequences and those herein. *R. cuneata* and *M. mercenaria* are inferred to have lost an adjacent amino acid, whereas *G. albus* is inferred to have lost 2 amino acids in this region. All sequences in this clade have high pairwise (uncorrected) distances from the remaining ingroup taxa (Table 1). Distances ranged from 0.296–0.463 among these taxa and the sipunculan taxa, alone, and from 0.258–0.463 among these taxa and all others. *Phascolion cryptum* had the highest pairwise distances among the sipunculans ranging from 0.325–0.402 be-

tween it and other sipunculans, but slightly lower distances between it and the aforementioned, problematic molluscan taxa (0.305–0.382). This resulted in large corrected distances and produced a basal position in NJ trees for *P. cryptum* joined with the bivalves and a more distant *G. albus* (Table 1; trees not shown). Under the assumption that sipunculans are monophyletic, I chose to eliminate *Phascolion* and the other associated clustered sequences from subsequent analyses as potential confounding factors.

The remaining 14 sipunculan taxa form a clade which is supported by MP amino acid analysis (Fig. 2), ML analysis on amino acids (Fig. 3), and NJ topologies on first and second nucleotide codon positions (Fig. 4). The topology of NJ for mean distances of amino acids is not shown but similar to the pattern in Fig. 2 except that it places *Antillesoma* closer to the two *Phascolosoma* species, places *Themiste lageniformis* as sister to the *Phascolopsis-T. alutacea* clade, and places *Siphonosoma cumanense* basal to the Sipunculidae. Likewise, a ML tree based on first and second codon positions (GTR + Γ model, not shown), constraining monophyly within groups and omitting deuterostomes, produces a similar tree to the NJ topology—except that *Oliva* groups with *Arctomelon* with *Katharina* basal to them, and *Urechis* is basal to the Polychaete/Hirudinea clade. Within sipunculans, the *Phascolosoma* species are grouped as sister taxa, *Siphonosoma* is basal to the Phascolosomatidea and *Phascolopsis* is basal to the Themistidae/Golfingidae clade.

Within the sipunculans, placement of *S. cumanense* differs the most among methods as either being basal to other sipunculans (Fig. 2, but note would be unresolved in a strict consensus tree and <50% bootstrap support), basal to the Phascolosomatidea (Fig. 3), or basal in the Sipunculidae (Fig. 4; again, <50% bootstrap support). Its placement and the arrangement of other sipunculan taxa resulting from NJ of LD/P distances (Fig. 4) coincides with the previous morphological hypothesis (C&G). The inconsistency of placement of *S. cumanense* among other sipunculan genera for different methods parallels the pattern noted previously for morphological data in the analysis by C&G.

In this analysis, sipunculans group with annelid taxa. Although annelid groups appear as paraphyletic with sipunculans (Figs. 2, 4), bootstrap support for this paraphyly is low for MP analysis of amino acids ($\leq 55\%$) and neighbor-joining methods for first and second codon positions (<50%). If a strict consensus tree is followed (i.e., poorly supported nodes are collapsed), the sipunculan/annelid group is an unresolved polytomy. However, support for the basal node group-

