Acta Oceanol. Sin., 2024, Vol. 43, No. 2, P. 137–150
https://doi.org/10.1007/s13131-023-2258-7
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

Gene characterization and phylogenetic analysis of four mitochondrial genomes in Caenogastropoda

Jiangyong Qu[1†], Wanqi Yang[1†], Xindong Teng[2], Li Xu[3], Dachuan Zhang[1], Zhikai Xing[1], Shuang Wang[1], Xiumei Liu[1], Lijun Wang[1*], Xumin Wang[1*]

[1] College of Life Science, Yantai University, Yantai 264005, China

[2] Qingdao International Travel Healthcare Center, Qingdao 266071, China

[3] Qingdao Dagang Customs, Qingdao 266071, China

Received 24 February 2023; accepted 17 July 2023

Abstract

Caenogastropoda is a highly diverse group, containing ~60% of all existing gastropods. Species in this subclass predominantly inhabit marine environments and have a high ecological and economic value. Owing to the increase in relevant phylogenetic studies, our understanding of between species relatedness in Caenogastropoda has improved. However, the biodiversity, taxonomic status, and phylogenetic relationships of this group remain unclear. In the present study, we performed next-generation sequencing of four complete mitochondrial genomes from three families (Buccinidae, Columbellidae, and Cypraeidae) and the four mitogenomes were classical circular structures, with a length of 16 177 bp in *Volutharpa ampullacea*, 16 244 bp in *Mitrella albuginosa*, 16 926 bp in *Mauritia arabica asiatica* and 15 422 bp in *Erronea errones*. Base composition analysis indicated that whole sequences were biased toward A and T. Then compared them with 171 complete mitochondrial genomes of Caenogastropoda. The phylogenetic relationship of Caenogastropoda derived from Maximum Likelihood (ML) and Bayesian Inference (BI) trees constructed based on CDS sequences was consistent with the results of traditional morphological analysis, with all three families showing close relationships. This study supported Caenogastropoda at the molecular level as a separate clade of Mollusca. According to our divergence time estimations, Caenogastropoda was formed during the middle Triassic period (~247.2–237 Ma). Our novel mitochondrial genomes provide evidence for the speciation of Caenogastropoda in addition to elucidating the mitochondrial genomic evolution of this subclass.

Key words: mitochondrial genome, phylogenetic analysis, Caenogastropoda

Citation: Qu Jiangyong, Yang Wanqi, Teng Xindong, Xu Li, Zhang Dachuan, Xing Zhikai, Wang Shuang, Liu Xiumei, Wang Lijun, Wang Xumin. 2024. Gene characterization and phylogenetic analysis of four mitochondrial genomes in Caenogastropoda. Acta Oceanologica Sinica, 43(2): 137–150, doi: 10.1007/s13131-023-2258-7

1 Introduction

Caenogastropoda, which belongs to the Phylum Mollusca, Class Gastropoda, has a wide variety of shellfish and contains approximately 60% of the living gastropod species (Colgan et al., 2007), comprising 41 superfamilies and 201 families (136 extant and 65 extinct) (Sun et al., 2012). Species of this subclass have strong adaptability, demonstrated by their ability to occupy various environments: including marine, terrestrial, and freshwater. Marine gastropods are found across a range of habitats, such as abyssal regions, intertidal zones, and reefs (Sigwart et al., 2021) and can adopt benthic, planktonic, and other lifestyles. Caenogastropoda shellfish are of high ecological and commercial importance due to their use in food, medicinal, and ornamental industries.

Comprehensive studies examining all Caenogastropoda species are lacking; however, some studies focus on specific species or taxonomic groups. For example, Simone (2004, 2005) carried out a phylogenetic analysis of various groups of Caenogastropoda, and Meyer (2003) studied Cypraeidae phylogeny. Consequently, the phylogenetic relationship between the family and superfamily of Caenogastropoda remains unclear (Osca et al., 2015). Current hypotheses regarding phylogenetic relationships are based on key morphological or anatomical features, such as shape and radula number (Gomes-dos-Santos et al., 2020), which supports the classification of Caenogastropoda as an independent branch. However, there is less strong empirical evidence from molecular research regarding the classification and interspecific phylogenetic relationships of Caenogastropoda, highlighting the need for further studies.

The mitochondrial genome is widely used in population genetics, pedigree geography, molecular evolution, genealogy, and comparison and evolutionary genomic research, due to its small molecular weight, stable gene composition, low recombination, and multiple copies in cells (He et al., 2011). Metazoan mtDNA is a linear, closed circular molecule, except in a small number of

Foundation item: Research and Development Program of Shandong Province, China (Major Science and Technology Innovation Project) under contract No. 2021CXGC011306; MNR Key Laboratory of Eco-Environmental Science and Technology, China under contract No. MEEST-2021-05; Natural Science Foundation of Shandong Province under contract No. ZR2020MD002; Doctoral Science Research Foundation of Yantai University under contract Nos SM15B01, SM19B70 and SM19B28; Double-Hundred Action of Yantai City under contract No. 2320004-SM20RC02.

*Corresponding author, E-mail: wanglijun@ytu.edu.cn; wangxm@ytu.edu.cn

†The two authors contributed equally to this paper.

aquatic animals (Bridge et al., 1992). The mitochondrial genome provides abundant molecular markers for phylogenetic research (Lee et al., 2019); thus, the classification, evolution and associated mechanisms of species depend heavily on the ancestry of the mitochondrial genome and its phylogeny.

In this study, we compared the complete, newly sequenced mitochondrial genomes of four gastropod species: *Volutharpa ampullacea* (Middendorff, 1848), *Mitrella albuginosa* (Reeve, 1859), *Mauritia arabica asiatica* (Schilder and Schilder, 1939), *Erronea errones* (Linnaeus, 1758). We reconstructed the phylogeny of Caenogastropoda-related species using 171 representative mitogenomes encompassing all four orders of this subclass and estimated the local divergence time of genetic events in the main branches. These data will provide important molecular evidence for the classification of Caenogastropoda species.

2 Materials and methods

2.1 *Sample collection and DNA extraction*

Volutharpa ampullacea, *M. albuginosa*, *Mauritia arabica asiatica* and *E. errones* were collected from Sanya, Hainan Province, China (18°20′52.52″N, 109°53′77.63″E) and used to sequence the complete mitochondrial genome. Portions of muscle tissue were dissected from each specimen (4 × 10 mg) and frozen at –80℃. Genomic DNA was extracted from the specimens following the operation steps detailed in the TIANamp Marine Animals DNA Kit DP324 instruction manual. DNA quality and quantity were measured using a NanoDrop ND1000 spectrophotometer (Thermo Fisher Scientific).

2.2 *High-throughput sequencing and assembly*

Library preparation was carried out using the Illumina TruSeq™ Nano DNA Sample Prep Kit (Williams et al., 2014). Mitochondrial genome sequencing of the four Caenogastropoda species was conducted by the Shanghai Lingen Biotechnology Corporation. A Covaris M220 ultrasonic crusher was used to segment fragments to lengths of 300–500 bp, and connectors were added at each end for bridge PCR amplification. Illumina HiSeq 4000 sequencing technology was used for paired-end sequencing of the sample DNA. It was expected that some of the Illumina original sequencing data would be of low quality; therefore, the quality of the original data was decreased to improve the accuracy of the subsequent assembly. SPAdes v3.10.1 (Bankevich et al., 2012) (http://bioinf.spbau.ru/spades) was used to splice and compare the clean data. GapCloser v1.12 (http://soap.genomics.org.cn/so-apdenovo.html) to fill and optimize the local holes in the assembly results. To obtain the final mitochondrial genome sequence, the mitochondrial assembly sequence was corrected by referencing the reference genome.

