doi:10.1038/nature10382

SUPPLEMENTARY METHODS

**Molecular techniques.** Slightly different methods were used by the Halanych, Lieb, Moroz and Todt labs to prepare cDNA for sequencing (Supplementary Table 2). For the Halanych Lab taxa, total RNA was extracted from frozen or RNAlater-fixed tissue using TRIzol (Invitrogen) and purified using the RNeasy kit (Qiagen) with on-column DNase digestion. Specimens of *Wirenia argentea* were starved for approximately 2 months prior to RNA extraction to reduce cnidarian contamination. First-strand cDNA was synthesized using the SMART cDNA library construction kit (Clontech). Full-length cDNA was then amplified using the Advantage 2 PCR system (Clontech) and normalized using the Trimmer-Direct kit (Evrogen). Normalized cDNA was sent to The University of South Carolina Environmental Genomics Core Facility (Columbia, SC, USA) for sequencing using 454 GS-FLX or Titanium (Roche). For the Lieb Lab taxa, total RNA was extracted from fresh or liquid nitrogen frozen tissue using TRIzol (Invitrogen) with DNase digestion using Nucleospin (Machery-Nagel). First-strand cDNA was prepared by the Max Planck Institute for Molecular Genetics using the Mini kit (Evrogen) followed by size selection with Chromaspin 1000 columns (Clontech). Full-length cDNA was amplified using PCR and normalized using the Trimmer-Direct kit (Evrogen). Normalized cDNA was sequenced using 454 Titanium (Roche). For the Moroz Lab taxa, total RNA was extracted with RNAqueous (Life Technologies) and reverse transcribed to cDNA using the Marathon cDNA amplification kit (BD Biosciences) and an oligo dT primer. Double-stranded cDNA was digested with a restriction enzyme followed by adaptor ligation to both ends. The adaptor ligated cDNA fragments were then amplified by PCR, purified, and sent to The University of Florida Interdisciplinary Center for Biotechnology Research (Gainesville, FL, USA) or SeqWright (Houston, TX, USA) for sequencing using 454 GS-FLX or Titanium. For the Todt Lab taxa, total RNA which was extracted using TRIzol (Invitrogen). For *Scutopus*, PolyA+ RNA was isolated using PolyATract (Promega) and for *Wirenia*, animals were

starved for approximately two months prior to total RNA isolation using the RNeasy kit (Qiagen) and cDNA library construction using the Creator SMART cDNA Library construction kit (Clontech) by GENterprise (Mainz, Germany) with directional cloning using a modified pSPORT vector. Around 1,000 clones were sequenced for both taxa using an ABI 3730 (Applied Biosystems).

**Orthology determination.** In order to identify as many genes suitable for phylogenetic analysis as possible, we initially attempted to construct a set of trochozoan core orthologs based on the genomes of *Lottia gigantea*, *Capitella teleta*, and *Helobdella robusta* with *Nematostella vectensis* or *Drosophila melanogaster* as an outgroup. Sets of core orthologs were constructed from the predicted transcripts of these taxa using inflation parameters ranging from 1.8 to 2.2. Surprisingly, these sets of core orthologs typically only included 500-600 genes versus the 1,032 genes in the HaMStR "model organisms" set of core othologs based on *Homo*, *Ciona*, *Drosophila*, *Caenorhabditis*, and *Saccharomyces*. This result is apparently due, at least in part, to a partial genome duplication in the leech *Helobdella*[40-41] which resulted in all genes duplicated in *Helobdella* (but single-copy in *Lottia* and *Capitella*) to be exluded by HaMStR. Notably, the presence of two inparalogs of many genes in the *Helobdella* genome was observed frequently during our manual evaluations of the alignments and ML trees. Additional high-quality, completely sequenced genomes from other taxa will undoubtedly aid future studies in identifying lineage-specific orthologous genes suitable for phylogenetic analysis.

To evaluate our orthology determination method that utilized *Drosophila* as the primer taxon, comparisons were made between *Lottia* sequences identified as orthologs to the *Drosophila* sequences using our methods relative to orthologs identified using the InParanoid 7 database. This revealed only 6 instances in which both methods identified one or more *Lottia* sequences as orthologs to a *Drosophila* sequence, but disagreed on

which sequence was the correct ortholog. There were 36 *Drosophila* genes for which InParanoid did not identify a *Lottia* ortholog but our methods did. Alternatively, there were 6 *Drosophila* sequences for which our methods did not identify a *Lottia* ortholog but InParanoid did. Although this may give the impression that our methods are less stringent than those of InParanoid, manual examination of these alignments revealed no obvious paralogous groups.

## SUPPLEMENTARY RESULTS

**Outgroup selection.** To further explore how outgroup choice affected topology, additional analyses with other sets of taxa as outgroups were performed. ML and BI analyses of a reduced dataset including all taxa listed in Supplementary Table 3 except the cnidarian *Nematostella* (Supplementary Figs 14-15) yielded the same relationships as Fig 2 thus providing additional support for the major findings of this study. However, support for Pleistomollusca was substantially decreased in ML (bs = 46) but not BI (pp = 100).

Because Entoprocta + Cycliophora were weakly placed as the sister taxon of Mollusca in the BI analysis including broad outgroups (Supplementary Fig 4), and because a close relationship between Mollusca and Entoprocta has been suggested before[42-43], we performed an ML analysis using Entoprocta + Cycliophora as the outgroup to Mollusca (Supplementary Fig 16a). Interestingly, in this tree the positions of Cephalopoda and Aculifera are reversed with respect to the annelid-rooted tree (Fig 2). One possible explanation for these results is homoplasy introduced by the long-branched cycliophoran *Symbion*. Therefore, we performed an additional ML analysis using only Entoprocta as the outgroup to Mollusca. Importantly, in the ML tree excluding *Symbion* (Supplementary Fig 16b), we recovered the same branching order among the major lineages of Mollusca as in Fig 2 with strong support at most nodes. This result suggests

that the long-branch taxon *Symbion* introduces homoplasy into the dataset and provides additional support for the Fig 2 topology.

**Undescribed *Neomenia* sp.** The *Neomenia* included in this study appears to be a new, undescribed species. Examination of three the paratypes of *Neomenia herwigi* Kaiser, 1976 (Zoological Museum Hamburg, accession numbers ZMH 1010/1, ZMH 1019) and a specimen of *Neomenia permagna* Salvini-Plawen, 1978 (Halanych collection) allowed for comparisons of sclerites and internal morphology. Although our *Neomenia* is found in the same geographical region as *N. herwigi* and shows some similarities to this species, substantial differences in the morphometric properties and abundances of certain sclerite types were observed. Also, our specimen is much smaller than the *N. herwigi* holotype and paratypes (about 1/4 of the length) and already in a reproductive state. However, it should be noted that neomenioids usually reach sexual maturity long before they stop growing.

SUPPLEMENTARY REFERENCES

40. Veenstra, J. A. Neuropeptide evolution: Neurohormones and neuropeptides predicted from the genomes of *Capitella teleta* and *Helobdella robusta*. Gen. Comp. Endocr. 171, 160-175 (2011).

41. Cho, S. J., Valles, Y., Giani, V. C., Seaver, E. C. & Weisblat, D. A. Evolutionary dynamics of the Wnt gene family: a lophotrochozoan perspective. *Mol. Biol. Evol.* **27**, 1645 (2010).

42. Wanninger, A. Shaping the things to come: ontogeny of lophotrochozoan neuromuscular systems and the Tetraneuralia concept. *Biol. Bull.* **216**, 293 (2009).

43. Wanninger, A., Fuchs, J. & Haszprunar, G. Anatomy of the serotonergic nervous system of an entoproct creeping-type larva and its phylogenetic implications.

*Invertebr. Biol.* **126**, 268–278 (2007).

44. Wanninger, A. & Haszprunar, G. Chiton myogenesis: Perspectives for the development and evolution of larval and adult muscle systems in molluscs. *J. Morphol.* **251**, 103-113 (2002).
45. Voss-Foucart, M.F., Fonze-Vignaux, M.T. & Jeuniaux, C. Systematic characters of some polychaetes (Annelida) at the level of the chemical composition of the jaws. *Biochem. Syst. Ecol.* **1**, 119–122 (1973).
46. Todt, C., Büchinger, T. & Wanninger, A. The nervous system of the basal mollusk *Wirenia argentea* (Solenogastres): a study employing immunocytochemical and 3D reconstruction techniques. *Mar. Biol. Res.* **4**, 290–303 (2008).
47. Hyman, L.H. *The Invertebrates. Vol. VI. Mollusca I.* 792 pp. (McGraw-Hill, New York, 1967).
48. Haszprunar, G. The first molluscs-small animals. *Italian Journal of Zoology* **59**, 1–16 (1992).
49. Lundin, K. & Schander, C. Ultrastructure of gill cilia and ciliary rootlets of *Chaetoderma nitidulum* Lovén 1844 (Mollusca, Chaetodermomorpha). *Acta Zool. Stockholm* **80**, 185–191 (1999).
50. Lundin, K. & Schander, C. Ciliary ultrastructure of neomeniomorphs (Mollusca, Neomeniomorpha= Solenogastres). *Invertebr. Biol.* **120**, 342–349 (2001).
51. Salvini-Plawen, L. The significance of the Placophora for molluscan phylogeny. *Venus* **65**, 1–17 (2006).
52. Todt, C. & von Salvini-Plawen, L. The digestive tract of *Helicoradomenia* (Solenogastres, Mollusca), aplacophoran molluscs from the hydrothermal vents of the East Pacific Rise. *Invertebr. Biol.* **124**, 230–253 (2005).

53. Lieb, B. & Todt, C. Hemocyanin in mollusks – a molecular survey and new data on hemocyanin genes in Solenogastres and Caudofoveata. Mol Phylogenet Evol. **49**. 382-385. (2008).

54. Lieb, B. & Wilson, N.G. The hemocyanin of *Laevipilina hyalina* (Monoplacophora). World Congrees of Malacology 2010 Poster Presentation (Phuket, Thailand, 2010).

55. Furuhashi, T., Schwarzinger, C., Miksik, I., Smrz, M. & Beran, A. Molluscan shell evolution with review of shell calcification hypothesis. *Comp. Biochem. Phys. B* **154**, 351–371 (2009).

56. Cruz, R., Lins, U. & Farina, M. Minerals of the radular apparatus of *Falcidens* sp. (Caudofoveata) and the evolutionary implications for the phylum Mollusca. *Biol. Bull.* **194**, 224-230 (1998).

**Supplementary Table 1 | Approximately Unbiased (AU) test results of Wilson et al. 2009 dataset.**

| Alternative Hypothesis | Ln Likelihood Score | P-value | Significantly Worse? |
|---|---|---|---|
| Best tree | -207,597.42 | | |
| All major lineages monophyletic, Mollusca monophyletic | -207,656.98 | 0.019 | Yes |
| Adenopoda (Chaetodermomorpha basal) | -207,688.19 | 0.014 | Yes |
| Hepagastralia (Neomeniomorpha basal) | -207,593.58 | 0.916 | No |
| Aculifera | -207,727.51 | 3e-04 | Yes |
| Aculifera + Monoplacophora | -207,712.06 | 3e-09 | Yes |
| Testaria (Polyplacophora + Conchifera) | -207,646.75 | 0.018 | Yes |
| Diasoma (Bivalvia + Scaphopoda) | -207,639.05 | 0.084 | No |
| Cyrtosoma (Gastropoda + Cephalopoda) | -207,626.04 | 0.231 | No |
| Cephalopoda + Scaphopoda | -207,627.16 | 0.229 | No |
| Gastropoda + Scaphopoda | -207,661.85 | 0.007 | Yes |

| | | | |
|---|---|---|---|
| Gastropoda + Bivalvia | -207,626.63 | 0.215 | No |

AU tests performed based on a tree inferred from a RAxML 7.2.7 re-analysis of the Muscle-aligned matrix of Wilson et al. 2009[3] (TreeBase matrix ID M4763) with joint branch length optimization of parameters under GTR + Γ model.

## Supplementary Table 2 | Specimen data for sequenced taxa.

| Species | Major Lineage | Tissue | Collection Location | Laboratory |
|---|---|---|---|---|
| *Antalis vulgaris* | Scaphopoda | Whole animal (no shell), starved 1 week | Roscoff, France | Lieb |
| *Dolabrifera dolabrifera* | Gastropoda | Neural tissue | Costa Rica | Moroz, Wright |
| *Hanleya nagelfar* | Polyplacophora | Foot tissue | Bergen, Norway | Halanych |
| *Helisoma trivolvis* | Gastropoda | Neural tissue | Biological Supplies | Moroz, Rehder |
| *Hermissenda crassicornis* | Gastropoda | Neural tissue | Friday Harbor, WA | Moroz |
| *Loligo opalescens* | Cephalopoda | Neural tissue | Friday Harbour, WA | Moroz |
| *Loligo pealei* | Cephalopoda | Neural tissue | Woods Hole, MA | Moroz |
| *Loligo vulgaris* | Cephalopoda | Neural tissue | Naples, Italy | Moroz |
| *Melanoides tuberculatus* | Gastropoda | Whole animal | Aquarium population Institute of Zoology, Mainz, Germany | Lieb |
| *Nautilus pompilius* | Cephalopoda | Neural tissue | Pacific | Moroz |
| *Neomenia* sp. | Neomeniomorpha | Mantle tissue | Trinity Peninsula, Antarctica (S63°23.05' W60°03.40'), 277 m | Halanych |
| *Nucula nitidosa* | Bivalvia | Whole animal, starved 1 week | Roscoff, France | Lieb |
| *Octopus vulgaris* | Cephalopoda | Neural tissue | Naples, Italy | Moroz, |

| | | | | |
|---|---|---|---|---|
| | | | | DiCosma |
| *Pleurobranchaea californica* | Gastropoda | Neural tissue | Monterey, CA | Moroz, Gillette |
| *Polyschides dalli antarcticus* | Scaphopoda | Whole animal | Eagle Island, Antarctica (S63°40.00', W57°19.75'), 335 m | Halanych |
| *Scutpus ventrolineatus* (454) | Chaetodermo-morpha | 2 adults | Skagerrak Strait (58°22.84',10°19.44'), 335 m | Halanych |
| *Scutpus ventrolineatus*[*] (Capillary) | Chaetodermo-morpha | Several adults | Bergen, Norway | Todt |
| *Solemya velum* | Bivalvia | Mantle tissue, starved 1 week | Woods Hole, MA | Lieb |
| *Theodoxus fluvatilis* | Gastropoda | Whole animal | Rhein River near Mainz, Germany | Lieb |
| *Tritonia diomedia* | Gastropoda | Neural tissue | Friday Harbor, WA | Moroz, Katz |
| *Wirenia argentea* (454) | Neomeniomorpha | Several adults, starved two months | Bergen, Norway | Halanych |
| *Wirenia argentea* (Capillary) | Neomeniomorpha | Several adults, starved two months | Bergen, Norway | Todt |