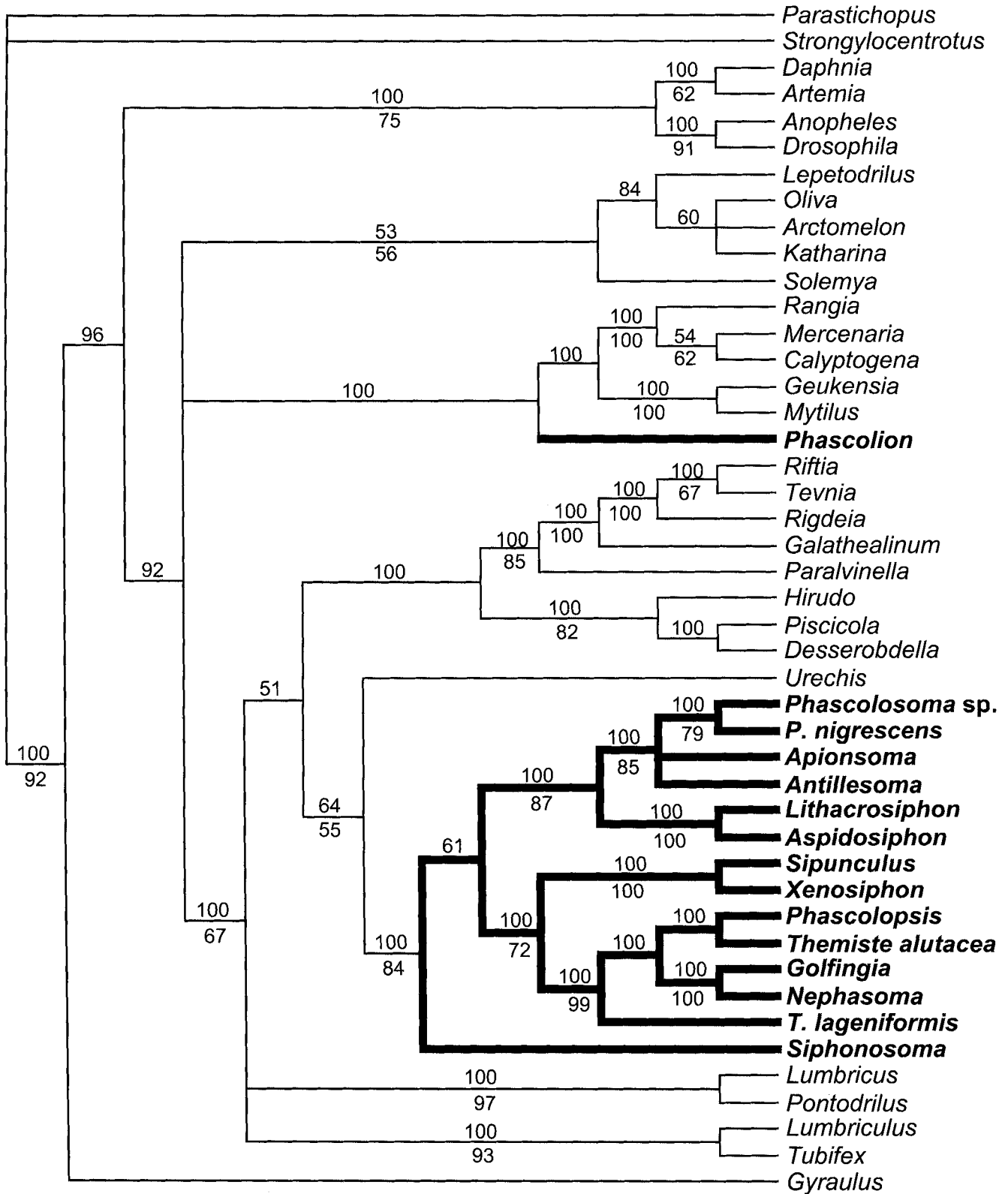


Fig. 2. Majority rule consensus tree (support >50% shown above branches) of a MP heuristic search using random addition for 100 replicate searches on inferred amino acids for 45 taxa. 100 random additions of taxa produced 300 equally parsimonious trees. Numbers below branches indicate parsimony support for each node for 100 bootstrap replicates of 100 random-addition full heuristic searches per bootstrap (support <50 not shown). Sipunculan taxon names are highlighted.

ing the sipunculan/annelid clade is stronger for analyses of both amino acids (67%; Fig. 2) and nucleotides (74%; Fig. 4).

The Kishino-Hasegawa (1989) procedure using maximum likelihood was performed on first and second position nucleotide data, excluding putative long-branch sequences in PAUP*. The searches (100 heuristic searches using random addition) found 12 most parsimonious trees (length = 911) when constrained to find ((sipunculan, annelid) other) trees and 4 most parsimonious trees (length = 918) when constrained to find ((sipunculan, mollusc) other) trees. The most parsimonious trees supporting the (sipunculan, annelid) relationship had better maximum likelihood scores (P ranging from 0.077–0.025) than the best-supported (sipunculan, mollusc) trees.

As a further test of the sipunculan/annelid relationship, results from likelihood mapping (Strimmer & von Haeseler 1997) were used. Here, previously mentioned problematic taxa as well as echinoderms were excluded from the analysis, with 36 taxa remaining. These taxa were constrained into 4 taxonomic groups as sipunculans, annelids (including the echiuran), molluscs, or arthropods. Using PUZZLE, I tested 10,000 quartets, randomly sampling one taxon from each of the four groups per replicate. A gamma distribution was used to approximate rate heterogeneities (4 categories) and the rate parameter was estimated by PUZZLE directly from the data set. While some sampling bias is present in the numbers of taxa compared within each group, >5 times as many quartets tested (70.4% for sipunculan/annelid vs. 12.3% for sipunculan/molluscan arrangements) favor a sipunculan/annelid relationship over that of sipunculans and molluscs.

Discussion

Sipunculans have essentially no fossil record, so understanding their history has been problematic. The morphologic hypothesis of C&G provides insight into patterns of evolution without having the ability to address the relative rate of their evolution. Although interpretations based on a strict molecular clock are highly suspect with the best of datasets, sipunculans have similar levels of molecular diversity to those found within annelids, molluscs, and arthropods (Fig. 4). Therefore, I conclude that the phylum Sipuncula is not of recent origin. Its genera possess the same level of molecular divergence among them as among groups of the annelids or even among the arthropods in this analysis, all of which have a substantial fossil record supporting their age of divergence back to the Cambrian or earlier.

The analysis for the 219 amino acids of CO1 agrees

with the morphological hypothesis of C&G, in large part. All analyses performed produced a grouping of Phascolosomatidea sister to all other sipunculans other than *Phascolion cryptum*. The NJ topology based on LD/P distances from first and second codon positions produced a result that is most consistent with prior results within sipunculans (C&G).

As noted in results, several mitochondrial sequences were greatly divergent from related taxa, including *Phascolion cryptum* and several molluscan species. Given this pattern in the mitochondrial gene sequence, branches of these taxa are “attracted” to one another by non-homologous, parallel similarities in their sequences (Felsenstein 1978, see Graybeal 1998 for discussion). However, a traditional solution of adding more taxa to the dataset (Hendy & Penny 1989) may not alleviate the problem. The mitochondrial gene order reported for the bivalve *Mytilus* (Hoffmann et al. 1982) and other molluscan species is highly derived relative to other members of the phylum and some species possess a unique mode of mitochondrial inheritance (doubly uniparental inheritance; Zouros et al. 1994). I infer that certain molluscan taxa represent unusual gene evolution for members of their phylum that precludes the use of CO1 to derive their evolutionary relationships.

It may surprise some that topological results within Sipuncula are largely congruent with the phylogeny presented by C&G. Their morphological analysis consisted of 12 characters, totaling 29 character states, used to evaluate relationships among 17 genera. However, an outgroup to polarize morphological characters within sipunculans was unknown. To overcome this difficulty, C&G proposed a hypothetical ancestral sipunculan to polarize characters into ancestral and derived states. A refined set of ancestral character states has since been discussed at length (Cutler 1994). The high degree of similarity between these results and C&G confirms that morphological characters, even if few, can provide an accurate estimate of systematic relationships among divergent taxa.

According to the NJ topology, this molecular phylogeny supports a grouping of the Phascolostomatidae and the Aspidosiphonidae to form the Phascolosomatidea (Fig. 4), one of two major lineages within sipunculans (C&G). The sequences analyzed also support the group Sipunculidea—with the Sipunculiformes sister to the Golfingiaformes (including Golfingiidae and Themistidae). The major difference between this analysis and that of C&G is the association of *Phascolopsis gouldii* within the Golfingiiformes as a relative of the Themistidae, as well the inconsistent placement of *Siphonosoma cumanese* (see later in discussion). In their original analysis, C&G show *P. gouldii* to orig-

Table 1. Pairwise distances between sipunculans and representative outgroup taxa for nucleotide data (including third positions and gaps) (upper right half—genetic distances for first and second positions corrected by the LogDet/paralinear method; lower left half—pairwise [uncorrected] distances). Columns and rows headed by bold numbers and containing italicized numbers are those of long-branch taxa.

