

OPEN

New data from Monoplacophora and a carefully-curated dataset resolve molluscan relationships

Kevin M. Kocot^{1*}, Albert J. Poustka^{2,3}, Isabella Stöger^{4,5}, Kenneth M. Halanych⁶ & Michael Schrödl^{4,5,7}

Relationships among the major lineages of Mollusca have long been debated. Morphological studies have considered the rarely collected Monoplacophora (Tryblidia) to have several plesiomorphic molluscan traits. The phylogenetic position of this group is contentious as morphologists have generally placed this clade as the sister taxon of the rest of Conchifera whereas earlier molecular studies supported a clade of Monoplacophora + Polyplacophora (Serialia) and phylogenomic studies have generally recovered a clade of Monoplacophora + Cephalopoda. Phylogenomic studies have also strongly supported a clade including Gastropoda, Bivalvia, and Scaphopoda, but relationships among these taxa have been inconsistent. In order to resolve conchiferan relationships and improve understanding of early molluscan evolution, we carefully curated a high-quality data matrix and conducted phylogenomic analyses with broad taxon sampling including newly sequenced genomic data from the monoplacophoran *Laevipilina antarctica*. Whereas a partitioned maximum likelihood (ML) analysis using site-homogeneous models recovered Monoplacophora sister to Cephalopoda with moderate support, both ML and Bayesian inference (BI) analyses using mixture models recovered Monoplacophora sister to all other conchiferans with strong support. A supertree approach also recovered Monoplacophora as the sister taxon of a clade composed of the rest of Conchifera. Gastropoda was recovered as the sister taxon of Scaphopoda in most analyses, which was strongly supported when mixture models were used. A molecular clock based on our BI topology dates diversification of Mollusca to ~546 MYA (+/- 6 MYA) and Conchifera to ~540 MYA (+/- 9 MYA), generally consistent with previous work employing nuclear housekeeping genes. These results provide important resolution of conchiferan mollusc phylogeny and offer new insights into ancestral character states of major mollusc clades.

Mollusca is the second most diverse animal phylum whose members exhibit an incredible array of body shapes and sizes. Many molluscs have important ecological roles in marine, freshwater, and terrestrial environments and others are culturally and/or economically important as a source of food, jewellery, or dye¹. Despite their diversity and importance, understanding of early molluscan evolution remains incomplete and several conflicting phylogenetic hypotheses¹⁻⁹ have been proposed regarding relationships among the eight major clades (i.e., classes): Bivalvia (clams, scallops, oysters, etc.), Caudofoveata (Chaetodermomorpha), Cephalopoda (octopuses, squids, and *Nautilus*), Gastropoda (snails and slugs), Monoplacophora (Tryblidia; deep-sea, limpet-like molluscs), Polyplacophora (chitons), Scaphopoda (tusk shells), and Solenogastres (Neomeniomorpha).

Within Conchifera (Bivalvia, Cephalopoda, Gastropoda, Monoplacophora, and Scaphopoda), the clade of molluscs with uni- or bivalved shells, the deep-sea limpet-like Monoplacophora has long been thought to be important to understanding early molluscan evolution^{5,10-14} with most morphology-based hypotheses placing Monoplacophora sister to a clade of all other conchiferans. However, no published molecular studies have

¹Department of Biological Sciences and Alabama Museum of Natural History, University of Alabama, Tuscaloosa, Alabama, 35487, USA. ²Evolution and Development Group, Max-Planck Institute for Molecular Genetics, Berlin, 14195, Germany. ³Environmental and Phylogenomics Group, Dahlem Centre for Genome Research and Medical Systems Biology, Berlin, 12489, Germany. ⁴SNSB-Bavarian State Collection of Zoology, Munich, 81247, Germany. ⁵Department Biology II, Ludwig-Maximilians-Universität, Planegg-Martinsried, 82152, Germany. ⁶Department of Biological Sciences, Auburn University, Auburn, Alabama, 36849, USA. ⁷GeoBio-Center LMU, München, 80333, Germany. *email: kmkocot@ua.edu

supported this topology to date (but see Philippe and Roure 2012¹⁵). Studies of molluscan phylogeny employing datasets dominated by nuclear ribosomal and mitochondrial genes have generally had poor resolution among major lineages^{10–14}. However, one finding of particular interest from these studies was the recovery of a close relationship of Monoplacophora and Polyplacophora (Serialia)^{16,17}. More recent studies employing PCR-amplified fragments of nuclear protein-coding “housekeeping” genes¹⁸ or nuclear protein-coding genes obtained from transcriptome and genome data^{19,20} have instead provided strong support for a clade called Aculifera, which groups Polyplacophora with Aplacophora (Caudofoveata + Solenogastres) to form a group of molluscs with calcareous sclerites.

Smith *et al.*¹⁹, the only published phylogenomic study to date focused on deep molluscan relationships to sample Monoplacophora (specifically *Laevipilina hyalina*), recovered it as the sister taxon of Cephalopoda. This result is inconsistent with the prevailing traditional morphological view placing Monoplacophora sister to all other conchiferans^{3,21–23}, but is consistent with some (but not all) palaeontological hypotheses on early molluscan diversification^{24–27}. Two subsequent studies included data from *L. hyalina* but focused on relationships within Gastropoda²⁸ or Bivalvia²⁹, and thus had limited taxon sampling outside of those clades. Kocot *et al.*³⁰ focused on among-phylum relationships within Lophotrochozoa but had relatively broad sampling of Mollusca. Most of those analyses recovered Monoplacophora as the sister taxon of Conchifera or Cephalopoda, but support for its placement was generally weak. Phylogenomic studies have also supported a clade including Gastropoda, Bivalvia, and Scaphopoda, although there has been inconsistency in recovered relationships among these taxa^{19,20,28,30}. Because conchiferan molluscs are well-represented in the early animal fossil record^{31,32}, understanding their phylogeny has important implications for understanding early animal evolution and the identity of enigmatic fossil taxa hypothesized to be stem-group molluscs.

