Improved resolution of recalcitrant nodes in the animal phylogeny through the analysis of genome gene content and morphology

Juravel, Ksenia[1]; Porras, Luis[1]; Höhna, Sebastian[1,2]; Pisani, Davide[3]; Wörheide, Gert[1,2,4, *]

[1] Department of Earth and Environmental Sciences, Paleontology & Geobiology, Ludwig-Maximilians-Universität München, Richard-Wagner-Str. 10, 80333 München, Germany

[2] GeoBio-Center, Ludwig-Maximilians-Universität München, Richard-Wagner-Str. 10, 80333 München, Germany

[3] School of Biological Sciences and School of Earth Sciences, University of Bristol, UK.

[4] SNSB-Bayerische Staatssammlung für Paläontologie und Geologie, Richard-Wagner-Str. 10, 80333 München, Germany

* corresponding author email: woerheide@lmu.de

Abstract

An accurate phylogeny of animals is needed to clarify their evolution, ecology, and impact on shaping the biosphere. Although datasets of several hundred thousand amino acids are nowadays routinely used to test phylogenetic hypotheses, key deep nodes in the metazoan tree remain unresolved: the root of animals, the root of Bilateria, and the root of Deuterostomia. To assess patterns of congruence with established amino-acid derived phylogenetic hypotheses for these problematic nodes, we independently and extensively analysed newly assembled genome gene content and morphological datasets. Our datasets strongly support sponges as the sister group of all the other animals, the worm-like bilaterian lineage Xenacoelomorpha as the sister group of the other Bilateria, and largely support monophyletic Deuterostomia. We conclude that the last common animal ancestor may have been a simple, microphagous organism without a nervous system and muscles, while the last common ancestor of Bilateria might have been a small, acoelomate-like worm without a through gut.

Introduction

Large multi-gene amino acid sequence (phylogenomic) datasets promised to achieve the phylogenetic resolution [1] needed to understand the evolution of life accurately [2]. These phylogenies enable inferences about the phenotype, physiology, and ecology of common ancestors of clades [3,4], and to test hypotheses about the emergence of key innovations such as the nervous- and digestive systems [5,6].

However, modelling the evolution of amino acid sequences is difficult [7,8]. Deep metazoan phylogenies reconstructed from alternative amino acid datasets, or even the same amino acid dataset analysed using different substitution models [4,9–11], as well as using different taxon samplings of the ingroup [12,13] and the outgroup [9,10], are frequently incongruent. This acknowledged model- and data dependency of phylogenomic analyses underpins the phylogenetic instability observed towards the root of the animal tree [e.g., 14].

Although the sister group of all animals is well established – the Choanoflagellata, a group of single-celled and sometimes colonial collared and flagellated eukaryotes [15] – three nodes towards the root of the animal tree are proving difficult to resolve using multi-gene amino acid datasets, hindering progress in understanding early animal evolution [16].

The first recalcitrant node in the animal tree is its root, and the discussion largely centres around the question of whether sponges (Porifera) or comb jellies (Ctenophora) are the sister group of all the other animals [17,18]. This controversy impinges on our understanding of the last common ancestor of Metazoa [19], and despite receiving much attention for more than a decade [9,10,12,13,18,20–27], it is not yet resolved.

Two other recalcitrant nodes have more recently been identified from alternative analyses of amino acid datasets that affect our understanding of the root of the Bilateria (all bilaterally

symmetrical animals, including humans). The first node involves the position of the worm-like Xenacoelomorpha, a bilaterian clade that unites the Acoelomorpha and Xenoturbellida [28]. With a few exceptions [29], Xenacoelomorpha are millimetre-sized and primarily benthic or sediment dwelling bilaterians devoid of a true body cavity and an anus. Xenacoelomorpha has been recovered in different positions in the animal tree: as the sister group of all other bilaterian animals (Nephrozoa) [4,29], or as the sister group of the Ambulacraria (Echinodermata+Hemichordata) constituting the clade Xenambulacraria [11,30]. The second node concerns the Deuterostomia, one of the two main bilaterian lineages ("Superphyla"). Bilateria have long been split into two lineages, Protostomia (Ecdysozoa + Spiralia [Lophotrochozoa]) and Deuterostomia (traditionally: Chordata + Ambulacraria [= Hemichordata + Echinodermata])[31], historically based on the different origins of the mouth and other features during development [32]. However, recent phylogenomic studies challenged the monophyly of Deuterostomia and recovered paraphyletic deuterostomes in conjunction with Xenambulacraria [33,34]. This combination of results, if confirmed, would have substantial implications for our understanding of the last common ancestor of all Bilateria, which might then have been a fairly large organism, with pharyngeal gill slits and other traits previously thought to represent apomorphies of Deuterostomia (see [34] for an in-depth discussion).

Accordingly, a stable resolution of the relationships of Xenacoelomorpha with reference to the deuterostomes is key to correctly infer the condition of the last common ancestor of the Bilateria – a small and simple organism if Xenacoelomorpha are the sister group to the Nephrozoa, or a larger and much more complex organism if Xenambulacraria is correct and Deuterostomia is not monophyletic.

Considering that phylogenomic analyses are model- and data dependent, we must employ different and novel approaches to select between phylogenetic hypotheses. One way is to use model fit- and model adequacy tests to discriminate between alternatives, favouring those derived using the best-fitting and most adequate model(s) [9,26]. Alternatively, simulations can be used to compare alternative tree topologies and their chance of being inferred under different models [30]. Finally, independent data sources can be used to "triangulate" conflicting hypotheses [35].

Here we use two independent data types, genome gene content ("gene content") data and morphology, to evaluate alternative hypotheses of animal relationships that emerged from previous analyses of amino acid sequence data and investigate their relative consilience [36,37]. We focus on the three recalcitrant nodes mentioned above: the relative relationships of sponges and comb jellies with respect to the other animals, the relationships of Xenacoelomorpha within the Bilateria, and the monophyly of Deuterostomia.

The phylogenetic analysis of gene content data relies on the proteomes derived from fully sequenced genomes and converts the presence or absence of gene families in the genomes of the terminals into a binary data matrix [9,25,38,39]. We constructed separate datasets for "Homogroups" (homologous gene families) and "Orthogroups" (orthologous gene families). The former include homologous proteins that are predicted to be inherited from a common ancestor and can contain orthologs, xenologs, and out-paralogs, whereas the latter contains only proteins predicted to be inherited from a common ancestor and separated by a speciation event (see Methods for details).

We assembled a large number of new gene content datasets (see Methods, Fig. 1) to extensively test the effect of different parameter combinations when identifying homogroups

and orthogroups, because this crucial step remains a challenge [40,41] and may influence the outcome of the downstream phylogenetic analysis [42].

We also compiled different datasets to extensively evaluate other potential sources of error, such as the so-called "long branch attraction" (LBA) artefact [43] (see Methods, Fig. 1). LBA occurs when two (or more) long branches in a phylogenetic tree group together without true relationship, generating "phylogenetic artefacts" [7]. Previous gene content analyses have focused on the root of the animals. Accordingly, here we primarily focus our LBA assessment on Xenacoelomorpha by performing taxon exclusion experiments in an approach similar to Philippe et al. [11].

Additionally, we carefully collated a 770-character morphological data matrix. As a starting point, we built on the classical work of Peter Ax [44] that was systematised by Deline et al. [45], and introduced additional information from two other reputable datasets [46,47] to build our matrix. All characters were meticulously reassessed before being included in our new dataset, and the coding of the base set was updated based on current morphological interpretations for groups such as Ecdysozoa and Xenacoelomorpha. In order to avoid artefacts caused by the lack of character comparability across the tree, we utilised two different coding strategies: non-additive and reductive coding (see Methods for details). Because the non-additive coding may be affected by taxa with many uncertain states, we ran the analyses with a reduced outgroup set, which retained only the Choanoflagellata, the sister group of animals [19]. Other taxa exclusion experiments include runs without the taxa that showed problematic behaviour in the gene content analyses, the longest branches in the morphological trees, and parts of Xenacoelomorpha to check robustness. Finally, we extensively explored several modelling assumptions of morphological character evolution to explore the robustness of our analyses (see Methods).

Altogether, our results provide strong support for the view that sponges are the sister group of the rest of the animals (consistent with the Porifera-sister hypothesis) and that Xenacoelomorpha are the sister group of the rest of the Bilateria (consistent with the Nephrozoa hypothesis). Monophyletic Deuterostomia is also largely supported.

Results

Genome gene content data analyses

47 genome-derived proteomes were used to initially generate and analyse a total of 190 gene content datasets of different taxon samplings and parameter combinations (see Methods and data repository for details). The datasets were partitioned into several groups due to the different approaches applied (see below), all taxon sub-samplings and different parameter combinations were done in parallel for homologous gene families ("homogroups") and orthologous gene families ("orthogroups") [38] (Fig. 1). To assess the reproducibility of the results, the construction and analysis of the different datasets was performed twice (for results of the replicated analyses see Supp. Fig. 5; see the data repository for a more detailed explanation).

