

Research



Cite this article: Drábková M *et al.* 2022

Different phylogenomic methods support monophyly of enigmatic 'Mesozoa' (Dicyemida + Orthonectida, Lophotrochozoa).

Proc. R. Soc. B **289**: 20220683.

<https://doi.org/10.1098/rspb.2022.0683>

Received: 8 April 2022

Accepted: 10 June 2022

Subject Category:

Evolution

Subject Areas:

evolution, genomics, taxonomy and systematics

Keywords:

Mesozoa, phylogenomics, Orthonectida, Dicyemida, Lophotrochozoa, LBA

Author for correspondence:

Marie Drábková

e-mail: marie.drabkova7@gmail.com

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6070258>.

Different phylogenomic methods support monophyly of enigmatic 'Mesozoa' (Dicyemida + Orthonectida, Lophotrochozoa)

Marie Drábková^{1,3}, Kevin M. Kocot⁴, Kenneth M. Halanych⁵, Todd H. Oakley⁶, Leonid L. Moroz⁷, Johanna T. Cannon⁶, Armand Kuris⁶, Ana Elisa Garcia-Vedrenne⁶, M. Sabrina Pankey⁸, Emily A. Ellis⁶, Rebecca Varney^{4,6}, Jan Štefka^{1,3} and Jan Zrzavý²

¹Department of Parasitology, and ²Department of Zoology, Faculty of Science, University of South Bohemia, České Budějovice 37005, Czech Republic

³Laboratory of Molecular Ecology and Evolution, Institute of Parasitology, Biology Centre CAS, České Budějovice 37005, Czech Republic

⁴Department of Biological Sciences, The University of Alabama, Campus Box 870344, Tuscaloosa, AL 35487, USA

⁵The Centre for Marine Science, University of North Carolina, Wilmington, 57000 Marvin K. Moss Lane, Wilmington, NC 28409, USA

⁶Department of Ecology, Evolution, and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106, USA

⁷Department of Neuroscience, and the Whitney Laboratory for Marine Bioscience, University of Florida, 9505 Ocean Shore Boulevard, St Augustine, FL 32080, USA

⁸Molecular, Cellular and Biomedical Sciences, University of New Hampshire, Durham, NH 03824, USA

MD, 0000-0003-1389-3605; KMK, 0000-0002-8673-2688; KMH, 0000-0002-8658-9674; THO, 0000-0002-4478-915X; LLM, 0000-0002-1333-3176; AEG-V, 0000-0002-3518-1564

Dicyemids and orthonectids were traditionally classified in a group called Mesozoa, but their placement in a single clade has been contested and their position(s) within Metazoa is uncertain. Here, we assembled a comprehensive matrix of Lophotrochozoa (Metazoa) and investigated the position of Dicyemida (= Rhombozoa) and Orthonectida, employing multiple phylogenomic approaches. We sequenced seven new transcriptomes and one draft genome from dicyemids (*Dicyema*, *Dicyemene*) and two transcriptomes from orthonectids (*Rhopalura*). Using these and published data, we assembled and analysed contamination-filtered datasets with up to 987 genes. Our results recover Mesozoa monophyletic and as a close relative of Platyhelminthes or Gnathifera. Because of the tendency of the long-branch mesozoans to group with other long-branch taxa in our analyses, we explored the impact of approaches purported to help alleviate long-branch attraction (e.g. taxon removal, coalescent inference, gene targeting). None of these were able to break the association of Orthonectida with Dicyemida in the maximum-likelihood trees. Contrastingly, the Bayesian analysis and site-specific frequency model in maximum-likelihood did not recover a monophyletic Mesozoa (but only when using a specific 50 gene matrix). The classic hypothesis on monophyletic Mesozoa is possibly reborn and should be further tested.

1. Introduction

Phylogenomic studies have dramatically improved our understanding of deep metazoan phylogeny, however, some key branches remain controversial (e.g. [1–9]). Incongruence in phylogenomic analyses could be owing to insufficient phylogenetic signal (e.g. owing to closely spaced branching events), limited taxon and/or gene sampling, or systematic error, particularly the long-branch attraction (LBA) artefact ([10]; reviewed by [11,12]).

The early radiation of Lophotrochozoa (549–535 Ma; [13]), a major clade of protostome metazoans including annelids, molluscs, brachiopods, bryozoans, flatworms and rotifers among others, has been especially challenging to uncover [5,8,9,14,15]. This is probably owing to relatively rapid diversification [16] and fast evolutionary rates in some lineages, leading to long terminal branches in phylogenetic analyses possibly prone to LBA. Differences in branch lengths can be caused by differences in the rate of molecular evolution and generation time, as both are known to be variable for invertebrate phyla [17], especially in parasites [18]. This may aggravate the inherent difficulty of resolving ancient radiations because of short internal branches among deep nodes [19].

The placement of Orthonectida and Dicyemida, two clades of enigmatic morphologically simple parasites, could be seen as examples of such a controversial case (for images of mesozoans, see electronic supplementary material, figure S1). Orthonectids (*ca* 25 spp. [20]) are parasites of a variety of marine invertebrates; their adults are free-swimming and produce ciliated larvae that enter the host tissues to form amoeboid trophic syncytia that effectively castrate their hosts [21]. Dicyemida (syn. Rhombozoa, used especially if including still unsequenced Heterocyemida; *ca* 100 spp. [22]) are symbionts inhabiting the renal organs of benthic cephalopods. Dicyemids were originally placed in Mesozoa by Van Beneden [23] who viewed them as a 'link' between Protozoa and Metazoa. By contrast, a hypothesis on mesozoan convergent evolution from more complex animals prevails in modern literature, even though their phylogenetic origins remain uncertain. Some early workers (reviewed in [24]) suggested mesozoans might be simplified platyhelminths, based on similarities in the life cycle and a putative resemblance of the mesozoan body structure to trematode miracidium larvae. Metschnikoff [25] derived orthonectids from annelids such as *Dinophilus*, which by themselves are secondarily simplified [26]. More recently, Slyusarev & Kristensen [27] found that external cells of Orthonectida are covered by a thin cuticle, which provided some support for the 'annelid hypothesis' of orthonectid origin.

Kozloff [28, p. 193], who worked on both mesozoan groups, viewed the similarities between orthonectids and dicyemids as superficial and stated that 'we do not know how either group has evolved, but it does seem clear that they are not closely related'. Phylogenetic analysis of complete 18S rRNA sequences recovered orthonectids and dicyemids as two separate clades of early branching animals [29], while analyses of morphology combined with 18S rRNA recovered Mesozoa as a monophyletic group close to the base of Bilateria [30]. Other attempts based on single genes supported a simplified bilaterian scenario for Dicyemida [31–33], consistently grouping it within Lophotrochozoa.

Phylogenomic analyses by Mikhailov *et al.* [34] (22 909 amino acid positions, 61 taxa), including the first published mesozoan genome (*Intoshia*, Orthonectida) recovered Orthonectida as a clade of uncertain position within Lophotrochozoa. Lu *et al.* [35] conducted phylogenomic analyses (58 124 amino acid positions, 29 taxa) with two genomes from both Orthonectida (*Intoshia*) and Dicyemida (*Dicyema*) and recovered Mesozoa as a monophyletic group close to gastrotrochs and/or platyhelminths. On the contrary, Schiffer *et al.* [36] proposed, based on both mitochondrial (2969 amino acid positions, 69 taxa) and nuclear (190 027 amino acid

positions, 45 taxa) data, that Mesozoa is not monophyletic: orthonectid *Intoshia* was recovered as nested within Annelida while the position of Dicyemida (three spp. of *Dicyema*) within Lophotrochozoa was considered unresolved. Zverkov *et al.* [15] reported similar results in their analysis of nuclear protein-coding genes (87 610 amino acid positions, 73 taxa including *Intoshia* and two spp. of *Dicyema*). Taken together, the phylogenomic analyses to date (and most molecular analyses in general) have converged to the scenario that both mesozoan groups represent secondarily simplified lophotrochozoans, but their phylogenetic position within Lophotrochozoa and whether or not they form a monophyletic group remain unresolved. Difficulties in the placement of these enigmatic marine groups may be owing to their fast rate of evolution, as evidenced by extremely long branches of both mesozoan groups in all analyses.