Results and Discussion

We sequenced a draft genome for the monoplacophoran *Laevipilina antarctica*. Unfortunately, because of the small size of this species, there was only adequate material for paired-end Illumina sequencing library preparation with insufficient material for mate pair, long-read, or transcriptome library preparation using techniques available at the time that this work was conducted. This resulted in a rather fragmented genome assembly (427,488 contigs >500 bp; N50 = 2,167 bp; 1.26 Gbp total assembly size). Assessment of this assembly with BUSCO³³ showed that it is rather incomplete with only 14.6% of the 978 metazoa_odb9 genes recovered as complete and another 17.9% recovered as fragmented. Nevertheless, aside from transcriptome data from *Laevipilina hyalina*, these represent the only available genome data from any monoplacophoran and are thus a valuable resource for testing the phylogenetic position of this group.

We curated a dataset of 257 genes totalling 54,596 amino acids in length with data from 49 taxa of which 32 represented ingroup species (Supplementary Table 1). Care was taken to exclude possible contamination and mistranslated sequence regions (see Methods) while minimizing the amount of missing data in the final matrix (27.86% missing data). Additionally, only genes with a sequence from *L. antarctica* were sampled. Phylogenetic analyses were conducted using maximum likelihood (ML) in RAxML^{8,34} with the best-fitting model for each gene, and in IQ-TREE using the posterior mean site frequency (LG + C60 + G + F) mixture model^{35–37}. A Bayesian inference (BI) analysis was conducted in PhyloBayes MPI³⁸ with the CAT-GTR mixture model³⁹.

ML analysis of the partitioned dataset in RAxML (Fig. 1A) recovered Monoplacophora sister to Cephalopoda with moderate bootstrap support (bs = 88), consistent with the results of Smith *et al.*¹⁹ and some interpretations of the fossil record¹¹. However, the ML analysis in IQ-TREE using the PMSF model (Fig. 1B) and the Bayesian inference analysis in PhyloBayes using the CAT-GTR model (Fig. 1C) recovered Monoplacophora sister to the rest of Conchifera with a bootstrap support value of 94 and posterior probability of 0.99 respectively, consistent with most morphology-based hypotheses of conchiferan relationships¹¹.

To examine support for Monoplacophora sister to Conchifera from individual partitions, we used a multi-species coalescent approach in ASTRAL 5.6.1⁴⁰. This analysis also recovered Monoplacophora sister to the rest of Conchifera (local posterior probability, lpp = 0.89; Fig. 1D).

Placement of Monoplacophora sister to all other conchiferans had a lower likelihood score than Monoplacophora + Cephalopoda in the RAxML analysis and could not be rejected by the Shimodaira-Hasegawa (SH) test ($p = 0.190$). This alternative topology was, however, rejected by the Approximately Unbiased (AU) test ($p = 0.001$). Both tests rejected the Serialia hypothesis (AU test $p = 0.001$; SH test $p = 0$).

A clade of all conchiferans except Monoplacophora, as recovered in most of our analyses, was originally proposed by morphologists and called Ganglionata (reviewed by Schrödl and Stöger 2014⁵). Despite the name, ganglia are neither restricted to Ganglionata nor do all species within Ganglionata show distinct pairs of ganglia^{41–43}. Kocot *et al.*²⁰ curated a morphological character matrix for Mollusca building on that of Haszprunar²¹ and conducted ancestral state reconstruction for key molluscan characters (see Methods) under a number of different phylogenetic scenarios including Monoplacophora sister to the rest of Conchifera. Our analyses placing Monoplacophora sister to the rest of Conchifera indicate that the only unambiguously apomorphic trait of Ganglionata is the reduction of adult dorsoventral muscle pairs from a hypothesized ancestral set of eight (or possibly seven⁴⁴). Monoplacophorans also differ from other conchiferans with respect to the arrangement and structure of mantle folds, anatomy of the shell gland, and structure of the shell²³, but whether these are retained conchiferan plesiomorphies or monoplacophoran apomorphies is ambiguous.

Relationships among Gastropoda, Bivalvia, and Scaphopoda, a clade of molluscs with relatively thick, multi-layered shells²⁷, have been the subject of debate^{3,5,7,8,31} due to incongruence among recent studies^{18–20,45,46}. Whereas our RAxML and ASTRAL analyses found poor support for relationships among these taxa, our IQ-TREE and PhyloBayes analyses using mixture models strongly supported Scaphopoda + Gastropoda with this clade sister to Bivalvia, consistent with Smith *et al.*¹⁹. Gastropoda is an extremely diverse, morphologically disparate, and ecologically variable group of species that inhabit almost all environments on land and in the sea.

