Phylogenetic position of the Pentastomida and [pan]crustacean relationships

Dennis V. Lavrov[1,2], Wesley M. Brown[1] and Jeffrey L. Boore[1,3]

[1] Department of Biology, University of Michigan, 830 North University Avenue, Ann Arbor, MI 48109-1048 USA

[2] Département de Biochimie, Université de Montréal, C.P. 6128, Montreal, Canada, QC H3C 3J7

[3] Department of Evolutionary Genomics, DoE Joint Genome Institute and Lawrence Berkeley National Laboratory, and Department of Integrative Biology, University of California Berkeley, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Corresponding author: Dennis Lavrov, Département de Biochimie, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Qué., H3C 3J7, Canada; Tel. (514) 343-6111 ext. 5188, Fax (514) 343-2210, email dlavrov@bch.umontreal.ca

SUMMARY

Pentastomids are a small group of vermiform animals with unique morphology and parasitic lifestyle. They are generally recognized as being related to the Arthropoda, however the nature of this relationship is controversial. We have determined the complete sequence of the mitochondrial DNA (mtDNA) of the pentastomid *Armillifer armillatus* and complete, or nearly complete, mtDNA sequences from representatives of four previously unsampled groups of Crustacea: Remipedia (*Speleonectes tulumensis*), Cephalocarida (*Hutchinsoniella macracantha*), Cirripedia (*Pollicipes polymerus*), and Branchiura (*Argulus americanus*). Analyses of the mtDNA gene arrangements and sequences determined in this study indicate unambiguously that pentastomids are a group of modified crustaceans likely related to branchiurans. In addition, gene arrangement comparisons strongly support an unforeseen assemblage of pentastomids with maxillopod and cephalocarid crustaceans, to the exclusion of remipedes, branchiopods, malacostracans and hexapods.

Keywords: Pentastomida; mitochondrial DNA; gene rearrangement; arthropod relationships; phylogenetic inference

1. INTRODUCTION

Pentastomids (tongue worms) comprise a small, entirely parasitic group of animals that includes about 130 extant species (Almeida & Christoffersen 1999). They are characterized by having an elongate, vermiform body, most often with pronounced annuli, and a sucking mouth flanked by two pairs of hooks. The hooks were originally misconstrued as additional mouths, hence the name of the group. As adults, pentastomids inhabit the respiratory tracts (mostly the lungs) of vertebrates; about 90% of the species infect reptiles (Riley 1986). Larval development generally occurs in intermediate hosts, for which a variety of vertebrate and invertebrate taxa serve, although a few pentastomids have a direct life cycle. Both adult and larval stages in a pentastomid life cycle are highly specialized for endoparasitism and lack internal organs for respiration, circulation and excretion (Riley 1986). This unusual morphology has confounded the understanding of pentastomid relationships to other animals. At different times the group has been allied with a variety of metazoan phyla, including Arthropoda, Tardigrada, Annelida, Platyhelminthes and Nematoda (reviewed in Almeida & Christoffersen 1999; Haugerud 1989; Riley *et al*. 1978). At present, their alliance with arthropods is generally accepted, although its nature is contentious (reviewed in Zrzavy 2001). Some authors group pentastomids with extant euarthropod lineages, with the most convincing case being made for a Pentastomida - Crustacea (or, more specifically, a Pentastomida – Branchiura) relationship (Abele *et al*. 1989; Abele *et al*. 1992; Giribet & Ribera 2000; Riley *et al*. 1978; Storch & Jamieson 1992; Wingstrand 1972). Others argue against affiliating them with any particular euarthropod taxon, citing the lack of shared characters, unusual cuticular β-chitin and the unique and likely primitive morphological features of the group

(Bockeler 1984; Chesunov 2002; Maas & Waloszek 2001; Walossek & Müller 1998). In particular, the discovery of characteristic pentastomid-like larvae from the Upper Cambrian Alum Shale of Sweden already adapted to a parasitic lifestyle, possibly in gill chambers or related cavities of early marine chordates, has been used as an argument against affiliating pentastomids with any group of extant arthropods (Walossek & Müller 1994; Walossek & Müller 1998). As a reflection of this phylogenetic uncertainty, the Pentastomida has been often treated as an "enigmatic" phylum related to Arthropoda (e.g. Brusca & Brusca 1990), or as a separate subphylum within the Arthropoda (NCBI taxonomic database, Wheeler *et al*. 2000). In an attempt to better resolve the phylogenetic position of the Pentastomida, we determined the mtDNA sequences of the pentastomid *Armillifer armillatus* and four additional crustaceans and performed comparative phylogenetic analysis using both the gene order and the inferred amino acid sequences.

2. MATERIALS AND METHODS

(a) *DNA extraction, amplification, sequencing and annotation*

Total DNA from individual specimens of *Armillifer armillatus*, *Argulus americanus*, *Hutchinsoniella macracantha*, *Pollicipes polymerus* and *Speleonectes tulumensis* was extracted according to (Saghai-Maroof *et al.* 1984). Primers designed to match generally conserved regions of the animal mtDNA were used to amplify short fragments from *cox1* (Folmer *et al.* 1994), *cob* (Boore & Brown 2000), *rns* (Hillis *et al.* 1996), *nad5* (nad5F: 5'-TWYTATTAGGKTGAGATGGKYTNGG-3', nad5-R: 5'-TARAAKCCWGMTATAAAWGGKAWWCC-3'), and the region comprising the 3' end

of *cox1* and the 5' end of cox2 (cox1-F: 5'-CCWCGWCGWTAYTCWGAYAYCCWGA-3', cox2-R: 5'-CWGAATATTCATAWSWTCARTATCATTG-3'). Specific primers (electronic appendix) were designed based on these sequences and used with Perkin Elmer's® XL or Takara® LA PCR kits to amplify the complete mtDNA of each specimen in two or three large overlapping fragments. PCR reaction products were purified by three serial passages through Ultrafree™ (30,000 NMWL) columns (Millipore) and used as templates in dye-terminator, cycle-sequencing reactions according to supplier's (Perkin Elmer®) instructions. Both strands of each amplification product were sequenced by primer walking, using an ABI 377 Automated DNA Sequencer. Sequences were assembled using Sequencing Analysis and Sequence Navigator software (ABI) and analyzed with MacVector 6.5 and GCG (Oxford Molecular Group) programs. Protein and ribosomal RNA gene sequences were identified by their similarity to published metazoan mtDNA sequences; tRNA genes were recognized initially by their potential to be folded into tRNA-like secondary structures, after which they were identified specifically by their anticodon sequences.

