A test of monophyly of the gutless Phallodrilinae (Oligochaeta, Tubificidae) and the use of a 573-bp region of the mitochondrial cytochrome oxidase I gene in analysis of annelid phylogeny

Johan A. A. Nylander, Christer Erséus & Mari Källersjö

Accepted 7 September 1998

Nylander, J. A. A., Erséus, C. & Källersjö, M. (1999) A test of monophyly of the gutless Phallodrilinae (Oligochaeta, Tubificidae) and the use of a 573 bp region of the mitochondrial cytochrome oxidase I sequences in analysis of annelid phylogeny. — *Zoologica Scripta 28*, 305–313.

A 573-bp region of the mitochondrial gene cytochrome c oxidase subunit I (COI) of two species of *Inanidrilus* Erséus and four species of *Olavius* Erséus (Phallodrilinae, Tubificidae) is used in a parsimony analysis together with a selection of 35 other annelids (including members of Polychaeta, Pogonophora, Aphanoneura, and the clitellate taxa Tubificidae, Enchytraeidae, Naididae, Lumbriculidae, Haplotaxidae, Lumbricidae, Criodrilidae, Branchiobdellida and Hirudinea), and with two molluscs as outgroups. The data support the monophyly of the *Olavius* and *Inanidrilus* group, with a monophyletic *Inanidrilus*. However, parsimony jackknife analyses show that most of the other groups are unsupported by the data set, thus revealing a large amount of homoplasy in the selected gene region. Practically no information is given of within/between family relationships except for a few, closely related species. This suggests that the analysed COI region is not useful, when used alone, for inferring higher level relationships among the annelids.

Johan A. A. Nylander, Department of Systematic Zoology, Evolutionary Biology Center, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden.

Christer Erséus, Department of Invertebrate Zoology, Swedish Museum of Natural History, Box 50007, SE-104 05 Stockholm, Sweden. E-mail: christer.erseus@nrm.se

Mari Källersjö, Laboratory of Molecular Systematics, Swedish Museum of Natural History, Box 50007, SE-104 05 Stockholm, Sweden

Introduction

Most of the marine Oligochaeta belong to the family Tubificidae, which is currently subdivided into five subfamilies: Tubificinae, Rhyacodrilinae, Telmatodrilinae, Limnodriloidinae and Phallodrilinae (Erséus 1990). Erséus pointed out, however, that the Tubificidae is likely to be paraphyletic and should include, at least, the Naididae.

Within the Phallodrilinae, a group of species completely lacking an intestinal system has been identified (Giere 1979, 1981; Erséus 1979; Giere *et al*. 1995). These gutless tubificids are nutritionally dependent on symbiotic bacteria living in the body wall (Giere 1981, 1985; Giere & Langheld 1987; Giere & Milligan 1989; Giere *et al*. 1995; Richards *et al*. 1982). It was first believed that some gutless phallodrilines had evolved in parallel (Erséus 1979, 1981, 1983), but Erséus (1984) proposed a monophyletic origin for this group, identifying two genera within it, *Inanidrilus* Erséus, 1979 and *Olavius* Erséus, 1984. Today, about 70 species have been described from, mostly, shallow marine waters of tropical and subtropical areas (Giere *et al*. 1995; Erséus & Giere 1995; Erséus 1997).

No complete phylogeny of the gutless species and their closest relatives exists and their position within the Phallodrilinae is unclear (Erséus 1992). Previous phylogenetic analyses of oligochaetes were all based on various morphological characters (e.g. Erséus 1984, 1990, 1992; Ferraguti & Erséus, *in press*; Jamieson *et al*. 1987; Brinkhurst 1988, 1989, 1991, 1994; Jamieson 1988; Erséus & Ferraguti 1995). The primary aim of the present study is to test the monophyly of the gutless group within the Phallodrilinae using mitochondrial DNA (mtDNA) sequence data.

The gene used, cytochrome *c* oxidase subunit I (COI), is

a protein-coding gene of the mtDNA. At the amino acid-level, it is generally regarded as one of the most conservative genes of the animal mitochondria (Kondo *et al.* 1993; Simon *et al.* 1994; Boore & Brown 1995; Cummings *et al.* 1995). At the same time, the COI gene contains more rapidly evolving nucleotide sites, particularly in third codon positions, which make it potentially useful for inferring more recent branching events. It has been used to assess relationships at practically all systematic levels, from phyla (Folmer *et al.* 1994; Hoeh *et al.* 1996) to within species (Sperling & Hickey 1994; Wüster *et al.* 1995). This study therefore has a secondary aim: to see whether COI-sequence data are useful also at different levels of annelid relationships. Species investigated are representatives from a wide range of taxa. They include species believed to be closely related within a genus, but we also do comparisons between taxa of much higher ranks.

Material and methods

The COI-sequence data analysed in this paper come from a variety of sources. Data for the six species of gutless Phallodrilinae and five other oligochaete species (see Table 1 and below) were sequenced by the first author following the procedures outlined below. Other sequence data were kindly provided by Professor Bent Christensen, Copenhagen (Christensen & Theisen, 1998; B. Christensen, unpublished; data available upon request) or were extracted from Genbank (See Table 1). Two molluscs *Katharina tunicata* (Polyplacophora) and *Oliva sayana* (Gastropoda) were used as outgroup taxa in the parsimony analyses.