## Supplementary Table 3 | Taxon sampling.

| OTU | Species | Data Type | Number of Reads | Number of Matrix Genes | Number of Amino Acids | Leaf Stability* | Data Source | NCBI Accession Number, Version Number, or Version Date |
|---|---|---|---|---|---|---|---|---|
| Alvinella | *Alvinella pompejana* | Sanger | 142,334 | 181 | 30,950 | 0.9919 | NCBI UniGene | March 31, 2009 |
| **Antalis** | ***Antalis vulgaris*** | **454** | **77,223** | **93** | **11,509** | **0.9662** | **This study** | **SRR108988.1** |
| Aplysia | *Aplysia californica* | Sanger | 250,102 | 210 | 37,730 | 0.9977 | NCBI UniGene | January 27, 2010 Version |
| Barentsia | *Barentsia elongata* | Sanger | 2,154 | 39 | 5,098 | | NCBI dbEST | FR837542 - FR837592 |
| Biomphalaria | *Biomphalaria glabrata* | 454 | 702,248 | 208 | 36,416 | 0.9977 | NCBI SRA | SRX000011, SRX001379, SRX001380, SRX014813, SRX014894-SRX014897 |
| Capitella | *Capitella teleta* | Genome | | 283 | 76,588 | 0.9924 | JGI | JGI Predicted Transcripts |
| Carinoma | *Carinoma mutabilis* | Sanger | 3,168 | 57 | 10,000 | | NCBI Trace | October 24, 2009 |
| Cerebratulus | *Cerebratulus lacteus* | Sanger | 6,144 | 65 | 11,463 | | NCBI Trace | October 21, 2009 |
| Chaetoderma | *Chaetoderma nitidulum* | Sanger | 1,632 | 36 | 5,417 | 0.9935 | NCBI Trace | October 21, 2009 |
| Chaetopterus | *Chaetopterus sp.* | Sanger | 3,360 | 80 | 14,311 | 0.9901 | NCBI Trace | October 24, 2009 |
| Chitonida | *Acanthopleura hirsuta* | Sanger | 498 | 61 | 8,731 | 0.9935 | NCBI dbEST | GO924863-GO943999 |
| | *Chaetopleura apiculata* | Sanger | 2,304 | | | | NCBI Trace | October 24, 2009 |
| | *Lepidochitona cinerea* | Sanger | 1,054 | | | | NCBI dbEST | FR836483 - FR837532 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Crassostrea | *Crassostrea gigas* | Sanger | 29,573 | 222 | 48,283 | 0.9976 | NCBI dbEST | AJ431681-AJ565846, AM237631-AM869575, BG467397-BQ427368, CB617326-CB617565, CD526814-CD526883, CF369125-CF369261, CK172301-CK172437, CU681473-CU999999, CX068761-CX069356, CX739338-CX739699, DV736295-DV736964, DW713815-DW714024, EE677412-EE677929, ES789087-ES789956, EW688558-EW779578, EX151492-EX151622, EX956364-EX956451, FC325715, FD483977-FD483996, FE192418-FE192425, FP000001-FP012228, FP089705-FP091223, GT052936-GT054201 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Crassostrea virginica* | Sanger | 14,560 | | | | NCBI dbEST | BG624106-BG624961, CD526707-CD650719, CK240390-CK240470, CV086962-CV172543, EH643873-EH649414 |
| Crepidula | *Crepidula fornicata* | 454 | 1,297,588 | 220 | 42,145 | 0.9978 | From authors | http://www.life.illinois.edu/henry/crepidula_databases.html |
| **Dolabrifera** | ***Dolabrifera dolabrifera*** | **454** | **371,556** | **113** | **15,718** | **0.9977** | **This study** | **SRR111921.2** |
| Dreissena | *Dreissena polymorpha* | Sanger | 998 | 41 | 7,555 | 0.9849 | NCBI dbEST | AJ517516-AJ517756, |
| | *Dreissena rostriformis* | Sanger | 3,429 | | | | NCBI dbEST | EY433616-EY437044 |
| Euprymna | *Euprymna scolopes* | Sanger | 35,420 | 204 | 37,219 | 0.9855 | NCBI Trace | DW251302-DW286722 |
| Haliotis | *Haliotis asinia* | Sanger | 8,335 | 151 | 18,847 | 0.9893 | NCBI dbEST | CF805554-CF805567, DY402832-DY403153, DW986183-DW986191, GD241824-GD272291, GT067284-GT067348, GT274080-GT277649 |
| | *Haliotis discus* | Sanger | 7,260 | | | | NCBI dbEST | CX725921-CX727313, DN856307-DN856389, |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | EG361618-EG363026, FE041029-FE042253 |
| | *Haliotis diversicolor* | Sanger | 7,100 | | | | NCBI dbEST | AY449735-AY449746, GT866281-GT873349, GW314901-GW314919 |
| **Hanleya** | ***Hanleya nagelfar*** | **454** | **149,253** | **120** | **17,403** | **0.9935** | **This study** | **SRR108987.1** |
| Helicoidea | *Helix aspersa* | Sanger | 216 | 40 | 7,141 | 0.9977 | NCBI dbEST | DR044213-DR044428 |
| | *Mandarina ponderosa* | Sanger | 312 | | | | NCBI dbEST | DR044429-DR044740 |
| | *Nesiohelix samarangae* | Sanger | 2,105 | | | | NCBI dbEST | DC603526-DC605630 |
| **Helisoma** | ***Helisoma trivolvis*** | **454** | **189,216** | **116** | **13,828** | **0.9977** | **This study** | **SRR108941.1** |
| **Hermissenda** | ***Hermissenda crassicornis*** | **454** | **88,881** | **69** | **7,961** | **0.9975** | **This study** | **SRR108974.1** |
| Hirudinea | *Helobdella robusta* | Genome | | 134 | 34,691 | 0.9924 | JGI | JGI Predicted Transcripts |
| | *Hirudo medicinalis* | Sanger | 26,833 | | | | NCBI Trace | EY478949-EY505781 |
| Hyriopsis | *Hyriopsis cumingii* | Sanger | 5,137 | 54 | 7,895 | 0.9986 | NCBI dbEST | EX828659-EX828681, FE968618-FE968692, FK026902-FK031940 |
| Idiosepius | *Idiosepius paradoxus* | Sanger | 9,079 | 107 | 20,461 | 0.9855 | NCBI dbEST | DB910977-DB920055 |
| Ilyanassa | *Ilyanassa obsoleta* | Sanger | 9,639 | 133 | 28,035 | 0.9978 | NCBI dbEST | EV825967-EV826048, FK710318-FK719874 |
| Laternula | *Laternula elliptica* | 454 | 1,033,858 | 139 | 17,238 | 0.9986 | NCBI SRA | SRX017389, SRX022359, SRX022360 |
| Littorina | *Littorina saxitilis* | 454 | 298,628 | 146 | 19,177 | 0.9978 | NCBI SRA | SRX023325, SRX023326 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Loligo | *Loligo bleekeri* | Sanger | 669 | 102 | 12,522 | 0.9855 | NCBI dbEST | FS372549-FS373217 |
| | ***Loligo opalescens*** | **454** | **3,258** | | | | **This study** | **SRR307161.2** |
| | ***Loligo pealei*** | **454** | **125,931** | | | | **This study** | **SRR108978.1** |
| | ***Loligo vulgaris*** | **454** | **26,015** | | | | **This study** | **SRR108981.1** |
| Lottia | *Lottia gigantea* | Genome | | 290 | 77,720 | 0.9933 | JGI | JGI Predicted Transcripts |
| Lymnaea | *Lymnaea stagnalis* | Sanger | 11,697 | 190 | 34,727 | 0.9977 | NCBI dbEST | CN809706-CN811025, ES291075-ES580561 |
| | ***Lymnaea stagnalis*** | **454** | **273,922** | | | | **This study** | **SRR108975.1** |
| **Melanoides** | ***Melanoides tuberculatus*** | **454** | **57,141** | **82** | **10,102** | **0.9978** | **This study** | **SRR108990.1** |
| Meretrix | *Meretrix meretrix* | Sanger | 2,111 | 60 | 12,279 | 0.9503 | NCBI dbEST | GR210953-GR212026, GR902434-GR903132, GT184089-GT184387 |
| Mytilus | *Mytilus californianus* | Sanger | 42,354 | 222 | 51,960 | 0.9973 | NCBI dbEST | ES387463.1-ES408175.1, ES735872.1-ES738966.1, FF339523.1-FF339585.1, GE747008.1-GE765490.1 |
| | *Mytilus galloprovincialis* | Sanger | 19,574 | | | | NCBI dbEST | AJ516092-AJ626468, EH662451-EH663597, FL488884-FL633565 |
| **Nautilus** | ***Nautilus pompilius*** | **454** | **549,720** | **152** | **17,264** | **0.9856** | **This study** | **SRR108979.1** |
| Nematostella | *Nematostella vectensis* | Genome | | 267 | 70,320 | | JGI | JGI Predicted Transcripts |
| **Neomenia** | ***Neomenia* sp.** | **454** | **126,484** | **66** | **6,083** | **0.9935** | **This study** | **SRR108985.1** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Nucula** | ***Nucula nitidosa*** | **454** | **75,202** | **97** | **12,282** | **0.9952** | **This study** | **SRR108989.1** |
| Octopus | *Octopus vulgaris* | Sanger | 16,432 | 218 | 33,349 | 0.9856 | From authors | http://www.cib.nig.ac.jp/dda/database/octopus.htm |
| | ***Octopus vulgaris*** | **454** | **882,605** | | | | **This study** | **SRR108980.1** |
| Oligochaeta | *Eisenia andrei* | Sanger | 2,400 | 156 | 34,650 | 0.9924 | NCBI dbEST | BB997898-BB999048, |
| | *Eisenia fetida* | Sanger | 3,935 | | | | NCBI dbEST | EH669363-EH672369, EL515444-EL515580, EY892395-EY893158, GO269559-GO269585, HO001170-HO001563, |
| | *Lumbricus rubellus* | Sanger | 19,934 | | | | From authors | http://xyala.cap.ed.ac.uk/Lumbribase/lumbribase_php/lumbribase.shtml |
| | *Tubifex tubifex* | Sanger | 17,014 | | | | NCBI dbEST | EY437148-EY454161 |
| Pectinidae | *Argopecten irradians* | Sanger | 7,057 | 118 | 22,150 | 0.9973 | NCBI dbEST | CB412266-CB417233, CF197421-CF197787, CK484086-CK484621, CN782333-CN783459, CV660837-CV660894, CV828452 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Argopecten purpuratus* | Sanger | 565 | | | | NCBI dbEST | ES469275-ES469694, FE895950-FE896091, FF147972-FF147974 |
| | *Chlamys farreri* | Sanger | 3,335 | | | | NCBI dbEST | DT716057-DT719391 |
| | *Mizuhopecten yessoensis* | Sanger | 3,011 | | | | NCBI dbEST | GH734852-GH736789, GR867007-GR868079, GT067737-GT067746, GT086406-GT087795, GT565072-GT570693 |
| | *Pecten maximus* | Sanger | 1,122 | | | | NCBI dbEST | DN793124-DN794245 |
| Pedicellina | *Pedicellina cernua* | Sanger | 5,184 | 62 | 12,755 | | NCBI Trace | October 24, 2009 |
| | *Pedicellina sp.* | Sanger | 2,668 | | | | NCBI Trace | October 4, 2009 |
| Pinctada | *Pinctada martensi* | Sanger | 7,130 | 93 | 17,760 | 0.9976 | NCBI dbEST | EY437147, FG396011, FG591193-FG598305, |
| | *Pinctada maxima* | Sanger | 7,096 | | | | NCBI dbEST | DV549057-DV549101, GH279961-GH738508 |
| **Pleurobranchae** | ***Pleurobranchaea*** | **454** | **255,718** | **96** | **11,488** | **0.9972** | **This study** | **SRR108976.1** |
| **Polyschides** | ***Polyschides dalli*** | **454** | **40,243** | **88** | **6,995** | **0.9662** | **This study** | **SRR108992.1** |
| **Scutopus** | ***Scutopus ventrolineatus*** | **Sanger** | **1,104** | **69** | **7,762** | **0.9935** | **This study** | **JG456490- JG456491** |
| | ***Scutopus ventrolineatus*** | **454** | **165,669** | | | | **This study** | **SRR108982.1** |
| Sinonovacula | *Sinonovacula constricta* | Sanger | 5,296 | 47 | 6,894 | 0.9754 | NCBI dbEST | GO308247-GO313553 |
| Sipuncula | *Sipunculus nudus* | Sanger | 2,207 | | | | NCBI dbEST | FR767771-FR770087 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Sipuncula | *Sipunculus nudus* | Sanger | 2,207 | 72 | 11,594 | 0.9901 | NCBI dbEST | FR767771-FR770087 |
| | *Themiste lageniformis* | Sanger | 2,640 | | | | NCBI Trace | October 21, 2009 |
| **Solemya** | ***Solemya velum*** | **454** | **67,786** | **149** | **19,683** | **0.9978** | **This study** | **SRR108983.1** |
| Strombus | *Strombus gigas* | 454 | 286,933 | 164 | 24,184 | 0.9978 | NCBI SRA | SRX017250 |
| Symbion | *Symbion pandora* | Sanger | 4,704 | 88 | 17,077 | | NCBI Trace | October 4, 2009 |
| Terebratalia | *Terebratalia transversa* | Sanger | 3,552 | 100 | 19,208 | | NCBI Trace | October 24, 2009 |
| **Theodoxus** | ***Theodoxus fluvatilis*** | **454** | **71,722** | **104** | **14,235** | **0.9942** | **This study** | **SRR108984.1** |
| **Tritonia** | ***Tritonia diomedia*** | **454** | **104,011** | **69** | **7,311** | **0.9975** | **This study** | **SRR108977.2** |
| Urechis | *Urechis caupo* | Sanger | 2,208 | 73 | 13,458 | 0.9924 | NCBI Trace | October 24, 2009 |
| Venerupis | *Venerupis decussatus* | Sanger | 4,645 | 112 | 19,543 | 0.9544 | NCBI dbEST | EY189760-EY255091, AM871090-AM871298, EL903765-EL903716 |
| | *Venerupis philippinarum* | Sanger | 5,657 | | | | NCBI dbEST | AM872010-AM877665 |
| **Wirenia** | ***Wirenia argentea*** | **Sanger** | **1,152** | **114** | **16,496** | **0.9935** | **This study** | **JG455978-JG454968** |
| | ***Wirenia argentea*** | **454** | **94,538** | | | | **This study** | **SRR108986.1** |

Taxa from which new data were collected are shown in blue. *Leaf stability scores are based on the annelid-rooted analysis shown in Fig 2. Taxa not included in Fig 2 do not have leaf stabilities reported in this table.

## Supplementary Table 4 | Genes selected for phylogenetic analyses.

| Gene ID* | Protein ID† | *Lottia* ID | Description | Len. | No. Tax. |
|---|---|---|---|---|---|
| 0007 | FBpp0072144 | 206392 | Eukaryotic translation initiation factor 6 | 239 | 10 |
| 0010 | FBpp0072801 | 186985 | 60S ribosomal protein L8 | 248 | 35 |
| 0011 | FBpp0085489 | 160203 | Succinate dehydrogenase iron-sulfur subunit, | 249 | 18 |
| 0015 | FBpp0075618 | 210271 | 40S ribosomal protein S4e | 235 | 32 |
| 0023 | FBpp0086063 | 203563 | Ribosome biogenesis protein RLP24 | 150 | 31 |
| 0026 | FBpp0087637 | 231795 | Nucleolar GTP-binding protein 1 | 578 | 20 |
| 0027 | FBpp0081401 | 213954 | T-complex protein 1 subunit eta | 511 | 13 |
| 0038 | FBpp0081947 | 157909 | Cytochrome c oxidase subunit 5A, mitochondrial | 122 | 25 |
| 0042 | FBpp0078227 | | Transmembrane protein 85 | 159 | 18 |
| 0044 | FBpp0077307 | 212802 | 40S ribosomal protein S21 | 81 | 26 |
| 0045 | FBpp0081305 | 206945 | Pre-mRNA-splicing factor ISY1 | 238 | 10 |
| 0048 | FBpp0080906 | 132027 | La ribonucleoprotein | 216 | 12 |
| 0053 | FBpp0085529 | 229894 | Ubiquitin-like protein 5 | 71 | 15 |
| 0061 | FBpp0078983 | 179268 | U1 small nuclear ribonucleoprotein 70 kDa | 194 | 12 |
| 0062 | FBpp0086599 | 126181 | Trafficking protein particle complex subunit 5 | 178 | 18 |
| 0063 | FBpp0081879 | 203563 | 60s ribosomal protein L24 | 155 | 33 |
| 0064 | FBpp0081185 | 235937 | Signal peptidase 18 kDa subunit | 172 | 10 |
| 0065 | FBpp0082571 | 200623 | Surfeit locus protein 4 | 250 | 14 |
| 0066 | FBpp0086387 | 65091 | Coiled-coil domain containing protein 94 | 225 | 15 |
| 0069 | FBpp0078894 | 210048 | PHD finger protein 5a | 111 | 20 |
| 0070 | FBpp0074532 | 199122 | N-alpha-acetyltransferase | 173 | 10 |
| 0072 | FBpp0083371 | 219397 | 40S ribosomal protein S20 | 105 | 29 |
| 0074 | FBpp0083244 | 185777 | RING finger protein 113A | 296 | 15 |
| 0077 | FBpp0082743 | 224262 | COP9 signalosome complex subunit 5 | 317 | 20 |
| 0080 | FBpp0078448 | 230810 | Proteasome subunit beta type-4 | 227 | 23 |
| 0081 | FBpp0082984 | 229754 | Ribosome production factor 2 | 292 | 12 |

| | | | | | |
|---|---|---|---|---|---|
| 0083 | FBpp0077979 | 93548 | DNAJ subfamily C member 2 | 477 | 15 |
| 0085 | FBpp0085314 | 226411 | 60S ribosomal protein L21 | 155 | 34 |
| 0087 | FBpp0075866 | 126569 | Serine/threonine-protein kinase RIO1 | 399 | 15 |
| 0090 | FBpp0075952 | 233453 | Ubiquitin fusion degradation protein 1 | 264 | 18 |
| 0091 | FBpp0087113 | 202957 | 40S ribosomal protein S11 | 135 | 35 |
| 0097 | FBpp0080121 | 197656 | DNA-directed RNA polymerase II subunit RPB3 | 253 | 17 |
| 0103 | FBpp0073686 | 236339 | Ubiquitin-conjugating enzyme E2 N | 144 | 18 |
| 0104 | FBpp0076602 | 70408 | 60S ribosomal protein L18 | 162 | 30 |
| 0110 | FBpp0082062 | 189380 | Proteasome 25 kDa subunit alpha type-2 | 220 | 29 |
| 0111 | FBpp0078134 | 237709 | 60S acidic ribosomal protein P0 | 242 | 34 |
| 0113 | FBpp0085379 | 111616 | Glutamate-rich WD repeat-containing protein 1 | 369 | 12 |
| 0119 | FBpp0084275 | | Isocitrate dehydrogenase subunit gamma 1 | 323 | 11 |
| 0122 | FBpp0073099 | 228476 | Zinc finger matrin-type protein 2 | 178 | 16 |
| 0125 | FBpp0080120 | 220690 | V-type proton ATPase subunit G | 111 | 20 |
| 0128 | FBpp0087347 | 134010 | H/ACA ribonucleoprotein complex subunit 3 | 56 | 21 |
| 0134 | FBpp0071846 | 110080 | 40S ribosomal protein S24 | 125 | 37 |
| 0136 | FBpp0076207 | 204040 | 40S ribosomal protein S17 | 113 | 33 |
| 0139 | FBpp0078086 | 109608 | DNA-directed RNA polymerase II polypeptide H | 148 | 14 |
| 0141 | FBpp0075257 | | Elongation Factor 1 Homolog | 82 | 13 |
| 0142 | FBpp0086654 | 149167 | Receptor expression-enhancing protein 5 | 172 | 19 |
| 0144 | FBpp0075246 | 185481 | Trafficking protein particle complex subunit 2 | 131 | 18 |
| 0149 | FBpp0084067 | 166689 | DNA-directed RNA polymerases I, II, and III subunit | 68 | 20 |
| 0156 | FBpp0088191 | 54396 | UV excision repair protein RAD23 | 306 | 14 |
| 0158 | FBpp0075990 | 197575 | Acireductone dioxygenase | 159 | 16 |
| 0162 | FBpp0076078 | 236022 | N-alpha-acetyltransferase | 167 | 14 |
| 0167 | FBpp0084959 | | 60S ribosomal protein L32 | 125 | 36 |
| 0168 | FBpp0087463 | 225615 | Alpha N-terminal protein methyltransferase 1A | 198 | 13 |
| 0170 | FBpp0074184 | 196504 | Actin-related protein 2/3 complex subunit 3 | 172 | 13 |
| 0172 | FBpp0081443 | 217311 | UPF0160 protein MYG1 | 328 | 16 |
| 0173 | FBpp0077912 | | Hydroxyacylglutathione hydrolase | 238 | 10 |
| 0182 | FBpp008653 | | Ras-like GTP-binding protein RHO | 170 | 22 |