Multiple approaches exist to tackle systematic errors in phylogenomics. One straightforward way to treat LBA and test leaf stability is to remove other long branches, thereby decreasing the 'pull' of these long branches and compare the resulting position of the taxon in question (e.g. the position of Orthonectida after exclusion of Dicyemida from the matrix, and vice versa [36]). However, if more than two long-branch taxa are present, interpretation of the results becomes difficult. The site-specific models of sequence evolution, such as the CAT + GTR model [37] which is implemented in a Bayesian framework in PHYLONOTES [38], and the PMSF model [39], implemented in a maximum-likelihood (ML) framework in IQ-TREE [40], have been purported to reduce LBA by modelling the actual complexity of the data resulting from biological processes. Other approaches thought to help mitigate LBA are based on selecting molecular markers based on properties such as branch-length homogeneity or compositional homogeneity. Balanced sequence composition has also been suggested to be a good predictor of phylogenetic signal (e.g. [41]), and taking steps to reduce compositional heterogeneity was advocated by Nesnidal *et al.* [42] in order to help overcome systematic artefacts affecting inference of relationships within Lophotrochozoa. Another approach for phylogeny reconstruction that may be useful for genomic datasets containing heterogeneous signal is the reconstruction of the species tree based on coalescent theory. This method has been argued to perform well in cases where heterogeneous single-gene trees are present in the analysed set [43,44].

Here, we sought to explore the phylogenetic position(s) of Orthonectida and Dicyemida in the context of lophotrochozoan phylogeny, focusing on extensive comparisons of results from different approaches purported to reduce LBA. To this end, we assembled a dataset with significantly increased taxon sampling for Mesozoa (genome/transcriptome data from eight dicyemids and three orthonectids) and diverse representatives of other lophotrochozoan phyla (32 spp.) that we carefully screened to exclude exogenous contamination. We constructed several data matrices from different sets of genes, selecting genes according to different criteria purported to help reduce LBA and other artefacts thought to have impacted earlier studies of lophotrochozoan phylogeny. We performed phylogenetic analyses on these datasets using the following approaches: (i) ML using the best-fitting site-homogeneous model for each gene as implemented in RAXML [45]; (ii) Bayesian inference (BI) using the CAT + GTR empirical profile mixture model to account for site-specific evolutionary rate heterogeneity [46]

as implemented in PHYLOBAYES [38]; (iii) ML inference using the site-specific PMSF profile mixture model [39] as implemented in IQ-TREE [40], enabling us to compare the use of site-specific models in both a ML and Bayesian framework; and (iv) a coalescent approach for phylogenomics as implemented in ASTRAL II [47].

2. Methods

(a) Taxon sampling

We sought to broadly sample the diversity of Lophotrochozoa with an emphasis on Dicyemida and Orthonectida, while avoiding including too many terminals in order to facilitate the computationally intensive analyses (e.g. PHYLOBAYES analyses with the site-heterogeneous CAT model). Thus, we opted to broadly sample each relevant phylum with the minimum number of taxa (typically 2–4 spp.) necessary to capture the deepest node in that phylum while using only high-quality (e.g. deeply sequenced and contamination-free) data. We constructed a dataset with seven newly sequenced dicyemid transcriptomes, one newly sequenced dicyemid genome, two newly sequenced orthonectid transcriptomes, and representatives of all other lophotrochozoan phyla, including three newly sequenced bryozoan transcriptomes (electronic supplementary material, table S1). To test the historical hypothesis that the mesozoans are closely related to flukes (Trematoda), *Schistosoma* was included as a representative of Neodermata (Platyhelminthes). Because orthonectids have been recovered as closely related to or even placed within Annelida, representatives of Palaeoannelida (*Magelona*, *Owenia*), Chaetopteriformia (*Phyllochaetopterus*), Sipunculida (*Phascolosoma*) and Pleistoannelida (*Capitella*) were included. Representatives of Cnidaria (*Nematostella*), Xenacoelomorpha (*Xenoturbella* and *Meara*), Ambulacraria (*Ptychodera* and *Strongylocentrotus*) and Ecdysozoa (*Priapulius*) were used as outgroups.

(b) Molecular laboratory techniques

To produce new transcriptome data for this study, total RNA was extracted from *Alcyonidium* sp. (Bryozoa), *Cristatella mucedo* (Bryozoa), *Pectinatella magnifica* (Bryozoa), two different collections of *Rhopalura* cf. *ophiocomae* (Orthonectida; both from ophiurid *Amphipholis squamata*), and dicyemids *Dicyema* sp. 1 (from octopus *Enteroctopus dofleini*), *Dicyema* sp. 2 (from *Octopus bimaculoides*), *Dicyemida* sp. 3 (from octopus *Megaleledone setebos*), *Dicyemida* sp. 4 (from *Octopus vulgaris*), and *Dicyemennae brevicephaloides* and *D. rossiae* (both from sepiolid *Rossia pacifica*). For assembly of dicyemid genome, individuals of *Dicyema moschatum* from octopus *Eledone moschata* were used. Dicyemid species were determined by combination of morphological characters (shape of the calotte *in vivo*, confocal fluorescent microscopy images; electronic supplementary material, figure S1) and host specificity. Specimen collection data, details of sample processing and sequence assembly (sample handling, extraction and cDNA library kits used, sequencing strategy, sequence assembly and annotation) for all newly sequenced taxa are provided in the electronic supplementary material, table S1.

(c) Contamination filtering

For the dicyemid genome assembly, only reads not mapping to the *Octopus bimaculoides* genome (NCBI GCF_001194135.1) were used in downstream analyses in order to remove host contamination (see the electronic supplementary material, S1 for details; e.g. figures S2 and S3 blobplots showing taxonomic assignment of sequences before and after decontamination step).

In the case of transcriptomes, sequences were compared to the NCBI Protein and Nucleotide databases (accessed March 2017) to check for possible contamination. The result was semi-manually evaluated, checking the similarity scores and taxonomy ID of hits. If the hits clearly originated from contamination (e.g., sequences >95% similar to bacterial sequences, parasitic protists and other nonmetazoans, or cephalopods in the case of the dicyemids), they were removed (for details see electronic supplementary material, S2). Contamination-filtered nucleotide sequences were then translated into amino acids with Transdecoder (<http://transdecoder.sf.net>; script available in electronic supplementary material, S2).

(d) Dataset construction

Translated transcripts for all taxa were searched against 2259 lophotrochozoan-specific core orthologous groups (OGs) as profile hidden Markov models (Lophotrochozoa-Kocot pHMMs [48] in HaMStR 13.2.6 [49]). Briefly, this dataset is based on genes identified to be single-copy in genomes or deeply sequenced transcriptomes from representatives of Annelida, Brachiopoda, Entoprocta, Mollusca, Nemertea, Platyhelminthes, Phoronida, and Rotifera. Sequences matching an OG's pHMM were compared to the proteome of *Lottia* using blastp. 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.

OGs were carefully filtered for missing data and checked for paralogs. The remaining sequences were aligned with MAFFT (for details see the electronic supplementary material, S3).

To reduce homoplasy that may be introduced by including distant outgroups, we excluded, for most analyses, all outgroups except the very slowly evolving ecdysozoan *Priapulius* (matrices with names ending in R have a Reduced outgroup, e.g. COMPR). We assembled several different matrices targeting genes with specific qualities such as missing data, phylogenetic signal, branch-length heterogeneity and compositional heterogeneity (electronic supplementary material, table S2, S4). We used Matrix Reduction (MARE [50]) to exclude OGs with low phylogenetic signal and to reduce missing data (a dataset called 'MARER' hereafter). OGs were sorted by branch length heterogeneity ('CxLBR' hereafter) and by compositional heterogeneity ('CxHR' hereafter) to select the best 50 and 100 OGs according to each criterion (i.e. C50LBR, C100LBR, C50HR, C100HR). We used TRESPEx 1.0 [51] and the single-gene ML trees described above to assess branch length heterogeneity of each OG and BaCoCA [52] to assess the compositional heterogeneity of each OG according to the relative composition frequency variability (RCFV) metric. We also assembled a matrix of the best 50 OGs of the MARE-reduced dataset (MARER) according to compositional heterogeneity (M50HR). Both C50LBR and MARER datasets were used for our numerous taxon removal experiments (see below) and Bayesian inference analyses using the computationally demanding CAT + GTR model. For an overview of matrices produced in this study, corresponding partition data, and a flowchart explaining data matrix production see the electronic supplementary material, S4.

