Phylogenomics of Neogastropoda: The Backbone Hidden in the Bush

Alexander E. Fedosov[1,2,*], Paul Zaharias[2], Thomas Lemarcis[2], Maria Vittoria Modica[2,3],
Mandë Holford[4,5,6], Marco Oliverio[2,7], Yuri I. Kantor[2,8], and Nicolas Puillandre[2]

[1]*Department of Zoology, Swedish Museum of Natural History, Box 50007, 10405 Stockholm, Sweden*
[2]*Institut de Systématique, Évolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE,
Université des Antilles, 57 rue Cuvier, CP 50, 75005 Paris, France*
[3]*Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Naples, Italy*
[4]*Department of Chemistry, Hunter College, Belfer Research Building, City University of New York, 413 E. 69th Street, BRB 424, New York,
NY 10021, USA*
[5]*Department of Invertebrate Zoology, the American Museum of Natural History, New York, NY 10024, USA*
[6]*PhD Programs in Biology, Biochemistry, and Chemistry, The Graduate Center of the City University of New York, New York,
NY 10016, USA*
[7]*Department of Biology and Biotechnologies "Charles Darwin," Sapienza University of Rome, Viale dell'Università 32, I-00185 Rome, Italy*
[8]*Department of Ecology and Morphology of Marine Invertebrates, A.N. Severtsov Institute of Ecology and Evolution,
Russian Academy of Sciences, Leninsky prospect, 33, 119071 Moscow, Russia*
[*]*Correspondence to be sent to: Department of Zoology, Swedish Museum of Natural History, Box 50007, 10405 Stockholm, Sweden;
E-mail: fedosovalexander@gmail.com.*

Abstract.—The molluskan order Neogastropoda encompasses over 15,000 almost exclusively marine species playing important roles in benthic communities and in the economies of coastal countries. Neogastropoda underwent intensive cladogenesis in the early stages of diversification, generating a "bush" at the base of their evolutionary tree, which has been hard to resolve even with high throughput molecular data. In the present study to resolve the bush, we use a variety of phylogenetic inference methods and a comprehensive exon capture dataset of 1817 loci (79.6% data occupancy) comprising 112 taxa of 48 out of 60 Neogastropoda families. Our results show consistent topologies and high support in all analyses at (super)family level, supporting monophyly of Muricoidea, Mitroidea, Conoidea, and, with some reservations, Olivoidea and Buccinoidea. Volutoidea and Turbinelloidea as currently circumscribed are clearly paraphyletic. Despite our analyses consistently resolving most backbone nodes, 3 prove problematic: First, the uncertain placement of Cancellariidae, as the sister group to either a Ficoidea-Tonnoidea clade or to the rest of Neogastropoda, leaves monophyly of Neogastropoda unresolved. Second, relationships are contradictory at the base of the major "core Neogastropoda" grouping. Third, coalescence-based analyses reject monophyly of the Buccinoidea in relation to Vasidae. We analyzed phylogenetic signal of targeted loci in relation to potential biases, and we propose the most probable resolutions in the latter 2 recalcitrant nodes. The uncertain placement of Cancellariidae may be explained by orthology violations due to differential paralog loss shortly after the whole genome duplication, which should be resolved with a curated set of longer loci. [Cancellariidae; marine mollusks; mollusca; phylogenomics; phylogenetic conflict; targeted enrichment.]

Understanding patterns of lineage relatedness is a fundamental task of life science and is the ultimate goal of the Tree of Life (TOL) initiative (Hinchliff et al. 2015). While the introduction of high-throughput sequencing technologies was initially believed to render TOL reconstruction a rather technical task depending mainly on adequate lineage sampling, it has become evident that the process is severely challenged by a phenomenon figuratively named "bushes in the tree of life" (Rokas and Carroll 2006). This pattern typically occurs in lineages that have undergone multiple cladogenesis events in a short time span (Rokas and Carroll 2006). Because the amount of phylogenetic signal is proportional to the TOL stem lengths, short stems require an increasingly large amount of data to be resolved (Lanyon 1988), and the inference of true topology in these segments of a tree is increasingly confounded by homoplasy (Takezaki et al. 2004). Nevertheless, quickly radiating lineages are among the most interesting to investigate, because intensive cladogenesis is a signature of evolutionary

success of the lineages (Hunter 1998), and understanding the origin of their prosperity requires a robust phylogenetic hypothesis (Whitfield and Lockhart 2007; Prum et al. 2015).

Being the second most species-rich phylum, Mollusca encompasses taxa with a remarkable diversity of body plans (Modica et al. 2019; Wanninger and Wollesen 2019; Kocot et al. 2020; Ponder et al. 2021) and unresolved or contentious relationships (Cunha et al. 2022; Uribe et al. 2022). Molluscan phylogenetics is challenged by the coexistence of uncertainties regarding the placement of ancient lineages, many of them being extinct (Sutton et al. 2016; Wanninger and Wollesen 2019), and a plethora of relatively recent successful radiations. The largest marine gastropod order, the Neogastropoda, is perhaps the most conspicuous example of the latter situation. Having radiated in the late Cretaceous and early Cenozoic, in the context of the Mesozoic Marine Revolution (Vermeij 1977), the Neogastropoda flourished in Cenozoic seas. Currently, the Neogastropoda exhibit a tremendous species richness

with over 15,000 species, corresponding to about one-fifth of the present-day molluskan diversity (MolluscaBase, available at https://www.molluscabase.org/). The vast majority of neogastropod species are carnivores. Being slow in motion, many lineages have developed a unique array of biochemical adaptations to mediate interactions with their prey and predators (Olivera et al. 2014; Ponte and Modica 2017; Kuznetsova et al. 2022). Deadly venoms of cone snails, comprising a high number of structurally and pharmacologically diversified neuropeptides referred to as conotoxins, are the best-known example of these biochemical adaptations. The unique pharmacological properties of conotoxins and their relevance for drug development (Safavi-Hemami et al. 2019) fuel the increasing multidisciplinary interest in Neogastropoda. However, the lack of a robust phylogenetic hypothesis of Neogastropoda is an impediment to the systematic investigation of the translational applications of their bioactive compounds. Therefore, reconstructing the phylogeny of the order will not only enable a reassessment of neogastropod systematics but also streamline evolutionary and biochemical research on this successful molluskan lineage.

The 60 currently recognized Neogastropoda families are classified into 7 superfamilies (Bouchet et al. 2017, with updates as per MolluscaBase); however, monophyly of the order remains questionable, and interrelationships among its main taxa are poorly understood. The published studies addressing Neogastropoda phylogenetics suffered complementary flaws. Morphology-based cladistic analyses (e.g., Riedel 2000; Simone 2011) were misled by the widespread homoplasies in character evolution. Molecular phylogenies based on the Sanger approach (e.g., Zou et al. 2011; Fedosov et al. 2019) lacked resolution at deep nodes, due to the clearly insufficient number of characters included. In turn, phylogenomic studies (Osca et al. 2015; Abdelkrim et al. 2018; Cunha and Giribet 2019; Lemarcis et al. 2022) had incomplete and unbalanced taxon sampling and/or suffered from the limitations inherent to mitogenome-based phylogenomics (Duchêne et al. 2011). The goal of this study is to resolve backbone Neogastropoda relationships through extensive lineage sampling and the application of a leading-edge phylogenomic approach to data generation and analysis. We successfully reconstructed a largely supported phylogenetic framework for the Neogastropoda, establishing for the first time affinities of previously enigmatic lineages. While our results suggest major revisions in the systematics of the Neogastropoda, their formal implementation extends beyond the scope of the present work.