2.3 *Annotation and comparative genome analysis*

In Geneious R10 (Biomatters Ltd, Auckland, New Zealand), protein-encoding genes, ribosomal (r) RNA genes, and transfer (t) RNA genes were annotated based on Caenogastropoda. This software was additionally used to calculate the AT and CG skew (Perna and Kocher, 1995).

$$\text{AT skew} = (\text{A}\% - \text{T}\%)/(\text{A}\% + \text{T}\%); \tag{1}$$

$$\text{GC skew} = (\text{G}\% - \text{C}\%)/(\text{G}\% + \text{C}\%). \tag{2}$$

MITOS (http://mitos.bioinf.uni-leipzig.de/index.py) was used to predict the coding protein, tRNA and rRNA genes of the mitochondrial genomes, and to predict the initial genomes. To obtain an accurate set of conserved genes, the start and stop codon positions of each gene were manually corrected (Bernt et al., 2013). Using MITOS to predict the secondary structure of tRNA, we first set the genetic code as "invertebrate", and then selected "RefSeq 89 metazoa" as the annotated reference data set. Physical maps of the four mitogenomes were prepared using Organellar Genome DRAW (OGDRAW) (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html) (Stephan et al., 2019). Based on MEGA 7.0, the alignment and base composition of the sequences were conducted (Sudhir et al., 2016). Use the codon usage function in MEGA 7.0 to get the relative synonymous codon usage (RSCU), and then draw the stacking histogram through R (Feng et al., 2021). The mitochondrial genomes of four species from the families Buccinidae, Columbellidae, and Cypraeidae were compared using Geneious' Mauve program (Drummond, 2012).

2.4 *Phylogenetic analysis*

Based on the 171 mitochondrial genomes obtained from GenBank in the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/) and four new mitochondrial genomes obtained in this study (Table 1), we conducted a phylogenetic analysis of Caenogastropoda using 13 mitochondrial protein-encoding genes. *Katharina tunicata* (GenBank accession number: NC_001636) was used as the outgroup

Table 1. List of study species and corresponding GenBank accession numbers

Order	Family	Species	Size/bp	Accession
Neogastropoda	Buccinidae	*Aeneator recens*	15 264	NC_039122
		Aeneator elegans	15 254	NC_039120
		Buccinulum pertinax	15 247	NC_039124
		Cominella adspersa	15 251	NC_039125
		Buccinum undatum	15 265	NC_040940
		Buccinum bayani	15 262	MW646292
		Buccinum pemphigus	15 265	NC_029373
		Volutharpa ampullacea	16 177	OP067009
		Volutharpa perryi	15 255	NC_028183
		Neptunea arthritica	15 256	KU246047
		Siphonalia subdilatata	15 393	MG827217
		Antarctoneptunea benthicola	15 229	NC_039119
		Antarctoneptunea aurora	15 227	NC_039117
		Kelletia lischkei	15 225	NC_039123
		Penion chathamensis	15 227	NC_039116

to be continued

continued from Table 1

Order	Family	Species	Size/bp	Accession
Neogastropoda	Buccinidae	*Penion sulcatus*	15 227	NC_037185
		Penion cuvierianus	15 235	NC_039118
		Penion ormesi	15 234	NC_039126
		Penion maximus	15 249	NC_037237
		Babylonia areolata	15 445	NC_023080
		Babylonia lutosa	15 346	NC_028628
		Babylonia zeylanica	16 181	MN604402
	Fasciolariidae	*Fusinus longicaudus*	16 319	NC_045906
	Melongenidae	*Brunneifusus ternatanus*	16 254	NC_059758
		Hemifusus tuba	15 483	MN462591
	Nassariidae	*Nassarius hepaticus*	15 337	NC_038169
		Nassarius conoidalis	15 332	NC_041310
		Nassarius siquijorensis	15 337	NC_048962
		Nassarius foveolatus	15 343	NC_041546
		Nassarius pullus	15 278	NC_041311
		Nassarius glans	15 296	NC_049091
		Nassarius javanus	15 325	NC_041547
		Nassarius sinarus	15 325	NC_041545
		Nassarius granifer	16 417	MT374078
		Nassarius variciferus	15 269	NC_029173
		Nassarius festivus	15 195	NC_037607
		Nassarius fraterculus	15 174	NC_037604
		Reticunassa hiradoensis	15 194	NC_037887
		Nassarius jacksonianus	15 234	NC_041548
		Tritia obsoleta	15 263	NC_007781
		Tritia reticulatus	15 271	NC_013248
	Columbellidae	*Columbella adansoni*	16 272	KP716637
		Mitrella albuginosa	16 244	MZ618619
	Olividae	*Amalda northlandica*	15 354	NC_014403
	Costellariidae	*Costapex baldwinae*	15 321	MW044625
	Conidae	*Bathytoma punicea*	16 037	NC_038182
		Californiconus californicus	15 444	NC_032377
		Conus betulinus	16 240	NC_039922
		Conus borgesi	15 536	NC_013243
		Conus infinitus	15 522	KY864967
		Conus guanche	15 506	KY801847
		Conus hybridus	15 507	KY801863
		Conus unifasciatus	15 506	KY801860
		Conus capitaneus	15 829	NC_030354
		Conus consors	16 112	NC_023460
		Conus striatus	15 738	NC_030536
		Conus gloriamaris	15 774	NC_030213
		Conus textile	15 562	NC_008797
		Conus quercinus	16 430	NC_035007
		Conus tribblei	15 570	NC_027957
		Conus tulipa	15 756	NC_027518
		Conus raulsilvai	15 534	MF491520
	Raphitomidae	*Phymorhynchus buccinoides*	15 764	MN583349
	Turridae	*Fusiturris similis*	15 595	NC_013242
		Lophiotoma cerithiformis	15 380	NC_008098
		Gemmuloborsonia moosai	15 541	NC_038183
	Terebridae	*Oxymeris dimidiata*	16 513	NC_013239
	Clavatulidae	*Turricula nelliae spurius*	16 453	MK251986
	Muricidae	*Bolinus brandaris*	15 380	NC_013250
		Murex trapa	15 408	MN462589
		Chicoreus torrefactus	15 359	NC_039164
		Chicoreus asianus	15 361	MN793976
		Concholepas concholepas	15 495	NC_017886

