Syst. Biol. 71(6):1271–1280, 2022

DOI:10.1093/sysbio/syac045
Advance Access Publication June 29, 2022

A Phylogenomic Backbone for Gastropod Molluscs

JUAN E. URIBE[1,2,*], VANESSA L. GONZÁLEZ[3], IKER IRISARRI[4,5], YASUNORI KANO[6], DAVID G. HERBERT[7], ELLEN E. STRONG[1], AND M.G. HARASEWYCH[1]

[1]Department of Invertebrate Zoology, MRC 163, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012, USA; [2]Department of Biodiversity and Evolutionary Biology, Museo Nacional de Ciencias Naturales (MNCN-CSIC), 28006 Madrid, Spain; [3]Global Genome Initiative, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, USA; [4]Department of Applied Bioinformatics, Institute for Microbiology and Genetics, University of Göttingen, and Campus Institute Data Science (CIDAS), Göttingen, Germany; [5]Leibniz Institute for the Analysis of Biodiversity Change (LIB), Zoological Museum Hamburg, 20146 Hamburg, Germany; [6]Department of Marine Ecosystems Dynamics, Atmosphere and Ocean Research Institute, The University of Tokyo, Kashiwa, Chiba, Japan; and [7]Department of Natural Sciences, National Museum Wales, Cardiff, CF10 3NP, UK

**Correspondence to be sent to: Department of Biodiversity and Evolutionary Biology, Museo Nacional de Ciencias Naturales (MNCN-CSIC), 28006 Madrid, Spain;*
E-mail: juanuribearboleda@gmail.com.

Received 29 June 2021; reviews returned 16 June 2022; accepted 24 June 2022
Associate Editor: Jamie Oaks

Abstract.—Gastropods have survived several mass extinctions during their evolutionary history resulting in extraordinary diversity in morphology, ecology, and developmental modes, which complicate the reconstruction of a robust phylogeny. Currently, gastropods are divided into six subclasses: Caenogastropoda, Heterobranchia, Neomphaliones, Neritimorpha, Patellogastropoda, and Vetigastropoda. Phylogenetic relationships among these taxa historically lack consensus, despite numerous efforts using morphological and molecular information. We generated sequence data for transcriptomes derived from 12 taxa belonging to clades with little or no prior representation in previous studies in order to infer the deeper cladogenetic events within Gastropoda and, for the first time, infer the position of the deep-sea Neomphaliones using a phylogenomic approach. We explored the impact of missing data, homoplasy, and compositional heterogeneity on the inferred phylogenetic hypotheses. We recovered a highly supported backbone for gastropod relationships that is congruent with morphological and mitogenomic evidence, in which Patellogastropoda, true limpets, are the sister lineage to all other gastropods (Orthogastropoda) which are divided into two main clades 1) Vetigastropoda *s.l.* (including Pleurotomariida + Neomphaliones) and 2) Neritimorpha + (Caenogastropoda + Heterobranchia). As such, our results support the recognition of five subclasses (or infraclasses) in Gastropoda: Patellogastropoda, Vetigastropoda, Neritimorpha, Caenogastropoda, and Heterobranchia. [Compositional heterogeneity; fast-evolving; long-branch attraction; missing data; Mollusca; phylogenetics; systematic error.]

With its origin in the Cambrian (~500 millions of years ago [Ma]; Fryda et al. 2008, figure 10.4) and an estimated ~90,000 living species (Ponder et al. 2019), Gastropoda is the second most diverse metazoan class. Gastropods have survived multiple mass extinctions during their evolutionary history and undergone independent ecological transitions into nearly all nonaerial environments (from marine, including deepest trenches, hydrothermal vents and cold seeps, to freshwater, estuarine, terrestrial, and even arboreal habitats). These successive radiation events have generated an extraordinary diversity in morphology, ecology, and developmental modes (Aktipis et al. 2008; Ponder et al. 2019). Yet, this also complicates the reconstruction of a robust phylogeny.

Earlier classifications of Gastropoda (Thiele et al. 1929–1931) were based on the morphology of their respiratory, nervous, and radular systems and recognized three major groups: Opisthobranchia, Pulmonata, and Prosobranchia (which included Archaeogastropoda, Mesogastropoda, and Stenoglossa). This tripartite subdivision of the class has long represented the basis for understanding gastropod diversity and evolution. Subsequent efforts to refine gastropod classification resulted in the recognition of a number of main clades: Caenogastropoda (Cox 1960), Neritimorpha (Golikov and Starobogatov 1975), Vetigastropoda (Salvini-Plawen 1980), Neomphalina (McLean 1981) (currently as Neomphaliones: Neomphalida + Cocculinida), Patellogastropoda (Lindberg 1986), Cocculiniformia (Haszprunar 1987b) (currently as the order Cocculinida), and Heterobranchia (Haszprunar 1985). These efforts culminated in a comprehensive cladistic analysis based on anatomical, ultrastructural, and developmental characters (Ponder and Lindberg 1997). Two main groups were recognized: Eogastropoda (Patellogastropoda, or true limpets, and their hypothetical coiled ancestors) and Orthogastropoda, which included all other gastropods, with the Neritimorpha + Cocculinida clades as sister to 1) Apogastropoda (Heterobranchia + Caenogastropoda) plus and 2) Vetigastropoda + Neomphalina (Fig. 1a). The morphology-based phylogeny of Sasaki (1998) supported the division of gastropods into Patellogastropoda and remaining gastropods (minus Heterobranchia): Caenogastropoda was sister to a (Neritimorpha, (Neomphalida, (Cocculinida, Vetigastropoda))) clade (Fig. 1b). The current consensus classification of Gastropoda recognizes six subclasses: Caenogastropoda, Heterobranchia, Neritimorpha, Patellogastropoda, Vetigastropoda, and Neomphaliones (including Neomphalida and Cocculinida) (Bouchet et al. 2017).

Many molecular phylogenetic studies have attempted to reconstruct the evolutionary relationships of Gastropoda. Earlier efforts used few molecular markers,

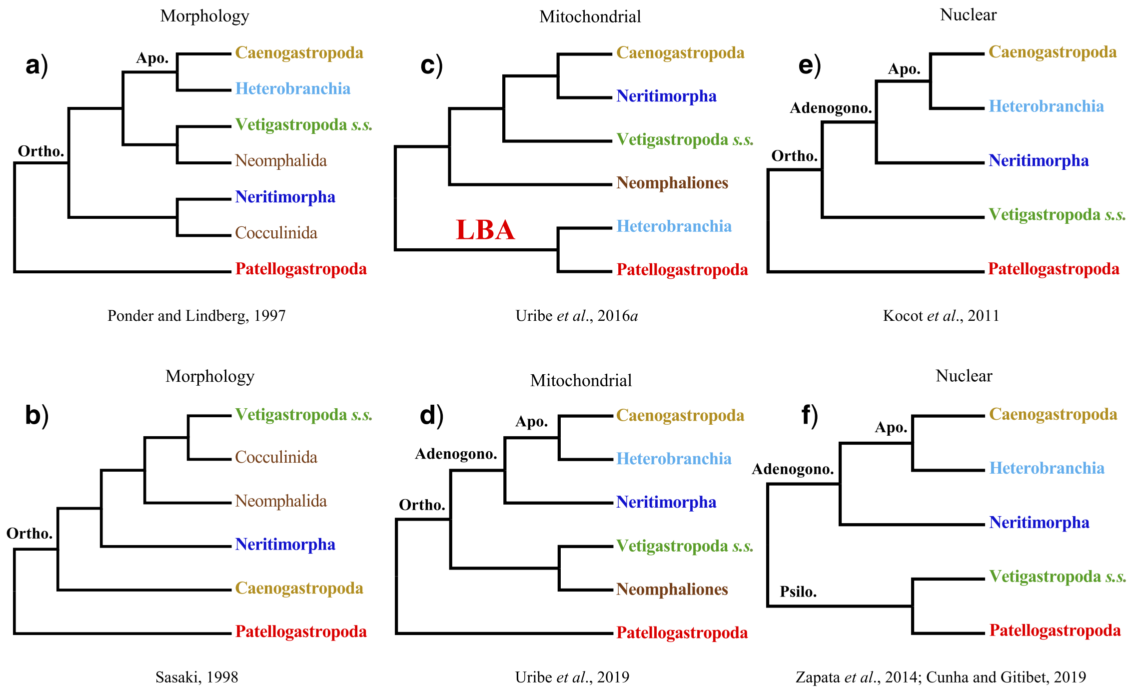

FIGURE 1. Main hypotheses of the internal relationships of Gastropoda. Clade names have been adapted to the current classification. Currently accepted subclasses are given in bold and the higher hierarchical level names are to be completed by adding "gastropoda" after the dot (e.g., Ortho = Orthogastropoda). LBA refers to long-branch attraction.