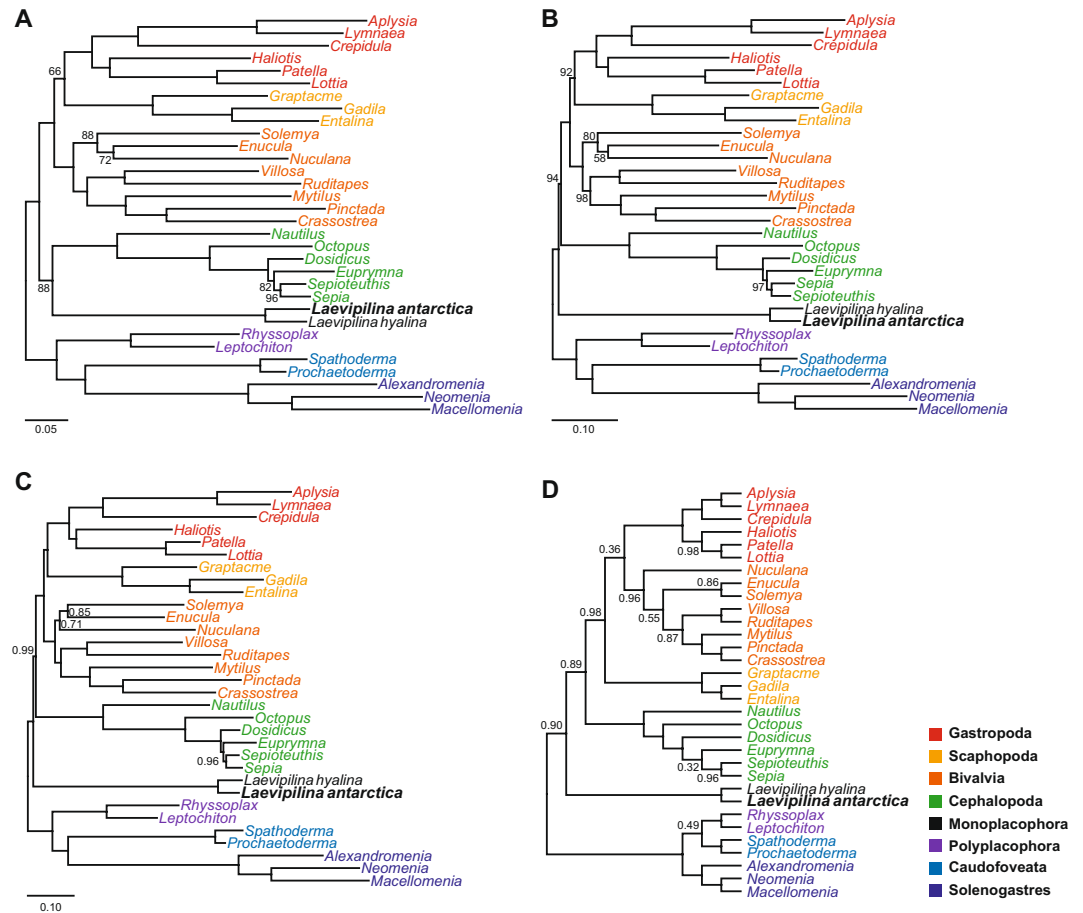


Figure 1. Results of phylogenetic analyses (outgroup taxa not shown). **(A)** RAXML maximum likelihood (ML) topology. Bootstrap support values below 100 shown. **(B)** IQ-TREE ML topology. Bootstrap support values below 100 shown. **(C)** PhyloBayes Bayesian inference (BI) topology. Posterior probabilities below 1.0 shown. **(D)** ASTRAL tree. Local posterior probabilities below 1.0 shown.

Scaphopoda, on the other hand, is a much less diverse group of relatively morphologically uniform animals that dig in marine sediments and prey upon foraminiferans and other infauna. This pair of unequal sister taxa contradicts the Cyrtosoma concept uniting Gastropoda and Cephalopoda (plus Monoplacophora by the original definition¹⁰; reviewed by Kocot²²). Interestingly, a close relationship of Scaphopoda and Gastropoda was proposed based on the pronounced dorsoventral axis⁴⁷ and recent work has confirmed the morphological ventral position of the scaphopod foot⁴⁸. Examination of published molluscan morphological data matrices^{20,21,49} reveals obvious symplesiomorphies shared between these taxa (e.g., external univalved shell), but we find no clear morphological synapomorphies for the gastropod-scaphopod clade.

Consistent with other phylogenomic studies^{18–20,50}, all of our analyses strongly support a molluscan dichotomy with two major clades: Conchifera and Aculifera⁵¹. Within Aculifera, we recovered chitons (Polyplacophora) sister to the vermiform, shell-less aplacophorans. Within Aplacophora, we recovered Solenogastres and Caudofoveata reciprocally monophyletic. Aculifera contradicts the classical morphology-based Testaria hypothesis⁵, which places chitons sister to Conchifera and the shell-less worm-like aplacophorans as an early-branching, paraphyletic grade. The Testaria hypothesis implies a progressive evolution from a simple unshelled worm-like ancestor towards chitons with shell plates and later with the uni- or bivalved conchiferans as the crown-group of Mollusca. Our results unequivocally reject this hypothesis (AU test p-value = 4.00E-56; SH test p-value = 0).

In light of support for placement of Monoplacophora sister to the rest of Conchifera and our earlier ancestral character state reconstruction analyses based on this phylogenetic hypothesis²⁰, we infer that the last common ancestor of extant molluscs was likely a dorsoventrally flattened animal that had a mantle, a dorsal cuticle, a broad foot, eight (or seven⁴⁴) pairs of dorsoventral muscles, a circumpedal or posterior mantle cavity with serially arranged gills, and a radula as part of a longitudinally arranged, regionalized digestive system. Whether or not the last common ancestor of extant molluscs had a single shell, multiple shell plates, or no shell is ambiguous²⁰. Possession of a single shell is clearly plesiomorphic for Conchifera but this was probably also the case in *Calvapilosa*, *Maikhanella*, and *Orthrozanclus*, fossil taxa inferred to be stem aculiferans⁵², suggesting that the last common molluscan ancestor may have been single-shelled. Additional studies comparing development, mineralogy, and other structural aspects of chiton shells, conchiferan shells, and aculiferan sclerites would be of great