| | | | | | |
|---|---|---|---|---|---|
| 0185 | FBpp0079445 | 202488 | Pescadillo | 460 | 14 |
| 0191 | FBpp0070943 | 207717 | 60S ribosomal protein L17 | 150 | 32 |
| 0193 | FBpp0077351 | 190601 | Seryl-tRNA synthetase | 493 | 15 |
| 0195 | FBpp0087095 | 150024 | T-complex protein 1 subunit epsilon | 534 | 17 |
| 0197 | FBpp0072072 | 207121 | Adenylate kinase 2, mitochondrial | 220 | 13 |
| 0200 | Fbpp0071052 | 209986 | 40S ribosomal protein S14 | 151 | 32 |
| 0203 | FBpp0073098 | 205433 | Ribosome biogenesis protein BRX1 | 236 | 17 |
| 0204 | FBpp0084759 | 192615 | Phosphoserine aminotransferase | 319 | 16 |
| 0205 | FBpp0082947 | 210506 | Beta-mannosyltransferase | 353 | 12 |
| 0206 | FBpp0074180 | 109561 | 40S ribosomal protein S5 | 193 | 32 |
| 0208 | FBpp0083801 | 150656 | Lipoyl synthase | 314 | 16 |
| 0212 | FBpp0083801 | 207505 | SEC13 | 294 | 14 |
| 0213 | FBpp0083802 | 223917 | 40S ribosomal protein S3 | 234 | 35 |
| 0215 | FBpp0072197 | 191571 | 26S proteasome non-ATPase regulatory subunit 7 | 309 | 22 |
| 0216 | FBpp0080854 | 238978 | Fructose-1,6-bisphosphatase 1 | 299 | 14 |
| 0217 | FBpp0086603 | 237412 | 26S proteasome non-ATPase regulatory subunit 11 | 408 | 15 |
| 0218 | FBpp0079187 | 184120 | Guanine nucleotide-binding protein subunit beta-2- | 258 | 31 |
| 0229 | FBpp0081437 | 200562 | Succinyl-CoA ligase subunit beta | 394 | 18 |
| 0230 | FBpp0080997 | 229010 | ATP-dependent RNA helicase DDX47 | 411 | 15 |
| 0231 | FBpp0085104 | 205749 | Casein kinase I delta | 318 | 11 |
| 0235 | FBpp0081767 | | Alcohol dehydrogenase class-3 | 373 | 16 |
| 0236 | FBpp0099974 | 219170 | Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit alpha | 570 | 15 |
| 0242 | FBpp0087760 | 217535 | Glucose-6-phosphate isomerase | 513 | 14 |
| 0243 | FBpp0083906 | 197242 | 26S protease regulatory subunit 4 | 437 | 20 |
| 0244 | FBpp0074756 | 203071 | RuvB-like helicase 2 | 452 | 18 |
| 0245 | FBpp0073290 | 237408 | 60 kDa heat shock protein, mitochondrial | 541 | 16 |
| 0246 | FBpp0082788 | 204318 | T-complex protein 1 subunit gamma | 514 | 21 |
| 0254 | FBpp0077066 | 211912 | NFU1 iron-sulfur cluster scaffold, mitochondrial | 202 | 12 |
| 0255 | FBpp0079233 | 239089 | Tumor suressor candidate 3 | 302 | 12 |
| 0262 | FBpp0082412 | 97481 | 39S ribosomal protein L11, mitochondrial | 175 | 16 |

| | | | | | |
|---|---|---|---|---|---|
| 0265 | FBpp0086895 | | Transcription factor BTF3 homolog 4 | 147 | 23 |
| 0268 | FBpp0080829 | 156082 | RNA-binding motif protein 13 | 147 | 12 |
| 0276 | FBpp0075307 | 232500 | Peptide methionine sulfoxide reductase | 219 | 13 |
| 0277 | FBpp0073844 | 232303 | 60S ribosomal protein L37e | 91 | 27 |
| 0283 | FBpp0086072 | 216998 | S-methyl-5'-thioadenosine phosphorylase | 266 | 12 |
| 0284 | FBpp0087704 | 203722 | Charged multivesicular body protein 4b | 200 | 18 |
| 0286 | FBpp0087608 | | 60S ribosomal protein L31 | 111 | 35 |
| 0288 | FBpp0075838 | 109978 | DNA-directed RNA polymerase III 25 kDa | 196 | 14 |
| 0290 | FBpp0080044 | 238952 | Glutaredoxin 3 | 250 | 19 |
| 0292 | FBpp0070832 | 219222 | Translation machinery-associated protein 3 | 171 | 23 |
| 0293 | FBpp0074516 | 236282 | RNA polymerase II subunit A C-terminal domain | 185 | 11 |
| 0295 | FBpp0086269 | 149778 | 40S ribosomal protein S15 | 145 | 39 |
| 0298 | FBpp0076652 | 206284 | Ribosome maturation protein SBDS | 239 | 15 |
| 0299 | FBpp0075700 | 233830 | Eukaryotic translation initiation factor 2 subunit 2 | 249 | 17 |
| 0300 | FBpp0086226 | 210661 | Superoxide dismutase, mitochondrial | 194 | 21 |
| 0310 | FBpp0085619 | 185272 | Proliferating cell nuclear antigen | 251 | 25 |
| 0316 | FBpp0077549 | 234041 | F-actin-capping protein subunit beta | 267 | 18 |
| 0317 | FBpp0072312 | 195818 | 60S ribosomal protein L19 | 196 | 40 |
| 0326 | FBpp0075595 | 216644 | GPN-loop GTPase 2 | 284 | 12 |
| 0333 | FBpp0087972 | 173095 | Cathepsin D | 339 | 23 |
| 0335 | FBpp0078689 | 200332 | Eukaryotic translation initiation factor 3 subunit I | 291 | 32 |
| 0340 | FBpp0072050 | 142681 | Transaldolase | 315 | 15 |
| 0341 | FBpp0073459 | 234495 | Pleiotropic regulator 1 | 405 | 13 |
| 0345 | FBpp0086223 | 114414 | FLAP endonuclease-1 | 371 | 15 |
| 0346 | FBpp0075861 | 153055 | Diphthamide biosynthesis protein 1 | 369 | 13 |
| 0347 | FBpp0086790 | 178160 | Elongation factor Tu | 376 | 20 |
| 0353 | FBpp0074278 | 114287 | Vacuolar protein sorting-associated protein 4 | 424 | 14 |
| 0354 | FBpp0081822 | 196203 | 60S ribosomal protein L3 | 364 | 35 |
| 0361 | FBpp0075386 | | Glycyl-tRNA synthetase | 646 | 13 |
| 0368 | FBpp0073017 | 232335 | Synaptic glycoprotein SC2 | 287 | 16 |
| 0370 | FBpp0070143 | 178960 | 60S ribosomal protein L22 | 116 | 30 |

| | | | | | |
|---|---|---|---|---|---|
| 0373 | FBpp0086795 | 114637 | GrpE 1 | 183 | 15 |
| 0376 | FBpp0083533 | 234160 | Vacuolar sorting protein SNF8 | 236 | 14 |
| 0377 | FBpp0075612 | 150191 | 40S ribosomal protein S12 | 128 | 31 |
| 0381 | FBpp0087511 | 56656 | mRNA turnover protein 4 | 207 | 25 |
| 0382 | FBpp0077740 | 102125 | Peptidase A22B, minor histocompatibility antigen | 262 | 14 |
| 0388 | FBpp0071047 | 134665 | Estradiol 17-beta-dehydrogenase | 270 | 12 |
| 0389 | FBpp0080724 | 236815 | 60S ribosomal protein L30 | 105 | 24 |
| 0392 | FBpp0070969 | 216779 | Coiled-coil domain-containing protein 25 | 173 | 15 |
| 0397 | FBpp0078463 | 132223 | Enolase-phosphatase E1 | 232 | 18 |
| 0399 | FBpp0079711 | 235879 | Transmembrane protein 111 | 251 | 16 |
| 0402 | FBpp0071427 | 182626 | Gamma-aminobutyric acid receptor-associated | 116 | 29 |
| 0405 | FBpp0075448 | 211297 | H/ACA ribonucleoprotein complex subunit 2 | 146 | 19 |
| 0406 | FBpp0077625 | 181759 | Vacuolar protein sorting-associated protein 29 | 173 | 11 |
| 0410 | FBpp0079812 | 182182 | Replication factor C 38kD subunit | 351 | 14 |
| 0412 | FBpp0083854 | 165837 | Oligoribonuclease | 175 | 12 |
| 0413 | FBpp0070924 | 154620 | Ubiquinone biosynthesis protein COQ7 | 191 | 18 |
| 0416 | FBpp0084434 | 156627 | Histone H2A.v | 127 | 29 |
| 0425 | FBpp0074129 | 172374 | Exosome complex exonuclease RRP45 | 261 | 11 |
| 0426 | FBpp0083457 | 171636 | Isopentenyl-diphosphate Delta-isomerase 1 | 216 | 12 |
| 0432 | FBpp0087186 | 224100 | Electron transfer flavoprotein subunit, mitochondrial | 319 | 17 |
| 0435 | FBpp0080020 | 177471 | Ethanolamine-phosphate cytidylyltransferase | 327 | 17 |
| 0440 | FBpp0088441 | 190901 | 40S ribosomal protein S7 | 177 | 30 |
| 0445 | FBpp0088818 | 155494 | Trafficking protein particle complex subunit 3 | 178 | 16 |
| 0449 | FBpp0079752 | 217766 | 60S ribosomal protein L9 | 172 | 39 |
| 0456 | FBpp0071808 | 128584 | 60S ribosomal protein L23 | 120 | 29 |
| 0457 | FBpp0084874 | 187625 | Eukaryotic initiation factor eIF-2B alpha subunit | 271 | 15 |
| 0458 | FBpp0083245 | 120210 | S-formylglutathione hydrolase | 270 | 20 |
| 0459 | FBpp0072833 | | Purine nucleoside phosphorylase | 272 | 10 |
| 0462 | FBpp0077029 | 205560 | Rab-10 | 198 | 14 |
| 0464 | FBpp0070723 | 208106 | Cyclin-dependent kinase 7 | 310 | 15 |
| 0471 | FBpp0085222 | 189149 | Deoxyhypusine hydroxylase | 279 | 16 |

| | | | | | |
|---|---|---|---|---|---|
| 0474 | FBpp0071198 | 184506 | RNA-binding motif protein 13 | 260 | 15 |
| 0477 | FBpp0080147 | 167800 | Actin-related protein 2/3 complex subunit 1A | 307 | 28 |
| 0481 | FBpp0084144 | 201859 | U3 small nucleolar ribonucleoprotein protein IMP4 | 283 | 12 |
| 0482 | FBpp0089041 | 212487 | Proteasome subunit alpha type-3 | 233 | 27 |
| 0483 | FBpp0076990 | 205824 | RNA-binding protein PNO1 | 203 | 28 |
| 0485 | FBpp0073947 | 132224 | Eukaryotic translation initiation factor 5 | 362 | 19 |
| 0489 | FBpp0073921 | 236462 | guanine nucleotide-binding protein subunit beta | 326 | 19 |
| 0492 | FBpp0084948 | 238326 | Triosephosphate isomerase | 216 | 26 |
| 0496 | FBpp0099686 | 239290 | 40S ribosomal protein S8 | 206 | 40 |
| 0499 | FBpp0080708 | 237446 | Prohibitin | 268 | 22 |
| 0505 | FBpp0072250 | 190068 | Inorganic pyrophosphatase | 274 | 14 |
| 0518 | FBpp0078278 | 212047 | 26S proteasome non-ATPase regulatory subunit 12 | 424 | 18 |
| 0520 | FBpp0081834 | 177527 | Dihydrolipoyllysine-residue succinyltransferase | 293 | 18 |
| 0521 | FBpp0086370 | | Glutamate oxaloacetate transaminase 1 | 387 | 17 |
| 0525 | FBpp0070333 | 115714 | Legumain | 377 | 12 |
| 0534 | FBpp0073120 | | Replication factor C subunit 2 | 327 | 17 |
| 0553 | FBpp0081592 | 205544 | Clathrin-associated adaptor complex AP-1 mu-1 | 418 | 18 |
| 0555 | FBpp0087764 | 189106 | T-complex protein 1 subunit theta | 538 | 24 |
| 0558 | FBpp0085265 | 199050 | Elongation factor 2 | 819 | 17 |
| 0566 | FBpp0078806 | 207101 | Initiation factor 4A | 390 | 22 |
| 0569 | FBpp0074906 | 219043 | Dihydrolipoyl dehydrogenase | 456 | 17 |
| 0571 | FBpp0070890 | 132169 | 26s protease regulatory subunit 10B | 389 | 15 |
| 0572 | FBpp0080705 | 200884 | Asparaginyl-tRNA synthetase | 497 | 22 |
| 0582 | FBpp0088021 | 203656 | 26S proteasome regulatory subunit 7 | 430 | 17 |
| 0583 | FBpp0071794 | 206617 | ATP synthase subunit alpha, mitochondrial | 530 | 32 |
| 0585 | FBpp0082140 | 119755 | V-type proton ATPase subunit B | 479 | 14 |
| 0586 | FBpp0085737 | 221335 | Succinate dehydrogenase flavoprotein subunit B, | 646 | 12 |
| 0588 | FBpp0073446 | 216416 | 78 kDa glucose-regulated protein | 613 | 14 |
| 0600 | FBpp0082110 | 229207 | Sorting nexin 12 | 157 | 19 |
| 0610 | FBpp0081787 | 182398 | Mediator of RNA polymerase II transcription subunit | 200 | 14 |
| 0612 | FBpp0083376 | 215959 | 40S ribosomal protein S30 | 121 | 37 |

| | | | | | |
|---|---|---|---|---|---|
| 0620 | FBpp0082522 | 183079 | ATP synthase subunit O, mitochondrial | 191 | 28 |
| 0621 | FBpp0082511 | 205756 | Histone-binding protein RBBP4 | 408 | 17 |
| 0624 | FBpp0083949 | | Signal peptidase complex subunit 3 | 170 | 13 |
| 0630 | FBpp0076238 | 156500 | 39S ribosomal protein L12, mitochondrial | 162 | 17 |
| 0637 | FBpp0073806 | 201172 | Glutaredoxin-related protein 5 | 130 | 15 |
| 0642 | FBpp0075967 | 218864 | Prefoldin subunit 2 | 127 | 21 |
| 0645 | FBpp0079746 | 196086 | Deoxyuridine 5'-triphosphate nucleotidohydrolase | 145 | 16 |
| 0646 | FBpp0078350 | 209321 | V-type proton ATPase subunit E | 207 | 23 |
| 0658 | FBpp0081882 | 202995 | Prefoldin subunit 3 | 174 | 21 |
| 0660 | FBpp0081528 | 201223 | 60S ribosomal protein L34b | 108 | 27 |
| 0662 | FBpp0080011 | 233682 | Golgi membrane protein YIPF4 | 209 | 15 |
| 0668 | FBpp0071223 | 98069 | Activator of basal transcription 1 | 186 | 13 |
| 0671 | FBpp0080446 | 219589 | Cytochrome c-2 | 104 | 29 |
| 0672 | FBpp0076607 | 185379 | Surfeit locus protein 1 | 207 | 14 |
| 0676 | FBpp0080890 | 220774 | Actin-related protein 2/3 complex subunit 2 | 291 | 15 |
| 0680 | FBpp0085166 | 207552 | 60S ribosomal protein L6 | 202 | 36 |
| 0682 | FBpp0085615 | 222194 | Phosphate carrier protein, mitochondrial | 329 | 16 |
| 0690 | FBpp0073316 | 204936 | Protein BUD31 homolog | 148 | 20 |
| 0691 | FBpp0075395 | 222708 | U4/U6 small nuclear ribonucleoprotein Prp31 | 458 | 12 |
| 0697 | FBpp0088242 | 202499 | 40S ribosomal protein S3a | 241 | 40 |
| 0699 | FBpp0077414 | 234443 | Carnitine acylcarnitine carrier protein, mitochondrial | 288 | 14 |
| 0701 | FBpp0079650 | 227857 | Ubiquitin carboxyl-terminal hydrolase 14 | 442 | 17 |
| 0702 | FBpp0086318 | 206255 | V-type proton ATPase subunit C 1-A | 350 | 12 |
| 0703 | FBpp0086096 | 101919 | U3 small nucleolar ribonucleoprotein protein IMP3 | 175 | 15 |
| 0705 | FBpp0075401 | 184255 | Protein disulfide-isomerase | 404 | 27 |
| 0706 | FBpp0078024 | 115372 | 26S proteasome non-ATPase regulatory subunit 4 | 333 | 23 |
| 0707 | FBpp0085223 | 205662 | Nucleoside diphosphate kinase | 145 | 32 |
| 0709 | FBpp0072020 | 209108 | Exosome complex exonuclease RRP4 | 270 | 13 |
| 0710 | FBpp0075693 | 171717 | Phosphomannomutase 2 | 245 | 13 |
| 0720 | FBpp0083687 | 200903 | 26S proteasome non-ATPase regulatory subunit 6 | 385 | 19 |
| 0722 | FBpp0079943 | 230007 | Ras-related protein Rab-6.1 (Rab-6A) | 196 | 14 |

| | | | | | |
|---|---|---|---|---|---|
| 0723 | FBpp0087958 | 222316 | Mitochondrial-processing peptidase subunit alpha | 456 | 13 |
| 0724 | FBpp0079606 | | Ubiquitin-40S ribosomal protein S27a | 159 | 41 |
| 0725 | FBpp0071600 | 216798 | mRNA export factor RAE1 | 351 | 14 |
| 0727 | FBpp0083673 | 213693 | Diphthine synthase | 256 | 14 |
| 0730 | FBpp0072151 | 207546 | FACT complex subunit SSRP1 | 488 | 13 |
| 0731 | FBpp0078633 | 196653 | DNA-directed RNA polymerases I and III subunit | 323 | 19 |
| 0732 | FBpp0088029 | 203916 | ATPase ASNA1 | 328 | 14 |
| 0733 | FBpp0082985 | 225558 | Malate dehydrogenase | 305 | 26 |
| 0734 | FBpp0071089 | 66846 | 40S Ribosomal protein S6 | 236 | 39 |
| 0735 | FBpp0082832 | 221202 | Ribosomal RNA methyltransferase CG5220 | 305 | 17 |
| 0736 | FBpp0081491 | 227198 | Ras-related GTP-binding protein A | 297 | 15 |
| 0740 | FBpp0110423 | 224562 | 60S ribosomal protein L5 | 274 | 42 |
| 0749 | FBpp0084617 | 219559 | 60S ribosomal protein L4 | 332 | 40 |
| 0752 | FBpp0088408 | 151663 | Obg-like ATPase 1 | 375 | 18 |
| 0753 | FBpp0073344 | | Glutamine synthetase 2, cytoplasmic | 331 | 22 |
| 0758 | FBpp0085919 | | Polyadenylate-binding protein | 512 | 30 |
| 0760 | FBpp0076460 | 193923 | Actin-related protein 3 | 416 | 15 |
| 0761 | FBpp0083611 | 199626 | Pyruvate kinase | 505 | 16 |
| 0767 | FBpp0073292 | 212332 | 26S proteasome regulatory subunit 6B | 400 | 23 |
| 0772 | FBpp0076649 | 110000 | DEAD box ATP-dependent RNA helicase | 291 | 18 |
| 0774 | FBpp0074662 | 138721 | 26S protease regulatory subunit 8 | 387 | 20 |
| 0781 | FBpp0077368 | 225601 | Peroxiredoxin 6 | 208 | 27 |
| 0783 | FBpp0076185 | 164153 | Phosphoserine phosphatase | 217 | 11 |
| 0784 | FBpp0078655 | 225011 | 60S ribosomal protein L37a | 90 | 23 |
| 0786 | FBpp0073853 | 165683 | UDP-N-acetylglucosamine transferase subunit | 173 | 11 |
| 0787 | FBpp0084929 | | Mitochondrial intermembrane space import and | 110 | 13 |
| 0788 | FBpp0071373 | 158030 | ATP synthase delta, mitochondrial | 154 | 32 |
| 0796 | FBpp0070150 | 192827 | 60S ribosomal protein L36 | 104 | 36 |
| 0800 | FBpp0074964 | 212293 | Charged multivesicular body protein 5 | 205 | 19 |
| 0804 | FBpp0074088 | 235941 | 40S ribosomal protein S19 | 144 | 34 |
| 0805 | FBpp0079642 | 196756 | Olygosaccharyltransferase alpha subunit | 586 | 16 |

