

Research



Cite this article: Laumer CE *et al.* 2019 Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proc. R. Soc. B* **286**: 20190831.
<http://dx.doi.org/10.1098/rspb.2019.0831>

Received: 9 April 2019

Accepted: 17 June 2019

Subject Category:

Evolution

Subject Areas:

taxonomy and systematics, evolution, genomics

Keywords:

phylogenomics, animal phylogeny, compositional bias, taxon sampling, matrix recoding

Author for correspondence:

Christopher E. Laumer
 e-mail: daumer@ebi.ac.uk

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.4552313>.

Revisiting metazoan phylogeny with genomic sampling of all phyla

Christopher E. Laumer^{1,2}, Rosa Fernández^{1,3}, Sarah Lemer^{1,4}, David Combosch^{1,4}, Kevin M. Kocot⁵, Ana Riesgo⁶, Sónia C. S. Andrade⁷, Wolfgang Sterrer⁸, Martin V. Sørensen⁹ and Gonzalo Giribet¹

¹Museum of Comparative Zoology (MCZ) and Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA

²EMBL-European Bioinformatics Institute (EBI), Wellcome Genome Campus, Hinxton CB10 1SD, UK

³Bioinformatics & Genomics Unit, Center for Genomic Regulation, Carrer del Dr. Aiguader 88, 08003 Barcelona (Spain)

⁴Marine Laboratory, University of Guam, UOG Station, Mangilao, Guam 96923, USA

⁵Department of Biological Sciences and Alabama Museum of Natural History, The University of Alabama, Campus Box 870344, Tuscaosa, AL 35487, USA

⁶Department of Life Sciences, Natural History Museum of London, Cromwell Road, London SW7 5BD, UK

⁷Departamento de Genética e Biologia Evolutiva, IB, Universidade de São Paulo, 05508090 São Paulo, SP, Brazil

⁸Bermuda Natural History Museum, PO Box FL 145, Flatts, FLBX, Bermuda

⁹Natural History Museum of Denmark, Universitetsparken 15, 2100 Copenhagen, Denmark

CEL, 0000-0001-8097-8516; RF, 0000-0002-4719-6640; SL, 0000-0003-0048-7296; DC, 0000-0001-7004-7435; KMK, 0000-0002-8673-2688; AR, 0000-0002-7993-1523; SCSA, 0000-0002-1302-5261; MVS, 0000-0002-0377-0276; GG, 0000-0002-5467-8429

Proper biological interpretation of a phylogeny can sometimes hinge on the placement of key taxa—or fail when such key taxa are not sampled. In this light, we here present the first attempt to investigate (though not conclusively resolve) animal relationships using genome-scale data from all phyla. Results from the site-heterogeneous CAT + GTR model recapitulate many established major clades, and strongly confirm some recent discoveries, such as a monophyletic Lophophorata, and a sister group relationship between Gnathifera and Chaetognatha, raising continued questions on the nature of the spiralian ancestor. We also explore matrix construction with an eye towards testing specific relationships; this approach uniquely recovers support for Panarthropoda, and shows that Lophotrochozoa (a subclade of Spiralia) can be constructed in strongly conflicting ways using different taxon- and/or orthologue sets. Dayhoff-6 recoding sacrifices information, but can also reveal surprising outcomes, e.g. full support for a clade of Lophophorata and Entoprocta + Cycliophora, a clade of Placozoa + Cnidaria, and raising support for Ctenophora as sister group to the remaining Metazoa, in a manner dependent on the gene and/or taxon sampling of the matrix in question. Future work should test the hypothesis that the few remaining uncertainties in animal phylogeny might reflect violations of the various stationarity assumptions used in contemporary inference methods.

1. Background

For over a decade, molecular phylogeneticists have enjoyed the use of automated methods to use shotgun DNA sequencing data to decipher the deepest relationships in the animal tree of life [1,2]. This paradigm has continued the disruptive tradition of molecular phylogenetics, allowing the placement of taxa whose morphology and embryology have proven uninformative or misleading in this regard, and demonstrating that early animal evolution resulted in considerably more flexibility in phenotypic evolution than initially expected [3–5]. Unfortunately, however, even with the availability of genome-scale data, the shift away from morphologically defined trees has not proceeded towards one consistent molecular tree. Controversies have abounded, including some ongoing

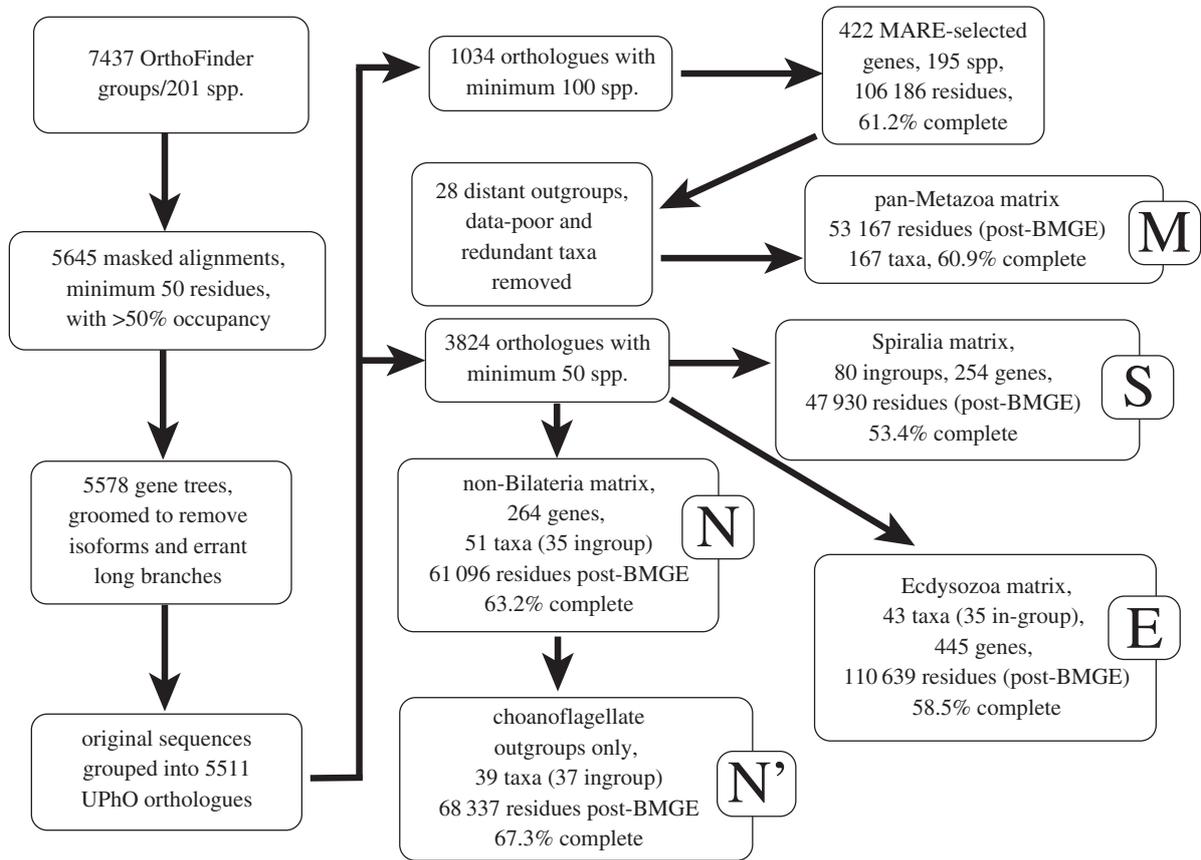


Figure 1. Schematic description of gene tree construction, orthologue assignment and matrix construction. Gene selection criteria for clade-specific matrix construction and other methodological details are discussed in the text.

ones, such as the contention that Xenacoelomorpha represent deuterostomes [6–8], or the debate over the earliest split in the animal tree [9–16]. Furthermore, although highly parallel short-read sequencing has essentially overtaken Sanger and competing second-generation sequencing technologies, there is still a dearth of genomic data from several phyla (e.g. Bryozoa, Loricifera, Kinorhyncha and Nematomorpha), and it has been a decade since the last major synoptic attempt to infer the relationships among all animal groups [2], despite exemplary recent analyses focused on specific clades [17,18]. Indeed, it seems there has not yet been an attempt to investigate the animal tree of life using genome-scale data from representatives of all metazoan phyla. Herein, collating a mixture of published (prior to 2018) and new transcriptome and genome data sequenced largely with Illumina technology, and employing numerous strategies to control the influence of systematic error [19] and to build both general and taxon-specific matrices from a single orthology assignment, we empirically review the signals for and robustness of most animal clades recognized in the recent era.