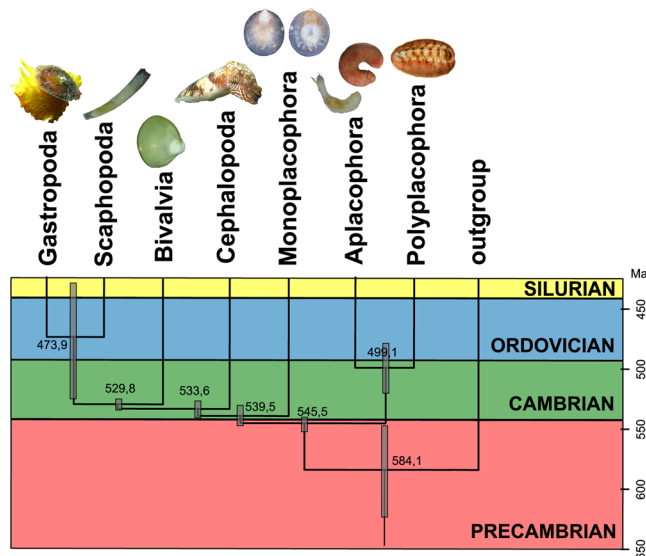


Figure 2. Summary of relaxed molecular clock analysis results. Numbers along y-axis are millions of years before present (Ma). Numbers at nodes represent the average age of the split; Error bars at nodes represent the height 95% HPD (highest posterior density). A detailed version of this tree is presented in Supplementary Fig. 1 and the raw data and uncollapsed tree are available via FigShare (see Data Availability section).

interest to further address this and other important questions about the origin(s) and homology of molluscan biomineralized structures³³.

Our molecular clock analysis (Fig. 2; Supplementary Fig. 1; Supplementary Table 3) indicates that the molluscan stem split from trochozoan relatives about 584 MYA (95% highest posterior density [HPD] = 547–623 MYA), Conchifera diversified 540 MYA (531–548 MYA), and Aculifera diversified 499 MYA (479–520 MYA), generally consistent with previous relaxed molecular clocks calculated from multilocus^{18,45,52,54} and phylogenomic data^{28,55}, showing the molluscan stem to be Precambrian in origin. The Ediacaran fossil genus *Kimberella* has been hypothesized to represent a stem-group mollusc by some^{31,56–58} but the molluscan affinity of *Kimberella* has been criticized by others who instead view it as an early-branching bilaterian³², in part because of its old age (~555 MYA). Although broad, our and other recent estimates for the divergence of molluscs are at least compatible with hypotheses regarding *Kimberella* as an early offshoot of the molluscan stemline^{31,45,59}. However, if *Kimberella* was indeed a mollusc, it differed from most extant molluscs in its lack of a shell (although sclerites may have been present) and, more significantly, a bizarre rake-like mode of feeding unlike that of any modern mollusc³².

Late Precambrian and Cambrian small shelly fossil (SSF) assemblages consist of abundant, diverse, and tiny (0.5–5 mm) animals⁶⁰ in strong contrast to the large-bodied Vendian *Kimberella*. Our time tree is consistent with the prevailing notion that SSFs such as helcionellids and other gastropod- and monoplacophoran-like fossils were conchiferan molluscs³², but relatively broad posterior densities preclude confident placement of these fossil taxa along any one branch. According to our time tree, molluscan SSFs would have been stem conchiferans, or less likely, belonged to the stem of Monoplacophora or the lineage that gave rise to the remaining conchiferans. As noted above, at least some fossil aculiferans had a single shell; at least some SSFs could conceivably have been aculiferans. Surprisingly, the split of gastropods and scaphopods is rather late according to our molecular clock analysis (474 MYA; 95% HDP = 479–520 MYA); this could mean that many Cambrian shells currently regarded to be gastropods were actually members of the gastropod-scaphopod stem lineage.

In conclusion, we analysed a high-quality and representative molluscan phylogenomic dataset and recovered a robust and intriguing hypothesis on molluscan class-level relationships. Analyses employing site-heterogeneous models and a coalescent approach provide support for a dichotomy dividing the molluscs into Aculifera and Conchifera, the latter with Monoplacophora sister to the rest of uni- or bivalved molluscs and gastropods sister to scaphopods, not bivalves. Our results contradict hypotheses such as Testaria, Serialia, and Monoplacophora + Cephalopoda, and have important consequences for reconstructing early molluscan evolution.

Methods

Molecular laboratory work. One specimen of *Laevipilina antarctica* (ZSM-Mol-20090330, DNABANK-Mol-MS-016) was collected with the *R/V Polarstern* in Antarctica between 70°24.00'S, 8°19.72'W and 70°23.86'S, 8°18.68'W at 597–602 m depth on 12 January 2008. DNA was extracted from the specimen using the NucleoSpin Tissue Kit (Macherey-Nagel, Düren, Germany). DNA (10 ng) was used for whole genome amplification using the Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare Life Sciences, Freiburg, Germany) followed by standard ethanol precipitation and re-purification using the Qiagen MinElute system (Qiagen, Hilden, Germany). Concentration was determined using a Qubit 2.0 Fluorometer, and 1 µg was used to create a sequencing library with the TruSeq DNA Sample Preparation Kit v2 (Illumina, San Diego, CA, USA) with an average insert size of

approximately 250 bp. Two lanes of 101 bp paired-end-reads were sequenced on the Illumina HiSeq 2000 system yielding about 90 Gbp. Raw reads were filtered for quality, PCR duplicates, and adapter sequences and corrected using SOAPfilter_v2.0 (https://github.com/tanghaibao/jcvi-bin/blob/master/SOAP/SOAPfilter_v2.0) using default settings.

Genome assembly and annotation. Reads retained by SOAPfilter_v2.0 were assembled *de novo* using SOAPdenovo2_v2.04⁶¹. Sparse_pregraph was used to construct the K-mer graph using the following settings: -K 31 -g 15 -z 2000000000 -d 1 -e 1 -r 0 -p 28. Contigs were computed using kmer iterations up to K = 63 (-M 3 -m 63 -p 30). The remapping step of SOAPdenovo was carried out using standard settings and the scaffolding step was used with parameters: -F -G 200 -p 28. Finally, additional gaps were filled using SOAP Gapcloser v1.12. Genescan⁶² was used to generate gene predictions resulting in 83 Mb of protein-coding sequences, which were subsequently used for phylogenomic analyses.