Specimens

Single specimens of the gutless phallodrilinae species *Olavius albidus*, *O. prodigus*, *O. tantulus*, *O. imperfectus*, *Inanidrilus reginae*, *I. leukodermatus*, two other phallodrilines (*Aktedrilus arcticus*, *Pirodrilus minutus*), two rhyacodrilines (*Heronidrilus heronae*, *Heterodrilus paucifascis*) and one enchytraeid (*Fridericia tuberosa*), were used for sequencing. Specimens were collected by Dr Emilia Rota, Siena, Italy (*F. tuberosa*) and the second author (all other taxa) and preserved in 80–96% ethanol.

Extraction and PCR amplification

Genomic DNA was extracted following a general extraction method as described by Palumbi *et al.* (1991). In brief, single alcohol-preserved specimens were ground in homogenizing buffer and incubated overnight with proteinase K. DNA was then isolated by phenol-chloroform extraction followed by ethanol precipitation. A 710-bp fragment of the COI gene was amplified using the flanking primers LCO 1490 (5′-GGTCAACAAATCATAAAGATATTGG-3′) and HCO 2198 (5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) (Folmer *et al.* 1994), applying the following PCR profile: (94°C/30 s – 49°C/30 s – 72°C/2 min) X 32 cycles. A typical PCR-protocol was as follows: Template DNA, 5 μL; AmpliTaq Polymerase, 1.25 units; Primer concentration, 20 pM; dNTP concentration, 200 μM; $MgCl_2$ concentration, 2 mM; and 5 μL 10X PCR buffer. Sterile water was added to give a total volume of 50 μL. PCR products were purified using the QIAquick spin PCR Purification Kit (QIAGEN, Santa Clarita, CA USA) following the protocol from the manufacturer.

Sequencing

Purified DNA was sequenced by the dideoxy method (Sanger *et al.* 1977) using both of the amplification primers and three additional, internal primers: C1-J-1751 (5′-GGATCACCTGATATAGCATTCCC-3′) of Simon *et al.* (1994); COI-355, alias LumberJack (5′-GGAACAGGATGAACAGTTTACCC-3′), a modification of the C1-J-1859 primer of Simon *et al.* (1994); and COI-355R, alias Reversed LumberJack (5′-GGGTAAACTGTTCATCCTGTTCC-3′). The COI-355R primer is the reverse-complement of the COI-355 primer and both were developed at the Laboratory of Molecular Systematics, Swedish Museum of Natural History. The number (355) refers to the position of the *D. yakuba* 5′ nucleotide (Clary & Wolstenholme 1985). For cycle sequencing, the Thermo Sequenase cycle sequencing with fluorescent 1 dye CY5-primer labelling protocol from Amersham Pharmacia Biotech (Uppsala, Sweden) was used, except that the samples were not ethanol precipitated after thermocycling. For each sequencing reaction 2 pmol primer (1.5 pmol/μL) and 5 μL PCR product were used with a cycle profile of 95°C for 2 min, then (95°C/30 s – 48°C/30 s – 72°C/60 s) X 30 cycles. Electrophoresis was carried out on a 6% acrylamide gel using an ALF-Express DNA sequencer from Amersham Pharmacia Biotech.

Alignments

Sequences were aligned with the sequence of *Lumbricus terrestris* (Boore & Brown 1995) using AssemblyLIGN v 1.0.7 (Oxford Molecular Group Inc., Campbell, CA, USA). Final alignment adjustments were done manually, comparing both nucleotides and amino acids. Nucleotides were translated into amino acid sequences using the *L. terrestris* genetic code (Boore & Brown 1995) for the annelids, and the Invertebrate Mitochondrial Code (Hoffmann *et al.* 1992) for the molluscs.

Sequence analyses

PAUP v 3.1.1 (Swofford 1991) was used for finding the most parsimonious trees. Heuristic searches with 100

Table 1 Table of included species, with references to COI gene sequences. Sequence sources: 1 = Boore & Brown (1994); 2 = Harasewych, M. G., Adamkewicz, S. L., Blake, J. A., Saudek, D. M., Spriggs, T. & Bult, C. J. direct GenBank submission; 3 = present study; 4 = Christensen & Theisen (1998); 5 = Christensen unpublished; 6 = Black, M. B., Hoeh, W. R., Hashimoto, J., Debruyeres, D., Lutz, R. A. & Vrijenhoek, R. C. direct GenBank submission; 7 = Boore & Brown (1995); 8 = Siddall, M. E. & Burreson, E. M. direct GenBank submission.