	1	2	3	4	5	6	7	8	9	10
1 <i>Rangia</i>	—	0.220	0.402	0.418	0.231	0.353	0.364	0.358	0.383	0.374
2 <i>Mercenaria</i>	0.268	—	0.438	0.482	0.224	0.388	0.362	0.366	0.351	0.375
3 <i>Geukensia</i>	0.379	0.385	—	0.234	0.495	0.379	0.393	0.375	0.380	0.366
4 <i>Mytilus</i>	0.410	0.426	0.316	—	0.510	0.414	0.399	0.410	0.397	0.410
5 <i>Calyptogenia</i>	0.297	0.278	0.391	0.438	—	0.368	0.362	0.378	0.415	0.392
6 <i>Gyraulus</i>	0.354	0.349	0.361	0.383	0.335	—	0.287	0.262	0.292	0.262
7 <i>Paralvinella</i>	0.402	0.408	0.415	0.402	0.408	0.348	—	0.173	0.210	0.204
8 <i>Urechis</i>	0.415	0.411	0.401	0.397	0.411	0.313	0.270	—	0.170	0.175
9 <i>Phascolosoma</i> sp.	0.424	0.404	0.413	0.419	0.424	0.354	0.333	0.283	—	0.058
10 <i>P. nigrescens</i>	0.406	0.410	0.394	0.406	0.392	0.309	0.315	0.258	0.224	—
11 <i>Sipunculus</i>	0.428	0.423	0.410	0.423	0.418	0.369	0.328	0.294	0.294	0.284
12 <i>Siphonosoma</i>	0.421	0.417	0.420	0.421	0.415	0.325	0.294	0.267	0.265	0.247
13 <i>Xenosiphon</i>	0.463	0.454	0.448	0.436	0.437	0.389	0.329	0.297	0.297	0.295
14 <i>Apionsoma</i>	0.400	0.419	0.393	0.407	0.398	0.332	0.307	0.272	0.242	0.231
15 <i>Lithacrosiphon</i>	0.337	0.368	0.353	0.376	0.355	0.296	0.327	0.308	0.293	0.259
16 <i>Phascolion</i>	0.382	0.357	0.355	0.365	0.345	0.305	0.381	0.361	0.370	0.365
17 <i>Phascolopsis</i>	0.393	0.392	0.403	0.413	0.402	0.340	0.301	0.266	0.279	0.284
18 <i>Them. alutacea</i>	0.417	0.413	0.433	0.436	0.431	0.382	0.311	0.302	0.303	0.289
19 <i>T. lageniformis</i>	0.416	0.430	0.422	0.437	0.421	0.358	0.311	0.302	0.293	0.291
20 <i>Golfingia</i>	0.393	0.403	0.414	0.424	0.390	0.339	0.301	0.299	0.313	0.286
21 <i>Nephasoma</i>	0.380	0.391	0.399	0.414	0.381	0.327	0.296	0.291	0.305	0.273
22 <i>Aspidosiphon</i>	0.335	0.370	0.351	0.373	0.357	0.299	0.324	0.303	0.291	0.263
23 <i>Antillesoma</i>	0.409	0.436	0.433	0.434	0.442	0.362	0.317	0.290	0.242	0.235

inate on their tree (C&G: p.170, their Fig. 3) between the Sipunculidae and the rest of the Sipunculidea. The presence of longitudinal muscles in a banded pattern along the body wall supports its placement in their study with the Sipunculiformes. C&G note that if occurrence of this character is homoplasious, *P. gouldii* would be considered more closely related to the Golfingiiformes than the Sipunculiformes, which is more similar to the present findings. The lack of coelomic extensions in *P. gouldii* would support the exclusion of *Phascolopsis* from the Sipunculidae on a morphological basis, which is consistent with this analysis. It is also of anecdotal interest that *P. gouldii* was formerly considered a species within *Golfingia* before its placement in its monotypic genus.

Of morphological characters used previously to analyze patterns within sipunculans (C&G, Cutler 1994), the type of tentacular arrangement, presence/absence of hooks in rings around the introvert, and characteristics of the spindle muscle are further supported by this analysis as synapomorphies separating the Phascolosomatidea from the Sipunculidea. In all Phascolosomatidea, there is a similar tentacular arrangement with the mouth positioned ventral to an apical circle of tentacles, whereas in the Sipunculidea, the mouth is

encircled, wholly or partially, by tentacles. Baltzer (1931) was the first to use tentacular arrangement as a systematically important taxonomic character to distinguish species of sipunculans.

All members of the Phascolosomatidea, except individuals of *Antillesoma antillarum*, possess complex introvert hooks that are arranged in orderly rings around the apical portion of the introvert, whereas in the Sipunculidea, for the most part, hooks in rings around the introvert are lacking. It is important to note that, in *A. antillarum*, *Phascolopsis gouldii*, *Phascolosoma meteori*, and *Themiste alutacea*, hooks have been found in early juveniles but do not persist in the adults (Cutler 1994). In other polymorphic genera, some species have papillae arranged in rings on the introvert not unlike the pattern of hooks found in the Phascolosomatidea (Cutler 1994). From these patterns, Cutler (1994) hypothesized that introvert hooks in an ordered array are plesiomorphic, citing the early appearance of similar hook structures in *Ottoia*, a fossil from the Burgess shale, and living acanthocephalan worms, kinorhynch, and priapulid worms. However, homology of anterior hooks in these widely divergent taxa is questionable, as some view *Ottoia* as a stem-group priapulid (e.g., Bruton 2001).