2. Methods

Detailed description of molecular methods for RNA and DNA sequencing and run parameters for all bioinformatic analyses are provided as electronic supplementary material.

(a) Orthologue assignment and matrix construction

Predicted proteomes derived from annotated genome and transcriptome assemblies were clustered into 7437 OrthoFinder groups (figure 1) comprising 201 spp. sampling all metazoan

phyla (except for Orthonectida, from which no genomic resources were available at the time of this work's inception [20,21], and which were recently shown to represent modified annelids [22]), plus a variety of opisthokont outgroups (electronic supplementary material, table S1). From these, we constructed 5578 maximum-likelihood (ML) gene family trees from a subset of well-aligned groups, and processed these to mask candidate redundant isoforms, remove isolated divergent sequences and further split trees into subfamilies subtended by long internal branches [23]. These groomed gene trees were parsed to extract 5511 orthologues by the criterion of unrooted phylogenetic orthology (UPhO) [24].

To construct a single supermatrix representing all Metazoa, we considered the set of 1034 orthologues with 100 or more representatives. This was done in part to ensure a matrix with high taxon occupancy, but also to limit the effects of cryptic horizontal gene-transfer, biological cross-contamination or index misassignment (of which libraries produced in this laboratory have, however, previously shown little evidence [25]): when selecting orthologues by parsing gene family trees with the species overlap algorithm, as done here, such processes should tend to split large orthologues into smaller groupings. From this set of 1034, we further reduced to 422 information-rich genes present in 195 taxa (figure 1). Our initial ML tree showed evidence of redundant and poorly placed individual taxa, as well as the presence of some clades previously shown to be driven by compositional bias (e.g. Polyzoa [26]; electronic supplementary material, figure S1). We therefore selected a set of 28 taxa to remove (see electronic supplementary material, table S1), including all non-choanoflagellate outgroups, following contemporaneous suggestions of compositionally driven effects from the inclusion of these clades [11,12], and trimmed the matrix of putatively saturated and compositionally biased sites with the BMGE tool [27]. Interestingly, this procedure

reduced the matrix from an initial 106 186 sites to 53 167 sites when these taxa were removed prior to BMGE, but to only 43 011 sites when all taxa are included prior to BMGE-trimming, indicating that more sites are detected as compositionally heterogeneous when distant outgroups are included in the test; we focus results from the 53 167-site matrix (M), but refer to those from the shorter matrix for comparison (see the electronic supplementary material figures).

We also constructed subclade-specific matrices with more limited taxon sampling (figure 1), meant to test specific relationships within these major subclades (e.g. Ecdysozoa, Spiralia), and with orthologues selected from within the set of 3824 with greater than 50 sequences each, to optimize representation of the clades in question. A 43-taxon (34 ingroups) Ecdysozoa matrix (E) was prepared with a MARE-reduced subset of 445 genes, from the set of those which had at least two representatives each of Kinorhyncha, Loricifera or Nematomorpha. An 80-taxon matrix (S) was constructed to test the position of Cycliophora and Entoprocta within Spiralia, composed of the 254 genes with representation of at least three each of Entoprocta + Cycliophora (no ecdysozoan outgroups were included); although relationships within Spiralia have been controversial in many respects, the position of Entoprocta and Cycliophora within this clade has been among the most difficult to assess, potentially due to compositional bias [17,26,28]. Indeed, an extensive recent analysis of Spiralian relationships [18] chose to avoid solving the position of Cycliophora entirely by trimming this group (and many other compositionally biased species) away entirely prior to phylogenetic inference. Here, we have favoured including all taxa possible, given the recognized importance of taxon sampling in accurate phylogenetic inference, choosing to mitigate compositional heterogeneity by trimming problematic alignment sites rather than entire taxa. A 51-taxon matrix (N) was finally constructed to evaluate relationships between Bilateria and the remaining four animal phyla and outgroups, including 264 genes selected by MARE from within the set of alignments that included the placozoan and at least one choanoflagellate, six sponges, three ctenophores and four cnidarians. We also constructed a version of this matrix (N') with all non-choanoflagellate outgroups removed, leaving 39 taxa; BMGE-trimming was performed after taxon deletion, yielding 68 337 and 61 096 residues for the 51 and 39 taxon matrices, respectively. These matrices allowed us to test whether a metazoan-wide matrix and those that are clade-specific and therefore more informative to the specific question, produce comparable results.

(b) Phylogenetic inference

We present principally results from Bayesian inference under the CAT + GTR + *I*4 model, which has been shown both theoretically and empirically to suppress long-branch attraction artefacts in heterogeneous matrices such as the ones presented here [29–31]. A minimum of four chains per matrix were run in PHYLOBAYES-MPI v. 1.6j for up to 1.5 years. Straightforward posterior consensus formation with conventional burn-ins indicated difficulties achieving convergence in many matrices, even following such long computation times, but we preferred, in contrast to some recent work [8,13], to exhaustively investigate the signal within one informative matrix (rather than jackknifing within a larger matrix [19]), and to employ the more general CAT + GTR + *I*4 model over the CAT + *I*4, which has been shown to be more susceptible to systematic error [32]. However, we determined that the apparent difficulties in achieving acceptable metrics of convergence were principally the result of isolated rogue taxa in 3 poorly taxonomically sampled ecdysozoan groups present in each matrix (annotated in electronic supplementary material, table S1). Sometimes, such taxa were represented sparsely in each matrix (e.g. only 1009 and 1554 occupied sites for the two kinorhynch species represented in the pan-Metazoa matrix); however, other rogue taxa

(particularly our single representative each of Nematomorpha and Loricifera) were represented in thousands of positions and nonetheless showed poor stability throughout CAT + GTR + *I*4 chains. Therefore, such rogue taxa, defined anew for each individual analysis, were masked prior to posterior consensus formation, resulting in acceptable (maxdiff < 0.2) maximum bipartition differences across chains and generally higher support values throughout, as has been seen previously [33]. Unpruned consensus summaries, fully labelled trees and ML analyses of each matrix are also presented in the electronic supplementary material which also includes a number of early analyses not further discussed (but described in caption, electronic supplementary material, figures S1–S24).

3. Results and discussion

(a) Pan-metazoan matrix (M)

In broad structure, the trees from both amino acid and Dayhoff-6 group CAT + GTR + *I*4 analyses of our most heavily analysed pan-metazoan matrix (figure 2), the BMGE-trimmed 53 167-site matrix M, recapitulate many deep relationships seen in molecular studies to date: Parahoxozoa, Planulozoa (= Cnidaria + Bilateria), Bilateria, Nephrozoa, Deuterostomia, Protostomia, Ecdysozoa and Spiralia all receive strong support. Ctenophora is recovered as the sister group of the remaining Metazoa, as is seen in many analyses [9]; however, unlike most other analyses to date, support for this split is poor in the amino acid analysis (figure 2*a,b*). Furthermore, contradicting previous observations that Dayhoff-6 group recoding tends to favour Porifera as the sister group to the remaining Metazoa, at least with this specific matrix, recoding instead maximizes support for Ctenophora in this position, even with only choanoflagellates as outgroups [11] (figure 2*c,d*). By contrast, an even more heavily reduced (43 011-site) version of matrix M, where taxon (particularly, outgroup) deletion was applied only after trimming putatively saturated and/or compositionally biased sites, still shows poor resampling support for the earliest bipartition in Metazoa in CAT + GTR analysis at the amino acid level (electronic supplementary material, figure S3). Here, however, Placozoa are recovered with full support as the sister group of Cnidaria (a result also recently reported elsewhere [34]).