To test whether the specific phylogenetic relationships of Xenacoelomorpha with reference to Deuterostomia were affected by LBA, different taxon sampling experiments, based on a core taxon set of 40 species, were performed by defining three groups of datasets (Fig. 1): the "Opi" (Opisthokonta) group that consisted of all the datasets scoring a complete set of 47 taxa, including full outgroups. The "Aco" group consisted of all datasets that excluded *Xenoturbella* from the Opi dataset, and the "Xen" group consisted of all datasets that excluded the Acoelomorpha from the Opi dataset. Opi, Aco, and Xen included datasets with different parameter combinations for orthogroups and homogroups, resulting in 120 datasets in total (Fig. 1, see Methods for details).

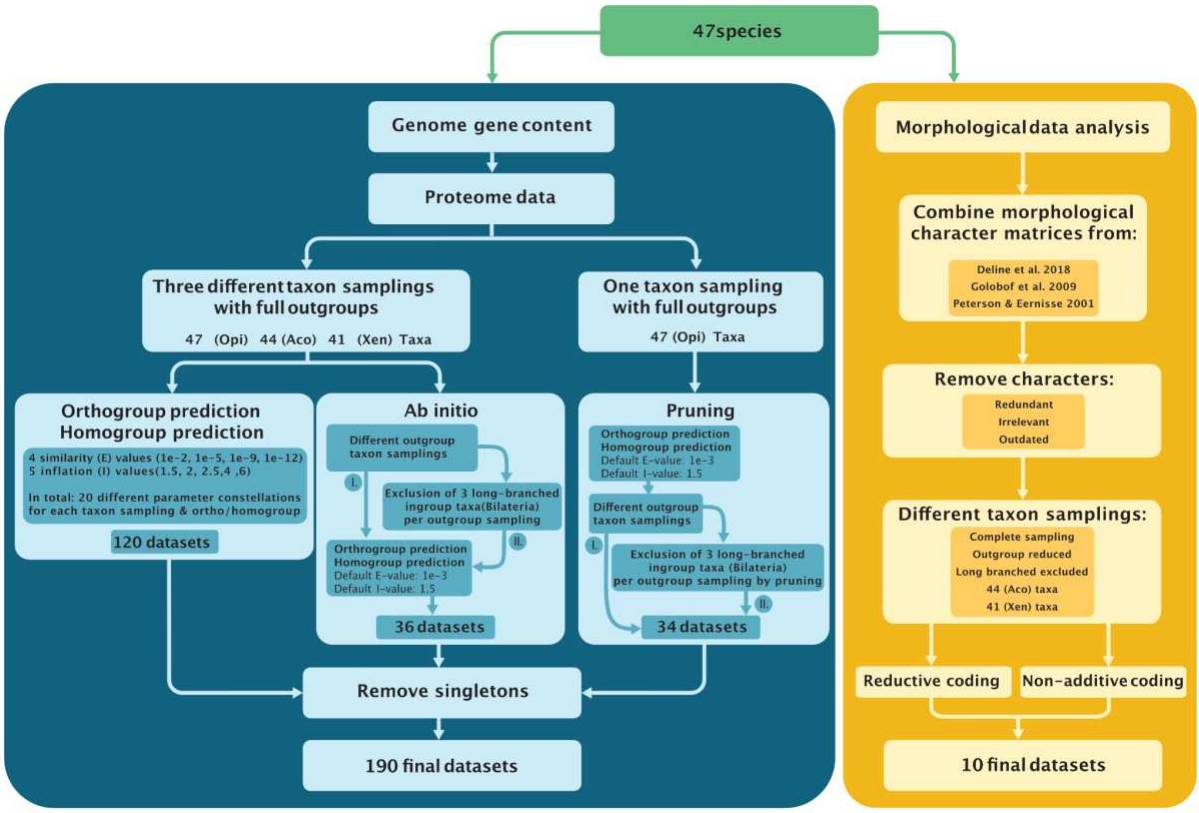


Figure 1: Concise graphical illustration of the methodology and workflow used for the creation of the different datasets analysed. *Left/Blue*: Genome Gene Content. "*Ab initio*" refers to dataset construction where the whole homo/orthogroup prediction was carried out *de novo* on the reduced taxon samplings, while "pruning" refers to the strategy where taxa are deleted from the full Opi homo/orthogroup data matrices which were constructed using default E (similarity) and I (inflation) values (see text for details). See data repository for the illustration of the complete steps of the gene content dataset creation; *Right/Yellow*: Morphology. The character list was assembled from three solid datasets that encompass the morphological disparity of the taxa in this study. Redundant characters were removed in addition to those that are not applicable to any of the terminals and historical ones that have been explicitly refuted in recent studies. The different taxon samplings mirror those of the gene content in addition to one in which the longest branches from the other morphological analyses were excluded.

With the same aim of LBA detection, additional 70 datasets were generated where distant outgroups (i.e., Fungi, Ichthyosporea) and the long-branched in-group (bilaterian) species *Caenorhabditis elegans* (Nematoda), *Pristionchus pacificus* (Nematoda), *and Schistosoma mansoni* (Platyhelminthes) were excluded, and different methods were used to construct the

data matrices. Datasets were assembled using two strategies. First, the "*ab initio*" strategy carried out the whole homo/orthogroup prediction *de novo* on the reduced taxon samplings. Second, the "pruning" strategy pruned taxa from the full Opi homo/orthogroup data matrices, which were constructed using default E (similarity) and I (inflation) values (Fig. 1, see Methods for details). The *ab initio* vs. pruning dataset constructions aimed to assess the effect of those two approaches on the dimensions (gene family number) of the resulting datasets and the topology of phylogenies estimated from them.

Topologies from the individual analyses were inspected manually (see Methods, Supp. Tables 2, 3 and Supp. Fig. 1). Additionally, Total Posterior Consensus Trees (TPCT; Supp. Data 4) were calculated for different datasets that summarise all trees sampled (after convergence) from all analyses with the exact same taxon sampling in a single majority rule consensus tree, therefore reflecting an averaging over all different E- and I-values used to reconstruct the different datasets. These trees are referred to as TPCT Opi (Fig. 2, Genome gene content), TPCT Opi-homo and Opi-ortho (Supp. Fig. 2, Fig. 5 A-B), TPCT Aco-homo and Aco-ortho (Supp. Fig. 3, Fig 5. C-D), and TPCT Xen-homo and Xen-ortho (Supp. Fig. 4, Fig. 5 E-F). Support for different hypotheses was then examined using statistical hypothesis testing [48,49] (see Supp. Fig. 12, 13).

Genome gene content supports Porifera as the sister group of the other animals.

Our new datasets provided the opportunity to investigate the most likely sister group of all other animals. In all 190 analyses, sponges emerged as a monophyletic group. The TPCT Opi (Fig. 2, genome gene content) indicates that the support across all analyses with a full taxon sampling is high with a Posterior Probability (PP) of 0.99 for the clade uniting all animals but the sponges, consistent with Porifera representing the sister group of the rest of the animals. Overwhelmingly strong statistical support was found for this analysis in our hypothesis testing (see Supp. Figs. 12, 13; Supp. Table 5).

Ctenophora invariably emerged as the sister group of all the animals except sponges in the TPCTs; however, the support for this node is variable in TPCTs derived from homogroups and orthogroups (PP=0.55-0.99; Supp. Figs. 2–4). The variable level of support indicates that some analyses found Ctenophora to be placed more crownward in the tree. Three alternative topologies were found for the placement of the Ctenophora when Porifera branched first (Supp. Fig. 1C, 2–5): Placozoa branches off before Ctenophora, the relationships between Ctenophora and Placozoa are not resolved, or Placozoa emerges as the sister group of Ctenophora. These three arrangements appear in very low numbers of trees, mostly derived from homogroup-based datasets (see Supp. Table 3 for details). In some cases, Placozoa emerges as the sister group of all animals (Supp. Table 3). Finally, Cnidaria appears as the sister group of the Bilateria in all analyses (PP=0.99).