Table 1. Extended.

11	12	13	14	15	16	17	18	19	20	21	22	23
0.381	0.397	0.417	0.366	0.302	0.378	0.371	0.369	0.368	0.357	0.341	0.300	0.394
0.375	0.362	0.412	0.378	0.329	0.369	0.365	0.349	0.372	0.361	0.344	0.344	0.407
0.370	0.394	0.394	0.376	0.333	0.403	0.397	0.389	0.396	0.427	0.402	0.328	0.407
0.389	0.433	0.412	0.395	0.383	0.384	0.412	0.431	0.428	0.440	0.422	0.378	0.457
0.382	0.374	0.383	0.379	0.341	0.394	0.358	0.343	0.351	0.350	0.333	0.341	0.411
0.269	0.254	0.284	0.280	0.253	0.288	0.294	0.295	0.288	0.287	0.269	0.262	0.297
0.197	0.202	0.210	0.212	0.208	0.319	0.194	0.190	0.197	0.209	0.187	0.206	0.223
0.168	0.162	0.172	0.176	0.195	0.276	0.155	0.174	0.183	0.207	0.197	0.191	0.181
0.149	0.124	0.153	0.078	0.139	0.296	0.156	0.159	0.167	0.195	0.181	0.145	0.067
0.148	0.136	0.149	0.076	0.115	0.287	0.165	0.150	0.171	0.201	0.185	0.115	0.062
—	0.132	0.044	0.160	0.164	0.277	0.117	0.110	0.121	0.183	0.159	0.159	0.140
0.261	—	0.137	0.142	0.164	0.332	0.141	0.139	0.145	0.189	0.177	0.167	0.131
0.225	0.274	—	0.154	0.186	0.304	0.130	0.123	0.120	0.195	0.170	0.184	0.137
0.304	0.268	0.289	—	0.135	0.305	0.158	0.147	0.170	0.197	0.187	0.130	0.081
0.312	0.300	0.337	0.269	—	0.280	0.182	0.175	0.194	0.170	0.157	0.007	0.136
0.377	0.387	0.402	0.366	0.326	—	0.270	0.287	0.294	0.265	0.268	0.287	0.337
0.241	0.265	0.267	0.270	0.323	0.339	—	0.048	0.061	0.105	0.095	0.187	0.153
0.261	0.287	0.288	0.286	0.329	0.395	0.206	—	0.049	0.116	0.100	0.170	0.146
0.272	0.278	0.266	0.291	0.336	0.395	0.210	0.200	—	0.105	0.092	0.194	0.156
0.319	0.293	0.304	0.292	0.293	0.345	0.241	0.252	0.251	—	0.034	0.190	0.197
0.307	0.283	0.286	0.287	0.282	0.343	0.230	0.239	0.245	0.034	—	0.172	0.185
0.309	0.304	0.334	0.264	0.012	0.325	0.323	0.328	0.332	0.299	0.286	—	0.139
0.305	0.268	0.292	0.247	0.290	0.384	0.275	0.286	0.275	0.294	0.289	0.290	—

The point of attachment for the spindle muscle, a thin muscle that runs through the central axis of the U-shaped gut, is on the posterior body wall in the Phascolosomatidea. Most members of the Sipunculidea lack this arrangement, except in *Siphonosoma*, the monotypic genus *Siphonomecus*, and those possessing a third arrangement (members of species in *Phascolion* and *Onchnesoma*) where the spindle muscle is absent except for a region of posterior attachment (C&G). The evolution of this arrangement cannot be addressed further by this analysis, given the extreme divergence of *Phascolion* from other sipunculans and the inconsistent placement of *Siphonosoma*.

Of the remaining 13 taxa, *Siphonosoma* is the least consistent in phylogenetic placement among these analyses (Figs. 2, 3, 4). *Siphonosoma cumanense* is basal in the sipunculan tree as suggested in Fig. 2 and by Cutler (1994). It possesses a mixture of traits apomorphic to both classes: tentacles that surround the mouth partially or wholly (Sipunculidea), lack of complex hooks in rings around the introvert (Sipunculidea), and posterior attachment of the spindle muscle (Phascolosomatidea). Only by examining other species in this genus, sequencing other genes, and possibly

evaluating ultrastructural traits can we gain more insight regarding this enigmatic sipunculan.

Other characters analyzed by C&G appear to have limited ability to resolve smaller groups of sipunculans (e.g., family-level; C&G; Cutler 1994). The off-center junction of the trunk and introvert is characteristic of the two species within the Aspidosiphonidae I sequenced, and indeed they are similar to one another genetically. In part, this morphological character is closely tied to the occurrence of an anal plate, or shield, in the two genera (noted in C&G), which is used to block the burrow opening in individuals of *Aspidosiphon* spp. and supports a calcified horn in individuals of *Lithacrosiphon* spp. The deflection of the introvert off-center may be a mechanical constraint of this morphologic feature. Indeed, in a specimen of the other aspidosiphonid genus that I could not amplify (*Cloeosiphon*), anal plates exist but they surround the introvert/trunk junction to arm this junction. When the introvert is fully extended, they do not deflect the introvert. A better representation of this character for the Aspidosiphonidae may be the presence of anal plates.