| | | | | | |
|---|---|---|---|---|---|
| 0806 | FBpp0082384 | 132099 | DNA-directed RNA polymerase II subunit RPB9 | 126 | 15 |
| 0807 | FBpp0075151 | 196960 | Multiprotein bridging factor 1 | 143 | 19 |
| 0809 | FBpp0078400 | 195680 | Splicing factor 3A subunit 3 | 479 | 14 |
| 0810 | FBpp0071461 | 196809 | F-actin-capping protein subunit alpha | 277 | 15 |
| 0812 | FBpp0078598 | 214378 | Vacuolar sorting protein VPS24 | 206 | 13 |
| 0824 | FBpp0070879 | 202003 | 60S ribosomal protein L7a | 258 | 38 |
| 0826 | FBpp0076124 | 223907 | Ubiquitin-conjugating enzyme E2-22 kDa | 193 | 13 |
| 0829 | FBpp0110169 | 214125 | NADH-cytochrome b5 reductase 3 | 290 | 18 |
| 0838 | FBpp0084905 | 235900 | ATP synthase subunit gamma, mitochondrial | 254 | 31 |
| 0839 | FBpp0083861 | 214609 | 26S proteasome non-ATPase regulatory subunit 13 | 362 | 19 |
| 0842 | FBpp0082464 | 223715 | V-type proton ATPase 21 kDa proteolipid subunit | 206 | 24 |
| 0843 | FBpp0073626 | 186221 | 40S ribosomal protein S15Aa | 130 | 35 |
| 0845 | FBpp0075382 | 231140 | Proteasome subunit beta type-7 | 260 | 20 |
| 0848 | FBpp0070434 | 233963 | V-type proton ATPase subunit D | 226 | 19 |
| 0851 | FBpp0086400 | 153858 | Proteasome subunit beta type-6 | 220 | 27 |
| 0852 | FBpp0085717 | 167500 | 60S ribosomal protein L11 | 173 | 37 |
| 0854 | FBpp0087806 | 227129 | Zinc finger CCCH domain-containing protein 15 | 286 | 16 |
| 0855 | FBpp0081283 | 182042 | Uncharacterized methyltransferase WBSCR22 | 251 | 23 |
| 0858 | FBpp0078896 | 194715 | Protein transport protein Sec61 subunit alpha | 469 | 20 |
| 0861 | FBpp0086066 | 232029 | Proteasome subunit alpha type-5 | 237 | 24 |
| 0867 | FBpp0070584 | 108853 | V-type proton ATPase subunit d 1 | 339 | 15 |
| 0868 | FBpp0081780 | 177209 | Arginine N-methyltransferase 1 | 329 | 19 |
| 0869 | FBpp0082459 | 105568 | Mitochondrial-processing peptidase subunit beta | 425 | 26 |
| 0881 | FBpp0081460 | 124007 | Tryptophanyl-tRNA synthetase | 397 | 16 |
| 0885 | FBpp0083395 | 220834 | Adenylosuccinate synthetase | 421 | 13 |
| 0887 | FBpp0084735 | 226825 | Pyruvate dehydrogenase E1 component subunit | 338 | 15 |
| 0888 | FBpp0074825 | 236064 | Catalase | 450 | 14 |
| 0891 | FBpp0083843 | 207066 | 26s protease regulatory subunit 6a | 422 | 13 |
| 0893 | FBpp0078997 | 144966 | Nucleolar protein 58 | 449 | 17 |
| 0900 | FBpp0073902 | 231565 | T-complex protein 1 subunit zeta | 515 | 18 |
| 0901 | FBpp0079992 | 199820 | T-complex protein 1 subunit delta | 512 | 16 |

| | | | | | |
|---|---|---|---|---|---|
| 0902 | FBpp0073847 | 184532 | Adenosylhomocysteinase | 416 | 19 |
| 0903 | FBpp0081704 | 201543 | RuvB-like helicase 1 (Pontin) | 452 | 17 |
| 0914 | FBpp0085966 | 140318 | Thymidylate kinase | 189 | 12 |
| 0915 | FBpp0072660 | 223843 | Hsp90 co-chaperone Cdc37 | 349 | 13 |
| 0919 | FBpp0083893 | 207726 | U6 snRNA-associated Sm-like protein LSm3 | 98 | 19 |
| 0923 | FBpp0077106 | 201479 | Protein geranyl transferase type-1 subunit alpha | 324 | 15 |
| 0928 | FBpp0071295 | 150772 | 40S ribosomal protein S28 | 65 | 29 |
| 0933 | FBpp0086474 | 205831 | V-type proton ATPase subunit F | 120 | 25 |
| 0934 | FBpp0080817 | 210079 | Dolichol-phosphate mannosyltransferase | 235 | 13 |
| 0937 | FBpp0070643 | 203874 | Transmembrane emp24 domain-containing | 184 | 13 |
| 0938 | FBpp0078009 | 114503 | UPF0368 protein Cxorf26 | 159 | 13 |
| 0943 | FBpp0086973 | 197780 | Nascent polypeptide-associated complex subunit | 169 | 36 |
| 0946 | FBpp0075111 | 156936 | Nonclathrin coat protein zeta | 176 | 15 |
| 0947 | FBpp0087901 | 228040 | Vacuolar protein sorting-associated protein 28 | 207 | 24 |
| 0953 | FBpp0080062 | 237076 | Exosome complex exonuclease RRP41 | 231 | 15 |
| 0954 | FBpp0075999 | 187336 | UPF0195 protein CG7949 | 149 | 16 |
| 0955 | FBpp0071279 | 197848 | Dolichyl-diphosphooligosaccharide--protein | 413 | 20 |
| 0964 | FBpp0071049 | 206537 | Synaptobrevin homolog YKT6 | 191 | 15 |
| 0965 | FBpp0086103 | 207423 | 60S ribosomal protein L18a | 174 | 37 |
| 0966 | FBpp0077750 | 200807 | Ribonuclease H2 subunit A | 281 | 12 |
| 0972 | FBpp0070047 | 229535 | 60S ribosomal protein L10 | 210 | 39 |
| 0978 | FBpp0084901 | 203788 | Electron transfer flavoprotein subunit beta | 238 | 21 |
| 0979 | FBpp0089135 | 225993 | Casein kinase II beta subunit | 215 | 22 |
| 0982 | FBpp0070871 | 170380 | Citrate synthase, mitochondrial | 418 | 23 |
| 0985 | FBpp0086701 | 202410 | 40S ribosomal protein S23 | 139 | 38 |
| 0987 | FBpp0072128 | 103941 | Nucleosome assembly protein 1 | 339 | 24 |
| 0988 | FBpp0081488 | 226791 | Proteasome subunit beta type-3 | 201 | 28 |
| 0991 | FBpp0072696 | 117504 | Methyltransferase-like protein | 268 | 18 |
| 0992 | FBpp0080691 | 207043 | 26S proteasome non-ATPase regulatory subunit 3 | 433 | 14 |
| 0993 | FBpp0075766 | 230263 | 60S ribosomal protein L10a | 209 | 41 |
| 1003 | FBpp0079447 | 190501 | cAMP-dependent protein kinase catalytic subunit | 343 | 14 |

| | | | | | |
|---|---|---|---|---|---|
| 1005 | FBpp0072904 | 161608 | Heat shock protein 90 alpha | 690 | 17 |
| 1007 | FBpp0081606 | | Serine/threonine-protein phosphatase 5 | 443 | 11 |
| 1013 | FBpp0083213 | 184295 | Developmentally-regulated GTP-binding protein 2 | 362 | 12 |
| 1030 | FBpp0088250 | 201878 | ATP synthase subunit beta, mitochondrial | 488 | 28 |

*Gene ID corresponds to the HaMStR model organisms core orthologs dataset numbering system. †Protein ID corresponds to the *Drosophila melanogaster* (HaMStR primer taxon) protein identification number used in the InParanoid database. Len, length of final Alicut/Aliscore[36] trimmed alignment; No. tax., number of taxa sampled. Data matrices including information on the best-fitting model used for each gene for all analyses are available from TreeBase under accession number S11762.

**Supplementary Table 5 | Approximately Unbiased (AU) test results.**

| Alternative Hypothesis | Ln Likelihood Score | P-value | Significantly Worse? |
|---|---|---|---|
| Best tree | -1048338.79 | | |
| Adenopoda (Chaetodermomorpha basal) | -1048458.37 | 0.001 | Yes |
| Hepagastralia (Neomeniomorpha basal) | -1048441.52 | 2e-04 | Yes |
| Testaria (Polyplacophora + Conchifera) | -1048370.95 | 0.015 | Yes |
| Diasoma (Bivalvia + Scaphopoda) | -1048350.63 | 0.626 | No |
| Cyrtosoma (Gastropoda + Cephalopoda) | -1048527.87 | 2e-08 | Yes |
| Cephalopoda + Scaphopoda | -1048363.12 | 0.416 | No |
| Gastropoda + Scaphopoda | -1048355.35 | 0.508 | No |

AU tests were performed using the matrix of all 308 genes with Annelida as the outgroup corresponding to the tree in Fig. 2. The best fitting AA substitution model for each gene was used in likelihood calculations in RAxML.

## Supplementary Table 6 | Clade support assessment across all analyses

| Tree: | Fig 2<br>All 308 genes | Figs S3-S4<br>All 308 genes broad outgroups | Fig S6<br>200 best genes | Fig S7<br>100 best genes | Fig S8<br>Non-ribosomal genes | Fig S9<br>Ribosomal genes | Fig S10<br>InParanoid / HaMStR genes | Fig S11 (bs support only)<br>All 308 genes WAG + CAT + F | Fig S12 (bs support only)<br>All 308 genes LG + CAT + F | Fig S13 (SH-like support)<br>All 308 genes FastTree | Fig S14-S15<br>No *Nematostella* | Figure S16a (bs support only)<br>All 308 genes | Figure S16b (bs support only)<br>All 308 genes |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Outgroup root: | A | N | A | A | A | A | A | A | A | A | E + C | E + C | E |
| Matches Fig 2 general topology? | Yes | Yes | Yes | Yes | No | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes |
| Aculifera | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 92 / 1.00 | 88 / 1.00 | 100 / 1.00 | 100 | 100 | 1.00 | 100 / 1.00 | 94 | 94 |
| Aplacophora | 100 / 0.99 | 100 / 0.99 | 100 / 0.99 | 100 / 0.99 | NR | 100 / 1.00 | 96 / 0.98 | 100 | 100 | 1.00 | 100 / 1.00 | 94 | 94 |
| Neomenio-morpha | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 97 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 | 100 | 1.00 | 100 / 1.00 | 100 | 100 |
| Chaetodermo-morpha | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 97 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 | 100 | 1.00 | 100 / 1.00 | 100 | 100 |
| Polyplacophora | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 99 / NR | 100 / 1.00 | 100 / 1.00 | 100 | 100 | 1.00 | 100 / 1.00 | 100 | 100 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Conchifera | 100 / 0.98 | 79 / 0.97 | 58 / 0.97 | 76 / 0.99 | 39 / NR | - / 0.97 | 50 / 0.96 | 78 | 95 | 0.85 | 100 / 1.00 | NR | 78 |
| Pleistomollusca | 100 / 1.00 | - / 1.00 | 88 / 1.00 | 93 / 1.00 | NR / NR | 80 / 0.99 | - / 0.99 | 85 | 97 | 0.98 | 46 / 1.00 | 98 | 97 |
| Gastropoda | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 | 100 | | 100 / 1.00 | 100 | 100 |
| Bivalvia | 100 / 1.00 | 69 / - | 99 / 1.00 | 100 / 1.00 | NR / NR | 99 / 1.00 | 52 / 1.00 | 98 | 100 | 1.00 | 65 / 1.00 | 100 | 100 |
| Scaphopoda + Pleistomollusca | 72 / 0.98 | 93 / 0.88 | 83 / 0.98 | 93 / 0.95 | 81 / 0.90 | 81 / 0.97 | 50 / 0.96 | 88 | 75 | 0.98 | 69 / 0.98 | 100 | 95 |
| Scaphopoda | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 98 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 | 100 | 1.00 | 100 / 1.00 | 100 | 100 |
| Cephalopoda | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 / 1.00 | 100 | 100 | 1.00 | 100 / 1.00 | 100 | 100 |

Values listed are bootstrap support (bs) / posterior probability (pp) unless otherwise. NR, not recovered; A = Annelida, N = *Nematostella*, E = Entoprocta; C= Cycliophora.

## Supplementary Table 7 | Morphological matrix for ancestral state reconstruction.

| # | Character | Annelida | Neomeniomorpha | Chaetodermomorpha | Polyplacophora | Monoplacophora | Bivalvia | Scaphopoda | Gastropoda | Cephalopoda | Character Coding | Plesiomorphic state of Mollusca (Monoplacophora not considered) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Cuticle | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | (0) Absent; (1) Present. | Present |
| 2 | Type of cuticle | 1 | 0 | 0 | 0 | x | x | x | x | x | (0) Chitinous; (1) Collagenous; (x) | Equivocal |
| 3 | Aragonitic sclerites | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | (0) Absent; (1) Present. | Absent |
| 4 | Shell | 0 | 0 | 0 | 1 | 2 | 2 | 2 | 2 | 2 | (0) Absent; (1) Shell(s) not by shell gland; (2) Shell(s) by shell gland. | Absent |
| **5** | **Periostracum** | **0** | **0** | **0** | **0** | **1** | **1** | **2** | **1** | **1** | **(0) Absent; (1) Present; (2) Non persistent. Scaphopods have a non-persisting periostracum that is secreted but erodes quickly. Chitons do not have a true periostracum[16].** | **Absent** |
| **6** | **Periostracal groove** | **x** | **x** | **x** | **x** | **1** | **1** | **1** | **1** | **1** | **(0) Absent; (1) Present; (x) Not applicable. Coding modified to reflect changes to character 5.** | **Absent** |
| 7 | Mantle papillae | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0/1 | 0 | (0) Absent; (1) Present. | Absent |
| 8 | Mantle cavity | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | (0) Absent; (1) Present. | Present |
| 9 | Position of mantle cavity | x | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | (0) Circumpedal; (1) Posterior; (x) No mantle cavity. | Equivocal |
| 10 | Ctenidia | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | (0) Absent; (1) Present. | Present |
| 11 | Number of ctenidial pairs | x | x | 0 | 3 | 2 | 0 | x | 0 | 0/1 | (0) 1 pair; (1) 2 pairs; (2) 3-6 pairs; (3) More than 6 pairs; (x) No ctenidia. | Equivocal |