Species; (Higher taxa); GenBank Accession number	Species; (Higher taxa); GenBank Accession number
Mollusca	Naididae
Katharina tunicata, (Wood, 1815)[2]; Polyplacophora; U09810	*Chaetogaster diastrophus* (Gruithuisen, 1828)[4]; AF054196
Oliva sayana Ravenel, 1834[1]; Gastropoda; U86333	*Stylaria lacustris* (Linnaeus, 1767)[4]; AF054194
Clitellata	*Nais barbata* Müller, 1773[4]; AF054193
Tubificidae, Phallodrilinae	*Dero digitata* (Müller, 1774)[4]; AF054195
Aktedrilus arcticus (Erséus, 1978)[3]; AF064042	Lumbriculidae
Inanidrilus leukodermatus (Giere, 1979)[3]; AF064040	*Rhynchelmis limosella* Hoffmeister, 1843[5]
Inanidrilus reginae Erséus, 1990[3]; AF064041	*Lumbriculus variegatus* (Müller, 1774)[8]; AF003257
Olavius albidus (Jamieson, 1977)[3]; AF064037	Lumbricidae
Olavius imperfectus Erséus, 1984[3]; AF064036	*Lumbricus terrestris* Linnaeus, 1758[7]; U24570
Olavius prodigus Erséus, 1993[3]; AF064038	Haplotaxidae
Olavius tantulus Erséus, 1984[3]; AF064039	*Haplotaxis gordioides* (Hartmann, 1821)[5]
Paraktedrilus bakeri (Koßmagk-Stephan & Erséus, 1985)[4]; AF054191	Propappidae
Pirodrilus minutus (Hrabe, 1973)[3]; AF064043	*Propappus volki* Michaelsen, 1916[5]
Tubificidae, Rhyacodrilinae	Branchiobdellida
Heronidrilus heronae (Erséus & Jamieson, 1981)[3]; AF064045	*Branchiobdella parasita* Henle, 1835[5]
Heterodrilus paucifascis Milligan, 1987[3]; AF064044	Euhirudinea
Monopylephorus rubroniveus Levinsen, 1884[5]	*Erpobdella octoculata* (Linnaeus, 1758)[8]; AF003274
Rhyacodrilus falciformis Bretscher, 1901[4]; AF054192	*Helobdella stagnalis* (Linnaeus, 1758)[5]
Tubificidae, Tubificinae	*Hirudo medicinalis* Linnaeus, 1758[8]; AF003272
Tubifex tubifex (Müller, 1774)[6]; U74076	Aphanoneura
Clitellio arenarius (Müller, 1776)[4]; AF054190	*Aeolosoma litorale* Bunke, 1967[4]; Aeolosomatidae; AF054188
Heterochaeta costata Claparède, 1863[4]; AF054189	Polychaeta
Enchytraeidae	*Amphisamyta galapagensis* Zottoli, 1983[6]; Ampharetidae; U74058
Buchholzia sp.[5]	*Hediste diversicolor* (Müller, 1776)[5]; Nereidae;
Enchytraeus albidus Henle, 1837[5]	*Paralvinella palmiformis* Desbruyères & Laubier, 1986[6]; Alvinellidae; U74070
Achaeta affinis Nielsen & Christensen, 1959[5]	Pogonophora
Henlea nasuta (Eisen, 1878)[5]	*Galathealinum brachiosum* Ivanov, 1961[6]; Polybrachiidae; U74066
Fridericia tuberosa Rota, 1995[3]; AF064047	Vestimentifera
Criodrilidae	*Oasisia alvinae* Jones, 1985[6]; Tevniidae; U74069
Criodrilus lacuum Hoffmeister, 1845[5]	

random addition sequences and TBR branch swapping were used, both for the nucleotide data and the amino acid data. All characters were initially treated as equally weighted.

Rapid changes in characters (leading to backward mutations or multiple hits) may introduce error to the analysis. For protein coding gene regions, the negative effect of convergent silent mutations (nucleotide substitutions that do not lead to amino acid replacements) can be minimized by differentially weighting of characters or performing the analyses at the amino acid level (for recent reviews of methods, see Simon *et al*. 1994; Swofford *et al*. 1996). In this paper, analyses are therefore conducted using both the nucleotide and amino acid data sets and subjected to different weighting schemes.

Character weighting according to codon positions were accomplished with MacClade vs. 3.07 (Maddison & Maddison 1992). Weights were based on the approximate inverse of frequencies of occurrence of informative sites (Brower 1994), and calculated from the most parsimonious trees from the equally weighted analysis in PAUP. Weights were scaled so that the lowest value = 1 and rounded to the nearest integer. Analyses were also done with all third codon positions excluded (weight = 0).

Character state transformation weighting was implemented by constructing a step matrix in PAUP assigning higher weight to transversions over transitions. The costs were based on the estimated transition:transversion (Ti:Tv) ratio from the initial most parsimonious tree, as calculated in MacClade. Both the weighting scheme

applied to codon positions and the estimation of the Ti:Tv ratio have the advantage of being based only on data at hand. However, the calculation of the Ti:Tv ratio is not straightforward (as recently discussed by Wakeley 1994, 1996; Yang 1996). Therefore, a range of cost ratios (Ti:Tv = 1:1.1, 1:2, 1:4 and 1:10), were also applied.

Nodal support for the equally weighted nucleotide and amino acid data were estimated by Parsimony Jackknifing (Farris *et al*. 1996) using pre-release versions of the programs Xac (for nucleotides) and Pax (for aminoacids). The programs were provided by James S. Farris, Laboratory of Molecular Systematics, Swedish Museum of Natural History. Data sets were jackknifed 10 000 times using branch swapping and 10 random addition sequences per pseudoreplicate.

The nodal support for the differentially weighted nucleotide data was estimated by bootstrapping (Felsenstein 1985), using PAUP (1000 pseudoreplicates, heuristic search and TBR branch swapping) since the Xac program does not handle differentially weighted data (according to codon positions or Ti:Tv ratios).

Results

Nucleotide data

After sequencing, a 573-bp fragment of the COI gene was comparable and aligned over all included taxa, corresponding to positions 88–660 in the *Lumbricus terrestris* sequence of Boore & Brown (1995). No insertions or deletions were detected. Of these positions, 413 (72.1%) were variable and 318 (55.5%) were informative (Table 2). There were considerable differences in frequency of change, Ti:Tv ratio and base composition between the codon positions (first, second and third, respectively). The third positions deviated considerably and showed the highest variation (only one third position invariant), the highest A + T bias (80.4% A + T content), and also a deviating transition bias (0.59, compared with 1.66 and 1.15 for first and second positions, respectively).