to be continued

continued from Table 1

Order	Family	Species	Size/bp	Accession
Neogastropoda	Muricidae	*Reishia clavigera*	15 285	NC_010090
		Thais luteostoma	15 301	NC_039165
		Indothais lacera	15 272	NC_037221
		Rapana venosa	15 271	NC_011193
		Menathais tuberosa	15 294	NC_031405
		Boreotrophon candelabrum	15 265	NC_046505
		Ceratostoma burnetti	15 334	NC_046569
		Ceratostoma rorifluum	15 338	NC_046526
		Ocinebrellus falcatus	15 326	NC_046052
		Ocinebrellus inornatus	15 324	NC_046577
	Volutidae	*Cymbiola nobilis*	16 314	MN871953
		Melo melo	15 721	MN462590
		Harpovoluta charcoti	15 487	MT232845
		Cymbium olla	15 375	NC_013245
		Neptuneopsis gilchristi	15 312	MN125492
	Cancellariidae	*Sydaphera spengleriana*	16 336	MZ934412
		Bivetiella cancellata	16 648	NC_013241
Littorinimorpha	Bursidae	*Bufonaria rana*	15 510	NC_054277
	Ranellidae	*Charonia lampas*	15 330	KU237290
		Charonia tritonis	15 346	MT043269
		Monoplex parthenopeus	15 270	NC_013247
	Cassidae	*Galeodea echinophora*	15 388	NC_028003
	Ficidae	*Ficus variegata*	15 736	NC_056153
	Cypraeidae	*Mauritia arabica asiatica*	16 926	MZ667219
		Erronea errones	15 422	OP021744
		Cypraea tigris	16 177	MK783263
		Monetaria annulus	16 087	LC469295
	Strombidae	*Conomurex luhuanus*	15 799	NC_035726
		Laevistrombus canarium	15 626	NC_053786
		Harpago chiragra	15,460	MH122656
		Strombus gigas	15 461	NC_024932
		Lambis lambis	15 481	MH115428
	Vermetidae	*Ceraesignum maximum*	15 578	NC_014583
		Dendropoma gregarium	15 641	NC_014580
		Eualetes tulipa	15 078	NC_014585
		Thylacodes squamigerus	15 544	NC_014588
	Naticidae	*Cryptonatica andoi*	15 302	NC_046598
		Cryptonatica janthostoma	15 235	NC_046704
		Naticarius hebraeus	15 384	NC_028002
		Paratectonatica tigrina	15 201	NC_050661
		Tanea lineata	15 156	NC_050662
		Euspira gilva	15 315	NC_046593
		Euspira pila	15 244	NC_046703
		Glossaulax reiniana	15 254	NC_041162
		Neverita didyma	15 629	NC_046594
		Mammilla kurodai	15 309	NC_046596
		Mammilla mammata	15 319	NC_046597
		Polinices sagamiensis	15 383	NC_046595
		Notocochlis gualtieriana	15 176	NC_046705
	Littorinidae	*Littorina brevicula*	16 356	NC_050987
		Littorina saxatilis	16 887	NC_030595
	Baicaliidae	*Baicalia turriformis*	15 127	NC_035869
		Maackia herderiana	15 154	NC_035871
		Godlewskia godlewskia	15 224	NC_035870
	Tateidae	*Potamopyrgus antipodarum*	15 149	MG979470
		Potamopyrgus estuarinus	15 120	NC_021595
	Pomatiopsidae	*Oncomelania hupensis*	15 182	NC_013073
		Oncomelania hupensis nosophora	15 182	LC276226

to be continued

continued from Table 1

Order	Family	Species	Size/bp	Accession
Littorinimorpha	Pomatiopsidae	*Oncomelania quadrasi*	15 184	LC276227
		Oncomelania hupensis robertsoni	15 191	NC_013187
		Tricula hortensis	15 179	NC_013833
Sorbeoconcha	Batillariidae	*Batillaria attramentaria*	16 095	MN557850
	Turritellidae	*Turritella bacillum*	15 868	NC_029717
	Pachychilidae	*Tylomelania sarasinorum*	16 632	NC_030263
	Potamididae	*Cerithidea obtusa*	15 708	NC_039951
		Cerithidea tonkiniana	15 617	MZ168697
	Pleuroceridae	*Koreanomelania nodifila*	15 737	NC_046494
		Koreoleptoxis globus ovalis	15 866	LC006055
	Semisulcospiridae	*Semisulcospira libertina*	15 432	NC_023364
		Semisulcospira gottschei	16 101	MK559478
	Thiaridae	*Melanoides tuberculata*	15 821	MZ321058
	Paludomidae	*Pseudocleopatra dartevellei*	15 368	NC_045095
Architaenioglossa	Viviparidae	*Bellamya aeruginosa*	17 013	NC_035735
		Bellamya purificata	17 310	NC_039097
		Bellamya quadrata	17 485	NC_031850
		Sinotaia quadrata	17 255	NC_056961
		Cipangopaludina ussuriensis	16 596	NC_035754
		Cipangopaludina cathayensis	17 157	NC_025577
		Cipangopaludina lecythis ampullacea	16 892	MZ488942
		Margarya melanioides	17 083	NC_035587
		Margarya oxytropoides	17 044	NC_035586
		Cipangopaludina longispira	17 356	NC_059028
		Margarya monodi	17 051	KY196441
		Cipangopaludina chinensis	17 009	NC_035734
		Rivularia auriculata	16 552	NC_051889
		Viviparus chui	16 392	NC_035733
	Cochlostomatidae	*Obscurella hidalgoi*	15 536	NC_028004
Unassigned	Provannidae	*Alviniconcha boucheti*	15 981	NC_049893
		Ifremeria nautilei	15 664	NC_024642

species. Nucleotide sequences were compared by using MEGA 7.0, and the serial alignment was completed by the ClustalW function. Using ProtTest 3.4.2, the Maximum Likelihood (ML) model was selected to construct the phylogenetic tree (Darriba et al., 2011), following that, ML bootstrapping analysis was performed using RAxML v8.2.12 to reconstruct a phylogenetic tree (Stamatakis, 2006), with 1000 replications under the MtMam + I + G + F model. Bayesian Inference (BI) was performed using MrBayes (Huelsenbeck and Ronquist, 2001). Phylogenetic analysis was conducted using four Markov chains for 1 000 000 generations across two independent runs. A sample of the output trees was taken every 100 generations. We removed the first 25% of the samples from the phylogenetic analysis as burn-in after the average standard deviation of split frequencies was <0.01. The phylogenetic tree was plotted and embellished in Figtree v1.4.4.

2.5 *Estimation of divergence times*

The divergence dates between Caenogastropoda clades were estimated using the 13 protein-coding genes (PCGs) at the nucleotide level and an uncorrelated relaxed molecular clock model in BEAST v1.10.4. The clock model was chosen, which allows rates to vary between branches without a priori assumptions regarding the autocorrelation between adjacent branches. The Yule process of speciation was employed for the prior tree. The best-fitting evolutionary model (GTR + I + G) was applied. The final Markov chain was run twice for 750 000 000 generations, with sampling every 1 000 generations. The first 10% of the samples were discarded as burn-in, according to the convergence of chains checked with Tracer v1.7.2. The ESS of all parameters was >200. TreeAnnotator v1.10.4 was used to generate the tree, and the divergence time was visualized using FigTree v1.4.4.

A lognormal posterior distribution was obtained to estimate the divergence times, based on the four calibration points for the divergence times of the respective splits. (1) Neogastropoda and Littorinimorpha diverged within the middle Triassic at 232.5 Ma (Bittner, 1912). (2) Neogastropoda had three important differentiation nodes in the late Cretaceous at 77.05 Ma (Kosnik, 2005) and 61 Ma (Zinsmeister et al., 1989). (3) The divergence of the genus Penion was restricted to 41.3–38 Ma (Foster et al., 2020). Using fossils as minimum bounds (offsets), we chose means and standard deviations (SD) to ensure that the 95% probability limit corresponded to a soft maximum limit.

3 Results

3.1 *Genome features*

There were four mitochondrial genomes with lengths of 15 422 bp (*E. errones*), 16 177 bp (*V. ampullacea*), 16 244 bp (*M. albuginosa*), and 16 926 bp (*Mauritia arabica asiatica*), that were assembled into circular molecules (Fig. 1). The average GC contents of the mitochondrial genomes were 35.6% (*Mauritia arabica asiatica*), 31.4% (*E. errones*), 29.1% (*M. albuginosa*), and 32.0% (*V. ampullacea*) (Table 2). As predicted, the mitochondrial genomes encoded 37 genes, including 13 protein-encoding genes, 22 tRNA genes, and 2 rRNA genes, consistent with those previously reported for Caenogastropoda species (Osca et al., 2015).