Materials and Methods

Bait Design and Taxa Sampling

Details of the probe kit design, taxonomic sampling, lab work, and a comprehensive account on the initial stages of the data analysis are provided in the Supplementary Material. Briefly, 46 transcriptomes of 32 caenogastropod species were used for bait design (Zaharias et al. 2020; Lemarcis et al. 2022). All transcriptomes were (re)-assembled as detailed in Fassio et al. (2019) and then aligned against the genome of *Lottia gigantea* to identify exon/intron boundaries (Abdelkrim et al. 2018). Then, we identified a subset of 4456 exons (>180-bp) spanning approximately 1.3 Mb that were present in at least 2 families of Conoidea, and in at least 3 non-conoidean transcriptomes. The empirical exon sequences (i.e., those present in analyzed transcriptomes) were used alongside reconstructed ancestral sequences (in fast-evolving loci) for probe design, producing a set of 42,011 2× tiling 100-bp baits. After duplicate removal, the final set comprised 40,040 baits developed into a MyBait generation-5 biotinylated probes kit (Mycroarray, Arbor Biosciences, CA).

We obtained ethanol-preserved tissue samples of 135 taxa, covering 51 families of Neogastropoda and related lineages (Fig. 1, Supplementary Table S1), and complemented these data with 12 transcriptomic data sets that had the highest BUSCO completeness (Waterhouse et al. 2018). Library preparation was performed in 3 batches: the protocol detailed in Abdelkrim et al. (2018) was used for the specimens in the first and second batches, while the KAPA protocol was used for the third batch specimens. The libraries were paired-end sequenced on Illumina HiSeq 4000 and Illumina NovaSeq platforms, with read lengths of 100 and 150 bp, respectively. The final number of reads per library ranged from 851,299 (MNHN IM-2013-43718, *Glabella rosadoi*) to 44,946,221 (MNHN IM-2013-48309, *Xenophora* sp.), with a median number of 11,4517,88 reads per library.

Data Assembly and Loci Recovery

Data assembly and processing generally followed Abdelkrim et al. (2018). To maximize recovery of targeted loci for exon capture datasets we used 2 assemblers, SPAdes (v3.14) (Bankevich et al. 2012) and TRINITY (v2.9) (Grabherr et al. 2011), whereas only TRINITY was used for the transcriptomes. For each exon-capture sample, SPAdes and TRINITY assemblies were merged and clustered by running CD-HIT (Fu et al. 2012) with 99% identity.

We associated assembled contigs with targets using BLASTn, (e-value 1e−20) and used Exonerate (v2.2.0) under the est2genome model to redefine boundaries of the targeted exons (Abdelkrim et al. 2018). For each sample, all contigs that generated BLAST hits against the exon library were extracted from the assembly. The quality-trimmed reads were mapped against the contigs of interest with bowtie (v2.2.7) (Langmead and Salzberg 2012) to assess capture efficiency. We used samtools (v1.9) and bcftools (v1.3) (Li et al. 2009) for single-nucleotide polymorphism calling, aiming to assess heterozygosity in the captured sequences. Sites with coverage <4 were masked as "N," followed by removal of short sequences (length ≤70% of target length), low-quality sequences ("N" comprising > 30% of
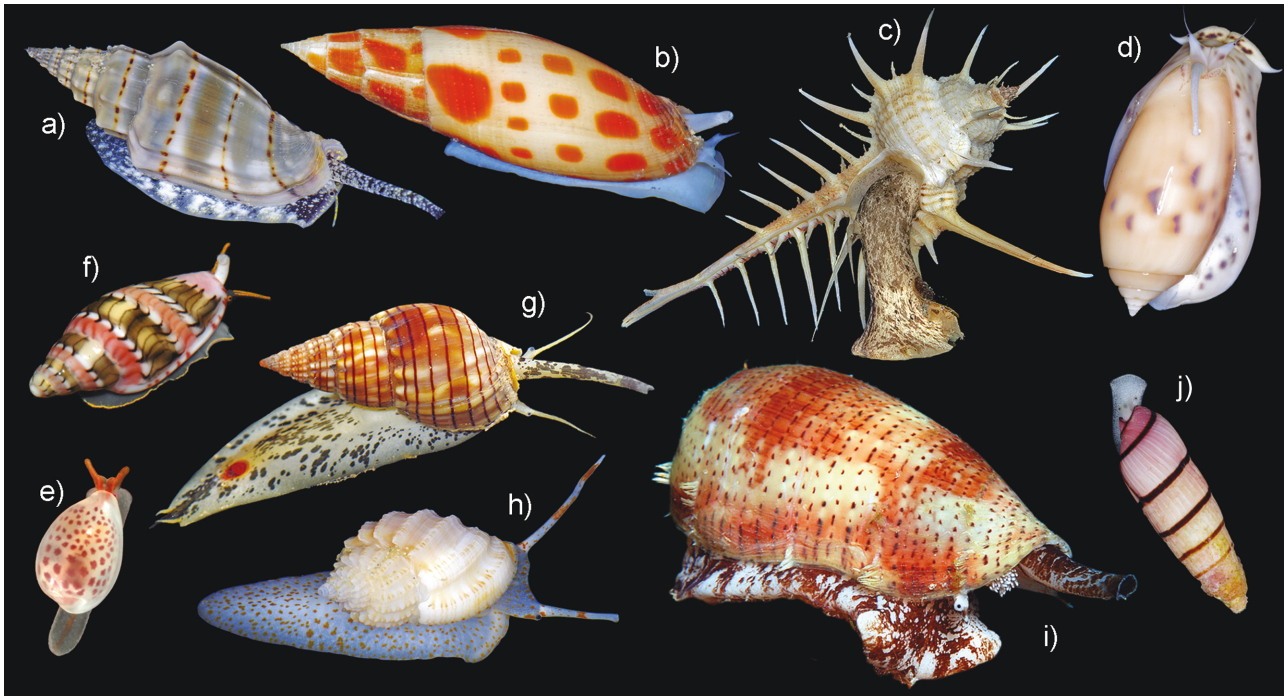
FIGURE 1. Living members of major Neogastropoda lineages. a) *Vexillum gruneri* (Costellariidae); b) *Mitra mitra* (Mitridae); c) *Murex tenuirostrum* (Muricidae), d) *Oliva amethystina* (Olividae); e) *Ticofurcilla* sp. (Cystiscidae), f) *Marginella festiva* (Marginellidae); g) *Nassarius glans* (Nassariidae); h) *Scalptia contabulata* (Cancellariidae), i) *Conus tulipa* (Conidae), and j) *Myurella pygmaea* (Terebridae). Photo credits: L. Charles, A. Ryansky, J. Johnson.

sequence length). Sequences with heterozygosity > 2 standard deviations from the mean were also removed.