obtaining diverse hypotheses that were not robustly supported (Winnepenninckx et al. 1998; McArthur and Koop 1999; Colgan et al. 2000, 2003; McArthur and Harasewych 2003). Mitogenomic data provided a second surge of phylogenetic analyses that increased resolution, but generally suffered from long-branch attraction and sequence saturation problems (Uribe et al. 2016; Fig. 1c). However, these limitations typically associated with mitochondrial genomes can be overcome using mixture models and amino acid (AA) recoding, together with a critical selection of slower-evolving taxa (Uribe et al. 2019; Fig. 1d). There have also been attempts to infer robust gastropod phylogenies based on transcriptomic data (Kocot et al. 2011; Fig. 1e; Zapata et al. 2014; Cunha and Giribet 2019; Fig. 1f). Overall, morphological, mitogenomic and nuclear phylogenomic approaches have led to consistent patterns regarding the relative phylogenetic position of some of the main clades, e.g. Apogastropoda (Caenogastropoda + Heterobranchia) (Ponder and Lindberg 1997; McArthur and Harasewych 2003; Castro and Colgan 2010; Kocot et al. 2011; Zapata et al. 2014: Cunha and Giribet 2019; Uribe et al. 2019; Fig. 1a, d, e, and f) and its sister relationship to Neritimorpha in the clade Adenogonogastropoda (Simone 2011) (=Angiogastropoda; Cunha and Giribet 2019) (Zapata et al. 2014; Cunha and Giribet 2019; Uribe et al. 2019, Fig. 1d, e, and f). However, existing phylotranscriptomic analyses conflict with morphological and mitogenomic trees regarding the earliest divergences within gastropods relating to Patellogastropoda, Vetigastropoda and allied taxa. It is worth noting that many of these conflicts can be attributed to the rooting of the gastropod phylogeny. Furthermore, it should be noted that Cocculinida and Neomphalida (subclass Neomphaliones) have yet to be included into a phylotranscriptomic analysis, and thus a phylogenomic analysis encompassing all major gastropod clades is still pending.

Vetigastropoda is a key clade to elucidate both the phylogeny and the evolutionary history of Gastropoda, given 1) its hypothesized age between ca. 485 Ma (Wagner 2002) and ca. 420 Ma (Fryda et al. 2008), 2) its position in some phylotranscriptomic trees contradicting the monophyly of the Orthogastropoda (Zapata et al. 2014; Cunha and Giribet 2019), and 3) the difficulty of defining its members (Bouchet and Rocroi 2005; Bouchet et al. 2017; Ponder et al. 2019). Vetigastropoda is exclusively marine and includes more than 4064 living species (MolluscaBase 2021). Since its original description by Salvini-Plawen (1980), numerous studies have debated both the composition of Vetigastropoda and the morphological synapomorphies that define it (Haszprunar 1987a; Salvini-Plawen and Haszprunar 1987; Ponder and Lindberg 1997; Sasaki 1998; Geiger and Thacker 2005; Bouchet et al. 2017; Ponder et al. 2019). Several works have demonstrated the inadequacy of some of the characters used to define its taxonomic classification (e.g., Geiger and Thacker 2005; Kano 2008; Uribe et al. 2016).

Here, we assembled a new nuclear phylogenomic data set from transcriptomic (RNAseq) data for 33 species, 12 of which were newly sequenced, and reconstructed a robust phylogenetic hypothesis for deep gastropod relationships with emphasis on the problematic Vetigastropoda and allied taxa. To achieve this aim: 1) we included phylogenomic data for the first time for two deep-sea clades (Neomphalida typically from hydrothermal vents and cold seeps, and Cocculinida from biogenic substrates), which are key to understand and circumscribe Vetigastropoda; and 2) we used methods to alleviate systematic biases and assess long-branch attraction artefacts (LBA; Felsenstein 1978). Our study underlines the importance of assessing systematic biases when reconstructing deep and rapidly occurring divergences.

MATERIALS AND METHODS

Taxon Sampling, Transcriptome Sequencing, and Assembly

A total of 33 transcriptomes were analyzed in this study, aiming for a balanced representation of all main gastropod clades, as well as eight outgroups belonging to Bivalvia, Scaphopoda, Cephalopoda, and Polyplacophora classes. The complete list of specimens, including information on origin, references, and vouchers in museum collections, as well as the details and criteria of the taxonomic sampling are provided as Supplementary Methods *a* and Table S1 (available on Dryad at http://dx.doi.org/10.5061/dryad.zcrjdfng6). After collection, all biological material was fixed in RNA*later* and stored at −20°C. Wet lab procedures (mRNA extraction, library preparation, transcriptome sequencing), read quality check and trimming, as well as transcriptome assembly are detailed in Supplementary Methods *a*.

Data Set Construction and Phylogenetic Analyses

Nuclear ribosomal RNA operons and mitochondrial genomes were retrieved and discarded from each assembled transcriptome. For nuclear transcripts, we predicted open reading frames and selected the longest isoform per gene. The inference of orthologous groups was performed with OMA (Orthologous Matrix; Altenhoff et al. 2019). To remove potential paralogs and problematic residues, individual genes were processed in five consecutive steps: 1) masking residues without evidence of positional homology using PREQUAL (Whelan et al. 2018); 2) multiple sequence alignment using MAFFT v. 7 (Katoh and Standley 2013); 3) filtering out high-entropy alignment blocks using Block Mapping and Gathering with Entropy (Criscuolo and Gribaldo 2010); 4) inference of individual gene trees using maximum-likelihood (ML) as implemented in IQ-TREE v.2 (Minh et al. 2020); 5) followed by removal of outlier long branches (potential paralogs) using TreeShrink (Mai and Mirarab 2018) (for additional details, see Supplementary Methods b). Resulting loci were concatenated into two final matrices: *Matrix*-1 (with at least 78.8% taxon occupancy or proportion of taxa per orthogroup), and *Matrix*-2 (with at least 88% taxon occupancy).

Concatenated matrices were analyzed by: ML with empirical mixture models using IQ-TREE v.2, 9- and 12-state AA recoding (as supplementary analyses), and retaining constant sites; Bayesian inference (BI) using PhyloBayes MPI v.1.5a (Lartillot et al. 2013) with infinite mixture model (CAT-GTR+Γ), and either with 6-state AA recoded or unrecoded or raw AA (as supplementary analyses), and removing constant sites. We tested the effect of gradually removing faster-evolving sites in the larger *Matrix*-1. Site rates were inferred using TIGER (Cummins and McInerney 2011) with the following command “-tiger -in *Matrix*-1.fas -b 20 -f s (sorted).” Additionally, we reconstructed a summary multispecies coalescent tree using ASTRAL-II (Mirarab and Warnow 2015). See details for matrix construction and phylogenetic inference in Supplementary Information (Supplementary Methods *b* and Table S2 available on Dryad).

