Received: 22 March 2022 | Revised: 24 May 2022 | Accepted: 13 June 2022

DOI: 10.1111/zsc.12552

ORIGINAL ARTICLE

Zoologica Scripta | WILEY

Neogastropod (Mollusca, Gastropoda) phylogeny: A step forward with mitogenomes

Thomas Lemarcis[1] | **Alexander E. Fedosov**[1,2] | **Yuri I. Kantor**[1,2] | **Jawad Abdelkrim**[3] | **Paul Zaharias**[1] | **Nicolas Puillandre**[1]

[1]Institut Systématique Evolution Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, EPHE, Sorbonne Université, Université des Antilles, Paris, France

[2]A. N. Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, Russia

[3]UAR 2700 'Acquisition et Analyse de Données pour l'Histoire naturelle', Service d'Analyse de Données, CNRS, Muséum National d'Histoire Naturelle, Sorbonne Universités, Paris Cedex, France

Correspondence
Thomas Lemarcis, Institut Systématique Evolution Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, 57 rue Cuvier, CP 51, Paris 75005, France.
Email: thomas.lemarcis@mnhn.fr

Funding information
H2020 European Research Council, Grant/Award Number: 865101; French Agence Nationale de la Recherche, France, Grant/Award Number: ANR-13-JSV7-0013-01

Abstract

The Neogastropoda (Mollusca, Gastropoda) encompass more than 15,000 described species of marine predators, including several model organisms in toxinology, embryology and physiology. However, their phylogenetic relationships remain mostly unresolved and their classification unstable. We took advantage of the many mitogenomes published in GenBank to produce a new molecular phylogeny of the neogastropods. We completed the taxon sampling by using an in-house bioinformatic pipeline to retrieve mitochondrial genes from 13 transcriptomes, corresponding to five families not represented in GenBank, for a final dataset of 113 taxa. Because mitogenomic data are prone to reconstruction artefacts, eight different evolutionary models were applied to reconstruct phylogenetic trees with IQTREE, RAxML and MrBayes. If the over-parametrization of some models produced trees with aberrant internal long branches, the global topology of the trees remained stable over models and softwares, and several relationships were revealed or found supported here for the first time. However, even if our dataset encompasses 60% of the valid families of neogastropods, some key taxa are missing and should be added in the future before proposing a revision of the classification of the neogastropods. Our study also demonstrates that even complex models struggle to satisfactorily handle the evolutionary history of mitogenomes, still leading to long-branch attractions in phylogenetic trees. Other approaches, such as reduced-genome strategies, must be envisaged to fully resolve the neogastropod phylogeny.

KEYWORDS
mitogenomes, Neogastropoda, phylogeny, systematics

1 | INTRODUCTION

Mitogenomes constitute one of the best compromises between informativeness and cost (both in terms of time and money) when it comes to selecting a suitable set of characters to resolve phylogenies (Zaharias et al., 2020). If it remains not informative enough for very deep relationships and is sometimes prone to long-branch artefacts (Schrödl

& Stöger, 2014; Uribe et al., 2019), it has been widely used in various groups of organisms, and in particular in molluscs, where it helped to clarify relationships from the species level (Abalde et al., 2017) to deeper relationships between lineages separated since several hundreds of MY (Uribe et al., 2019; Williams et al., 2014). And even in an era when datasets based on transcriptomes or reduced genomes are becoming more and more common, new mitogenomes are being produced regularly. Furthermore, in gastropods, the publication of a new mitogenome(s) provides often the opportunity to produce a new phylogeny by combining the newly produced mitogenome(s) with previously published ones (Barghi et al., 2016; Choi et al., 2021; Harasewych et al., 2019; Huang et al., 2021; Wang et al., 2021).

However, such publications generally limit the scope of the phylogeny to a few mitogenomes among those available in public databases, and the phylogenetic analyses conducted are often limited to one method, and do not test for various methods and/or models in order to more thoroughly reconstruct phylogenetic relationships. One of the groups that would benefit from such a more complete analysis are the neogastropods. This crown group of the caenogastropods is composed of more than 15,000 described species, mostly predators, and their evolutionary success has regularly been suspected to be linked to their capacity to produce various molecular compounds to capture their prey, such as anaesthetics, anticoagulants and neurotoxins (Bose et al., 2017; Modica et al., 2018; Olivera et al., 2017). However, the lack of a clear and robust phylogenetic context for the group, illustrated by their unstable superfamily and family-level classification, makes it difficult to test such hypothesis (Kuznetsova et al., 2022).

The current state-of-the-art of the neogastropod classification, as provided in WoRMS (WORMS, 2018, consulted on the 14 March 2022), is mostly based on the molecular phylogenies published in the last 10 years and the associated taxonomic revisions. Thus, the superfamilies Conoidea (Abdelkrim et al., 2018; Bouchet et al., 2011; Kantor et al., 2012; Puillandre et al., 2011; Yang et al., 2021), Buccinoidea (Couto et al., 2016; Galindo et al., 2016; Kantor et al., 2021; Oliverio & Modica, 2010), Mitroidea (Fedosov et al., 2015, 2018; Kantor et al., 2014), Olivoidea (Kantor et al., 2017), Muricoidea (Barco et al., 2010), the families Cancellariidae (Modica et al., 2011), Costellariidae (Fedosov et al., 2015, 2017) and the marginellids (Cystiscidae, Granulinidae, Marginellidae, Marginellonidae—Fedosov et al., 2019) have been found monophyletic, and the family- and/or genus-level classifications have been deeply revised. In some cases, the reconstructed phylogenetic relationships necessitated the introduction of a large number of new family-level taxa, in particular within the Conoidea and Buccinoidea, that switched from 7 and 9 families, respectively, to 17 and 20. The superfamily Turbinelloidea still awaits to be revised, although Fedosov et al. (2015, 2017) already started the work, and only a few families remain virtually untouched by molecular phylogeneticians (Volutidae, Harpidae, Babyloniidae and Strepsiduridae—but see Ravitchandirane and Sukumar (2013) for the Babyloniidae).

But the main challenge of the neogastropod phylogenetic reconstruction is to clarify the relationships at the superfamily level: three families (Harpidae, Babyloniidae and Strepsiduridae) remain unassigned in a superfamily and the relationships between the seven currently recognized superfamilies (and the three unassigned families) are virtually unknown. All the previously published phylogenetic studies that, at least partly, dealt with the neogastropod superfamily relationships failed to resolve most of these deep relationships. Two notable exceptions are (a) the recovery of the sister-taxa Tonnoidea (also revised at the family level—Strong et al., 2019) and Ficoidea, two superfamilies currently not recognized as a Neogastropoda, within the neogastropod clade (as sister to all the other neogastropods but the Cancellariidae; Colgan et al., 2007; Fourdrilis et al., 2018; Harasewych et al., 2019; Machkour-M'Rabet et al., 2021; Osca et al., 2015; Wang et al., 2017, 2021); (b) the monophyly of a group that includes all the neogastropods except the Volutoidea (Cancellariidae, marginellids, Volutidae), the Tonnoidea and the Ficoidea (Abdelkrim et al., 2018; Choi et al., 2021; Cunha et al., 2009; Fourdrilis et al., 2018; Harasewych et al., 2019; Machkour-M'Rabet et al., 2021; Osca et al., 2015; Uribe et al., 2021; Wang et al., 2017, 2021).

In order to further improve our understanding of the neogastropod phylogenetic relationships, we propose to take advantage of the many neogastropod mitogenomes available in GenBank (359 on January 28th, 2022), complemented with several mitogenomes extracted from transcriptomic data obtained from GenBank or other projects of our team (unpub.), to reconstruct the most complete phylogeny of neogastropods published to date. Even if the sampling remains incomplete, as several families of neogastropod are not represented in this mitogenome dataset, this phylogeny certainly constitutes a step forward, and highlights some previously unnoticed relationships that allow us discussing the evolution of morpho-anatomical characters and elaborating further on the evolutionary success of the neogastropods.