The more common effect of Dayhoff-6 recoding in matrix M is to reduce support for many clades, exemplified within Ecdysozoa. In the amino acid analysis (figure 2*a*), strong support is seen for an arthropod–onychophoran clade, as well as for a tardigrade–nematode clade, both of which are common outcomes of ecdysozoan molecular phylogenies [35–37], although the latter contradicts the clade Panarthropoda, also recovered in some genome-scale phylogenies [35,38]. Support for both of these clades is eroded under Dayhoff-6 recoding (figure 2*c*); indeed, the only ecdysozoan split for which strong support remains robust to this reduced alphabet is the division between Priapulida and the remaining Ecdysozoa. Unfortunately, within this analysis, our only representatives of the phyla Kinorhyncha, Nematomorpha and Loricifera behaved as rogue taxa (electronic supplementary material, figure S15–S17), and were therefore masked prior to posterior consensus summary. Ecdysozoa, one of the best-supported metazoan clades, therefore continues to be poorly resolved and understood [37].

Within Spiralia, our thorough taxon sampling permits interrogation of relationships within this challenging clade.

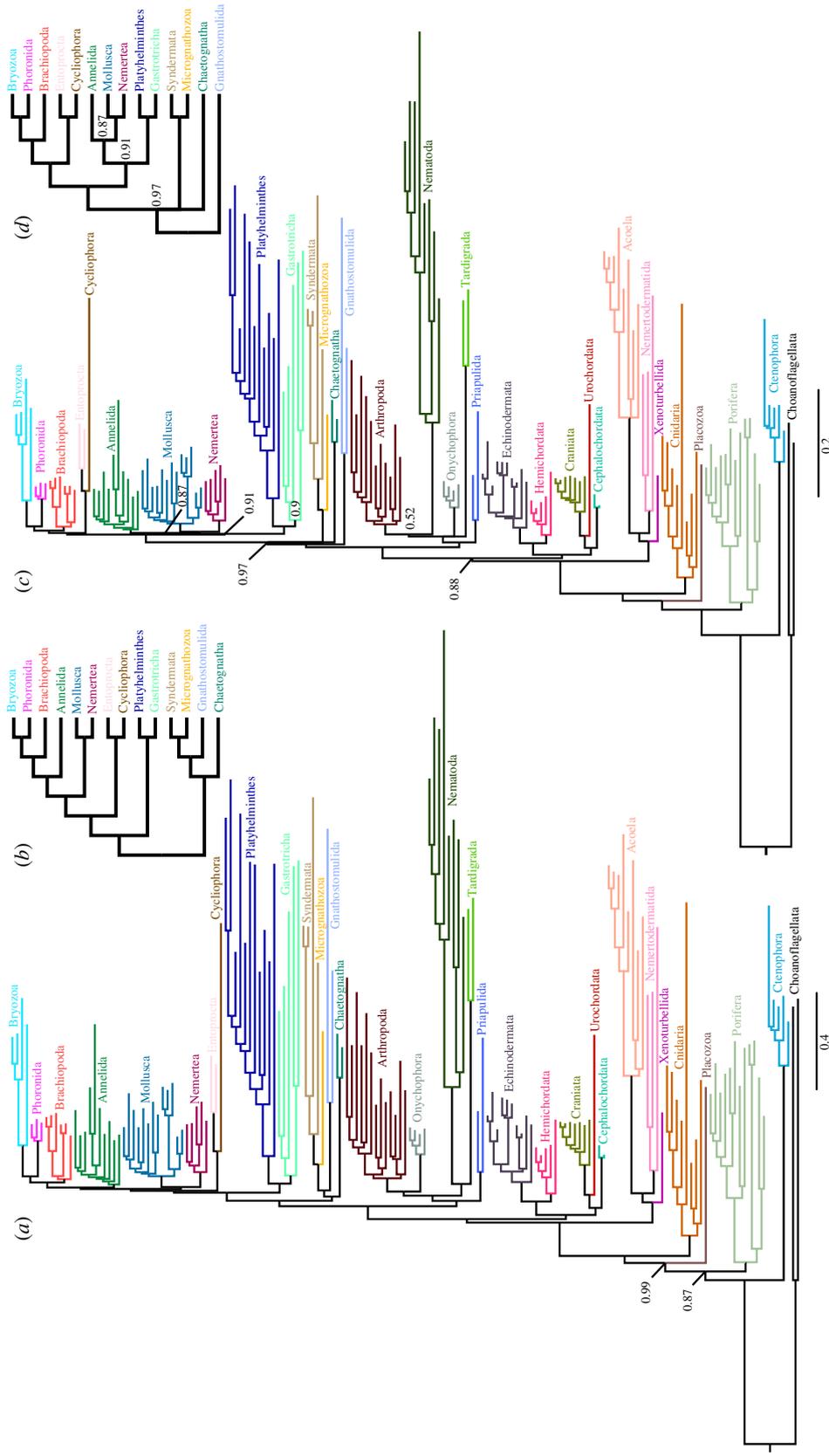


Figure 2. (a) Posterior consensus summary of CAT + GTR + I4 analysis of reduced, BMGE-trimmed pan-Metazoa matrix in amino acid space, trimmed of four rogue taxa prior to summary. (b) Cladogram depiction of relationships and support within Spiralia recovered in the previous analysis, shown to improve readability. (c) Posterior consensus summary of CAT + GTR + I4 analysis of the same matrix, recoded into Dayhoff-6 groups and trimmed of six rogue taxa prior to summary. (d) Cladogram depiction of relationships and support within Spiralia recovered in the previous analysis. Nodal support values are posterior probability; unlabelled nodes received full support. Relationships within labelled clades (phyla) are not annotated with support values to improve visualization, except in the case that the monophyly of the clade in question received less than full posterior probability. (Online version in colour.)

Our results resemble most recent studies in recovering a mostly macrofaunal clade (Lophotrochozoa), sister group to a flatworm–gastrotrich clade (sometimes termed Rouphozoa [39]), both of which are sister to a clade populated by Gnathifera (Gnathostomulida, Micrognathozoa and Syndermata, inclusive of Rotifera and Acanthocephala) [17,28,39,40]. These results, however, differ markedly from a recent analysis also focusing on spiralian relationships, which found Rouphozoa to be non-monophyletic, with Gastrotricha and Platyhelminthes nested separately within Lophotrochozoa [18]. The contrasting approaches taken to mitigate compositional bias in this paper and our own (see Methods) may underlay this discrepancy. However, our analyses are in agreement with this paper in also finding Gnathifera as the sister group of Chaetognatha with full support in the displayed posterior consensus summary (figure 2). We emphasize, however, that support for this split is obtained only when our rogue ecdysozoan taxa are masked; in the unmasked posterior consensus (electronic supplementary material, figure S6), the loriciferan *Armorloricus elegans* is unexpectedly recovered within Gnathifera, as the sister group to Syndermata + Micrognathozoa, albeit its general instability breaks support for this relationship, as well as support for more basal nodes from this point down to the origin of Protostomia. This effect seems idiosyncratic to CAT + GTR + *I*4 analysis of this particular amino acid matrix, as it is not observed in the conjugate Dayhoff-6 recoded analysis of this matrix (electronic supplementary material, figure S8), in ML analysis of the same matrix (electronic supplementary material, figure S5), or in CAT + GTR + *I*4 analyses of the post-BMGE taxon-pruned version of this matrix (electronic supplementary material, figures S3 and S4). However, even when rogue taxa are masked, recoding modifies support for Chaetognatha as sister group to a monophyletic Gnathifera, although all constituent groups are still recovered as spiralian groups falling outside the clade formed by Platyhelminthes, Gastrotricha and Lophotrochozoa (figure 2c). The new relationship between Gnathifera and Chaetognatha is thus supported here, and in all the analyses of Marlétaz *et al.* [18], and was anticipated based on Hox presence data in Rotifera and Chaetognatha [41]. It has also been endorsed by the homology of the jaw elements of chaetognaths to those of the Middle Cambrian Burgess Shale fossil *Amiskwia sagittiformis*, which has a chaetognath-like body plan, but a jaw apparatus reminiscent of gnathiferans [42,43]. In view of the alternating support for Chaetognatha as the sister group to Gnathifera versus being nested within this clade—a duality found both by Marlétaz *et al.* [18] and ourselves in different analyses—we view the decision to declare chaetognaths as crown-group gnathiferans as premature, if not necessarily incorrect.

