Author for correspondence:
Ruolin Wu
e-mail: ruolin.wu@bristol.ac.uk

THE ROYAL SOCIETY
PUBLISHING

Phylogeny

The unbearable uncertainty of panarthropod relationships

Ruolin Wu[1,2], Davide Pisani[1,3] and Philip C. J. Donoghue[1,2]

[1]Bristol Palaeobiology Group, [2]School of Earth Sciences and [3]School of Biological Sciences, Life Sciences Building, University of Bristol, Tyndall Avenue, Bristol BS8 1TQ, UK

RW, 0000-0002-6782-873X; DP, 0000-0003-0949-6682; PCJD, 0000-0003-3116-7463

Panarthropoda, the clade comprising the phyla Onychophora, Tardigrada and Euarthropoda, encompasses the largest majority of animal biodiversity. The relationships among the phyla are contested and resolution is key to understanding the evolutionary assembly of panarthropod bodyplans. Molecular phylogenetic analyses generally support monophyly of Onychophora and Euarthropoda to the exclusion of Tardigrada (Lobopodia hypothesis), which is also supported by some analyses of morphological data. However, analyses of morphological data have also been interpreted to support monophyly of Tardigrada and Euarthropoda to the exclusion of Onychophora (Tactopoda hypothesis). Support has also been found for a clade of Onychophora and Tardigrada that excludes Euarthropoda (Protarthropoda hypothesis). Here we show, using a diversity of phylogenetic inference methods, that morphological datasets cannot discriminate statistically between the Lobopodia, Tactopoda and Protarthropoda hypotheses. Since the relationships among the living clades of panarthropod phyla cannot be discriminated based on morphological data, we call into question the accuracy of morphology-based phylogenies of Panarthropoda that include fossil species and the evolutionary hypotheses based upon them.

1. Introduction

Euarthropods (Chelicerata, Myriapoda and Pancrustacea—also referred to as Arthropoda [1]) dominate animal biodiversity but the origin of their bodyplans remains unclear due to the uncertainty of phylogenetic relationships with their nearest living relatives. Euarthropods are members of Ecdysozoa, a clade composed of Scalidophora (Kinorhyncha, Lorcifera and Priapulida), Nematoida (Nematoda and Nematomorpha) and Panarthropoda (Euarthropoda, Onychophora and Tardigrada). Conventionally, molecular [2–6] and some morphological [7–13] phylogenetic analyses have supported the Lobopodia hypothesis (=Arthropoda of [1]) in which Euarthropoda and Onychophora are closest relatives; however, this has been challenged by morphology-based phylogenetic analyses that instead support a sister-group relationship between Euarthropoda and Tardigrada (Tactopoda hypothesis) [10,14–20]. The Protarthropoda hypothesis (a clade of onychophorans and tardigrades) is a third rival that has been supported by both molecular [21,22] and morphological [19,23] data. These competing hypotheses impact upon attempts to resolve the relationships of fossil and living ecdysozoans and, consequently, result in contrasting scenarios for the evolutionary assembly of panarthropod bodyplans.

Since support for Tactopoda is rooted in morphology and attempts to resolve bodyplan evolution require integrated phylogenetic analysis of living and fossil taxa, here we explore support for these competing phylogenetic hypotheses

within morphological datasets that have recovered Lobopodia [8,9,13] and Tactopoda [16–18]. Morphology-based phylogenetic analyses are particularly sensitive to taxon and character sampling, as well as methods of phylogenetic inference, principally because of their small size. Through application of parsimony, maximum likelihood and Bayesian phylogenetic inference methods as well as standard statistical tests of phylogenetic support, we show that morphological datasets cannot discriminate among the three competing phylogenetic hypotheses of panarthropod relationships. As such, we should anticipate that the relationships of their fossil relatives and the evolutionary narratives based upon them, are even more weakly supported.

2. Material and methods

To determine whether morphological datasets can discriminate among the three competing phylogenetic hypotheses, we analysed morphological datasets that have been used previously to support them [9,13,18] using parsimony, maximum likelihood and Bayesian methods of phylogenetic inference. Within these frameworks, we explored the sensitivity of the optimal topologies to the data on which the hypotheses are based. We did this by constraining the phylogenetic analyses to the Tactopoda, Lobopodia and Protarthropoda hypotheses.

We evaluated statistical support for these competing topologies using Kishino-Hasegawa (KH) [24], Shimodaira-Hasegawa (SH) [25], approximately unbiased (AU) tests [26] and Bayes factors (BFs) implemented using the stepping stone method [27] or harmonic mean [28]. These tests determine whether, given the data and the model, phylogenetic hypotheses can be distinguished from one another. This approach is particularly important in morphological and palaeontological datasets because of their comparatively small size relative to molecular datasets and the expectation that decisiveness correlates with dataset size [29].