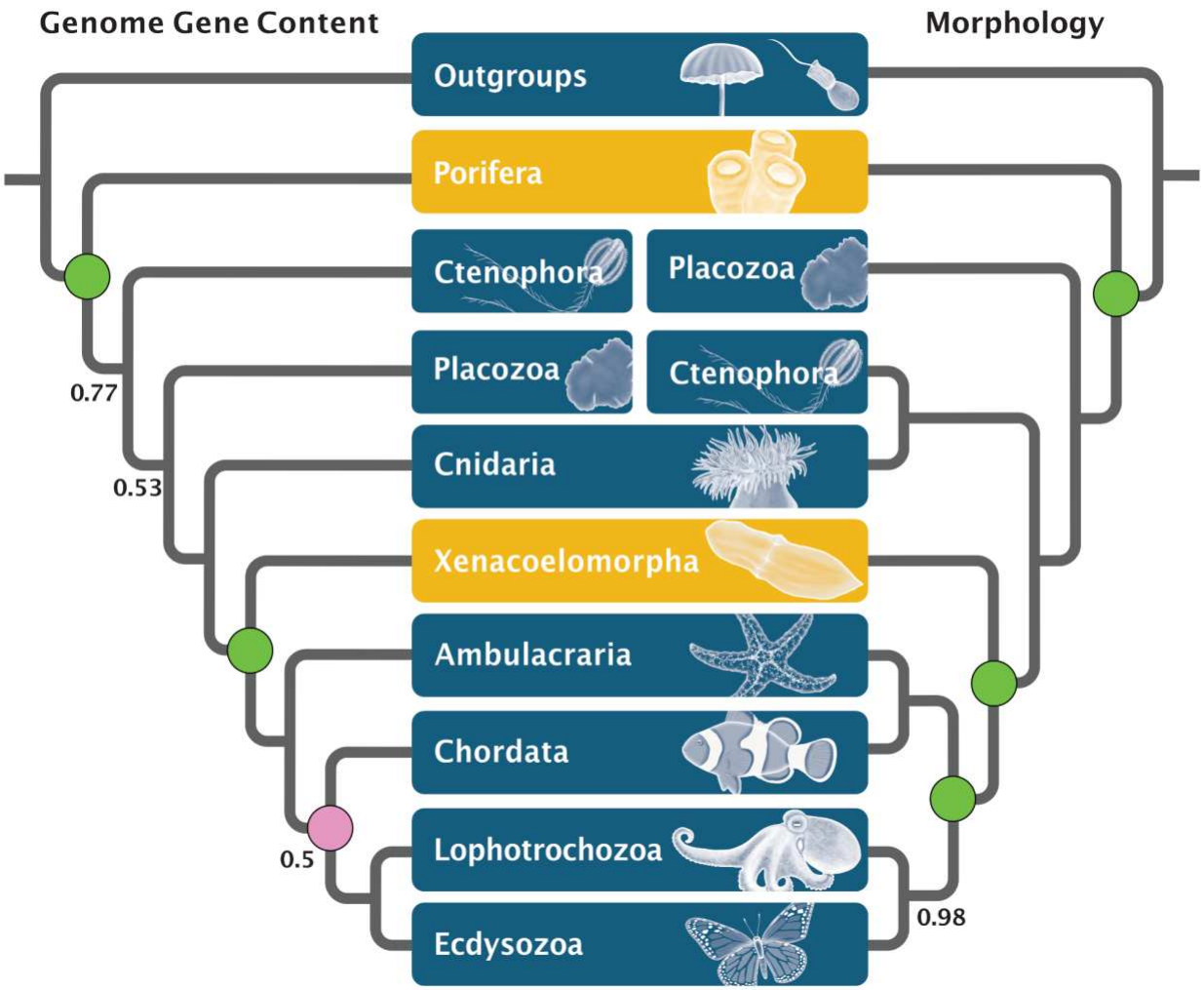


Figure 2: **Reconstruction of animal phylogeny with 47 species (Opi taxon sampling) based on gene content datasets (TPCT) and morphological data.** *Left*: Total consensus tree of >10.5 million individual tree samples from analyses using datasets of homogroups and orthogroups of all the different E- and I-values for genome gene content (for details see Materials and Methods, see Supp. Data 1 for details of analytical settings). *Right*: morphology-based phylogeny based on the non-additive coding scheme.

Note the different positions of Ctenophora. Second to branch off in gene content and sister group to Cnidaria in morphology (i.e., Coelenterata) analyses. The monophyly of Deuterostomia is strongly supported by morphology but around 50% by gene content datasets.

Posterior probabilities lower than 0.99 are indicated on both phylogenies.

Statistical hypothesis tests of focal nodes: *Green circle* = node is strongly supported in the majority of tests conducted; *Purple circle* = node is not strongly supported in the majority of tests conducted (see Supp. Figs 9, 12, 13 for details).

Genome gene content supports Xenacoelomorpha as the sister group of the other Bilateria

The 47-genomes Opi dataset included five Xenacoelomorpha species and the full outgroup taxon sampling (Fig. 1, see Methods). With these datasets, Xenacoelomorpha was recovered as the highly-supported sister group of the rest of the Bilateria (Fig. 2, Genome gene content), consistent with the Nephrozoa hypothesis, irrespective of whether homogroups or orthogroups were used, and with different inflation values and different outgroup sampling. Statistical hypothesis tests provided very strong support for the Nephrozoa hypothesis in 96% of the Opi, Aco and Xen datasets (Supp. Figs. 12, 13). Similarly, datasets in the Aco group (those in which *Xenoturbella* was excluded) placed Acoelomorpha as the sister group of the rest of the Bilateria (both based on homogroups and orthogroups, Supp. Fig. 3). The overwhelming majority of the 41-genome datasets in the Xen group (those where Acoelomorpha were excluded) also resolved *X. bocki* as the sister group of the rest of the Bilateria (Supp. Fig. 3, 5, Supp. Table 3). Finally, in the TPCT Opi-ortho, deuterostome paraphyly is supported but with low posterior probability (PP=0.77). Statistical hypothesis test support for deuterostome monophyly is strong from most Opi, Aco and Xen homogroup datasets, but not so from orthogroup datasets (see Supp. Figs. 12, 13).

Parameter changes affect mainly the final topologies from homogroup-based datasets.

Different Similarity (E) and Inflation (I) values were used to construct the gene content datasets and evaluate their influence on dataset construction and downstream phylogeny estimation. Parameter changes resulted in final homo- and orthogroup matrices with different numbers of characters, but always in the range of 20,000 to 80,000 genes (Supp. Table 2, 3). The choice of E-values did not significantly affect matrix reconstruction, but by contrast, the choice of I-values and whether homo- or orthogroups were used when defining matrices had significant but predictable effects.

It was expected that Orthogroup-based datasets contain a larger number of characters than the corresponding homogroup-based datasets (Supp. Fig. 1 A, B), because homogroups include multiple orthogroups. Furthermore, higher inflation values resulted in the identification of a higher number of smaller homo- and orthogroups, which translated into matrices with more characters. In datasets Opi, Aco, and Xen, the lower I-values resulted in phylogenies favouring the Porifera-sister hypothesis, Xenacoelomorpha as the sister group of the Nephrozoa, and monophyletic Deuterostomia; this trend is stronger for the orthology-based datasets (see Supp. Fig. 1C).

Phylogenies based on homogroups exhibit more variability in the resulting tree topologies than phylogenies based on orthogroups. However, while the overwhelming majority of homogroup-based trees were consistent with the Porifera-sister hypothesis, 11.1% of all those trees showed Placozoa as the sister group of all the other animals. From all homogroup-based analyses that showed Porifera-sister, less than 25% of datasets constructed using high I-values placed *X. bocki* within Deuterostomia (see Supp. Fig. 1C and Supp. Table 3). Up to 75% of homogroup-based datasets have consistent support for the Nephrozoa hypothesis, independent of inflation values.

Paraphyletic Deuterostomia appears in around 25% of the trees estimated from data sets constructed with high inflation values (Supp. Fig. 1C), while in the rest of the treatments it appears in less than 25% of the trees. The variability of the phylogenies obtained with high inflation values is also reflected in the statistical hypothesis tests performed, where high granularity of homogroups did not support any of the tested constraints (Supp. Data 5). The prediction of homo- or orthogroups appears to affect the support for deuterostome paraphyly; orthogroups favour it, while homogroup-based datasets do not (Supp. Fig. 1-4).

The Porifera-sister hypothesis is robust to outgroup sampling in both homogroup- and orthogroup-based phylogenies, as indicated by their very strong statistical hypothesis test support (see Supp. Table 5). Similarly, the Nephrozoa hypothesis received very strong support from the reduced outgroup sampling datasets in our statistical hypothesis tests (see Supp. Table 5), and all reduced taxon-sampling phylogenies where Porifera branched first supported monophyletic Deuterostomia (Supp. Fig. 1C).

The different taxon exclusion schemes showed high variations in the number of characters in the final homogroup- and orthogroup-based data matrices (Supp. Fig. 1A). However, only minor topological changes were observed in phylogenies reconstructed with different numbers of characters, compared to the phylogeny displayed in Figure 2 (Genome gene content). *Xenoturbella bocki* was only recovered in an intra-nephrozoan location in three analyses, all were from the orthogroup-based Holozoa datasets (Supp. Table 3).

Morphological data analyses

The morphological data sets constructed here are the first to include state-of-the-art knowledge about shared characters across Xenacoelomorpha. Two different coding schemes, i.e., non-additive and reductive coding (Methods; Fig. 1, Supp. Data 1) were applied to the morphological dataset. In addition to the different coding schemes, four taxon exclusion experiments were performed: a version with a reduced outgroup, where all the non-metazoan outgroups except the choanoflagellates were excluded from the taxon sampling, two matrices with the 41 and 44 taxon samplings (the core 40 taxa plus *Xenoturbella bocki* and the four species of Acoelomorpha, respectively) and a set without the three taxa with the longest morphological branches (dataset name Morphology Long Branches, MLB) in the previous analyses (*Ixodes scapularis* [Arthropoda], *Danio rerio*, *Gallus gallus* [both Chordata]). All ten analyses resulted in similar topologies (see data repository for details). The analysis of the non-additive matrices exhibits heterogeneous branch lengths and high node support across

289 the phylogeny (Fig. 2, Morphology; Supp. Fig. 6). The phylogeny resulting from the datasets
290 applying reductive coding has lower node support, with three polytomies in the ingroup
291 (within echinoderms, chordates and the sponge classes; Supp. Fig. 7).

292 The only notable difference between the results of these analyses are the relationships within
293 Porifera. In all phylogenies, sponges branched off first (Fig. 3 Morphology; Supp. Fig. 7, 9).
294 However, in the reductive-coding datasets, sponges are paraphyletic, with demosponges
295 branching off first and the Homoscleromorpha and Calcarea in a polytomy with the rest of
296 the animals. In both datasets, Placozoans branched off next and are the sister group of the
297 traditional Eumetazoa (PP=1.0 for non-additive coding, and PP=0.89 for reductive coding).
298 Within eumetazoans, ctenophores are the sister group of the Cnidaria (Coelenterata) (PP=1.0
299 for non-additive coding, and PP=0.65 for reductive coding).