Orthology Assessment

The sequences of interest were sorted by target identity and then aligned using MAFFT (v7.407) (Katoh and Standley 2013) with *G-INS-i*, and *–adjust_direction* option enabled. The alignments were then translated using MACSE (v2.06) (Ranwez et al. 2018), and the obtained amino-acid sequences sorted back by sample for orthogroup identification with ORTHOFINDER (v2.5.4) (Emms and Kelly 2019). The 30 most complex orthogroups comprising multiple sequences for nearly all samples were removed. Gene trees were reconstructed for the remaining 3000 orthogroups with ≥65 samples represented, using RAxML (v8.2.12) under the GTRGAMMA model with 100 bootstrap replicates (Stamatakis 2006). When a sample was represented by multiple sequences in an orthogroup alignment, we first used a custom Python script S10-3 to remove residual cross-contamination based on the orthogroup tree topology, coverage data, and sequences lengths, and then selected the largest 1:1 ortholog subtree using PhyloPyPruner (v1.2.6) (Thalén 2018). We removed terminal long branches using the custom Python script S10-4 and end-trimmed the alignments using TRIMAL (v1.2) (Capella-Gutiérrez et al. 2009). The 112 taxa with the highest data occupancy (i.e., the smallest amount of missing data (*highDO* taxa)) were retained for downstream analyses.

Matrix Assembly and Phylogenetic Analyses

The 1817 orthogroup alignments comprising ≥35 aa sites with ≥70 highDO taxa included generated the matrix NEO70 (total 125,508 sites, 20.4% missing data). To further reduce missing data, a subset of 731 alignments comprising ≥95 highDO taxa were selected to build the matrix NEO95 (total 52,805 sites, 11.4% missing data). We used RAxML with a PROTGAMMALG4X model and 20 rapid bootstraps (Cunha et al. 2022) for a second-round gene tree reconstruction, and further subsampled the matrix NEO95 using GenesortR (Mongiardino Koch 2021). GenesortR first scores all loci based on 7 parameters reflecting phylogenetic "usefulness," so the loci that could bias phylogenetic reconstructions can be removed. The obtained matrix NEO95-GSR500 consisted of the 500 "best" loci and included 37,958 aligned amino-acid sites.

Multispecies coalescent phylogenies were reconstructed from 3 respective sets of gene trees by using both ASTRAL III (v5.6.3) (Zhang et al. 2018) and ASTEROID (v1.0) (Morel et al. 2023). Maximum Likelihood phylogenies were reconstructed with IQ-TREE (v.2.2.1) (Minh et al. 2020), performed on both gene-partitioned (IQ-part) and on unpartitioned matrices with best-fit profile mixture models (IQ-PMM). In partitioned analyses, best-fit models were estimated for edge-unlinked partitions, and the partitions with compatible model parameters merged prior to the tree search (**-st AA -msub nuclear -ninit 10 -bb 1500 -sp partition_file -m MFP + MERGE -rcluster 10 -madd LG4M,LG4X -mrate G,R,E**). Due

to the prohibitive runtimes of the MFP-MERGE mode, partition merging was not performed on the NEO70 dataset. In the IQ-PMM analyses, the command line of Cunha et al. (2022) was run to identify the best-fit exchange matrix (**-st AA—msub nuclear -ninit 10 -bb 1500 -m MFP -mset LG,WAG -rcluster 10 -mfreq F + C40/60 -mrate G,R**). Sixty mixture classes (C60) were enabled for NEO95 matrices. In contrast, we only allowed 40 mixture classes (C40) for NEO70 due to 1 TB RAM limitation of our phylogenetic server.

We performed Bayesian inference by running PhyloBayes (v4.1) (Lartillot et al. 2013) on the 2 NEO95 matrices, under CAT-GTR model, disregarding constant sites. Each analysis was run in 4 chains and terminated once convergence criteria (accessed with tracecomp) were achieved for at least 2 chains (8771 and 11,760 generations for NEO95 and NEO95-GSR500 matrices, respectively).

To visualize overall similarities among the obtained tree topologies, we first ran a custom Python script S10-6 to retrieve all unique clades comprising 2 or more taxa from the trees from the analyses described above (Fig. 2a), and then compiled a clade presence–absence (coded as 1 and 0) matrix for these 14 trees. This matrix was subjected to principal component analysis (PCA) using PAST (Hammer et al. 2001).

Our phylogenetic analyses repeatedly recovered alternative topologies at three backbone nodes. The nodes that produced conflicting topologies define (i) the placement of Cancellariidae (referred to as baseNEO), (ii) the first offshoot of Core Neogastropoda (baseCore) (iii) the affinities of early branching Buccinoidea (base-Buc) (Fig. 2a–e). To understand the source of support for these conflicting hypotheses, for each contradictory relationship, we performed site-wise phylogenetic signal measures (ΔSLS) as detailed by Shen et al. (2017). First, 6 analyses under constrained topologies were run on the matrix NEO95 (two for each node, one under ML-PMM, another with partitions) to obtain best-scoring alternative topologies. Then for each pair of alternative trees, SLS (per site likelihood score) was calculated under respective model (PROTGAMMALG4X was run as unpartitioned model), by running RAxML with –f G option.

We calculated site-wise phylogenetic signal (ΔSLS) by subtracting an alternative topology' SLS from the main topology' SLS (therefore, positive ΔSLS values are those supporting the main topology). We employed Approximately Unbiased (AU) test in CONSEL (Shimodaira and Hasegawa 2001) to check if one topology is significantly better than the alternative. By summing up ΔSLS values for each locus, we computed ΔGLS values as a proxy of gene-wise phylogenetic signal (custom Python script S10-7). In addition to ΔGLS, we calculated standard deviation for ΔSLS values of each locus, and we used proportion of SD(ΔSLS) to ΔGLS as a measure of noise in the phylogenetic signal. If this proportion exceeded 10 for a locus (suggesting highly dissimilar site-wise signals, summing up to

close-to-zero ΔGLS), this locus was excluded as bearing a contradictory signal. The remaining loci were divided into 3 subsets: 10% loci with the lowest ΔGLS, 10% loci with the highest ΔGLS, and the remaining 80%. Then we performed a t-test to find out whether there was a significant difference among the subsets with respect to potential biases (Saturation, Compositional heterogeneity, Evolutionary rate), assessed by GenesortR.

Results

Support for Neogastropoda Superfamilies and Families

The composition and relationships within the superfamily level clades are highly congruent among the 17 trees reconstructed from the 3 analyzed datasets (Supplementary Figs. S2–S15). All our analyses support the monophyly of Muricoidea (=Muricidae), Mitroidea, and Conoidea. The remaining 4 superfamilies are consistently recovered as paraphyletic. Volutoidea comprises 2 unrelated clusters: Cancellariidae and Volutidae plus marginelliform gastropods (Cystiscidae and Marginellidae). The Panamanian species *Triumphis distorta* traditionally placed in Pseudolividae but unequivocally recovered within Buccinoidea, violates reciprocal monophyly of Olivoidea and Buccinoidea. Whereas Olivoidea excluding *Triumphius* is consistently monophyletic, relationships at the base of Buccinoidea are contradictory (see below). The Turbinelloidea taxa form five unrelated highly supported clades: (i) Volutomitridae plus *Exilioidea* (Ptychatractidae), (ii) Costellariidae plus *Exilia* (Ptychatractidae), (iii) Columbariidae, (iv) Vasidae, and (v) Turbinellidae. The extant families Harpidae and Babyloniidae, currently not assigned to superfamilies (Mollusca base accessed on 17 July 2023), represented by respectively 3 and 1 species in our dataset, do not show consistent affinities with any other lineage. Of the 25 tonnoidean and neogastropod families represented by two or more species in our dataset, monophyly is consistently rejected for Pseudolividae (see above) and Ptychatractidae (with *Exilioidea* always being a sister group to Volutomitridae, and *Exilia* the sister group to Costellariidae). Furthermore, in 5 analyses, Volutidae is retrieved as paraphyletic in relation to the marginelliform clade (Fig. 2a), and in 6 (all coalescence-based) Nassariidae is paraphyletic in relation to other Buccinoidea (for support values of the Volutidae and Nassariidae nodes see Fig. 2a). One important finding among the family level relationships is the placement of Columbellidae within the Buccinoidea as a sister to the Colubrariidae-Colidae-Prosiphonidae-Eosiphonidae clade, which was strongly supported in all our analyses.