The heuristic search analysis of the equally weighted nucleotide data yielded seven most parsimonious trees, each 3466 steps long (CI (informative sites only) = 0.21, RI = 0.31). The consensus tree (Fig. 1A) is to a large extent unresolved, although the two gutless genera (*Olavius* spp., *Inanidrilus* spp.) appears monophyletic and comprise a clade of their own within the annelids. With regard to other ingroup relationships, the topology shows unexpected features. For instance, the aphanoneuran (*Aeolosoma*) is placed as the sister taxon of all remaining annelids, and monophyly of the Clitellata is not supported: the two pogonophorans (*Galathealinum-Oasisia*) comprise a sister group of the medical leech (*Hirudo*), and the polychaetes (*Paralvinella-Amphisamytha* and *Hediste*) are positioned within the polytomy comprising most of the oligochaetes. Moreover, the branchiobdellid (*Branchiobdella*) together with the leech *Helobdella* constitute the sister group of the gutless clade.

The jackknife analysis, however, shows that most of the structure of the consensus tree is insufficiently supported by the nucleotide sequences. The jackknife tree (Fig. 1B) contains a basal ingroup polytomy, and it reveals virtually no pattern of relationships of the major groups of clitellates and non-clitellate annelids. Although the gutless clade as a whole is well supported (91%), the sistergroup relationship between the gutless genera from Fig. 1A is not retained. The monophyly of *Olavius* is not supported and *Inanidrilus* receives only the lowest support (53%). *Olavius albidus* and *O. prodigus* form a reasonably well supported (75%) subclade within the gutless group.

Two naidids (*Dero-Nais*) and the pogonophorans (*Galathealinum-Oasisia*) comprise two clades with great support (100% and 97%, respectively). Low support is also given to the *Haplotaxis-Lumbricus* group (59%) whereas, more surprisingly, moderate support (67%) is given to a clade consisting of one naidid (*Stylaria*) and the propappid (*Propappus*). Moreover, there is low (56%) but unexpected support for a clade comprising the phallodriline *Pirodrilus* and its sistergroup containing the two tubificines *Clitellio* and *Heterochaeta*.

Parsimony analyses of differentially weighted data gave different results depending on the particular weighting scheme used. However, the resulting trees (individual results not shown) differed greatly from one another, and no clear tendency of stability or coherence in grouping could be seen. Applying the inferred character weights for

Table 2 Table of nucleotide composition and estimated substitution pattern for COI mtDNA sequence data (average over taxa). Number of steps, informative positions and Ti/Tv ratios were inferred from the most parsimonious trees. Abbreviations: Ti/Tv = transitions over transversions; MPTs = most parsimonious trees; A = adenosine; C = cytidine; G = guanosine; T = thymidine.

Codon position	Total number of bases	%A	%C	%G	%T	Ti/Tv	Number of variable positions	Number of informative positions	Steps in MPTs
all	573	29.2	21.8	15.6	33.4	0.79	413	318	3466
1st	191	26.7	24.1	26.8	22.4	1.66	126	92	810
2nd	191	14.2	26.5	15.4	43.9	1.15	97	37	235
3rd	191	46.7	14.8	4.8	33.7	0.59	190	189	2421