To compare differences in the phylogenetic usefulness of the data sets, we calculated seven properties related to potential phylogenetic informativeness and bias using the package *genesortR* (Mongiardino 2021). Specifically, we measured four potential biases (average pairwise patristic distance, compositional heterogeneity, saturation, and root-to-tip variance) and three proxies for phylogenetic signal (Robinson-Foulds similarity to a target topology, average bootstrap support, proportion of variable sites). We also quantified the per-site and per-gene support for two competing hypotheses using the method of Shen et al. (2017).

RESULTS

A total of 32 phylogenomic trees were reconstructed by ML, BI, and coalescent analysis for the two data sets, as well as from shorter data sets after stepwise removal of faster-evolving sites in the larger data set *Matrix*-1. Data set features, best-fit evolutionary models, composition homogeneity chi-square tests, and all phylogenetic trees are provided in Figs. 2(a, b) and 3(a–f), Supplementary Figs. S1–S4, and S6, and Tables S2 and S3 (available on Dryad). All 32 trees recovered monophyletic Gastropoda and topologies were largely concordant in the shallow divergences, but differed at the deepest nodes. All gastropod subclasses except for Vetigastropoda *s.s.* were recovered as monophyletic. The relative position of Neomphaliones as sister to Pleurotomariida (here represented by *Bayerotrochus midas*) rendered Vetigastropoda *s.s.* paraphyletic in most trees except in the coalescent species tree (Supplementary Fig. S2a–d available on Dryad) and ML trees with 9 and 12-state AA recoding (Supplementary Fig. S3a–d available on Dryad).

Phylogenetic inference under ML or BI based on *Matrix*-1 recovered Patellogastropoda sister to Vetigastropoda *s.l.* (Pleurotomariida and Neomphaliones), and Neritimorpha sister to Apogastropoda (Caenogastropoda and Heterobranchia) (Fig. 2a). ML and BI trees inferred from *Matrix*-2 recovered a deep split between Patellogastropoda and Orthogastropoda. Within Orthogastropoda, Neritimorpha + Apogastropoda was sister to Vetigastropoda *s.l.* (Fig. 2b). Despite the topological

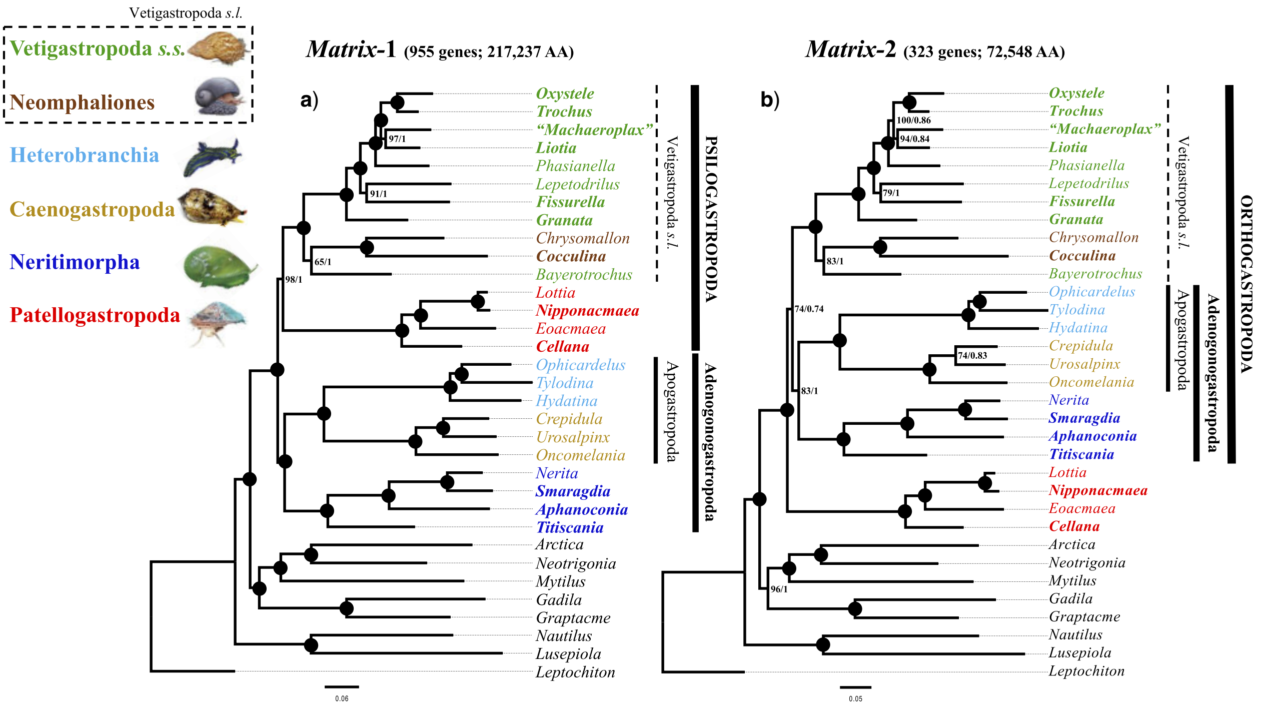

FIGURE 2. Phylogenetic hypotheses for deep gastropod relationships based on phylotranscriptomic datasets: (a) *Matrix*-1 with 955 genes and 217,237 amino acid sites and (b) *Matrix*-2 with 323 genes and 72,548 amino acid sites. Trees were reconstructed using maximum-likelihood (ML; topologies shown) and Bayesian inference (BI) using site-heterogeneous models and six-state recoding. Numbers at nodes are statistical support values for ML (bootstrap proportions)/BI (posterior probabilities). Black dots indicate branches with maximum bootstrap and posterior probability values (100/1, respectively). The BI trees were constructed without taking into account constant sites. Scale bar is in expected substitutions/site. The specimens sequenced in this study are highlighted in bold font.

congruence with ML, we note that BI analyses failed to converge after >20,000 cycles (constant sites were also removed, see Materials and Methods with extension in Supplementary Methods b). Three out of four summary-coalescent trees based on the two matrices converged on the same topology, with Patellogastropoda as a poorly supported sister to (Vetigastropoda *s.l.* (Apogastropoda + Neritimorpha)) (Supplementary Fig. S2a–c available on Dryad). One ASTRAL tree recovered Caenogastropoda as sister to Neritimorpha + (Vetigastropoda *s.l.* + Patellogastropoda) (Supplementary Fig. S2d available on Dryad).

To reduce the effect of AA compositional heterogeneity and saturation, rapidly evolving sites were discarded progressively. Five topologies were inferred (Fig. 3b–f). Removal of 20%, 25%, or 30% of rapidly evolving sites resulted in Patellogastropoda being recovered as the sister lineage to Orthogastropoda (Fig. 3d–f), the same topology as derived from *Matrix*-2 (Fig. 2b). The clade Pleurotomariida + Neomphaliones as sister to the remaining Vetigastropoda was most strongly supported by discarding 30% of rapidly evolving sites (Fig. 3f).

DISCUSSION

We investigated the deep phylogenetic structure of Gastropoda with a balanced sampling of all subclasses, including for the first time Neomphalida and Cocculinida, two key lineages from deep-sea habitats. Preliminary phylogenetic analyses (*Matrix*-1) seemed to confirm previous phylotranscriptomic studies (Zapata et al. 2014; Cunha and Giribet 2019) that contradicted morphological and mitogenomic evidence. However, after carefully accounting for key methodological factors such as matrix completeness, compositional heterogeneity, and proportion of rapidly evolving sites, some of our phylogenetic hypotheses are more congruent with earlier morphological and mitogenomic evidence, as detailed below.