Taxon sampling and data preparation for phylogenomic analysis. Taxon sampling (Supplementary Table 1) was selected to broadly span the diversity of Mollusca including at least two representatives of each major lineage and at least two representatives of each phylum considered a candidate for the sister taxon of Mollusca⁶³. Publicly available protein sequences from complete genomes and assembled transcriptomes were downloaded when available. Dataset assembly and processing built on our established and routinely used bioinformatic pipeline^{30,64–67} with a number of modifications to help reduce possible exogenous contamination and low quality data (e.g., incorrectly translated gene predictions from Genescan; see below). Unassembled publicly available transcriptome data were digitally normalized and assembled using Trinity⁶⁸. Transcriptome assemblies were translated with TransDecoder (<https://sourceforge.net/p/transdecoder/>), keeping only amino acid (AA) sequences longer than 100 AAs.

Orthology inference. For orthology inference, we employed HaMStR 13⁶⁹, which infers orthology based on predefined sets of orthologous groups (OGs). We employed the Trochozoa custom core-ortholog set of Kocot *et al.*³⁰. Translated transcripts for all taxa were then searched against the 2,259 Trochozoa pHMMs. Sequences matching an OG's pHMM were then compared to the proteome of *Lottia gigantea* using BLASTP⁷⁰ with the -strict option. If the *Lottia* amino acid sequence contributing to the pHMM was the best BLASTP hit in each of these back-BLASTs, the sequence was then assigned to that OG.

Dataset processing. Sequences shorter than 100 amino acids were deleted and OGs sampled for fewer than 35 taxa were discarded. Redundant identical sequences were removed with UniQhaplo (<http://raven.iab.alaska.edu/~ntakebay/>). In cases where one of the first or last 20 characters of an amino acid sequence was an X, all characters between the X and that end of the sequence were deleted and treated as missing data. Each OG was then aligned with MAFFT⁷¹ (mafft-auto-localpair-maxiterate 1000). Alignments were then trimmed with Aliscore⁷² and Alicut⁷³ to remove ambiguously aligned regions. Next, a consensus sequence was inferred for each alignment using the EMBOSS program infoalign⁷⁴. For each sequence in each single-gene amino acid alignment, the percentage of positions of that sequence that differed from the consensus of the alignment were calculated using the infoalign's "change" calculation. Any sequence with a "change" value greater than 75 was deleted. Subsequently, a custom script (AlignmentCompare; https://github.com/kmkocot/basal_metazoan_phylogenomics_scripts_01-2015) was used to delete any likely mistranslated sequence regions of 20 or fewer amino acids in length surrounded by ten or more gaps on either side. Next, alignment columns with fewer than four non-gap characters were deleted. At this point, alignments shorter than 50 amino acids in length were discarded. Lastly, sequences that did not overlap with all other sequences in the alignment by at least 20 amino acids were deleted, starting with the shortest sequences not meeting this criterion.

In some cases, a taxon was represented in an OG by two or more sequences (splice variants, lineage-specific gene duplications [=inparalogs], overlooked paralogs, or exogenous contamination). In order to select the best sequence for each taxon and exclude any paralogs or exogenous contamination, we built trees in FastTree 2⁷⁵ and used PhyloTreePruner⁷⁶ to select the best sequence for each taxon. OGs sampled for fewer than 35 taxa and OGs lacking a sequence from *Laevipilina antarctica* were discarded. The remaining alignments were manually screened to identify and remove putative contamination or mistranslated sequences. Sequences that were obviously very different from the majority of the sequences in the alignment were blasted against NCBI NR using BLASTP and sequences that did not return an animal as the top hit were discarded. Finally, remaining OGs were then concatenated using FASconCAT⁷⁷.

Phylogenetic analyses. Maximum likelihood analyses were conducted in RAxML 8.2.4³⁴ and IQ-TREE 1.5.5³⁵. For the RAxML analysis, matrices were partitioned by gene with the PROTGAMMAAUTO model (the best-fitting model for each gene) used for all partitions. The tree with the best likelihood score after 10 random addition sequence replicates was retained and topological robustness (i.e., nodal support) was assessed with 100 replicates of fast bootstrapping (the -f a command line option was used). For the IQ-TREE analysis, we used the posterior mean site frequency (PMSF) model³⁷, which is an approximation to full empirical profile mixture models for ML analysis. Specifically, the LG + C60 + G + F model was specified. Because this approach requires a guide tree to infer the site frequency model, we used the previously generated RAxML tree. Nodal support was assessed with 1000 replicates of ultrafast bootstrapping (-bb 1000). Bayesian Inference analysis was conducted with PhyloBayes 4.1b⁷⁸ using the site-heterogeneous CAT-GTR model. Two chains were run for 14,143 and 13,400 generations, respectively with the first 2,000 trees from each chain discarded as burn-in. A bpcomp maxdiff value of 0.28 indicated that the chains had converged.

To examine support for key hypotheses from individual partitions, we made trees for each gene in RAxML using the best-fitting model, used these as guide trees for IQ-TREE analyses with the LG + C20 + G + F model,

and inferred a supertree using a multi-species coalescent model in ASTRAL 5.6.1⁴⁰. Weakly-supported nodes (bs < 50) were collapsed as advocated by Zhang *et al.*⁴⁰. Hypothesis testing using the Approximately Unbiased test⁷⁹ and the Shimodaira Hasegawa test⁸⁰ was conducted using RAxML 8.2.4³⁴ and CONSEL⁸¹ based on the RAxML analysis.