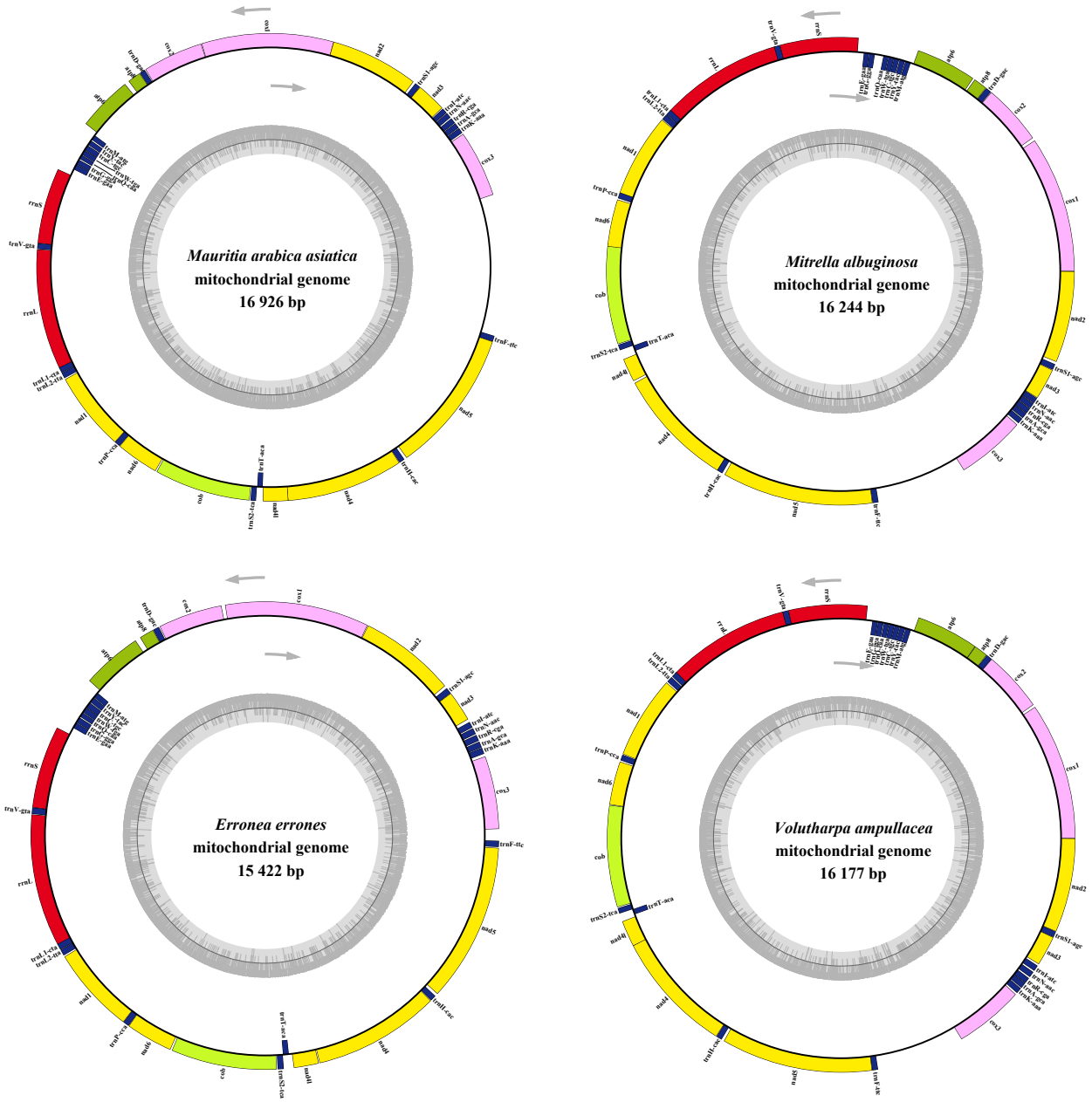

Fig. 1. Gene map of the complete mitochondrial genomes of *Mauritia arabica asiatica*, *Erronea errone*s, *Mitrella albuginosa* and *Volutharpa ampullacea*. Colored regions represent protein-coding and rRNA genes, black regions are tRNAs, the outer circle is the plus (+) strand, and the inner circle is the minus (–) strand.

Gene distributions in both the heavy and light strands were conserved among the four genomes. The heavy strand encodes most genes, and the light strand encodes only eight tRNA genes (*trnM*, *trnY*, *trnC*, *trnW*, *trnQ*, *trnG*, *trnE*, and *trnT*), which is a typical feature of Caenogastropoda (Peretolchina et al., 2020). All tRNA gene lengths ranged from 57 bp to 71 bp. Among the 13 protein coding genes, the longest was *nad5* (1 722 bp), and the shortest was *atp8* (159 bp). The length of *rrnS* ranged from 882 bp to 889 bp and *rrnL* from 1 387 bp to 1 424 bp (Table 3). The two genes were located between *trnE* and *trnL1*, and separated by *trnV*. Through the above description, it showed that the mitochondrial genome structure of the entire Caenogastropod is relatively stable and highly conserved. Differences in mitochondrial genomes may be due to differences in intergenic or non-coding regions.

3.2 *Protein-coding genes*

The 13 PCGs were represented in the mitogenomes of the four newly sequenced species and were H-strand encoded, which included three subunits of cytochrome coxidase (*cox1-3*), seven subunits of NADH dehydrogenase (*nad1-6*, *nad4L*), a ubichinol cytochrome c reductase (*cytb*), and two subunits of ATP synthases (*atp6*, *atp8*). It is common for protein-coding genes to use ATG as the start codon in the four mitochondrial genomes. ATA, ATT, and ATC were used as start codons in addition to ATG (Table 3). ATC was used as the start codon for *nad4* in the mitochondrial genome of *E. errones*. Based on the high percentage values in the mitochondrial genomes of all three typical stop codons (TAA, TAG, and TGA), a preference for TAA was clearly evident for *M. arabica asiatica* (8, 61.54%), *E. errones* (9, 69.23%), *M. albuginsa* (11, 84.62%), and *V. ampullacea*

Table 2. General features of the complete mitochondrial genomes of *Volutharpa ampullacea*, *Mitrella albuginosa*, *Mauritia arabica asiatica* and *Erronea errones*

Species	*Mauritia arabica asiatica*	*Erronea errones*	*Mitrella albuginosa*	*Volutharpa ampullacea*
Accession number	MZ667219	OP021744	MZ618619	OP067009
Genome size/bp	16 926	15 422	16 244	16 177
GC content/%	35.6	31.4	29.1	32.0
A/%	26.9	28.6	28.1	28.0
T/%	36.9	39.2	41.9	39.6
C/%	17.9	15.6	12.8	16.0
G/%	18.4	16.6	17.2	16.4
GC skew	0.013 8	0.031	0.147	0.012 3
AT skew	–0.157	–0.156	–0.197	–0.172
rRNA/tRNA/CDS content	2/22/13	2/22/13	2/22/13	2/22/13
Protein-coding/%	65.95	71.68	68.59	69.14
Gene total length/bp	11 163	11 055	11 142	11 184
Spacer content/%	11.62	3.77	8.33	7.73
Protein-coding/%	65.95	71.68	68.59	69.14

Table 3. Summary of the gene features of *Volutharpa ampullacea*, *Mitrella albuginosa*, *Mauritia arabica asiatica* and *Erronea errones*