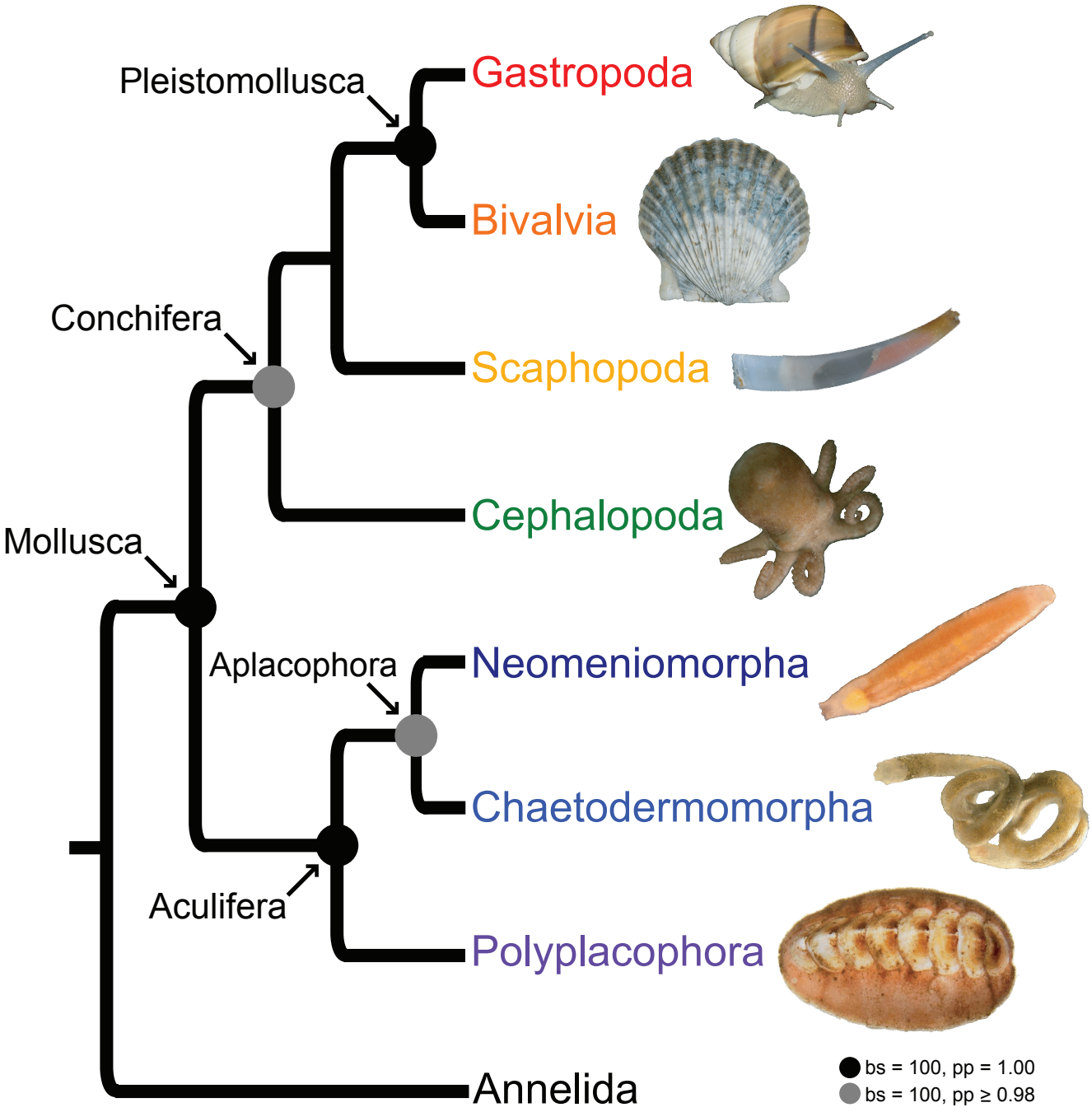
| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **12** | **Body wall musculature** | **0** | **0** | **0** | **0** | **1** | **1** | **1** | **1** | **1** | **(0) Circular/diagonal/longitudinal; (1) Otherwise. Circular/diagonal/longitudinal body wall musculature is found in larval chitons[44].** | **Circular / diagonal / longitudinal** |
| **13** | **Structure of longitudinal muscles of body** | **1** | **0** | **0** | **?** | **x** | **x** | **x** | **x** | **x** | **(0) Smooth; (1) Striated; (x) No longitudinal muscles. The condition in Polyplacophora has not been** | **Equivocal** |
| 14 | Intercrossing of the inner dorsoventral musculature (IDVM) | 0 | 1 | 0/1 | 1 | 1 | 1 | 1 | 1 | ? | (0) Absent; (1) Present. | Present |
| **15** | **Number of dorsoventral muscle pairs** | **x** | **0** | **0** | **0** | **0** | **1** | **1** | **1** | **1** | **(0) Eight or more; (1) Less than eight; (x) no IVDM. This character is coded with fewer states than in Haszprunar (2000)[4].** | **Equivocal (at least some)** |
| 17 | Specific head retractor muscles | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | (0) Absent; (1) Present. | Absent |
| 20 | Prepedal cirri | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | (0) Absent; (1) Present. | Absent |
| 21 | Coelomic cavities | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | (0) Absent; (1) Present. | Present |
| 22 | Eucoelomic condition | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | (0) Absent; (1) Present. | Absent |
| 23 | Heart in pericardium | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | (0) Absent; (1) Present. The absence of this character in scaphopods is widely | Present |
| 24 | Circulatory system | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | (0) Pseudovessels; (1) Mainly sinuisal; (2) Mainly endothelial. | Mainly sinuisal |
| 25 | Pericardioduct | x | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | (0) Absent; (1) Present; (x) No pericardium. | Present |
| **26** | **Formation of coelomoducts** | **0** | **1** | **?** | **1** | **?** | **1** | **?** | **1** | **1** | **(0) Ingrowth; (1) Outgrowth; (?) Unknown. Neomenioid coelomoducts form via outgrowth[15].** | **Outgrowth** |
| 27 | Number of coelomoduct pairs | 0/2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0/1 | (0) One; (1) Two; (2) More than two; (x) No coelomoduct. | One |
| 28 | Podocytes | 0/1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | (0) Absent; (1) Present. | Present |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **29** | **Protonephridia** | **0/1** | **1** | **1** | **1** | **?** | **1** | **1** | **1** | **0** | **(0) Absent; (1) Present. *Chaetoderma*[14] and *Wirenia*[15] have protnephridia.** | **Present** |
| 30 | Rhogocytes | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | (0) Absent; (1) Present. | Present |
| 31 | Number of gonads | 2/4 | 2 | 2 | 2 | 2/3 | 2 | 1 | 0 | 1/2/3 | (0) Single "right" (pretorsonal left); (1) Single right (2) One pair; (3) two pairs; (4) More than two pairs. | One pair |
| 32 | Position of gonad | ? | 0 | 0 | 0 | 1 | 0 | 0 | 0/1 | 0 | (0) Dorsal of gut; (1) Ventral of gut; (?) Equivocal. | Dorsal of gut |
| 34 | Release of gametes through pericardium | x | 0/1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | (0) Absent; (1) Present; (x) No pericardium. | Absent |
| **36** | **Jaws** | **1** | **0** | **0** | **0** | **2** | **0** | **2** | **2** | **2** | **(0) Absent; (1) Scleroproteinaceous; (2) Chitinous. Polychaete jaws are composed of scleroproteins while mollusc jaws are chitinous[45].** | **Equivocal** |
| 37 | Radula | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | (0) Absent; (1) Present. | Present |
| **38** | **Radular membrane** | **x** | **1** | **1** | **1** | **1** | **x** | **1** | **1** | **1** | **(0) Absent; (1) Present; (x) Radula lacking. Neomenioids have a radular membrane[24].** | **Present** |
| 39 | Radular type | x | 0 | 0 | 1 | 1 | x | 1 | 1 | 1 | (0) Basically distichous/bifid; (1) Basically rasping; (x) No radula. | Basically rasping |
| 40 | Buccal cartilages | x | 0 | 0 | 1 | 1 | x | 1 | 1 | 1 | (0) Absent; (1) Present; (x) Radula lacking. | Present |
| 41 | Oesophageal pouches | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | (0) Absent; (1) Present. | Absent |
| 42 | Highly glandular midgut | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | (0) Absent; (1) Present. | Present |
| 43 | Subdivided midgut | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | (0) Absent; (1) Present. | Present |
| 44 | Bilobed midgut gland | x | x | 0 | 1 | 1 | 1 | 1 | 1 | 1 | (0) Absent; (1) Present; (x) No midgut gland. | Equivocal |
| 46 | Intestinal loops | 0/1 | 0 | 0 | 3 | 2 | 3 | 3 | 3 | 3 | (0) Absent; (1) Along longitudinal axis; (2) Unidirectional; (3) True bidirectional looping. | Equivocal |
| 47 | Position of anus | 0/1 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | (0) Opposite of oral opening; (1) Near mouth opening at dorsal side; (2) Near | Opposite of oral opening |
| 48 | Tetraneury | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | (0) Absent; (1) Present. | Present |
| 49 | Precerebral ganglia | 0 | 0/1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | (0) Absent; (1) Present. | Absent |

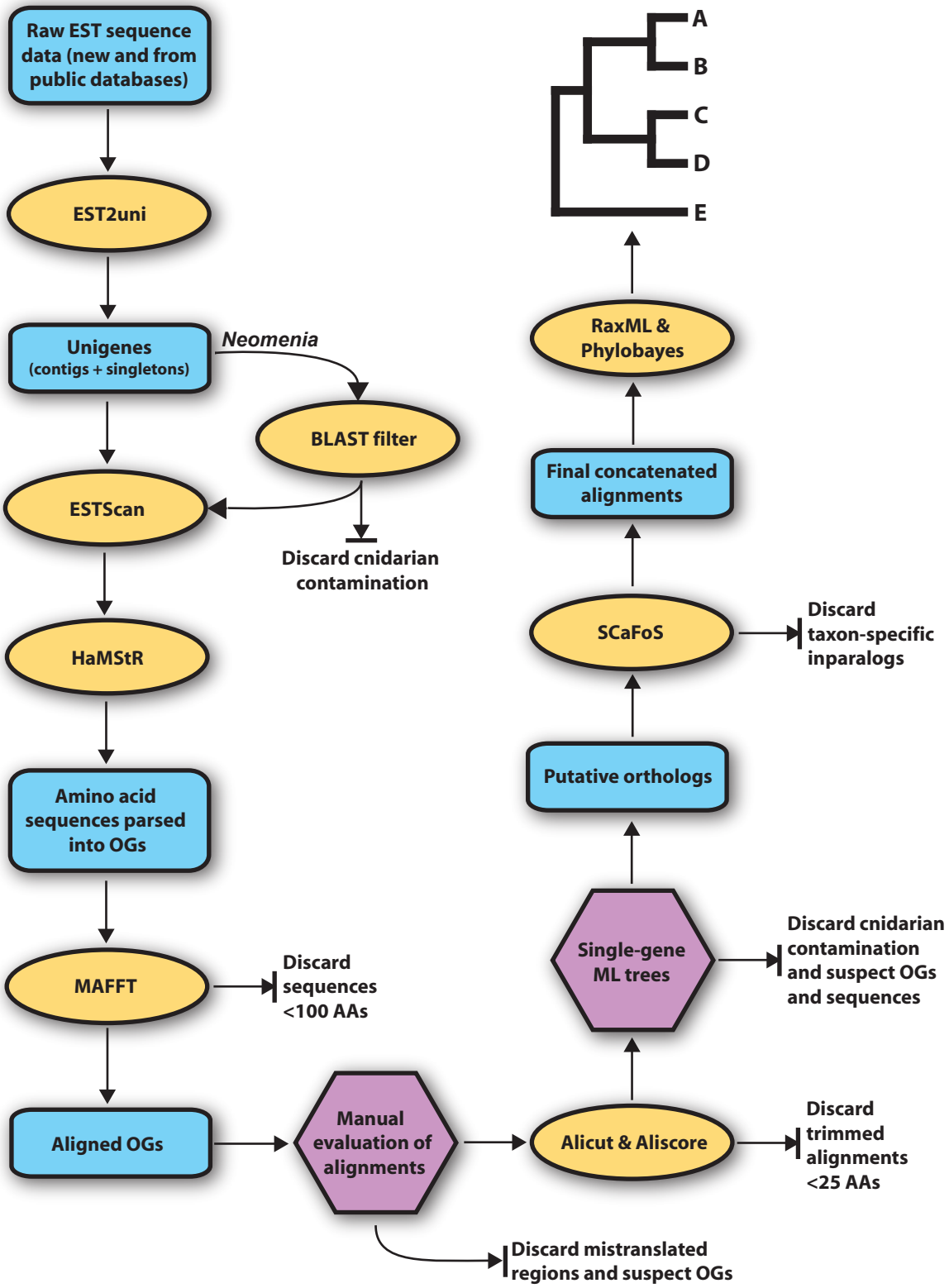
| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **50** | **Pedal ganglia** | **x** | **1** | **0** | **0** | **0** | **1** | **1** | **1** | **1** | **(0) Absent; (1) Present. Neomenioids[46] and gastropods have pedal (ventral) ganglia.** | **Equivocal** |
| 51 | Position of visceral loop | x | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 2 | (0) Between DVM; (1) Outwards DVM; (2) Inwards DVM; (x) No DVM. | Outwards of DVM |
| 52 | Position of visceral commisure | x | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | (0) Suprarectal; (1) Subrectal; (x) Homology unclear. | Equivocal |
| 53 | Innervation of the shell margin | x | x | x | 0 | 0 | 0 | 0 | 1 | 1 | (0) Cerebropleural and visceral; (1) Only cerebropleural; (x) No shell(s). | Equivocal |
| 54 | Cerebral (pretrochal) eyes | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | (0) Absent; (1) Present. | Equivocal |
| 55 | Paired statocysts | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | (0) Absent; (1) Present. | Absent |
| 57 | Position of osphradium | x | 1 | 1 | 0 | x | 0 | x | 0 | 0 | (0) Pallial; (1) Extrapallial; (x) No osphradium. | Pallial |
| **60** | **Foot** | **0** | **1** | **0** | **1** | **1** | **1** | **1** | **1** | **1** | **(0) Absent; (1) Present.** | **Present** |
| **61** | **Foot intrinsic musculature** | **x** | **0** | **x** | **0** | **1** | **1** | **1** | **1** | **1** | **(0) Absent; (1) Present; (x) No foot. The musculature of the chiton foot is built up of dorsoventral muscles exclusively[47]. Monoplacophorans have a weak intrinsic musculature in the rim of the foot[48].** | **Equivocal** |
| **62** | **Internal fertilization** | **0/1** | **1** | **0** | **0** | **0** | **0** | **0** | **0/1** | **1** | **(0) Absent; (1) Present.** | **Absent** |
| **63** | **Secondary (anterior) ciliary rootlet** | **0/1** | **1** | **1** | **1** | **?** | **0** | **1** | **0** | **?** | **(0) Absent; (1) Present. Coding as per Lundin and Schander (1999)[49], Lundin and Schander (2001)[50], and Lundin et al. (2008)[19].** | **Present** |
| **64** | **Adult excretory organs** | **x** | **?** | **1** | **0** | **2** | **0** | **0** | **0** | **0** | **(0) Connected to the pericardium; (1) Integrated within the peridardioducts; (2) Other. Coding as per Salvini Plawen (2006)[51].** | **Connected to the pericardium** |
| **65** | **Subradular membrane** | **0** | **0** | **0** | **1** | **1** | **x** | **1** | **1** | **1** | **(0) Absent; (1) Present; (x) Radula lacking. Coding as per Todt and Salvini-Plawen (2005)[52].** | **Equivocal** |
| **66** | **Locomotory cilia in foregut** | **1** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **(0) Absent; (1) Present.** | **Absent** |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **67** | **Hemocyanin** | **0** | **0** | **1** | **1** | **1** | **0/1** | **0** | **1** | **1** | **(0) Absent; (1) Present. Coding as per Lieb and Todt (2008)[53] and Lieb and Wilson (2010)[54].** | **Present** |
| **68** | **Sulfated groups in cuticle staining with DMMB** | **?** | **1** | **1** | **0** | **x** | **x** | **x** | **x** | **x** | **(0) Absent; (1) Present; (x) No cuticle. Coding as per Furuhashi et al. (2009)[55].** | **Absent** |
| **69** | **Apatite in radula** | **x** | **?** | **1** | **1** | **0** | **x** | **0** | **0** | **0** | **(0) Absent; (1) Present; (x) Radula lacking. Coding as per Cruz et al. (1998)[56].** | **Absent** |

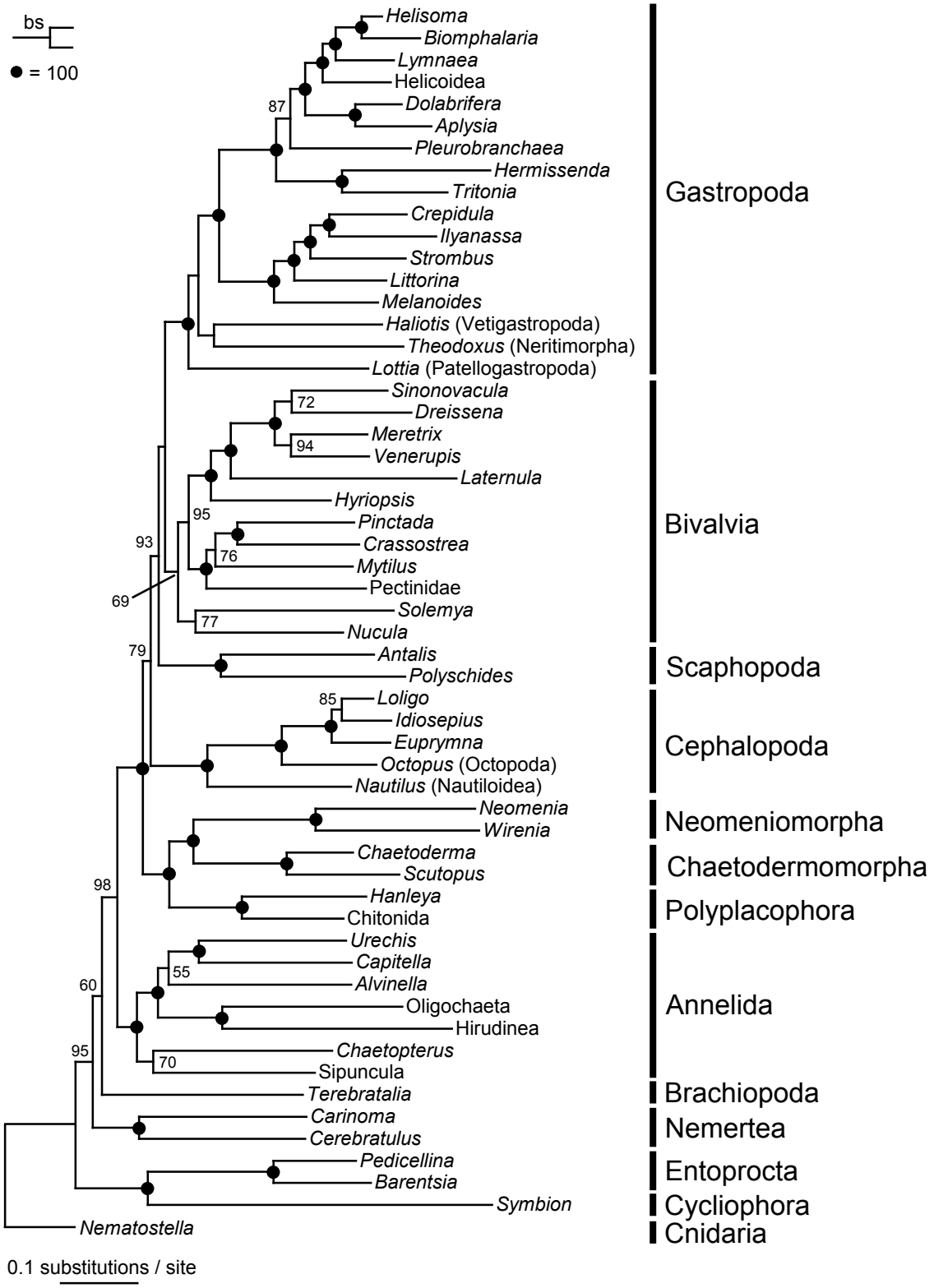
Character numbers and coding based on Haszprunar (2000)[4]; citations are provided for changed or new characters indicated by bold text. Note that several characters suspected of being homoplasious were omitted.