2 | MATERIAL AND METHODS

2.1 | Sampling

All the neogastropod, tonnoidean and ficoidean mitogenomes from GenBank were downloaded on 28 January

2022. A selection of mitogenomes of Littorinimorpha, and in particular the most closely related lineages to Neogastropoda (Osca et al., 2015) were added as outgroups. Following Osca et al. (2015), *Ifremeria nautilei* (Provannidae) was used to root all the phylogenetic trees. The family, genus and species names were updated according to WoRMS. To complement the taxonomic sampling, mitochondrial gene sequences extracted from publicly available transcriptomes or transcriptomic data produced in the framework of other projects by our team were added to the dataset. The complete list of mitogenomes from GenBank and mitochondrial gene sequences extracted from transcriptomes is provided in Appendix S1.

2.2 | Mitogenomes extracted from transcriptomic data

To complete the taxonomic sampling, 13 partial mitogenomes were extracted from transcriptomes (see Appendix S1). We designed a reference-based approach with one to four mitogenomes used as reference for each transcriptome. A suite of custom python scripts to reproduce the pipeline is available at https://github.com/Hyperdiverseproject/Neo_mitogenomes/.

First, a BLASTn (Camacho et al., 2009) search of the transcriptomes against the reference mitogenomes was performed and the contig with the best hit (e-value < 1e-10) was retained. In parallel, a BLASTp search was also performed. Assembled contigs were translated from transcriptomes using open reading frames (ORF)finder (Rombel et al., 2002), and translated genes were retrieved from reference mitogenomes directly from GenBank.

To ensure contig quality and avoid contaminations, raw reads from transcriptomes were mapped against the best hit contigs from BLASTn and BLASTp searches using Bowtie2 (Langmead & Salzberg, 2012). The depth of coverage was calculated with the option 'coverage' of samtools (Danecek et al., 2021). If two contigs from the same transcriptome aligned in the same region of the reference mitogenome, the contig with the highest depth of coverage was selected. Hit contigs were then merged and aligned against the mitogenome references with multiple alignment using Fast Fourier Transform (MAFFT) and 'localpair' and 'addfragments' options (Katoh & Frith, 2012).

A second round of filtering was performed by calculating the number of differences ('genetic distance') between the contig sequence and the sequence of the closest reference mitogenome: if the contig had genetic distance >50% of the sequence length, the contig was removed from the alignment. Remaining contigs were realigned against the reference mitogenomes and contigs not aligned with a mitochondrial gene were removed from the alignment. The pipeline resulted in 13 partial mitogenomes extracted from transcriptomes (Appendix S1).

2.3 | Subsampling and data cleaning

Single mitochondrial genes (13 protein-coding genes and ribosomal genes) were extracted from GenBank mitogenomes using AnnotationBustR (Borstein & O'Meara, 2018). Given that some lineages (species, genera) are overrepresented in the GenBank mitogenome dataset, we first performed a phylogenetic analysis (Neighbour-Joining with MEGA—Kumar et al., 2016) to confirm that conspecific and congeneric mitogenomes were clustering together, and then to select one mitogenome per genus, the one corresponding to the type species and/or with less missing data. One exception is the genus *Turricula*, for which two available (partial) mitogenomes do not cluster together: one (extracted from a transcriptome) cluster with the other Clavatulidae, the other (*Turricula nelliae spurius*, MK251986) within the Pseudomelatomidae. Given the morphological resemblance of *Turricula nelliae* with some members of Pseudomelatomidae (e.g. *Comitas*), we suspect here a misidentification. In any case, both mitogenomes were retained in the analyses. The resulting dataset comprised 114 taxa, including 10 outgroups.

Each gene was then aligned using MAFFT v7.490 (Katoh & Standley, 2013) auto mode and phylogenetic trees were reconstructed with IQ-TREE v2.1.3 (Minh et al., 2020) with General Time Reversible (GTR) + Gamma (G) in order to detect potential misalignments and contaminations. Indels in coding genes (when not a multiple of 3) were manually removed, as well as sequence fragments at the beginning or end of the alignment that were out of the gene ORF. Two gene fragments were also removed from the dataset: the first 600 nucleotides (NTs) of the cytochrome *c* oxidase subunit III (*cox3*) gene of JQ446041 (*Concholepas concholepas*) had no blast hit in GenBank; the second part of the cytochrome *c* oxidase subunit I (*cox1*) of MW316798 (*Aspa marginata*) was obviously misaligned, leading to a very long branch. We also removed the mitogenome of *Ceraesignum maximum* (Vermetidae, HM174253), because it constituted a very long branch and was not crucial for neogastropod phylogeny, as an early branching taxa within Caenogastropoda (Osca et al., 2015; Rawlings et al., 2010), thus leaving 113 taxa in the final dataset.

2.4 | Phylogenetic analyses

Sequences from coding genes were translated from NTs to amino acids (AAs) using the translateNT2AA program

implemented in MACSE (Ranwez et al., 2011) and then aligned with MAFFT using the G-INS-i algorithm. We used MAFFT E-INS-i for non-coding mitochondrially encoded 12S and 16S RNA. The program reportGapsAA2NT in Multiple Alignment of Coding SEquences Accounting (MACSE) was used to derive each NT alignment from each MAFFT AA alignment. Finally, the 15 genes were concatenated in two ways: only NTs (all 15 genes) or AAs for the 13 coding genes + 12S and 16S in NTs. We will refer to the two datasets as the NT matrix or AA + NTs matrix.

We performed a first round of analyses using IQ-TREE to determine the best fit model strategy for the NT matrix. In all cases, ModelFinder Plus (MFP) was used to select the best model of substitution (Kalyaanamoorthy et al., 2017). The dataset was partitioned in eight different ways (Table 1): a single partition for the entire NT ('single-partition + MFP'), partitioning by gene with edge-proportional ('gene-partitioned + MFP) and edge-unlinked ('gene partitioned + MFP + edge-unlinked') partition models (Chernomor et al., 2016), partitioning by codon (except for 12S and 16S) with edge-proportional ('codon partitioned + MFP') and edge-unlinked ('codon partitioned + MFP + edge-unlinked') partition models, partitioning by codon (except for 12S and 16S) with a selection of the best-fit partitioning scheme by merging partitions with edge-proportional ('codon partitioned + MFP + MERGE') and edge-unlinked ('codon partitioned + MFP + MERGE + edge-unlinked') partition models, by using the GHOST model with four classes, unlinked branch lengths, substitution rates and inferred base frequencies ('GHOST'; Crotty et al., 2019). The GHOST model was specifically designed to take into account heterotachy, that is, variation of the evolutionary rate of site through time (Lopez et al., 2002). Heterotachy is a process that is likely to have occurred in large groups and in fast-evolving genomes such as the mitochondrial genome. Each analysis was run with 1000 standard bootstraps, except for GHOST due to numerical underflow issues. Because the number of free parameters can greatly change from one model to another (Table 1), the Log-Likelihood scores alone are not appropriate to compare the different models. Instead, we ranked the Akaike information criterion (AICc) and Bayesian information criterion (BIC) scores outputted by IQ-TREE for each analysis and compared them (Table 1). The best AICc score was found for the codon partitioned + MFP analysis while the best BIC score was found for the codon partitioned + MFP + MERGE. Since there was no strong argument to support one or the other, we decided to run RAxML-ng and MrBayes on the NT matrix using both the codon partitioned + MFP and the codon partitioned + MFP + MERGE.