Gene	Strand	Size/bp	Start codon	Stop codon	Anticodon	GC-content/%
cox1	+	1 533^{b}/1 536^{a,d}/1 545^{c}	ATA^{c}/ATG^{a,b,d}	TAA^{a,b,c}/TAG^{d}		32.2^{b}/34.8^{a}/35.7^{d}/37.3^{c}
cox2	+	687^{a,b,d}/711^{c}	ATT^{c}/ATG^{a,b,d}	TAA^{a,b,c,d}		30.4^{b}/33.2^{a}/33.6^{d}/38.0^{c}
trnD	+	68^{a,b,c,d}			GTC	17.6^{b}/22.1^{a}/23.5^{d}/26.5^{c}
atp8	+	159 ^{a,b,c,d}	ATG^{a,b,c,d}	TAA^{a,b,c,d}		23.9^{d}/26.4^{b}/28.9^{a,c}
atp6	+	696^{c,d}/702^{a,b}	ATA^{a,b}/ATG^{c,d}	TAA^{a,b,d}/TAG^{c}		29.2^{b}/30.9^{a}/31.2^{d}/36.6^{c}
trnM	–	66^{a,d}/67^{b,c}			CAT	29.9^{b}/30.3^{a}/33.3^{d}/35.8^{c}
trnY	–	65^{a}/66^{b,c,d}			GTA	39.4^{b,d}/40.9^{c}/41.5^{a}
trnC	–	64^{a,b}/69^{d}/71^{c}			GCA	40.6^{b}/42.2^{a}/42.3^{c}/43.5^{d}
trnW	–	66^{a,b,c,d}			TCA	21.2^{c,d}/24.2^{b}/25.8^{a}
trnQ	–	57^{c,d}/62^{a,b}			TTG	40.3^{b}/42.1^{c,d}/43.5^{a}
trnG	–	67 ^{a,b,c,d}			TCC	25.4^{c,d}/28.4^{a}/29.9^{b}
trnE	–	65^{b}/67^{a}/70^{c,d}			TTC	21.5^{b}/24.3^{d}/31.3^{a}/31.4^{c}
rrnS	+	882^{a,b}/883^{d}/889^{c}				30.6^{b}/32.3^{a}/34.2^{c}/34.7^{d}
trnV	+	68^{a}/69^{b,c}/70^{d}			TAC	16.2^{a}/18.6^{d}/18.8^{b,c}
rrnL	+	1 387^{a}/1 393^{b}/1 420^{d}/1 424^{c}				25.8^{b}/26.8^{d}/27.3^{a}/29.3^{c}
trnL1	+	69^{a,c,d}/70^{b}			TAG	25.7^{b}/27.5^{a}/29.0^{d}/31.9^{c}
trnL2	+	69^{a,b}/70^{c,d}			TAA	29.0^{a,b}/32.9^{c}/34.3^{d}
nad1	+	940^{b}/942^{a,c,d}	ATG^{a,b,c,d}	TAA^{a,c}/TAG^{b,d}		30.7^{a}/31.8^{b}/32.5^{a}/34.4^{c}
trnP	+	67^{d}/68^{a,c}/69^{b}			TGG	31.3^{d}/32.4^{c}/36.8^{a}/37.7^{b}
nad6	+	501^{a}/504^{c,d}/546^{b}	ATG^{a,b,c,d}	TAA^{a,b,c,d}		26.6^{b}/28.4^{d}/27.5^{a}/33.5^{c}
cob	+	1 053^{b}/1 140^{a,c}/1 149^{d}	ATG^{a,c}/ATT^{b}/ATA^{d}	TAA^{b,c,d}/TAG^{a}		30.6^{b}/33.9^{d}/34.7^{a}/38.9^{c}
trnS2	+	64^{a}/65^{b}/66^{c,d}			TGA	34.4^{a}/35.4^{b}/37.9^{d}/40.9^{c}
trnT	–	66^{a}/67^{b,d}/68^{c}			TGT	31.3^{d}/34.3^{b}/38.2^{c}/39.4^{a}
nad4l	+	297 ^{a,b,c,d}	ATG^{a,b,c,d}	TAG^{a,b,c,d}		27.9^{b}/29.6^{d}/31.0^{a}/34.0^{c}
nad4	+	1 356^{d}/1 371^{c}/1 374^{a}/1 377^{b}	ATG^{a,b,c}/ATC^{d}	TAA^{b}/TAG^{a,c,d}		27.8^{b}/30.0^{d}/31.7^{a}/35.4^{c}
trnH	+	66^{a,b,d}/68^{c}			GTG	25.8^{b}/28.8^{a}/30.3^{d}/30.9^{c}
nad5	+	1 722 ^{a,b,c,d}	ATG^{a,b,c,d}	TAA^{b,d}/TAG^{a,c}		28.3^{b}/30.3^{d}/30.5^{a}/35.6^{c}
trnF	+	67^{a}/68^{b,d}/69^{c}			GAA	36.8^{b}/40.3^{a}/51.5^{d}/52.2^{c}
cox3	+	780^{a,b}/786^{c}/792^{d}	ATA^{c}/ATG^{a,b}/ATT^{d}	TAA^{a,b,c,d}		35.3^{b}/36.5^{a}/36.7^{d}/41^{c}
trnK	+	67^{b}/68^{a,c}/69^{d}			TTT	29.4^{a}/31.3^{b}/36.8^{c}/43.5^{d}
trnA	+	67^{b,c}/68^{a,d}			TGC	22.1^{d}/23.9^{b}/26.9^{c}/30.9^{a}
trnR	+	69 ^{a,b,c,d}			TCG	34.8^{b}/39.1^{a}/40.6^{d}/42.0^{c}
trnN	+	67^{c}/68^{b}/69^{a,d}			GTT	30.9^{b}/33.3^{d}/36.2^{a}/40.3^{c}
trnI	+	66^{c}/67^{b,d}/69^{a}			GAT	36.2^{a}/38.8^{b}/40.3^{d}/42.4^{c}
nad3	+	354 ^{a,b,c,d}	ATG^{a,b,c,d}	TAA^{b,d}/TAG^{a,c}		29.7^{b}/31.1^{d}/34.3^{a}/37.0^{c}
trnS1	+	68 ^{a,b,c,d}			GCT	32.4^{b}/33.8^{a}/35.3^{c,d}
nad2	+	1 056^{a,b}/1 062^{c,d}	ATG^{a,b,c,d}	TAA^{a,b,c,d}		27.3^{b}/28.4^{d}/29.2^{a}/33.6^{c}

^{a}*Volutharpa ampullacea*, ^{b}*Mitrella albuginosa*, ^{c}*Mauritia arabica asiatica*, ^{d}*Erronea errones*.

(9, 69.23%) (Table 3).

Conversely, all other amino acids are encoded by either two or four codons, with the exception of Ser and Leu, which are encoded by eight and six codons, respectively (Yang et al., 2018).

The most frequently used PCGs were Leu (UUA), Ser (UCU), Pro (CCU), and Ala (GCU). A or U codons were more frequent in the third position compared with C or G codons, according to the relative synonymous codon usage values. Therefore, NNA and NNU were the majority codons, whereas NNC and NNG were the minority codons (Fig. 2).