(b) ***Phylogenetic analysis***

In order to infer which shared patterns of gene arrangement are derived (i.e. synapomorphies), the newly determined mitochondrial gene arrangements were compared to that of the chelicerate *Limulus polyphemus*, which has previously been inferred to be the ancestral (plesiomorphic) gene order for arthropods. That inference is based on the comparisons of mitochondrial gene orders among all of the four major lineages of arthropods: hexapods, crustaceans, chelicerates and myriapods. There is a

single difference, in the position *trnL(uaa)*, between the mitochondrial gene orders of *L. polyphemus* and several hexapods and crustaceans (Boore 1999). By comparing to outgroup taxa, it was shown that the position of *trnL(uaa)* is derived in hexapods and crustaceans and ancestral in *L. polyphemus* (Boore *et al*. 1998). Similarly, the only rearrangement that differentiates mitochondrial gene orders of L. polyphemus and the centipede *Lithobius forficatus* (the translocation of *trnC*, (Lavrov *et al*. 2000)) is unique for *L. forficatus* and does not appear in the mtDNAs of other myriapods (Lavrov *et al*. 2002). Since the positions of all other genes are identical in mtDNAs of *L. polyphemus* and at least some representatives of hexapods, crustaceans and myriapods, they are parsimoniously inferred to be plesiomorphic for the arthropods. The ancestral status of the gene order found in *L. polyphemus* mtDNA is also supported by our unpublished data on priapulid, tardigrade and onychophoran mitochondrial genomes.

Rearrangements were considered to be synapomorphies only if taxa share both the loss of a gene from its ancestral position and its gain at an identical position elsewhere in the genome, as defined by the two neighboring genes, or if the gene has remained in its original position but its transcriptional polarity has been inverted. Four such shared, derived rearrangements have been identified among the analyzed species and have been used as independent characters for phylogenetic reconstruction based on parsimony.

The species used for sequence-based analysis and their GenBank accession numbers are listed in table 1. Two sets of sequences were analyzed. The first is comprised of the concatenated amino acid sequences inferred from 12 mitochondrial genes (all but *atp8*, which is absent in some of the taxa studied) from 24 species and provides a broad sampling of animal groups outside the Arthropoda. The second contains 22 species and is focused on arthropod relationships. For each of these sets, the amino

acid sequences of individual proteins were aligned three times using the ClustalW (1.82) program (Thompson *et al.* 1994) with different combinations of opening/extension gap penalties: 10/0.2 (default), 12/4 and 5/1. For the last alignment no increased gap penalties near existing gaps, no reduced gap penalties in hydrophilic stretches and no residue-specific penalties were applied. The alignments were compared using the SOAP program (Löytynoja & Milinkovitch 2001) and the positions which were identical among them were concatenated and included in the phylogenetic analysis. The final alignments for the two datasets were 2245 and 2629 amino acids in length, respectively.

We performed maximum likelihood (ML) searches for the best tree using the PROML program within the Phylip 3.6a3 package (Felsenstein 2002) with gamma distributed rates, the JTT matrix of amino acid substitutions and four categories of substitution rates. Alternative topologies were compared using CODEML (Yang 1997) and CONSEL (Shimodaira & Hasegawa 2001) programs. Distances were calculated using the TREE-PUZZLE 5.0 program (Strimmer & von Haeseler 1996), using the mtREV24 matrix, observed frequencies of amino-acids, gamma distributed rates with 8 categories, and an α parameter estimated from the dataset. The distance tree topology was inferred using the WEIGHBOR program (Bruno *et al.* 2000). The bootstrap datasets of 100 (ML) or 1000 (distance) replicates were generated by the SEQBOOT program within the PHYLIP package (Felsenstein 2002). Bootstrap analysis for ML phylogenies was conducted using the PROML program; that for distance-based phylogeny was performed using the "puzzleboot script" by Mike Holder and Andrew Roger (http://hades.biochem.dal.ca/Rogerlab/Software/software.html) and the distance programs listed above.

3. RESULTS

(a) *Pentastomids are related to crustaceans and hexapods*

The mtDNA of *A. armillatus* contains the 37 genes typical of metazoan mtDNAs. The gene arrangement differs from that inferred to be ancestral for arthropods by the positions of five tRNA genes: *trnK*, *trnL(uag)*, *trnL(uaa)*, *trnQ* and *trnS(uga)* (figure 1*a*). One of these differences, the derived location of *trnL(uaa)* between *cox1* and *cox2*, has been previously found in hexapods and crustaceans and used to support their placement in a monophyletic taxon (Pancrustacea (Zrzavy & Stys 1997) or Tetraconata (Dohle 2001)), to the exclusion of myriapods and chelicerates (Boore *et al.* 1998). The finding of the same gene rearrangement in *A. armillatus* strongly supports the inclusion of pentastomids within the Arthropoda as either a sister group to or as a part of the Tetraconata.

Phylogenetic analysis based on the concatenated amino acid sequences from mitochondrial protein coding genes places pentastomid in a well-supported monophyletic group with arthropods and nematodes, as the sister group to the latter taxon. While the placement of nematodes within the Arthropoda would be consistent with some previous analyses based on 18S rDNA sequences (Aguinaldo *et al*. 1997) and while there are some striking similarities in the structural RNAs encoded by pentastomid and nematode mtDNAs (Lavrov 2001), we are reluctant to overemphasize this association, due to the presence of long branches leading to *A. armillatus* and to the nematode species used in this analysis. In addition, when the ML tree topology is changed such that nematodes form a sister group either to the Arthropoda or to the Protostomia, these topologies are not rejected by the approximately unbiased (AU) test (Shimodaira & Hasegawa 2001) ($p = 0.053$ and 0.057, respectively). By contrast, the topologies that have *A. armillatus*

either alone or together with nematodes as the sister taxon to all protostomes are rejected by the AU test (p=0.004 and 0.013, respectively). Aside from this pentastomid-nematode association, there is little resolution within arthropods in either distance-based or ML trees. Since the position of nematodes cannot be resolved based on mitochondrial sequence data and in order to eliminate the potential bias in the phylogenetic reconstruction that fast evolving nematode sequences would cause (Felsenstein 1978), nematode sequences were excluded from subsequent analyses.