The AA matrix data were run with the same strategy for IQ-TREE, RAxML and MrBayes; a gene partitioned approach was used and ModelFinder identified the best fit model for each partition. Unfortunately, the AA matrix did not work with 12S and 16S in IQ-TREE due to a software error mentioning running out of RAM (and reported to the IQ-TREE google group). For both the NT and AA matrix, when the model of substitution selected by ModelFinder did not exist in either RAxML or MrBayes, we converted the model into the equivalent model available in each respective software.

RAxML was run using the 'all-in-one' analysis flag implying the use of bootstopping criterion (Pattengale et al., 2010). The bayesian analyses run in MrBayes v3.2 (Ronquist et al., 2012) consisted of three parallel analyses, each with eight Markov chains of 50,000,000 generations and a sampling frequency of one tree each 10,000 generations. The number of swaps was set to 5, and the chain temperature at 0.02. We evaluated the convergence of each analyses using Tracer v1.7.1 (Rambaut et al., 2018).

TABLE 1 Log-likelihood, number of free parameters, AICc and BIC scores for the IQ-tree analyses performed with eight different partitions and models

IQ-TREE model	Log-likelihood	Nb free parameters	AICc	BIC
Single partition + MFP	−664,013.774	249	1,328,534.31	1,330,412.72
Gene partitioned + MFP	−660,006.477	540	1,321,134.938	1,325,185.616
Gene partitioned + MFP + edge-unlinked	−658,844.169	3316	1,326,294.878	1,349,452.313
Codon partitioned + MFP	−638,265.299	817	**1,278,262.591**	1,284,356.644
Codon partitioned + MFP + edge-unlinked	−632,616.015	9102	1,314,381.131	1,352,420.118
Codon partitioned + MFP + MERGE	−639,582.443	510	1,280,222.257	**1,284,050.178**
Codon partitioned + MFP + MERGE + edge-unlinked	−639,462.959	1219	1,281,588.603	1,290,602.724
GHOST	−790,917.6179	927	1,583,816.399	1,590,714.972

The best AICc and BIC scores are in bold.

Abbreviations: AICc, Akaike information criterion; BIC, Bayesian information criterion; MFP, ModelFinder Plus.

Analyses were stopped before reaching 50,000,000 generations to limit computation time, but not before the ESS values were all superior to 200, thus leading to a burnin of 24%, 43% and 4% for the NT codon partitioned + MFP, NT codon partitioned +MFP + MERGE and the AA matrix, respectively.

3 | RESULTS

The final dataset included sequences of 10 outgroups, 11 Tonnoidea (5 families), 1 Ficoidea (1 family) and 91 neogastropod (36 families) representatives. Among them, incomplete mitogenomes (between 8 and 13 genes each—Genbank accession numbers: see Appendix S1) were recovered from transcriptomes for 13 taxa, representing 10 families, including 5 (Colubrariidae, Olividae, Personidae, Pisaniidae, Turbinellidae) that were not represented in GenBank. However, 26 families of neogastropods and 4 families of Tonnoideans considered as valid in WoRMS are still not represented in our dataset.

Eight trees were obtained and will be compared (Figures 1 and S1): NT codon partitioned + MFP with IQ-TREE, RAxML and MrBayes, NT codon partitioned + MFP + MERGE with IQ-TREE, RAxML and MrBayes and AA with RAxML and MrBayes. The trees obtained with the same matrix, method or partition model are almost identical, but more differences (with only a few supported nodes—Bootstraps B < 80, Posterior Probabilities [PP] < 0.95) are found when comparing trees obtained with different matrices, methods and/or substitution models. The two NT IQTREE are identical, except for the position of Pisaniidae and Cominellidae (unsupported in both cases) within the Buccinoidea. Similarly, the two NT RAxML trees are identical, with the only difference being the Tonnoidea + Ficoidea clade sister to the Calyptraeidae + Neogastropoda clade in the codon + MFP + Merge tree whereas the Calyptraeidae are sister to the Cancellariidae in the codon + MFP tree (as in the six other trees), but these nodes are not supported. Except a few other unsupported differences, the RAxML trees are also very similar to the IQTREE trees, and all of them are also similar to the MrBayes trees. In general, the support values for the IQTREE and RAxML trees are lower than for the MrBayes trees. In the following section, the description of the results will focus on the MrBayes trees, and the IQTREE and RAxML trees (Figures 1 and S1) will be discussed only to point at the differences (or similarities) with the MrBayes trees.

In all trees, all the families represented by several taxa are monophyletic, except the Turridae (the genera *Gemmuloborsonia* and *Lucerapex* are sister to the Pseudomelatomidae + Drilliidae clade in the MrBayes AA tree with PP = 0.99, thus not clustering with the rest of the Turridae, as for example, in the MrBayes NT trees with B = 1) and the Tudiclidae (non-monophyly supported in most trees). Within families, the relationships are very consistent among trees. The relationships are less stable within superfamilies, with some unstable taxa such as the Clathurellidae and Mitromorphidae within the Conoidea, and several families within the Buccinoidea (but again, the corresponding nodes are most often not supported). For example, the Columbellidae are sister to the rest of the Buccinoidea in the NT trees (not supported), but sister to the Colubrariidae in the AA trees (PP = 0.99, B = 52).

The Buccinoidea, Conoidea, Tonnoidea, Olivoidea, Muricoidea, Mitroidea, (the latter two being represented by only one family each) are always found monophyletic, with high support. However, the superfamily Turbinelloidea, although represented by two taxa only, is never found monophyletic, with *Vasum* (Turbinellidae) included with high support in a clade together with Conoidea, Mitroidea, Buccinoidea, Olivoidea and Babyloniidae, thus excluding Costellariidae. The position of *Vasum* is not stable among trees: it is sister to the Buccinoidea (e.g. in the MrBayes NT codon + MFP tree − PP = 0.87) or sister to the Babyloniidae (e.g. in the MrBayes AA tree − PP = 1). Similarly, the Volutoidea are never found monophyletic, with the Volutidae being sister to a clade including Conoidea, Buccinoidea, Olivoidea, Turbinelloidea, Muricoidea and Babyloniidae, and the Cancellariidae being sister to the Calyptraeidae with high support in most trees.

In all trees, the clade Conoidea + Mitroidea + Buccinoidea + Olivoidea + Babyloniidae + Turbinellidae is always supported. The Muricidae, Costellariidae, Volutidae and Tonnoidea + Ficoidea are then successively branching as sister clades to the previous one, always with good support with both the AA and NT matrices (0.98 < PP < 1). The only exception is in the RAxML codon + MFP + Merge tree, as detailed above. Finally, the Cancellariidae are always sister to the Calyptraeidae with high support (0.96 < PP < 1), and all the neogastropods (except Cancellariidae) + Tonnoidea/Ficoidea are monophyletic (PP = 1).