(a) Datasets

A diversity of morphological datasets have been used to resolve panarthropod relationships, but most of these are members of three dataset families, two supporting Lobopodia, the other supporting Tactopoda. As exemplar Lobopodia-supporting datasets, we used Legg *et al.* [9] (henceforth 'Legg dataset'), updated from Legg *et al.* [8] and Rota-Stabelli *et al.* [22], and Aria *et al.* [9] (henceforth 'Aria dataset'), modified from Aria [25]. The Legg dataset is composed of 311 taxa and 753 characters, including 90 extant euarthropods, two extant onychophorans and two extant tardigrades, plus *Caenorhabditis* and *Priapulus* as outgroup taxa. The Aria dataset is composed of 111 taxa and 276 characters, including 36 extant euarthropods, plus Nematoda and Priapulida as the outgroup; the clades of onychophorans and tardigrades are distinguished as 'Onychophora' and 'Tardigrada'. As an exemplar Tactopoda-supporting dataset, we used Yang *et al.* [18], updated from Yang *et al.* [17] and Smith & Ortega-Hernandez [16] (henceforth 'Yang dataset'). The Yang dataset is composed of 50 taxa and 95 characters, including two extant euarthropods, three extant onychophorans and five extant tardigrades, plus *Tubiluchus troglodytes* as an outgroup.

(b) Phylogenetic methods

To control for the impact of competing phylogenetic inference methods, we used PAUP* 4.0 [30] to perform parsimony analyses; Iqtree 2.1.3 [31] to perform maximum-likelihood analyses; and MrBayes v. 3.2.7a [32] to perform Bayesian analyses. For parsimony, characters are unordered and equally weighted. For maximum likelihood, we used the Mk + FQ + R3 model for

the Legg dataset, the Mk + FQ model for the Yang dataset and the Mk + FQ + ASC + G4 model for the Aria dataset. For Bayesian, we used the Mkv+Γ model (Mk and Mkv model see [33]). These maximum likelihood models are the best-fitting models identified for each dataset by Modelfinder [34] in Iqtree according to Bayesian information criterion (for analytical detail see the electronic supplementary material).

(c) Topology tests and model selection method

In an attempt to discriminate among the competing hypotheses, we first conducted unconstrained phylogenetic analyses of the morphological datasets using each of the phylogenetic inference methods. We then undertook three constrained analyses in which partial (backbone) topology constraints were imposed upon the relationships of the living species only; positions of fossil species were unconstrained in all instances. These topological constraints were implemented to be compatible with the three competing hypotheses of panarthropod relationships (figure 1).

To determine whether the data can discriminate decisively between hypotheses, we investigated support levels across the considered topologies and implemented tests to compare pairs of alternative tree topologies. Node support was estimated using bootstrap (1000 replicates) under parsimony and maximum likelihood. For the Bayesian analyses, node support was estimated using posterior probabilities.

To test alternative tree topologies under parsimony and maximum likelihood, we used the KH, SH and AU tests, using $p = 0.05$ as a significance threshold. For the Bayesian analyses, BFs were used to test the relative fit of alternative tree topologies to the data (given the model). For Yang and Aria datasets, we ran stepping-stone analyses [27] to obtain the marginal likelihood values and calculated BFs from their ratio. This approach could not be applied to the larger Legg dataset because of computational tractability and so for this dataset we estimated BFs from harmonic means [28]. BFs were interpreted following Kass & Raftery [35].

3. Results

(a) Unconstrained phylogenetic analyses

Unconstrained parsimony analysis of the three datasets recovered the results reported in the papers from which the datasets were derived, Lobopodia [9,13] and Tactopoda [18] respectively. Maximum-likelihood and Bayesian analyses recovered compatible topologies, though the consensus trees from the Bayesian analyses were often less well resolved (as expected [36,37]) than the maximum-likelihood trees (see electronic supplementary material).

Bootstrap analyses of the two morphological datasets under parsimony and maximum likelihood generated highly unresolved consensus trees, as expected given the Bayesian analysis results (figure 2). In particular, bootstrap support values for the Yang dataset suggest that it is not possible to discriminate the relationships among the three panarthropod phyla. Bootstrap analysis of the Legg dataset recovers Lobopodia (BS-MP = 79%, BS-ML = 59%; BS: bootstrap support. Figure 2*e*,*f*), but the internal relationships are poorly resolved. The bootstrap consensus from the Aria dataset is highly unresolved.