Backbone Relationships of Neogastropoda

To address backbone Neogastropoda relationships, we select the IQ-PMM tree obtained from the NEO95 matrix (Fig. 2a), which features the most frequently
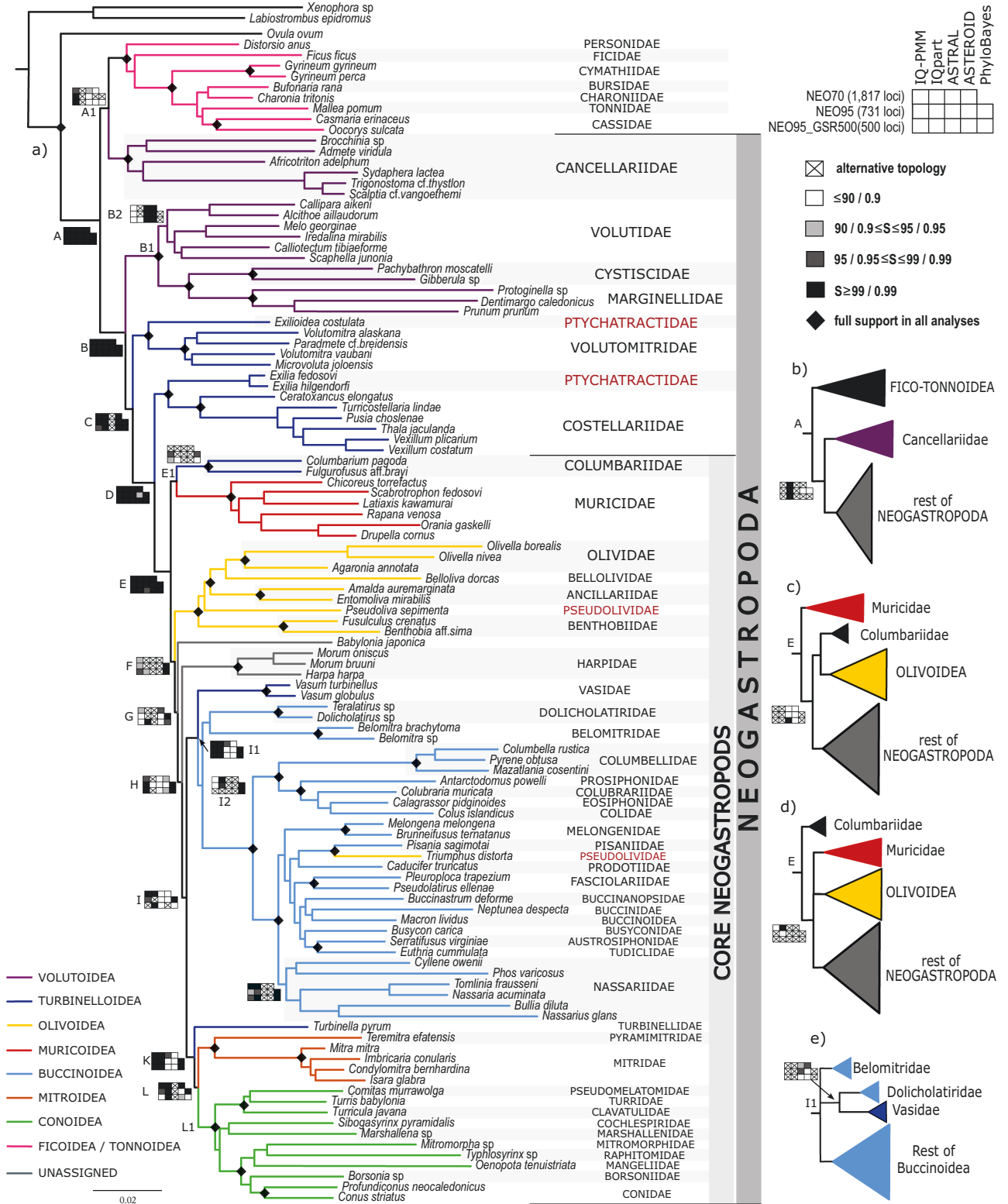
FIGURE 2. a) IQ-TREE-PMM tree generated with the NEO95 matrix; Ptychatractidae and Pseudolividae consistently recovered paraphyletic. Node supports for deep nodes summarized from 14 analyses as shown on top-right inset, IQ-PMM, IQ tree under Profile Mixture Model; tree branches are color-coded according to the current superfamily classification—bottom-left inset; scale as probability of substitution per site; b—e) alternative topologies, and their support; b) placement of Cancellariidae; c and d) base of the "core Neogastropoda"; and e) base of Buccinoidea.

sampled topology at each backbone node (denoted as A-L), and at the base of Neogastropoda (node A1). We recognize as problematic nodes those with a consistently sampled alternative topology, criteria for consistency being: recovered (i) in at least 3 analyses, (ii) with at least 2 different inference methods, and (iii) with moderate or high support in at least one analysis. The first such problematic node concerns the placement of the family Cancellariidae at the base of Neogastropoda, either as a sister group to the Ficoidea–Tonnoidea clade (Fig. 2a) or to the rest of the Neogastropoda (Fig. 2b). The first topology receives high support in all IQ-PMM analyses, the second in the partitioned IQ-Tree analyses, whereas the coalescence-based and PhyloBayes inferences lack support for the placement of Cancellariidae.

The remaining neogastropod taxa always form a maximally supported clade (node B), and the topology at the three deepest nodes D–E is consistent and highly supported across most analyses. These nodes correspond to the consecutively branching off (i) Volutidae plus marginelliform gastropods (C), (ii) Volutomitridae plus *Exilioidea* (D), and (iii) Costellariidae plus *Exilia* (E). The remaining taxa are always recovered in a highly supported cluster (node E), which we refer to from here onwards as "core Neogastropoda." This clade comprises 7 major lineages corresponding to (i) family Columbariidae, (ii) family Muricidae, (iii) superfamily Olivoidea (except *Triumphius*), (iv) family Babyloniidae, (v) family Harpidae, (vi) BV clade (Buccinoidea including *Triumphius* and Vasidae), and (vii) TMC clade, (Turbinellidae, (Mitroidea, Conoidea)).

Three conflicting topologies at the base of core Neogastropoda (nodes E1, F) correspond to either Muricidae (Fig. 2c), or Columbariidae (Fig. 2d), or Muricidae plus Columbariidae (most consistently recovered, Fig. 2a), being the sister group to all other core lineages. The latter topology is supported by 2 IQ-PMM and both PhyloBayes analyses and invariably places the Olivoidea as the next branching lineage. In contrast, the partitioned IQ-TREE analyses (except the NEO95-500 matrix) favor Columbariidae as the first branching core lineage (Fig. 2c), whereas all coalescence-based analyses place Muricidae at the base of the core radiation, and Columbariidae as a sister group to Olivoidea, though usually without support.