4 | DISCUSSION

4.1 | A note on model selection for the NT matrix

While selecting a model of sequence evolution is facilitated by using model-selection methods (e.g. ModelFinder), choosing the right model-selection strategy on complex datasets such as mitochondrial genomes remains challenging. The use of information-theoretic

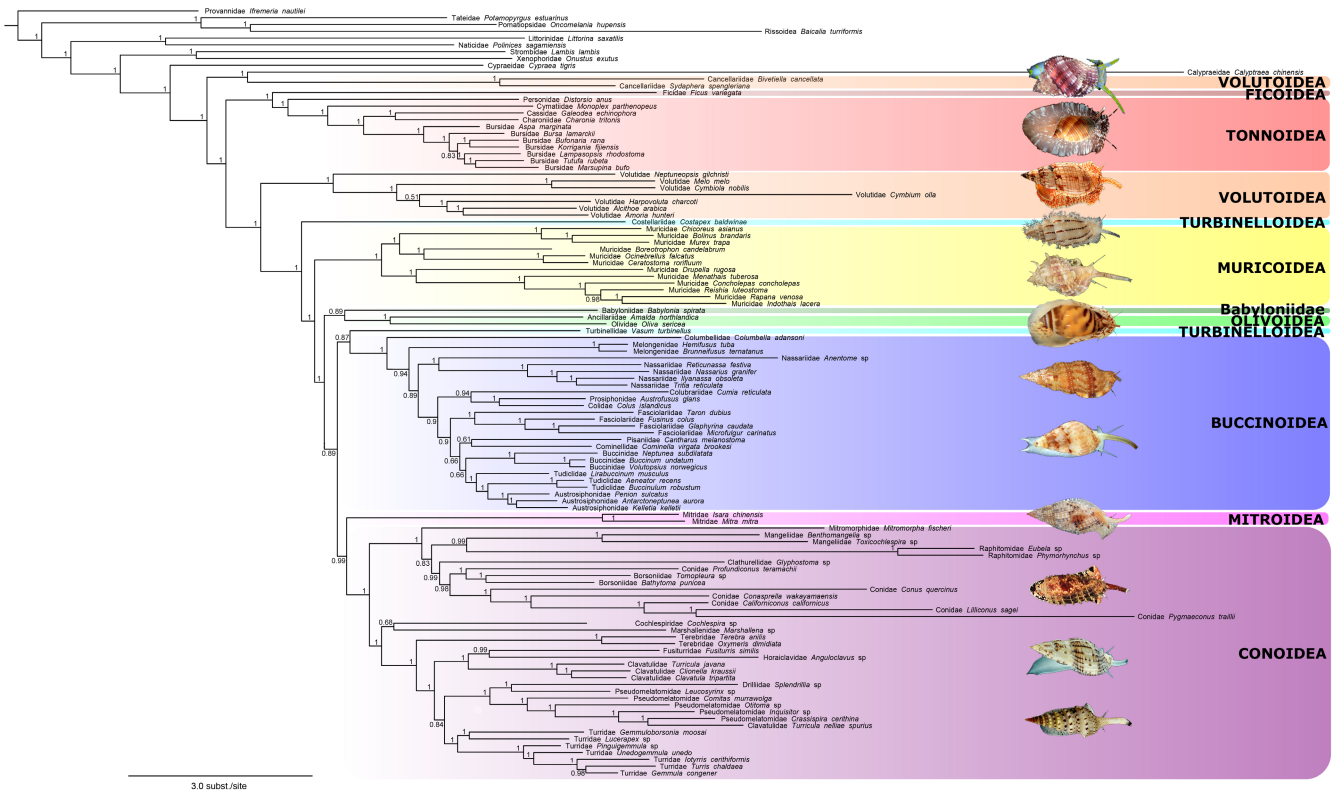
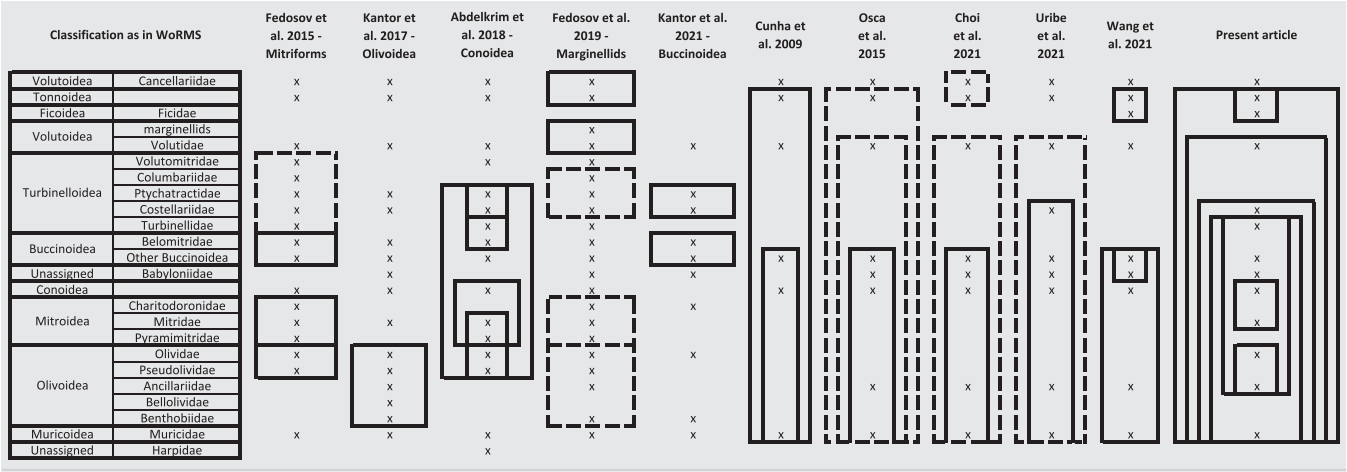


FIGURE 1 Phylogenetic tree obtained with the NT 'codon + MFP' matrix, analysed with MrBayes. The bootstraps values and posterior probabilities (>0.5) are given for each node. Superfamilies are highlighted with colour boxes. Illustrations from top to down: Cancellariidae, Tonnidae, Volutidae, Costellariidae, Muricidae, Olividae, Colubrariidae, Nassariidae, Mitridae, Conidae, Raphitomidae, Turridae *(*Credits: Philippe Maestrati, Laurent Charles /MNHN). MFP, ModelFinder Plus; NT, nucleotide

methods (Posada & Buckley, 2004) such as corrected AICc and BIC is now common practice to distinguish and select several phylogenetic models. In the first phase of our phylogenetic analyses using IQ-TREE, we used eight different strategies for analyzing the NT matrix and rank each phylogenetic model using either AICc or BIC criteria (Table 1). The implementation of the GHOST model has led to the lowest supports and topologies highly different from the others, possibly due to repeated numerical underflow errors. But GHOST is also a mixture model and some recent discussions (Crotty & Holland, 2022) questioned the use of information-theory methods to compare partition and mixture models. Similarly, models involving edge-unlinked partitions have low AICc or BIC scores, likely due to the great amount (several thousands) of parameters to evaluate. Interestingly enough, GHOST and edge-unlinked models are the only two strategies that can take heterotachy (Lopez et al., 2002) into account. Our results stress out the need for alternative heterotachy models that are less parameter-rich and more suited for datasets with high number of taxa. Obviously, using a single partition over the entire NT matrix is unrealistic, even using a FreeRate model (Le et al., 2012) with 10 categories to take more site heterogeneity into account. Thus, our options for model selection were reduced to a medium-level parametrized gene-partition strategy, a highly parametrized codon-partition strategy or an alternate medium-level parametrized strategy where codon-partitions are merged using the MERGE option in IQ-TREE. The 'codon partitioned + MF' had the best AICc score, while the 'codon partitioned + MF + MERGE' strategy had the best BIC score. Hence, we decided to keep both for subsequent analysis with RAxML and MrBayes. While the two codon-partitioned strategies were the ones favoured by information-theory criteria, we must note that all trees inferred with this partition schemes reveal unrealistic branch lengths, that is, with some internal branches having on average more than seven substitutions per site, regardless of the reconstruction method used (i.e. RAxML, IQ-TREE or MrBayes). We suspect that for some partitions, that is, the ones with high-evolving rate sites such as third codon positions, the methods fail to correctly assess the substitution rates. A careful look at some partition parameters in MrBayes show that a lot of them are not converging, even at the end of the analyses. In spite of the obviously overparametrized model that was used and unrealistic branch lengths, the topologies (which

TABLE 2 Comparison of neogastropod molecular phylogenies



A 'x' means that the taxon is represented in the phylogeny; if several samples of this taxon were included, it also means that the taxon has been recovered monophyletic. Full rectangles correspond to supported clades (bootstraps > 80 and/or posterior probabilities > 0.95); dashed rectangles correspond to unsupported clades. The five firsts phylogenies were based on few mitochondrial and nuclear genes (typically: cox-1, 16S, 12S, 28S), and for each of them the focus taxon is indicated; the five others were based on mitogenomes; the last one corresponds to the phylogeny on the Figure 1. Superfamilies for which the monophyly is supported in several phylogenies (Conoidea, Tonnoidea, Buccinoidea except Belomitridae) are not detailed at the family level

are the prime interest of our study) remain stable across analyses or within the tree space explored by MrBayes.