With adequate transcriptome representation from Entoprocta, Cycliophora, and with greatly improved sampling in Bryozoa, we see remarkable dynamics concerning the relative positions of these groups under different analytical conditions. Even trimmed of many compositionally biased sites, with ML analysis (electronic supplementary material, figures S2 and S10) we recover support for Polyzoa [2] (Entoprocta + Cycliophora + Bryozoa), a grouping previously suspected to represent a compositional artefact [18,26,40,44] as sister group to the remaining Lophotrochozoa. In CAT + GTR + *I*4 analysis of the amino acid matrix (figure 2a), we instead see Bryozoa placed within a lophophorate clade, as the sister group to Phoronida, with Entoprocta + Cycliophora forming the sister group to Lophotrochozoa. However, CAT +

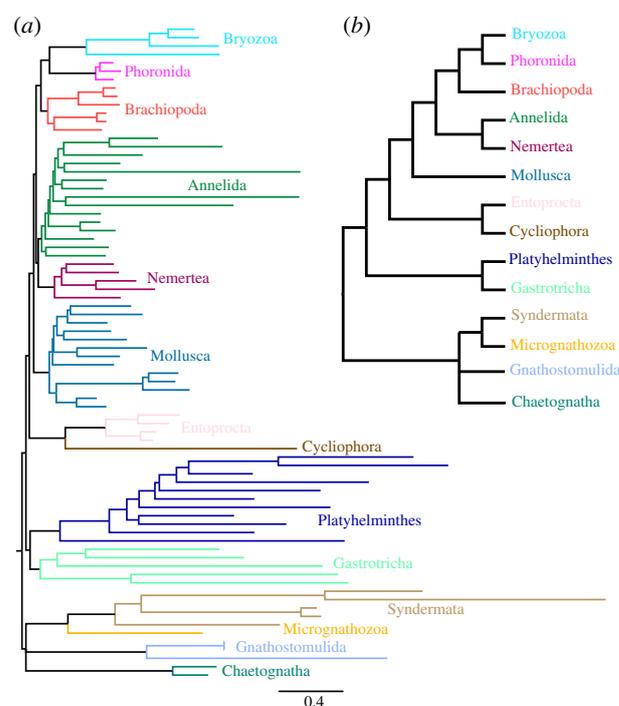


Figure 3. (a) Posterior consensus summary of CAT + GTR + *I*4 analysis of BMGE-trimmed Spiralia-specific matrix in amino acid space, trimmed of 2 rogue taxa prior to summary; the phylogram is drawn with the position of the root taken from the pan-Metazoa results shown in figure 2b. Cladogram depiction of the same, given to improve readability. Criteria for nodal annotation are as in figure 2; in this case, no internodes outside the labelled phyla received less than full support.

GTR + *I*4 analysis of recoded Dayhoff-6 groups yields yet a third possibility, to our knowledge not yet recovered in any molecular phylogeny—a clade of Entoprocta + Cycliophora sister group to a monophyletic Lophophorata, although the position of this clade within Spiralia at large is uncertain in this recoded analysis. On the contrary, Marlétaz *et al.* [18] found Entoprocta to be the sister group of Mollusca, recovering the traditional clade Lacunifera, based on the supposed haemocoel of entoprocts that has been interpreted by some authors as a lacunar circulatory system, similar to that of molluscs [45,46]. Prior analyses have shown that the position of Entoprocta is highly dependent on the presence or absence of Cycliophora, which was pruned prior to analysis in the study of Marlétaz *et al.* [18].

(c) Spiralian matrix (S)

A parallel CAT + GTR + *I*4 analysis of a separate matrix constructed to optimize representation of Entoprocta + Cycliophora, in particular, provides another test of spiralian relationships. Amino acid level results (figure 3) are remarkably similar to the picture seen in the pan-Metazoa matrix, from which the position of the root in this outgroup-lacking tree is taken. Indeed, the only major differences are the lack of support for a monophyletic Gnathifera in this analysis (as it includes Chaetognatha, a result also sometimes found by Marlétaz *et al.* [18]), and relationships within Lophotrochozoa, here with nemerteans and annelids forming a sister group to the lophophorates, with molluscs branching immediately prior to this clade, whereas in the pan-Metazoa analysis (figure 2a) molluscs and nemerteans constitute the sister group to an annelid–lophophorate clade. Both of these scenarios differ

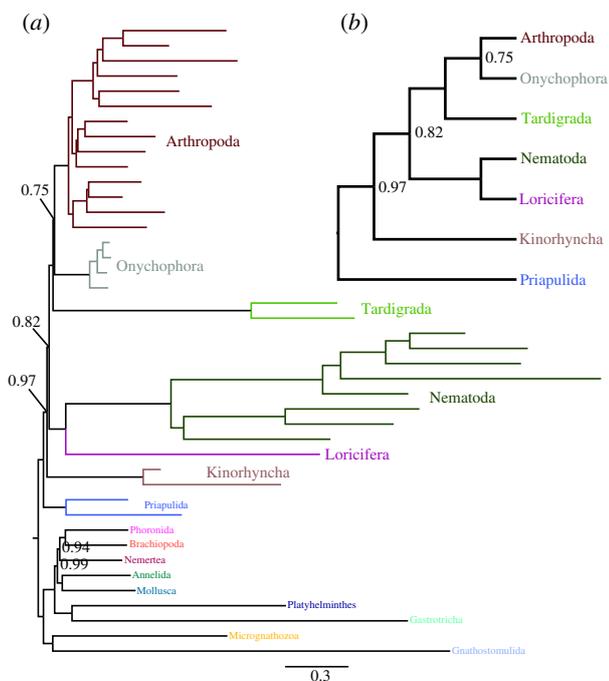


Figure 4. (a) Posterior consensus summary of CAT + GTR + I_4 analysis of BMGE-trimmed Ecdysozoa-specific matrix in amino acid space, trimmed of *Nectonema* sp. (Nematomorpha) due to its behaviour as a rogue taxon prior to summary. (b) Cladogram depiction of ingroup ecdysozoan relationships within this phylogram, given to improve readability. Criteria for nodal annotation are as in figure 2. (Online version in colour.)

from the result recovered by a similar CAT + GTR + I_4 analysis published in 2015 [40]. Dayhoff-6 recoding in this matrix, in contrast to the pan-Metazoa matrix, does not recover a clade of Lophophorata and Entoprocta + Cycliophora; instead, the overall topology is identical to that of the amino acid matrix but in this relatively small matrix support diminishes throughout (electronic supplementary material, figure S13).

(d) Ecdysozoan matrix (E)

Due to the prevalence of rogue ecdysozoan taxa in our pan-Metazoa matrix, we constructed an Ecdysozoa-focused matrix to optimize gene sampling in these species. CAT + GTR + I_4 analysis of this matrix (figure 4) gives preliminary positions for these rogue taxa: for instance, our representative of Loricifera is strongly supported here as an ingroup ecdysozoan sister group to Nematoda. As in the pan-Metazoa matrix, Priapulida falls as the sister group to the remaining members of Ecdysozoa. Kinorhyncha is not recovered as a sister group of Priapulida, instead falling out as the sister group the non-priapulid ecdysozoans (albeit with marginal support); however, we emphasize that although this matrix contains more occupied sites from Kinorhyncha than the pan-Metazoa matrix, both representatives are still only occupied in just over 4000 sites, and thus their position should be taken with caution. The sole representative of Nematomorpha in our dataset, unfortunately, still exhibits rogue taxon behaviour in CAT + GTR + I_4 chains even in this matrix (and is therefore masked in the posterior summary shown in figure 4). However, in mixture-model ML analysis, it is strongly supported as a member of a clade including both Nematoda and Loricifera (electronic supplementary material, figure S18). Interestingly, although this matrix was constructed without reference to this clade, the results from its