The affinities among the 4 remaining lineages are generally more consistent and suggest a sister relationship between the BV and TMC clades, with Harpidae being a sister group to (BV, TMC), and Babyloniidae a sister group to (Harpidae, (BV, TMC)). Two further problematic nodes, I2 and L1, concern relationships at the base of the buccinoidean and conoidean radiations, respectively. The conflicting topologies at the base of Buccinoidea concern affinities of the early branching buccinoidean families Belomitridae and Dolicholatiridae. In all coalescence-based and some ML inferences, either both these families, or only Dolicholatiridae appear more closely related to Vasidae than to the rest of the

Buccinoidea (Fig. 2e, coalescence-based analyses, moderately supported, or lacking support).

Finally, fourth major uncertainty affects relationships at the base of Conoidea, where our analyses were inconclusive in defining the earliest branching lineage. A topology in which Cochlespiridae is the sister group to all other conoideans (Abdelkrim et al. 2018) is recovered in both PhyloBayes analyses, ASTEROID and IQ-PMM (both on the matrix NEO95_500), whereas the majority of the analyses suggest a sister relationship between Cochlespiridae and Marshallenidae. Possibly, this persistent grouping is an LBA artifact that is efficiently countered by CAT GTR model in Phylobayes (Uribe et al. 2018). Since the present study focuses on the relationships among the major Neogastropoda lineages, and the relationships within Conoidea have recently been addressed with phylogenomics (Abdelkrim et al. 2018), we have reduced the taxon coverage in this lineage. Having noted the robustly supported monophyly of the Conoidea in all analyses, we did not examine the sources of conflict among its lineages.

Sources of Phylogenetic Conflict

The PCA performed on the matrix summarizing clade presence–absence (Supplementary Fig. S16) shows that topology at conflicting nodes depends more on the phylogenetic inference method than on the matrix used. The two first principal components explained 45.7% of the observed variation. The first PC clearly separates the coalescence-based and concatenation-based analyses, indicating that ASTEROID trees are overall slightly more congruent with the ML- and Bayesian trees. The second PC separates the partitioned IQ-TREE trees (on top of the plot), the IQ-PMM trees, and the PhyloBayes trees, but also bears some signal of the matrix analyzed: for each inference method, NEO95-500 trees are placed on the diagram lower than the trees obtained from the larger datasets NEO70 and NEO95.

The AU tests on the ΔSLS values calculated under GAMMALG4X did not prefer one of the conflicting topologies over another in any comparison (Supplementary Table S2). For the partitioned data, only the main topology at the nodes I1/I2 (monophyletic Buccinoidea) fits the data significantly better than the respective alternative topology. The t-test suggests that regardless of the query node, the loci with a strong ΔGLS on average show a higher evolutionary rate (the reason why they offer some resolution), and under a partitioned model, are more likely to be affected by both compositional heterogeneity and saturation (Supplementary Table S2, Fig. 3). Under GAMMALG4X, the loci with a strong signal favoring affinity of the Cancellariidae with Ficoidea-Tonnoidea are clearly fewer, and they are significantly more affected by both saturation and high compositional heterogeneity (Fig. 3d and j). Furthermore, loci with strong signal favoring Vasidae–Dolicholatiridae–Belomitridae affinity at the base of Buccinoidea show
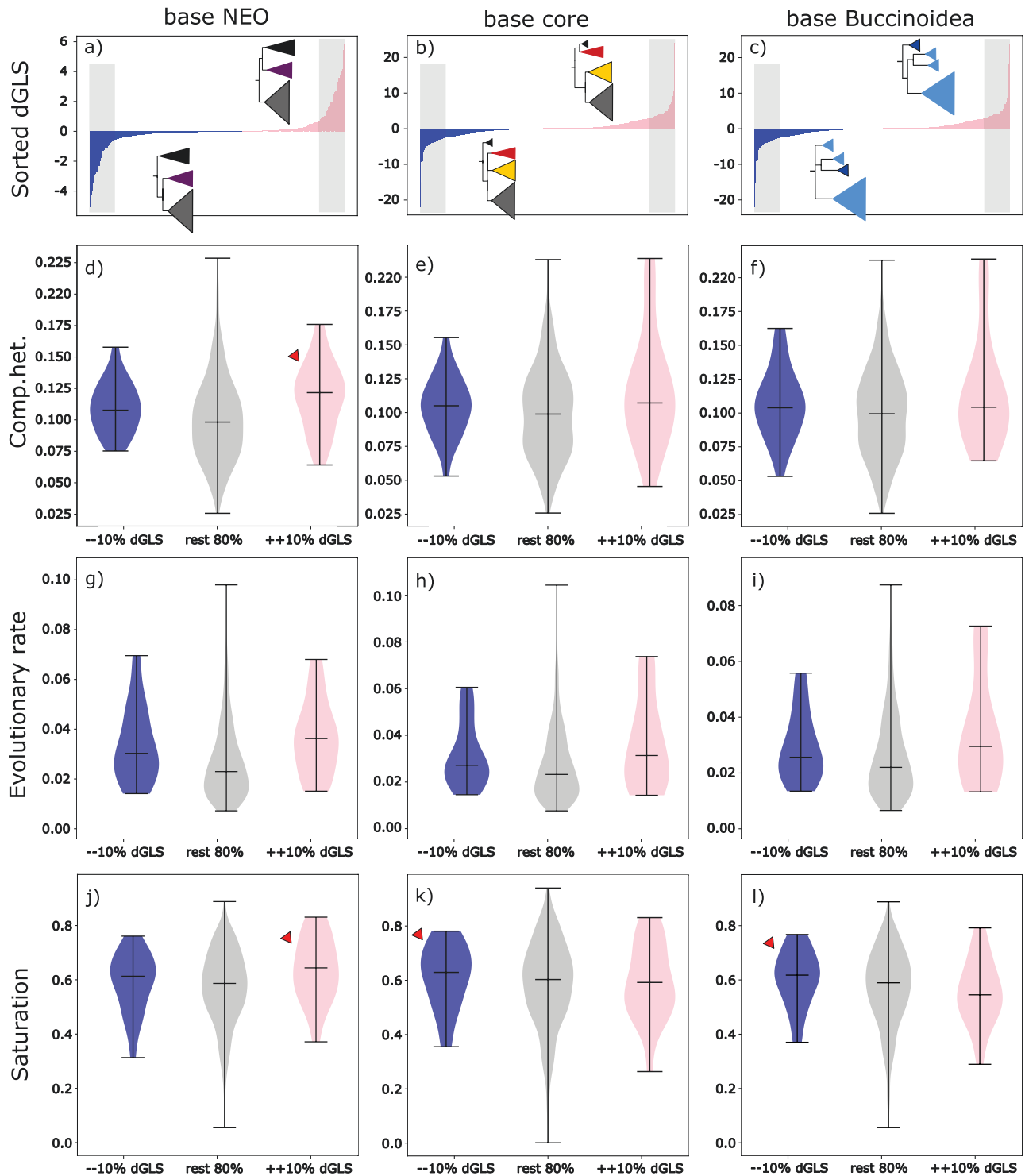
FIGURE 3. Phylogenetic signal (ΔGLS) supporting alternative topologies in 3 contradictory nodes. a–c) Distribution of ΔGLS values in the 731 loci of the NEO95 matrix under GAMMAPROTLG4X model; bars to the right representing loci supporting major topology (recovered in IQ-PMM analysis); those to the left supporting the alternative topology (from constrained topology in IQ-PMM analysis)—both respective topologies are shown; gray zones mark 10% of loci with strongest ΔGLS signal for one or another topology. d–l) Loci metrics, compositional heterogeneity (second row), evolutionary rate (third row), and saturation at the third codon position (bottom row) in 3 groups of loci by ΔGLS: 10% loci with strongest ΔGLS support for the alternative topology (left), 80% of loci with weak ΔGLS values, irrespective of supported topology (center), 10% loci with strongest ΔGLS support for the main topology (right). Pointers mark values mentioned in the text.

higher levels of saturation compared to those with strong support for the main topology (Fig. 3l).