It is now generally accepted that over-parameterization should be favoured over under-parameterization (Abadi et al., 2019; Fabreti & Höhna, 2022), and studies with similar datasets (i.e. mitochondrial genomes spanning a large taxonomic diversity) have also faced the same issues of unrealistic branch lengths but reliable topologies (e.g. Song et al., 2016; Uribe & Zardoya, 2017). Our eight NT matrix topologies (Figures 1 and S1) show few very important differences (i.e. apart from already known unstable regions of the tree) but over-partitioned datasets and overly complex models can return unrealistic branch lengths, even when AICc or BIC scores are good. Caution must be used before running any kind of meta-phylogeny analysis (e.g. dating or diversification analyses) that will use branch length information and we recommend carefully checking the branch lengths to eventually detect unrealistic branch lengths.

4.2 | Input for neogastropod classification

With 42 families of neogastropods and tonnoideans, representing almost 60% of the families currently considered as valid in WoRMS, our dataset is the most complete published so far. Many of the previously recovered relationships are recovered here and new ones are revealed (Table 2).

At the family level, all the families represented by at least two species are recovered monophyletic, except two. The first exception is the Tudiclidae (Buccinoidea), with *Lirabuccinum musculus* not clustering with the two others (*Aeneator recens* and *Buccinulum robustum*). Kantor et al. (2021) already pointed at the radula of *Lirabuccinum musculus* being different from the other Tudiclidae species, suggesting that it might not be a Tudiclidae. The second exception is the Turridae (Conoidea), with *Lucerapex* sp. and *Gemmuloborsonia moosai* not clustering with the others in most trees. This non-monophyly (supported only in the MrBayes AA tree) was already found in Uribe et al. (2018), also based on mitogenomes, but has been contradicted by the exon-based phylogeny of Abdelkrim et al. (2018), with the Turridae being monophyletic and highly supported.

At the superfamily level, some superfamilies are recovered monophyletic and supported, often confirming previously published phylogenies, although based on a much more reduced sampling (Choi et al., 2021; Osca et al., 2015; Uribe et al., 2021; Wang et al., 2021). For example, the Conoidea, represented by many mitogenomes covering almost all the family-level diversity of the group, is found monophyletic. However, some unsupported relationships between families are probably resulting from long-branch attraction (LBA), as suggested in Uribe et al. (2018), and are contradicted both by exon-based phylogenies and anatomical characters (Abdelkrim et al., 2018). For example, the Cochlespiridae, found here closely related to the Marshallenidae in several trees, is more probably sister to all the other Conoidea. The Olivoidea, here for the first time in a mitogenome phylogeny represented by several species, are found monophyletic, confirming the results obtained previously (Kantor et al., 2017). Similarly, the Buccinoidea, recently revised in Kantor et al. (2021), are found monophyletic. However, the Belomitridae, not

included here, might not be closely related to the other Buccinoidea, as suggested by Abdelkrim et al. (2018) and Fedosov et al. (2019). Furthermore, the relationships between the Buccinoidea families are not supported, and sometimes contradict previously published results: for example, the Columbellidae are sister to the Colubrariidae in the MrBayes AA tree with high support, whereas they were found embedded within the Nassariidae in Kantor et al. (2021). The Mitroidea, represented here by a single family (Mitridae), also includes the Pyramimitridae and Charitodoronidae, and these three families were constituting a clade in previously published phylogenies (Fedosov et al., 2018). The Turbinelloidea, represented here by two families (Costellariidae and Turbinellidae), are not monophyletic. The other Turbinelloidea families are sometimes represented by one or a few samples in previously published phylogenies, but their close relationship is never supported, and this superfamily clearly constitutes a potentially non-monophyletic taxon. The Volutoidea, represented here by the Volutidae and Cancellariidae, are not monophyletic. The family Volutidae is sister to all the other neogastropods (except Cancellariidae), and was found to form a clade with the marginellids (Cystiscidae, Granulinidae, Marginellidae and Marginellonidae) in Fedosov et al. (2019). However, as illustrated in our tree, the Cancellariidae never cluster with the rest of the Volutoidea (except in Wang et al. (2021), although not highly supported), and is either found to be sister to all the other Neogastropoda + Tonnoidea/Ficoidea (Cunha et al., 2009; Osca et al., 2015) or to Tonnoidea (Choi et al., 2021; Fedosov et al., 2019). Here, it is generally sister to *Calyptraea chinensis* (Calyptraeidae), a taxon not included in previously published neogastropod phylogenies. Morpho-anatomical data do not support this relationship (Simone, 2002), and the long branch leading to *Calyptraea chinensis* would suggest the occurrence of a phenomenon of LBA. It is important to note that the Cancellariidae have been placed within the Volutoidea only recently (Bouchet et al., 2017), while previously it belonged to its own superfamily Cancellarioidea. Thus, the position of the Cancellariidae, as sister to Calyptraeidae or to the Neogastropoda/Tonnoidea/Ficoidea clade remains to be determined. Finally, the Tonnoidea are found monophyletic and sister to the Ficoidea, a result already supported in Strong et al. (2019) and Wang et al. (2021).

Several supported relationships between the superfamilies (and with the Babyloniidae) are here recovered. The Conoidea and Mitroidea (represented only by the Mitridae) form a clade, as suggested also in Abdelkrim et al. (2018). While the relationships between the clade Conoidea + Mitroidea with Buccinoidea, Turbinellidae, Olivoidea and Babyloniidae are not resolved, they all form a well-supported clade. The deeper relationships are also always supported, with the respective position of the Muricoidea, Costellariidae, Volutidae and Tonnoidea + Ficidae always branching in the same order in our trees, with high support. Nevertheless, a number of phylogenetically important neogastropod taxa remain unrepresented in our dataset (Volutomitridae, Harpidae), or underrepresented (Cancellariidae, Turbinelloidea, Olivoidea). Inclusion of these currently missing lineages may lead to the modification of the tree topology, and therefore, we consider the current tree still too preliminary to induce revisions in systematics or to redefine apomorphies of the currently revealed clades. Further increase in the taxonomic sampling to eventually include all crucial neogastropod lineages will be vital to build a solid basis for the order's reclassification.