Independent origins of several morphological features are now supported (C&G, Cutler 1994, this

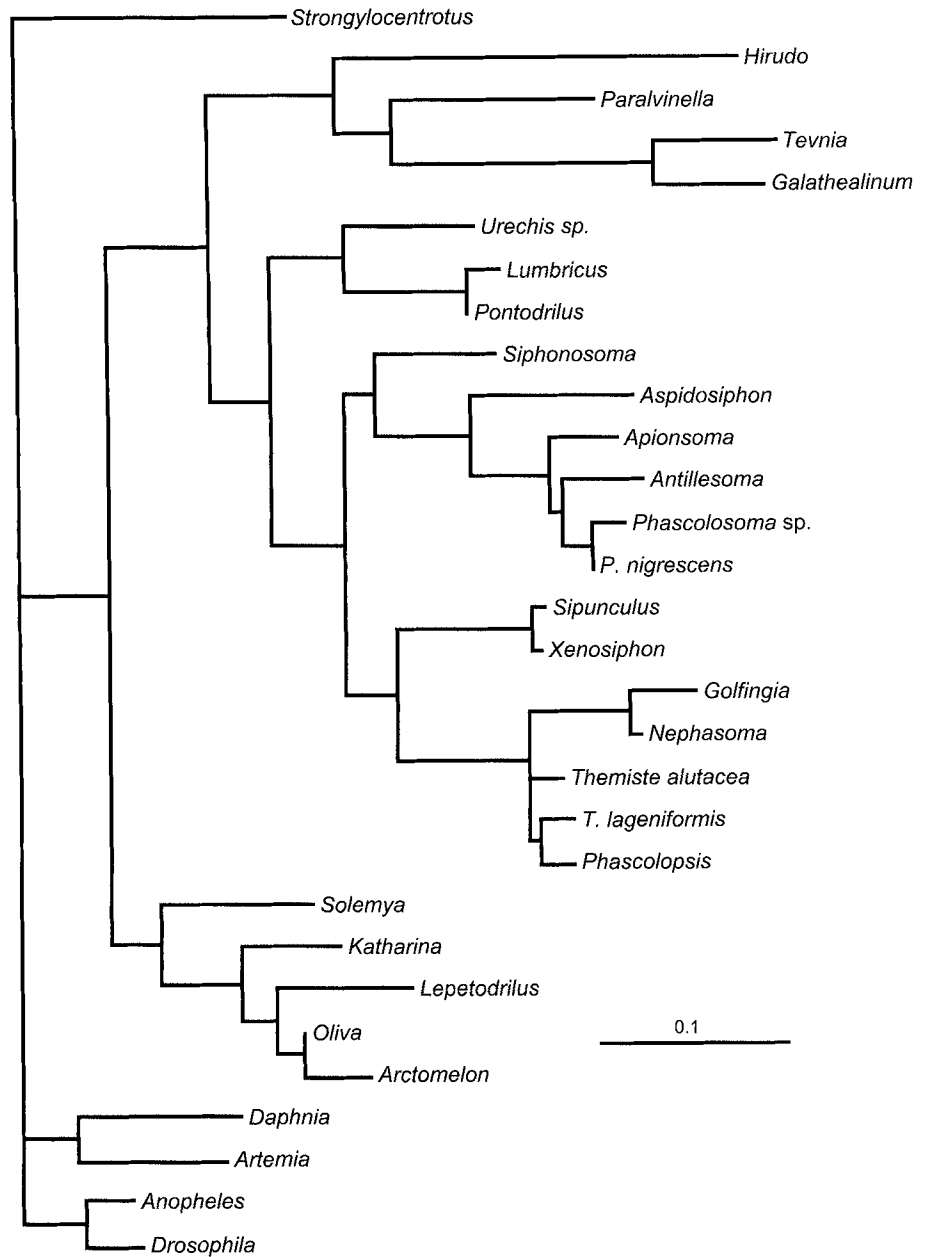


Fig. 3. A protein maximum likelihood tree ($-\ln$ likelihood score = 2886.62) based on inferred amino acid sequences using PAML (Yang 1997). Tree excludes long-branch taxon, *Lithacrosiphon cristatus* (identical in amino acid sequence to *Aspidosiphon laevis*) and randomly excludes members of outgroup clades to accommodate program limits (maximum number of taxa = 30). The kappa parameter was estimated from the data set and no molecular clock was assumed. The tree was subsequently rooted with the *Strongylocentrotus* sequence.

study). The body wall musculature of *Phascolopsis* is not only misleading in suggesting affinities to the Sipunculidea, but similar arrangements appear to have evolved independently in *Antillesoma*, *Phascolosoma*, *Lithacrosiphon*, and some *Aspidosiphon* species. Evolutionary trends in the arrangement of introvert retractor muscles remain uncertain, although the plesiomorphic condition reported in C&G is supported (two pairs, a ventral and dorsal set) by its occurrence in Phascolosomatidea and Sipunculidea. However, one cannot suggest a transformation series for the derived conditions save that dorsal retractors appear to be lost in some animals.

In this study, I have assumed that these species typify their respective genera. Yet, the sequences of *Siphonosoma cumanense* and *Phascolion cryptum* leave their positions at odds with either consistent placement (in the case of *S. cumanense*) or a monophyletic Sipunculidea (in the case of *P. cryptum*). It is possible that these species are not typical of their genera. *Siphonosoma* is a monospecific genus, so other representatives are not available to compare. *P. cryptum* is noted to be morphologically and ecologically quite different from other members of the genus; it possesses unique primary tentacles in the crown and secondary tentacles where hooks are present in other species (Cutler 1994,