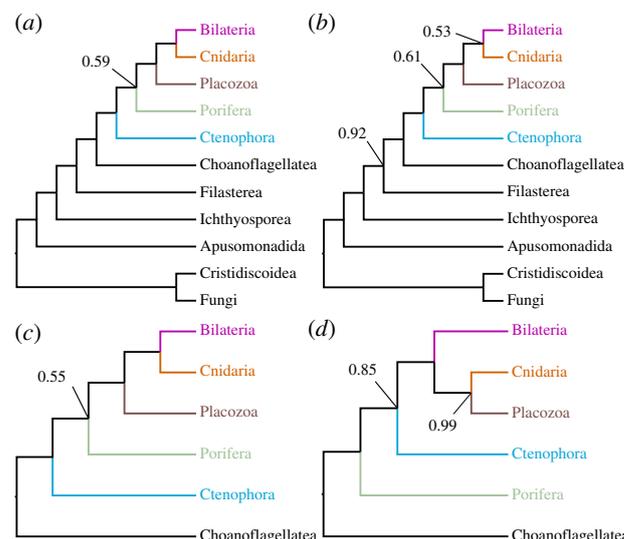


Figure 5. Cladograms depicting metazoan relationships outside Bilateria, summarizing relationships from CAT + GTR + I_4 analysis of the non-Bilateria specific matrix, varied as follows: (a) amino acid matrix with all sampled opisthokont outgroups; (b) the same, recoded into Dayhoff-6 groups; (c) amino acid matrix including only Choanoflagellata as outgroups; (d) the same, recoded into Dayhoff-6 groups. On the matrix pruned of non-choanoflagellate outgroups, BMGE-trimming was performed after pruning, yielding a matrix of 68 337 sites, in contrast to the 61 096 sites retained when all outgroups are included prior to trimming. Trees have been arbitrarily drawn with the root between Apusomonadida and Opisthokonta. (Online version in colour.)

analysis under CAT + GTR + I_4 recover Panarthropoda (Arthropoda + Onychophora + Tardigrada) with full support (figure 4). Curiously, however, support for Arthropoda + Onychophora within this panarthropod clade is lower than in previous studies. Furthermore, under ML, even with a profile mixture model (although less complex than the general CAT + GTR model) we fail to recover Panarthropoda, with Tardigrada strongly supported as the sister group to the Nematoda + Nematomorpha + Loricifera clade (electronic supplementary material, figure S18).

(e) Non-bilaterian matrices (N and N')

Recent work on metazoan relationships outside Bilateria has famously shown contrasting strong support for either Ctenophora or Porifera in the position of the sister group to the rest of Metazoa. It has been claimed that support for Ctenophora in this position even in taxonomically well-sampled datasets [9,10] is an artefact, which can be ameliorated by some combination of using adequately complex substitution models (such as CAT + GTR) [11], deleting compositionally biased, distant outgroups, and/or recoding amino acids into simpler alphabets [12,34]. To test these claims, we constructed a matrix to optimize balanced sampling of non-bilaterians, pruned of detectably compositionally biased and saturated sites (figure 1), and analysed it under CAT + GTR + I_4 in both amino acid and Dayhoff-6 group codings, with or without non-choanoflagellate outgroups. Remarkably, using this matrix, we find no particular support in any of these conditions for either Porifera or Ctenophora as the sister group to the remaining Metazoa (figure 5)—unlike the strong support claimed in many recent analyses addressing this particular issue. While the combination of both Dayhoff-6 recoding and distant outgroup pruning does increase support for Porifera in this position, the posterior probability for this

split is still well below most reasonable significance thresholds (figure 5*d*). Given that other matrices we have analysed under these conditions have recovered strong support for either Ctenophora (figure 2*c*) or Porifera [34] as the deepest-splitting animal phylum, it would appear that the effects of removing distant outgroups and using reduced amino acid alphabets on this problem are matrix-specific, contradicting the assertion that these approaches, in general, lessen systematic error and reveal the true phylogenetic signal in the data. It is, however, interesting to observe that in this matrix, with Dayhoff-6 recoding and removal of distant outgroups, we see strong support for Placozoa as the sister taxon to Cnidaria (figure 5*d*), mirroring the results of a separate recent study undertaken in parallel [34]. This result seems to require both factors, but is possibly more influenced by the recoding, given that support for Planulozoa (= Bilateria + Cnidaria, contra other uses [47]) is still complete in the outgroup-reduced amino acid analysis (figure 5*c*), but heavily diminished in the recoded, outgroup-unpruned analysis (figure 5*b*).

(f) Implications and further directions for metazoan phylogenetics

Relationships within Ecdysozoa to date have not received much attention with genome-scale molecular data, perhaps owing to the rarity and/or limited nucleic acid yield in individuals of such key taxa such as Nematomorpha, Loricifera and Kinorhyncha [37,48]. Our combination of published data with new transcriptomes from representatives of all ecdysozoan taxa aspired to combat this deficiency; however, owing to our minimal taxon sampling of especially long-branched taxa such as Loricifera and Nematomorpha, combined with e.g. limited library complexity from the unamplified kinorhynch cDNA, the conclusions we can make with the data at hand are limited at best and difficult to ameliorate with new data [36] due to the duration of this study, with some analyses running longer than one year. The evidence for a monophyletic Panarthropoda found in CAT + GTR analysis of the Ecdysozoa-specific matrix (figure 4), albeit not in other more general analyses, is in close accord with other molecular evidence for the monophyly of this clade, long supported by uncontroversial morphological apomorphies such as lateral appendages [35,38]. Our results also question the notion of a monophyletic Scalidophora (Loricifera, Kinorhyncha and Priapulida), a morphologically disparate clade united only by the shared presence of innervated scalids, as the sister group to the remaining Ecdysozoa. The non-monophyly of a putative clade composed of Kinorhyncha and Priapulida in our Ecdysozoa-specific mixture model analyses (figure 4; electronic supplementary material, figure S15–S18) should be seen as at best modestly supported, and we emphasize that the matrix occupancy for our sampled kinorhynchs is no more than approximately 4% in any matrix we have analysed. By contrast, despite a still-modest matrix occupancy from the amplified cDNA library of *Armorloricus elegans*, we see strong support for a sister group relationship between Nematoda and this scalidophoran taxon (figure 4), or in trees including them, for a clade comprising Nematoda, Nematomorpha and Loricifera (electronic supplementary material, figures S15 and S18). This result recalls a previous result from ‘universal-marker’ phylogenetics in which Loricifera was recovered as the sister group to Nematomorpha, this clade itself sister to Nematoda [49]. That we report genome-scale evidence for a very similar

position might, therefore, bolster the suggested homologies between larval and adult loriciferans and the nematomorph gordiid larva. However, we emphasize that this result should be seen as provisional, pending a synoptic analysis of Ecdysozoa with good gene and taxon sampling, especially within Nematomorpha and all members of Scalidophora.

Relationships within the spiralian subclade Lophotrochozoa, despite our excellent gene and taxon sampling of this clade, have continued to be volatile in our analyses and contrast with the recent study by Marlétaz *et al.* [18]. One consistent element is the support (figures 2*a* and 3) for a monophyletic Lophophorata, which validates the homology of the lophophore and associated structures in these taxa. The specific topology we recover within Lophophorata, with Phoronida and Byrozoa being well-supported sister taxa, similar to other recent phylogenies that control systematic error [26,50], hearkens to earlier morphological hypotheses of lophophorate phylogeny which homologized the epistome of phylactolaemate bryozoans with that of phoronids [51], for example, and contradicts the assertion principally founded by rRNA phylogenetics that Phoronida represents a subtaxon of Brachiopoda, sister group to Inarticulata [52–54]. In a recoded analysis (figure 2*c*), we have also recovered, for the first time to our knowledge in a molecular phylogeny, strong support for this monophyletic lophophorate clade as the sister group to Entoprocta/Cycliophora. The existence of such a clade, which might imply that the long-branched Entoprocta/Cycliophora are being driven in amino acid analyses outside Lophotrochozoa towards the platyzoan taxa, certainly requires further validation, especially in the context of the recent analysis that, when excluding Cycliophora, place Entoprocta with Mollusca [18]. If corroborated by other analyses, however, our clade may resurrect the core aspect of the Polyzoa hypothesis: that Bryozoa and Entoprocta descend from a common ancestor with asexual budding and/or coloniality [1]. Indeed, considering that the funnel replacement mechanism of cycliophorans is homologous to the process of budding and that many phoronid species also reproduce asexually by transverse fission or budding [55], this would make modern Brachiopoda the only taxon of this sessile clade with U-shaped guts to lack asexual reproduction. However, comparisons between modern taxa may be misleading without reference to the fossil record, which is rich for this lineage. Perhaps most worthy of consideration in light of this topology is *Cotyledion*, interpreted as a macrofaunal, solitary entoproct whose calyx was armed by mineralized sclerites, implying that the minute size and pseudocoelomate nature of modern entoprocts, as well as the absence of sclerites, may be derived features [56]. A sister group relationship between Entoprocta + Cycliophora and Lophophorata would be consistent with the possibility that not only the sessile habit and U-shaped gut of these taxa are homologous, but also that the sclerites of *Cotyledion* and presumably other stem entoprocts might be homologous to those of other fossil lophophorates (e.g. tomotiids [54]), or indeed more deeply to those of other ‘small shelly fossil’ taxa assigned to Lophotrochozoa (e.g. halkieriids [57], possibly cancelloriids [58]). Given the putative spiral cleavage [59] and putative trochophore larva [60] of some species of Entoprocta, this would further bolster the notion that Lophophorata have lost these developmental modes inherited from at least the ancestor of Lophotrochozoa, a scenario which is also becoming increasingly clear on developmental grounds alone [61]. The fact that *Cotyledion* to all appearances is a solitary animal would indicate that either the budding of modern