The majority of tRNA genes fold into typical clover structures (Wolstenholme, 1992). The dihydrouridine (DHU) loop of all tRNAs is a large ring as opposed to a conservative stem-ring structure. However, this is a feature of a typical metazoan mito-

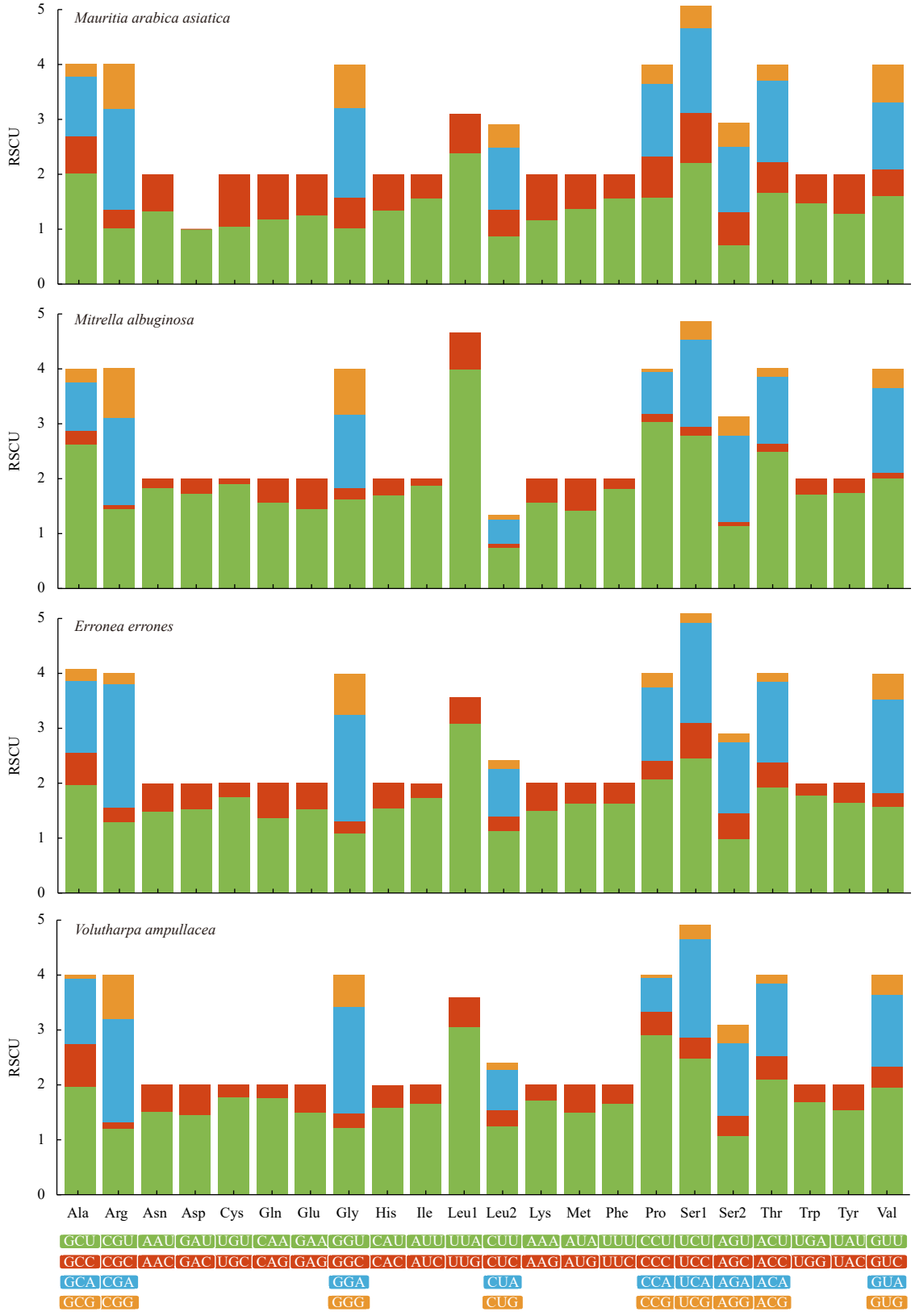

Fig. 2. The relative synonymous codon usage (RSCU) in the mitochondrial genomes of four species.

chondrial genome (Navajas et al., 2002). We found that almost all tRNAs had the same length of the anticodon (AC) loop (7 bp), except *trnC* and *trnH*. The AC loop was 9 bp long in all four species. Among all tRNA genes, the amino acid receptor (AA) stems were conserved at 7 bp, excluding *trnQ* (Fig. 3). The anticodon and dihydrouridine arms are generally conserved sections in tRNA secondary structures, and the total length of tRNA is commonly determined by the variable and the size of the D-loop (Pu et al., 2017).

3.3 *Mitochondrial genome collinearity analysis*

We compared the mitochondrial genomes of four species and their respective families. We detected the mitogenome structures of 28 species from three families and found that their gene sequences all matched (Fig. 4), and there was no inversion, reversal and other phenomena between genes. This suggests that the mitochondrial genomes of these species are highly conserved. The alignment display was organized into one horizontal "panel" per input genome sequence, whereby each panel of the genome contains the name of the genome sequence, which acts as a scale to show the sequence coordinates for that genome. In Fig. 4, the green region is tRNA, the red region is rRNA, and the white region is the protein-coding sequence. We also compared two taxonomically unclear species and found them to be very similar to the species gene arrangement of Littorinimorpha and Neogastropoda, with only the *trnW* position being altered.

3.4 *Phylogenetic relationship*

We collected 171 protein-coding sequences from the mitochondrial genomes of gastropods and obtained 13 protein-coding genes to construct ML and BI phylogenetic trees (Table S1), using *K. tunicata* as the outgroup. We found that the phylogenetic trees inferred by ML and BI methods were highly concordant and divided all species into four clades, with one for each order: Neogastropoda, Littorinimorpha, Sorbeoconcha, and Architaenioglossa (Fig. 5). The newly sequenced species, *V. ampullacea* was the closest relative to *V. perryi*. The experimental species, *M. albuginosa* and *Columbella adansoni* had the closest relationship. In the family Cypraeidae, the sequenced species *Mauritia arabica asiatica* was the closest relative of *Cypraea tigris* and clustered with the experimental species *E. errones*, followed by *Monetaria annulus*. There were high bootstrap support and posterior probability values in the trees, except for the branches of the Sorbeoconcha and Architaenioglossa orders, and the clade of families Buccinidae and Conidae.

3.5 *Divergence times*

A chronogram was inferred based on nucleotide sequences and fossil records. Time-calibrated phylogeny indicated that Neogastropoda and Littorinimorpha diverged approximately 127.97 Ma [95% highest posterior density (HPD) interval = 120.19–139.44 Ma]. The Cypraeidae node was estimated to be formed ~ 109.99 Ma (with a 95% HPD: 110.91–132.92 Ma) and the Columbellidae node at ~ 78.42 Ma (with a 95% HPD: 73.09–82.99 Ma). The divergence time of *Mauritia arabica* and *E. errones* was 36.38 Ma (with a 95% HPD: 30.94–41.55 Ma) (Fig. 6).