(b) Constrained phylogenetic analyses and topology tests

Under parsimony, constrained analyses of the Yang dataset yielded 4978 most parsimonious trees (MPTs) when
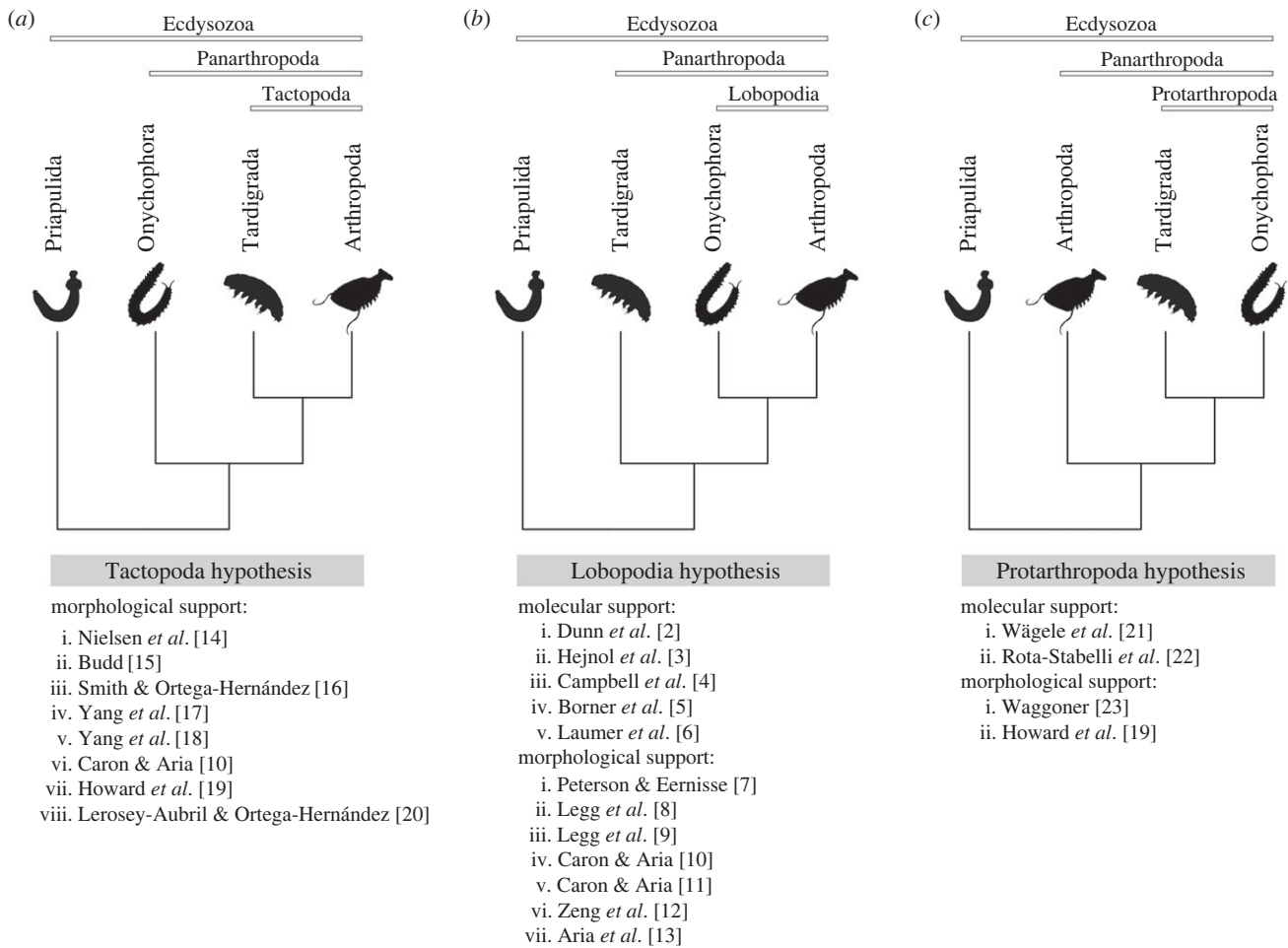
Figure 1. Competing hypotheses within panarthropod lineages and a summary of the support for the topologies of the hypotheses. (a) Tactopoda hypothesis, (b) Lobopodia hypothesis, (c) Protarthropoda hypothesis. Major related studies are listed and details see corresponding references.

Tactopoda was enforced, 6434 when Lobopodia was enforced, and 4034 when Protarthropoda was enforced. The KH test cannot discriminate among hypotheses. Using the KH test, under the Tactopoda constraint, all hypotheses have $p = 1$. Under Lobopodia, p-values range 0.3199–0.6971. Under Protarthropoda, p-values range 0.3199–0.6721 (table 1f). Under maximum likelihood, KH, SH and AU tests yielded p-values ranging 0.145–1.000 under all constraints (table 1g). Stepping stone analysis yielded BFs close to 1 for all, with a difference of less than 0.01, indicating no significant statistical difference between the competing topologies (table 1h and electronic supplementary material).