Bryozoa (and possibly Phoronida) and Entoprocta + Cycliophora are convergent reproductive modes, or that modern Brachiopoda and possibly the stem leading to *Cotyledion* have convergently lost this mode of reproduction. Further knowledge and phylogenetic placement of problematic fossils which show evidence of clonal reproduction, which have also been attributed to total group Lophophorata, may help further clarify this question [62,63].

The recognition that the mostly meiofaunal, acoelomate members of Spiralia, namely Platyhelminthes (which despite its many macrofaunal lineages, was ancestrally surely meiofaunal [64,65]), Gastrotricha and Gnathifera, comprise two separate, deeply splitting lineages within this clade (Rouphozoa and Gnathifera), had been previously used to argue that the spiralian ancestor may, perhaps inheriting traits from an earlier bilaterian ancestor [7], plausibly have been itself a relatively simple, microscopic, acoelomate worm [39,40]. Our analyses complicate this picture. Simply the recognition that Chaetognatha, a group of coelomate macroscopic worms with a complete gut, may form a sister group to Gnathifera, makes it more difficult to reconstruct the spiralian ancestor as an organism superficially similar to a modern platyhelminth or gnathostomid. The placement of Platyhelminthes as sister group to Nemertea, and of Gastrotricha as sister group to Lophophorata within the traditional coelomate spiralians by Marlétaz *et al.* [18]—who, however, chiefly arrive at this result through taxon deletion as a means of mitigating compositional bias—would challenge this hypothesis even further. Indeed, simply considering only our own results, it may be worthwhile to question even the monophyly of Lophotrochozoa (exclusive of Platyhelminthes and Gastrotricha), another foundation of the ‘platyzoan paraphyly’ hypothesis for which support is lacking in both of the recoded analyses including spiralian taxa that we analyse here (figure 2c; electronic supplementary material, figure s13).

4. Conclusion

Resolving the most ancient relationships among animals with large-scale molecular datasets continues to present several frustrating paradoxes, often not recognized in recent publications, which tend to claim resolution with strong support for clades that remain in conflict in this thorough study. As the number of taxa for which sequence data are available grows, at least partly as a result of the prevalence of lineage-specific genes, the number of large orthologues available to study deep relationships with balanced matrix occupancy diminishes. Researchers may feel the need to even further reduce the number of sites in a matrix as the number of taxa increases, since only the most complicated, computationally demanding models yield reasonable results on datasets that span many billions of years of collective evolutionary divergence. Even with tens of thousands of well-aligned sites, stringently validated orthology, and the most flexible, descriptive site-heterogeneous models available today, different gene sets can give full support to conflicting phylogenies. Indeed, even examination of a single matrix under a single model—but using different taxon sets, or masking all but a subset of the recorded substitutions—can yield strong conflicts.

We view such conflicts as reasons for optimism—in true Socratic fashion, they let us know what we do not yet know.

Compared to the situation a decade ago, it is now a relatively small list. We consider the chief outstanding goals to be:

- Verifying the status of Porifera or Ctenophora as the sister group of the remaining metazoans.
- Defining Chaetognatha as a member of Gnathifera or as the sister group of this clade.
- Clarifying which taxon is the sister group of Cnidaria (an issue recently complicated by both conventional molecular phylogenetic analysis [34], analysis of gene family gain and loss [16], and new fossil evidence [66]).
- Continuing to interrogate the position of Xenoacoelomorpha (not addressed here).
- Testing the monophyly of Scalidophora, Panarthropoda and Lophotrochozoa.
- Within the latter, precisely defining the relationships among a monophyletic Lophophorata, Entoprocta + Cycliophora, and the remaining three trochozoan phyla (Annelida, Mollusca and Nemertea).

Defining such fixed-scope problems provides a powerful approach moving forward: clade-specific matrices constructed to test a minimal number of relationships among taxa already demonstrated to be monophyletic allows a much larger number of genes and sites to be examined. Matrices made within such gene sets also are less likely to violate—or perhaps simply less strongly violate—the stationarity assumptions still made by almost all practical phylogenetic inference software: that a single frequency vector can describe composition across the tree [67,68]; that rates of evolution at given sites do not differ among taxa [69]; and that a single substitution matrix accurately describes evolution at a single site among distantly related clades [70]. We hypothesize that such model violations are likely to eventually explain many of the conflicts we and others have seen in metazoan molecular phylogenies. In the near term, we see hope for controlling such violations by limiting the summed patristic distance of a matrix to the minimum required to test relationships with good taxon sampling of the clades in question, and by using sensitive statistical tests to detect and remove sites and genes that show evidence of non-stationarity. Using reduced amino acid alphabets may also mask some forms of non-stationary substitution without removing the site outright, and such reduced state matrices have the advantage of being computationally much simpler to model in useful timeframes. However, the theoretical properties of different recoding schemes remain barely understood [71].

The call to punctiliously discard data which violate the stationarity assumptions used to infer phylogenies will be much easier to meet when highly contiguous, well-annotated genomes—now routinely and economically generated with third generation sequencing [72,73]—are used exclusively; it may be the beginning of the end of the days of using incomplete transcriptome assemblies as an interim approximation to genomes, as done here. Such datasets will also make it much more straightforward to detect genomic changes that bear phylogenetic signal besides those observable in multiple sequence alignments [74,75]. In the long term, however, perhaps the best hope for resolving persistent phylogenetic conflicts in Metazoa and elsewhere will come not from the generation of more data [29], but from analysis of such data with practical, computationally scalable [76] software that

flexibly describes heterogeneity in sequence evolution not only among sites, but also through time.

Data accessibility. All newly sequenced libraries have been given NCBI accessions reported in electronic supplementary material, table S1. Predicted peptides from each species, primary OrthoFinder outputs, the 5511 UPhO groups used for gene selection, all matrices provided as phylip files, and all phylogenetic trees and associated files, are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.293kp3d> [77].

Authors' contributions. Study design and discussion: G.G., S.L., R.F. and C.E.L. Phylogenetic analyses: C.E.L. Writing: C.E.L., G.G. Generated

and assembled molecular data: C.E.L., R.F., S.L., D.C., A.R. and S.A. Obtained specimens: G.G., C.E.L., K.K., A.R., M.V.S. and W.S.

Competing interests. We declare we have no competing interests.

Funding. This research has been supported by projects DEB-1457539 and DEB-0732903 from the US National Science Foundation to G.G. and by internal funds from the Museum of Comparative Zoology and from the Faculty of Arts and Sciences, Harvard University. This article was published as open access through an MCZ grant from the Wetmore Colles Fund.

Acknowledgements. Kira Treibergs is thanked for helping to collect and identify new representatives of Bryozoa sequenced here.