4 Discussion

4.1 *Genome features*

In this study, the complete mitochondrial genomes of *M. arabica asiatica*, *E. errones*, *M. albuginosa,* and *V. ampullacea* from the subclass Caenogastropoda expanded the available mitochon-

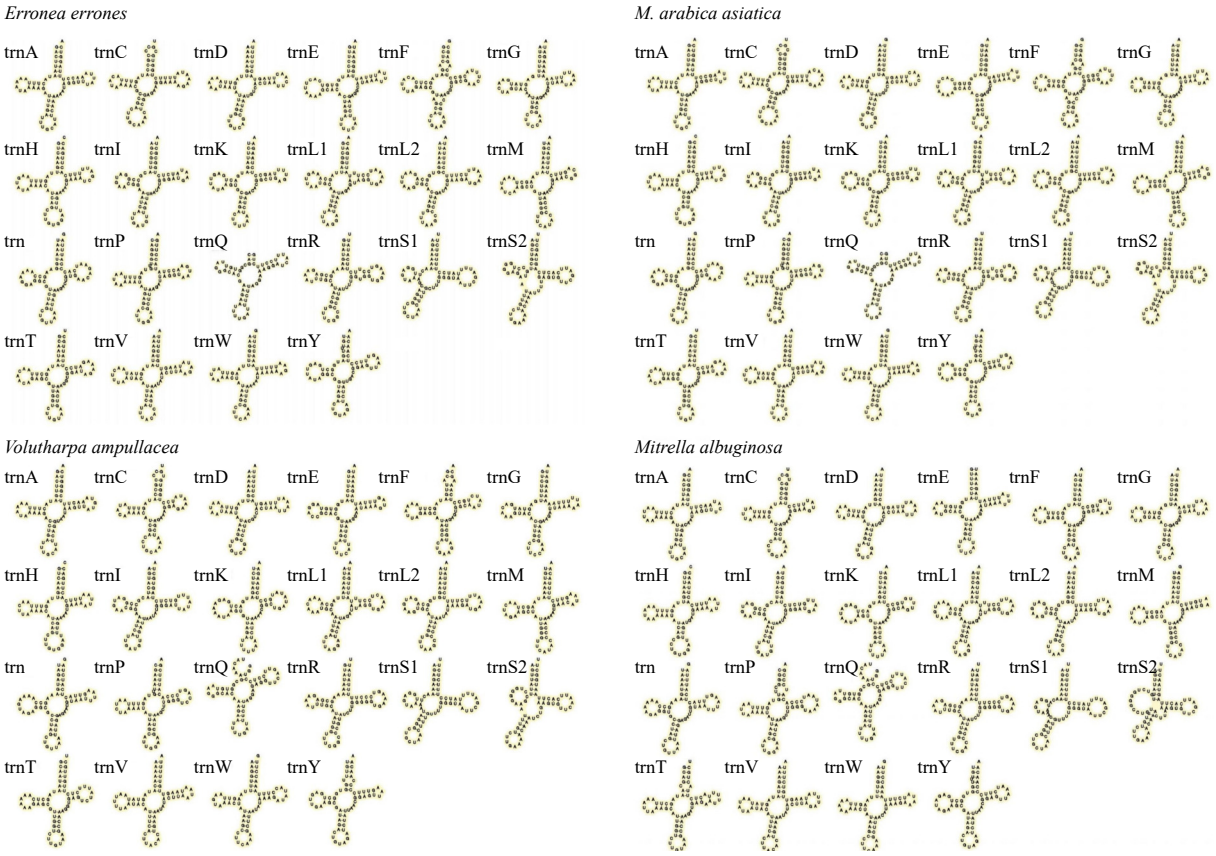

Fig. 3. Map of the tRNA secondary structure in the four studied mitochondrial genomes.

Fig. 4. Multiple genome wide alignment of 28 mitochondrial genomes of Caenogastropoda.

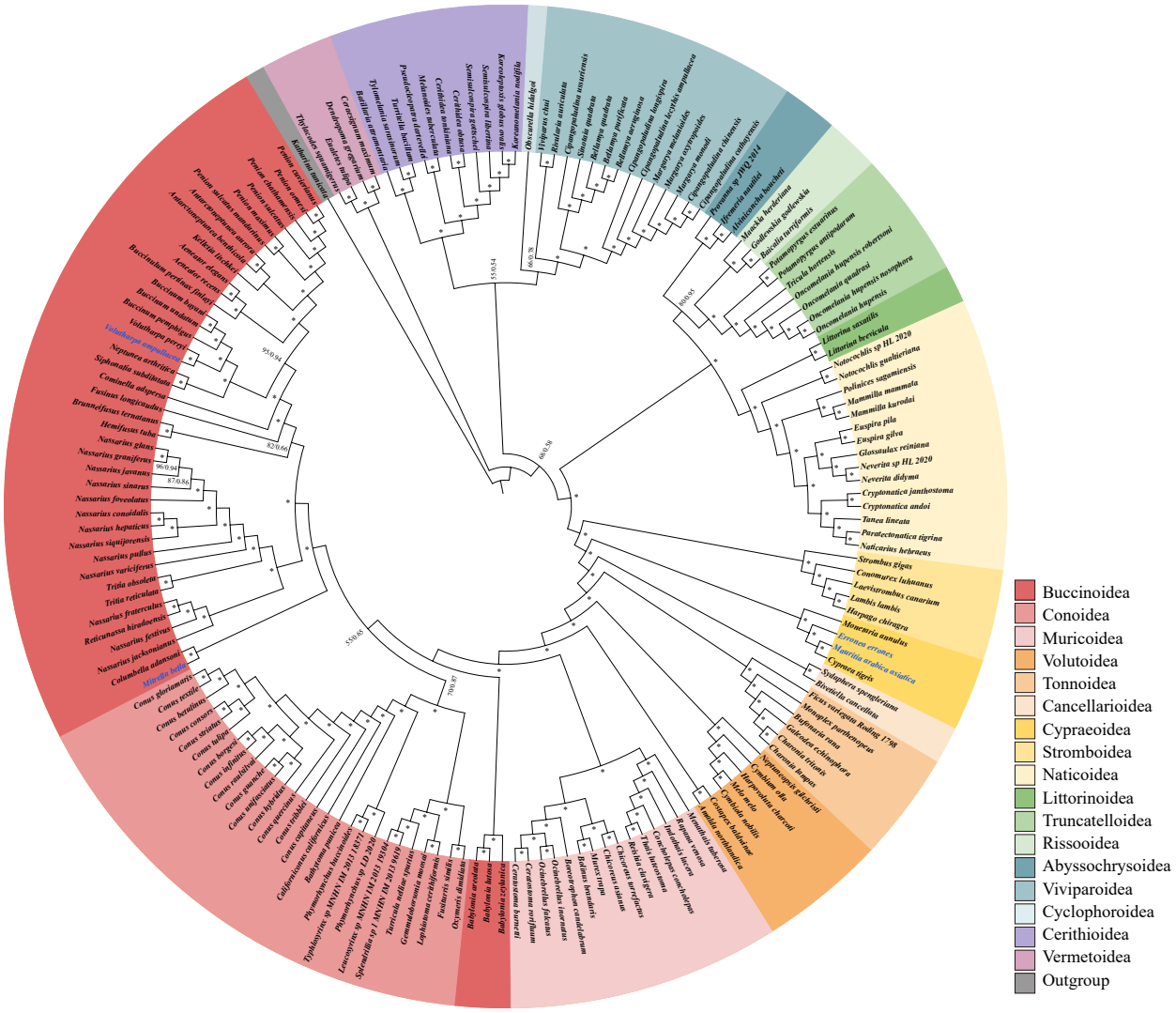

Fig. 5. Phylogenetic tree inferred using Bayesian Inference (BI) and Maximum Likelihood (ML) methods based on tandem sequences of 13 PCGs from 171 gastropod mitotic genomes. *Katharina tunicata* (NC_001636) is the outgroup. "*" means full nodal support by both inferences (1/100).

drial pool of Caenogastropoda. Typically, gastropod mitochondrial genomes exhibit high rates of gene rearrangement between the major lineages. However, genome composition is relatively stable across major lineages, and rearrangements are generally limited to tRNA genes (Grande et al., 2008). The newly sequenced mitochondrial genomes have relatively consistent gene sequences, including tRNA genes, and conform to the consensus gene sequence shared by most gastropod mitochondrial genomes. All the PCGs analyzed in this study used conventional start codons. ATG and ATA have been reported in several mollusk genomes, and TAA and TAG are the most common stop codons in mollusks (Xu et al., 2012). Incomplete stop codons may function through polyadenylation of subsequently transcribed mRNA.