Divergence time estimates (Supplementary Table 3) were obtained in BEAST2 v.2.4.6⁸² on the CIPRES Science Gateway (<https://www.phylo.org/>) with a log-normal relaxed clock and the WAG model of substitution. The topology of the tree was manually constrained *a priori* by defining the major splits of the BI tree analysed herein. Fossil calibrations^{83–89} are presented in Supplementary Table 4. The analysis was executed for 180 million generations sampling a tree every 1,000 generations. After discarding the first 3,600 trees as burn-in, 14,401 trees were analysed with TreeAnnotator 2.4.5 to build the summary tree.

Ancestral character state reconstruction. Ancestral character state reconstruction was performed previously by Kocot *et al.*²⁰ using an updated and modified version of the morphological matrix of Haszprunar²¹. Because this analysis was already performed in light of numerous alternative hypotheses of molluscan class-level phylogeny including Monoplacophora sister to the remainder of Conchifera, it was not re-done here. The data matrix analysed is available via FigShare at <https://figshare.com/s/934e61a053acd8d37c1>.

Data availability

Illumina paired-end genomic data for *L. antarctica* were submitted to NCBI SRA under accession number SRR6506080. The assembled *L. antarctica* genome, assembly statistics, Genescan output, molecular and morphological data matrices analysed, and other data files associated with results presented herein were submitted to FigShare: <https://figshare.com/s/934e61a053acd8d37c1>. Sources of publicly available datasets used herein are presented in Supplementary Table 1.

Received: 9 July 2019; Accepted: 12 December 2019;

Published online: 09 January 2020

References

- Ponder, W. & Lindberg, D. Molluscan Evolution and Phylogeny: An Introduction. In *Phylogeny and Evolution of the Mollusca* 1–17 (University of California Press, 2008).
- Haszprunar, G., Schander, C. & Halaných, K. M. Relationships of Higher Molluscan Taxa. In *Phylogeny and Evolution of the Mollusca* 19–32 (University of California Press, 2008).
- Schrödl, M. & Stöger, I. A review on deep molluscan phylogeny: old markers, integrative approaches, persistent problems. *J. Nat. Hist.* **48**, 2773–2804 (2014).
- Salvini-Plawen, L. V. On the origin and evolution of the Mollusca. *Atti Conv. Lincei* **49**, 235–293 (1981).
- Salvini-Plawen, L. V. & Steiner, G. The Testaria concept (Polyplacophora+ Conchifera) updated. *J. Nat. Hist.* **48**, 2751–2772 (2014).
- Sigwart, J. D. & Lindberg, D. R. Consensus and confusion in molluscan trees: evaluating morphological and molecular phylogenies. *Syst. Biol.* **64**, 384–395 (2014).
- Telford, M. J. Mollusc evolution: seven shells on the sea shore. *Curr. Biol.* **23**, R952–R954 (2013).
- Telford, M. J. & Budd, G. E. Invertebrate evolution: bringing order to the molluscan chaos. *Curr. Biol.* **21**, R964–R966 (2011).
- Wanninger, A. & Wollesen, T. The Evolution of Molluscs. *Biol. Rev.* **94**, 102–115 (2018).
- Runnegar, B. & Pojeta, J. Jr. Molluscan phylogeny: the paleontological viewpoint. *Science* **186**, 311–317 (1974).
- Pojeta, J. & Runnegar, B. The paleontology of rostroconch mollusks and the early history of the phylum Mollusca. *US Geol. Surv. Prof. Pap.* **986**, 1–88 (1976).
- Salvini-Plawen, L. v. Early evolution and the primitive groups. In *The Mollusca vol. 10 Evolution* 59–150 (Academic Press, 1985).
- Trueman, E. R. & Brown, A. C. The mechanism of shell elevation in Haliotis (Mollusca: Gastropoda) and a consideration of the evolution of the hydrostatic skeleton in Mollusca. *J. Zool.* **205**, 585–594 (1985).
- Salvini-Plawen, L. V. & Steiner, G. Synapomorphies and plesiomorphies in higher classification of Mollusca. In *Origin and Evolutionary Radiation of the Mollusca* 29–51 (Oxford University Press, 1996).
- Philippe, H. & Roure, B. What does phylogenomics really tell us about molluscan evolution? In *Abstracts of the 2012 Society for the Study of Evolution Meeting* (2012).
- Giribet, G. *et al.* Evidence for a clade composed of molluscs with serially repeated structures: Monoplacophorans are related to chitons. **103**, 7723–7728 (2006).
- Wilson, N. G., Rouse, G. W. & Giribet, G. Assessing the molluscan hypothesis Serialia (Monoplacophora+ Polyplacophora) using novel molecular data. *Mol. Phylogenet. Evol.* **54**, 187–193 (2010).
- Vinther, J., Sperling, E. A., Briggs, D. E. & Peterson, K. J. A molecular palaeobiological hypothesis for the origin of aplacophoran molluscs and their derivation from chiton-like ancestors. *Proc. R. Soc. B Biol. Sci.* **279**, 1259–1268 (2012).
- Smith, S. A. *et al.* Resolving the evolutionary relationships of molluscs with phylogenomic tools. *Nature* **480**, 364–367 (2011).
- Kocot, K. M. *et al.* Phylogenomics reveals deep molluscan relationships. *Nature* **477**, 452–456 (2011).
- Haszprunar, G. Is the Aplacophora monophyletic? A cladistic point of view. *Am. Malacol. Bull.* **15**, 115–130 (2000).
- Kocot, K. M. Recent advances and unanswered questions in deep molluscan phylogenetics. *Am. Malacol. Bull.* **31**, 1–14 (2013).
- Haszprunar, G. Monoplacophora (Tryblidia). In *Phylogeny and Evolution of the Mollusca* 97–104 (University of California Press, 2008).
- Yochelson, E. L. Alternative approach to interpretation of phylogeny of ancient mollusks. *Malacologia* **17**, 165–191 (1978).
- Yochelson, E. L. Early radiation of Mollusca and mollusc-like groups. *Orig. Major Invertebr. Groups* **12**, 323–358 (1979).
- Yochelson, E. L., Flower, R. H. & Webers, G. F. The bearing of the new Late Cambrian monoplacophoran genus *Knightoconus* upon the origin of the Cephalopoda. *Lethaia* **6**, 275–309 (1973).
- Runnegar, B. Early evolution of the Mollusca: The fossil record. In *Origin and Evolutionary Radiation of the Mollusca* 77–87 (Oxford University Press, 1996).
- Zapata, F. *et al.* Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. *Proc. R. Soc. B Biol. Sci.* **281**, 20141739 (2014).
- González, V. L. *et al.* A phylogenetic backbone for Bivalvia: an RNA-seq approach. *Proc. R. Soc. B Biol. Sci.* **282**, 20142332 (2015).
- Kocot, K. M. *et al.* Phylogenomics of Lophotrochozoa with consideration of systematic error. *Syst. Biol.* **66**, 256–282 (2017).
- Vinther, J. The origins of molluscs. *Palaeontology* **58**, 19–34 (2015).
- Parkhaev, P. Y. Origin and the early evolution of the phylum Mollusca. *Paleontol. J.* **51**, 663–686 (2017).
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).