300 In our Bayesian analyses, the hypothesis that Xenacoelomorpha is the sister group of the
301 Nephrozoa is fully supported in the non-additive coded dataset (Supp. Fig. 9) and the
302 outgroup-reduced reductive coded dataset (Supp. Fig. 8), but slightly less supported in the
303 complete sample reductive-coded phylogeny (PP=0.9) (Supp. Fig. 7). The internal
304 relationships of Bilateria show monophyletic Nephrozoa, Deuterostomia, Protostomia,
305 Ecdysozoa, and Spiralia in all the coding schemes applied. In order to further corroborate the
306 results of our Bayesian analyses of the morphological data, we also analysed the set with both
307 codings under maximum parsimony using TNT [50]. The resulting phylogenies from both
308 codings are congruent with the other Bayesian ones (Supp. Figs. 10 and 11). The differences
309 between codings mirror the ones seen from the Bayesian analyses. The reductive coding
310 shows paraphyletic Porifera and much lower bootstrap support overall. The only topological
311 difference between the analyses is the support for a clade of ctenophores and cnidarians in the
312 reductive coding. Instead of being the sister group of ctenophores, cnidarians appear in a
313 polytomy with bilaterians and ctenophores (Supp. Fig. 11).

The statistical hypothesis tests found strong to very strong support for the topology displayed in Fig. 2 (Morphology) for the three different taxon samplings (Opi, Aco and Xen; Supp. Fig. 9). The Nephrozoa hypothesis and the Porifera-sister hypothesis have consistent very strong support. Deuterostome monophyly has strong support in the reductive coding, but very strong support in the non-additive coding (see Supp. Table 5 for the exact values). This statistical support was robust over all different assumed models of morphological character evolution. However, the coding, non-additive vs. reductive, yielded different strengths of support, with the reductive coding producing weak to strong statistical support, whereas the non-additive coding produced very strong support in all scenarios (Supp. Fig. 9). Interestingly, the assumption of a fixed prior distribution over a hyperprior approach for the branch lengths reduced the strength of support in some cases (Supp. Fig. 9). None of the other modelling assumptions had any impact on the estimated strength of support for the different tested hypotheses.

Statistical hypothesis tests support monophyletic Deuterostomia

Although the gene content TPCT displayed in Fig. 2 shows paraphyletic Deuterostomia, this tree topology received only low support (PP=0.5). Statistical hypothesis tests (Supp. Fig. 13, and details above) showed that monophyletic Deuterostomes was consistently and very strongly supported in the majority of datasets analysed, except for orthogroup taxon sampling Opi with inflation values other than the default value of 1.5, and homogroup taxon sampling Opi with higher inflation values of 4 and 6, as well as taxon sampling Xen with an inflation value of 6. The statistical hypothesis tests of the morphological data (Supp. Fig. 9) provided strong to very strong support for monophyletic Deuterostomes.

Discussion

We analysed new genome gene content datasets constructed under various settings and with various taxon samplings, and newly assembled and curated morphological character matrices.

In contrast to primary sequence-based phylogenies, the use of gene content in phylogenetics is a comparably recent development [9,25,38,39] and has been advocated to complement amino acid phylogenomic analyses [14]. This approach relies on the correct estimation of the underlying ortho- and homogroups, which is affected by the tool- and parameter choices [51].

In order to gain an understanding of the effect of different parameter combinations on the prediction of ortho- and homogroups in gene content-based phylogenies, we tested a variety of similarity (E) and inflation (I) values. The differences in the numbers of characters in our datasets, as parameters change, is consistent with the observation that the identification and delimitation of gene families is difficult [40,41]. However, we observed good congruence across datasets over the topology in Fig. 2 (Genome Gene Content), indicating that errors induced by misidentifications of orthogroups were negligible (contra [42]), while homogroup-based topologies were less congruent mostly when high inflation values were used for the predictions.

Potential biases can be induced in the results of gene content analyses when the available genomes are fragmented. While we strived to use high quality genomes only, some were still fragmented, and even recent "chromosome-level" genome assemblies can not guarantee a complete and unfragmented set of the gene content of a species. For example, the genome of *Ephydatia muelleri*, not available at the time we assembled our data set in 2018, is dispersed over 1419 scaffolds, even though about 84% of it was contained in the 24 largest scaffolds, encompassing 22 of the 23 chromosomes [52]. Virtually complete chromosome scale genome assemblies of non-bilaterians are only now starting to appear, i.e., the ctenophore *Hormiphora californensis*, where 99.47% of the genome are contained in 13 scaffolds [53].

While the ascertainment bias correction introduced and used in the gene content analyses of Pisani et al. [9] and Pett et al. [38] accounts for unobserved genes in all species, no correction

363 currently exists to account for unobserved genes in individual species, the type of bias that
364 may be induced by incomplete genomes. However, we used ortholog and homolog
365 identification methods that are standard in the field (see Methods) and those do not rely on
366 complete genes, but assess the given sequence. Nonetheless, developing additional
367 corrections to account for potential errors introduced during *in silico* genome assembly and
368 annotation could be a fruitful avenue for future research.

369 Considerable attention was given to the investigation of putative long-branch attraction
370 artefacts (LBA) that might have caused a placement of Xenacoelomorpha at the root of
371 Bilateria and the sponges at the root of the animals. To achieve this goal we performed taxon
372 exclusion experiments, similar to Pisani et al. [9] and Philippe et al. [11]. Based on our tests,
373 where we do not see taxa changing position as the ingroup and the outgroup are subsampled,
374 we suggest that the placement of Porifera and Xenoacoelomorpha in our trees does not seem
375 to be affected by LBA.

376 Based on multi-gene alignments, several studies showed that the evolutionary model used can
377 affect the inferred topologies [10,22,23,26,27,30]. For the burgeoning field of the phylogenetic
378 analysis of gene content data, model development is still limited. Pett et al. [38] applied both the
379 Dollo model, in which, if applied to gene content data, each gene family may be gained only
380 once on a tree, and a reversible binary substitution model, in which a gene family may be
381 gained more than once on a tree. Both models recovered identical topologies, but the
382 reversible binary substitution model, also used here, was shown to have the best fit for this
383 type of data. In any case, additional and more biologically realistic evolutionary models need
384 to be developed to analyse genome gene content data that may show better fit and adequacy.

385 The independently estimated phylogeny from the morphological dataset is fully consistent
386 with the results from the gene content analyses concerning the placement of Porifera and

Xenacoelomorpha. A notable difference concerns the position of Ctenophora, which appears as the sister group of Cnidaria, forming the classic Coelenterata [54] (Fig. 2, Morphology). Deuterostomes are recovered as monophyletic in the morphology-based phylogeny, different from their paraphyly as recovered in a few gene content analyses.

Our genomic and morphological results agree with each other, with previous genome content analyses [9,38], and with phylogenetic trees of amino acid datasets supporting the Nephrozoa [4,29] and Porifera-sister hypotheses [9,12,20,22,26,27,30,55]. Our results on the other hand are in disagreement with studies that identified Ctenophora as the sister of all the other animals [10,13,21,23,25,56–58], and Xenambulacraria [11,28,30,34,59].

Nonetheless, irrespective of the arrangement of the lineages towards the root of the animal tree, the transition to animal multicellularity from a unicellular last common ancestor was marked by an expansion of a preexisting genetic toolkit to enable multicellularity [60]. The functionalities necessary for this transition, such as cell adhesion, were already present in the closest protist relatives of animals, the Choanoflagellata [19]. Additionally, new protein domains evolved in the Urmetazoan that enabled more complex traits [5,61–63], for example novel signalling pathways, such as tyrosine kinases signal transduction cascades [62] and many components of *Wnt* pathway [61], and transcription factors, such as the common glutamate GABA-like receptors [64].