(b) ***The position of pentastomids inside Tetraconata***

To further clarify pentastomids' position within Arthropoda, we have determined complete or nearly complete mtDNA sequences from representatives of four previously unsampled groups of Crustacea: Remipedia (*Speleonectes tulumensis*), Cephalocarida (*Hutchinsoniella macracantha*), Cirripedia (*Pollicipes polymerus*) and Branchiura (*Argulus americanus*) (figure 1*b*). Each of these genomes is characterized by a moderate number of gene rearrangements compared to the ancestral arthropod arrangement, most of which are limited to tRNA translocations. Comparisons of arthropod mitochondrial gene arrangements showed that three derived arrangements found in *A. armillatus* are also present in other crustaceans:

i. *trnL(uaa)* is located between *cox1* and *cox2* in all except *H. macracantha* and some highly rearranged genomes;

ii. *trnK* is located between *trnR* and *trnN* in *A. americanus* and *H. macracantha*;

iii. *trnQ* is located between *trnY* and *trnC* in *A. americanus*.

An additional shared and derived arrangement, which appears to be a nearest neighbor exchange between *trnP* and *trnT*, is present in the cirripede *P. polymerus* and the cephalocarid *H. macracantha*. All other derived gene arrangements found in the newly sequenced genomes are autapomorphic and thus not informative for phylogenetic analysis. We regarded each of the rearrangements shared between at least two species as an independent character and constructed a matrix where ancestral arrangements were encoded as 0, and derived gene arrangements either as 1, for synapomorphies, or as 1/0, for autapomorphies. We used this matrix in phylogenetic reconstructions based on the maximum parsimony criterion as implemented in the PAUP* 4.0b10 package (Swofford 2002). Eight most parsimonious trees were found, each 4 steps long, with a CI = 1. The strict consensus of these trees and the gene arrangements informative for phylogenetic reconstruction are presented in figure 3. This analysis places pentastomids in a clade with cephalocarid and maxillopod crustaceans to the exclusion of remipedes, branchiopods, malacostracans and insects.

Both maximum-likelihood and distance analyses based on the restricted dataset (figure 2 *c*, *d*) show medium to high bootstrap support for the monophyly of Collembola, Insecta, Malacostraca (Decapoda), Branchiopoda, Diplopoda and for the association between the pentastomid *A. armillatus* and the branchiuran *A. americanus*. In addition both analyses provide medium to weak bootstrap support for the Tetraconata (77 and 52%, respectively), although the best tree in distance-based reconstruction does not recover this group as monophyletic. By contrast none of the relationships among major tetraconate lineages has been supported with bootstrap values greater than 50%. This lack of resolution contrasts with the previous analyses based on mitochondrial sequence

data (Hwang *et al*. 2001; Nardi *et al*. 2003), which showed high support for several arthropod assemblages, but is similar to the situation with 18S dataset, where addition of extra sequences had a destabilizing effect on phylogenetic trees (Spears & Abele 1998). It should also be noted, that the newly added sequences of pentastomid, branchiuran and remipede, have a significantly deviant amino acid composition as estimated by Tree-Puzzle program. Considering the general lack of resolution in our analysis, the high bootstrap support for the relationship between *A. armillatus* and *A. americanus* should be treated with caution. It should also be noted that these taxa form the longest branches on our phylogenetic trees and therefore may be susceptible to the long-branch attraction artifact (Felsenstein 1978).

4. DISCUSSION

The affinity of pentastomids with branchiuran crustaceans was originally suggested based on the striking similarity in their sperm morphology (Storch & Jamieson 1992; Wingstrand 1972) and has found some additional support from the analysis of 18S rDNA sequences (Abele *et al.* 1989; Abele *et al.* 1992; but see Spears & Abele 1998). Both of these lines of evidence have been questioned, however, by other studies that have pointed to the widespread convergence in sperm structures among invertebrates and the potential problems with 18S rDNA data (Chesunov 2002; Walossek & Müller 1994) and have placed pentastomids as an outgroup to "true" arthropods (Euarthropoda) (Maas & Waloszek 2001). The present analysis of mitochondrial DNA of the pentastomid *A. armillatus*, revealed three derived gene arrangements informative for phylogenetic position of pentastomids. These rearrangements indicate strongly that the Pentastomida

is not an early "pro-arthropod" lineage, but should be placed inside the assemblage of maxillopod and cephalocarid crustaceans, within the Tetraconata. In addition, the ML and distance analyses based on amino acid sequences of mitochondrial protein coding genes support a close relationship between the pentastomid *A. armillatus* and the branchiuran *A. americanus*.

The placement of the Cephalocarida in an assemblage with maxillopod crustaceans and pentastomids, to the exclusion of Branchiopoda, Malacostraca, Hexapoda and Remipedia provides an unexpected hypothesis for the phylogenetic position of this taxon within Crustacea. Cephalocarida, a small group of minute epibenthic crustaceans, has been originally described as a basal lineage of crustaceans (Sanders 1957) but is now thought to be a sister group either to Branchiopoda (Schram 1986), or to a higher taxon that includes Branchiopoda, (Branchiopoda+Malacostraca (Hessler 1992) or Branchiopoda+Maxillopoda (Walossek 1999)). We are aware of only two other studies that have suggested a possible direct link between cephalocarid and maxillopod (copepod) crustaceans (Ito 1989; Spears & Abele 1999), although in both cases the results were inconclusive.

For someone unfamiliar with mitochondrial gene arrangement data, the support of a lineage provided by a single or few mitochondrial gene rearrangements may seem inadequate. However, as should be the case with any character, our confidence in the inference drawn from the gene arrangement data is in reverse proportion to the probability of convergence in gene arrangements. While convergent gene rearrangements in metazoan mtDNA have been observed, both the theoretical considerations (Boore & Brown 1998) and the analysis of ca. 300 complete and many

times this number of partially-sequenced metazoan mitochondrial genomes (Boore 1999, unpublished data) indicates that they are extremely rare. In fact, the known examples of convergent gene rearrangements are limited to the exchange in the positions of two adjacent tRNA genes in several groups of insects (Dowton & Austin 1999; Flook *et al*. 1995). Since such rearrangements can be explained by a relatively simple mechanism, it was suggested that inferences based on nearest-neighbor exchange between adjacent genes and on potentially problematic short-distance rearrangements near the origin of replication should be downweighted in phylogenetic analysis (Boore & Brown 1998). By contrast, the translocations of *trnL(uaa)* and *trnK*, supporting the Tetraconata and the Pentastomida+Maxillopoda+Cephalocarida groups, respectively, occurred over large and otherwise well-conserved regions of the mitochondrial genome and so are unlikely to happen convergently.