(e) Phylogenomic analyses

(i) Maximum-likelihood analyses

Phylogenetic analyses of all datasets were conducted using ML with RAXML 8.2.4 [45]. Matrices were partitioned by gene (=orthologous group or OG) and the AUTO option was used to select the best-fitting model for each partition. For each ML analysis, the tree with the best likelihood score after 10 random addition sequence replicates was retained and topological robustness (i.e. nodal support) was assessed by the optimal

number of rapid bootstraps. Furthermore, some matrices were analysed with posterior mean site frequency (PMSF) model in ML (LG + C60 + G + F [39]) as implemented in IQ-TREE 1.5.3 [40], as an approximation of the CAT site-specific model [37,53].

(ii) Bayesian inference

Bayesian analyses (BI) were performed in PHYLOBAYES MPI 1.5 or 1.7 [38] using the CAT + GTR model [37,53]. Because of the computationally intensive nature of PHYLOBAYES analyses, only datasets with a highly reduced number of genes sampled were analysed (see Results). Each analysis was run simultaneously in four chains. Analyses were periodically checked for progress and convergence among chains was assessed with the bpcomp program from the PHYLOBAYES package after discarding a burn-in of 1500 cycles. Convergence was assumed if the maxdiff dropped below 0.3, as advocated in the PHYLOBAYES manual. Chains were run until convergence or for at least 12 000 generations, but often for longer if resources were available (for details see the electronic supplementary material, S8).

(iii) Comparison of maximum likelihood and Bayesian inference tree-building methods

In order to inspect differences between tree-building methods without the confounding factor of the long-branch mesozoans, we analysed the C50LBR matrix (the only BI analysis that successfully converged) by RAXML, IQ-TREE (PMSF model), and BI (CAT model). This data matrix was also modified by the exclusion of Dicyemida, Orthonectida, or both.

(f) Taxon removal experiments

To test the robustness of relationships within Lophotrochozoa, we carried out a series of taxon removal experiments based on the MARER dataset. Topological effects of the presence/absence of several clades of interest or their combinations (Platyhelminthes, Gastrotricha, Rousphozoa, Gnathifera, Gnathifera + Platyhelminthes, Entoprocta + Cyclophora, Bryozoa and Annelida, i.e. all possible sister groups of Mesozoa and/or clades containing long-branch taxa; for details see the electronic supplementary material, S5) were tested in ML (IQ-TREE).

(g) Gene signal analysis

We examined conflicting signal among genes by the difference in the log-likelihood score (difference in gene likelihood score (Δ GLS)) between two competing tree topologies (as in [54,55]). The two topologies correspond to the unconstrained tree resulting from ML analysis (T1; including monophyletic Mesozoa) and the constrained tree (T2) where Orthonectida were forced to group with Annelida. ML trees were estimated from the 202-gene (MARER) dataset under an automatically selected model per partition in RAXML. Site likelihood scores for contrasting gene topologies needed for Δ GLS computation were estimated under the LG model in RAXML. We also applied the approximately unbiased (AU) test in the software package CONSEL v. 0.20. The AU test was conducted using the multi-scale bootstrap technique based on the site-wise log-likelihood scores.

(h) Coalescent approach

Coalescent phylogenetic analysis of single-gene trees generated in RAXML (987 gene trees) was run in ASTRAL-II 4.11.1 [47] with default settings. Supports were computed by gene resampling as suggested by Simmons *et al.* [56] in 1000 replicates. The presented majority rule consensus tree was computed by PHYUTILITY [57].

(i) Tree plotting

All trees were plotted with ETE Toolkit v. 3 [58] and FIGTREE (<http://tree.bio.ed.ac.uk/software/figtree/>) and adjusted in INKSCAPE (<https://inkscape.org/nl/>) or VECTORNATOR (<https://www.vectornator.io>). All scripts used for phylogenetic analyses and for plotting trees are provided in the electronic supplementary material, S6.

3. Results

(a) Matrix construction

Our bioinformatic pipeline for orthologue clustering with strict paralogy filtering and decontamination resulted in a matrix with 49 taxa and 987 OGs that was 171 791 amino acid positions long, with 31% missing data (COMP). This set of OGs was further reduced to test the effects of analysing subsets of the best OGs in terms of phylogenetic signal, resulting in a MARE-reduced matrix (MARER) of 202 OGs, as well as branch-length heterogeneity and compositional heterogeneity, resulting in matrices composed of the best 100, and 50 OGs according to these criteria (C100LBR, C50LBR, C100HR and C50HR matrices; see the electronic supplementary material, S4).

(b) Comparison of tree reconstruction approaches

(i) Maximum likelihood

ML analysis of the complete matrix (COMP) in RAXML, using the best-fitting model for each gene, recovered Mesozoa monophyletic with maximal support and sister to Gnathostomulida with strong bootstrap support (BS = 98; electronic supplementary material, figure S4 in the electronic supplementary material, S7). Overall, Lophotrochozoa was split into two subclades: Trochozoa (Annelida, Brachiopoda, Phoronida, Mollusca and Nemertea) and a clade we refer provisionally to as 'Platyzoa *s.l.*', which includes platyzoan as well as polyzoan phyla. Polyzoa (Bryozoa + Entoprocta-Cyclophora) and Gnathifera (Gnathostomulida + Micrognathozoa-Rotifera) were not recovered monophyletic. Removal of distant outgroup taxa in favour of the short-branch ecdysozoan *Priapulid* did not have any effect on the resulting branching order and support values were comparable (electronic supplementary material, figure S5 in S7). Thus, all outgroup taxa except *Priapulid* were excluded from subsequent analyses to decrease computational demands.

To examine the effect of using a site-heterogeneous model in a ML framework, we analysed the COMP matrix with the posterior mean site frequency model (PMSF [39]) with 60 rate categories as implemented in IQ-TREE 1.5.3 [40]. The resulting tree topology (figure 1) slightly differed from that of the RAXML analysis: when the site-specific PMSF model was used, monophyletic Mesozoa was again maximally supported but recovered sister to Platyhelminthes (BS = 90), with that clade sister to monophyletic Gnathifera (BS = 99). For comparison of site-specific models in ML and BI we also analysed a smaller matrix (C50LBR that was used also for BI analyses) with the PMSF model. The resulting tree recovered Mesozoa as two separate clades in Platyzoa *s.l.*: Orthonectida were sister to Platyhelminthes and Dicyemida sister to Gnathifera (albeit both with low support: BS = 49 and 54, respectively; electronic supplementary material, figure S6 in S7).