Constrained analyses of the Legg dataset using parsimony yielded four MPTs when enforcing Tactopoda, nine when enforcing Lobopodia, and 12 when enforcing Protarthropoda. Trees obtained under Lobopodia are shorter. Again, KH tests could not discriminate among hypotheses. p-Values for the Tactopoda trees vary 0.4114–0.5741, while the Protarthropoda trees range 0.2280–0.4331 (table 1j). KH, SH and AU tests could not discriminate among the Tactopoda-, Lobopodia- or Protarthropoda-constrained maximum-likelihood trees, yielding p-values ranging 0.346–1.000. p-Values for Protarthropoda are always the highest for each of the topology tests (table 1k). BFs from constrained analyses of the Legg dataset are all close to 1, indicating no significant difference between the competing topologies (table 1l and electronic supplementary material).

Constrained parsimony analyses of the Aria dataset yielded one MPT when enforcing Tactopoda, three for Lobopodia and two for Protarthropoda. Under maximum parsimony, KH tests cannot discriminate among hypotheses, while all MPTs for Lobopodia have $p = 1$ and two MPTs for Protarthropoda are 0.8353 and 0.8478 (table 1b). However, under maximum likelihood, KH, SH and AU tests discriminate the Lobopodia ($p = 1$) from the other ($p < 0.05$, table 1c). BFs for the three hypotheses are close to 1, meaning that no significant difference could be identified between any two of them (table 1d and electronic supplementary material).

4. Discussion

Resolving evolutionary relationships among the panarthropod phyla is integral to understanding the evolutionary assembly of the arthropod bodyplans, from genomic, developmental and phenotypic perspectives. The pattern of character assembly for the arthropod bodyplans is perhaps the best resolved for all fossil groups, thanks to Cambrian fossil Konservat-Lagerstätten which are dominated by euarthropods [38–40]. This is largely a consequence of the biased preservation of their recalcitrant cuticles and because many key features of the euarthropod bodyplan are represented in their cuticular skeletons [41,42]. However, it is difficult to rationalize the evolutionary significance of such data without first resolving the phylogenetic relationships among the fossil species and this depends, in turn, upon resolution of the evolutionary relationships among their extant relatives. At present, panarthropod evolution is interpreted within three competing and mutually exclusive hypotheses, Lobopodia, Tactopoda and
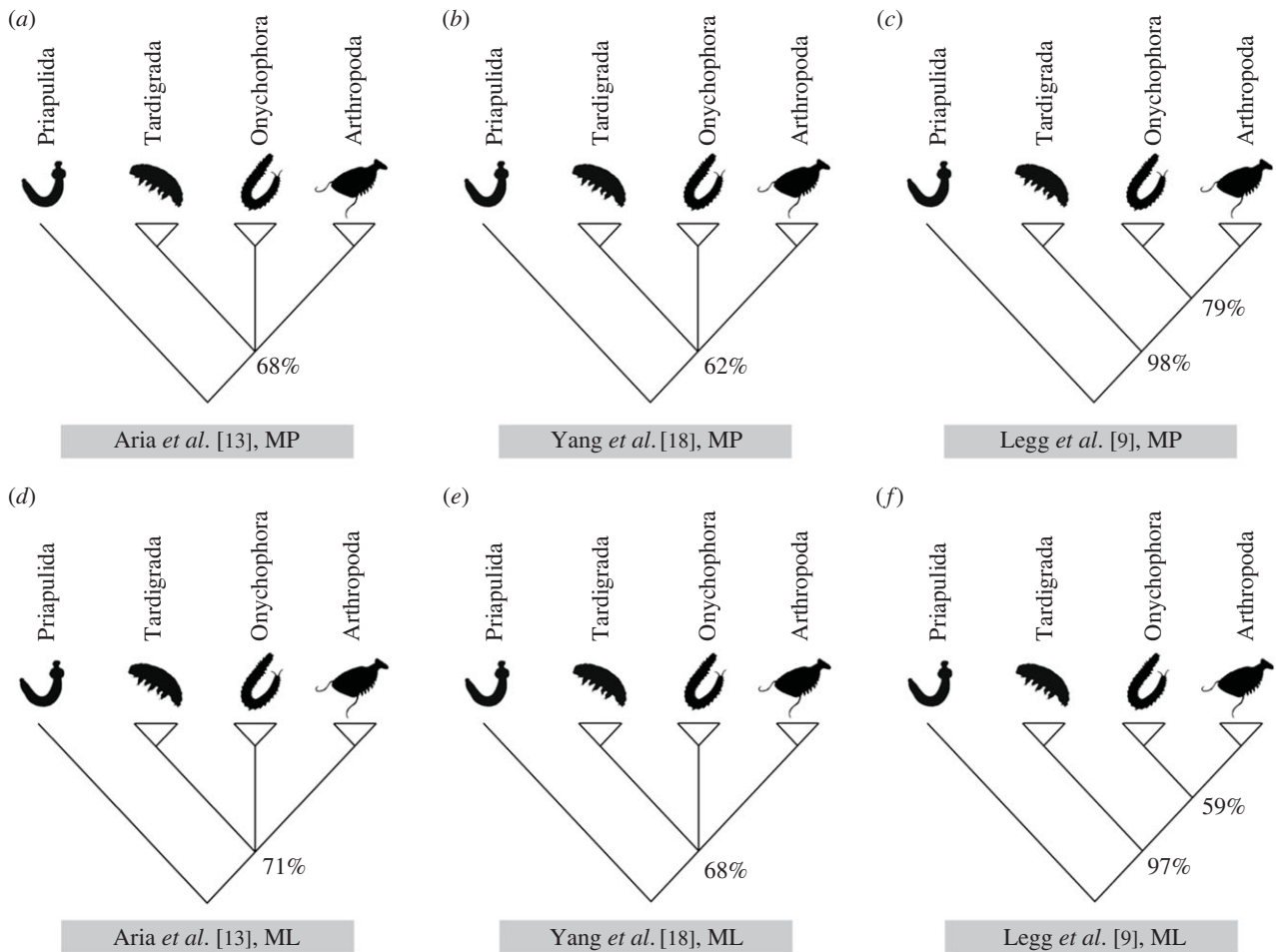
Figure 2. Bootstrap results of the Aria, Yang, and Legg dataset under parsimony and maximum-likelihood criterion. Numbers adjacent to nodes reflect clade support bootstrap results under the parsimony criterion for (*a*) Aria (*c*) Yang and (*e*) Legg dataset; under the parsimony criterion the bootstrap results for (*b*) Aria (*d*) Yang and (*f*) Legg dataset.