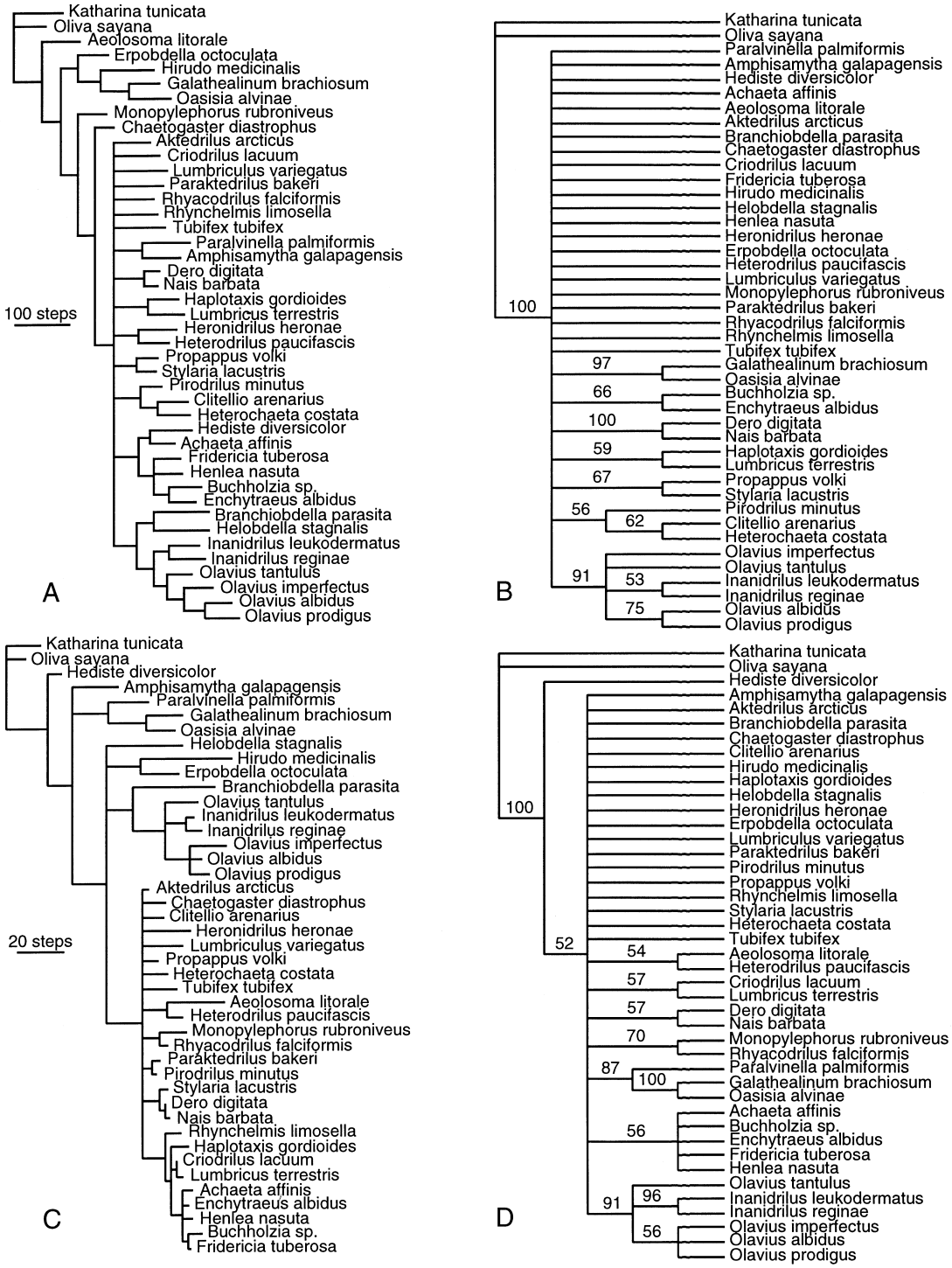

Fig. 1 *A-C*. Results from parsimony analysis of COI data. —*A*. Consensus of seven most parsimonious trees from nucleotide analysis (l = 3466, CI (informative sites only) = 0.21, RI = 0.31). Bar indicates branch length (ACCTRAN optimization in PAUP). —*B*. Jackknife tree of nucleotide data. Values above branches indicate support (10 000 replicates, 10 random addition sequences per replicate using branch swapping; using Jax). —*C*. Consensus of 720 most parsimonious trees for amino acid data (l = 611, CI (informative sites only) = 0.49 and RI = 0.60). Bar indicates branch length (ACCTRAN optimization in PAUP). —*D*. Jackknife tree of amino acid data. Values above branches indicate support (10 000 replicates, 10 random addition sequences per replicate with branch swapping; using Pax).

codon positions of 3:10:1 (1st, 2nd and 3rd position, respectively) for instance, showed a more basal position of the gutless group placed as sistergroup to a clade comprising the leeches *Hirudo* and *Erpobdella* together with the aphanoneuran *Aeolosoma*. While the inferred Ti:Tv weighting of 1.3:1 (which assigns greater weight to transitions than transversions) placed the gutless group in a distal position in the tree with a clade consisting of *Branchiobdella* and the leech *Helobdella* as its sistergroup.

Bootstrap analyses of the weighted data sets gave the same overall picture as the equally weighted analysis: a few well supported smaller groups or pairs of species, such as *Dero-Nais*, *Galathealinum-Oasisia*, *Propappus-Stylaria* and the *Inanidrilus-Olavius* group with *O. albidus* and *O. prodigus* as close relatives. *Inanidrilus* appears monophyletic, although with low to moderate support (50–77%), under all weighting schemes except when applying the inferred Ti:Tv weighting.

Amino acid data

Translation of the nucleotides into amino acids resulted in a matrix with a total of 191 characters, 142 (74.3%) of them variable and 89 (46.6%) parsimony informative. Three amino acid positions were ambiguous or coded as missing (= ?) due to uncertainties in the underlying nucleotide sequence.

The heuristic search analysis of the amino acid data yielded 720 most parsimonious trees, each of length 611 steps, CI (informative sites only) = 0.49 and RI = 0.60.

The consensus tree (Fig. 1C) is to a large extent collapsed, and the few groups that remain differ considerably from the ones found among the shortest trees from the nucleotide data (Fig. 1A). In fact, the only groups in common are the monophyletic gutless clade as a whole, the two *Inanidrilus*, three *Olavius* (*imperfectus-albidus-prodigus*), two naidids (*Dero-Nais*) and the two pogonophorans (*Galathealinum-Oasisia*).

The jackknife analysis of the amino acid data resulted, as for the nucleotide data, in a basal collapse of the annelid group with support for only a few terminal branches (Fig. 1D). The remaining groups differ somewhat from the nucleotide tree and only four groups are shared by the two different jackknife trees (cf. Fig. 1B,D). The polychaete *Hediste* is placed as sistergroup to all remaining annelids, although with minimum support (52%). The pogonophorans are grouped with the polychaete *Paralvinella*, and there are low, but notable, support values for two naidids (*Dero-Nais*), two rhyacodriline tubificids (*Monopylephorus-Rhyacodrilus*) and, a clade comprising all included enchytraeid taxa, *Buchholzia*, *Enchytraeus*, *Achaeta*, *Henlea*, and *Fridericia* (support values 87%, 57%, 70% and 56%, respectively, see Fig. 1D). The criodrilid *Criodrilus* groups together with the lumbricid *Lumbricus*, and more unexpectedly, the rhyacodriline *Heterodrilus* together with the aphanoneuran *Aeolosoma*, although both clades have low support, 57% and 54%, respectively. The gutless group has good support (91%), and also two subgroups within it are supported: there is an assuring jackknife value for the *Inanidrilus* spp. group (96%), but a considerable lower value for the group consisting of *O. imperfectus*, *O. albidus* and *O. prodigus* (56%). Monophyly of *Olavius* as a whole, however, is not supported just as in the unweighted and weighted nucleotide analyses. The amino acid data thus provide support for a slightly different tree than that indicated by the nucleotide data, and they neither support any close relationship between *Propappus* and *Stylaria*, *Haplotaxis* and *Lumbricus*, nor between the phallodriline Pirodrilus and the two tubificines *Clitellio* and *Heterochaeta* (cf. Fig. 1B).