4.3 | Future priorities

In future studies, priority should be given to completing the taxon sampling. Among the missing families, some were previously confidently placed in superfamilies in published molecular phylogenies, and do not constitute priority targets: Laubierinidae, Ranellidae, Thalassocyonidae and Tonnidae in Tonnoidea (Strong et al., 2019), Buccinanopsidae, Busyconidae, Chauvetiidae, Eosiphonidae, Prodotiidae and Retimohniidae in Buccinoidea (Kantor et al., 2021), Bouchetispiridae and Conorbidae in Conoidea (Abdelkrim et al., 2018; Kantor et al., 2012), Charitodoronidae and Pyramimitridae in Mitroidea (Fedosov et al., 2015) and Bellolividae, Benthobiidae and Pseudolividae in Olivoidea (Kantor et al., 2017). However, the addition of these families in a phylogeny may improve the overall quality of the tree, either by strengthening the support for the monophyly of the corresponding superfamilies or by clarifying the relationships between the superfamilies. However, priority should be given to those families that have never, or rarely, been included in molecular phylogenies, and whose superfamily membership remain dubious. Within the Buccinoidea, the Belomitridae and Dolicholatiridae have been shown to be sister to the rest of the Buccinoidea in Kantor et al. (2021), but the Belomitridae were not recovered as sister to the other Buccinoidea in Abdelkrim et al. (2018). The four families of marginellids (Cystiscidae, Granulinidae, Marginellidae and Marginellonidae) have been shown to be monophyletic and most probably sister to the Volutidae (thus forming the Volutoidea, excepted Cancellariidae) in Fedosov et al. (2019); however, this remains to be confirmed in a large-scale phylogeny of neogastropods. The likely non-monophyletic Turbinelloidea probably constitutes the main gap in the neogastropod sampling, since several

published molecular phylogenies (including the present one) do not support its monophyly, although three out of five families (Columbariidae, Ptychatractidae, Volutomitridae) have almost never been included in a molecular phylogeny targeting the Neogastropoda. The only exception is Fedosov et al. (2015), in which these five lineages were forming a clade, although not supported. Finally, two unassigned neogastropod families, Harpidae and Strepsiduridae, remain to be included in a molecular phylogeny, and also constitute priorities for future studies.

Furthermore, the monophyly of most, if not all, neogastropod families remain to be tested by including more representatives in each of them, and it is not impossible that some of them would not cluster with the type-genera of the corresponding family. On top of that, and this is true for the whole neogastropods, sequencing more genera, and even more species within each genus, may reveal family-level taxa that remained undetected so far, that is happening regularly when (super)family-level classifications are revised with molecular data (e.g. Buccinoidea – (Kantor et al., 2021); Conoidea – (Abdelkrim et al., 2018)). Thus, as a general rule, sampling effort should first focus on type taxa, in order to ascertain the link with available family-level names, and second on genus or species-level lineages whose family membership is dubious, either because of divergent DNA sequences (typically: *cox-1*) or peculiar morpho-anatomical characters. It should be noted that some of these taxa are rare and difficult to collect, especially if the type taxa are targeted. Given that sampling in the field becomes more and more difficult, because of the legislative barriers reinforced in many countries, museum collections certainly represent the most promising source of material to complete the sampling. However, sequencing DNA from not freshly collected material is more challenging, and requires adequate sequencing strategies (Raxworthy & Smith, 2021).

Indeed, the second priority to improve the phylogeny of neogastropods is to develop next-generation sequencing (NGS)-based methods to recover genetic information of sufficient quality and quantity to resolve the deeper nodes of the phylogeny. As discussed before, sequencing mitogenomes constitutes a good compromise in terms of time and money, but it is prone to artefacts (LBA) and might not be resolutive enough for the deeper nodes of the neogastropod phylogeny. Reduced-genome (UCE, Exon-capture) or transcriptome-based phylogenies have recently shown a high potential in resolving deep nodes in molluscan phylogenies, and are of course promising for neogastropod. If transcriptomic data require fresh samples, exon-capture approaches, by targeting short exons spread over the whole genome, can cope with degraded DNA (e.g. Abdelkrim et al., 2018; Moles & Giribet, 2021). It thus constitutes, in our opinion, the preferential strategy that must be applied to resolve the neogastropod phylogeny.

ACKNOWLEDGEMENTS

The material used to produce the new transcriptomes was collected during the KANACONO expedition in New Caledonia (convention MNHN-Province Sud, APA_NCPS_2016_012; PI N. Puillandre and S. Samadi), with support from the Laboratoire d'Excellence Diversités Biologiques et Culturelles (LabEx BCDiv, ANR-10-LABX-0003-BCDiv) and the project CONOTAX, funded by the French Agence Nationale de la Recherche, France (ANR-13-JSV7-0013-01). This expedition operated under the regulations then in force in the countries in question and satisfy the conditions set by the Nagoya Protocol for access to genetic resources. Additional samples were collected during the expedition to Central Vietnam, supported by the Coastal branch of Russian-Vietnamese Tropical Research and Technology Center (Nha-Trang). The authors are thankful to the staff of the Tropical Center for assistance in the field sampling and for lending laboratory equipment. We thank Laetitia Aznar-Cormano for her help with the RNA extraction and Juliette Gorson and Mandë Holford for their help in the RNA sequencing. The analyses were performed on the Plateforme de Calcul Intensif et Algorithmique PCIA (UAR2700 2AD, MNHN). The present study was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 865101) to N.P.

ORCID

Thomas Lemarcis https://orcid.org/0000-0002-7099-1984
Alexander E. Fedosov https://orcid.org/0000-0002-8035-1403
Yuri I. Kantor https://orcid.org/0000-0002-3209-4940
Jawad Abdelkrim https://orcid.org/0000-0001-7996-8660
Paul Zaharias https://orcid.org/0000-0003-3550-2636
Nicolas Puillandre https://orcid.org/0000-0002-9797-0892

REFERENCES

Abadi, S., Azouri, D., Pupko, T., & Mayrose, I. (2019). Model selection may not be a mandatory step for phylogeny reconstruction. *Nature Communications*, *10*(1), 934. https://doi.org/10.1038/s41467-019-08822-w

Abalde, S., Tenorio, M. J., Afonso, C. M. L., & Zardoya, R. (2017). Mitogenomic phylogeny of cone snails endemic to Senegal. *Molecular Phylogenetics and Evolution*, *112*, 79–87. https://doi.org/10.1016/j.ympev.2017.04.020

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). *Molecular Biology and Evolution*, *35*, 2355–2374.

Barco, A., Claremont, M., Reid, D. G., Houart, R., Bouchet, P., Williams, S. T., Cruaud, C., Couloux, A., & Oliverio, M. (2010). A molecular phylogenetic framework for the Muricidae, a diverse family of carnivorous gastropods. *Molecular Phylogenetics and Evolution*, *56*(3), 1025–1039.

Barghi, N., Concepcion, G. P., Olivera, B. M., & Lluisma, A. O. (2016). Characterization of the complete mitochondrial genome of *Conus tribblei* Walls, 1977. *Mitochondrial DNA Part A*, *27*(6), 4451–4452. https://doi.org/10.3109/19401736.2015.1089566

Borstein, S. R., & O'Meara, B. C. (2018). *AnnotationBustR*: An R package to extract subsequences from GenBank annotations. *PeerJ*, *6*, e5179. https://doi.org/10.7717/peerj.5179

Bose, U., Wang, T., Zhao, M., Motti, C., Hall, M., & Cummins, S. (2017). Multiomics analysis of the giant triton snail salivary gland, a crown-of-thorns starfish predator. *Scientific Reports*, *7*(1), 6000.

Bouchet, P., Kantor, Y., Sysoev, A., & Puillandre, N. (2011). A new operational classification of the Conoidea (Gastropoda). *Journal of Molluscan Studies*, *77*, 273–308.