The combination of the fossil record and the strong support for the pentastomid-crustacean relationship found in this and in some previous studies presents an interesting paradox for the evolution of pentastomids, and for crustaceans in general. On the one hand, both crown-group crustaceans and pentastomids appear virtually simultaneously in the fossil record at about 500 Mya (Müller 1983; Walossek 1993; Walossek & Müller 1998), and their separation from the oldest crustacean fossils known is by only 10-20 My (Chen *et al*. 2001; Siveter *et al*. 2001). On the other hand, Pentastomida, a morphologically unique group of animals, appears to be related to a particular, subordinate crustacean taxon, the Branchiura. This may point either to the incompleteness of the current crustacean fossil record, or to the possibility that Cambrian "pentastomids" are not directly related to extant pentastomids and that their similarity is

due to convergence. Further studies may help to decide between these two hypotheses, but in the absence of molecular data we believe that it will be extremely difficult either to prove or to disprove convergence between modern and fossil organisms.

ACKNOWLEDGEMENTS

We are grateful to John Riley, William Poly, Jerome Regier, Trisha Spears and Lawrence Abele for the specimens used in this study and to Franz Lang, Hervé Philippe and two anonymous reviewers for helpful suggestions on the preliminary version of this article. This work was supported by National Science Foundation Dissertation Improvement Grant DEB 9972712 to W.M.B. and D.V.L., National Science Foundation Grant DEB 9807100 to W.M.B and J.L.B., and a University of Michigan predoctoral fellowship to D.V.L.). Most of the data analysis was conducted in the laboratory of Dr. B. Franz Lang at the Université de Montréal; part of the work on the manuscript was performed under the auspices of the U.S. Department of Energy, Office of Biological and Environmental Research, in the University of California, Lawrence Berkeley National Laboratory, under contract No. DE-AC03-76SF00098.

REFERENCES

Abele, L. G., Kim, W. & Felgenhauer, B. E. 1989 Molecular evidence for inclusion of the phylum Pentastomida in the Crustacea. *Mol. Biol. Evol.* **6**, 685-691.

Abele, L. G., Spears, T., Kim, W. & Applegate, M. 1992 Phylogeny of selected maxillopodan and other crustacean taxa based on 18S ribosomal nucleotide sequences: a preliminary analysis. *Acta Zool.* **73**, 373-382.

Aguinaldo, A. M. A., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A. & Lake, J. A. 1997 Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**, 489-493.

Almeida, W. D. & Christoffersen, M. L. 1999 A cladistic approach to relationships in Pentastomida. *J. Parasitol.* **85**, 695-704.

Bockeler, W. 1984 Embryogenese und ZNS-Differenzierung bei Reighardia sternae. Licht- und elektronenmikroskopische Untersuchungen zur Tagmosis und systematischen Stellung der Pentastomiden. *Zool. Jahrb. Abt. Anat. Ontogenie Tiere* **111**, 297-342.

Boore, J. L. 1999 Animal mitochondrial genomes. *Nucl. Acids Res.* **27**, 1767-1780.

Boore, J. L. & Brown, W. M. 1998 Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. *Curr. Opin. Genet. Dev.* **8**, 668-674.

Boore, J. L. & Brown, W. M. 2000 Mitochondrial genomes of *Galathealinum*, *Helobdella*, and *Platynereis*: sequence and gene arrangement comparisons indicate that Pogonophora is not a phylum and Annelida and Arthropoda are not sister taxa. *Mol. Biol. Evol.* **17**, 87-106.

Boore, J. L., Lavrov, D. V. & Brown, W. M. 1998 Gene translocation links insects and crustaceans. *Nature* **392**, 667-668.

Bruno, W. J., Socci, N. D. & Halpern, A. L. 2000 Weighted neighbor joining: a likelihood-based approach to distance-based phylogeny reconstruction. *Mol. Biol. Evol.* **17**, 189-197.

Brusca, R. C. & Brusca, G. J. 1990 *Invertebrates*. Sunderland, Massachusetts: Sinauer Associates, Inc.

Chen, J. Y., Vannier, J. & Huang, D. Y. 2001 The origin of crustaceans: new evidence from the Early Cambrian of China. *Proc. R. Soc. Lond. B* **268**, 2181-2187.

Chesunov, A. V. 2002 A case of tongueworms (Pentastomida): a specific problem in context of the modern phylogenetics. *Zh. Obshch. Biol.* **63**, 209-226.

Dohle, W. 2001 Are the insects terrestrial crustaceans? A discussion of some new facts and arguments and the proposal of the proper name 'Tetraconata' for the monophyletic unit Crustacea + Hexapoda. *Ann. Soc. Entomol. Fr. (N.S.)* **37**, 85-103.

Dowton, M. & Austin, A. D. 1999 Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. *Mol. Biol. Evol.* **16**, 298-309.

Felsenstein, J. 1978 Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* **27**, 401-410.

Felsenstein, J. 2002 *PHYLIP (Phylogeny Inference Package). Version 3.6a3*. Seattle: University of Washington.

Flook, P., Rowell, H. & Gellissen, G. 1995 Homoplastic rearrangements of insect mitochondrial tRNA genes. *Naturwissenschaften* **82**, 336-337.

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. Biotechnol.* **3**, 294-299.

Giribet, G. & Ribera, C. 2000 A Review of Arthropod Phylogeny: New Data Based on Ribosomal DNA Sequences and Direct Character Optimization. *Cladistics* **16**, 204-231.

Haugerud, R. E. 1989 Evolution in the pentastomids. *Parasitol. Today* **5**, 126-132.

Hessler, R. R. 1992 Reflections on the phylogenetic position of the Cephalocarida. *Acta Zool.* **73**, 315-316.

Hillis, D. M., Moritz, C. & Mable, B. K. (ed.) 1996 *Molecular Systematics*. Sunderland, MA: Sinauer Associates.

Hwang, U. W., Friedrich, M., Tautz, D., Park, C. J. & Kim, W. 2001 Mitochondrial protein phylogeny joins myriapods with chelicerates. *Nature* **413**, 154-7.