Discussion

The gutless Phallodrilinae

Both the nucleotide and amino acid sequences provide evidence that *Olavius* and *Inanidrilus* have a common origin (Fig. 1A-D); the analysed data also corroborate that *Inanidrilus* is a natural group within the gutless assemblage. Whether *Olavius* is paraphyletic or a proper sister group of *Inanidrilus*, however, is not conclusive on the basis of the gene fragment studied here.

From a morphological point of view, the gutless Phallodrilinae show great similarities in somatic as well as genital features (Erséus 1984), and structural uniformity characterizes their symbiotic bacteria too (Giere *et al*. 1995). It is therefore noteworthy that, in the most parsimonious trees, the branch lengths are considerably greater for the gutless group and its individual subclades than for most of the other oligochaete nodes (Fig. 1A,C). This seems to indicate that gutless oligochaetes were either evolutionary separated from other tubificids a relatively long time ago, and/or that the rate of nucleotide/amino acid substitutions has been unusually high in the gutless clade.

Other annelids

Apart from the ingroup as a whole and the gutless clade with the monophyletic *Inanidrilus*, only two groups in common for the most parsimonious trees from the nucleotide and amino acid data were supported after jackknifing. These were the seemingly closely related naidid species *Dero-Nais* and the pogonophoran *Galathealinum* together with the vestimentiferan *Oasisia*. All other groupings seemed to be too weakly or ambiguously supported by the data. The low resolution in the jackknife trees (Fig. 1B, D), together with the variation showed between codon positions (Table 2) strongly suggest that the investigated fragment of the COI gene has evolved too rapidly to

enable a phylogenetic assessment at the higher levels of clitellate and other annelid relationships. The data are even too ambiguous for a recognition of a coherence among all studied members of the subfamily Phallodrilinae.

Some of the inconsistencies could be caused by saturation by multiple substitutions. The chance of such substitutions is correlated with frequency of change and is likely to be high in the analysed data, especially in third codon positions, as only 0.5% of third positions are invariant compared with 34.6% and 50.8% of first and second positions, respectively. Furthermore, the Ti:Tv ratio of third positions (0.57) is much lower than that of first and second positions (1.67 and 1.16, respectively). A drop in this ratio is expected as data approach saturation (e.g. Holmquist 1983; DeSalle *et al*. 1987). Moreover, the overall Ti:Tv ratio (0.78), which shows a negative transition bias, indicates saturation of the analysed gene region.

Weighting

Ad hoc weighting (*sensu* Allard & Carpenter 1996) is frequently applied by molecular systematists in attempts to surpress homoplasy or 'noise' in their data. It could be argued that a result that is insensible to differential weighting of the data, is to be regarded as a more robust hypothesis of phylogeny than results that vary with the weighting schemes used. As mentioned above, the most parsimonious trees based on our annelid data are indeed sensitive to differential weighting. The results are also sensitive to the inclusion and/or exclusion of taxa (results not shown); depending on what weighting scheme is used, or what set of taxa is analysed (very) different statements of sistergroup-relationships could be made. However, if one chooses to focus on supported groups, as we prefer to do, there is no major difference between the weighted, *supported*, results. Thus, *ad hoc* weighting did not seemed to provide any further help in clarifying relationships of the selected taxa.

An intriguing effect of the Ti:Tv weighting applied here (when inferring substitutions from a most parsimonious tree) is that, although applying greater weight to less frequent substitution types, it downweights transversions in favour of transitions. That transitions occur more often than transversions, especially in the animal mitochondria, is a generally accepted property of DNA-sequence evolution (e.g. Brown *et al*. 1982; Irwin *et al*. 1991; Kocher & Wilson 1991).

We used, admittedly, weighting procedures that are 'crude simplifications of more intricate weighting possibilities' (Simon *et al*. 1994). More sophisticated methods of weighting are proposed, such as trying to use only 'conservative' nucleotide substitutions (e.g. Irwin *et al*. 1991; see also review in Swofford *et al*. 1996). Such weighting schemes are based on a priori 'knowledge' of the nature or pattern of evolution for the gene at hand. However, this pattern may be dependent on the organisms included and at what taxonomical level the analysis is conducted. Experience from one group of organisms is not necessarily easily applied to another. Moreover, if prior knowledge simply does not exist, as in the case of annelid COI evolution, then using any weighting scheme may be just as good a guess as any other.

Conclusions

The investigated fragment of the COI gene supports monophyly of the gutless group within the Phallodrilinae, and also a monophyletic status of *Inanidrilus* (although the analysis so far has only included two representatives of this genus). However, with the exception of a few small groups of very closely related species, other parts of the phylogeny of the Annelida are not resolved by the COI gene, at least not if restricted to the fragment examined here. Differential character weighting or conducting analyses at the amino acid level did not improve the analytical results. Evidently, other genes will have to be explored, either alone or analysed together with the COI data, to test the monophyly of *Olavius* and other higher level taxa within the Clitellata and the Annelida as a whole.