Discussion

Relationships at the Base of Neogastropoda

Recent studies on various metazoan lineages shared the common conclusion that the presence of conflicting signals is an inherent property of phylogenomic datasets (Betancur-R. et al. 2019; Parins-Fukuchi et al. 2021; Cunha et al. 2022; Mongiardino Koch et al. 2023) and suggested that the true topology could be identified by accounting for technical errors and exploring sources of the conflicts (but see Mongiardino Koch et al. 2023). In our analyses, the most challenging conflict concerns the placement of Cancellariidae, either as a sister to the rest of Neogastropoda or to Ficoidea-Tonnoidea, as supported by partitioned and ML-PMM analyses, respectively. We demonstrate that the latter topology may, at least partly, be driven by loci with high levels of saturation and compositional heterogeneity.

One further factor adding to uncertainty at this node is the inevitably difficult orthology inference due to the whole genome duplication (WGD) event that pre-dated the neogastropod radiation, confirmed by karyological (Hallinan and Lindberg 2011) and whole genome data (Pardos-Blas et al. 2021; Farhat et al. 2023). Although redundant gene copies are usually quickly lost, the clades that have diverged shortly after a WGD event may differentially retain paralogs, leading to inaccurate phylogeny estimates (Xiong et al. 2022). Hence, the contradictory signals regarding the placement of Cancellariidae may be explained by a differential pattern of paralogs loss in Tonnoidea, Cancellariidae, and the rest of Neogastropoda, as the separation of these lineages was likely one of the first major splits that followed the WGD (Hallinan and Lindberg 2011; Farhat et al. 2023). Because the probability of the gene loss is proportional to the internal branch length (Xiong et al. 2022), the extent of the differential gene loss should be less in the pair Cancellariidae/Ficoidea-Tonnoidea compared to the pair Cancellariidae/rest of Neogastropoda, as the lineages in the latter pair are invariably separated by a higher sum of branch lengths (custom Python script S10-8, Supplementary Table S3). As a result, there would exist a pool of loci alignments where gene copies in Cancellariidae are orthologous to those in Ficoidea-Tonnoidea but not in the rest of Neogastropoda, and these alignments would expectedly favor the (Cancellariidae, (Ficoidea, Tonnoidea)) grouping.

It is noteworthy that none of our analyses recovered Cancellariidae as a sister to Tonnoidea plus Neogastropoda, a placement supported by recent mitogenomic phylogenies (Osca et al. 2015; Lemarcis et al. 2022) but based on a very limited sampling of Cancellariidae. Morphological data generally supports monophyly of Neogastropoda. However, the key anatomical traits for understanding Neogastropoda evolution, radula, and valve of Leiblein, are highly aberrant in Cancellariidae (Modica et al. 2011), and they do not provide any clues on the affinities of this enigmatic lineage. Further genomic data, whole genome assemblies, or a carefully curated set of longer loci would be instrumental for disentangling the relationships at the base of Neogastropoda radiation.

Relationships Within the Core Neogastropoda

We examined deep relationships within the order Neogastropoda based on both an unprecedented taxonomic coverage (112 neogastropod taxa representing 48 families) and a representative genomic sampling (from 1817 loci with ~20.4% of missing data to 731 loci with 11.6% missing data only). Although we failed to recover a single topology for the Neogastropoda tree, high support was retrieved for most backbone nodes, allowing to localize uncertainty to 4 specific nodes. Three of them are associated with the origin of remarkably species-rich radiations: the core Neogastropoda, the superfamily Buccinoidea, and the superfamily Conoidea.

Within core Neogastropoda, the sister relationship of Columbariidae and Muricidae is morphologically plausible, albeit their similarities are mainly limited to shared plesiomorphies (Kantor 2002). The most frequently sampled alternative topology (Fig. 2c) results from the coalescence-based analyses, however, with low support values at query nodes. Furthermore, recent findings casted doubts on the ability of summary-based approaches to accurately resolve deep and intricate phylogenies (e.g., Gatesy and Springer 2014). Therefore, we regard this alternative topology as rather unlikely. Similarly, the only analysis supporting Columbariidae as a sister group to all other core lineages (NEO70, partitioned IQ-TREE; Fig. 1d) relies on a larger proportion of missing data, with inference performed on very short loci, resulting in the overall unrealistically high bootstrap support values (Thomson and Brown 2022). Therefore, we consider the topology where the Columbariidae-Muricidae lineage represents the first offshoot within core Neogastropoda as the most probable. This topology is most frequently sampled and is supported in nearly half of our concatenation-based inferences.

Two very short branches at the base of the Buccinoidea separate first the Vasidae and then the Dolicholatiridae and Belomitridae from the main stem of Buccinoidea. Vasidae, Dolicholatiridae, and Belomitridae share somewhat similar radulae, with bicuspidate lateral teeth (Medinskaya et al. 1996). However, Dolicholatiridae and Belomitridae, similarly to all other Buccinoidea, lack accessory salivary glands and an anal gland, whereas the latter is present in Vasidae. While it is tempting to speculate that the loss of accessory salivary glands and anal gland in Dolicholatiridae, Belomitridae, and all other Buccinoidea is a result of a single evolutionary event supporting their affinity, a shared loss of a trait cannot be considered evidence

of affinity (Strong and Lipscomb 1999). Therefore, the anatomical evidence is inconclusive, as to whether Dolicholatiridae–Belomitridae are closer to the Vasidae or to the major Buccinoidea clade. Phylogenetic uncertainty here is likely due to the series of very short branches followed by a longer one leading to the major Buccinoidea. Topology resolution in proximity of such patterns is susceptible to a biased signal from loci affected by saturation (Breinholt and Kawahara 2013), and indeed we detected higher levels of saturation in loci with a strong signal for Vasidae–Dolicholatiridae–Belomitridae grouping. This result, and the generally consistent support for monophyletic Buccinoidea in our concatenation-based analyses, prompt us to consider this topology (Fig. 2a) as the most probable.

Rapid diversification of the core Neogastropoda coincided with the dramatic paleo-climatic events of the late Cretaceous and the K-Pg boundary (Vermeij 1977). This period was marked by the origin of many lineage-specific morphological innovations, mainly associated with the dynamic evolution of foregut underpinning the diversification of feeding strategies in Neogastropoda (Ponder 1973; Kantor 2002). Parins-Fukuchi et al. (2021) suggested that the complex evolutionary patterns of genes linked to bursts of morphological disparity could also complicate phylogenetic inference. Similar to Neogastropoda, the evolutionary histories of two iconic vertebrate radiations—birds and mammals—suffer from a lack of resolution at the phylogenetic splits typically aligned with the K-Pg boundary. Remarkably, even with significantly more genomic resources available in these lineages, certain relationships remain challenging to address due to pervasive phylogenomic conflicts. Nonetheless, we anticipate that the present phylogeny will serve as a valuable guide for the future expansion of genomic resources for Neogastropoda. This expansion is crucial for understanding the evolutionary history of this remarkable group of marine invertebrates.