Overcoming Systematic Biases

Effect of matrix occupancy.—Several phylogenomic studies of deep evolutionary relationships have demonstrated that varying data set integrities can result in incongruent tree topologies (Philippe et al. 2011; Yang and Smith 2014; Ballesteros et al. 2019). Other studies argued that a wide diversity of genes, despite low matrix occupancy, may be more accurate for resolving deep relationships (Hejnol et al. 2009; Fernandez et al. 2016; Kulkarni et al. 2021). It was recently argued that requiring high matrix occupancy could lead to the overall loss of phylogenetic signal (Mongiardino 2021). However, matrix occupancy thresholds seem to differently affect each particular phylogenetic problem (Hosner et al. 2016; Streicher et al. 2016).

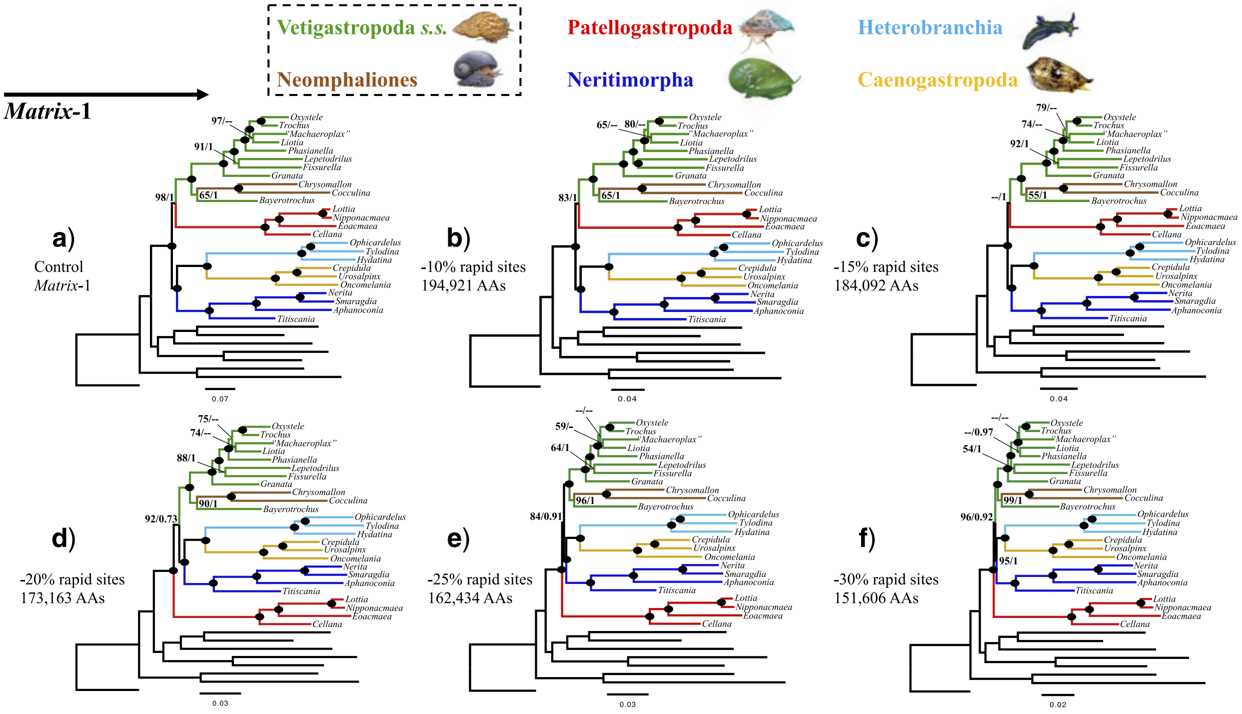

FIGURE 3. Tree topologies (a–f) differ in the progressive deletion of rapidly evolving sites. Trees were reconstructed using maximum-likelihood (ML; topologies shown; All BI trees are shown in Supplementary Fig. S4a–e available on Dryad) and Bayesian inference (BI) using site-heterogeneous models and six-state recoding. Numbers at nodes are statistical support values for ML (bootstrap proportions)/BI (posterior probabilities). Black dots indicate branches with maximum bootstrap and posterior probability values (100/1, respectively). The BI trees were constructed without taking into account the constant sites. Scale bar is in expected substitutions/site.

We assessed the role of matrix occupancy using looser (78.8%; *Matrix*-1) and stricter (with 88%; *Matrix*-2) occupancy thresholds that were chosen experimentally on the basis of the concatenated ML topologies made previously (see Supplementary Fig. S1 available on Dryad). These two matrices were analyzed using complex (site-heterogeneous) models and AA recoding, two methods that have been useful in several studies for mitigating LBA (Rota-Stabelli et al. 2013; Irisarri and Meyer 2016; Simion et al. 2020). The two matrices resulted in incongruent topologies (see Fig. 2a and b; see discussion below). Potential biases were explored using a recent approach implemented in *genesortR* (Mongiardino 2021). Our results suggest that *Matrix*-1 is more prone to systematic bias than *Matrix*-2 due to the higher root-to-tip variance, level of saturation, average patristic distance, and compositional heterogeneity—four sources of systematic bias largely explained by the first principal component of our *genesortR* analyses (Supplementary Fig. S5a and b, respectively, available on Dryad).

Some studies showed that heterogeneous missing data distribution can hamper the reconstruction of complex evolutionary histories (Lemmon et al. 2009; Simmons 2012; Branstetter et al. 2017; Simion et al. 2020) and could even exacerbate LBA (Roure et al. 2013; Philippe et al. 2017). In fact, the heterogeneous or nonrandom distribution of missing data has been generally considered problematic in the phylogenetics literature (Wiens and Morrill 2011; Roure et al. 2013). It is interesting to note that Vetigastropoda and allied taxa show fewer aligned AA positions in the less complete *Matrix*-1 (with an average of 72.0% vs. 79.4% in the remaining clades) and a greater increase in AA completeness in *Matrix*-2 when compared with other clades (10.22% vs. 5.12%; see Fig. 4). This could suggest that the biased distribution of missing data can negatively affect phylogenetic reconstruction, particularly when missing data are disproportionally high in key lineages whose phylogenetic position is disputed, as is the case if Vetigastropoda here. Thus, missing data in *Matrix*-1 compared with *Matrix*-2 likely compromises the recovery of robust relationships along the backbone of the gastropod phylogeny, that is, between Patellogastropoda, Vetigastropoda and Adenogonogastropoda.

Effect of removing rapidly evolving sites.—Under the rationale of eliminating positions that have undergone multiple substitutions and are thus prone to homoplasy and LBA (Felsenstein 1978; Brinkmann and Philippe 1999; Philippe et al. 2000; Brinkmann et al. 2005), several works have implemented methods for the removal of rapidly evolving sites (Pisani 2004; Brinkmann et al. 2005; Rodríguez-Ezpeleta et al. 2007; Simmons and Gatesy 2016; Ballesteros et al. 2019; Simmons et al. 2019; Strassert et al. 2019). This has been shown to reduce