Acknowledgements

We wish to thank Bent Christensen, Department of Population Biology, Copenhagen University, Denmark for using his unpublished sequence data, Gullevi Bergqvist, Klas Nyblom and Michael Norén (all at the Laboratory of Molecular Systematics, Swedish Museum of Natural History) for help with laboratory work, and two anonymous reviewers for their comments on an earlier version of this paper. This research was supported by the Swedish Natural Science Research Council (Project no. B-BU 3825).

References

Allard, M. W. & Carpenter, J. M. (1996). On weighting and congruence. *Cladistics*, *12*, 183–198.

Boore, J. L. & Brown, W. M. (1994). Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. *Genetics*, *138*, 423–443.

Boore, J. L. & Brown, W. M. (1995). Complete sequence of the mitochondrial DNA of the annelid worm *Lumbricus terrestris*. *Genetics*, *141*, 305–319.

Brinkhurst, R. O. (1988). A taxonomic analysis of the Haplotaxidae (Annelida, Oligochaeta). *Canadian Journal of Zoology*, *66*, 2243–2252.

Brinkhurst, R. O. (1989). A phylogenetic analysis of the Lumbriculidae (Annelida, Oligochaeta). *Canadian Journal of Zoology*, *88*, 392–397.

Brinkhurst, R. O. (1991). A phylogenetic analysis of the Tubificinae (Oligochaeta, Tubificidae) *Canadian Journal of Zoology*, *69*, 97–110.

Brinkhurst, R. O. (1994). Evolutionary relationships within the Clitellata: an update. *Megadrilogica*, *5*, 109–116.

Brower, A. V. Z. (1994). Phylogeny of *Heliconius* butterflies inferred from mitochondrial DNA sequences (Lepidoptera: Nymphalidae). *Molecular Phylogenetics and Evolution*, *3*, 159–174.

Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. (1982). Mitochondrial DNA sequences of primates: tempo and mode of evolution. *Journal of Molecular Evolution*, *18*, 225–239.

Christensen, B. & Theisen, B. F. (in press). Phylogenetic status of the family Naididae (Oligochaeta, Annelida) as inferred from DNA analyses). *Journal of Zoological Systematics and Evolutionary Research 36*, 169–172.

Clary, D. O. & Wolstenholme, D. R. (1985). The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization and genetic code. *Journal of Molecular Evolution*, *22*, 252–271.

Cummings, M. P., Otto, S. P. & Wakeley, J. (1995). Sampling properties of DNA sequence data in phylogenetic analysis. *Molecular Biology and Evolution*, *12*, 814–823.

DeSalle, R., Freedman, T., Prager, E. M. & Wilson, A. C. (1987). Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian Drosophila. *Journal of Molecular Evolution*, *26*, 157–164.

Erséus, C. (1979). Taxonomic revision of the marine genus *Phallodrilus* Pierantoni (Oligochaeta, Tubificidae) with descriptions of thirteen new species. *Zoologica Scripta*, *8*, 187–208.

Erséus, C. (1981). Taxonomic studies of Phallodrilinae (Oligochaeta, Tubificidae) from the Great Barrier Reef and the Comoro Islands with descriptions of ten new species and one new genus. *Zoologica Scripta*, *10*, 15–31.

Erséus, C. (1983). Deep-sea *Phallodrilus* and *Bathydrilus* (Oligochaeta, Tubificidae) from the Atlantic Ocean, with descriptions of ten new species. *Cahiers de Biologie Marine*, *24*, 125–146.

Erséus, C. (1984). Taxonomy and phylogeny of the gutless Phallodrilinae (Oligochaeta, Tubificidae), with descriptions of one new genus and twenty-two new species. *Zoologica Scripta*, *13*, 239–272.

Erséus, C. (1990). Cladistic analysis of the subfamilies within the Tubificidae (Oligochaeta). *Zoologica Scripta*, *19*, 57–63.

Erséus, C. (1992). A generic revision of the Phallodrilinae (Oligochaeta, Tubificidae). *Zoologica Scripta*, *21*, 5–48.

Erséus, C. (1997). Marine Tubificidae (Oligochaeta) from the Montebello and Houtman Abrolhos Islands, Western Australia, with descriptions of twenty-three new species. In: F. E. Wells (Ed.). *The Marine Flora and Fauna of the Houtman Abrolhos Islands, Western Australia*. pp. 389–458, Western Australian Museum, Perth.

Erséus, C. & Ferraguti, M. (1995). The use of spermatozoal ultrastructure in phylogenetic studies of Tubificidae (Oligochaeta). In: B. G. M. Jamieson, J. Ausio, &, J.-L. Justine (Eds). *Advances in Spermatozoal Phylogeny and Taxonomy*. pp. 189–201, Mémoirs de Muséum National d'Histoire Naturelle, 166, Paris.

Erséus, C. & Giere, O. (1995). *Olavius nicolae*, a new gutless marine tubificid species (Oligochaeta) from Belize. *Proceedings of the Biological Society of Washington*, *108*, 491–495.

Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D. & Kluge, A. G. (1996). Parsimony jackknifing outperforms neighbor-joining. *Cladistics*, *12*, 99–124.

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

Ferraguti, M. & Erséus, C. (in press). Sperm types and their use for a phylogenetic analysis of aquatic clitellates. *Hydrobiologia*, in press.

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. *Molecular Marine Biology and Biotechnology*, *3*, 294–299.