34. Stamatakis, A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
35. Nguyen, L.-T., Schmidt, H. A., Haeseler, Avon & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
36. Chernomor, O., von Haeseler, A. & Minh, B. Q. Terrace aware data structure for phylogenomic inference from supermatrices. *Syst. Biol.* **65**, 997–1008 (2016).
37. Wang, H.-C., Minh, B. Q., Susko, E. & Roger, A. J. Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Syst. Biol.* **67**, 216–235 (2017).
38. Lartillot, N., Rodrigue, N., Stubbs, D. & Richer, J. PhyloBayes MPI. Phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* **62**, 611–615 (2013).
39. Lartillot, N. & Philippe, H. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21**, 1095 (2004).
40. Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**, 153 (2018).
41. Moroz, L. L. On the independent origins of complex brains and neurons. *Brain. Behav. Evol.* **74**, 177–190 (2009).
42. Sumner-Rooney, L. & Sigwart, J. D. Do chitons have a brain? New evidence for diversity and complexity in the polyplacophoran central nervous system. *J. Morphol.* **279**, 936–949 (2018).
43. Sigwart, J. D. *et al.* A new sensory organ in “primitive” molluscs (Polyplacophora: Lepidopleurida), and its context in the nervous system of chitons. *Front. Zool.* **11**, 7 (2014).
44. Scherholz, M., Redl, E., Wollesen, T., Todt, C. & Wanninger, A. Aplacophoran mollusks evolved from ancestors with polyplacophoran-like features. *Curr. Biol.* **23**, 2130–2134 (2013).
45. Stöger, I. *et al.* The continuing debate on deep molluscan phylogeny: evidence for serialia (Mollusca, Monoplacophora). *BioMed Res. Int.* **2013**, 407072 (2013).
46. Smith, S. A. *et al.* Corrigendum: Resolving the evolutionary relationships of molluscs with phylogenomic tools. *Nature* **493**, 708–708 (2013).
47. Plate, L. H. Über den Bau und die Verwandtschaftsbeziehungen der Solenoconchen. *Zool. Jahrb. Anat.* **5**, 301–386 (1892).
48. Sigwart, J. D., Sumner-Rooney, L. H., Dickey, J. & Carey, N. The scaphopod foot is ventral: more evidence from the anatomy of *Rhabdus rectus* (Carpenter, 1864) (Dentaliida: Rhabdidae). *Molluscan Res.* **37**, 79–87 (2017).
49. Bieler, R. *et al.* Investigating the Bivalve Tree of Life—an exemplar-based approach combining molecular and novel morphological characters. *Invertebr. Syst.* **28**, 32–115 (2014).
50. Osca, D., Irisarri, I., Todt, C., Grande, C. & Zardoya, R. The complete mitochondrial genome of *Scutopus ventrolineatus* (Mollusca: Chaetodermomorpha) supports the Aculifera hypothesis. *BMC Evol. Biol.* **14**, 197 (2014).
51. Scheltema, A. H. Aplacophora as progenetic aculiferans and the coelomate origin of mollusks as the sister taxon of Sipuncula. *Biol. Bull.* **184**, 57–78 (1993).
52. Vinther, J., Parry, L., Briggs, D. E. & Van Roy, P. Ancestral morphology of crown-group molluscs revealed by a new Ordovician stem aculiferan. *Nature* **542**, 471 (2017).
53. Scheltema, A. & Schander, C. Exoskeletons: tracing molluscan evolution. *Venus* **65**, 19–26 (2006).
54. Erwin, D. H. *et al.* The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science* **334**, 1091–1097 (2011).
55. Dohrmann, M. & Wörheide, G. Dating early animal evolution using phylogenomic data. *Sci. Rep.* **7**, 3599 (2017).
56. Fedonkin, M. A. & Waggoner, B. M. The Late Precambrian fossil *Kimberella* is a mollusc-like bilaterian organism. *Nature* **388**, 868–871 (1997).
57. Fedonkin, M. A., Simonetta, A. & Ivantsov, A. Y. New data on *Kimberella*, the Vendian mollusc-like organism (White Sea region, Russia): palaeoecological and evolutionary implications. *Geol. Soc. Lond. Spec. Publ.* **286**, 157–179 (2007).
58. Ivantsov, A. Y. New reconstruction of *Kimberella*, problematic Vendian metazoan. *Paleontol. J.* **43**, 601–611 (2009).
59. Vinther, J. A molecular palaeobiological perspective on aculiferan evolution. *J. Nat. Hist.* **48**, 2805–2823 (2014).
60. Parkhaev, P. Y. The early Cambrian radiation of Mollusca. In *Phylogeny and Evolution of the Mollusca* 33–70 (University of California Press, 2008).
61. Luo, R. *et al.* SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* **1**, 18 (2012).
62. Burge, C. & Karlin, S. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78–94 (1997).
63. Kocot, K. M. On 20 years of Lophotrochozoa. *Org. Divers. Evol.* **16**, 329–343 (2016).
64. Hall, M. R. *et al.* The crown-of-thorns starfish genome as a guide for biocontrol of this coral reef pest. *Nature* **544** (2017).
65. Kocot, K. M., Tassia, M. G., Halanych, K. M. & Swalla, B. J. Phylogenomics offers resolution of major tunicate relationships. *Mol. Phylogenet. Evol.* **121**, 166–173 (2018).
66. Smythe, A. B., Holovachov, O. & Kocot, K. M. Improved phylogenomic sampling of free-living nematodes enhances resolution of higher-level nematode phylogeny. *BMC Evol. Biol.* **19**, 121 (2019).
67. Kocot, K. M., Todt, C., Mikkelsen, N. T. & Halanych, K. M. Phylogenomics of Aplacophora (Mollusca, Aculifera) and a solenogaster without a foot. *Proc. R. Soc. B* **286**, 20190115 (2019).
68. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
69. Ebersberger, I., Strauss, S. & Von Haeseler, A. HaMStR: profile hidden Markov model based search for orthologs in ESTs. *BMC Evol. Biol.* **9**, 157 (2009).
70. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
71. Katoh, K., Kuma, K., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518 (2005).
72. Misof, B. & Misof, K. A monte carlo approach successfully identifies randomness in multiple sequence alignments: a more objective means of data exclusion. *Syst. Biol.* **58**, 21–34 (2009).
73. Kück, P. ALICUT: a PerlScript which cuts ALIScore identified RSS. *Dep. Bioinforma. Zool. Forschungsmuseum Koenig ZFMK Bonn Ger. Version 2* (2009).
74. Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European molecular biology open software suite. *Trends Genet.* **16**, 276–277 (2000).
75. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 – approximately maximum-likelihood trees for large alignments. *Plos One* **5**, e9490 (2010).
76. Kocot, K. M., Citarella, M. R., Moroz, L. L. & Halanych, K. M. PhyloTreePruner: a phylogenetic tree-based approach for selection of orthologous sequences for phylogenomics. *Evol. Bioinformatics* **9**, 429 (2013).
77. Kück, P. & Meusemann, K. FASconCAT: Convenient handling of data matrices. *Mol. Phylogenet. Evol.* **56**, 1115–1118 (2010).
78. Lartillot, N., Lepage, T. & Blanquart, S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* **25**, 2286 (2009).
79. Shimodaira, H. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* **51**, 492–508 (2002).
80. Shimodaira, H. & Hasegawa, M. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**, 1114–1116 (1999).