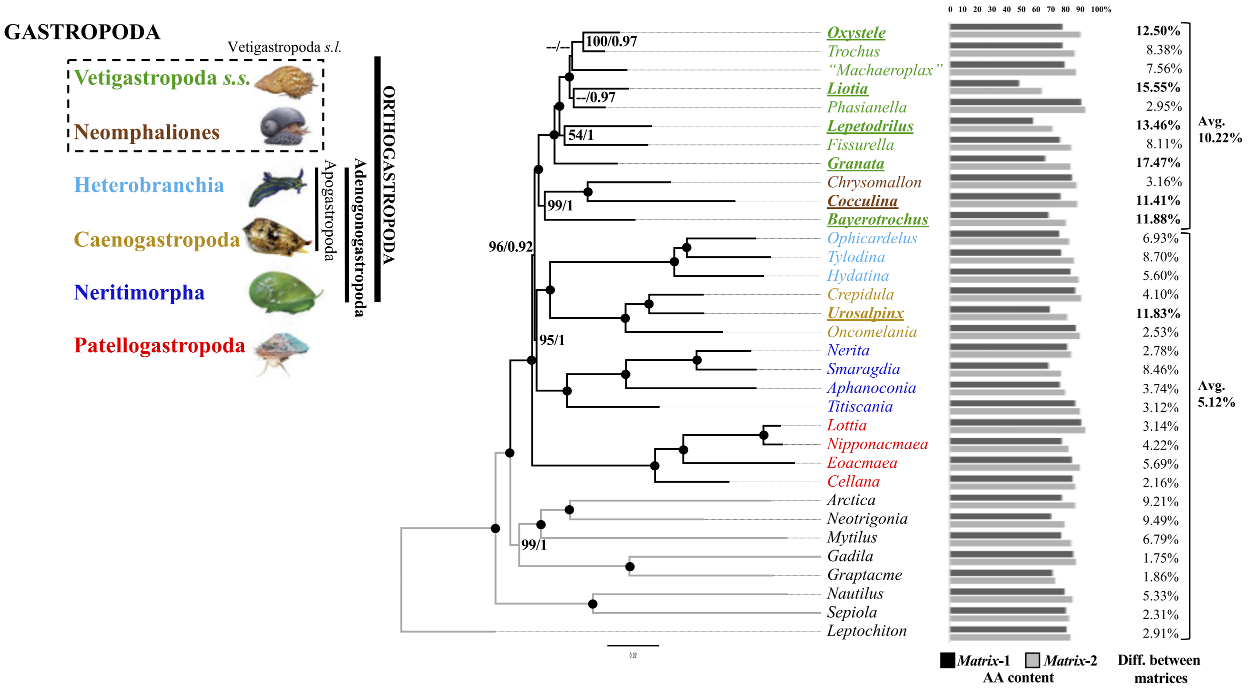

FIGURE 4. Gastropod classification proposed here (left) based on the best tree reached in this study (right, *Matrix*-1 discarding 30% of rapidly evolving sites; Fig. 3f). We recognize five subclasses, namely Patellogastropoda, Vetigastropoda (s.l., including Neomphaliones), Neritimorpha, Caenogastropoda, and Heterobranchia. Numbers at tree nodes denote bootstrap proportions and posterior probabilities; black dots indicate maximum values (100/1). Matrix completeness (amino acid content) in *Matrix*-1 and *Matrix*-2 is shown for each terminal at the right of the tree. Taxa with >10% difference between the two matrices are highlighted with bold letters; distribution of missing data in *Matrix*-1 is biased with a more pronounced increase in Vetigastropoda s.l. (10.22%) compared with the (5.12%).

noise when resolving ancient phylogenetic relationships (Brinkmann and Philippe 1999), as is the case for Gastropoda, a clade of Cambrian age (Ponder et al. 2019). Given the incongruence among our initial trees (Fig. 2) and after further consideration of the results from the more complete data set (*Matrix*-2; 88% taxon occupancy; 72,548 AA positions) we tested the effect of removing rapidly evolving sites from *Matrix*-1 aiming to reduce the effect of saturation and compositional heterogeneity and thereby possible LBA artefacts (Phillips et al. 2004). The progressive deletion of rapidly evolving sites showed that the topology obtained with *Matrix*-1 (217,237 AAs) was strongly influenced by relatively few rapidly evolving sites (Fig. 3) and the initial topology collapsed when 20% of the fastest sites were discarded (173,163 AAs). Support values for key branches were highest when 25% or 30% of the rapidly evolving sites were removed (162,434 or 151,606 AA positions, respectively; Fig. 3e–f). Despite criticisms that TIGER might sometimes infer biased site rates (against sites with symmetric distributions, greater character space, or higher character conflict; Simmons and Gatesy 2016), we observed that the removal of faster evolving sites produced topologies congruent with those from the more complete *Matrix*-2, as well as with previous morphology-based phylogenies (Ponder and Lindberg 1997; Sasaki 1998). This suggests that TIGER can still capture rapidly evolving sites (including many sites likely showing high compositional heterogeneity) and can help counteract phylogenetic artifacts in *Matrix*-1. Interestingly, site-wise log likelihoods of fast-evolving sites showed a similar distribution in the full *Matrix*-1 and the shorter *Matrix*-2, with ∼66% supporting the topology in Fig. 2a and ∼33% supporting that of Fig. 2b (see Supplementary Table S4 available on Dryad). Although the employed method of Shen et al. (2017) does not consider alternative topologies to those being tested, this result suggests that the decisive factor for inferring one topology or another must lie in the combined effect of specific alignment sites.

Independent of the strategy used (matrix occupancy or discarding rapidly evolving sites), six-state AA recoding has been recommended to reduce compositional heterogeneity (e.g., Rota-Stabelli et al. 2013; Laumer et al. 2018; Cunha and Giribet 2019; Philippe et al. 2019). This was corroborated in chi-square tests for composition homogeneity of *Matrix*-1 with six-state recoding versus *Matrix*-1 without recoding (Supplementary Table S3 available on Dryad). However, the loss of phylogenetic information entailed in the six-state recoding might outweigh its benefits and could lead to erroneous topology reconstruction, and thus a comparison recodings with more (e.g., 9 and 12) AA states and the original data (without recoding) is desirable (Hernandez and Ryan 2021). Our tree reconstruction without recoding (for *Matrix*-1, 2, and 1*f*) and with 9- and 12-state recoding (for *Matrix*-1 and 2) corroborates the same deep phylogenetic relationships of gastropods reached using six-state AA recoding (see Supplementary Methods *b* and Figs. S6a–c, S3a–d, respectively, available on Dryad). Yet, the higher statistical support for a putatively incorrect (Patellogastropoda + Vetigastropoda) relationship obtained with six-state recoding (Fig. 2a) compared with the original data (Supplementary Fig. S6a available on Dryad) could suggest that six-state recoding might indeed suffer from loss of phylogenetic signal.

Summary-coalescent analyses.—As an alternative to data concatenation, we analyzed the two data sets using the summary coalescent method ASTRAL-II (Mirarab and Warnow 2015). When gene trees were inferred with RAxML, both data sets recovered the same topology as the concatenated analysis of *Matrix*-1 (Fig. 2a) (Patellogastropoda + Vetigastropoda) but with low bootstrap support (see Supplementary Fig. S2a and c available on Dryad). When gene trees were inferred with IQ-TREE for both data sets, *Matrix*-1 also supported (Patellogastropoda + Vetigastropoda) as the sister clade of the remaining gastropods, however, *Matrix*-2 recovered Caenogastropoda as sister of ((Vetigastropoda + Patellogastropoda) Neritimorpha) (Supplementary Fig. S2d available on Dryad). The observed differences are possibly the result of the irreproducibility of the ML gene trees fed into summary coalescent methods (Shen et al. 2020). The use of summary coalescent analyses for reconstructing ancient relationships has also been criticized on the basis of the accumulation of stochastic error that arises from a limited amount of data in single loci (Telford et al. 2015; Simion et al. 2020).

Deep Phylogeny of Gastropoda and Its Systematic Implications

Patellogastropoda and Orthogastropoda.—Our most reliable phylogenetic hypothesis (Figs. 2b and 3d–f) recovered Patellogastropoda as sister to Orthogastropoda, in agreement with the current classification and morphological synapomorphies (Ponder and Lindberg 1997; Sasaki 1998) as well as one mitogenomic analysis (Uribe et al. 2019) and one nuclear phylogenomic data set (Kocot et al. 2011). Ponder et al. (2019, Chapter 18.2) have further reinforced this division by describing numerous synapomorphies for Orthogastropoda. Other phylogenomic studies have rejected the monophyly of Orthogastropoda (Zapata et al. 2014) and proposed a classification (Figs. 1f and 2a) uniting Vetigastropoda and Patellogastropoda (defined as Psilogastropoda on the basis of plesiomorphic characters; Ponder et al. 2019), as the sister lineage to the remaining gastropods (Zapata et al. 2014, Cunha and Giribet 2019). Interestingly, this topology was also recovered in analyses of *Matrix*-1, but we consider this to be an artifact of missing data and rapidly evolving sites.