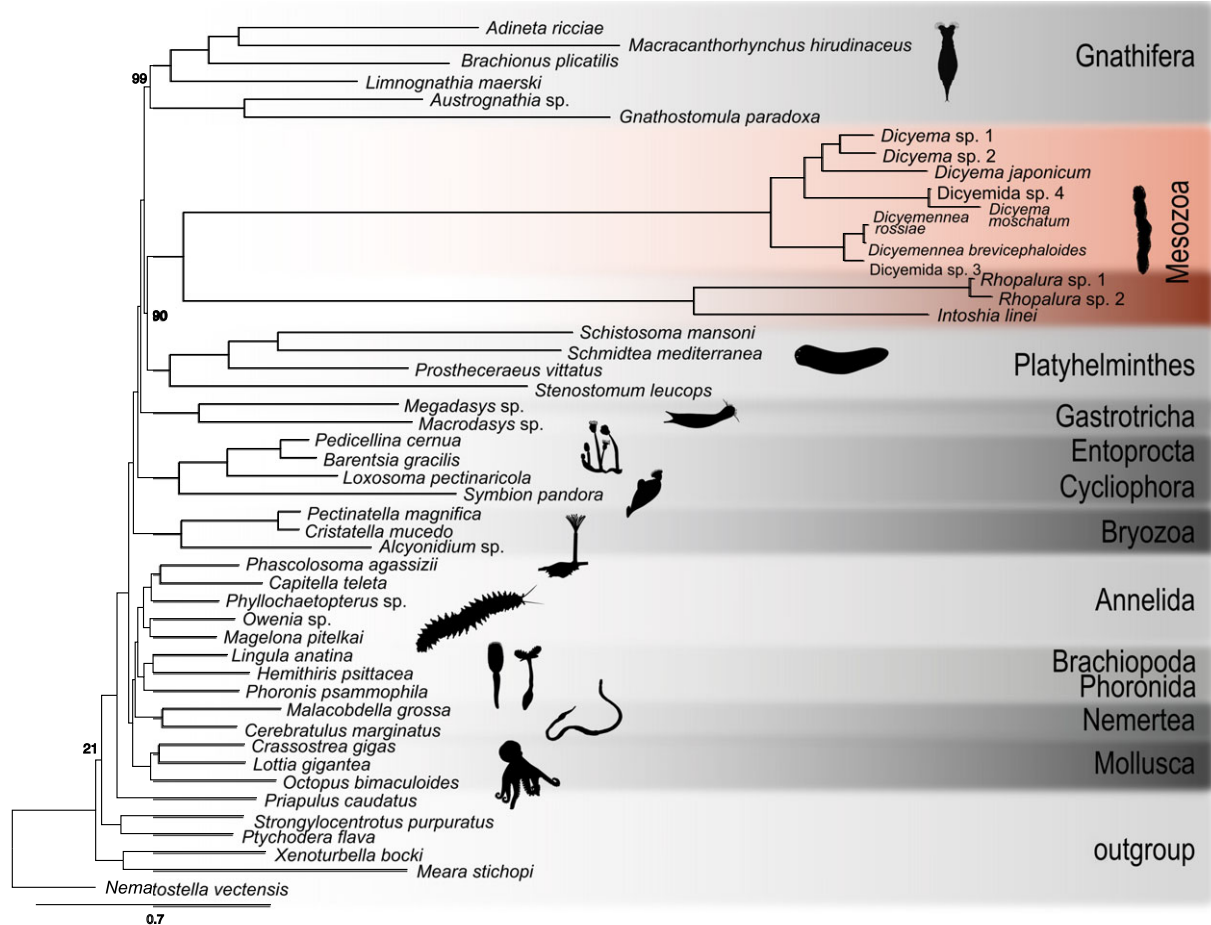


Figure 1. Tree representing mesozoan position in Lophotrochozoa based on 987 genes, computed in maximum-likelihood framework with site-specific model with 60 categories (IQ-TREE c60 PMSF COMP), only bootstrap supports lower than 100 showed. (Online version in colour.)

ML analysis of the MARE-reduced matrix (MARER; electronic supplementary material, figure S7 in S7) also recovered Mesozoa as a maximally supported monophyletic group within paraphyletic Gnathifera, as the sister group of Gnathostomulida (BS=92). The high similarity of the MARER and COMPR topologies suggests that the reduced matrix represents well the whole dataset and could substitute it where the lower number of genes is essential to carry out the analysis.

In an attempt to reduce potential sources of systematic error, we assembled datasets based on reduced subsets of the best 100 genes with the lowest branch heterogeneity (C100LBR; electronic supplementary material, figure S8 in S7) and the lowest compositional heterogeneity (C100HR; electronic supplementary material, figure S9 in S7), and then compare RAXML trees derived from both matrices. Results of both analyses recovered Mesozoa monophyletic within Platyzoa, either as a sister group of Platyhelminthes, with monophyletic Gnathifera sister to the Mesozoa–Platyhelminthes clade (C100LBR), or as a sister group of Gnathostomulida, and Platyhelminthes and Micrognathozoa–Rotifera as successive sister groups to the Mesozoa–Gnathostomulida clade (C100HR).

(ii) Bayesian inference

As the site-heterogeneous CAT + GTR model has been suggested to be more robust against long-branch attraction artefacts than site-homogeneous models [37] (but see [59]), we performed BI analyses with this model in PHYLOBAYES. Matrices with reduced numbers of genes (C50LBR,

C100LBR, C100HR, and M50HR) and just *Priapul* as the outgroup were chosen for BI, owing to the computational demands of the complex CAT + GTR model. In all of the analyses of C50LBR, the four chains converged on one solution, according to the PHYLOBAYES bpcomp maxdiff statistic [38], but the other analyses failed to converge after greater than 30 000 generations (see the electronic supplementary material S8 for details and resulting trees). The BI analysis of the C50LBR dataset (electronic supplementary material, figure S10 in S7) recovered Mesozoa polyphyletic but within monophyletic Platyzoa, with Dicyemida in a poorly supported polytomy with Gnathostomulida and Micrognathozoa–Rotifera, and Orthonectida sister to Platyhelminthes (with posterior probability (PP)=0.98, a moderate support by Bayesian standards).

(iii) Comparison of maximum likelihood and Bayesian inference tree-building methods

In order to inspect differences between tree-building methods, we analysed the C50LBR matrix (the only BI analysis that successfully converged) with or without Dicyemida and Orthonectida, respectively, by RAXML (or with similar setting in IQ-TREE with automatically selected gene models), IQ-TREE (PMSF model) and BI (CAT model in PHYLOBAYES). Both ML trees with excluded mesozoans were identical (electronic supplementary material, figures S11 and S12 in S7): Platyhelminthes and Gnathifera formed a clade, and Entoprocta–Cyclophora, Bryozoa–Gastrotricha, Nemertea, Annelida–Brachiozoa and Mollusca were

successive sister groups to it (i.e. there were monophyletic Platyzoa *s.l.* within paraphyletic Trochozoa). On the contrary, in the BI tree excl. Mesozoa (electronic supplementary material, figure S13 in S7) there were monophyletic Trochozoa and Platyzoa *s.l.*, both weakly supported, and within Platyzoa *s.l.*, a very weakly supported basalmost position of Gnathostomulida (i.e. polyphyletic Gnathifera). When only Orthonectida were included, they were recovered as a sister group of Platyhelminthes in all trees (electronic supplementary material, figures S14–S16, in S7). Dicyemida alone were recovered as a sister group of Gnathifera as a whole (IQ-TREE, electronic supplementary material, figure S17 in S7), or grouped with Gnathostomulida within monophyletic Gnathifera (RAxML, electronic supplementary material, figure S18 in S7), or in an unresolved trichotomy with Gnathostomulida and Micrognathozoa–Rotifera (BI; electronic supplementary material, figure S19 in S7). When both mesozoan groups were added to the analyses, they became diphylectic, both placed at the same positions as individual mesozoan clades alone (Dicyemida next to Gnathifera, or in a polytomy with their subclades, Orthonectida sister to Platyhelminthes; electronic supplementary material, figures S6 and S10 in S7). Only in the RAxML tree the dicyemids were apparently attracted toward orthonectids to form monophyletic Mesozoa sister to Platyhelminthes (electronic supplementary material, figure S20 in S7).

(iv) Coalescent approach

The coalescent tree (electronic supplementary material, figure S21 in S7) recovered mesozoans monophyletic with maximal support. They were nested in polytomy with Gnathostomulida, Rotifera–Micrognathozoa and Platyhelminthes with relatively high support (BS = 97). Overall, the coalescent tree showed a similar branching pattern as ML and BI analyses (Lophotrochozoa split into Platyzoa *s.l.* and Trochozoa). In this case, however, Bryozoa was not associated with Platyzoa *s.l.* as in ML and BI trees, but was recovered as an early branching clade of Trochozoa.

For comparison of topological differences in trees resulting from ML, BI and coalescent methods (matrices COMP, MARER, C50LBR) see the electronic supplementary material, table S3.