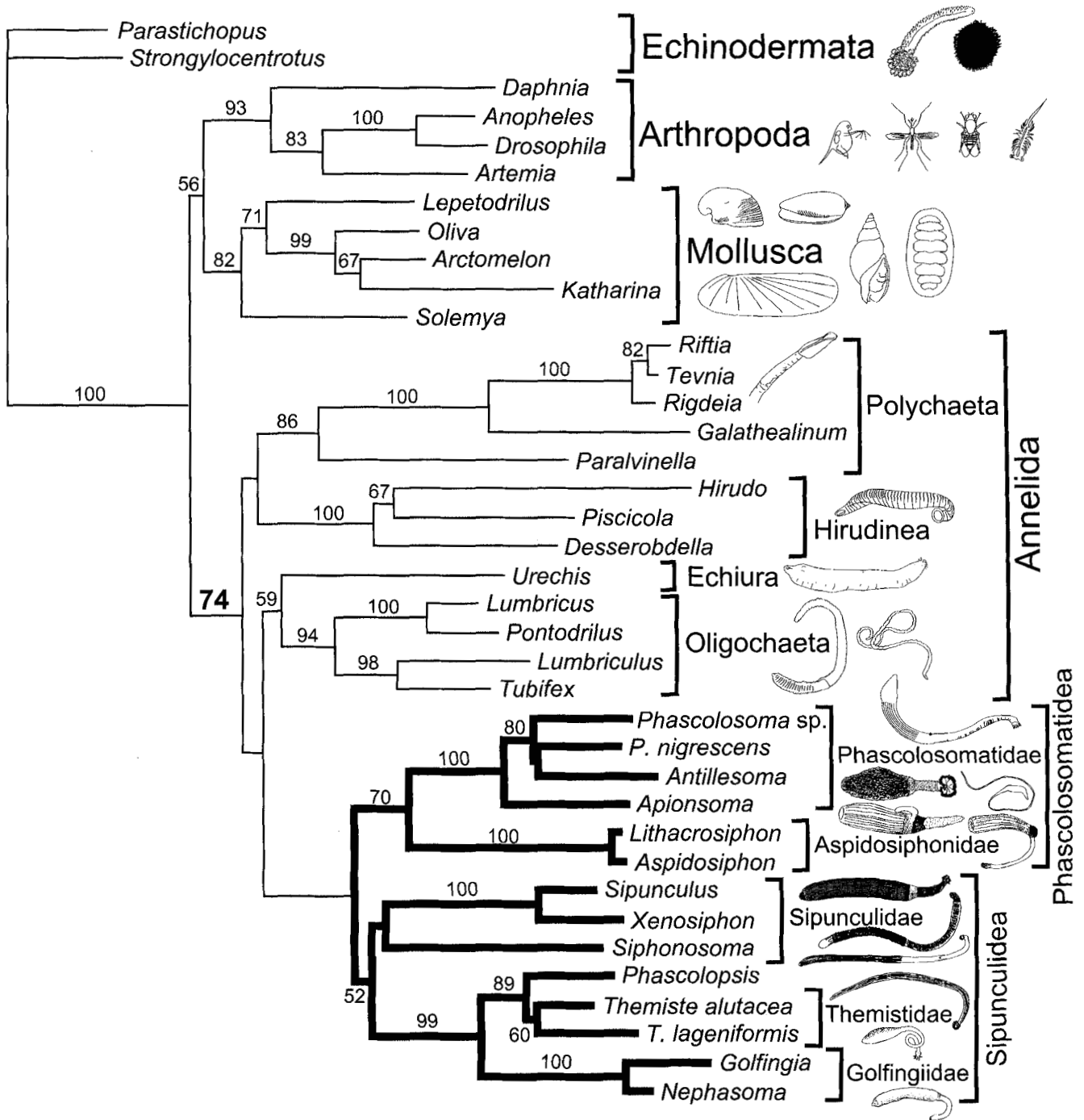


Fig. 4. Neighbor-joining tree calculated for 38 taxa (deleting presumed long branch taxa) using LD/P DNA distances on first and second codon positions. Numbers at branches are support values reported for branches represented in greater than 50% of trees for 500 bootstrap replicates on DNA distances.

Hendrix 1975). Different species for these genera should be sampled and other genes analyzed to address these limitations of the present study.

Excluding those problematic taxa, the results of MP and ML analyses on amino acid data (Figs. 2, 3) and NJ analysis (Fig. 4) on LD/P distances support a close

relationship between sipunculans and other coelomate (non-hemichordate) worms, including oligochaetes, leeches, polychaetes, pogonophorans, and an echiuran. The latter two groups have been recently placed within the phylum Annelida by molecular methods (Black et al. 1997; McHugh 1997). The results of the

Kishino-Hasegawa test and likelihood mapping using CO1 data give statistical support for a sipunculan/annelid relationship, rather than the alternative sipunculan/molluscan relationship. This is consistent with analyses on elongation factor 1 α (McHugh 1997) and mitochondrial gene order (Boore & Staton 2002). The broad pattern of association here is consistent also with earlier published studies (Winnepenninckx et al. 1995; Giribet et al. 2000 [their 18S rDNA tree, only]; Regier & Shultz 2000), which group sipunculans more closely with annelids than molluscs.

To be sure, this data is in conflict with myriad other groupings resulting from the analyses on the small ribosomal (18S) subunit. Sipunculans can be basal to annelids and molluscs (e.g., Giribet et al. 2000 [combined trees], sister to echiurans and pogonophorans to the exclusion of a polychaete (Aleshin et al. 1998), sister to a nemertine worm (Turbeville et al. 1992), sister to entoprocts (Mackey et al. 1996; Winnepenninckx et al. 1996), or grouped with an oligochaete and a platyhelminth but excluding an echiuran and polychaete (Mallatt & Winchell 2002). Such a list is not exhaustive but a sampling of the diverse placement of sipunculans under previous analyses using this single gene.

The mitochondrial cytochrome c oxidase I gene may only have limited utility in resolving higher-level systematics. Indeed in this analysis, CO1 yields different paraphyletic assemblage of annelids at the amino acid and DNA levels, although the support for these nodes among annelid groups is low (Figs. 2, 4). The paraphyly that has been reported for both morphological (Rouse & Fauchald 1997) and elongation factor 1 α data (McHugh 1997) within the annelids suggests possible rapid radiation within worm groups that may prove difficult to resolve. Also note that, for these data, intraphylum relationships of molluscs yield a paraphyletic gastropod and chiton assemblage for MP of amino acids and first and second codon positions under NJ (Fig. 4) and MP analyses (trees not shown), as well as paraphyly for crustaceans and insects under NJ (Fig. 4) and MP analyses (trees not shown). However, the node grouping sipunculans (minus *Phascolion*) and annelids is supported by bootstrap analyses of codon positions 1 and 2 of DNA data by NJ for LD/P distances (Fig. 4), also for NJ bootstrap for uncorrected distances (77% support for 500 replicates—first and second codon positions for 38 taxa, tree not shown) and MP methods (61% support for 500 replicates of 100 stepwise-additions per replicate—first and second codon positions for 38 taxa, tree not shown).