Ito, T. 1989 Origin of the basis in copepod limbs, with reference to remipedian and cephalocarid limbs. *J. Crust. Biol.* **9**, 85-103.

Lavrov, D. V. 2001 Arthropod phylogeny based on gene arrangement and other characters from mitochondrial DNA, pp. 196. Ann Arbor, MI: University of Michigan.

Lavrov, D. V., Boore, J. L. & Brown, W. M. 2002 Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: duplication and nonrandom loss. *Mol. Biol. Evol.* **19**, 163-169.

Lavrov, D. V., Brown, W. M. & Boore, J. L. 2000 A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proc. Natl. Acad. Sci. USA* **97**, 13738-13742.

Löytynoja, A. & Milinkovitch, M. C. 2001 SOAP, cleaning multiple alignments from unstable blocks. *Bioinformatics* **17**, 573-574.

Maas, A. & Waloszek, D. 2001 Cambrian Derivatives of the Early Arthropod Stem Lineage, Pentastomids, Tardigrades and Lobopodians - An 'Orsten' Perspective. *Zool. Anz.* **240**, 451-459.

Müller, K. J. 1983 Crustacea with preserved soft parts from the Upper Cambrian of Sweden. *Lethaia* **16**, 93-109.

Nardi, F., Spinsanti, G., Boore, J. L., Carapelli, A., Dallai, R. & Frati, F. 2003 Hexapod origins: monophyletic or paraphyletic? *Science* **299**, 1887-9.

Riley, J. 1986 The biology of pentastomids. *Adv. Parasitol.* **25**, 45-128.

Riley, J., Banaja, A. A. & James, J. L. 1978 The phylogenetic relationship of the Pentastomida: the case for their inclusion within the Crustacea. *Int. J. Parasitol.* **8**, 245-254.

Saghai-Maroof, M. A., Soliman, K. M., Jorgensen, R. A. & Allard, R. W. 1984 Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci. USA* **81**, 8014-8018.

Sanders, H. L. 1957 The Cephalocarida and crustacean phylogeny. *Syst. Zool.* **6**, 112-129.

Schram, F. R. 1986 *Crustacea*. New York: Oxford University Press.

Shimodaira, H. & Hasegawa, M. 2001 CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**, 1246-1247.

Siveter, D. J., Williams, M. & Waloszek, D. 2001 A phosphatocopid crustacean with appendages from the Lower Cambrian. *Science* **293**, 479-481.

Spears, T. & Abele, L. G. 1998 Crustacean phylogeny inferred from 18S rDNA. In *Arthropod relationships* (ed. R. A. Fortey & R. H. Thomas), pp. 171-189. London: Chapman and Hall.

Spears, T. & Abele, L. G. 1999 Phylogenetic relationships of crustaceans with foliaceous limbs: an 18S rDNA study of Branchiopoda, Cephalocarida, and Phyllocarida. *J. Crust. Biol.* **19**, 825-843.

Storch, V. & Jamieson, B. G. M. 1992 Further spermatological evidence for including the Pentastomida (tongue worms) in the Crustacea. *Int. J. Parasitol.* **22**, 95-108.

Strimmer, K. & von Haeseler, A. 1996 Quartet puzzling: A quartet maximum likelihood method for reconstructing tree topologies. Mol. Biol. Evol. 13: 964-969.

Swofford, D. L. 2002 *PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4*. Sunderland, Massachusetts: Sinauer Associates.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673-4680.

Walossek, D. 1993 The Upper Cambrian *Rehbachiella kinneullensis* and the phylogeny of Branchiopoda and Crustacea. *Fossils and Strata* **32**, 1-202.

Walossek, D. 1999 On the Cambrian diversity of Crustacea. In *Crustaceans and the Biodiversity Crisis: Proceedings of the Fourth International Crustacean Congress, Amsterdam, the Netherlands, July 20-24, 1998* (ed. F. R. Schram & J. C. von Vaupel Klein), pp. 3-27. Leiden: Brill Academic Publishers.

Walossek, D. & Müller, K. J. 1994 Pentastomid parasites from the Lower Paleozoic of Sweden. *Trans. R. Soc. Edinb. -Earth Sci.* **85**, 1-37.

Walossek, D. & Müller, K. J. 1998 Cambrian 'Orsten'-type arthropods and the phylogeny of Crustacea. In *Arthropod relationships* (ed. R. A. Fortey & R. H. Thomas), pp. 139-153. London: Chapman and Hall.

Wheeler, D. L., Chappey, C., Lash, A. E., Leipe, D. D., Madden, T. L., Schuler, G. D., Tatusova, T. A. & Rapp, B. A. 2000 Database resources of the National Center for Biotechnology Information. *Nucl. Acids Res.* **28**, 10-14.

Wingstrand, K. G. 1972 Comparative spermatology of a pentastomid, *Raillietiella hemidactyli*, and a branchiuran crustacean, *Argulus foliaceus*, with a discussion of pentastomid relationships. *Kong. Dansk. Videnskab. Selskab Biol. Skrifter* **19**, 1-72.

Yang, Z. 1997 PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555-556.

Zrzavy, J. 2001 The interrelationships of metazoan parasites: a review of phylum-and higher-level hypotheses from recent morphological and molecular phylogenetic analyses. *Folia Parasitol. (Praha)* **48**, 81-103.

Zrzavy, J. & Stys, P. 1997 The basic body plan of arthropods: insights from evolutionary morphology and developmental biology. *J. Evol. Biol.* **10**, 353-367.