(c) Taxon removal experiments

In order to examine the effect of taxon sampling on the position of long-branch clades, we performed a series of taxon removal experiments, based on the 202-gene MARER dataset, analysed by both RAxML and IQ-TREE (electronic supplementary material, figures S22–S40 in S5). The IQ-TREE-based tree with both mesozoan groups excluded was split to Trochozoa and Platyzoa *s.l.*; the latter group included Platyhelminthes and Gnathifera as sister groups, followed by Gastrotricha, Entoprocta–Cycliophora and Bryozoa. Orthonectida alone groups with Gnathostomulida (within Gnathifera), Dicyemida alone with Platyhelminthes (or as a sister group of Catenulida within Platyhelminthes in RAxML tree). When both mesozoan clades were present they formed monophyletic groups, sister to Gnathostomulida (i.e. at the same position as Orthonectida alone). However, all deeper nodes were weakly supported. Subsequent exclusion of Platyhelminthes, Gastrotricha, Rouphozoa (Gastrotricha + Platyhelminthes), Gnathifera, Gnathifera + Platyhelminthes,

Entoprocta–Cycliophora, Bryozoa and Annelida had no effects on the tree topology. Only when Gnathostomulida were excluded, Mesozoa did not stay as a sister group of the rest of Gnathifera but transferred towards Platyhelminthes. On the contrary, removing the Micrognathozoa and Rotifera had no effect and Mesozoa remained close to the Gnathostomulida. Importantly, in all experiments Mesozoa remained a clade within Platyzoa *s.l.* Only when all Platyzoa and Dicyemida were excluded (merely 17 spp. of Trochozoa and Orthonectida plus *Priapulid* remained), Orthonectida was recovered as nested within Annelida, albeit with low support (electronic supplementary material, figure S37 in S5).

(d) Gene-conflict analyses

The likelihood score of the constrained topology T2 (Orthonectida grouping with Annelida), compared by AU test, was significantly worse than that of the unconstrained topology with monophyletic Mesozoa (T1) in the case of the MARER matrix. Out of 202 genes in the data matrix, 148 support topology T1 and only 54 support the constrained topology T2 based on Δ GLS (electronic supplementary material, figure S41 in S9). However, the RAxML tree based on genes supporting T1 includes, as usually, monophyletic Mesozoa within Gnathifera (as a sister group of Gnathostomulida), but the tree based on T2-supporting genes includes Dicyemida as a sister group of Catenulida within Platyhelminthes, Orthonectida sister to the Platyhelminthes–Dicyemida clade, followed by Gastrotricha, and monophyletic Gnathifera are sister to the whole Rouphozoa (Gastrotricha + Platyhelminthes) clade (the latter including both mesozoan groups). The T1-supporting genes are insignificantly longer than T2-compatible ones (mean 159 > 139); also taxonomic representativeness is comparable in both groups of genes (mean 36.5 : 34.5) but compositional heterogeneity is slightly higher in T1-supporting genes (T1 0.163, T2 0.149).

4. Discussion

(a) Phylogenetic relationships

Consistent with previous studies [29,35,36], all our results confirm that both Orthonectida and Dicyemida are secondarily morphologically simplified bilaterians, not a ‘primitive missing link’ between Protozoa and Metazoa as occasionally suggested earlier [23,60]. Both Orthonectida and Dicyemida were recovered as monophyletic groups belonging to Lophotrochozoa in all analyses performed in this study. In general, our results represent a ‘traditional view’ on lophotrochozoan relationships [5] with small- and simple-bodied phyla forming a clade (Platyzoa *s.l.*) apart from Trochozoa; the newly sequenced transcriptomes supported isolated position of Bryozoa next to other Platyzoa *s.l.* rather than monophyletic Lophophorata within Trochozoa [61]. Importantly, this pattern was not affected by the presence/absence of mesozoans.

Dicyemids fundamentally differ from orthonectids by having no cuticle, nervous system, muscles, or true tissues at all. Furthermore, life-histories of the two groups are extremely derived yet quite different (host-parasite relationships, host spectrum, position of parasitic stage in the life cycle, etc.). The only potentially shared ultrastructural character, a unique type of ciliary rooting [62], has been challenged by

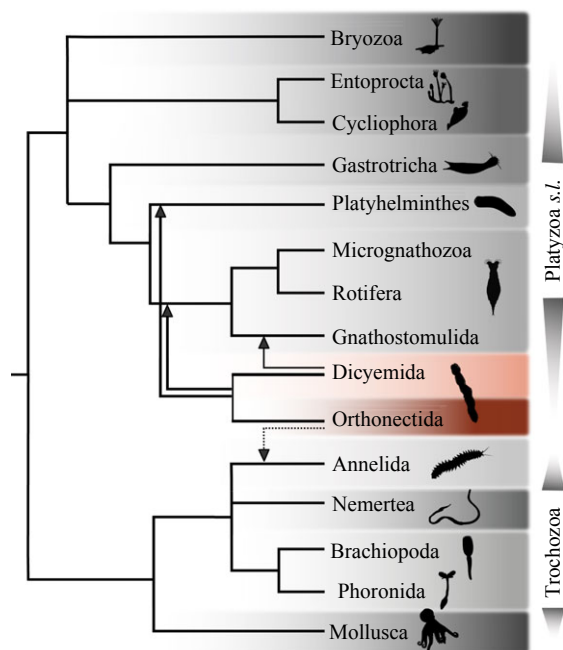


Figure 2. Overview of possible scenarios of mesozoan evolutionary history. Animal silhouette images are from Phylopic.com. (Online version in colour.)

the discovery of more diverse ciliary rootlets in Orthonectida [27]. Consequently, the monophyletic phylum Mesozoa has recently been doubted or rejected by most authors. However, the genomic data so far provided equivocal results (monophyly: [35]; polyphyly: [15,36]). Nevertheless, the old hypothesis about monophyly of Mesozoa seems to be (quite surprisingly) supported by most of our phylogenomic analyses. All ML and coalescent analyses of the complete 987-gene (COMP) and 202-gene (MARER) datasets (BI analyses did not converge) recovered Mesozoa monophyletic, usually with strong support, and close to Gnathifera or Gnathostomulida (COMP: RAxML, MARER: RAxML, IQ-TREE, coalescent), to Platyhelminthes (COMP: IQ-TREE), or in the polytomy with both (COMP: coalescent; figure 2 summarizing the most probable position of Mesozoa in Lophotrochozoa). The same applied also to RAxML analysis of the C50LBR dataset: Mesozoa were found monophyletic and sister to Platyhelminthes. Both gene subsampling strategies (branch length homogeneity and compositional homogeneity), used in an attempt to alleviate LBA, were not able to disintegrate monophyletic Mesozoa or to reduce its strong support.

Contrarily, the IQ-TREE and BI analyses (with site-specific models) of the C50LBR data matrix recovered Mesozoa split into two separate clades. Dicyemida was found either within Gnathifera in a poorly supported polytomy with Gnathostomulida and Micrognathozoa–Rotifera (BI, with only very low support $pp=0.62$), or as a sister group of Gnathifera (IQ-TREE, again with very low support $BS=54$). The position of Orthonectida as a sister group of Platyhelminthes was relatively well supported in the BI tree ($PP=0.98$) but not in the IQ-TREE one ($BS=49$).