References

- Dunn CW *et al.* 2008 Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452**, 745–749. (doi:10.1038/nature06614)
- Hejnol A *et al.* 2009 Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proc. R. Soc. B* **276**, 4261–4270. (doi:10.1098/rspb.2009.0896)
- Adoutte A, Balavoine G, Lartillot N, Lespinet O, Prud'homme B, De Rosa R. 2000 The new animal phylogeny: reliability and implications. *Proc. Natl Acad. Sci. USA* **97**, 4453–4456. (doi:10.1073/pnas.97.9.4453)
- 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)
- Giribet G. 2016 New animal phylogeny: future challenges for animal phylogeny in the age of phylogenomics. *Org. Divers. Evol.* **16**, 419–426. (doi:10.1007/s13127-015-0236-4)
- Philippe H *et al.* 2011 Acoelomorph flatworms are deuterostomes related to Xenoturbella. *Nature* **470**, 255–258. (doi:10.1038/nature09676)
- Cannon JT, Vellutini BC, Smith J, Ronquist F, Jondelius U, Hejnol A. 2016 Xenacoelomorpha is the sister group to Nephrozoa. *Nature* **530**, 89–93. (doi:10.1038/nature16520)
- Philippe H *et al.* 2019 Mitigating anticipated effects of systematic errors supports sister-group relationship between Xenacoelomorpha and Ambulacraria. *Curr. Biol.* **29**, 9. (doi:10.1016/j.cub.2019.04.009)
- Whelan NV, Kocot KM, Moroz TP, Mukherjee K, Williams P, Paulay G, Moroz LL, Halanych KM. 2017 Ctenophore relationships and their placement as the sister group to all other animals. *Nat. Ecol. Evol.* **1**, 1737. (doi:10.1038/s41559-017-0331-3)
- Whelan NV, Kocot KM, Moroz LL, Halanych KM. 2015 Error, signal, and the placement of Ctenophora sister to all other animals. *Proc. Natl Acad. Sci. USA* **112**, 5773–5778. (doi:10.1073/pnas.1503453112)
- Pisani D, Pett W, Dohrmann M, Feuda R, Rota-Stabelli O, Philippe H, Lartillot N. 2015 Genomic data do not support comb jellies as the sister group to all other animals. *Proc. Natl Acad. Sci. USA* **112**, 15 402–15 407. (doi:10.1073/pnas.1518127112)
- Feuda R, Dohrmann M, Pett W, Philippe H, Rota-Stabelli O, Lartillot N, Pisani D. 2017 Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. *Curr. Biol.* **27**, 3864–3870. (doi:10.1016/j.cub.2017.11.008)
- Simion P *et al.* 2017 A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. *Curr. Biol.* **27**, 958–967. (doi:10.1016/j.cub.2017.02.031)
- Pick KS *et al.* 2010 Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. *Mol. Biol. Evol.* **27**, 1983–1987. (doi:10.1093/molbev/msq089)
- King N, Rokas A. 2017 Embracing uncertainty in reconstructing early animal evolution. *Curr. Biol.* **27**, R1081–R1088. (doi:10.1016/j.cub.2017.08.054)
- Pett W, Adamski M, Adamska M, Francis WR, Eitel M, Pisani D, Dudley J. 2019 The role of homology and orthology in the phylogenomic analysis of metazoan gene content. *Mol. Biol. Evol.* **36**, 643–649. (doi:10.1093/molbev/msz013)
- Kocot KM *et al.* 2017 Phylogenomics of lophotrochozoa with consideration of systematic error. *Syst. Biol.* **66**, 256–282.
- Marlétaz F, Peijnenburg KTC.A., Goto T, Satoh N, Rokhsar DS. 2019 A new spiralian phylogeny places the enigmatic arrow worms among gnathiferans. *Curr. Biol.* **29**, 312–318. (doi:10.1016/j.cub.2018.11.042)
- Laumer CE. 2018 Inferring ancient relationships with genomic data: a commentary on current practices. *Integr. Comp. Biol.* **58**, 623–639. (doi:10.1093/icb/icy075)
- Mikhailov KV *et al.* 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)
- Lu T-M, Kanda M, Satoh N, Furuya H. 2017 The phylogenetic position of dicyemid mesozoans offers insights into spiralian evolution. *Zool. Lett.* **3**, 6. (doi:10.1186/s40851-017-0068-5)
- Schiffer PH, Robertson HE, Telford MJ. 2018 Orthonectids are highly degenerate annelid worms. *Curr. Biol.* **28**, 1970–1974. (doi:10.1016/j.cub.2018.04.088)
- Yang Y, Smith SA. 2014 Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: improving accuracy and matrix occupancy for phylogenomics. *Mol. Biol. Evol.* **31**, 3081–3092. (doi:10.1093/molbev/msu245)
- Ballesteros JA, Hormiga G. 2016 A new orthology assessment method for phylogenomic data: unrooted phylogenetic orthology. *Mol. Biol. Evol.* **33**, 2117–2134. (doi:10.1093/molbev/msw069)
- Simion P *et al.* 2018 A software tool 'CroCo' detects pervasive cross-species contamination in next generation sequencing data. *BMC Biol.* **16**, 28. (doi:10.1186/s12915-018-0486-7)
- 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)
- Crisuolo A, Gribaldo S. 2010 BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, 210. (doi:10.1186/1471-2148-10-210)
- Kocot KM. 2016 On 20 years of Lophotrochozoa. *Org. Divers. Evol.* **16**, 329–343. (doi:10.1007/s13127-015-0261-3)
- Philippe H, Brinkmann H, Lavrov DV, Littlewood DTJ, Manuel M, Baurain D. 2011 Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biol.* **9**, e1000602. (doi:10.1371/journal.pbio.1000602)
- 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**, S4. (doi:10.1186/1471-2148-7-S1-S4)
- 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)
- Whelan NV, Halanych 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)
- Aberer AJ, Krompass D, Stamatakis A. 2013 Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. *Syst. Biol.* **62**, 162–166. (doi:10.1093/sysbio/sys078)