Bouchet, P., Rocroi, J.-P., Hausdorf, B., Kaim, A., Kano, Y., Nützel, A., Parkhaev, P., Schroedl, M., & Strong, E. E. (2017). Revised classification, nomenclator and typification of gastropod and Monoplacophoran Families. *Malacologia*, *61*(1–2), 1–526. https://doi.org/10.4002/040.061.0201

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Chernomor, O., von Haeseler, A., & Minh, B. Q. (2016). Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices. *Systematic Biology*, *65*(6), 997–1008. https://doi.org/10.1093/sysbio/syw037

Choi, E. H., Choi, N. R., & Hwang, U. W. (2021). The mitochondrial genome of an Endangered freshwater snail *Koreoleptoxis nodifila* (Caenogastropoda: Semisulcospiridae) from South Korea. *Mitochondrial DNA Part B*, *6*(3), 1120–1123. https://doi.org/10.1080/23802359.2021.1901626

Colgan, D. J., Ponder, W. F., Beacham, E., & Macaranas, J. (2007). Molecular phylogenetics of Caenogastropoda (Gastropoda: Mollusca). *Molecular Phylogenetics and Evolution*, *42*, 717–737.

Couto, D. R., Bouchet, P., Kantor, Y. I., Simone, L. R. L., & Giribet, G. (2016). A multilocus molecular phylogeny of Fasciolariidae (Neogastropoda: Buccinoidea). *Molecular Phylogenetics and Evolution*, *99*, 309–322. https://doi.org/10.1016/j.ympev.2016.03.025

Crotty, S. M., & Holland, B. R. (2022). Comparing partitioned models to mixture models: Do information criteria apply? *Systematic Biology*, syac003. https://doi.org/10.1093/sysbio/syac003

Crotty, S. M., Minh, B. Q., Bean, N. G., Holland, B. R., Tuke, J., Jermiin, L. S., & Haeseler, A. V. (2019). GHOST: Recovering historical signal from heterotachously evolved sequence alignments. *Systematic Biology*, *69*, 249–264. https://doi.org/10.1093/sysbio/syz051

Cunha, R. L., Grande, C., & Zardoya, R. (2009). Neogastropod phylogenetic relationships based on entire mitochondrial genomes. *BMC Evolutionary Biology*, *9*, 210.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., SA, M. C., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), giab008. https://doi.org/10.1093/gigascience/giab008

Fabreti, L. G., & Höhna, S. (2022). Bayesian inference of phylogeny is robust to substitution model over-parameterization. *Evolutionary Biology*. https://doi.org/10.1101/2022.02.17.480861

Fedosov, A., Puillandre, N., Herrmann, M., Kantor, Y., Oliverio, M., Dgebuadze, P., Modica, M. V., & Bouchet, P. (2018). The collapse of Mitra: Molecular systematics and morphology of the Mitridae (Gastropoda: Neogastropoda). *Zoological Journal of the Linnean Society*, *183*(2), 253–337.

Fedosov, A., Puillandre, N., Kantor, Y., & Bouchet, P. (2015). Phylogeny and systematics of mitriform gastropods (Mollusca: Gastropoda: Neogastropoda). *Zoological Journal of the Linnean Society*, *175*(2), 336–359.

Fedosov, A. E., Caballer Gutierrez, M., Buge, B., Sorokin, P. V., Puillandre, N., & Bouchet, P. (2019). Mapping the missing branch on the neogastropod tree of life: Molecular phylogeny of marginelliform gastropods. *Journal of Molluscan Studies*, *85*, 440–452.

Fedosov, A. E., Puillandre, N., Herrmann, M., Dgebuadze, P., & Bouchet, P. (2017). Phylogeny, systematics, and evolution of the family Costellariidae (Gastropoda: Neogastropoda). *Zoological Journal of the Linnean Society*, *179*(3), 541–626.

Fourdrilis, S., de Frias Martins, A. M., & Backeljau, T. (2018). Relation between mitochondrial DNA hyperdiversity, mutation rate and mitochondrial genome evolution in Melarhaphe neritoides (Gastropoda: Littorinidae) and other Caenogastropoda. *Scientific Reports*, *8*(1), 17964. https://doi.org/10.1038/s41598-018-36428-7

Galindo, L. A., Puillandre, N., Utge, J., Lozouet, P., & Bouchet, P. (2016). The phylogeny and systematics of the Nassariidae revisited (Gastropoda, Buccinoidea). *Molecular Phylogenetics and Evolution*, *99*, 337–353.

Harasewych, M. G., Sei, M., Wirshing, H. H., & Uribe, J. E. (2019). The complete mitochondrial genome of Neptuneopsis gilchristi G.B. Sowerby III, 1898 (Neogastropoda: Volutidae: Calliotectinae). *The Nautilus*, *133*, 7.

Huang, H., Yang, H., Li, J., Guo, B., & Ye, Y. (2021). *Complete mitochondrial genomes of Babylonia formosae and Babylonia zeylanica (Neogastropoda: Babyloniidae) and increased sampling give new insights into neogastropoda phylogeny*. In Review. https://doi.org/10.21203/rs.3.rs-948577/v1

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589.

Kantor, Y., Lozouet, P., Puillandre, N., & Bouchet, P. (2014). Lost and found: The Eocene family Pyramimitridae (Neogastropoda) discovered in the Recent fauna of the Indo-Pacific. *Zootaxa*, *3754*(3), 239–276. https://doi.org/10.11646/zootaxa.3754.3.2

Kantor, Y. I., Fedosov, A., Puillandre, N., Bonillo, C., & Bouchet, P. (2017). Returning to the roots: Morphology, molecular phylogeny and classification of the Olivoidea (Gastropoda: Neogastropoda). *Zoological Journal of the Linnean Society*, *180*(3), 493–541.

Kantor, Y. I., Strong, E. E., & Puillandre, N. (2012). A new lineage of Conoidea (Gastropoda: Neogastropoda) revealed by morphological and molecular data. *Journal of Molluscan Studies*, *78*, 246–255.

Kantor, Y. I., Fedosov, A. E., Kosyan, A. R., Puillandre, N., Sorokin, P. A., Kano, Y., Clark, R. N., & Bouchet, P. (2021). Molecular phylogeny and revised classification of the Buccinoidea (Neogastropoda). *Zoological Journal of the Linnean Society*, *194*, 789–857. https://doi.org/10.1093/zoolinnean/zlab031

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Katoh, K., & Frith, M. C. (2012). Adding unaligned sequences into an existing alignment using MAFFT and LAST. *Bioinformatics*, *28*(23), 3144–3146. https://doi.org/10.1093/bioinformatics/bts578

Kumar, S., Stecher, G., & Tamura, K. (2016). mega7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://doi.org/10.1093/molbev/msw054

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: A novel class of conotoxin-like venom peptides from predatory gastropods of the genus* Vexillum [Preprint]. *Biochemistry*. https://doi.org/10.1101/2022.01.15.476460

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Le, S. Q., Dang, C. C., & Gascuel, O. (2012). Modeling protein evolution with several amino acid replacement matrices depending on site rates. *Molecular Biology and Evolution*, *29*(10), 2921–2936. https://doi.org/10.1093/molbev/mss112

Lopez, P., Casane, D., & Philippe, H. (2002). Heterotachy, an important process of protein evolution. *Molecular Biology and Evolution*, *19*(1), 1–7. https://doi.org/10.1093/oxfordjournals.molbev.a003973

Machkour-M'Rabet, S., Hanes, M. M., Martínez-Noguez, J. J., Cruz-Medina, J., & García-De León, F. J. (2021). The queen conch mitogenome: Intra- and interspecific mitogenomic variability in Strombidae and phylogenetic considerations within the Hypsogastropoda. *Scientific Reports*, *11*(1), 11972. https://doi.org/10.1038/s41598-021-91224-0

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. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Modica, M. V., Bouchet, P., Cruaud, C., Utge, J., & Oliverio, M. (2011). Molecular phylogeny of the nutmeg shells (Neogastropoda, Cancellariidae). *Molecular Phylogenetics and Evolution*, *59*, 685–697.