Protarthropoda, which have different implications for the assembly of panarthropod bodyplans [39].

We have attempted to discriminate among these competing hypotheses using morphological datasets that have supported Lobopodia and Tactopoda. While it is clearly possible to recover Tactopoda from the parsimony analyses of the Yang dataset and Lobopodia from the Legg and Aria datasets, the phylogenetic signal for these hypotheses is weak in all datasets. The consensus tree from bootstrap analysis of the Yang dataset is uninformative on the relationships of arthropods, onychophorans and tardigrades. The Legg dataset also has a poorly resolved bootstrap consensus but it retains support for Lobopodia (BS-MP = 79%, BS-ML = 59%). Bootstrap analyses of the Aria dataset reveal that it is uninformative under parsimony and maximum likelihood. Bootstrap analyses are harsh tests of the signal in morphological datasets which are usually small in comparison to molecular datasets. Consistent with results from the bootstrap analyses, the results of our topology tests, designed to determine whether a dataset can statistically discriminate between alternative hypotheses, are even less encouraging. We employed KH, SH, AU and BFs to determine whether the Yang, Legg and Aria datasets can discriminate statistically between Tactopoda, Lobopodia and Protarthopoda. Our results show that the Yang and Legg datasets cannot discriminate statistically between any of the prior hypotheses, while the Aria dataset enjoys statistically significant support for Lobopodia, but only under maximum likelihood.

If the relationships of living panarthropods cannot be resolved using morphological data, what store should be put in phylogenies of their fossil relatives, on which our understanding of the assembly of panarthropod bodyplans is based? To be sure, fossils are integral to this endeavour, revealing cryptic homologies among living relatives and, therefore, informing on their evolutionary relationships [43]. Fossils also inform on the pattern and sequence of character evolution, as well as enriching understanding of historical biogeography otherwise based solely on living species [43]. However, this counts for naught unless there is a robust phylogenetic framework for living and fossil species on which to base evolutionary inferences. Two out of the three tested datasets could not resolve panarthropod relationships under any circumstances, suggesting that, minimally, morphological datasets should routinely be tested to evaluate their decisiveness, particularly when used to propose new phylogenetic hypotheses.