Relationships of Neogastropoda and Their Implications for Systematics

Our findings unequivocally support the monophyly of 5 Neogastropod superfamilies: Conoidea, Muricoidea, Mitroidea, Olivoidea, and Buccinoidea (with the reassignment of *Triumphius* from the Olivoidea to the Buccinoidea). Within Buccinoidea, we confidently place the previously disputed Columbellidae as the sister group to the Colubrariidae-Colidae-Prosiphonidae-Eosiphonidae clade. Furthermore, all the concatenation-based inferences confirmed the monophyly of Nassariidae, questioned by Kantor et al. (2022). Notably, we identify Vasidae for the first time as the sister group to the Buccinoidea. Indeed, the affinity of Vasidae and Buccinoidea *sensu* Kantor et al. (2022) is recovered in all our analyses and has a much stronger support than the Buccinoidea clade itself. Based on this outcome, we propose the inclusion of Vasidae in the superfamily Buccinoidea.

The scope of the superfamily Volutoidea must be restricted to the content of the clade including Volutidae and marginelliform gastropods (Fedosov et al. 2019). Future investigations are required to validate the monophyly of Volutidae and ascertain the placement of enigmatic taxa such as the families Granulinidae and Marginellonidae. The family Cancellariidae should definitely be assigned to a separate superfamily Cancellarioidea, as previously proposed by Ponder (1973) and Bouchet and Rocroi (2005).

Our analyses reveal the polyphyly of Turbinelloidea (*sensu* Fedosov et al. 2017), a result that necessitates profound revisions to neogastropod systematics. Some changes, such as the inclusion of Columbariidae in Muricoidea, and Vasidae in the Buccinoidea, can be readily inferred from the present phylogeny, others yet to be proposed. The existing scheme with 8 superfamilies leaves out of superfamilies at least 4 major lineages retrieved in our analyses. Therefore, the establishment of 4 new superfamilies to accommodate (i) Volutomitridae plus *Exilioidea*, (ii) Costellariidae plus *Exilia*, (iii) Babyloniidae, and (iv) Harpidae, emerges as the most reliable systematic arrangement based on the reconstructed tree topology.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.8931zcrx5.

Data Accessibility

Raw read data (both transcriptomic and genomic) are available under the NCBI Bioproject PRJNA885117. Phylogenetic matrices, gene alignments and trees, output of the orthology inference software, as well as the original scripts, are available as supplementary data at Dryad https://doi.org/10.5061/dryad.8931zcrx5.

References

Abdelkrim J., Aznar-Cormano L., Fedosov A., Kantor Y., Lozouet P., Phuong M., Zaharias P., Puillandre N. 2018. Exon-capture based phylogeny and diversification of the venomous gastropods (Neogastropoda, Conoidea). Mol. Biol. Evol. 35:2355–2374.

Bankevich A., Nurk S., Antipov D., Gurevich A.A., Dvorkin M., Kulikov A.S., Lesin V.M., Nikolenko S.I., Pham S., Prjibelski A.D., Pyshkin A.V., Sirotkin A.V., Vyahhi N., Tesler G., Alekseyev M.A., Pevzner P.A. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19:455–477.

Betancur-R R., Arcila D., Vari R.P., Hughes L.C., Oliveira C., Sabaj M.H., Ortí G. 2019. Phylogenomic incongruence, hypothesis testing, and taxonomic sampling: the monophyly of characiform fishes. Evolution. 73:329–345.

Bouchet P., Rocroi J.–P. 2005. Classification and nomenclator of gastropod families. Malacologia. 47:1–397.

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. 61:1–526.

Breinholt J.W., Kawahara A.Y. 2013. Phylotranscriptomics: saturated third codon positions radically influence the estimation of trees based on next-gen data. Genome Biol. Evol. 5:2082–2092.

Capella-Gutiérrez S., Silla-Martínez J.M., Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25:1972–1973.

Cunha T.J., Giribet G. 2019. A congruent topology for deep gastropod relationships. Proc. Biol. Sci. 286:20182776.

Cunha T.J., Reimer J.D., Giribet G. 2022. Investigating sources of conflict in deep phylogenomics of vetigastropod snails. Syst. Biol. 71:1009–1022.

Duchêne S., Archer F.I., Vilstrup J., Caballero S., Morin P.A. 2011. Mitogenome phylogenetics: the impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation. PLoS One 6:e27138.

Emms D.M., Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20:238.

Farhat S., Modica M.V., Puillandre N. 2023. Whole genome duplication and gene evolution in the hyperdiverse venomous gastropods. Mol. Biol. Evol. 40:msad171.

Fassio G., Modica M.V., Mary L., Zaharias P., Fedosov A.E., Gorson J., Kantor Y.I., Holford M., Puillandre N. 2019. Venom diversity and evolution in the most divergent cone snail genus *Profundiconus*. Toxins. 11:623.

Fedosov A.E., Caballer Gutierrez M., Buge B., Boyer F., Sorokin P.V., Puillandre N., Bouchet P. 2019. Mapping the missing branch on

Neogastropoda tree of life: molecular phylogeny of marginelliform gastropods. J. Moll. Stud. 85:439–451.

Fedosov A.E., Puillandre N., Herrmann M., Dgebuadze P., Bouchet P. 2017. Phylogeny, systematics and evolution of the family Costellariidae (Gastropoda: Neogastropoda). Zool. J. Linn. Soc. 179:541–626.

Fu L., Niu B., Zhu Z., Wu S., Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28:3150–3152.

Gatesy J., Springer M.S. 2014. Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Mol. Phylogenet. Evol. 80:231–266.

Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N., Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29:644–652.

Hallinan N.M., Lindberg D.R. 2011. Comparative analysis of chromosome counts infers three paleopolyploidies in the mollusca. Genome Biol. Evol. 3:1150–1163.

Hammer O., Harper D.A.T., Ryan P.D. 2001. PAST: paleontological statistics software package for education and data analyses. Paleontol. Electron 4:1–9.

Hinchliff C.E., Smith S.A., Allman J.F., Burleigh J.G., Chaudhary R., Coghill L.M., Crandall K.A., Deng J., Drew B.T., Gazis R., Gude K., Hibbett D.S., Katz L.A., Iv H.D.L., McTavish E.J., Midford P.E., Owen C.L., Ree R.H., Rees J.A., Soltis D.E., Williams T., Cranston K.A. 2015. Synthesis of phylogeny and taxonomy into a comprehensive tree of life. Proc. Natl. Acad. Sci. U.S.A. 112:12764–12769.

Hunter J.P. 1998. Key innovations and the ecology of macroevolution. Trends Ecol. Evol. 13:31–36.

Kantor Y.I. 2002. Morphological prerequisites for understanding neogastropod phylogeny. Boll. Malacol (Suppl. 4):161–174.