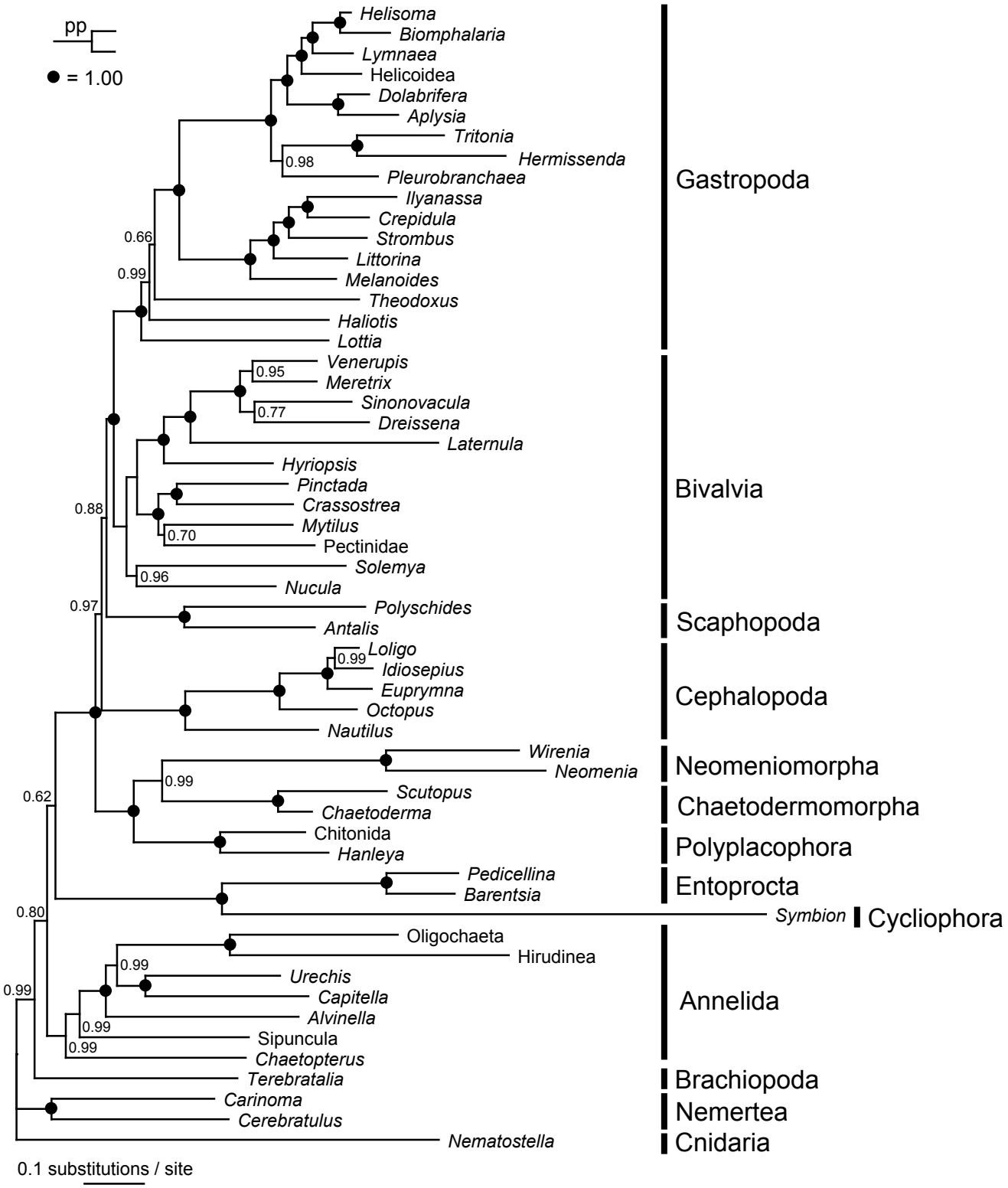

**Supplementary Figure 1 | Deep molluscan phylogeny as inferred in the present study.** Black circles represent nodes with bs = 100 and pp = 1.00. Gray circles represent nodes with bs = 100 and pp ≥ 0.98. The actual specimens of Polyschides and Hanleya used in this study are shown. Photos are not to scale. This is a full-page version of Figure 4 from the main text.



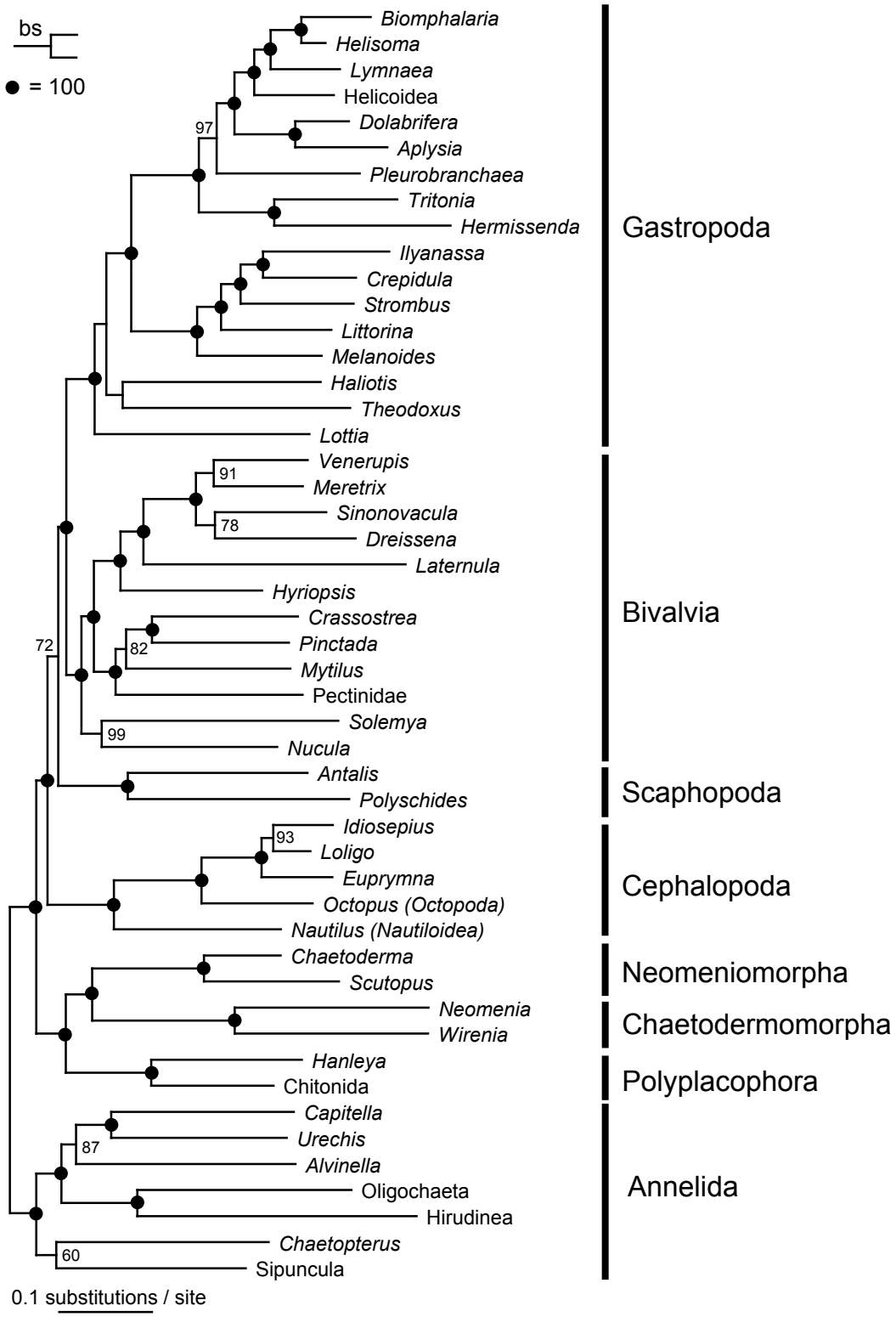

**Supplementary Figure 2 | Flow chart of bioinformatics pipeline.** Rounded blue rectangles represent input / output files, tan ovals represent programs or scripts, and violet hexagons represent steps involving manual evaluation.



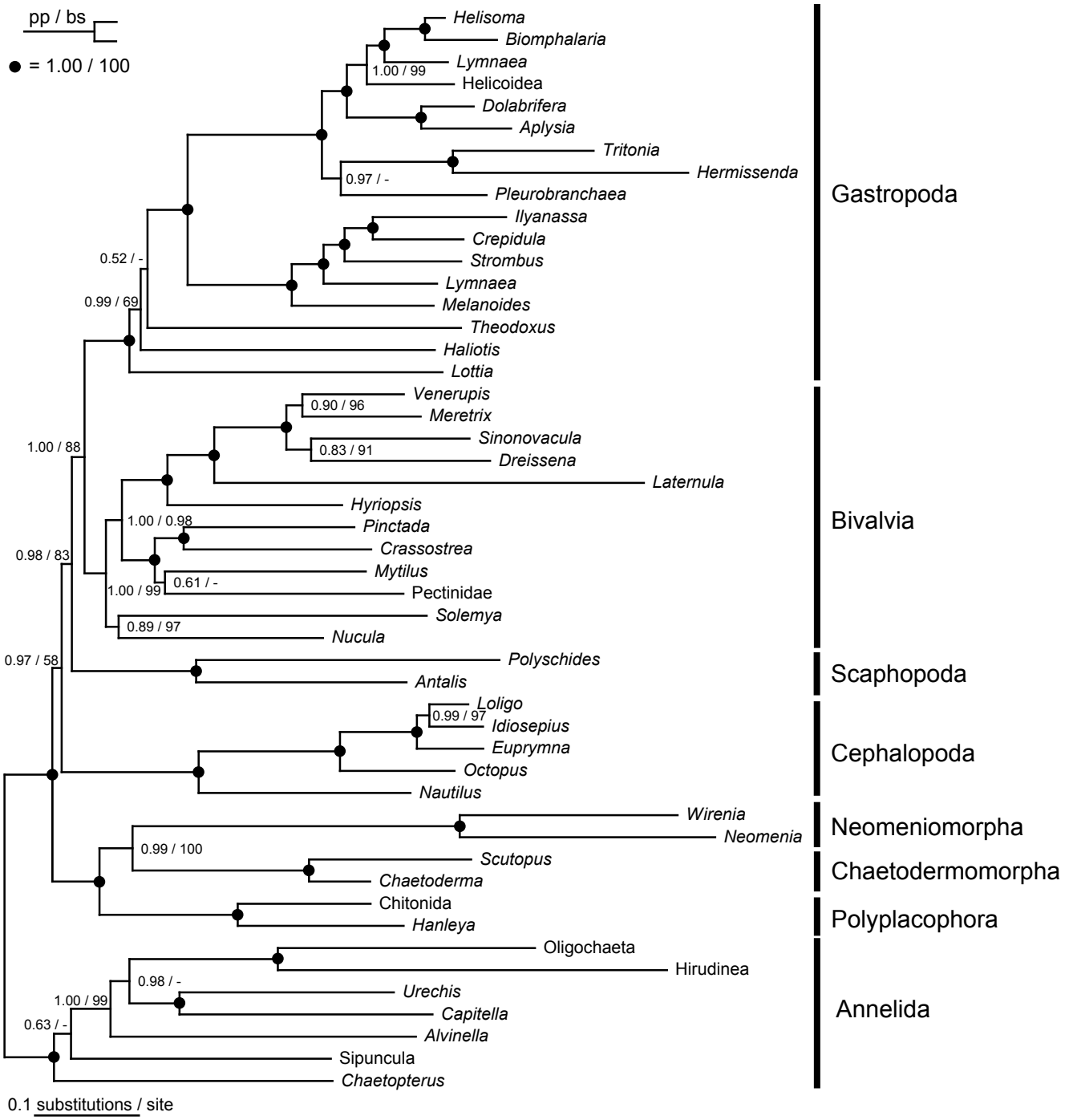

**Supplementary Figure 3 | Maximum likelihood topology based on 308 genes with broad outgroup sampling.** The most-likely tree (log likelihood = -1,197,496.85) sampled in RAxML using the best-fitting AA substitution model for each gene is shown. ML bootstrap (bs) support values >50 are listed at each node. Filled circles represent nodes with bs = 100. Average percent of genes sampled per taxon is 40% and overall matrix completeness is 26%. The length of the matrix is 84,614 AAs.



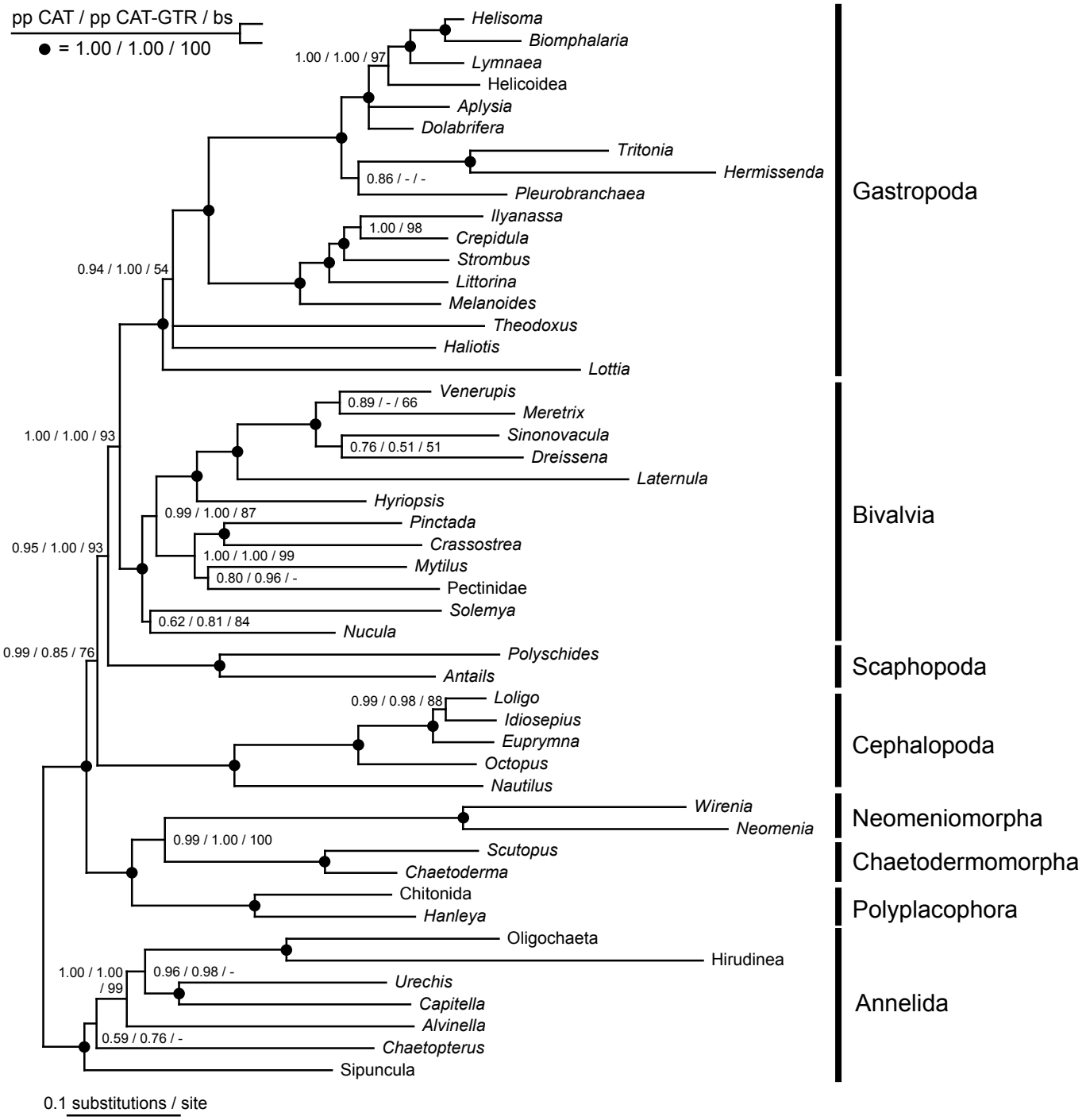

**Supplementary Figure 4 | Bayesian inference topology based on 308 genes with broad out-group sampling.** Fifty percent majority-rule consensus tree shown. Posterior probabilities (pp) >0.50 are listed at each node. Filled circles represent nodes with pp = 1.00. The average percent of genes sampled per taxon is 40% and overall matrix completeness is 26%. The length of the matrix is 84,614 AAs.



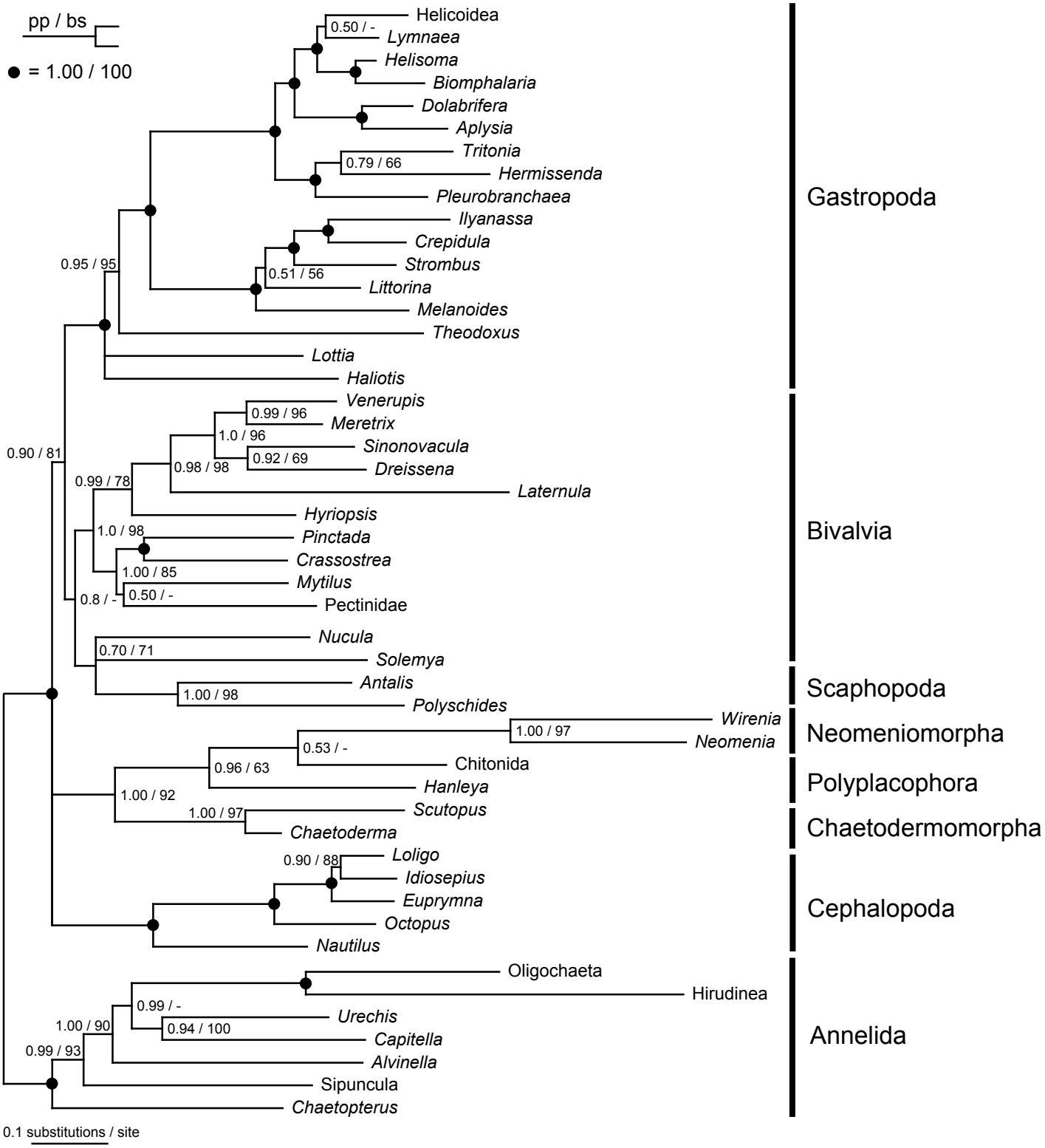

**Supplementary Figure 5 | Maximum likelihood topology based on 308 genes with Annelida outgroup.** The most-likely tree (log likelihood = -1,048,338.79) sampled in RAxML using the best-fitting AA substitution model for each gene is shown. ML bootstrap (bs) support values >50 are listed at each node. Filled circles represent nodes with bs = 100. The average percent of genes sampled per taxon is 41% and overall matrix completeness is 26%. The length of the matrix is 84,614 AAs.