These data strongly support the previous morphological analysis put forth in C&G and support similar conclusions as some previous molecular systematic in-

vestigations at higher levels. Many investigators of molecular systematics assume that only molecules can provide rigorous results; this then demonstrates the power of a careful morphological analysis. Within the phylum, this independent test demonstrates the overall power of a careful morphological investigation within a phylum (C&G). In the case of higher-level systematics, these data demonstrate that some molecular datasets can potentially provide inference at several levels, simultaneously. Although the support of these data is not strong in and of itself, these findings serve to reinforce other findings (e.g., Boore & Staton 2002), adding to the preponderance of evidence for a systematic relationship between sipunculans and annelids excluding molluscs.

Acknowledgments. This material is based upon work supported by a NSF Postdoctoral Fellowship grant DEB-9303301 awarded to JL Staton to train in the lab of WM Brown. Partial sequencing support was provided by NASA Exobiology Program grant NAGW 4223 to DK Jacobs. I thank P Gibbs for supplying me with ethanol-preserved specimens of *Golfingia elongata* and *Nephasoma rimicola*. I would like to thank S Reed for help in obtaining most specimens of sipunculans, and M Rice for hosting me as a postdoctoral fellow and later a visiting researcher at the Smithsonian Marine Station. This is contribution number 545 from the Smithsonian Marine Station at Ft. Pierce, FL. I thank JM Turbeville, TM Collins, and F Kraus for early training in DNA methodology. I thank K Halanych for help in implementing some of the analyses. This research was supported, in part, by a MarCraig Memorial fellowship at the Museum of Comparative Zoology; comments of D Campbell, EB Cutler, G Giribet, D McHugh, P Waddell, and one anonymous reviewer greatly improved the manuscript.

References

- Åkesson B 1958. A study of the nervous system of the sipunculoideae, with some remarks on the development of the two species *Phascolion strombi* Montagu and *Golfingia minuta* Keferstein. Undersökningar över Oresund (Lund.) 39: 1–249.
- Aleshin VV, Milyutina IA, Kedrova OS, Vladychenskaya NS, & Petrov NB 1998. Phylogeny of Nematoda and Cephalorhyncha derived from 18S rDNA. J. Mol. Evol. 47: 597–605.
- Baird WB 1868. Monograph on the species of worms belonging to the subclass Gephyreae. Proc. Zool. Soc. Lond. 1868: 77–114.
- Baltzer F 1931. Sipunculida. In: Handbuch der Zoologie, Vol 2. Kükenthal W & Krumbach T, eds., pp. 15–61. Walter De Gruyter, Berlin.
- Black MB, Halanych KM, Maas PAY, Hoeh WR, Hashimoto J, Desbruyères D, Lutz RA, & Vrijenhoek RC 1997. Molecular systematics of vestimentiferan tubeworms from hydrothermal vents and cold-water seeps. Mar. Biol. 130: 141–149.