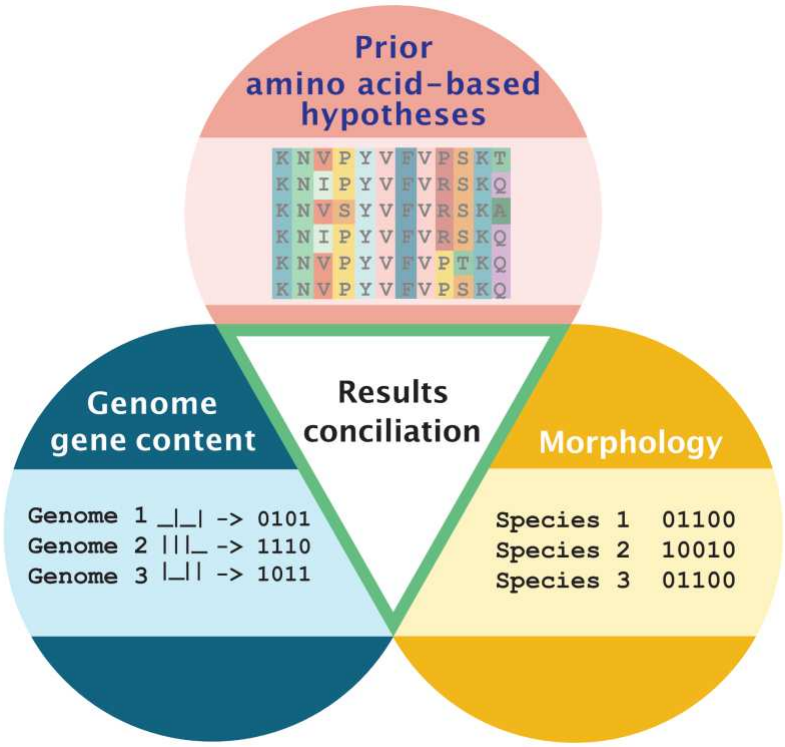
In any case, our conciliated results allow for addressing more conclusively questions about early animal evolution. If we accept that sponges are the sister group of the rest of the animals (Fig. 2), it can not be excluded that the last common animal ancestor (the urmetazoan) may have been a sponge-like organism that fed using choanocyte-type cells [65]. However, the homology of the collar apparatus in the Choanoflagellata, the sister group of animals, with the one of the choanocyte in sponges is currently disputed [66–68]. In spite of that,

411 whatever the true phenotype and metabolic capacities [69] of this urmetazoan were, the key
412 innovations required for animal multicellularity must have happened along the stem lineage
413 towards this urmetazoan. Furthermore, if the Porifera-sister hypothesis is correct, the last
414 common ancestor of animals might have lacked most recognizable metazoan cell types and
415 organ systems, despite having the capacity to transit between different cell states similar to
416 stem cells [67].

417 If we accept that Xenacoelomorpha is the sister group of the rest of the Bilateria (Nephrozoa)
418 and Deuterostomia is monophyletic, the urbilaterian (the last common ancestor of Bilateria)
419 might have been an acoelomate worm [4]. This contrasts scenarios [70] that posit a very complex
420 urbilaterian that could have possessed a coelom, metameric segmentation, and many other
421 bilaterian organ systems. The most notable feature of the urbilaterian would be the lack of
422 any ultrafiltration organs or cell types [4,71]. This lack has been argued to be primary because
423 most xenacoelomorphs are predators and a system for nitrogen excretion is very beneficial
424 for animals with protein-rich diets [6]. Other notable aspects would be the presence of a blind
425 stomach without an anus and their simple gonads which would have been more similar to
426 those of most non-bilaterians. Nevertheless, the high morphological disparity present within
427 extant xenacoelomorphs introduces some uncertainty about the plesiomorphic status of many
428 features. Their nervous systems, for example, are extremely varied [72] and the presence of
429 eyes in their last common ancestor can not be established with confidence [6].

430 Elucidating the origin of bilaterians is also fundamental for our understanding of the early
431 history of our biosphere. The precise sequence of character acquisition is important because it
432 can be correlated with the appearance of more complex body plans and new metazoan
433 ecological guilds such as burrowers and grazers. For example, in the early Cambrian fossil
434 record, it has been postulated that the rising abundance of burrowing bilaterian animals led to

435 the decline of the dominant Precambrian bacterial mats and an initial diversification of
436 ecological interactions – the "agronomic revolution" [73].



437

438 **Figure 3: Illustration of the different data sources used in this study to conciliate results.**
439 In circles are the different data sources. *Top/Red*: These data (amino acid sequence-based multi gene
440 alignments) are not used here but the competing hypotheses about the relationships towards the root
441 of the animal tree of life assessed in this study are derived from previous publications that used this
442 data type. *Left/Blue*: Genome gene content. These data are used here. *Right/Yellow*: Morphological
443 characters. These data are used here.
444 *Middle triangle*: The outcome of independent sources of information allows the conciliation of the
445 results.
446

447 In summary, we independently analysed two lines of evidence, i.e., novel gene content and
448 morphological data matrices, and investigated the robustness of different parameter
449 constellations, including taxon sampling, on the resulting phylogenies. Our results provide
450 further evidence to resolve recalcitrant nodes in the animal phylogeny.

With reference to the root of the animals, where the debate is quite mature, and many contributions from different fields exist [9,10,12,13,17,18,20–23,25–27,30,54–57], our results further strengthen the view that sponges are the sister group of all the other animals. However, resolving the exact relationships of the Ctenophora and Placozoa with respect to the Cnidaria and the Bilateria remains a future challenge.

With reference to the phylogenetic placement of the Xenoacoelomorpha, our analyses favour the Nephrozoa hypothesis. However, the debate on the placement of the Xenoacolemorpha is much less developed [4,11,28–30,34,59], with some key new hypotheses (e.g., the non-monophyly of Deuterostomia) recently emerging [33,34]. Clearly, more studies, using different datasets and methods, as well as the development of more sophisticated evolutionary models for the analysis of gene content data, are necessary to more firmly establish the relationships at the root of the Bilateria.

3. Methods

Data set creation

1. *The general strategy for assembly of the genome gene content datasets*

Publically available proteomes derived from full genome sequences of 47 species were collected in 2018 (Supp. Table 1), representing 17 phyla, to create a balanced taxon sampling across animal phyla, supplementing the taxon sampling of Pett et al. [38]. The collection of proteomes also included non-metazoan outgroups sampled across Opisthokonta (Fungi + Ichthyosporea + Choanoflagellates + Metazoa; Supp. Data 1).

The core taxon set includes 40 species (bold in Supp. Table 1), from which additional taxon samplings were created. The 47-species Opisthokonta (Opi) taxon set contained the full set of species, and is the largest genome gene presence/absence dataset to date. Two additional taxon sets (see Fig. 1; Supp. Data 1, 2) with different taxon samplings of Xenacoelomorpha were assembled adding species to the 40-species core set: a 44-species dataset that had four Acoelomorpha species and no *Xenoturbella bocki* (specified with "Aco" in the dataset name) and a 41 species dataset that had only *X. bocki* and no Acoelomorpha (specified with "Xen" in the dataset name). The rationale behind this taxon-pruning approach was to test for long-branch attraction artefacts in the ingroup (following Philippe et al. [11]) that may impact the relationships of Xenacoelomorpha.

For each taxon sampling strategy two datasets were generated. The first coded the presence/absence of homogroups (i.e., protein families as defined by the output of the Orthofinder-1 pipeline [74]) across taxa. This coding strategy uses the shared presence of a protein family as phylogenetic evidence. The second coded the presence/absence of orthogroups. When this second coding strategy is used, individual orthogroups within each

protein family are treated as individual characters. This is the same strategy introduced and justified by Pett et al. [38]

Homology searches were performed using different parameters of similarity (E-value) in DIAMOND and granulation (Inflation value; I) in the MCL algorithm. Granulation affects the cluster size, i.e., the number of the predicted clusters (orthogroups) that will be considered members of the same homogroup (i.e., of the same protein family). Small I-values indicate coarse-grained clustering resulting in larger clusters (i.e., larger protein families with many paralogs, i.e., orthogroups). Large I-values will lead to fine-grained clustering, chopping bigger clusters into smaller ones, including fewer paralogs (i.e., fewer orthogroups) [75]. Increasing the inflation value (I) therefore leads to homogroup-based datasets with more characters.

For all species in the dataset where only coding sequences (CDS) were available, transdecoder [76] was used to extract the best possible prediction of open reading frames (ORF) and corresponding proteins. All proteomes were analysed using a general approach similar to Pett et al. [38], but with different tools. A homology search of the individual proteomes against each other was conducted with a combination of four different E-values. The search was performed using Diamond v0.9.22.123 [77] for the E-values of 1e-2, 1e-5, 1e-9, and 1e-12. To obtain orthogroups, we used OrthoFinder v2.3.7 [78] with the Diamond option. To establish the homogroup datasets, we used homomcl [38] with a Diamond search. MCL v14-137 [75] was used to cluster the different gene sets with five I parameters: 1.5 (default), 2, 2.5, 4, and 6 [79,80]. Similar to Pett et al. [38], we applied a correction for the ascertainment bias in our phylogenetic model and removed all singletons (i.e., sequences that appear to be present in only one genome) from each presence/absence matrix (gene groups represented by single species). Both homogroup and orthogroup datasets therefore do not contain any single species homo- or orthogroups (singletons), i.e., proteins need to be shared by at least two species and at most

all but two species. The final matrices of homogroup/orthogroup presence/absence for phylogenetic analyses were generated with custom python and BASH scripts. For the dataset naming convention used here, see Supp. Table 4.

All steps of the analysis (dataset construction, phylogenetic analyses) were performed twice to ensure reproducibility, resulting in a total of 380 different datasets analysed.

Datasets to test for long-branch attraction artefacts (LBA)

Using the default E-value of 1e-3 and I-value of 1.5 in OrthoFinder, Diamond, and MCL, we further tested the outcome of different species combinations. The complete taxon sampling of the 47 Opisthokonta (Opi) species and the two subsets Aco and Xeno were used to construct further reduced datasets for two different approaches (see Fig. 2). These are divided into two sub-categories to test for putative long-branch attraction artefacts by either outgroup taxa exclusion or by excluding long-branched ingroup taxa from the taxon sampling.

Taxa exclusion experiments

We tested the effect of reducing taxa in two different ways: first we excluded taxa before running homology searches. When this approach is used, taxa are excluded before the datasets are generated, this is the *ab initio* approach (see Fig. 1 and Supp. Data 1). The second approach, here called “pruning” (see Fig. 1 and Supp. Data 1) simply removed taxa from the datasets. The latter significantly reduces computational time.