81. Shimodaira, H. & Hasegawa, M. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**, 1246–1247 (2001).
82. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).
83. Passamaneck, Y. J., Schander, C. & Halanych, K. M. Investigation of molluscan phylogeny using large-subunit and small-subunit nuclear rRNA sequences. *Mol. Phylogenet. Evol.* **32**, 25–38 (2004).
84. Meyer, A., Todt, C., Mikkelsen, N. & Lieb, B. Fast evolving 18S rRNA sequences from Solenogastres (Mollusca) resist standard PCR amplification and give new insights into mollusk substitution rate heterogeneity. *BMC Evol. Biol.* **10**, 70 (2010).
85. Stöger, I. & Schrödl, M. Mitogenomics does not resolve deep molluscan relationships (yet?). *Mol. Phylogenet. Evol.* **69**, 376–392 (2013).
86. Nishiguchi, M. K. & Mapes, R. H. Cephalopoda. *Phylogeny Evol. Mollusca* 163–199 (2008).
87. Vendrasco, M. J. & Runnegar, B. Late Cambrian and early Ordovician stem group chitons (Mollusca: Polyplacophora) from Utah and Missouri. *J. Paleontol.* **78**, 675–689 (2004).
88. Nützel, A., Erwin, D. H. & Mapes, R. H. Identity and phylogeny of the late Paleozoic Subulitoidea (Gastropoda). *J. Paleontol.* **74**, 575–598 (2000).
89. Yochelson, E. L. Scaphopoda. In *Functional Morphology of the Invertebrate Skeleton* 363–367 (Wiley, 1999).

Acknowledgements

Laevipilina antarctica was collected during the ANDEEP-SYSTCO expedition on R/V Polarstern; thanks go to the AWI, the Captain and the Crew, to the cruise leader Prof. Angelika Brandt (Frankfurt), and to Enrico Schwabe (ZSM) for facilitating specimen collection. We thank Peter Kohnert and Katharina Jörger for providing photographs. We thank Rebecca Varney and Meghan Yap Chiongo for providing helpful feedback on an earlier version of this manuscript. This project was supported by the Deep Metazoan and the Antarctic Priority Programmes of the German Research Foundation (DFG SCHR667/9-1 and SCHR667/15-1), an LMU graduate stipend to I.S., and NSF DEB 1845174 and University of Alabama start-up funds to K.M.K. This is Auburn University Marine Biology Program contribution #194 and Molette Lab contribution #95.

Author contributions

M.S. and A.J.P. conceived the project. M.S. collected the specimen of *L. antarctica*. A.J.P. prepared sequencing libraries, assembled and annotated the genome, and conducted other bioinformatic analyses. K.M.K. conducted dataset preparation for phylogenomic analysis, orthology inference, dataset processing, and most phylogenetic analyses. I.S. conducted molecular clock analyses. K.M.K., M.S., K.M.H., I.S. and A.J.P. wrote the manuscript and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-56728-w>.

Correspondence and requests for materials should be addressed to K.M.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020