Giere, O. (1979). Studies on marine Oligochaeta from Bermuda, with emphasis on new *Phallodrilus* species (Tubificidae). *Cahiers de Biologie Marine*, *20*, 301–314.

Giere, O. (1981). The gutless marine oligochaete *Phallodrilus leukodermatus*. Structural studies on an aberrant tubificid associated with bacteria. *Marine Ecology Progress Series*, *5*, 353–357.

Giere, O. (1985). The gutless marine tubificid *Phallodrilus planus*, a flattened oligochaete with symbiontic bacteria. Results from morphological and ecological studies. *Zoologica Scripta*, *14*, 279–286.

Giere, O. & Langheld, C. (1987). Structural organization, transfer and biological fate of endosymbiontic bacteria in gutless oligochaetes. *Marine Biology*, *93*, 641–650.

Giere, O. & Milligan, M. R. (1989). The gutless, bacteria-symbiotic tubificid *Inanidrilus bulbosus*. Structural analysis of an oligochaete with aberrant reproductive organs. *Zoologischer Anzeiger*, *223*, 174–185.

Giere, O., Nieser, C., Windoffer, R. & Erséus, C. (1995). A comparative structural study on bacterial symbioses of Caribbean gutless Tubificidae (Annelida, Oligochaeta). *Acta Zoologica*, *76*, 281–290.

Hoeh, W. R., Stewart, D. T., Sutherland, B. W. & Zouros, E. (1996). Cytochrome c oxidase sequence comparisons suggest an unusually high rate of mitochondrial evolution in *Mytilus* (Mollusca: Bivalvia). *Molecular Biology and Evolution*, *13*, 418–421.

Hoffmann, R. J., Boore, J. L. & Brown, W. M. (1992). A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis*. *Genetics*, *131*, 397–412.

Holmquist, R. (1983). Transition and transversions in evolutionary descent: an approach to understanding. *Journal of Molecular Evolution*, *19*, 134–144.

Irwin, D. M., Kocher, T. D. & Wilson, A. C. (1991). Evolution of the cytochrome b gene of mammals. *Journal of Molecular Evolution*, *32*, 128–144.

Jamieson, B. G. M. (1988). On the phylogeny and higher classification of the Oligochaeta. *Cladistics*, *4*, 367–410.

Jamieson, B. G. M., Erséus, C. & Ferraguti, M. (1987). Parsimony analysis of the phylogeny of some Oligochaeta (Annelida) using spermatozoal ultrastructure. *Cladistics*, *3*, 145–155.

Kocher, T. D. & Wilson, A. C. (1991). Sequence evolution of mitochondrial DNA in humans and chimpanzees: control region and a protein-coding region. In: S. Osawa, &, T. Honjo (Eds). *Evolution of Life*. pp. 391–413, Springer-Verlag, Tokyo.

Kondo, R., Satoshi, H., Satta, Y. & Takahata, N. (1993). Evolution of hominoid mitochondrial DNA with special reference to the silent substitution rate over the genome. *Journal of Molecular Evolution*, *36*, 517–531.

Maddison, W. P. & Maddison, D. R. (1992). *MacClade*, *Version 3.06 (Software and Documentation)*. Sinauer, Sunderland, Massachusetts.

Palumbi, S. R., Martin, A. P., Romano, S., McMillan, W. O., Stice, L. & Grabowski, G. (1991). *The simple fool's guide to PCR*, *Version 2.0. Department of Zoology and Kewalo Marine Laboratory*. University of Hawaii, Honolulu.

Richards, K. S., Fleming, T. P. & Jamieson, B. G. M. (1982). An ultrastructural study of the distal epidermis and the occurrence of subcuticular bacteria in the gutless tubificid *Phallodrilus albidus* (Oligochaeta: annelida). *Australian Journal of Zoology*, *30*, 327–336.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, *USA*, *74*, 5463–5467.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. & Flook, P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, *87*, 651–701.

Sperling, F. A. H. & Hickey, D. A. (1994). Mitochondrial DNA sequence variation in the spruce bud worm species complex (Choristoneura: Lepidoptera). *Molecular Biology and Evolution*, *11*, 656–665.

Swofford, D. L. (1991). *PAUP: Phylogenetic Analysis Using Parsimony*, Version 3.1.1. Champaign, Illinois Natural History Survey, Illinois.

Swofford, D. L., Olsen, G. J., Waddell, P. J. & Hillis, D. M. (1996). Phylogenetic inference. In: Hillis, D. M., Moritz, C. & Marble, B. K. (Eds). *Molecular Systematics* 2nd edn. pp. 407–514, Sinauer, Sunderland, Massachusetts.

Wakeley, J. (1994). Substitution-rate variation among sites and the estimation of transition bias. *Molecular Biology and Evolution*, *11*, 436–442.

Wakeley, J. (1996). The excess of transitions among nucleotide substitutions: new methods of estimating transition bias underscore its significance. *Trends in Ecology and Evolution*, *11*, 158–163.

Wüster, W., Thorpe, R. S., Cox, M. J., Jintakune, P. & Nabhitabhata, J. (1995). Population systematics of the snake genus *Naja* (Reptilia: Serpentes: Elapidae) in Indochina: multivariate morphometrics and comparative mitochondrial DNA sequencing (cytochrome oxidase I). *Journal of Evolutionary Biology*, *8*, 493–510.

Yang, Z. (1996). Among-site rate variation and its impact on phylogenetic analyses. *Trends in Ecology and Evolution*, *11*, 367–372.