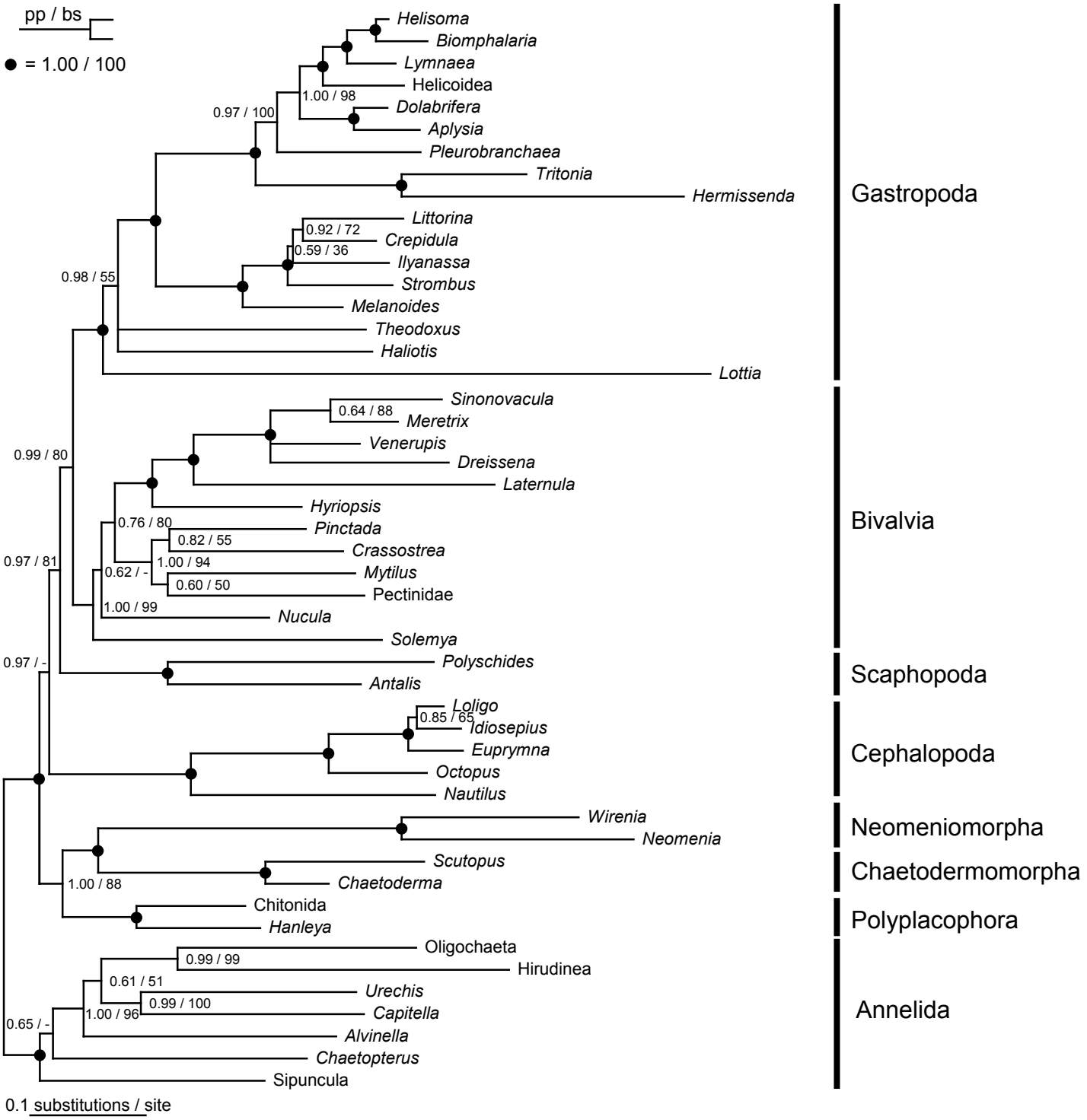

**Supplementary Figure 6 | Bayesian inference topology based on 200 best-sampled genes.** Fifty percent majority-rule consensus tree shown. Posterior probabilities (pp) >0.50 and bootstrap support values (bs) >50 are listed at each node. Filled circles represent nodes with pp = 1.00 and bs = 100. The average percent of genes sampled per taxon is 48% and overall matrix completeness is 31%. The length of the matrix is 52,686 AAs.



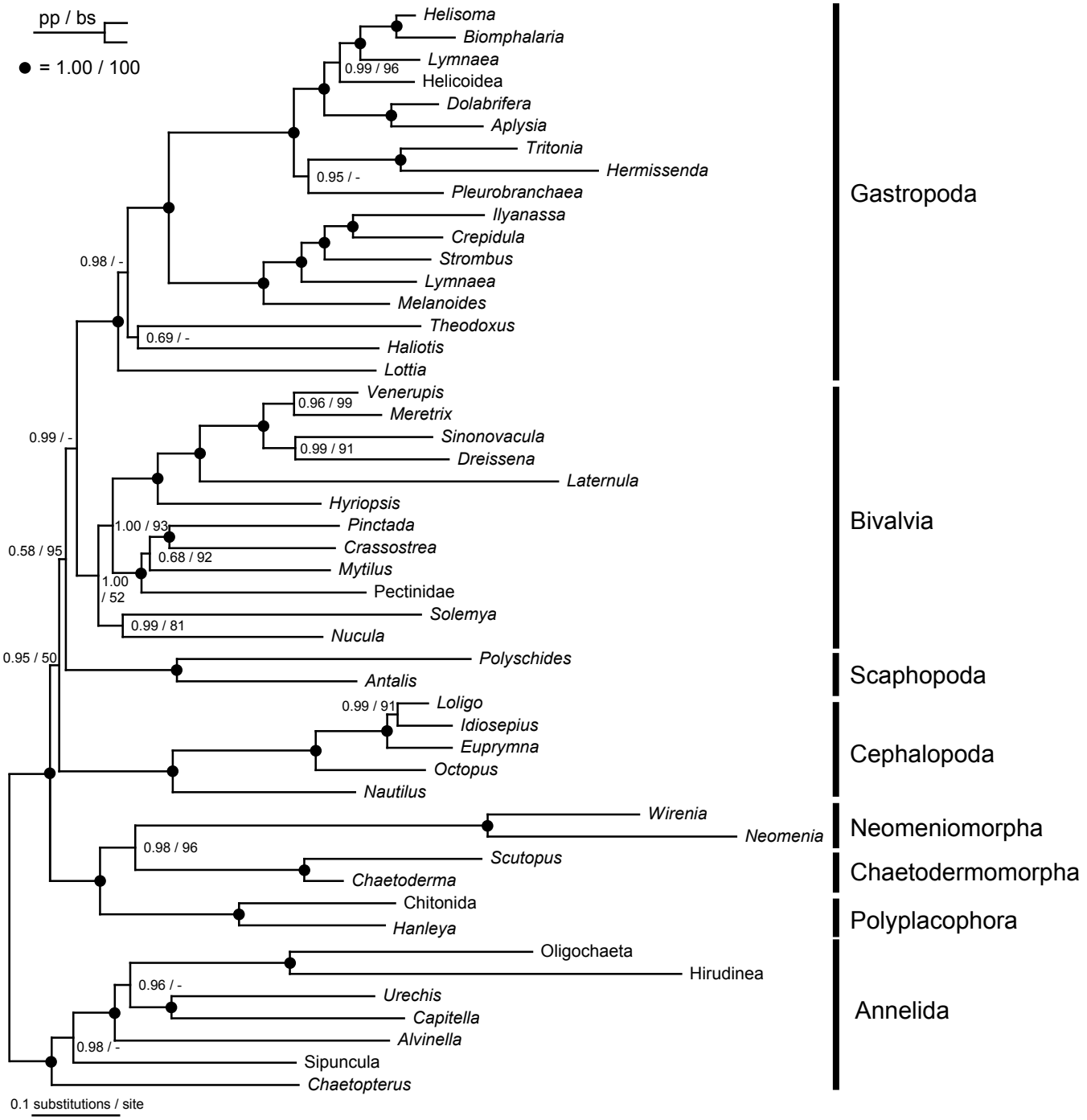

**Supplementary Figure 7 | Bayesian inference topology based on 100 best-sampled genes.** Fifty percent majority-rule consensus tree inferred using CAT model shown with CAT and CAT-GTR posterior probabilities (pp) >0.50 and bootstrap support values (bs) >50 listed at each node. Filled circles represent nodes with pp CAT = 1.00, pp CAT-GTR = 1.00, and bs = 100. The average percent of genes sampled per taxon is 61% and overall matrix completeness is 44%. The length of the matrix is 22,053 AAs.



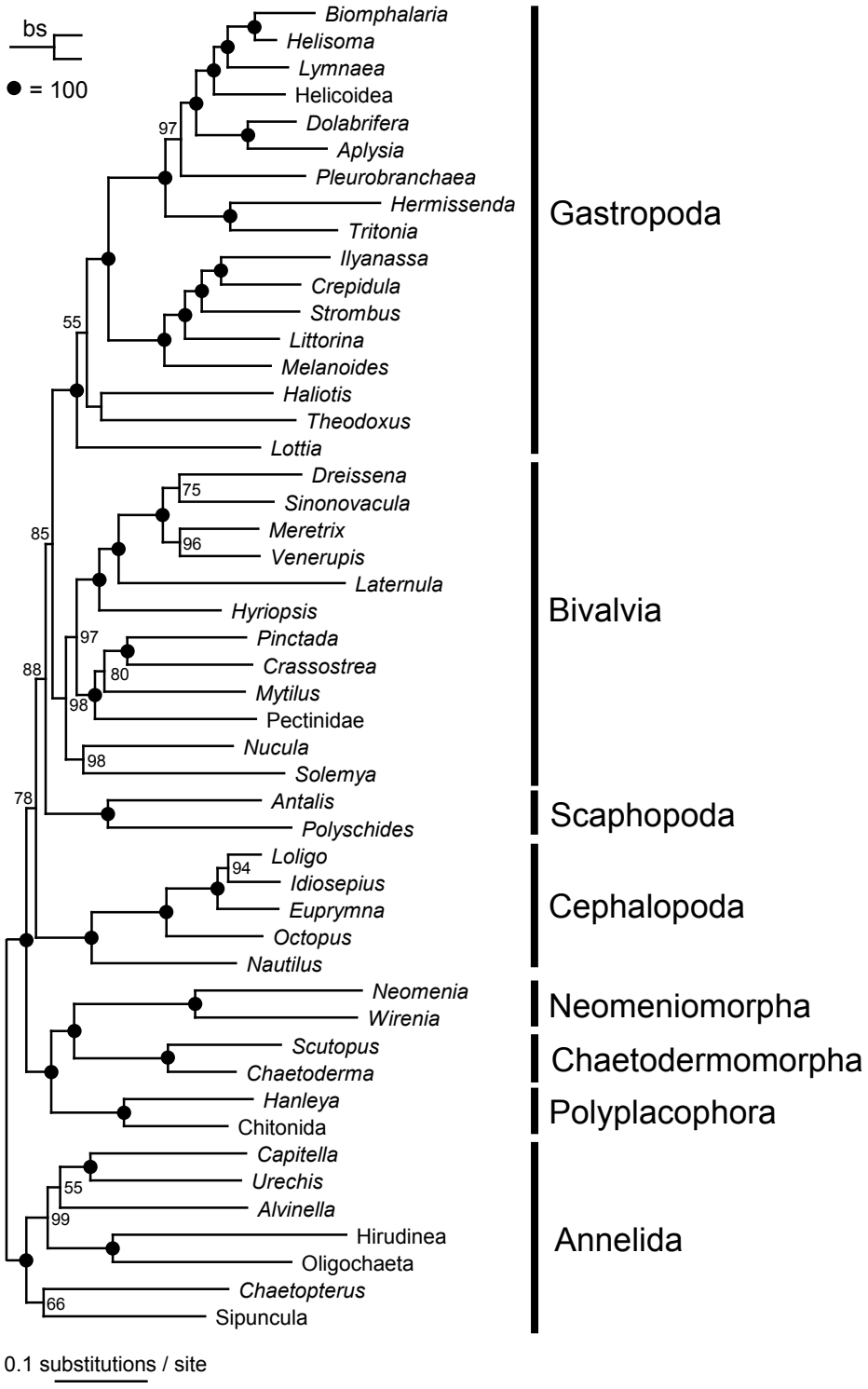

**Supplementary Figure 8 | Bayesian inference topology based on non-ribosomal proteins.** Fifty percent majority rule consensus of approximately 3,000 trees per chain (5 chains) after discarding the first 5,000 trees as burn-in. Posterior probabilities (pp) >0.50 and bootstrap support values (bs) >50 are listed at each node. Filled circles represent nodes with pp = 1.00 and bs = 100. The average percent of genes sampled per taxon is 30% and overall matrix completeness is 22%. The length of the matrix, which includes 260 genes, is 76,527 AAs.



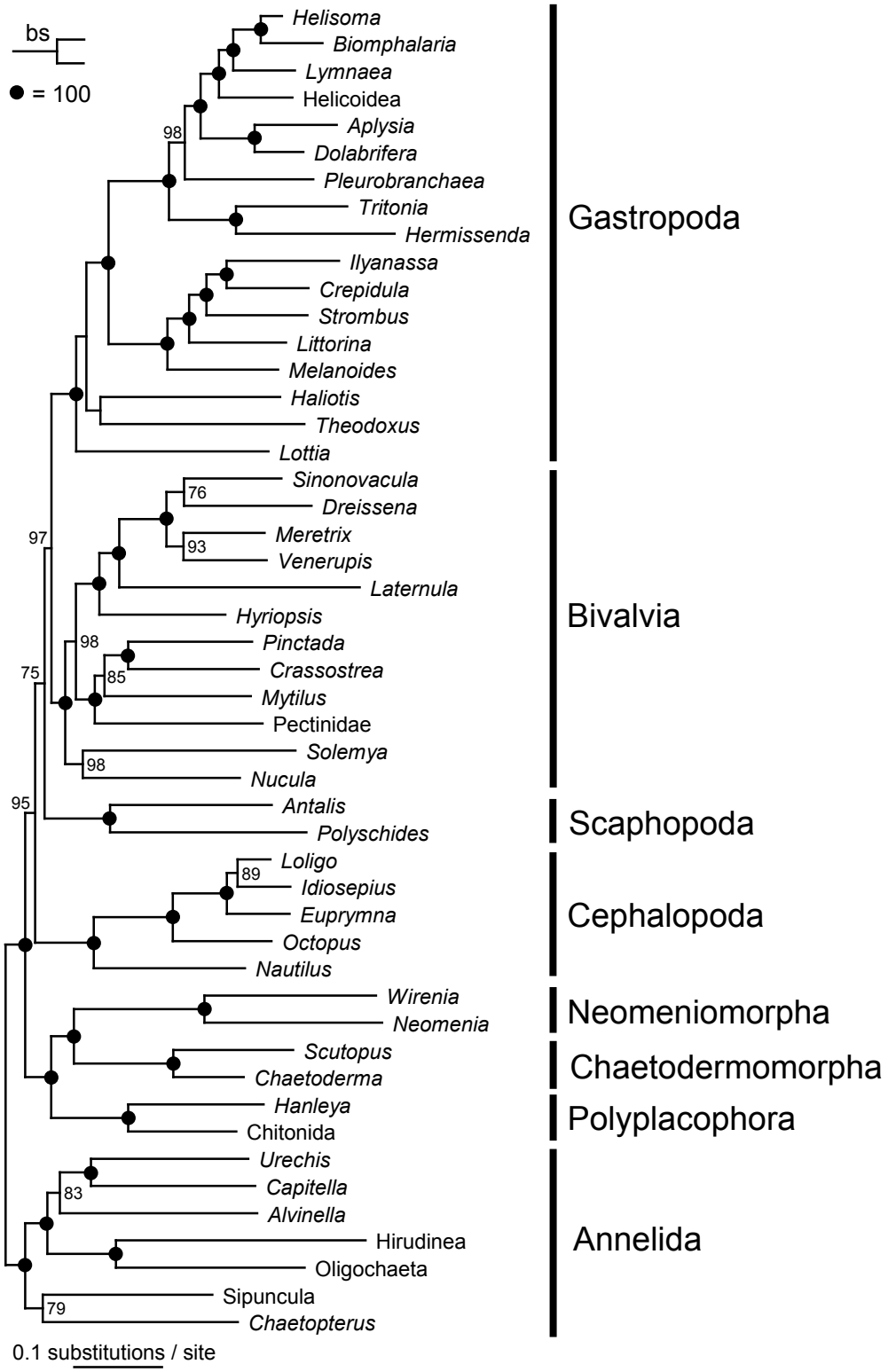

**Supplementary Figure 9 | Bayesian inference topology based on ribosomal proteins.** Fifty percent majority-rule consensus tree shown. Posterior probabilities (pp) >0.50 and bootstrap support values (bs) >50 are listed at each node. Filled circles represent nodes with pp = 1.00 and bs = 100. The average percent of genes sampled per taxon is 67% and overall matrix completeness is 59%. The length of the matrix, which includes 49 genes, is 8,087 AAs.