Vetigastropoda s.l.: *Neomphaliones and Pleurotomariida*.—The main disagreement in the phylogeny and classification of clades allied to Vetigastropoda concerns understudied lineages from deep-sea habitats, that is, Cocculinida, Neomphalida, Lepetellida (including Lepetodriloidea and Lepetelloidea). The physiological and morphological adaptations developed independently in these clades have often led to cases of convergence being interpreted as synapomorphies. For example, Cocculinoidea and Lepetelloidea, which inhabit mostly biogenic substrates in the deep ocean, were initially classified within the suborder Cocculiniformia (Haszprunar 1988), but are now recognized as members of Neomphaliones and Vetigastropoda, respectively (Bouchet and Rocroi 2005; Bouchet et al. 2017). Recent morphological and molecular phylogenetic analyses have shown that many of these deep-sea clades represent independent colonization of similar habitats (Ponder and Lindberg 1997; Colgan et al. 2003; McArthur and Harasewych 2003; Geiger and Thacker 2005; Kano 2008) and that caution is needed when interpreting certain anatomical characters at higher taxonomic levels.

Since the original definition of Vetigastropoda (Salvini-Plawen 1980), opinions concerning the circumscription of the taxon have differed (e.g., Haszprunar 1987b; Ponder and Lindberg 1997; Sasaki 1998). Bouchet and Rocroi (2005), Bouchet et al. (2017) and more recently Ponder et al. (2019), proposed three classifications that are the most relevant with regard to Vetigastropoda. These three schemes differ mainly in the inclusion (Bouchet and Rocroi 2005; Ponder et al. 2019) or exclusion (Bouchet et al. 2017) of Neomphaliones (Cocculinida and Neomphalida) within Vetigastropoda, as well as the taxonomic ranking of the main clades. Bouchet et al. (2017) favored a conservative concept of Vetigastropoda (*sensu stricto* which includes Lepetellida, Seguenziida, Trochida, Pleurotomariida), and argued that a solid phylogenomic framework including Neomphaliones was still needed. Most analyses in the present study, which included both Cocculinida and Neomphalina for the first time, recovered the close association of non-pleurotomariid Vetigastropoda and Neomphaliones with high statistical support (Figs. 2 and 3; Supplementary Figs. S1–S4, and S6 available on Dryad). Similar results have been reported in several previous studies (Sasaki 1998; Geiger and Thacker 2005; Uribe et al. 2016, 2019). In contrast, a multilocus (three nuclear and two mitochondrial loci) phylogeny recovered non-pleurotomariid Vetigastropoda more closely related to Patellogastropoda than to Pleurotomariida or Neomphaliones (Aktipis and Giribet 2012), whereas another study combining mitogenomic and nuclear-gene data recovered Vetigastropoda *s.s.* closer to Apogastropoda than to Neomphaliones (Lee et al. 2019). However, these two studies based on ribosomal sequences may suffer LBA problems similar to earlier analyses (McArthur and Koop 1999; Colgan et al. 2003; McArthur and Harasewych 2003). Interestingly, Vetigastropoda *s.s.* was recovered with moderate support as the sister clade of Neomphaliones using only the mitogenomic data set (figure 1 in Lee et al. 2019). However, Lee et al. (2019) did not include Patellogastropoda and Heterobranchia, which exhibit very long branches in mitogenomic phylogenies (e.g., Grande et al. 2008).

Pleurotomariida, a putative member of Vetigastropoda (Bouchet and Rocroi 2005; Bouchet et al. 2017), are today represented by specialist sponge-feeders (Harasewych 2002) distributed from the continental shelf to the bathyal slope (100–1000 m; Anseeuw and Goto 1996). There are 36 extant species included in four genera in a single family (MolluscaBase 2021), with a tentative origin in the Early Triassic (Harasewych 2002), although the fossil record of related extinct families goes back to the Late Ordovician or Late Silurian (Ponder et al. 2019). The epipodial sense organ is the most restrictive synapomorphy of Vetigastropoda *s.s.* (Salvini-Plawen and Haszprunar 1987; Sasaki 1998). In 20 out of 28 concatenated trees (except Supplementary Figs. S2—summary-coalescence trees with the *Matrix*-1 and 2—and S3 available on Dryad—ML analysis with 12 and 9-state recoding for *Matrix*-1 and 2), Vetigastropoda *s.s.* was not recovered monophyletic due to the sister-group relationship of Pleurotomariida and Neomphaliones (e.g. Fig. 3f).

Regardless of whether Vetigastropoda *s.s.* is monophyletic or not, our study shows its close affinity to Neomphaliones with high statistical support, thus favoring the concept of Vetigastropoda including Neomphalida and Cocculinida (Ponder et al. 2019). Future approaches that increase taxon sampling around the early divergences of Vetigastropoda and allied taxa, particularly slowly evolving taxa (Brinkmann et al. 2005), will surely contribute to increasing resolution not just within Vetigastropoda *s.l.*, but also among deep gastropod relationships.

Apogastropoda sister to Neritimorpha.—All but one topology (Supplementary Fig. S2d available on Dryad; Summary-coalescence tree using the *Matrix*-2) recovered Neritimorpha as the sister clade of Apogastropoda (Heterobranchia + Caenogastropoda) with high statistical support, in agreement with several previous studies (Ponder and Lindberg 1997; Kocot et al. 2011; Zapata et al. 2014; Cunha and Giribet 2019; Uribe et al. 2019). Together, this grouping was named Adenogonogastropoda (Simone 2011) based on several developmental and anatomical synapomorphies (e.g., acquisition of planktotrophy and multispiral protoconch, eyes with lens, and glandular pallial oviduct).

CONCLUDING REMARKS

In an effort to produce a stable evolutionary framework for the Gastropoda, this study reconstructed phylogenetic relationships among its six subclasses *sensu* Bouchet et al. (2017). The transcriptomic data set was filtered employing two methods that sought to minimize the effect of missing data, homoplasy, and compositional heterogeneity. Our results agree with a recently proposed classification (Ponder et al. 2019) that assigned the deep-sea Neomphalida and Cocculinida to Vetigastropoda (*s*.1.) resulting in five subclasses or infraclasses of Gastropoda (Patellogastropoda, Vetigastropoda, Neritimorpha, Caenogastropoda and Heterobranchia). Matrices with more rapidly evolving sites and heterogeneous missing data distribution have likely caused LBA artifacts and tree instability with regard to the placement of Patellogastropoda and Vetigastropoda. This underlines the importance of implementing methods that strive to alleviate biases when attempting to reconstruct deep evolutionary relationships.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.zcrjdfng6.

ACKNOWLEDGMENTS

All laboratory work, sequencing, and analyses were conducted in and with the support of the Laboratories of Analytical Biology (LAB) facilities of the National Museum of Natural History. Computing was conducted on the Smithsonian High Performance Cluster (https://doi.org/10.25572/SIHPC), Smithsonian Institution at https://researchcomputing.si.edu/high-performance-computing-cluster. Our gratitude goes to Rafael Zardoya, Silvia Perea, and Daniel Pérez-Gelabert for comments on an earlier draft. We thank the Escobar Biological Station (Cazalla de la Sierra, Spain) for support during the COVID-19 confinement. Many thanks also to the Evolutionary Genomics Training on the Caribbean Coast of Colombia (#EGT3C_2019) organized by Andrew Crawford (UNIANDES) and co. for the bases and concepts learned in this course.

REFERENCES

Aktipis S.W., Giribet G. 2012. Testing relationships among the vetigastropod taxa: a molecular approach. J. Mollusc. Stud. 78(1):12–27.

Aktipis S.W., Giribet G., Lindberg D.R., Ponder W.F. 2008. Gastropoda. In: Ponder W.F., Lindberg D.R., editors. Phylogeny and evolution of the Mollusca. Berkeley: University of California Press. p. 201–237.