Modica, M. V., Sánchez, J. R., Pasquadibisceglie, A., Oliverio, M., Mariottini, P., & Cervelli, M. (2018). Anti-haemostatic compounds from the vampire snail Cumia reticulata: Molecular cloning and in-silico structure-function analysis. *Computational Biology and Chemistry.*, *75*, 168–177.

Moles, J., & Giribet, G. (2021). A polyvalent and universal tool for genomic studies in gastropod molluscs (Heterobranchia). *Molecular Phylogenetics and Evolution*, *155*, 106996. https://doi.org/10.1016/j.ympev.2020.106996

Olivera, B. M., Fedosov, A., Imperial, J. S., & Kantor, Y. (2017). Physiology of envenomation by conoidean gastropods. In *Physiology of Molluscs: A Collection of Selected Reviews* (pp. 153–188). Apple Academic Press.

Oliverio, M., & Modica, M. V. (2010). Relationships of the haematophagous marine snail Colubraria (Rachiglossa: Colubrariidae), within the neogastropod phylogenetic framework. *Zoological Journal of the Linnean Society*, *158*, 779–800.

Osca, D., Templado, J., & Zardoya, R. (2015). Caenogastropod mitogenomics. *Molecular Phylogenetics and Evolution*, *93*, 118–128.

Pattengale, N. D., Alipour, M., Bininda-Emonds, O. R., Moret, B. M., & Stamatakis, A. (2010). How many bootstrap replicates are necessary? *Journal of Computational Biology*, *17*(3), 337–354.

Posada, D., & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and bayesian approaches over likelihood ratio tests. *Systematic Biology*, *53*, 793–808.

Puillandre, N., Kantor, Y. I., Sysoev, A. V., Couloux, A., Meyer, C. P., Rawlings, T. A., Todd, J. A., & Bouchet, P. (2011). The dragon tamed? A molecular phylogeny of the Conoidea (Gastropoda). *Journal of Molluscan Studies*, *77*(3), 259–272. https://doi.org/10.1093/mollus/eyr015

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904.

Ranwez, V., Harispe, S., Delsuc, F., & Douzery, E. J. P. (2011). MACSE: Multiple Alignment of Coding SEquences Accounting for Frameshifts and Stop Codons. *PLoS One*, *6*(9), e22594. https://doi.org/10.1371/journal.pone.0022594

Ravitchandirane, V., Sukumar, V., & Thangaraj, M. (2013). Phylogenetic status of babylonia Zeylanica (family Babyloniidae) based on 18S rRNA gene Fragment. *Annales of West University of Timisoara. Series of Biology*, *16*(2), 135.

Rawlings, T. A., MacInnis, M. J., Bieler, R., Boore, J. L., & Collins, T. M. (2010). Sessile snails, dynamic genomes: Gene rearrangements within the mitochondrial genome of a family of caenogastropod molluscs. *BMC Genomics*, *11*(1), 440. https://doi.org/10.1186/1471-2164-11-440

Raxworthy, C. J., & Smith, B. T. (2021). Mining museums for historical DNA: Advances and challenges in museomics. *Trends in Ecology & Evolution*, *36*(11), 1049–1060. https://doi.org/10.1016/j.tree.2021.07.009

Rombel, I. T., Sykes, K. F., Rayner, S., & Johnston, S. A. (2002). ORF-FINDER: A vector for high-throughput gene identification. *Gene*, *282*(1–2), 33–41. https://doi.org/10.1016/S0378-1119(01)00819-8

Ronquist, F., Teslenko, M., Van Den mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, *61*(3), 539–542.

Schrödl, M., & Stöger, I. (2014). A review on deep molluscan phylogeny: Old markers, integrative approaches, persistent problems. *Journal of Natural History*, *48*(45–48), 2773–2804.

Simone, L. R. L. (2002). Comparative morphological study and phylogeny of representatives of the superfamily Calyptraeoidea (including Hipponicoidea) (Mollusca, Caenogastropoda). *Biota Neotropica*, *2*(2), 1–137. https://doi.org/10.1590/S1676-06032002000200013

Song, F., Li, H., Jiang, P., Zhou, X., Liu, J., Sun, C., Vogler, A. P., & Cai, W. (2016). Capturing the phylogeny of Holometabola with mitochondrial genome data and Bayesian site-heterogeneous mixture models. *Genome Biology and Evolution*, *8*(5), 1411–1426. https://doi.org/10.1093/gbe/evw086

Strong, E. E., Puillandre, N., Beu, A. G., Castelin, M., & Bouchet, P. (2019). Frogs and tuns and tritons–A molecular phylogeny and revised family classification of the predatory gastropod superfamily Tonnoidea (Caenogastropoda). *Molecular Phylogenetics and Evolution*, *130*, 18–34.

Uribe, J. E., Fedosov, A. E., Murphy, K. R., Sei, M., & Harasewych, M. G. (2021). The complete mitochondrial genome of *Costapex baldwinae* (Gastropoda: Neogastropoda: Turbinelloidea: Costellariidae) from the Caribbean Deep-Sea. *Mitochondrial DNA Part B*, *6*(3), 943–945. https://doi.org/10.1080/23802359.2021.1889408

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. *Molecular Phylogenetics and Evolution*, *133*, 12–23.

Uribe, J. E., & Zardoya, R. (2017). Revisiting the phylogeny of Cephalopoda using complete mitochondrial genomes. *Journal of Molluscan Studies*, *83*(2), 133–144. https://doi.org/10.1093/mollus/eyw052

Uribe, J. E., Zardoya, R., & Puillandre, N. (2018). Phylogenetic relationships of the conoidean snails (Gastropoda: Caenogastropoda) based on mitochondrial genomes. *Molecular Phylogenetics and Evolution*, *127*, 898–906.

Wang, J.-G., Zhang, D., Jakovlić, I., & Wang, W.-M. (2017). Sequencing of the complete mitochondrial genomes of eight freshwater snail species exposes pervasive paraphyly within the Viviparidae family (Caenogastropoda). *PLoS One*, *12*(7), e0181699. https://doi.org/10.1371/journal.pone.0181699

Wang, Q., Liu, H., Yue, C., Xie, X., Li, D., Liang, M., & Li, Q. (2021). Characterization of the complete mitochondrial genome of *Ficus variegata* (Littorinimorpha: Ficidae) and molecular phylogeny of Caenogastropoda. *Mitochondrial DNA Part B*, *6*(3), 1126–1128. https://doi.org/10.1080/23802359.2021.1901628

Williams, S. T., Foster, P. G., & Littlewood, D. T. J. (2014). The complete mitochondrial genome of a turbinid vetigastropod from MiSeq Illumina sequencing of genomic DNA and steps towards a resolved gastropod phylogeny. *Gene*, *533*, 38–47.

WoRMS Editorial Board. (2022). *World Register of Marine Species*. Available from www.marinespecies.org. Accessed 14 March, 2022.

Yang, M., Dong, D., & Li, X. (2021). The complete mitogenome of Phymorhynchus sp. (Neogastropoda, Conoidea, Raphitomidae) provides insights into the deep-sea adaptive evolution of Conoidea. *Ecology and Evolution*, *11*, 7518–7531. https://doi.org/10.1002/ece3.7582

Zaharias, P., Pante, E., Gey, D., Fedosov, A. E., & Puillandre, N. (2020). Data, time and money: Evaluating the best compromise for inferring molecular phylogenies of non-model animal taxa. *Molecular Phylogenetics and Evolution*, *142*, 106660.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lemarcis, T., Fedosov, A. E., Kantor, Y. I., Abdelkrim, J., Zaharias, P., & Puillandre, N. (2022). Neogastropod (Mollusca, Gastropoda) phylogeny: A step forward with mitogenomes. *Zoologica Scripta*, *51*, 550–561. https://doi.org/10.1111/zsc.12552