Rather than present counsel despair, we highlight some effective solutions to the unbearable uncertainty of panarthropod relationships. First, researchers could explore whether statistically decisive datasets of morphological characters can be assembled; not all of the datasets we analysed were designed explicitly to address panarthropod relationships and so there may be a realistic prospect that progress is possible (indeed, there is some statistical support for Lobopodia in the Aria dataset). Second, rather than striving for fully resolved but weakly supported phylogenetic hypotheses, palaeontologists and

Table 1. Results of constrained analyses and topology tests by using the Aria, Yang and Legg dataset. (a–d) Results for Aria dataset, (e–h) results for Yang dataset, (i–l) results for Legg dataset. (a, e, i) The number of the extant species of different clades in the dataset of three datasets, (b, f, j) KH test result under parsimony trees of three various hypotheses, (c, g, k) KH, SH and AU test results under maximum likelihood criterion of hypotheses, and (d, h, l) the marginal-likelihood values and Bayes factors calculated from them; '/' means division, T, L, P: Tactopoda, Lobopodia, Protarthropoda hypothesis.

analyses using dataset from Aria et al. [13]

(a) information of the dataset

classification	number of extant species
outgroup	2
Onychophora	1
Tardigrada	1
Arthropoda	36

(b) tests under parsimony

hypothesis	number of parsimony trees	range of the p-value of Kishino-Hasegawa (KH) test
Tactopoda	1	(shortest length tree)
Lobopodia	3	all are 1.0000
Protarthropoda	2	0.8353, 0.8478

(c) tests under maximum likelihood

hypothesis	p-value KH test	p-value SH test	p-value AU test
Tactopoda	0	5.00×10^{-5}	3.07×10^{-6}
Lobopodia	1	1	1
Protarthropoda	0	0	2.23×10^{-8}

(d) tests under Bayesian

hypothesis	marginal likelihood	Bayes factor ratio of models	Bayes factor value
Tactopoda	−4547.380	T/L	≈1.001
Lobopodia	−4543.720	L/P	≈0.999
Protarthropoda	−4547.410	T/P	≈1.000

analyses using dataset from Yang et al. [18]

(e) information of the dataset

classification	number of extant species
outgroup	1
Onychophora	3
Tardigrada	5
Arthropoda	2

(f) tests under parsimony

hypothesis	number of parsimony trees	range of the p-value of Kishino-Hasegawa (KH) test
Tactopoda	4978	All are 1.0000
Lobopodia	6434	0.3199−0.6971
Protarthropoda	4034	0.3199−0.6721

(g) tests under maximum likelihood

hypothesis	p-value KH test	p-value SH test	p-value AU test
Tactopoda	0.743	1.000	0.856
Lobopodia	0.257	0.257	0.145
Protarthropoda	0.257	0.257	0.145

(h) tests under Bayesian

hypothesis	marginal likelihood	Bayes factor ratio of models	Bayes factor value
Tactopoda	−708.740	T/L	≈0.997
Lobopodia	−710.730	L/P	≈1.002
Protarthropoda	−709.240	T/P	≈0.999

analyses using dataset from Legg et al. [9]

(i) information of the dataset

classification	number of extant species
outgroup	2
Onychophora	2
Tardigrada	2
Arthropoda	90

(j) tests under parsimony

hypothesis	number of parsimony trees	range of the p-value of Kishino-Hasegawa (KH) test
Tactopoda	4	0.4114−0.5741
Lobopodia	9	all are 1.0000
Protarthropoda	12	0.2280−0.4331

(k) tests under maximum likelihood

hypothesis	p-value KH test	p-value SH test	p-value AU test
Tactopoda	0.346	0.553	0.380
Lobopodia	0.378	0.488	0.403
Protarthropoda	0.654	1.000	0.637

(l) tests under Bayesian

hypothesis	harmonic mean	Bayes factor ratio of models	Bayes factor value
Tactopoda	−9441.53	T/L	≈1.001
Lobopodia	−9432.82	L/P	≈1.000
Protarthropoda	−9435.85	T/P	≈1.001

other morphologists should embrace uncertainty, inferring evolutionary history based on those statistically robust relationships that can be resolved. Third, morphology is not the only source of pertinent data and phylogenetic analyses of living and fossil panarthropods can be conducted within the constraint of phylogenetic hypotheses informed by molecular data. Molecular phylogenetic analyses have supported a diversity of hypotheses of ecdysozoans intra-relationships, particularly in terms of the phylogenetic position Tardigrada (sister to Lobopodia [4,6] versus sister to Nematoda [2,3,5,6]). However, Lobopodia is almost universally supported by current phylogenomic datasets, irrespective of whether or not Tardigrada is recovered as a member of Panarthropoda [2–6]. Based on current evidence we suggest that Tactopoda should be considered unsupported and Lobopodia, which is supported by multiple lines of evidence, should be the preferred working hypotheses for panarthropod relationships. Finally, perhaps the most progressive and effective solution is to stop discriminating between molecular and morphological data and instead marshall all data relevant to the phylogenetic question. Methods for combined analysis of morphological and molecular data are now widely available (e.g. [31]) and can be used to provide an integrated understanding of the evolutionary relationships and evolutionary history of Panarthropoda.