Kantor Y.I., Fedosov A.E., Kosyan A.R., Puillandre N., Sorokin P.A., Kano Y., Clark R., Bouchet P. 2022. Molecular phylogeny and revised classification of the Buccinoidea (Neogastropoda). Zool. J. Linn. Soc. 194:789–857.

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

Kocot K.M., Poustka A.J., Stöger I., Halanych K.M., Schrödl M. 2020. New data from Monoplacophora and a carefully-curated dataset resolve molluscan relationships. Sci. Rep. 10:101.

Kuznetsova K.G., Zvonareva S.S., Ziganshin R., Mekhova E.S., Dgebuadze P., Yen D.T.H., Nguyen T.H.T., Moshkovskii S.A., Fedosov A.E. 2022. Vexitoxins: conotoxin-like venom peptides from predatory gastropods of the genus *Vexillum*. Proc. Biol. Sci. 289:20221152.

Langmead B., Salzberg S.L. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9:357–359.

Lanyon S.M. 1988. The stochastic mode of molecular evolution: what consequences for systematic investigations? Auk 105:565–573.

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

Lemarcis T., Fedosov A.E., Kantor Y.I., Abdelkrim J., Zaharias P., Puillandre N. 2022. Neogastropod (Mollusca, Gastropoda) phylogeny: a step forward with mitogenomes. Zool. Scr. 51:550–561.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R.; 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.

Medinskaya A.I., Harasewych M.G., Kantor Y.I. 1996. On the anatomy of *Vasum muricatum* (Born, 1778) (Neogastropoda, Turbinellidae). Ruthenica 5:131–138.

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. 37:1530–1534.

Modica M.V., Bouchet P., Cruaud C., Utge J., Oliverio M. 2011. Molecular phylogeny of the nutmeg shells (Neogastropoda, Cancellariidae). Mol. Phylogenet. Evol. 59:685–697.

Modica M.V., Gorson J., Fedosov A.E., Malcolm G., Terryn Y., Puillandre N., Holford M. 2019. Macroevolutionary analyses suggest environmental factors, not venom apparatus, play key role in Terebridae marine snail diversification. Syst. Biol. 69:413–430.

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

Mongiardino Koch N., Tilic E., Miller A.K., Stiller J., Rouse G.W. 2023. Confusion will be my epitaph: genome-scale discordance stifles phylogenetic resolution of Holothuroidea. Proc. Biol. Sci. 290:20230988.

Morel B., Williams T.A., Stamatakis A. 2023. Asteroid: a new algorithm to infer species trees from gene trees under high proportions of missing data. Bioinformatics 39:btac832.

Olivera B.M., Showers Corneli P., Watkins M., Fedosov A. 2014. Biodiversity of cone snails and other venomous marine gastropods: evolutionary success through neuropharmacology. Annu. Rev. Anim. Biosci. 2:487–513.

Osca D., Templado J., Zardoya R. 2015. Caenogastropod mitogenomics. Mol. Phylogenet. Evol. 93:118–128.

Pardos-Blas J.R., Irisarri I., Abalde S., Afonso C.M.L., Tenorio M.J., Zardoya R. 2021. The genome of the venomous snail *Lautoconus ventricosus* sheds light on the origin of conotoxin diversity. GigaScience 10:giab037.

Parins-Fukuchi C., Stull G.W., Smith S.A. 2021. Phylogenomic conflict coincides with rapid morphological innovation. Proc. Natl. Acad. Sci. U.S.A. 118:e2023058118.

Ponder W.F. 1973. The origin and evolution of Neogastropoda. Malacologia. 12:295–338.

Ponder W.F., Lindberg D.R., Ponder J.M. 2021. Chapter 3: Shell, body, and muscles. In: Biology and evolution of the mollusca. Vol. 1. Boca Raton, FL: CRC Press. p. 1–900.

Ponte G., Modica M.V. 2017. Salivary glands in predatory mollusks: evolutionary considerations. Fronti. Physiol. 8:1–8.

Prum R.O., Berv J.S., Dornburg A., Field D.J., Townsend J.P., Lemmon E.M., Lemmon A.R. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526:569–573.

Ranwez V., Douzery E.J.P., Cambon C., Chantret N., Delsuc F. 2018. MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. Mol. Biol. Evol. 35:2582–2584.

Riedel F. 2000. Ursprung und Evolution der "höheren" Caenogastropoda. Berlin. 1–265.

Rokas A., Carroll S.B. 2006. Bushes in the tree of life. PLoS Biol. 4:e352–1904.

Safavi-Hemami H., Brogan S.E., Olivera B.M. 2019. Pain therapeutics from cone snail venoms: from Ziconotide to novel non-opioid pathways. J. Proteomics 190:12–20.

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. 1:0126.

Shimodaira H., Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17:1246–1247.

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

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Strong E.E., Lipscomb D. 1999. Character coding and inapplicable data. Cladistics 15:363–371.

Sutton M., Perales-Raya C., Gilbert I. 2016. A phylogeny of fossil and living neocoleoid cephalopods. Cladistics 32:297–307.

Takezaki N., Figueroa F., Zaleska-Rutczynska Z., Takahata N., Klein J. 2004. The phylogenetic relationship of tetrapod, coelacanth, and lungfish revealed by the sequences of forty-four nuclear genes. Mol. Biol. Evol. 21:1512–1524.

Thalén F. 2018. PhyloPyPruner: tree-based orthology inference for phylogenomics with new methods for identifying and excluding contamination. Lund University Student Papers. 8963554.

Thomson R.C., Brown J.M. 2022. On the need for new measures of phylogenomic support. Syst. Biol. 71:917–920.

Uribe J.E., González V.L., Irisarri I., Kano Y., Herbert D.G., Strong E.E., Harasewych M.G. 2022. A phylogenomic backbone for gastropod molluscs. Syst. Biol. 71:1271–1280.

Uribe J.E., Zardoya R., Puillandre N. 2018. Phylogenetic relationships of the conoidean snails (Gastropoda: Caenogastropoda) based on mitochondrial genomes. Mol. Phylogenet. Evol. 127:898–906.

Vermeij G.J. 1977. The Mesozoic marine revolution: evidence from snails, predators and grazers. Paleobiology 3:245–258.

Wanninger A., Wollesen T. 2019. The evolution of molluscs. Biol. Rev. Camb. Philos. Soc. 94:102–115.

Waterhouse R.M., Seppey M., Simão F.A., Manni M., Ioannidis P., Klioutchnikov G., Kriventseva E.V., Zdobnov E.M. 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. Mol. Biol. Evol. 35:543–548.

Whitfield J.B., Lockhart P.J. 2007. Deciphering ancient rapid radiations. Trends Ecol. Evol. 22:258–265.

Xiong H., Wang D., Shao C., Yang X., Yang J., Ma T., Davis C.C., Liu L., Xi Z. 2022. Species tree estimation and the impact of gene loss following whole-genome duplication. Syst. Biol. 71:1348–1361.

Zaharias P., Pante E., Gey D., Fedosov A., Puillandre N. 2020. Data, time and money: evaluating the best compromise for inferring molecular phylogenies of non-model animal taxa. Mol. Phylogenet. Evol. 142:106660.

Zhang C., Rabiee M., Sayyari E., Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinf. 19:153.

Zou S., Li Q., Kong L. 2011. Additional gene data and increased sampling give new insights into the phylogenetic relationships of Neogastropoda, within the caenogastropod phylogenetic framework. Mol. Phylogenet. Evol. 61:425–435.