34. Laumer CE, Gruber-Vodicka H, Hadfield MG, Pearse VB, Riesgo A, Marioni JC, Giribet G. 2018 Support for a clade of Placozoa and Cnidaria in genes with minimal compositional bias. *eLife* **7**, e36278. (doi:10.7554/eLife.36278)
35. Borner J, Rehm P, Schill RO, Ebersberger I, Burmester T. 2014 A transcriptome approach to ecdysozoan phylogeny. *Mol. Phylogenet. Evol.* **80**, 79–87. (doi:10.1016/j.ympev.2014.08.001)
36. Yoshida Y *et al.* 2017 Comparative genomics of the tardigrades *Hypsibius dujardini* and *Ramazzottius varieornatus*. *PLoS Biol.* **15**, e2002266. (doi:10.1371/journal.pbio.2002266)
37. Giribet G, Edgecombe GD. 2017 Current understanding of Ecdysozoa and its internal phylogenetic relationships. *Integr. Comp. Biol.* **57**, 455–466. (doi:10.1093/icb/ixc072)
38. Campbell LI *et al.* 2011 MicroRNAs and phylogenomics resolve the relationships of Tardigrada and suggest that velvet worms are the sister group of Arthropoda. *Proc. Natl Acad. Sci. USA* **108**, 15 920–15 924. (doi:10.1073/pnas.1105499108)
39. Struck TH *et al.* 2014 Platyzoan paraphyly based on phylogenomic data supports a noncoelomate ancestry of Spiralia. *Mol. Biol. Evol.* **31**, 1833–1849. (doi:10.1093/molbev/msu143)
40. Laumer CE *et al.* 2015 Spiralian phylogeny informs the evolution of microscopic lineages. *Curr. Biol.* **25**, 2000–2006. (doi:10.1016/j.cub.2015.06.068)
41. Fröbicus AC, Funch P. 2017 Rotiferan Hox genes give new insights into the evolution of metazoan bodyplans. *Nat. Commun.* **8**, 9. (doi:10.1038/s41467-017-00020-w)
42. Vinther J, Parry LA. 2019 Bilateral jaw elements in *Amiskwia sagittiformis* Bridge the morphological gap between gnathiferans and chaetognaths. *Curr. Biol.* **29**, 881–888. (doi:10.1016/j.cub.2019.01.052)
43. Caron J-B, Cheung B. 2019 *Amiskwia* is a large Cambrian gnathiferan with complex gnathostomulid-like jaws. *Commun. Biol.* **2**, 164. (doi:10.1038/s42003-019-0388-4)
44. Nesnidal MP, Helmkampf 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)
45. Haszpruner G, Wanninger A. 2008 On the fine structure of the creeping larva of *Loxosomella murmanica*: additional evidence for a clade of Kamptozoa (Entoprocta) and Mollusca. *Acta Zool.* **89**, 137–148. (doi:10.1111/j.1463-6395.2007.00301.x)
46. Ax P. 1999 *Das System der Metazoa II. Ein Lehrbuch der phylogenetischen Systematik*. Stuttgart, Germany: Gustav Fischer Verlag.
47. Wallberg A, Thollesson M, Farris JS, Jondelius U. 2004 The phylogenetic position of the comb jellies (Ctenophora) and the importance of taxonomic sampling. *Cladistics* **20**, 558–578. (doi:10.1111/j.1096-0031.2004.00041.x)
48. Giribet G. 2016 Genomics and the animal tree of life: conflicts and future prospects. *Zool. Scr.* **45**, 14–21. (doi:10.1111/zsc.12215)
49. Sørensen MV, Hebsgaard MB, Heiner I, Glenner H, Willerslev E, Kristensen RM. 2008 New data from an enigmatic phylum: evidence from molecular sequence data supports a sister-group relationship between Loricifera and Nematomorpha. *J. Zool. Syst. Evol. Res.* **46**, 231–239. (doi:10.1111/j.1439-0469.2008.00478.x)
50. Nesnidal MP, Bruchhaus I, Ebersberger I, Hausdorf B. 2014 Lophophorata monophyletic—after all. In *Deep metazoan phylogeny: the backbone of the tree of life; new insights from analyses of molecules, morphology, and theory of data analysis* (eds JW Wägele, T Bartolomaeus), pp. 127–142. Berlin, Germany: Walter De Gruyter.
51. Zimmer RL. 1973 Morphological and developmental affinities of the lophophorates. In *Living and fossil bryozoa: recent advances in research* (ed. GP Larwood), pp. 593–599. Boca Raton, FL: Academic Press.
52. Sperling EA, Pisani D, Peterson KJ. 2011 Molecular paleobiological insights into the origin of the Brachiopoda. *Evol. Dev.* **13**, 290–303. (doi:10.1111/j.1525-142X.2011.00480.x)
53. Cohen BL. 2013 Rerooting the rDNA gene tree reveals phoronids to be 'brachiopods without shells'; dangers of wide taxon samples in metazoan phylogenetics (Phoronida; Brachiopoda). *Zool. J. Linn. Soc.* **167**, 82–92. (doi:10.1111/j.1096-3642.2012.00869.x)
54. Carlson SJ. 2016 The evolution of Brachiopoda. *Annu. Rev. Earth Planet. Sci.* **44**, 409–438. (doi:10.1146/annurev-earth-060115-012348)
55. Emig CC. 1974 The systematics and evolution of the phylum Phoronida. *J. Zool. Syst. Evol. Res.* **12**, 128–151. (doi:10.1111/j.1439-0469.1974.tb00161.x)
56. Zhang Z *et al.* 2013 A sclerite-bearing stem group entoproct from the early Cambrian and its implications. *Sci. Rep.* **3**, 1066. (doi:10.1038/srep01066)
57. Zhao F, Smith MR, Yin Z, Zeng H, Li G, Zhu M. 2017 *Orthozandrus elongata* n. sp. and the significance of sclerite-covered taxa for early trochozoan evolution. *Sci. Rep.* **7**, 16232. (doi:10.1038/s41598-017-16304-6)
58. Porter SM. 2008 Skeletal microstructure indicates Chancelloriids and halkieriids are closely related. *Palaeontology* **51**, 865–879. (doi:10.1111/j.1475-4983.2008.00792.x)
59. Merkel J, Wollesen T, Lieb B, Wanninger A. 2012 Spiral cleavage and early embryology of a loxosomatid entoproct and the usefulness of spiralian apical cross patterns for phylogenetic inferences. *BMC Dev. Biol.* **12**, 11. (doi:10.1186/1471-213X-12-11)
60. Nielsen C. 2005 Trochophora larvae: cell-lineages, ciliary bands and body regions. 2. Other groups and general discussion. *J. Exp. Zool. B Mol. Dev. Evol.* **304B**, 401–447. (doi:10.1002/jez.b.21050)
61. Vellutini BC, Martín-Durán JM, Hejnol A. 2017 Cleavage modification did not alter blastomere fates during bryozoan evolution. *BMC Biol.* **15**, 33. (doi:10.1186/s12915-017-0371-9)
62. Zhuravlev AY, Wood RA, Penny AM. 2015 Ediacaran skeletal metazoan interpreted as a lophophorate. *Proc. R. Soc. B* **282**, 20151860. (doi:10.1098/rspb.2015.1860)
63. Vinn O, Zatoń M. 2012 Phenetic phylogenetics of tentaculitoids—extinct, problematic calcareous tube-forming organisms. *GFF* **134**, 145–156. (doi:10.1080/11035897.2012.669788)
64. Laumer CE, Giribet G. 2014 Inclusive taxon sampling suggests a single, stepwise origin of ectolecithality in Platyhelminthes. *Biol. J. Linn. Soc.* **111**, 570–588. (doi:10.1111/bij.12236)
65. Laumer CE, Hejnol A, Giribet G. 2015 Nuclear genomic signals of the 'microturbellarian' roots of platyhelminth evolutionary innovation. *eLife* **4**, e05503. (doi:10.7554/eLife.05503)
66. Zhao Y, Vinther J, Parry LA, Wei F, Green E, Pisani D, Hou X, Edgecombe GD, Cong P. 2019 Cambrian sessile, suspension feeding stem-group ctenophores and evolution of the comb jelly body plan. *Curr. Biol.* **29**, 1112–1125. (doi:10.1016/j.cub.2019.02.036)
67. Foster PG, Hickey DA. 1999 Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. *J. Mol. Evol.* **48**, 284–290. (doi:10.1007/PL00006471)
68. Foster PG, Schultz T. 2004 Modeling compositional heterogeneity. *Syst. Biol.* **53**, 485–495. (doi:10.1080/10635150490445779)
69. Philippe H, Zhou Y, Brinkmann H, Rodrigue N, Delsuc F. 2005 Heterotachy and long-branch attraction in phylogenetics. *BMC Evol. Biol.* **5**, 50. (doi:10.1186/1471-2148-5-50)
70. Roure B, Philippe H. 2011 Site-specific time heterogeneity of the substitution process and its impact on phylogenetic inference. *BMC Evol. Biol.* **11**, 17. (doi:10.1186/1471-2148-11-17)
71. Susko E, Roger AJ. 2007 On reduced amino acid alphabets for phylogenetic inference. *Mol. Biol. Evol.* **24**, 2139–2150. (doi:10.1093/molbev/msm144)
72. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017 Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736. (doi:10.1101/gr.215087.116)
73. Grohme MA *et al.* 2018 The genome of *Schmidtea mediterranea* and the evolution of core cellular mechanisms. *Nature* **554**, 56–61. (doi:10.1038/nature25473)
74. Rokas A, Holland PWH. 2000 Rare genomic changes as a tool for phylogenetics. *Trends Ecol. Evol.* **15**, 454–459. (doi:10.1016/S0169-5347(00)01967-4)
75. Boore JL. 2006 The use of genome-level characters for phylogenetic reconstruction. *Trends Ecol. Evol.* **21**, 439–446. (doi:10.1016/j.tree.2006.05.009)
76. Dang T, Kishino H. 2019 Stochastic variational inference for Bayesian phylogenetics: a case of CAT Model. *Mol. Biol. Evol.* **36**, 825–833. (doi:10.1093/molbev/msz020)
77. Laumer CE *et al.* 2019 Data from: Revisiting metazoan phylogeny with genomic sampling of all phyla. Dryad Digital Repository. (doi:10.5061/dryad.293kp3d)