References

1. Ortega-Hernández J. 2016 Making sense of 'lower' and 'upper' stem-group Euarthropoda, with comments on the strict use of the name Arthropoda von Siebold, 1848. *Biol. Rev.* **91**, 255–273. (doi:10.1111/brv.12168)

2. Dunn CW *et al.* 2008 Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452**, 745–749. (doi:10.1038/nature06614)

3. 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)

4. 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**, 15920–15924. (doi:10.1073/pnas.1105499108)

5. 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)

6. 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)

7. Peterson KJ, Eernisse DJ. 2001 Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences. *Evol. Dev.* **3**, 170–205. (doi:10.1046/j.1525-142x.2001.003003170.x)

8. Legg DA, Sutton MD, Edgecombe GD, Caron J-B. 2012 Cambrian bivalved arthropod reveals origin of arthrodization. *Proc. R. Soc. B* **279**, 4699–4704. (doi:10.1098/rspb.2012.1958)

9. Legg DA, Sutton MD, Edgecombe GD. 2013 Arthropod fossil data increase congruence of morphological and molecular phylogenies. *Nat. Commun.* **4**, 2485. (doi:10.1038/ncomms3485)

10. Caron J-B, Aria C. 2017 Cambrian suspension-feeding lobopodians and the early radiation of panarthropods. *BMC Evol. Biol.* **17**, 29. (doi:10.1186/s12862-016-0858-y)

11. Caron J, Aria C. 2020 The Collins' monster, a spinous suspension-feeding lobopodian from the Cambrian Burgess Shale of British Columbia. *Palaeontology* **63**, 979–994. (doi:10.1111/pala.12499)

12. Zeng H, Zhao F, Niu K, Zhu M, Huang D. 2020 An early Cambrian euarthropod with radiodont-like raptorial appendages. *Nature* **588**, 101–105. (doi:10.1038/s41586-020-2883-7)

13. Aria C, Zhao F, Zhu M. 2021 Fuxianhuiids are mandibulates and share affinities with total-group Myriapoda. *J. Geol. Soc.* **178**, jgs2020-246. (doi:10.1144/jgs2020-246)

14. Nielsen C, Scharff N, Eibye-Jacobsen D. 1996 Cladistic analyses of the animal kingdom. *Biol. J. Linn. Soc.* **57**, 385–410. (doi:10.1111/j.1095-8312.1996.tb01857.x)

15. Budd GE. 2001 Tardigrades as 'stem-group arthropods': the evidence from the Cambrian fauna. *Zool. Anz. - J. Comp. Zool.* **240**, 265–279. (doi:10.1078/0044-5231-00034)

16. Smith MR, Ortega-Hernández J. 2014 Hallucinaut's onychophoran-like claws and the case for Tactopoda. *Nature* **514**, 363–366. (doi:10.1038/nature13576)

17. Yang J, Ortega-Hernández J, Gerber S, Butterfield NJ, Hou J, Lan T, Zhang X. 2015 A superarmored lobopodian from the Cambrian of China and early

disparity in the evolution of Onychophora. *Proc. Natl Acad. Sci. USA* **112**, 8678–8683. (doi:10.1073/pnas.1505596112)

18. Yang J, Ortega-Hernández J, Butterfield NJ, Liu Y, Boyan GS, Hou J, Lan T, Zhang X. 2016 Fuxianhuiid ventral nerve cord and early nervous system evolution in Panarthropoda. *Proc. Natl Acad. Sci. USA* **113**, 2988–2993. (doi:10.1073/pnas.1522434113)

19. Howard RJ, Hou X, Edgecombe GD, Salge T, Shi X, Ma X. 2020 A tube-dwelling early Cambrian lobopodian. *Curr. Biol.* **30**, 1529–1536. (doi:10.1016/j.cub.2020.01.075)

20. Lerosey-Aubril R, Ortega-Hernández J. 2022 A new lobopodian from the middle Cambrian of Utah: did swimming body flaps convergently evolve in stem-group arthropods? *Pap. Palaeontol.* **8**, e1450. (doi:10.1002/spp2.1450)

21. Wägele JW, Erikson T, Lockhart P, Misof B. 1999 The Ecdysozoa: artifact or monophylum? *J. Zool. Syst. Evol. Res.* **37**, 211–223. (doi:10.1111/j.1439-0469.1999.tb00985.x)