The possibility that Mesozoa are two separate groups attracted to each other owing to long-branch artefact was tested by including just one mesozoan group in absence of the other. Both Dicyemida and Orthonectida consistently grouped within the Gnathifera–Platyhelminthes clade. The taxon-removal analyses suggest that Gnathostomulida (but not Micrognathozoa and Rotifera) is the only attractor strong

enough to change position of mesozoans from Platyhelminthes to Gnathifera. Thus, we can assume that gnathiferan or even within-gnathiferan affinities of Mesozoa are probably artificial, and the alternative, i.e. close relationship between Mesozoa (or at least Dicyemida; see below) and Platyhelminthes, is more probable. The hypothesis considering Mesozoa as extremely simplified parasitic flatworms (namely, Neodermata, the group including flukes and tapeworms) has actually been considered for a long time (see [63]), based on putative similarities between trematode miracidia and mesozoan infective larvae and more recently on the ciliary rootlet system putatively shared by Neodermata and Mesozoa (but see above). On the contrary, Mesozoa has the orthodox mitochondrial genetic code rather than the apomorphic unorthodox code characteristic for Rhabditophora (that includes Neodermata), which suggests the idea that they might be derived flatworms is unlikely [64]. Even if Dicyemida were occasionally recovered as platyhelminths here (T2 tree in Δ GLS analysis or IQ-TREE analysis of the MARER dataset with orthonectids excluded), they were a sister group of Catenulida, not an in-group of Rhabditophora. Hence, parasitism of the two mesozoan groups has certainly evolved independently from the parasitism of neodermatan platyhelminths, even if they share evolutionary origins with flatworms.

Whereas phylogenetic position of dicyemids alone was usually compatible with the position of the whole Mesozoa, higher instability of position of Orthonectida was evident in published studies. Mikhailov *et al.* [34] sequenced the genome of *Intoshia linei* and recovered Orthonectida as a sister group of Platyhelminthes in Bayesian analysis with the CAT + GTR model as well as in ML analysis, yet the use of CAT model with a flat rate in the Bayesian analysis resulted in Orthonectida sister to Annelida. With the addition of a dicyemid genome, Lu *et al.* [35] recovered Orthonectida (together with Dicyemida in monophyletic Mesozoa) as close relatives of Gastrotricha and/or Platyhelminthes, while Schiffer *et al.* [36] and Zverkov *et al.* [15] found them as early branching annelids, unrelated to Dicyemida, and Bondarenko *et al.* [65] even as a sister group of leeches (Annelida: Pleistoannelida: Clitellata).

The annelid affinities of the Orthonectida, based on morphology, have been proposed previously, starting with Metschnikoff [25]. In more recent studies, it has been hypothesized on the basis of annelid-like microvilli-formed cuticle in the free-living stages of Orthonectida [27]. Orthonectids (contra dicyemids) also have a true epidermis (but their epidermal basal lamina is reduced or absent). A possible trace of an ancestral segmented body plan is seen as the series of regularly spaced circular muscles along the anterior-posterior axis of Orthonectida (along with repeated bands of cilia and paired serotonin-like immunoreactive nerve cells [66]). However, orthonectids use cilia for locomotion, and their muscle system is reduced to serve almost exclusively for copulation and hatching of larvae. The annelid muscular pattern consists of two muscle layers, external circular/transverse and internal longitudinal. In Orthonectida, the circular muscles are situated inversely, inside the longitudinal muscles. The only annelid with external longitudinal muscles sunken into the epidermis is non-segmented meiofaunal *Lobatocerebrum* [67], a member of Dinophiliformia, one of the basal annelid clades [26]. The possible relationships of Orthonectida and basal annelids are compatible with the results published by Schiffer *et al.* [36] who found in the Orthonectida a short

stretch of mitochondrial genes (*nad1*, *nad6*, *cob*) in the same order as in *Owenia* but not as in the pleistoannelids.

In our study, the scenario of Orthonectida related or belonging to Annelida was recovered only in one unconverged BI analysis (C100HR matrix) and in a taxon removal experiment where no dicyemid and platyzoan taxa were present, both results being extremely problematic to interpret. The AU test showed that topology with orthonectids related to annelids is significantly worse based on our data than topology with monophyletic Mesozoa within Platyzoa. Furthermore, ΔGLS analysis revealed that monophyly of Mesozoa is supported by more genes than its polyphyly with annelid affinities of Orthonectida (148 : 54). Moreover, even in the genes compatible with the annelid hypothesis, the 'annelid signal' is minor, and the resulting tree included within-platyzoan Mesozoa, albeit diphylectic.

(b) Different methods, different views

Two main approaches are typically used in analysing phylogenomic matrices, BI and ML, which have a wide, but not fully overlapping range of available models of sequence evolution. Site-specific frequency models (or empirical profile mixture models) based on modelling of each alignment site individually are considered to better model sequence evolution and therefore be more reliable in obtaining correct phylogeny in difficult cases [46,68]. Until recently, the site-specific frequency model (CAT [53]) was available only in the Bayesian framework, but an approximation of such a model (posterior mean site frequency model [39]) has recently been implemented in IQ-TREE in the ML framework.

In the case of Mesozoa, almost all maximum-likelihood-based methods, regardless of the number of genes used, recovered a similar topology; contrarily, Bayesian analyses tended to recover a different tree. Based on the 50-gene matrix (C50LBR, the only matrix that converged in BI), both BI and IQ-TREE analyses recovered diphylectic Mesozoa (Dicyemida close to Gnathifera, Orthonectida to Platyhelminthes), while in the RAxML tree Mesozoa is monophyletic and sister to Platyhelminthes. It could suggest that the site-specific model is a key for recovery of this specific topology in the 50 gene matrix. However, when the same model was used in ML on bigger matrices, it again consistently showed Mesozoa as monophyletic.

Similar inconsistency was recovered by Zverkov *et al.* [15]. They presented BI phylogeny with polyphyletic Mesozoa and paraphyletic Platyzoa (basal Gnathifera, followed by the Dicyemida–Entoprocta–Cycliophora–Gastrotricha–Platyhelminthes clade, the latter sister to Trochozoa incl. Bryozoa and Orthonectida) but they recovered monophyletic Mesozoa within monophyletic Platyzoa in RAxML tree (their electronic supplementary material, figure S4) and monophyletic Mesozoa as a sister group of Rousphozoa (Gastrotricha + Platyhelminthes) in IQ-TREE topology (electronic supplementary material, figure S5). Both Zverkov *et al.*'s and the present analyses suggest a consistent conflict between ML and BI topologies. Surprisingly, ML methods with both classic and site-specific PMSF models showed similar scenarios.

The inference of relationships among lophotrochozoan clades is thought to be hindered by the effect of LBA owing to (sometimes dramatic) differences in molecular evolutionary rate, the application of inaccurate models of sequence evolution, and the inherent difficulty of resolving ancient radiations because of short stem branches in deep nodes [19]. Using genes with specific characteristics (branch length homogeneity, compositional homogeneity) has been suggested to be more suited for phylogenetic inference of taxa with uneven branch lengths [41,69]. However, targeting genes with these specific qualities did not have a clear effect in the case of mesozoan phylogeny, similarly as in an earlier phylogenomic analysis of Lophotrochozoa [48]. Also, in our case, the phylogenetic position of the long-branch mesozoan groups was not affected by taxon-removal experiments. This suggests that either LBA is not the main player causing significant errors in the case of lophotrochozoan and mesozoan phylogeny, or that LBA in the case of Mesozoa is too strong to be overcome by the currently used methods.

Data accessibility. Electronic supplementary material can be found online [70].

Authors' contributions. M.D.: data curation, investigation, methodology, project administration, resources, software, visualization, writing—original draft, writing—review and editing; K.M.K.: conceptualization, data curation, funding acquisition, methodology, project administration, resources, supervision, writing—review and editing; K.M.H.: conceptualization, writing—review and editing; T.H.O.: conceptualization, funding acquisition, writing—review and editing; A.E.G.-V.: data curation, writing—review and editing; M.S.P.: resources; R.V.: resources; L.L.M.: conceptualization, funding acquisition, resources, writing—review and editing; J.T.C.: resources, writing—review and editing; A.K.: conceptualization; E.A.E.: resources; J.S.: conceptualization, funding acquisition, project administration, resources, supervision, writing—review and editing; J.Z.: conceptualization, funding acquisition, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was supported by the Czech Science Foundation (GACR grant nos. 15-08717S to J.Z. and M.D., and 19-28399X to J.Š.) and, in part, by the United States National Science Foundation (grant nos. 1146575, 1557923, 1548121 and 1645219 to L.L.M. and 1846174 to K.M.K.).