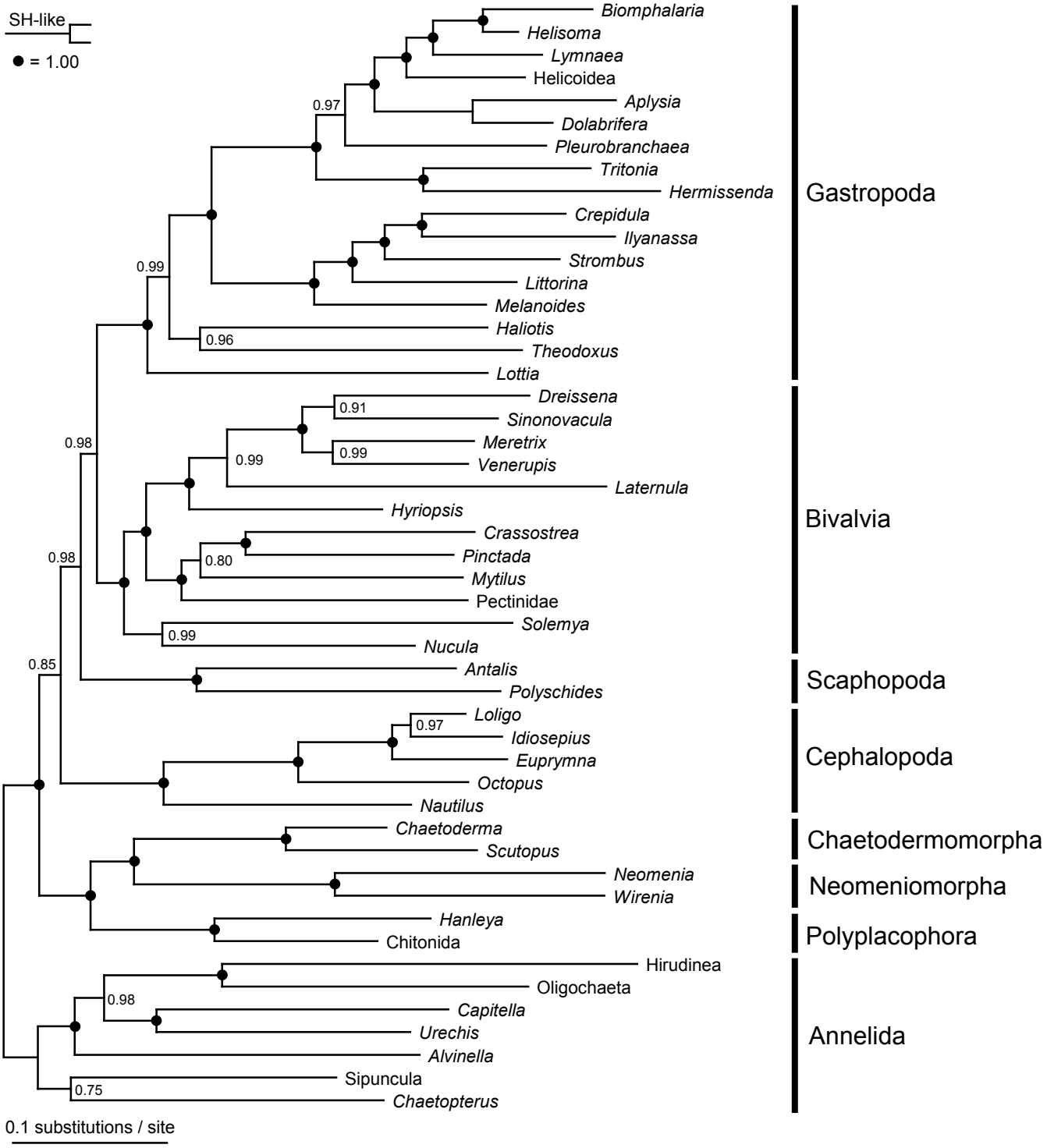

**Supplementary Figure 10 | Bayesian inference topology based on genes our method and InParanoid identify the same Lottia sequence as an ortholog to the primer taxon (Drosophila) sequence.** Fifty percent majority-rule consensus tree shown. Posterior probabilities (pp) >0.50 and bootstrap support values (bs) >50 are listed at each node. Filled circles represent nodes with pp = 1.00 and bs = 100. The average percent of genes sampled per taxon is 40% and overall matrix completeness is 25%. The length of the matrix, which includes 243 genes, is 66,821 AAs.



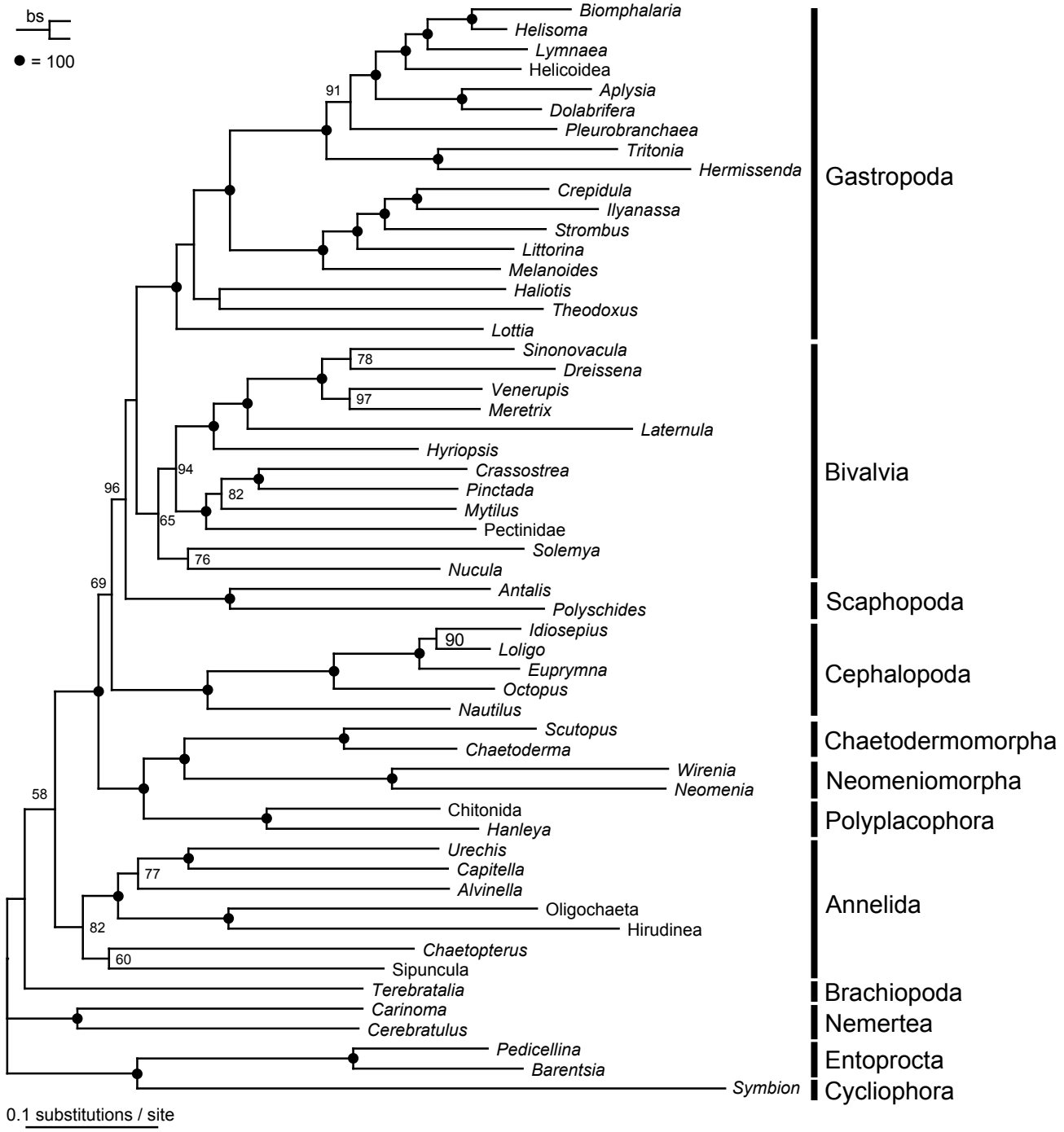

**Supplementary Figure 11 | Maximum likelihood topology based on all 308 genes using the WAG + CAT + F model.** The most likely tree (log likelihood = -1,055,336.03) sampled in RAxML is shown. Bootstrap support (bs) values >50 are listed at each node. Filled circles represent nodes with bs = 100. The average percent of genes sampled per taxon is 41% and overall matrix completeness is 26%. The length of the matrix is 84,614 AAs.



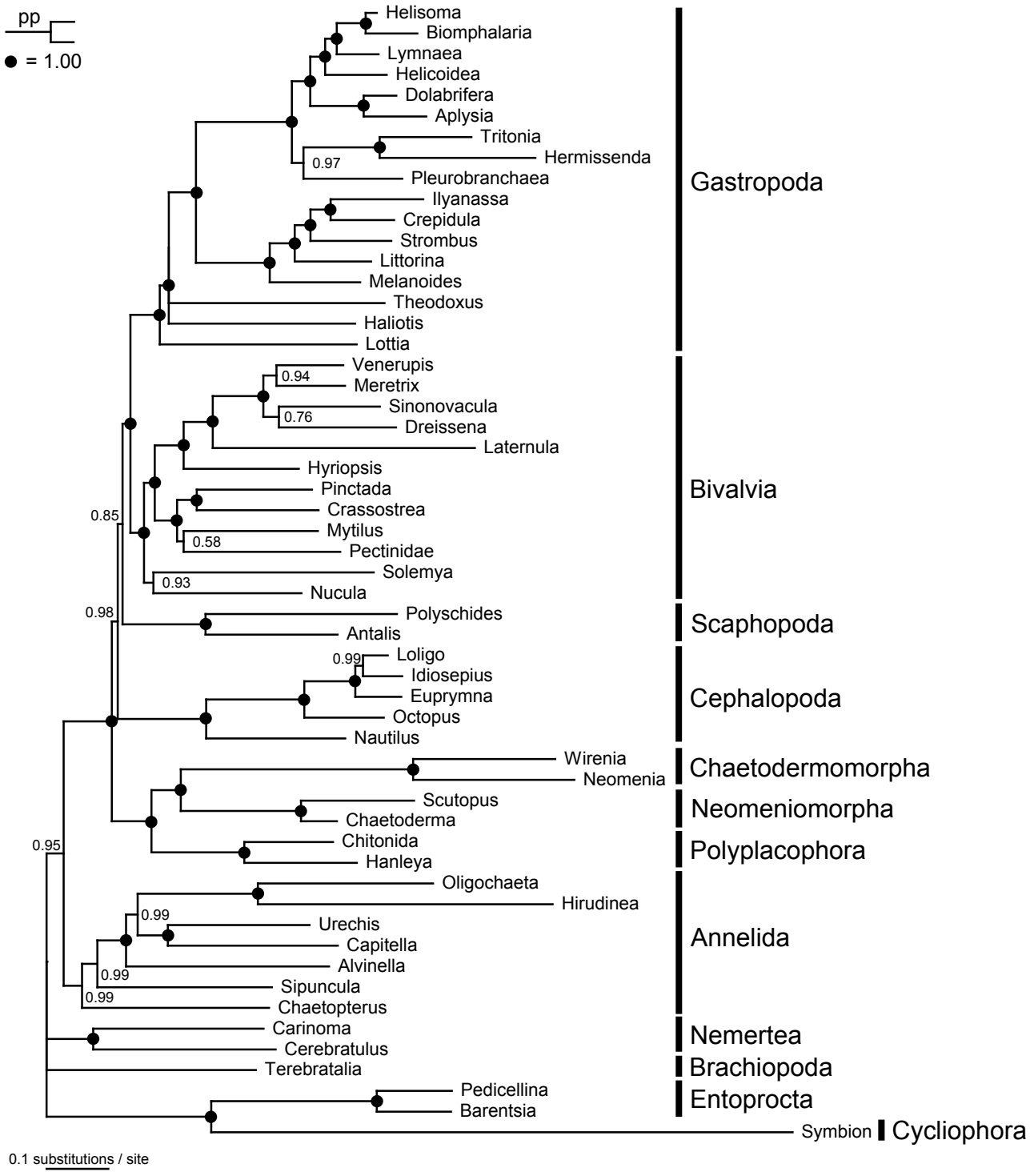

**Supplementary Figure 12 | Maximum likelihood topology based on all 308 genes using the LG + CAT + F model.** Most likely tree sampled in RaxML shown (log likelihood = -1,052,785.42). Bootstrap support values >50 are listed at each node. Filled circles represent nodes with bs = 100. The average percent of genes sampled per taxon is 41% and overall matrix completeness is 26%. The length of the matrix is 84,614 AAs.



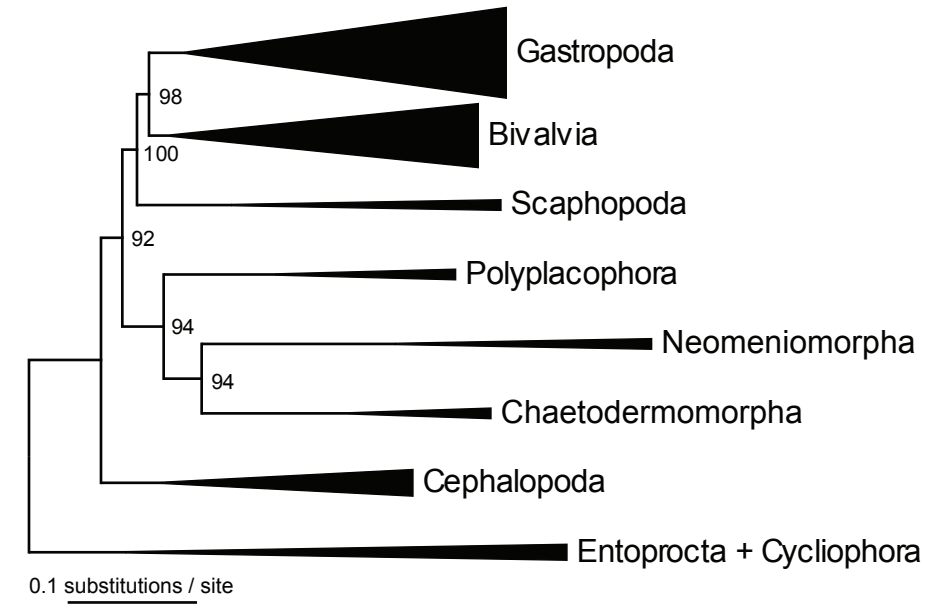

**Supplementary Figure 13 | FastTree topology based on all 308 genes using the JTT + CAT model.** The most likely tree sampled in FastTree is shown. SH-like support values >50 are listed at each node. Filled circles represent nodes with SH-like support values of 100. The average percent of genes sampled per taxon is 41% and overall matrix completeness is 26%. The length of the matrix is 84,614 AAs.



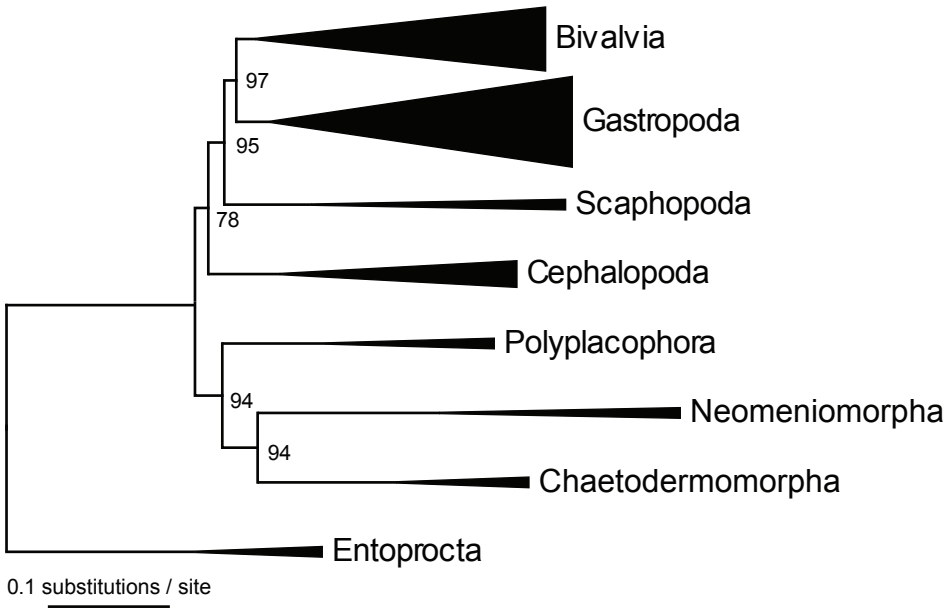

**Supplementary Figure 14 | Maximum likelihood topology based on 308 genes with broad outgroup sampling excluding Nematostella.** The most-likely tree (log likelihood = -1,124,009.87) sampled in RAxML using the best-fitting AA substitution model for each gene is shown. ML bootstrap (bs) support values >50 are listed at each node. Filled circles represent nodes with bs = 100. Average percent of genes sampled per taxon is 39% and overall matrix completeness is 25%. The length of the matrix is 84,614 AAs. This tree has been drawn to be consistent with Fig 2 but it should be considered as an unrooted tree.




**Supplementary Figure 15 | Bayesian inference topology based on 308 genes with broad outgroup sampling excluding Nematostella.** Fifty percent majority-rule consensus tree shown. Posterior probabilities (pp) >0.50 and bootstrap support values (bs) >50 are listed at each node. Filled circles represent nodes with pp = 1.00 and bs = 100. Average percent of genes sampled per taxon is 39% and overall matrix completeness is 25%. The length of the matrix is 84,614 AAs. This tree has been drawn to be consistent with Fig 2 but it should be considered as an unrooted tree.




a




b

**Supplementary Figure 16 | Maximum likelihood topologies based on 308 genes with Entoprocta + Cycliophora or Entoprocta as outgroup.** Most-likely tree sampled in RaxML with a, Entoprocta + Cycliophora as outgroup and b, only Entoprocta as the outgroup. All major lineages collapsed for simplicity. Bootstrap support values (bs) >50 are listed at each node.