22. Rota-Stabelli O, Campbell L, Brinkmann H, Edgecombe GD, Longhorn SJ, Peterson KJ, Pisani D, Philippe H, Telford MJ. 2011 A congruent solution to arthropod phylogeny: phylogenomics, microRNAs and morphology support monophyletic Mandibulata. *Proc. R. Soc. B* **278**, 298–306. (doi:10.1098/rspb.2010.0590)

23. Waggoner BM. 1996 Phylogenetic hypotheses of the relationships of arthropods to Precambrian and Cambrian problematic fossil taxa. *Syst. Biol.* **45**, 190–222. (doi:10.1093/sysbio/45.2.190)

24. Kishino H, Hasegawa M. 1989 Evaluation of the maximum likelihood estimate of the evolutionary

tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* **29**, 170–179. (doi:10.1007/BF02100115)

25. Shimodaira H, Hasegawa M. 1999 Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**, 1114–1116. (doi:10.1093/oxfordjournals.molbev.a026201)

26. Shimodaira H. 2002 An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* **51**, 492–508. (doi:10.1080/10635150290069913)

27. Xie W, Lewis PO, Fan Y, Kuo L, Chen M-H. 2011 Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst. Biol.* **60**, 150–160. (doi:10.1093/sysbio/syq085)

28. Nylander JAA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey J. 2004 Bayesian phylogenetic analysis of combined data. *Syst. Biol.* **53**, 47–67. (doi:10.1080/10635150490264699)

29. Puttick MN *et al.* 2017 Uncertain-tree: discriminating among competing approaches to the phylogenetic analysis of phenotype data. *Proc. R. Soc. B* **284**, 20162290. (doi:10.1098/rspb.2016.2290)

30. Swofford DL. 2003 PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, MA: Sinauer Associates.

31. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020 IQ-TREE 2: new models and efficient methods for phylogenetic inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534. (doi:10.1093/molbev/msaa015)

32. Ronquist F *et al.* 2012 MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542. (doi:10.1093/sysbio/sys029)

33. Lewis PO. 2001 A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* **50**, 913–925. (doi:10.1080/106351501753462876)

34. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589. (doi:10.1038/nmeth.4285)

35. Kass RE, Raftery AE. 1995 Bayes factors. *J. Am. Stat. Assoc.* **90**, 773–795. (doi:10.1080/01621459.1995.10476572)

36. O'Reilly JE, Puttick MN, Parry L, Tanner AR, Tarver JE, Fleming J, Pisani D, Donoghue PCJ. 2016 Bayesian methods outperform parsimony but at the expense of precision in the estimation of phylogeny from discrete morphological data. *Biol. Lett.* **12**, 20160081. (doi:10.1098/rsbl.2016.0081)

37. Puttick MN *et al.* 2017 Parsimony and maximum-likelihood phylogenetic analyses of morphology do not generally integrate uncertainty in inferring evolutionary history: a response to Brown *et al.*

Proc. R. Soc. B **284**, 20171636. (doi:10.1098/rspb.2017.1636)

38. Aria C. 2022 The origin and early evolution of arthropods. *Biol. Rev.* **97**, 1786–1809.

39. Edgecombe GD. 2020 Arthropod origins: integrating paleontological and molecular evidence. *Annu. Rev. Ecol. Evol. Syst.* **51**, 1–25. (doi:10.1146/annurev-ecolsys-011720-124437)

40. Ortega-Hernández J, Janssen R, Budd GE. 2019 The last common ancestor of Ecdysozoa had an adult terminal mouth. *Arthropod Struct. Dev.* **49**, 155–158. (doi:10.1016/j.asd.2018.11.003)

41. Butterfield NJ. 2003 Exceptional fossil preservation and the cambrian explosion. *Integr. Comp. Biol.* **43**, 166–177. (doi:10.1093/icb/43.1.166)

42. Butler AD, Cunningham JA, Budd GE, Donoghue PCJ. 2015 Experimental taphonomy of *Artemia* reveals the role of endogenous microbes in mediating decay and fossilization. *Proc. R. Soc. B* **282**, 20150476. (doi:10.1098/rspb.2015.0476)

43. Patterson C. 1981 Significance of fossils in determining evolutionary relationships. *Annu. Rev. Ecol. Syst.* **12**, 195–223. (doi:10.1146/annurev.es.12.110181.001211)

44. Wu R, Pisani D, Donoghue PCJ. 2023 The unbearable uncertainty of panarthropod relationships. Figshare. (doi:10.6084/m9.figshare.c.6340227)