Acknowledgements. We would like to thank colleagues and institutions who let us use their facilities for sample processing, namely Graziano Fiorito and Giovanna Ponte (Stazione Zoologica Anton Dohrn, Naples, Italy); Andrea De Lucia, Andrea Camedda and Stefano Marra (Institute for Coastal Marine Environment, CNR, Oristano, Sardinia, Italy) and we acknowledge the Laboratory of Microscopy and Histology of the Biology Centre CAS for technical help with the guiding through immunohistochemical visualization and for allowing the use of their facilities. We thank Deb Crocker and Robert Griffin for support of the University of Alabama High-Performance Computing cluster (UAHPC). Computational resources were also supplied by the project 'e-Infrastruktura CZ' (e-INFRA LM2018140) provided within the programme Projects of Large Research, Development and Innovations Infrastructures, as well as the National Science Foundation (CNS-1725797) administered by UCSB's Center for Scientific Computing (CSC).

References

1. Dunn CW, Giribet G, Edgecombe GD, Hejnol A. 2014 Animal phylogeny and its evolutionary implications. *Annu. Rev. Ecol. Evol. Syst.* **45**, 371–395. (doi:10.1146/annurev-ecolsys-120213-091627)
2. Wägele WBT. 2014 *Deep metazoan phylogeny: the backbone of the tree of life*. Berlin–Boston: de Gruyter.

3. Telford MJ, Budd GE, Philippe H. 2015 Phylogenomic insights into animal evolution. *Curr. Biol.* **25**, R876–R887. (doi:10.1016/j.cub.2015.07.060)
4. Halanich KM. 2016 How our view of animal phylogeny was reshaped by molecular approaches: lessons learned. *Org. Divers. Evol.* **16**, 319–328. (doi:10.1007/s13127-016-0264-8)
5. Kocot KM. 2016 On 20 years of Lophotrochozoa. *Org. Divers. Evol.* **16**, 329–343. (doi:10.1007/s13127-015-0261-3)
6. Wanninger A. 2016 Twenty years into the ‘new animal phylogeny’: changes and challenges. *Org. Divers. Evol.* **16**, 315–318. (doi:10.1007/s13127-016-0277-3)
7. Giribet G, Edgecombe GD. 2017 Current understanding of ecdysozoa and its internal phylogenetic relationships. *Integr. Comp. Biol.* **57**, 455–466. (doi:10.1093/icb/ix072)
8. Bleidorn C. 2019 Recent progress in reconstructing lophotrochozoan (spiralian) phylogeny. *Org. Divers. Evol.* **19**, 557–566. (doi:10.1007/s13127-019-00412-4)
9. Marlétaz F, Peijnenburg KTCA, Goto T, Satoh N, Rokhsar DS. 2019 A new spiralian phylogeny places the enigmatic arrow worms among gnathiferans. *Curr. Biol.* **29**, 312–318.e3. (doi:10.1016/j.cub.2018.11.042)
10. Felsenstein J. 1978 Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* **27**, 401–410. (doi:10.2307/2412923)
11. Bergsten J. 2005 A review of long-branch attraction. *Cladistics* **21**, 163–193. (doi:10.1111/j.1096-0031.2005.00059.x)
12. Philippe H, Vienne Dd, Ranwez V, Roure B, Baurain D, Delsuc F. 2017 Pitfalls in supermatrix phylogenomics. *Eur. J. Taxon.* **283**, 1–25. (doi:10.5852/ejt.2017.283)
13. Qun Y, Junye M, Xiaoyan S, Peiyun C. 2007 Phylogenetic of early metazoans: combined evidence from molecular and fossil data. *Geol. J.* **42**, 281–295. (doi:10.1002/gj.1074)
14. Laumer CE *et al.* 2019 Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proc. R. Soc. B* **286**, 20190831. (doi:10.1098/rspb.2019.0831)
15. Zverkov OA *et al.* 2019 Dicyemida and Orthonectida: two stories of body plan simplification. *Front. Genet.* **10**, 443. (doi:10.3389/fgene.2019.00443)
16. Rokas A, Kruger D, Carroll SB. 2005 Animal evolution and the molecular signature of radiations compressed in time. *Science* **310**, 1933–1938. (doi:10.1126/science.1116759)
17. Thomas JA, Welch JJ, Lanfear R, Bromham L. 2010 A generation time effect on the rate of molecular evolution in invertebrates. *Mol. Biol. Evol.* **27**, 1173–1180. (doi:10.1093/molbev/msq009)
18. Haraguchi Y, Sasaki A. 1996 Host–parasite arms race in mutation modifications: indefinite escalation despite a heavy load? *J. Theor. Biol.* **183**, 121–137. (doi:10.1006/jtbi.1996.9999)
19. Budd GE, Jackson ISC. 2016 Ecological innovations in the Cambrian and the origins of the crown group phyla. *Phil. Trans. R. Soc. B* **371**, 20150287. (doi:10.1098/rstb.2015.0287)
20. WoRMS Editorial Board 2018: van der Land, Furuya and Decock. 2018 World Register of Marine Species. See <http://www.marinespecies.org> at VLIZ (accessed 18 May 2018).
21. Deheyn D, Watson NA, Jangoux M. 1998 Symbioses in *Amphipholis squamata* (Echinodermata, Ophiuroidea, Amphiuridae): geographical variation of infestation and effect of symbionts on the host’s light production. *Int. J. Parasitol.* **28**, 1413–1424. (doi:10.1016/S0020-7519(98)00119-2)
22. Catalano SR. 2012 A review of the families, genera and species of Dicyemida Van Beneden, 1876. *Zootaxa* **32**, 1–32. (doi:10.11646/zootaxa.3479.1.1)
23. Van Beneden ME. 1876 Recherches sur le Dicyemides, survivants actuels d’un embranchement des Mésozoaires. *Bull. Académie R. Sci. Lett. B-arts Belg.* 1876; Bruxelles: 1160–250.
24. Nouvel H. 1948 Les Dicyemides 2em partie: Infusoriforme, Teratologie, Specificité du parasitisme, Affinités. *Arch. Biol. Liege* **59**, 147–223.
25. Metschnikoff E. 1881 Untersuchungen über Orthonectiden. *Z. Für. Wiss. Zool.* **32**, 349–387.
26. Martín-Durán JM *et al.* 2021 Conservative route to genome compaction in a miniature annelid. *Nat. Ecol. Evol.* **5**, 231–242. (doi:10.1038/s41559-020-01327-6)
27. Slyusarev GS, Kristensen R. 2003 Fine structure of the ciliated cells and ciliary rootlets of *Intoshia variabilis* (Orthonectida). *Zoomorphology* **122**, 33–39. (doi:10.1007/s00435-002-0065-9)
28. Kozloff EN. 1969 Morphology of the Orthonectid *Rhopalura ophiocoma*. *J. Parasitol.* **55**, 171–195. (doi:10.2307/3277368)
29. Pawlowski J, Montoya-Burgos JL, Fahrni JF, Wüest J, Zaninetti L. 1996 Origin of the Mesozoa inferred from 18S rRNA gene sequences. *Mol. Biol. Evol.* **13**, 1128–1132. (doi:10.1093/oxfordjournals.molbev.a025675)
30. Zrzavy J, Mihulka S, Kepka P, Bezdek A, Tietz D. 1998 Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. *Cladistics* **14**, 249–285. (doi:10.1111/j.1096-0031.1998.tb00338.x)
31. Kobayashi M. 1999 Dicyemids are higher animals. *Nat. Macmillan Mag. Ltd.* **40**, 762–763.
32. Aruga J, Odaka YS, Kamiya A, Furuya H. 2007 *Dicyema* Pax6 and Zic: tool-kit genes in a highly simplified bilaterian. *BMC Evol. Biol.* **7**, 201. (doi:10.1186/1471-2148-7-201)
33. Suzuki TG, Ogino K, Tsuneki K, Furuya H. 2010 Phylogenetic analysis of dicyemid mesozoans (phylum Dicyemida) from innexin amino acid sequences: dicyemids are not related to Platyhelminthes. *J. Parasitol.* **96**, 614–625. (doi:10.1645/GE-2305.1)
34. Mikhailov KV, Slyusarev GS, Nikitin MA, Logacheva MD, Penin AA, Aleoshin VV, Panchin YV. 2016 The genome of *Intoshia linei* affirms orthonectids as highly simplified spiralian. *Curr. Biol.* **26**, 1768–1774. (doi:10.1016/j.cub.2016.05.007)
35. Lu TM, Kanda M, Satoh N, Furuya H. 2017 The phylogenetic position of dicyemid mesozoans offers insights into spiralian evolution. *Zool. Lett.* **3**, 1–9. (doi:10.1186/s40851-017-0062-y)
36. Schiffer PH, Robertson HE, Telford MJ. 2018 Orthonectids are highly degenerate annelid worms. *Curr. Biol.* **28**, 1970–1974. e3. (doi:10.1016/j.cub.2018.04.088)
37. Lartillot N, Philippe H. 2004 A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21**, 1095–1109. (doi:10.1093/molbev/msh112)
38. Lartillot N, Rodrigue N, Stubbs D, Richer J. 2013 PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* **62**, 611–615. (doi:10.1093/sysbio/syt022)
39. Wang HC, Minh BQ, Susko E, Roger AJ. 2018 Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Syst. Biol.* **67**, 216–235. (doi:10.1093/sysbio/syx068)
40. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015 IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274. (doi:10.1093/molbev/msu300)
41. Shen XX, Salichos L, Rokas A. 2016 A genome-scale investigation of how sequence, function, and tree-based gene properties influence phylogenetic inference. *Genome Biol. Evol.* **8**, 2565–2580. (doi:10.1093/gbe/evw179)
42. Nesnidal MP, Helmkamp M, Bruchhaus I, Hausdorf B. 2010 Compositional heterogeneity and phylogenomic inference of metazoan relationships. *Mol. Biol. Evol.* **27**, 2095–2104. (doi:10.1093/molbev/msq097)
43. Liu L, Yu L, Kubatko L, Pearl DK, Edwards SV. 2009 Coalescent methods for estimating phylogenetic trees. *Mol. Phylogenet. Evol.* **53**, 320–328. (doi:10.1016/j.ympev.2009.05.033)
44. Kapli P, Yang Z, Telford MJ. 2020 Phylogenetic tree building in the genomic age. *Nat. Rev. Genet.* **21**, 428–444. (doi:10.1038/s41576-020-0233-0)
45. Stamatakis A. 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. (doi:10.1093/bioinformatics/btu033)
46. Lartillot N, Brinkmann H, Philippe H. 2007 Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. *BMC Evol. Biol.* **7**(Suppl. 1), S4. (doi:10.1186/1471-2148-7-S1-S4)
47. Mirarab S, Warnow T. 2015 ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* **31**, i44–i52. (doi:10.1093/bioinformatics/btv234)
48. Kocot KM *et al.* 2017 Phylogenomics of lophotrochozoa with consideration of systematic error. *Syst. Biol.* **66**, 256–282.
49. Ebersberger I, Strauss S, von Haeseler A. 2009 HaMStR: profile hidden Markov model based search for orthologs in ESTs. *BMC Evol. Biol.* **9**, 157. (doi:10.1186/1471-2148-9-157)