Table 1. *Species, taxonomic classification, and accession numbers*

Species	Classification	Accession number
Armillifer armillatus	Pentastomida	AY456186*
Argulus americanus	Crustacea, Maxillopoda	AY456187*
Pollicipes polymerus	Crustacea, Maxillopoda	AY456188*
Tigriopus japonicus	Crustacea, Maxillopoda	AB060648
Hutchinsoniella macracantha	Crustacea, Cephalocarida	AY456189*
Speleonectes tulumensis	Crustacea, Remipedia	AY456190*
Artemia franciscana	Crustacea, Branchiopoda	X69067
Daphnia pulex	Crustacea, Branchiopoda	AF117817
Triops cancriformis	Crustacea, Branchiopoda	AB084514
Pagurus longicarpus	Crustacea, Malacostraca	AF150756
Panulirus japonicus	Crustacea, Malacostraca	AB071201
Penaeus monodon	Crustacea, Malacostraca	AF217843
Portunus trituberculatus	Crustacea, Malacostraca	AB093006
Tetrodontophora bielanensis	Hexapoda, Collembola	AF272824
Gomphiocephalus hodgsoni	Hexapoda, Collembola	AY191995
Drosophila yakuba	Hexapoda, Insecta	X03240
Locusta migratoria	Hexapoda, Insecta	X80245
Tricholepidion gertschi	Hexapoda, Insecta	AY191994
Ixodes hexagonus	Chelicerata, Arachnida	AF081828
Limulus polyphemus	Chelicerata, Merostomata	AF216203

Lithobius forficatus	Myriapoda, Chilopoda	AF309492
Narceus annularus	Myriapoda, Diplopoda	AY055727
Thyropygus sp.	Myriapoda, Diplopoda	AY055728
Trichinella spiralis	Nematoda, Enoplea	AF293969
Onchocerca volvulus	Nematoda, Chromadorea	AF015193
Lumbricus terrestris	Annelida, Olygochaeta	U24570
Platynereis dumerilii	Annelida, Polychaeta	AF178678
Katharina tunicata	Mollusca, Polyplacophora	U09810
Loligo bleekeri	Mollusca, Cephalopoda	AB029616
Homo sapiens	Chordata, Mammalia	AF347015
Mustelus manazo	Chordata, Chondrichthyes	AB015962
Asterina pectinifera	Echinodermata, Asterozoa	D16387
Florometra serratissima	Echinodermata, Crinoidea	AF049132
Metridium senile	Cnidaria, Anthozoa	AF000023
Sarcophyton glaucum	Cnidaria, Anthozoa	AF064823, AF063191

*New sequence

Figure captions

Figure 1. Comparison of gene arrangements in the mtDNA of the pentastomid *Armillifer armillatus* with those inferred for the ancestral pancrustacean, the ancestral arthropod (*a*), and from species representing the crustacean classes Maxillopoda, Cephalocarida and Remipedia (*b*). Genes are not drawn to scale; protein and rRNA genes are indicated by larger boxes, tRNA genes by smaller boxes, and non-coding regions by checkered boxes. Genes are transcribed from left to right except when underlined; underlining indicates the opposite transcriptional polarity. Genes that have been translocated compared to their position in the *L. polyphemus* genome are in grey. Protein and ribosomal RNA gene abbreviations: *atp6*, *8* – subunits 6 and 8 of F0 ATPase; *cob* – apocytochrome *b*; *cox1-3* – cytochrome *c* oxidase subunits 1-3; *nad1-6* and *nad4L* – NADH dehydrogenase subunits 1-6 and 4L; *rns* and *rnl* – small and large subunit rRNAs. tRNA gene abbreviations use the one letter amino acid code; the two leucine and two serine tRNA genes are differentiated by their anticodon sequences, with *trnL(uag)* marked as L_1, *trnL(uaa)* - as L_2, *trnS(ucu)* - as S_1 and *trnS(uga)* - as S_2. The second copy of *trnH* in the *S. tulumensis* genome is marked as H'. Two tandemly repeated copies of *trnC* found in *P. polymerus* genome are not shown.

Figure 2. Phylogenetic position of the Pentastomida based on maximum likelihood (*a*, *c*) and distance (*b*, *d*) analyses of amino acid sequences inferred from mitochondrial protein coding genes. Numbers above the branches indicate bootstrap support percentages based on the analysis of 100 (ML) or 1000 (distance) bootstrap replicates. Only the values

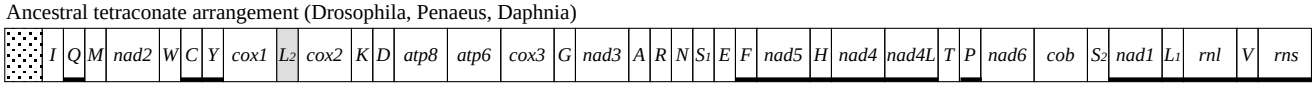
relevant to the position of the Pentastomida are shown in (*a*) and (*b)*; those above 50% are shown in (*c*) and (*d*).

Figure 3. Phylogenetic analysis based on mitochondrial gene arrangements. The tree represents a strict consensus of 8 MP trees. Four informative gene rearrangements identified among different "classes" of Crustacea have been used for the analysis all based on tRNA translocation. Plesiomorphic and synapomorphic positions of these tRNA genes are shown by blocks; autapomorphic positions are indicated by small vertical bars. Genes are designated as in figure 1. Transcription is from left to right except for the genes marked with minus.

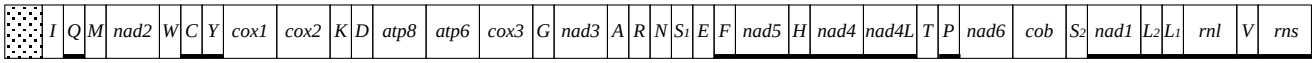
(*a*)

Armillifer armillatus (Pentastomida)

Ancestral tetraconate arrangement (Drosophila, Penaeus, Daphnia)

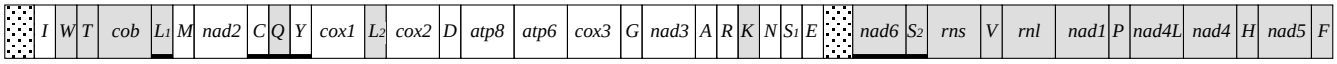


Ancestral arthropod arrangement (Limulus)

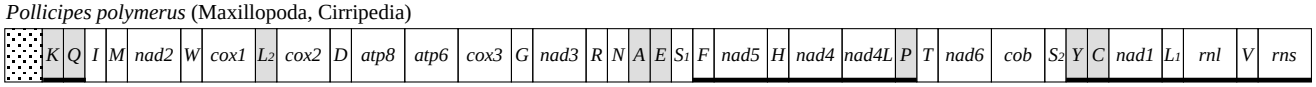


(*b*)

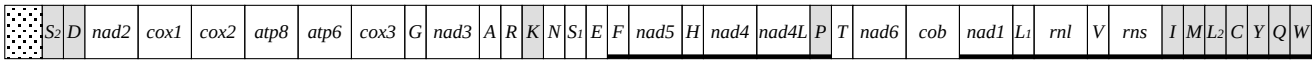
Argulus americanus (Maxillopoda, Branchiura)