1. Outgroup taxon exclusion:

i.) All outgroups but the Choanoflagellates, the sister group of the Metazoa, were successively excluded from the full 47-species Opisthokonta (Opi) taxon set, and a

new OrthoFinder search was conducted to create three different taxon sets, namely ii) Ichthyosporea + Choanoflagellata + Metazoa (= Holozoa; dataset prefix Holo), and iii) Choanoflagellata + Metazoa (= Choanozoa; dataset prefix Cho) [81]; see Supp. Data 1 for more details.

ii.) All outgroups but the Choanoflagellates were pruned from the whole taxon set above. However, the initial character matrix derived from the full Opi dataset was used (no new OrthoFinder search), deleting new singletons and orphans (that resulted from taxon deletion) instead of re-running OrthoFinder; see Supp. Data 1 for more details.

2. Exclusion of long-branched ingroup taxa:

i.) The long-branched species *Caenorhabditis elegans* (Nematoda), *Pristionchus pacificus* (Nematoda), and *Schistosoma mansoni* (Platyhelminthes) were excluded from each of the different taxon sets described above. The complete analysis of ortho- and homogroups estimation was rerun from start to end (*ab initio*). The datasets analysed were Opi-homo/ortho-Ab, Hol-homo/ortho-Ab, and Cho-homo/ortho-Ab, where Ab refers to *ab initio*; see Supp. Data 1 for more details.

ii.) The long-branched species *Caenorhabditis elegans* (Nematoda), *Pristionchus pacificus* (Nematoda), and *Schistosoma mansoni* (Platyhelminthes) were excluded from the final matrix of 47 species together with the outgroups, but without re-running the complete analysis of ortho- and homogroups estimation from start to end, creating three more datasets: Opi-homo/ortho-P, Hol-homo/ortho-P, and Cho-homo/ortho-P, where P refers to *pruning*; see Supp. Data 1 for more details.

555 Overall, 70 datasets were generated combining alternative taxon sampling and character
556 coding (homogroups and orthogroups) strategies. For a full illustrated explanation of the
557 different datasets created, see Fig. 1 (main manuscript) and Figure "All_graph.p.pdf" of the
558 data repository in folder "Additional information".

559 ***Phylogenetic analysis based on genome gene content data matrices***

560 All matrices were analysed with the MPI version of RevBayes v1.0.14 [82,83]. The reversible
561 binary substitution model [84,85] was used for phylogenetic analysis, as it was found to have the
562 best fit to gene content data in Pett et al. [38] (for details see Supp. Data 6). Each run was
563 conducted with four replicated MCMC runs of 50,000 to 80,000 generations to achieve full
564 convergence. Convergence of the four runs was assessed with bpcomp and tracecomp of
565 PhyloBayes v4.1c [86]. An ESS value >300 and bpdiff values <0.3 were used as thresholds to
566 indicate convergence.

567 Majority rule consensus trees were calculated with bpcomp of PhyloBayes v4.1c [86] for each
568 dataset and i) from the individual four MCMC runs of each of the matrices that achieved
569 convergence; ii) from all posterior trees from all converged MCMC runs of homo- and
570 orthogroup datasets, all different E-value (similarity) and inflation value (I) constellations
571 with the same taxon samplings. The resulting phylogeny thus represents the total majority
572 rule consensus tree of all posterior trees / samples from all the different MCMC simulations
573 (TPCT). For a detailed methodological explanation of Total Posterior Consensus Tree
574 (TPCT) see Supp. Data 4. The final trees were visualised with Figtree v1.4.4 [87], all the trees
575 were rooted with the most distant outgroup (Supp. Table 1).

576

577

Phylogenetic analysis based on morphological characters

The taxon sampling of the morphological data matrix was tailored to be identical to the 47-taxon Opi gene content dataset to make the results fully comparable (see data repository). The set of 770 morphological characters is a curated combination of three different previously published datasets: 1) Dataset 1 [46] was used due to its broad eukaryotic sampling, including some fungi and non-metazoan holozoans needed for the coding of the outgroups. 2) Dataset 2 [45] represented the animal backbone as the most comprehensive and exhaustive source of general animal morphological characters. 3) Dataset 3 [47] was added because it included more up-to-date interpretations of some morphological features. Although Dataset 2 [45] is an extensive dataset, it is based on the classical work of Peter Ax from 1996 [44] and, consequently, some well-established changes in the scoring of some characters were needed. For example, characters regarding cuticles and moulting not known at the time of Ax's work to define the Ecdysozoa [88] were coded independently for "nemathelminthes" and arthropods in the original dataset.

The final character list analysed here (Supp. Data 3) was constructed by first combining the character lists of the publications as mentioned above. Then, the combined list was manually checked, and some characters were removed based on four criteria: 1) characters that were redundant (i.e., that reference the same information); 2) characters that only make reference to the specific morphology of clades that were not included in the sample; 3) highly debated characters where the homology was uncertain and has been questioned through independent lines of research, like the homology of "articulatan" (the classical grouping of annelids and arthropods) features [88]; and 4) characters that would have to be coded as unknown for most taxa because we are coding at the species level (i.e., reproductive, developmental and molecular).

In addition to the full 47 taxa set, four taxon sampling experiments were performed by pruning taxa from the full taxon samplings similar to the gene content analyses: two datasets without the two problematic/unresolved echinoderms and a subsample of Xenacoelomorpha (only Xenoturbella and only Acoelomorpha, respectively); a dataset without long branches observed in preliminary morphological analyses (*Danio rerio*, *Gallus gallus*, *Ixodes scapularis*); and lastly a dataset excluding all outgroups except the two choanoflagellates. All morphological data matrices are available in the data repository.

Modelling morphological evolution by using stochastic processes is more intricate than modelling molecular sequence evolution because it cannot be assumed that the same evolutionary process is acting on all characters identically. Stochastic processes for molecular evolution have extensively been studied and extended in the last three decades but stochastic processes for morphological character evolution are only recently catching up. Therefore, we explored several recently developed stochastic processes to test for potential biases in our phylogenetic estimates due to model assumptions. All our stochastic processes are variants of the Markov *k* (Mk) model, where *k* represents the number of states for a character, to model transitions between character states [89,90]. First, we explored the impact of ascertainment bias by either assuming that invariant characters were removed (Mkv model [89,90]) or by assuming that parsimony non-informative characters (i.e., autapomorphies) were removed. We expect that this ascertainment bias primarily influences branch length estimates but not topology estimates [91,92]. Second, we explored whether assuming a fixed exponential prior distribution with a mean of 0.1 expected substitutions per site per branch or a hyperprior distribution on the branch lengths has an impact on the estimated tree topology [93]. Third, we explored whether assuming that all morphological characters are evolving according to the same shared rate or if there is rate variation that can be modelled using four quantiles of gamma distribution [94]. Finally, we explored whether the assumption that all binary characters either

627 share equal rates of transitions or the 0 and the 1 state occur in different frequencies by using
628 a symmetric mixture model with four or five categories [95].
629 We explored all possible combinations of model assumptions (2 ascertainment bias
630 corrections x 2 branch length priors x 2 models of rate variation across characters x 3 models
631 of transition rate variation = 24 models per dataset) for each of the 10 morphological datasets
632 (see Fig. 1; 240 analyses in total). These analyses were run in the Bayesian phylogenetic
633 inference software RevBayes [83] using the MPI version. We used MCMC simulations to
634 approximate the posterior distribution and ran two replicated MCMC simulations per analysis
635 to check for convergence. Each MCMC simulation was run for 250,000 iterations with, on
636 average, 150 moves per iteration. Furthermore, we used the Metropolis-Coupled MCMC
637 extension with one cold and three heated chains to improve convergence.
638 Additionally, to validate our analyses with more commonly applied phylogenetic inference of
639 morphological characters, we performed a single Bayesian analysis in MrBayes [89] and a
640 parsimony analysis per dataset. In the MrBayes analyses, we used a Markov *k* (Mk) model,
641 where *k* represents the number of states for a character, to model transitions between
642 character states [89,90]. Additionally, we assumed that only variable characters (Mkv model)
643 were used and therefore applied the commonly used ascertainment bias correction [89,90]. We
644 ran two replicate MCMC analyses with two million iterations per chain for each dataset. The
645 reductive-coded Opi and Aco sets were run for 10 million because they had not fully
646 converged after the initial two million generations. We checked for convergence using Tracer
647 v1.7.1 [96]. The parsimony analyses were performed on TNT v1.5 using the New Technology
648 search option [50] and 100 bootstrap replicates.

649

650

Hypothesis testing

We used posterior odds [48,49] to test statistical support for three competing hypotheses: (1) the Porifera-sister vs Ctenophora-sister hypotheses, (2) Nephrozoa vs Xenambulacraria hypotheses, and (3) Deuterostome monophyly vs Deuterostome paraphyly. Specifically, we computed the statistical support in favour of the null model M_0 over the alternative model M_1. Following standard statistical practice [48], we used the log-posterior odds of larger than 1 as substantial support, larger than 3 as strong support, and larger than 5 as very strong support. For a detailed explanation of the statistical hypothesis tests carried out see Supp. Data 5.

Code availability

All data and code necessary to reproduce results are available in a public repository https://github.com/PalMuc/triangulation.