Altenhoff A.M., Levy J., Zarowiecki M., Tomiczek B., Vesztrocy A.W., Dalquen D.A., Müller S., Telford M.J., Glover N.M., Dylus D., Dessimoz C. 2019. OMA standalone: orthology inference among public and custom genomes and transcriptomes. Genome Res. 29(7):1152–1163.

Anseeuw P., Goto Y. 1996. The living Pleurotomariidae. Osaka: Elle Scientific Publications, 202 pp.

Ballesteros J.A., Santibanez Lopez C.E., Kováč Ľ., Gavish-Regev E., Sharma P.P. 2019. Ordered phylogenomic subsampling enables diagnosis of systematic errors in the placement of the enigmatic arachnid order Palpigradi. Proc. R. Soc. B. 286(1917):20192426.

Bouchet P., Rocroi J.P. 2005. Classification and nomenclator of gastropod families. Malacologia 47:397.

Bouchet P., Rocroi J.P., Hausdorf B., Kaim A., Kano Y., Nützel A., Parkhaev P., Schrödl M., Strong E.E. 2017. Revised classification, nomenclator and typification of gastropod and monoplacophoran families. Malacologia 61(1–2):1–526.

Branstetter M.G., Longino J.T., Ward P.S., Faircloth B.C. 2017. Enriching the ant tree of life: Enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. Methods Ecol. Evol. 8(6):768–776.

Brinkmann H., Philippe H. 1999. Archaea sister group of Bacteria? Indications from tree reconstruction artifacts in ancient phylogenies. Mol. Biol. Evol. 16(6):817–825.

Brinkmann H., Van der Giezen M., Zhou Y., De Raucourt G.P., Philippe H. 2005. An empirical assessment of long-branch attraction artefacts in deep eukaryotic phylogenomics. Syst. Biol. 54(5):743–757.

Castro L.R., Colgan D.J. 2010. The phylogenetic position of Neritimorpha based on the mitochondrial genome of *Nerita melanotragus* (Mollusca: Gastropoda). Mol. Phylogenet. Evol. 57(2): 918–923.

Colgan D.J., Ponder W.F., Beacham E., Macaranas J.M. 2003. Gastropod phylogeny based on six segments from four genes representing coding or non-coding and mitochondrial or nuclear DNA. Molluscan Res. 23(2):123–148.

Colgan D.J., Ponder W.F., Eggler P.E. 2000. Gastropod evolutionary rates and phylogenetic relationships assessed using partial 28S rDNA and histone H3 sequences. Zool. Scr. 29(1):29–63.

Cox L. R. 1960. Thoughts on the classification of the Gastropoda. J. Mollusc. Stud. 33(6):239–261.

Criscuolo A., Gribaldo S. 2010. BMGE (block mapping and gathering with entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol. Biol. 10(1):210.

Cummins C.A., McInerney J.O. 2011. A method for inferring the rate of evolution of homologous characters that can potentially improve phylogenetic inference, resolve deep divergence and correct systematic biases. Syst. Biol. 60(6):833–844.

Cunha T.J., Giribet G. 2019. A congruent topology for deep gastropod relationships. Proc. R. Soc. B. 286(1898):20182776.

Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Biol. 27(4):401–410.

Fernandez R., Edgecombe G.D., Giribet G. 2016. Exploring phylogenetic relationships within Myriapoda and the effects of matrix composition and occupancy on phylogenomic reconstruction. Syst. Biol. 65(5):871–889.

Fryda J.R., Nützel A., Wagner P.J. 2008. Paleozoic Gastropoda, pp. 239–270. In: Ponder W.F., Lindberg D.R., editors. Phylogeny and evolution of the Mollusca, University of California Press, Berkeley, CA.

Geiger D.L., Thacker C.E. 2005. Molecular phylogeny of Vetigastropoda reveals non-monophyletic Scissurellidae, Trochoidea, and Fissurelloidea. Molluscan Res. 25(1):47–55.

Golikov A.N., Starobogatov Y.I. 1975. Systematics of prosobranch gastropods. Malacologia 15:185–232.

Grande C., Templado J., Zardoya R. 2008. Evolution of gastropod mitochondrial genome arrangements. BMC Evol. Biol. 8(1):61.

Harasewych M.G. 2002. Pleurotomarioidean gastropods. Adv. Mar. Biol. 42:237–294.

Haszprunar G. 1985. The Heterobranchia – a new concept of the phylogeny of the higher Gastropoda. J. Zool. Syst. Evol. Res. 23(1):15–37.

Haszprunar G. 1987a. The fine structure of the ctenidial sense organs (bursicles) of Vetigastropoda (Zeugobranchia, Trochoidea) and their functional and phylogenetic significance. J. Mollusc. Stud. 53(1):46–61.

Haszprunar G. 1987b. Anatomy and affinities of cocculinid limpets (Mollusca, Archaeogastropoda). Zool. Scr. 16:305–324.

Haszprunar G. 1988. Comparative anatomy of cocculiniform gastropods and its bearing on archaeogastropod systematics. Malacol. Rev. 4:64–84.

Hejnol A., Obst M., Stamatakis A., Ott M., Rouse G.W., Edgecombe G.D., Martinez P., Baguñà J., Bailly X., Jondelius U., Wiens M., Müller W.E.G., Seaver E., Wheeler W.C., Martindale M.Q., Giribet G., Dunn C.W. 2009. Assessing the root of bilaterian animals with scalable phylogenomic methods. Proc. R. Soc. B. 276(1677):4261–4270.

Hernandez A.M., Ryan J.F. 2021. Six-state amino acid recoding is not an effective strategy to offset compositional heterogeneity and saturation in phylogenetic analyses. Syst. Biol. 70(6): 1200–1212.

Hosner P.A., Faircloth B.C., Glenn T.C., Braun E.L., Kimball R.T. 2016. Avoiding missing data biases in phylogenomic inference: an empirical study in the landfowl (Aves: Galliformes). Mol. Biol. Evol. 33(4):1110–1125.

Irisarri I., Meyer A. 2016. The identification of the closest living relative (s) of tetrapods: phylogenomic lessons for resolving short ancient internodes. Syst. Biol. 65(6):1057–1075.

Kano Y. 2008. Vetigastropod phylogeny and a new concept of Seguenzioidea: independent evolution of copulatory organs in the deep-sea habitats. Zool. Scr. 37(1):1–21.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30(4):772–780.

Kocot K.M., Cannon J.T., Todt C., Citarella M.R., Kohn A.B., Meyer A., Santos S.R., Schander C., Moroz L.L., Lieb B., Halanych K.M. 2011. Phylogenomics reveals deep molluscan relationships. Nature 477(7365):452–456.

Kulkarni S., Kallal R.J., Wood H., Dimitrov D., Giribet G., Hormiga G. 2021. Interrogating genomic-scale data to resolve recalcitrant nodes in the Spider Tree of Life. Mol. Biol. Evol, 38(3):891–903.

Lartillot N., Rodrigue N., Stubbs D., Richer J. 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Syst. Biol. 62(4):611–615.

Laumer C.E., Gruber-Vodicka H., Hadfield M.G., Pearse V.B., Riesgo A., Marioni J.C., Giribet G. 2018. Support for a clade of Placozoa and Cnidaria in genes with minimal compositional bias. Elife 7:e36278.

Lee H., Chen W.J., Puillandre N., Aznar-Cormano L., Tsai M.H., Samadi S. 2019. Incorporation of deep-sea and small-sized species provides new insights into gastropods phylogeny. Mol. Phylogenet. Evol. 135:136–147.

Lemmon A.R., Brown J.M., Stanger-Hall K., Lemmon E.M. 2009. The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. Syst. Biol. 58(1):130–145.

Lindberg D.R. 1986. Radular evolution in the Patellogastropoda. Am. Malacol. Union Bull. Abs. 4:115.