Pollicipes polymerus (Maxillopoda, Cirripedia)



Hutchinsoniella macracantha (Cephalocarida)



Speleonectes tulumensis (Remipedia)

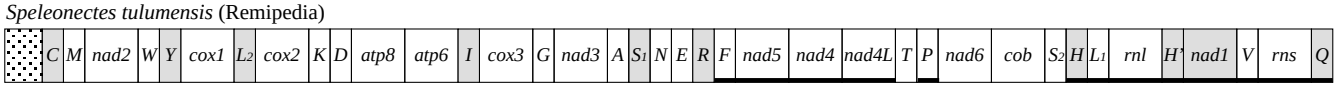


Figure 1

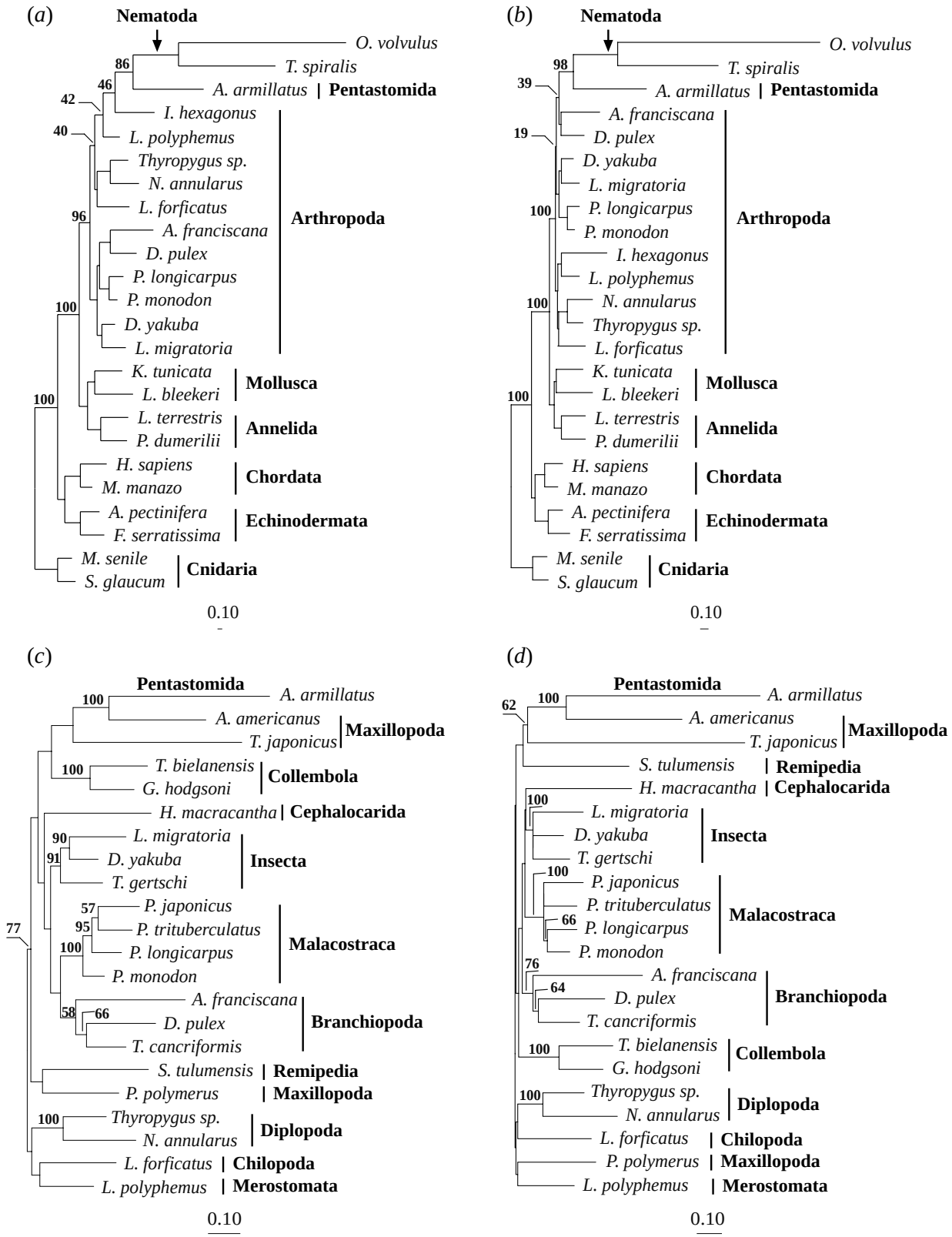

(a)
Nematoda
O. volvulus
T. spiralis
A. armillatus | Pentastomida
I. hexagonus
L. polyphemus
Thyropygus sp.
N. annularus
L. forficatus
A. franciscana
D. pulex
P. longicarpus
P. monodon
D. yakuba
L. migratoria
Arthropoda
K. tunicata
L. bleekeri
Mollusca
L. terrestris
P. dumerilii
Annelida
H. sapiens
M. manazo
Chordata
A. pectinifera
F. serratissima
Echinodermata
M. senile
S. glaucum
Cnidaria
86
46
42
40
96
100
100
0.10
(b)
Nematoda
O. volvulus
T. spiralis
A. armillatus | Pentastomida
A. franciscana
D. pulex
D. yakuba
L. migratoria
P. longicarpus
P. monodon
I. hexagonus
L. polyphemus
N. annularus
Thyropygus sp.
L. forficatus
Arthropoda
K. tunicata
L. bleekeri
Mollusca
L. terrestris
P. dumerilii
Annelida
H. sapiens
M. manazo
Chordata
A. pectinifera
F. serratissima
Echinodermata
M. senile
S. glaucum
Cnidaria
98
39
19
100
100
100
0.10
(c)
Pentastomida
A. armillatus
A. americanus
T. japonicus
Maxillopoda
T. bielanensis
G. hodgsoni
Collembola
H. macracantha | Cephalocarida
L. migratoria
D. yakuba
T. gertschi
Insecta
P. japonicus
P. trituberculatus
P. longicarpus
P. monodon
Malacostraca
A. franciscana
D. pulex
T. cancriformis
Branchiopoda
S. tulumensis | Remipedia
P. polymerus | Maxillopoda
Thyropygus sp.
N. annularus
Diplopoda
L. forficatus | Chilopoda
L. polyphemus | Merostomata
100
100
90
91
57
95
100
66
58
77
100
0.10
(d)
Pentastomida
A. armillatus
A. americanus
T. japonicus
Maxillopoda
S. tulumensis | Remipedia
H. macracantha | Cephalocarida
L. migratoria
D. yakuba
T. gertschi
Insecta
P. japonicus
P. trituberculatus
P. longicarpus
P. monodon
Malacostraca
A. franciscana
D. pulex
T. cancriformis
Branchiopoda
T. bielanensis
G. hodgsoni
Collembola
Thyropygus sp.
N. annularus
Diplopoda
L. forficatus | Chilopoda
P. polymerus | Maxillopoda
L. polyphemus | Merostomata
62
100
100
100
66
76
64
100
100
0.10