- Boore JL & Staton JL 2002. The mitochondrial genome of the sipunculid *Phascolopsis gouldii* supports its association with Annelida rather than Mollusca. *Mol. Biol. Evol.* 19: 127–137.
- Bruton DL 2001. A death assemblage of priapulid worms from the middle Cambrian Burgess Shale. *Lethaia* 34: 163–167.
- Collins TM, Wimberger PH, & Naylor GJP 1994. Compositional bias, character-state bias, and character-state reconstruction using parsimony. *Syst. Biol.* 43: 482–496.
- Collins TM, Frazer K, Palmer AR, Vermeij GJ, & Brown WM 1996. Evolutionary history of northern hemisphere *Nucella* (Gastropoda Muricidae): molecular, morphological, ecological, and paleontological evidence. *Evolution* 50(6): 2287–2304.
- Cutler EB 1994. The Sipuncula: their systematics, biology, and evolution. Cornell University Press, Ithaca, NY. 453 pp.
- Cutler EB & Gibbs PE 1985. A phylogenetic analysis of higher taxa in the phylum Sipuncula. *Syst. Zool.* 34: 162–173.
- Delle Chiaje S 1823. Memorie sulla storia e notomina degli animali senza vertebre del Regno di Napoli, Vol. I. Fratelli Fernandes, Naples. 184 pp.
- Felsenstein J 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27: 401–410.
- Florkin M 1976. Biochemical evidence for the phylogenetic relationships of the Sipuncula. In: Proceedings of the International Symposium on the Biology of the Sipuncula and Echiura, Vol. II. Rice ME & Todorovic M, eds., pp. 95–108. Naucno Delo Press, Belgrad.
- Folmer O, Black M, Hoeh W, Lutz R, & Vrijenhoek R 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotech.* 3(5): 294–299.
- Gerould JH 1903. The development of *Phascolosoma*. Studies on the embryology of the Sipunculidae I. *Zool. Jahrb., Anat.* 23: 77–162.
- Gibbs PE 1987. A new species of *Phascolosoma* (Sipuncula) associated with a decaying whale's skull trawled at 880 m depth in the south-west Pacific. *New Zeal. J. Zool.* 14: 135–137.
- Gibbs PE & Cutler EB 1987. A classification of the phylum Sipuncula. *Bull. Brit. Mus. Nat. Hist. (Zool.)* 52(1): 43–58.
- Giribet G, Distel DL, Polz M, Sterrer W, & Wheeler WC 2000. Triploblastic relationships with emphasis on the acoelomates and the position of Gnathostomulida, Cycliophora, Plathelminthes, and Chaetognatha: a combined approach of 18S rDNA sequences and morphology. *Syst. Biol.* 49: 539–562.
- Goldman N, Anderson JP, & Rodrigo AG 2000. Likelihood-based tests of topologies in phylogenetics. *Syst. Biol.* 49(4): 652–670.
- Graybeal A 1998. Is it better to add taxa or characters to a difficult phylogenetic problem? *Syst. Biol.* 47(1): 9–17.
- Heath H 1899. The development of *Ischnochiton*. *Zool. Jahrb., Anat.* 12: 567–656.
- Hendrix GY 1975. A review of the genus *Phascolion* (Sipuncula) with the descriptions of two new species from the western Atlantic. In: Proceedings of the International Symposium on the Biology of the Sipuncula and Echiura, Vol. I. Rice ME & Todorovic M, eds., pp. 117–137. Naucno Delo Press, Belgrad.
- Hendy MD & Penny D 1989. A framework for the quantitative study of evolutionary trees. *Syst. Zool.* 38: 297–309.
- Henry RP 1987. Invertebrate red blood cell carbonic anhydrase. *J. Exp. Zool.* 242: 113–116.
- Hoffmann RJ, Boore JL, & Brown WM 1992. A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis*. *Genetics* 131: 397–412.
- Jones DT, Taylor WR, & Thornton JM 1992. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8: 275–282.
- Kishino H & Hasegawa M 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* 29: 170–179.
- Lake JA 1994. Reconstructing evolutionary trees from DNA and protein sequences: Paralinear distances. *Proc. Natl. Acad. Sci. USA* 91: 1455–1459.
- Lamarck JA 1816. Histoire naturelle des animaux sans vertèbres, Vol. 3. Radiaires, vers, crustacés, insectes. Verdrière, Paris. 586 pp.
- Mackey LY, Winnepenninckx B, De Wachter R, Backeljau T, Emschermann P, & Garey JR 1996. 18S rRNA suggests that Entoprocta are protostomes, unrelated to Ectoprocta. *J. Mol. Evol.* 42: 552–559.
- Mallatt J & Winchell CJ 2002. Testing the new animal phylogeny: first use of combined large-subunit and small-subunit rRNA gene sequences to classify the protostomes. *Mol. Biol. Evol.* 19: 289–301.
- McHugh D 1997. Molecular evidence that echiurans and pogonophorans are derived annelids. *Proc. Natl. Acad. Sci. USA* 94: 8006–8009.
- Pickford GE 1947. Echiurida and Sipunculida. *Encyclopaedia Britannica*. 20: 717–718.
- Posada D & Crandall KA 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14(9): 817–818.
- Quatrefages MA de 1847. Études sur les types inférieurs de l'embranchement des Annelés. Mémoire sur l'échiure de Gaertner (*Echiurus Gaertnerii* NOB.). *Annales des Sciences Naturelles, ser. 3, Zoologie (Paris)* 7: 307–343.
- Regier JC & Shultz JW. 1998. Molecular phylogeny of arthropods and the significance of the Cambrian 'explosion' for molecular systematics. *Am. Zool.* 38: 918–928.
- Rice ME 1975. Sipuncula. In: *Reproduction of Marine Invertebrates*, Vol. 2. Entoprocta and lesser coelomates. Giese A & Pearse J, eds., pp. 67–127. Academic Press, New York.
- 1985. Sipuncula: developmental evidence for phylogenetic inference. In: *The Origins and Relationships of Lower Invertebrates*. Conway Morris S, George JD, Gib-

- son R, & Platt HM, eds., pp. 274–296. Oxford University Press, Oxford.
- Rouse GW & Fauchald K 1997. Cladistics and polychaetes. *Zoologica Scripta* 26(2): 139–204.
- Saitou N & Nei M 1987. The neighbor-joining method: a new method for constructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406–425.
- Scheltema AH 1993. Aplacophora as progenetic aculiferans and the coelomate origin of mollusks as the sister taxon of Sipuncula. *Biol. Bull.* 184: 57–78.
- 1996. Phylogenetic position of the Sipuncula, Mollusca and the progenetic Aplacophora. In: *Origin and Evolutionary Radiation of the Mollusca*. Taylor JD, ed., pp. 53–58. Oxford University Press, Oxford.
- Sedgwick A 1898. Sipunculoidea (Gephrea, Achaeta). In: *A Student's Textbook of Zoology*. Sedgwick A, ed., pp. 534–539. Swan Sonnenschein, London.
- Steel M 1994. Recovering a tree from the Markov leaf colourations it generates under a Markov model. *Appl. Math. Lett.* 7: 19–23.
- Stephen AC & Edmonds SJ 1972. The phyla Sipuncula and Echiura. *Trustees Brit. Mus. (Nat. Hist.) London*. 528 pp.
- Strimmer K & von Haeseler A 1996. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* 13: 964–969.
- 1997. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proc. Natl. Acad. Sci. USA*. 94: 6815–6819.
- Strimmer K, Goldman N, & von Haeseler A 1997. Bayesian probabilities and quartet puzzling. *Mol. Biol. Evol.* 14: 210–211.
- Swofford DL 2000. PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Turbeville JM, Field KG, & Raff RA 1992. Phylogenetic position of phylum Nemertini, inferred from 18S rRNA sequences: molecular data as a test of morphological character homology. *Mol Biol Evol* 9: 235–249.
- Winnepenninckx B, Backeljau T, & De Wachter R 1995. Phylogeny of protostome worms derived from 18S rRNA sequences. *Mol. Biol. Evol.* 12: 641–649.
- 1996. Investigation of molluscan phylogeny on the basis of 18S rRNA sequences. *Mol. Biol. Evol.* 13: 1306–1317.
- Yang ZH 1994. Maximum-likelihood phylogenetic estimation from DNA sequences with variable rates over sites—approximate methods. *J. Mol. Evol.* 39(3): 306–314.
- 1997. *Phylogenetic analysis by maximum likelihood (PAML)*, Version 1.3. Department of Integrative Biology, University of California at Berkeley.
- Zouros E, Ball AO, Saavedra C, & Freeman KR 1994. An unusual type of mitochondrial DNA inheritance in the blue mussel *Mytilus*. *Proc. Natl. Acad. Sci. USA* 91: 7463–7467.