Mai U., Mirarab S. 2018. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. BMC Genomics 19(5):23–40.

McArthur A.G., Harasewych M.G. 2003. Molecular systematics of the major lineages of the Gastropoda. In: Lydeard C., Lindberg D.R., editors. Molecular systematics and phylogeography of Mollusks. Washington (DC): Smithsonian Books, p. 140–160.

McArthur A.G., Koop B.F. 1999. Partial 28S rDNA sequences and the antiquity of hydrothermal vent endemic gastropods. Mol. Phylogenet. Evol. 13(2):255–274.

McLean J.H. 1981. The Galapagos Rift limpet *Neomphalus*: relevance to understanding the evolution of a major Paleozoic-Mesozoic radiation. Malacologia 21:291–336.

Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., 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(5):1530–1534.

Mirarab S., Warnow T. 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31(12):i44–i52.

MolluscaBase, editor. 2021. MolluscaBase. Gastropoda. World Register of Marine Species. Available from: URL http://www.marinespecies.org/aphia.php?p=taxdetails&id=101.

Mongiardino K.N. 2021. Phylogenomic subsampling and the search for phylogenetically reliable loci. Mol. Biol. Evol. 38(9):4025–4038.

Oehlmann J., Schulte-Oehlmann U. 2003. Molluscs as bioindicators. In: Trace metals and other contaminants in the environment. Oxford: Elsevier. Vol. 6, p. 577–635.

Philippe H., Brinkmann H., Lavrov D.V., Littlewood D.T.J., Manuel M., Wörheide G., Baurain D. 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol. 9(3):e1000602.

Philippe H., de Vienne D.M., Ranwez V., Roure B., Baurain D., Delsuc F. 2017. Pitfalls in supermatrix phylogenomics. Eur. J. Taxon. 283:1–25.

Philippe H., Lopez P., Brinkmann H., Budin K., Germot A., Laurent J., Moreira D., Múller M., Guyader H.L. 2000. Early – branching or fast – evolving eukaryotes? An answer based on slowly evolving positions. Proc. R. Soc. B. 267(1449):1213–1221.

Philippe H., Poustka A.J., Chiodin M., Hoff K.J., Dessimoz C., Tomiczek B., Schiffer P.H., Müller S., Domman D., Horn M., Kuh H., Timmermann B., Satoh1 N., Hikosaka-Katayama T., Nakano H., Rowe M.L., Elphick M.R., Thomas-Chollier M., Telford M.J. 2019. Mitigating anticipated effects of systematic errors supports sister-group relationship between Xenacoelomorpha and Ambulacraria. Curr. Biol. 29(11):1818–1826.

Phillips M.J., Delsuc F., Penny D. 2004. Genome-scale phylogeny and the detection of systematic biases. Mol. Biol. Evol. 21(7):1455–1458.

Pisani D. 2004. Identifying and removing fast-evolving sites using compatibility analysis: an example from the Arthropoda. Syst. Biol. 53(6):978–989.

Ponder W.F., Lindberg D.R. 1997. Towards a phylogeny of gastropod molluscs: an analysis using morphological characters. Zool. J. Linnean Soc. 119:83–265.

Ponder W.F., Lindberg D.R., Ponder J.M. 2019. Gastropoda I: introduction and the stem groups. In: Biology and evolution of the Mollusca. Raton, FL: CRC Press. Vol. II, p. 289–363.

Rodríguez-Ezpeleta N., Brinkmann H., Roure B., Lartillot N., Lang B.F., Philippe H. 2007. Detecting and overcoming systematic errors in genome-scale phylogenies. Syst. Biol. 56(3):389–399.

Rota-Stabelli O., Lartillot N., Philippe H., Pisani D. 2013. Serine codon-usage bias in deep phylogenomics: pancrustacean relationships as a case study. Syst. Biol. 62(1):121–133.

Roure B., Baurain D., Philippe H. 2013. Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. Mol. Biol. Evol. 30(1):197–214.

Salvini-Plawen L.V. 1980. A reconsideration of systematics in the Mollusca. Malacologia 19:249–278.

Salvini-Plawen L.V., Haszprunar G. 1987. The Vetigastropoda and the systematics of streptoneurous Gastropoda (Mollusca). J. Zool. 211:747–770.

Sasaki T. 1998. Comparative anatomy and phylogeny of the recent Archaeogastropoda (Mollusca: Gastropoda). Univ. Mus. Univ. Tokyo Bull. 38:1–223.

Shen X.X., Hittinger C.T., Rokas A. 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nat. Ecol. Evol. 1(5):1–10.

Shen X.X., Li Y., Hittinger C.T., Chen X.X., Rokas A. 2020. An investigation of irreproducibility in maximum likelihood phylogenetic inference. Nat. Commun. 11(1):1–14.

Simion P., Delsuc F., Philippe H. 2020. To what extent current limits of phylogenomics can be overcome? In: Phylogenetics in the genomic era (Chapter 2.1). Hyper Articles en Ligne (HAL).

Simmons M.P. 2012. Radical instability and spurious branch support by likelihood when applied to matrices with non-random distributions of missing data. Mol. Phylogenet. Evol. 62(1): 472–484.

Simmons M.P., Gatesy J. 2016. Biases of tree-independent-character-subsampling methods. Mol. Phylogenet. Evol. 100: 424–443.

Simmons M.P., Sloan D.B., Springer M.S., Gatesy J. 2019. Gene-wise resampling outperforms site-wise resampling in phylogenetic coalescence analyses. Mol. Phylogenet. Evol, 131:80–92.

Simone L.R.L. 2011. Phylogeny of the Caenogastropoda (Mollusca), based on comparative morphology. Arq. Zool. 42(4):161–323.

Strassert J.F., Jamy M., Mylnikov A.P., Tikhonenkov D.V., Burki F. 2019. New phylogenomic analysis of the enigmatic phylum Telonemia further resolves the eukaryote tree of life. Mol. Biol. Evol. 36(4):757–765.

Streicher J.W., Schulte J.A., Wiens J.J. 2016. How should genes and taxa be sampled for phylogenomic analyses with missing data? An empirical study in iguanian lizards. Syst. Biol. 65(1):128–145.

Telford M.J., Budd G.E., Philippe H. 2015. Phylogenomic insights into animal evolution. Curr. Biol. 25(19):R876–R887.

Thiele J. (1929–1931). *Handbuch der systematischen Weichtierkunde*. Jena: Fischer. Vols 1 & 2, p. 1–778.

Uribe J.E., Irisarri I., Templado J., Zardoya R. 2019. New patellogastropod mitogenomes help counteracting long-branch attraction in the deep phylogeny of gastropod mollusks. Mol. Phylogenet. Evol. 133:12–23.

Uribe J.E., Kano Y., Templado J., Zardoya R. 2016. Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. Zool. Scr. 45(2):145–159.

Wagner P.J. 2002. Phylogenetic relationships of the earliest anisostrophically coiled gastropods. Smithsonian Contrib. Paleobiol. 88:1–152.

Whelan S., Irisarri I., Burki F. 2018. PREQUAL: detecting non-homologous characters in sets of unaligned homologous sequences. Bioinformatics 34(22):3929–3930.

Wiens J.J., Morrill M.C. 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. Syst. Biol. 60(5):719–731.

Winnepenninckx B., Steiner G., Backeljau T., De Wachter R. 1998. Details of gastropod phylogeny inferred from 18S rRNA sequences. Mol. Phylogenet. Evol. 9(1):55–63.

Yang Y., Smith S.A. 2014. Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: improving accuracy and matrix occupancy for phylogenomics. Mol. Biol. Evol, 31(11):3081–3092.

Zapata F., Wilson N.G., Howison M., Andrade S.C., Jörger K.M., Schrödl M., Goetz F.E., Giribet G., Dunn C.W. 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. Proc. R. Soc. B. 281(1794):20141739.