50. Meyer B, Meusemann K, Misof B. 2011 MARE v0.1.2-rc.xxx. See <https://bonn.leibniz-lib.de/en/research/research-centres-and-groups/mare>.
51. Struck TH. 2014 Trespex-detection of misleading signal in phylogenetic reconstructions based on tree information. *Evol Bioinforma.* **10**, 51–67. (doi:10.4137/EBO.S14239)
52. Kück P, Struck TH. 2014 BaCoCa – a heuristic software tool for the parallel assessment of sequence biases in hundreds of gene and taxon partitions. *Mol. Phylogenet. Evol.* **70**, 94–98. (doi:10.1016/j.ympev.2013.09.011)
53. Si Quang L, Gascuel O, Lartillot N. 2008 Empirical profile mixture models for phylogenetic reconstruction. *Bioinformatics* **24**, 2317–2323. (doi:10.1093/bioinformatics/btn445)
54. Shen XX, Hittinger CT, Rokas A. 2017 Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nat. Ecol. Evol.* **1**, 0126. (doi:10.1038/s41559-017-0126)
55. Ballesteros JA, Sharma PP. 2019 A critical appraisal of the placement of xiphosura (chelicerata) with account of known sources of phylogenetic error. *Syst. Biol.* **68**, 896–917. (doi:10.1093/sysbio/syz011)
56. Simmons MP, Sloan DB, Springer MS, Gatesy J. 2019 Gene-wise resampling outperforms site-wise resampling in phylogenetic coalescence analyses. *Mol. Phylogenet. Evol.* **131**, 80–92. (doi:10.1016/j.ympev.2018.10.001)
57. Smith SA, Dunn CW. 2008 Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics* **24**, 715–716. (doi:10.1093/bioinformatics/btm619)
58. Huerta-Cepas J, Serra F, Bork P. 2016 ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* **33**, 1635–1638. (doi:10.1093/molbev/msw046)
59. Whelan NV, Halanaych KM. 2017 Who Let the CAT Out of the Bag? Accurately dealing with substitutional heterogeneity in phylogenomic analyses. *Syst. Biol.* **66**, 232–255. (doi:10.1093/sysbio/syw084)
60. Czaker R. 2000 Extracellular matrix (ECM) components in a very primitive multicellular animal, the dicyemid mesozoan *Kantharella antarctica*. *Anat. Rec.* **259**, 52–59. (doi:10.1002/(SICI)1097-0185(20000501)259:1<52::AID-AR6>3.0.CO;2-J)
61. Nesnidal MP *et al.* 2013 New phylogenomic data support the monophyly of Lophophorata and an Ectoproct-Phoronid clade and indicate that Polyzoa and Kryptrochozoa are caused by systematic bias. *BMC Evol. Biol.* **13**, 253. (doi:10.1186/1471-2148-13-253)
62. Ax P. 1996 *Multicellular animals: a new approach to the phylogenetic order in nature*, vol. 1. Berlin, Germany: Springer.
63. Whitman CO. 1883 *A contribution to the embryology, life-history, and classification of the Dicyemids*. Leipzig, Germany: Wilhelm Engelmann.
64. Telford MJ, Herniou EA, Russell RB, Littlewood DT. 2000 Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms. *Proc. Natl Acad. Sci. USA* **97**, 11 359–11 364. (doi:10.1073/pnas.97.21.11359)
65. Bondarenko N, Bondarenko A, Starunov V, Slyusarev G. 2019 Comparative analysis of the mitochondrial genomes of Orthonectida: insights into the evolution of an invertebrate parasite species. *Mol. Genet. Genomics.* **294**, 715–727. (doi:10.1007/s00438-019-01543-1)
66. Slyusarev GS, Bondarenko NI, Skalon EK, Rappoport AK, Radchenko D, Starunov VV. 2022 The structure of the muscular and nervous systems of the orthonectid *Rhopalura litoralis* (Orthonectida) or what parasitism can do to an annelid. *Org. Divers. Evol.* **22**, 35–45. (doi:10.1007/s13127-021-00519-7)
67. Kerbl A, Bekkouche N, Sterrer W, Worsaae K. 2015 Detailed reconstruction of the nervous and muscular system of Lobatocerebridae with an evaluation of its annelid affinity. *BMC Evol. Biol.* **15**, 277. (doi:10.1186/s12862-015-0531-x)
68. Roure B, Baurain D, Philippe H. 2013 Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. *Mol. Biol. Evol.* **30**, 197–214. (doi:10.1093/molbev/mss208)
69. Struck TH. 2013 The impact of paralogy on phylogenomic studies – a case study on annelid relationships. *PLoS ONE* **8**, e62892. (doi:10.1371/journal.pone.0062892)
70. Drábková M *et al.* 2022 Different phylogenomic methods support monophyly of enigmatic ‘Mesozoa’ (Dicyemida + Orthonectida, Lophotrochozoa). Figshare. (doi:10.6084/m9.figshare.c.6070258)