Acknowledgements

GW, KJ, and DP acknowledge funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 764840 (ITN IGNITE). GW and LP acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG) through Project FLAGSHIP to GW (Wo896/20-1), and SH through a DFG Emmy-Noether Research Group (HO6201/1-1). GW and SH acknowledge funding through the Ludwig-Maximilians-Universität Munich (LMU) Munich's Institutional Strategy LMUexcellent within the framework of the German Excellence Initiative. We also acknowledge Julie Johnson (Life Science Studios) for assistance with Figs. 1–3 illustrations.

Author contributions

K.J.: co-designed the study, assembled the gene content data sets and carried out the gene content analyses, drafted and revised the manuscript; **L.P.:** assembled the morphological

676 dataset and carried out the morphological analysis, revised the manuscript; **S.H.:** co-
677 supervised the study and developed code for data analyses, revised the manuscript; **D.P.:** co-
678 supervised the study, revised the manuscript; **G.W.:** conceived, designed and supervised the
679 study, acquired the funding and provided the infrastructure, drafted and revised the
680 manuscript.
681
682 **Data availability**
683 All data are available in a public repository https://github.com/PalMuc/triangulation.
684
685 **Competing interests**
686 The authors do not declare competing interests.
687
688 **References:**
689

690 1. Rokas, A., Williams, B. L., King, N. & Carroll, S. B. Genome-scale approaches to resolving
691 incongruence in molecular phylogenies. *Nature* **425**, 798–804 (2003).
692 2. Gaucher, E. A., Kratzer, J. T. & Randall, R. N. Deep phylogeny--how a tree can help
693 characterize early life on Earth. *Cold Spring Harb. Perspect. Biol.* **2**, a002238 (2010).
694 3. Schierwater, B. *et al.* Never Ending Analysis of a Century Old Evolutionary Debate: 'Unringing'
695 the Urmetazoon Bell. *Front. Ecol. Evol.* **4**, 5 (2016).
696 4. Cannon, J. T. *et al.* Xenacoelomorpha is the sister group to Nephrozoa. *Nature* **530**, 89–93
697 (2016).
698 5. Marlow, H. & Arendt, D. Evolution: ctenophore genomes and the origin of neurons. *Curr. Biol.*
699 **24**, R757–61 (2014).
700 6. Haszprunar, G. Review of data for a morphological look on Xenacoelomorpha (Bilateria incertae
701 sedis). *Org. Divers. Evol.* **16**, 363–389 (2016).
702 7. Philippe, H. *et al.* Resolving difficult phylogenetic questions: Why more sequences are not

enough. *PLoS Biol.* **9**, e1000602 (2011).

8. Tihelka, E. *et al.* The evolution of insect biodiversity. *Curr. Biol.* **31**, R1299–R1311 (2021).
9. Pisani, D. *et al.* Genomic data do not support comb jellies as the sister group to all other animals. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15402–15407 (2015).
10. Whelan, N. V., Kocot, K. M., Moroz, L. L. & Halanych, K. M. Error, signal, and the placement of Ctenophora sister to all other animals. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 5773–5778 (2015).
11. Philippe, H. *et al.* Mitigating Anticipated Effects of Systematic Errors Supports Sister-Group Relationship between Xenacoelomorpha and Ambulacraria. *Curr. Biol.* **29**, 1818–1826.e6 (2019).
12. Pick, K. S. *et al.* Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. *Mol. Biol. Evol.* **27**, 1983–1987 (2010).
13. Dunn, C. W. *et al.* Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452**, 745–749 (2008).
14. Dunn, C. W., Giribet, G., Edgecombe, G. D. & Hejnol, A. Animal Phylogeny and Its Evolutionary Implications. *Annu. Rev. Ecol. Evol. Syst.* **45**, 371–395 (2014).
15. King, N. *et al.* The genome of the choanoflagellate Monosiga brevicollis and the origin of metazoans. *Nature* **451**, 783–788 (2008).
16. Jékely, G. & Budd, G. E. Animal Phylogeny: Resolving the Slugfest of Ctenophores, Sponges and Acoels? *Curr. Biol.* **31**, R202–R204 (2021).
17. Dohrmann, M. & Wörheide, G. Novel scenarios of early animal evolution--is it time to rewrite textbooks? *Integr. Comp. Biol.* **53**, 503–511 (2013).
18. Telford, M. J., Moroz, L. L. & Halanych, K. M. Evolution: A sisterly dispute. *Nature* **529**, 286–287 (2016).
19. Ros-Rocher, N., Pérez-Posada, A., Leger, M. M. & Ruiz-Trillo, I. The origin of animals: an ancestral reconstruction of the unicellular-to-multicellular transition. *Open Biol.* **11**, 200359 (2021).
20. Philippe, H. *et al.* Phylogenomics Revives Traditional Views on Deep Animal Relationships. *Curr. Biol.* **19**, 706–712 (2009).

731 21. Whelan, N. V. *et al.* Ctenophore relationships and their placement as the sister group to all other
732 animals. *Nat Ecol Evol* **1**, 1737–1746 (2017).
733 22. Redmond, A. K. & McLysaght, A. Evidence for sponges as sister to all other animals from
734 partitioned phylogenomics with mixture models and recoding. *Nat. Commun.* **12**, 1–14 (2021).
735 23. Li, Y., Shen, X.-X., Evans, B., Dunn, C. W. & Rokas, A. Rooting the animal tree of life. *Mol.*
736 *Biol. Evol.* **38**, 4322–4333 (2021).
737 24. Shen, X.-X., Hittinger, C. T. & Rokas, A. Contentious relationships in phylogenomic studies can
738 be driven by a handful of genes. *Nat Ecol Evol* **1**, 126 (2017).
739 25. Ryan, J. F. *et al.* The genome of the ctenophore Mnemiopsis leidyi and its implications for cell
740 type evolution. *Science* **342**, 1242592 (2013).
741 26. Feuda, R. *et al.* Improved Modeling of Compositional Heterogeneity Supports Sponges as Sister
742 to All Other Animals. *Curr. Biol.* **27**, 3864–3870.e4 (2017).
743 27. Simion, P. *et al.* A Large and Consistent Phylogenomic Dataset Supports Sponges as the Sister
744 Group to All Other Animals. *Curr. Biol.* **27**, 958–967 (2017).
745 28. Philippe, H. *et al.* Acoelomorph flatworms are deuterostomes related to Xenoturbella. *Nature*
746 **470**, 255–258 (2011).
747 29. Rouse, G. W., Wilson, N. G., Carvajal, J. I. & Vrijenhoek, R. C. New deep-sea species of
748 Xenoturbella and the position of Xenacoelomorpha. *Nature* **530**, 94–97 (2016).
749 30. Kapli, P. & Telford, M. J. Topology-dependent asymmetry in systematic errors affects
750 phylogenetic placement of Ctenophora and Xenacoelomorpha. *Sci Adv* **6**, (2020).
751 31. Ruggiero, M. A. *et al.* A higher level classification of all living organisms. *PLoS One* **10**,
752 e0119248 (2015).
753 32. Hyman, L. H. *The invertebrates: smaller coelomate groups, Chaetognatha, Hemichordata,*
754 *Pogonophora, Phoronida, Ectoprocta, Brachipoda, Sipunculida, the coelomate Bilateria*. vol. 5
755 (New York: McGraw-Hill Book Company Inc., 1959).
756 33. Marlétaz, F., Peijnenburg, K. T. C. A., Goto, T., Satoh, N. & Rokhsar, D. S. A New Spiralian
757 Phylogeny Places the Enigmatic Arrow Worms among Gnathiferans. *Curr. Biol.* **29**, 312–318.e3
758 (2019).

759 34. Kapli, P. *et al.* Lack of support for Deuterostomia prompts reinterpretation of the first Bilateria.
760 *Sci Adv* **7**, (2021).
761 35. Munafò, M. R. & Davey Smith, G. Robust research needs many lines of evidence. *Nature* **553**,
762 399–401 (2018).
763 36. Rota-Stabelli, O. *et al.* A congruent solution to arthropod phylogeny: phylogenomics,
764 microRNAs and morphology support monophyletic Mandibulata. *Proc. Royal Soc. B* **278**, 298–
765 306 (2011).
766 37. Campbell, L. I. *et al.* MicroRNAs and phylogenomics resolve the relationships of Tardigrada and
767 suggest that velvet worms are the sister group of Arthropoda. *Proceedings of the National*
768 *Academy of Sciences* **108**, 15920–15924 (2011).
769 38. Pett, W. *et al.* The Role of Homology and Orthology in the Phylogenomic Analysis of Metazoan
770 Gene Content. *Mol. Biol. Evol.* **36**, 643–649 (2019).
771 39. Leclère, L. *et al.* The genome of the jellyfish Clytia hemisphaerica and the evolution of the
772 cnidarian life-cycle. *Nat Ecol Evol* **3**, 801–810 (2019).
773 40. Lunter, G. *et al.* Uncertainty in homology inferences: assessing and improving genomic sequence
774 alignment. *Genome Res.* **18**, 298–309 (2008).
775 41. Frech, C. & Chen, N. Genome-wide comparative gene family classification. *PLoS one* **5**, e13409
776 (2010).
777 42. Natsidis, P., Kapli, P., Schiffer, P. H. & Telford, M. J. Systematic errors in orthology inference
778 and their effects on evolutionary analyses. *iScience* **24**, (2021).
779 43. Felsenstein, J. Cases in which Parsimony or Compatibility Methods will be Positively
780 Misleading. *Syst. Biol.* **27**, 401–410 (1978).
781 44. Ax, P. *Multicellular Animals: A new Approach to the Phylogenetic Order in Nature*. vol. 1
782 (Springer, 1996).
783 45. Deline, B. *et al.* Evolution of metazoan morphological disparity. *Proc. Natl. Acad. Sci. U. S. A.*
784 **115**, E8909–E8918 (2018).
785 46. Goloboff, P. A. *et al.* Phylogenetic analysis of 73 060 taxa corroborates major eukaryotic groups.
786 *Cladistics* **25**, 211–230 (2009).