Figure2

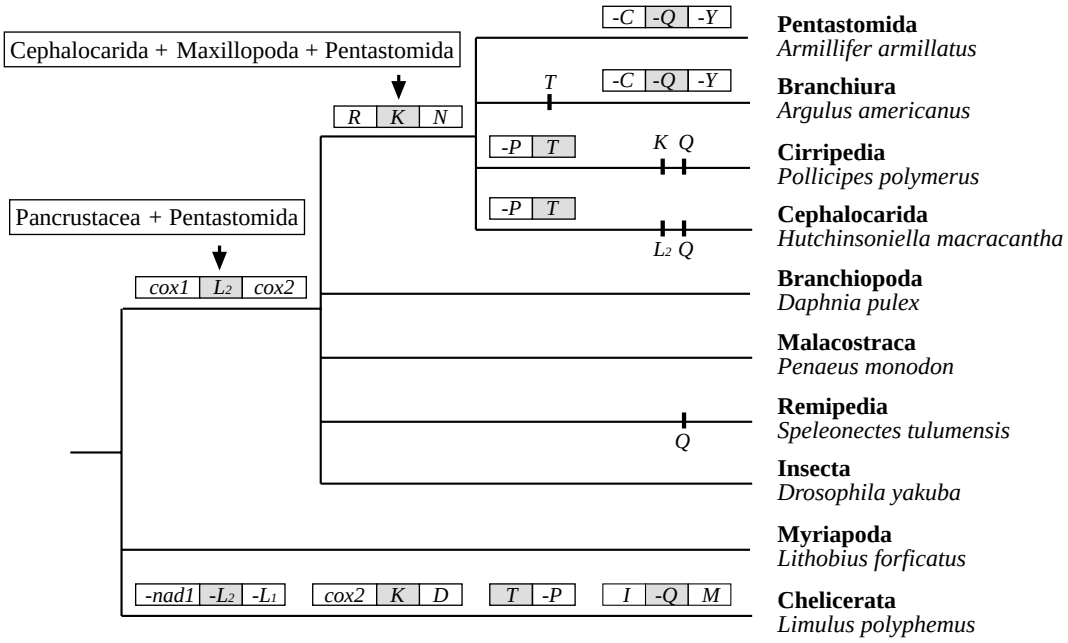

Cephalocarida + Maxillopoda + Pentastomida
R K N
Pancrustacea + Pentastomida
cox1 L2 cox2
-C -Q -Y
T
-C -Q -Y
-P T
K Q
-P T
L2 Q
Q
-nad1 -L2 -L1
cox2 K D
T -P
I -Q M
Pentastomida
Armillifer armillatus
Branchiura
Argulus americanus
Cirripedia
Pollicipes polymerus
Cephalocarida
Hutchinsoniella macracantha
Branchiopoda
Daphnia pulex
Malacostraca
Penaeus monodon
Remipedia
Speleonectes tulumensis
Insecta
Drosophila yakuba
Myriapoda
Lithobius forficatus
Chelicerata
Limulus polyphemus

Figure 3

Electronic appendix. Specific primers used for mtDNA amplification

Primer	Sequence
Armillifer-cox1-F1	5'-CAACCATCCTTAACATACACTCC-3'
Armillifer-cox1-R1	5'-TCATAGGGATGAGTCAGTTCCCG-3'
Armillifer-cob-F1	5'-GATTCCTTCACAGCAATCACCAC-3'
Armillifer-cob-R1	5'-GATGTTGTGTCCACCTCATAGTC-3'
Pollicipes-cox1-F1	5’-TGATCAGCGATAGTTGGAACAGC-3’
Pollicipes-cox1-R1	5’-CTTCCAGGTTGACCTAGTTCTGC-3’
Pollicipes-cob-F1	5’-TGCAGTAGATAATGCTACCTTGAC-3’
Pollicipes-cob-R1	5’-AGCAGGTGGATAACAGCTATTGC-3’
Speleonectes-cox2-F1	5’-ATCAGACGAGCCCATATCACCAC-3’
Speleonectes-cox2-R1	5’-AAGCCTTCTGCACTAGTCTGCCA-3’
Speleonectes-cob-F1	5’-CCATTCCATCCATACTTCCTAAC-3’
Speleonectes-cob-R1	5’-GTCTTTCCAATGTAAGGTACTGC-3’
Argulus-cox1-F1	5’-CAGTAGAGTCTGGAGCTGGAAC-5’
Argulus-cox2-R1	5-AGTATGTCTAATAGGAGGAGCTC-3’
Argulus-rns-F1	5’-AACGATAATCCACGAGTATTCTAC-3’
Argulus-rns-R1	5’-ACTTTTACTTCTAAATCCTCCTTCA-3’
Argulus-nad5-F1	5’-TTCAGCCTGACTTCCTGCTGC-3’
Argulus-nad5-R1	5’-TTCATAACATCCAGATAACCAGA-3’
Hutchinsoniella-cox1-F1	5’-CTTTGCCTATTTTGGCTGGGGCTA-3’
Hutchinsoniella-cox1-R1	5’-CCGCCTGACCTAACTCAGATCGA-3’

Hutchinsoniella-nad5-F1	5’-TTCTCCGCTTGATTACCCATGGC-3’
Hutchinsoniella-nad5-R1	5’-CTCAGATAAACACCCGCTGTGAC-3’
Hutchinsoniella-rns-F1	5’-TTGGTTAAACATGTGCCAGCTAC-3’
Hutchinsoniella-rns-R1	5’-TTAGCTATACCATGGCTTGACATC-3’