787 47. Peterson, K. J. & Eernisse, D. J. Animal phylogeny and the ancestry of bilaterians: inferences
788 from morphology and 18S rDNA gene sequences. *Evol. Dev.* **3**, 170–205 (2001).
789 48. Kass, R. E. & Raftery, A. E. Bayes Factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).
790 49. Bergsten, J., Nilsson, A. N. & Ronquist, F. Bayesian tests of topology hypotheses with an
791 example from diving beetles. *Syst. Biol.* **62**, 660–673 (2013).
792 50. Goloboff, P. A., Farris, J. S. & Nixon, K. C. TNT, a free program for phylogenetic analysis.
793 *Cladistics* **24**, 774–786 (2008).
794 51. Remm, M., Storm, C. E. & Sonnhammer, E. L. Automatic clustering of orthologs and in-
795 paralogs from pairwise species comparisons. *J. Mol. Biol.* **314**, 1041–1052 (2001).
796 52. Kenny, N. J. *et al.* Tracing animal genomic evolution with the chromosomal-level assembly of
797 the freshwater sponge Ephydatia muelleri. *Nat. Commun.* **11**, 1–11 (2020).
798 53. Schultz, D. T. *et al.* A chromosome-scale genome assembly and karyotype of the ctenophore
799 Hormiphora californensis. *G3* (2021) doi:10.1093/g3journal/jkab302.
800 54. Zhao, Y. *et al.* Cambrian Sessile, Suspension Feeding Stem-Group Ctenophores and Evolution of
801 the Comb Jelly Body Plan. *Curr. Biol.* **29**, 1112–1125.e2 (2019).
802 55. Nosenko, T. *et al.* Deep metazoan phylogeny: When different genes tell different stories. *Mol.*
803 *Phylogenet. Evol.* **67**, 223–233 (2013).
804 56. Hejnol, A. *et al.* Assessing the root of bilaterian animals with scalable phylogenomic methods.
805 *Proceedings Of The Royal Society B-Biological Sciences* **276**, 4261–4270 (2009).
806 57. Moroz, L. L. *et al.* The ctenophore genome and the evolutionary origins of neural systems.
807 *Nature* **510**, 109–114 (2014).
808 58. Chang, E. S. *et al.* Genomic insights into the evolutionary origin of Myxozoa within Cnidaria.
809 *Proc. Natl. Acad. Sci. U. S. A.* **112**, 14912–14917 (2015).
810 59. Bourlat, S. J. *et al.* Deuterostome phylogeny reveals monophyletic chordates and the new
811 phylum Xenoturbellida. *Nature* **444**, 85–88 (2006).
812 60. Sebé-Pedrós, A. & de Mendoza, A. Transcription Factors and the Origin of Animal
813 Multicellularity. in *Evolutionary Transitions to Multicellular Life: Principles and mechanisms*
814 (eds. Ruiz-Trillo, I. & Nedelcu, A. M.) 379–394 (Springer Netherlands, 2015).

815 61. Nichols, S. A., Dirks, W., Pearse, J. S. & King, N. Early evolution of animal cell signaling and
816 adhesion genes. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12451–12456 (2006).
817 62. Radha, V., Nambirajan, S. & Swarup, G. Association of Lyn tyrosine kinase with the nuclear
818 matrix and cell-cycle-dependent changes in matrix-associated tyrosine kinase activity. *Eur. J.*
819 *Biochem.* **236**, 352–359 (1996).
820 63. Adamska, M. *et al.* Wnt and TGF-beta expression in the sponge Amphimedon queenslandica and
821 the origin of metazoan embryonic patterning. *PLoS One* **2**, e1031 (2007).
822 64. Müller, W. E. G. Review: How was the metazoan threshold crossed? The hypothetical
823 Urmetazoa. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **129**, 433–460 (2001).
824 65. Nielsen, C. Six major steps in animal evolution: are we derived sponge larvae? *Evol. Dev.* **10**,
825 241–257 (2008).
826 66. Mah, J. L., Christensen-Dalsgaard, K. K. & Leys, S. P. Choanoflagellate and choanocyte collar-
827 flagellar systems and the assumption of homology. *Evol. Dev.* **16**, 25–37 (2013).
828 67. Sogabe, S. *et al.* Pluripotency and the origin of animal multicellularity. *Nature* **510**, 519–522
829 (2019).
830 68. Pozdnyakov, I. R. & Karpov, S. A. Flagellar apparatus structure of choanocyte in Sycon sp. and
831 its significance for phylogeny of Porifera. *Zoomorphology* **132**, 351–357 (2013).
832 69. Mills, D. B. *et al.* The last common ancestor of animals lacked the HIF pathway and respired in
833 low-oxygen environments. *Elife* **7**, e31176 (2018).
834 70. Balavoine, G. & Adoutte, A. The Segmented Urbilateria: A Testable Scenario1. *Integr. Comp.*
835 *Biol.* **43**, 137–147 (2003).
836 71. Perea-Atienza, E. *et al.* The nervous system of Xenacoelomorpha: a genomic perspective. *J. Exp.*
837 *Biol.* **218**, 618–628 (2015).
838 72. Jondelius, U., Raikova, O. I. & Martinez, P. Xenacoelomorpha, a Key Group to Understand
839 Bilaterian Evolution: Morphological and Molecular Perspectives. in *Evolution, Origin of Life,*
840 *Concepts and Methods* (ed. Pontarotti, P.) 287–315 (Springer International Publishing, 2019).
841 73. Seilacher, A. Biomat-related lifestyles in the Precambrian. *Palaios* **14**, 86–93 (1999).
842 74. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative

genomics. *Genome Biol.* **20**, 238 (2019).

75. Enright, A. J., Van Dongen, S. & Ouzounis, C. A. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584 (2002).
76. Haas, B. J. TransDecoder. https://github.com/TransDecoder/TransDecoder/ (2017).
77. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
78. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157 (2015).
79. Ballesteros, J. A. & Sharma, P. P. A Critical Appraisal of the Placement of Xiphosura (Chelicerata) with Account of Known Sources of Phylogenetic Error. *Syst. Biol.* **68**, 896–917 (2019).
80. van Dongen, S. & Abreu-Goodger, C. Using MCL to extract clusters from networks. *Methods Mol. Biol.* **804**, 281–295 (2012).
81. Torruella, G. *et al.* Phylogenetic relationships within the Opisthokonta based on phylogenomic analyses of conserved single-copy protein domains. *Mol. Biol. Evol.* **29**, 531–544 (2012).
82. Höhna, S., Landis, M. J. & Huelsenbeck, J. P. Parallel power posterior analyses for fast computation of marginal likelihoods in phylogenetics. *PeerJ* **9**, e12438 (2021).
83. Höhna, S. *et al.* RevBayes: Bayesian Phylogenetic Inference Using Graphical Models and an Interactive Model-Specification Language. *Systematic Biology* **65**, 726–736 (2016).
84. Felsenstein, J. Phylogenies from restriction sites: A maximum-likelihood approach. *Evolution* **46**, 159–173 (1992).
85. Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
86. Lartillot, N., Lepage, T. & Blanquart, S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* **25**, 2286–2288 (2009).
87. Rambaut, A. FigTree v1. 4. http://tree.bio.ed.ac.uk/ (2012).
88. Schmidt-Rhaesa, A., Bartolomaeus, T., Lemburg, C., Ehlers, U. & Garey, J. R. The position of

the Arthropoda in the phylogenetic system. *J. Morphol.* **238**, 263–285 (1998).

89. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).

90. Lewis, P. O. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* **50**, 913–925 (2001).

91. Matzke, N. J. & Irmis, R. B. Including autapomorphies is important for paleontological tip-dating with clocklike data, but not with non-clock data. *PeerJ* **6**, e4553 (2018).

92. Nylander, J. A. A., Ronquist, F., Huelsenbeck, J. P. & Nieves-Aldrey, J. L. Bayesian phylogenetic analysis of combined data. *Syst. Biol.* **53**, 47–67 (2004).

93. Rannala, B., Zhu, T. & Yang, Z. Tail paradox, partial identifiability, and influential priors in Bayesian branch length inference. *Mol. Biol. Evol.* **29**, 325–335 (2012).

94. Wagner, P. J. Modelling rate distributions using character compatibility: implications for morphological evolution among fossil invertebrates. *Biol. Lett.* **8**, 143–146 (2012).

95. Wright, A. M., Lloyd, G. T. & Hillis, D. M. Modeling Character Change Heterogeneity in Phylogenetic Analyses of Morphology through the Use of Priors. *Syst. Biol.* **65**, 602–611 (2016).

96. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).