The gene order of the mitochondrial genomes of *Alviniconcha boucheti* and *Ifremeria nautilei* was compared to that of other gastropods, which was found to contain similar gene arrangements to that reported for Caenogastropoda. Mitochondrial genome organization in gastropods is prone to rearrangements between main lineages but is relatively stable, with changes being restricted to tRNA genes (Osca et al., 2014). Therefore, a mitochondrial gene arrangement can be used as a fair proxy to assign any gastropod to one of the main lineages, such as the assignment of *A. boucheti* and *I. nautilei* to Caenogastropoda. Within Caenogastropods, the closest gene order is that shared by the mitochondrial genomes of the family Baicaliidae and Pomatiopsidae in the order Littorinimorpha, and the family Muricidae in the order Neogastropoda (Cunha et al., 2009). The relative position of the *trnW* gene in these families differs from that of *A. boucheti* and *I. nautilei*, which is located between the *trnC* and *trnQ* genes, and in the latter is found between the *trnM* and *trnY* genes (Fig. 7).

4.2 *Phylogenetic analyses*

The phylogenetic relationships within Caenogastropoda remain unclear due to differences in results between studies (Ponder et al., 2008). In this study, a large amount of sequence data (mitochondrial genome) was added to resolve the remaining taxonomic controversies in the Caenogastropoda phylogeny (Osca et al., 2015). For classification, our conclusion is consistent with the traditional taxonomic view that the subclass Caenogastropoda can be roughly divided into four orders: Littorinimorpha, Architaenioglossa, Sorbeoconcha, and Neogastropoda, com-

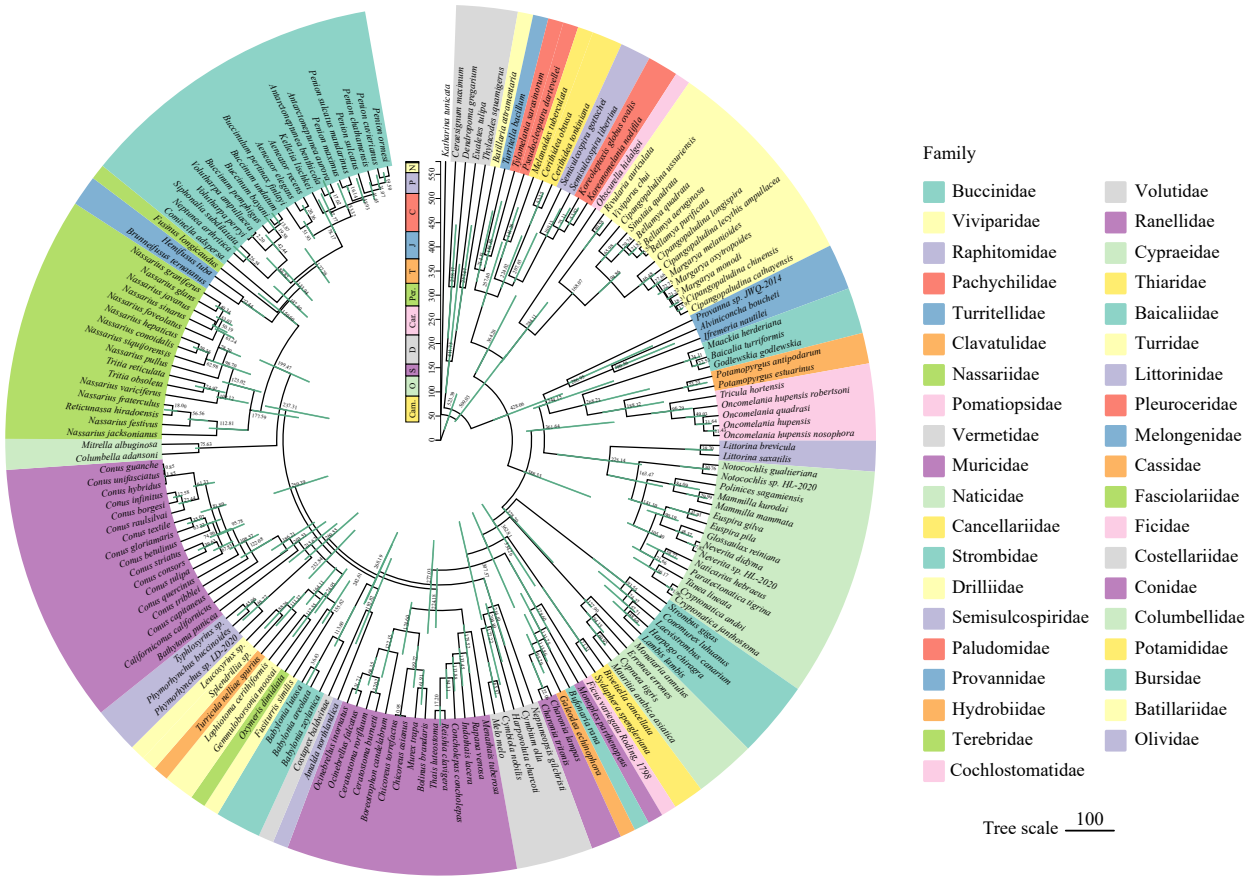

Fig. 6. Divergence time estimates for 171 Caenogastropoda species inferred using the Bayesian Relaxation Dating Method (BEAST) based on the nucleotide sequences of 13 PCGs. The divergence time is estimated using four calibration points.

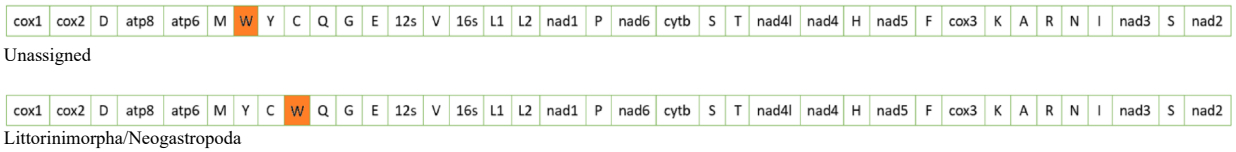

Fig. 7. The mitochondrial genome composition and arrangement of main lineages of Littorinimorpha and Neogastropoda.

prised of 17 superfamilies and 38 families. The results of molecular phylogenetic analysis were consistent with those of previous study on Caenogastropoda (Sun et al., 2012). This study supports the Caenogastropoda subclass as an independent branch at the molecular level, based on a DNA barcoding approach of the COI gene. In the phylogenetic tree, the genus *Volutharpa* was attributed to the family Buccinidae, and the genera *Aeneator* and *Buccinulum* had the closest relationship with high node support rate, similar to that reported by Vaux et al. (2018). Nassariidae is a large family of Gastropoda, and the external shapes of shells of different species are too similar between species to make distinctions (Wilke and Falniowski, 2001), whereas the phylogenetic analysis of mitochondrial sequences can provide a highly accurate classification relationship. However, this includes an unclassified order. *Alviniconcha boucheti* and *I. nautilei* are two species of the superfamily Abyssochrysoidea. According to the phylogenetic tree, they are closely related to the Littorinoidea and Truncatelloidea superfamilies of Littorinimorpha.

Based on our findings, the divergence time between genera *Nassarius* and *Tritia* is 34.33 Ma, which is consistent with the research results of Yang et al. (2019). Most species of the family Muricidae formed during the Cretaceous period (Sohl, 1969), after which Neogastropoda shellfish infiltrated the majority of oceans worldwide (Ponder, 1973). There was a diversification of Caenogastropoda's crown group of ~238 Ma, and a radiation pattern (accelerated rates of diversification) was detected between 172 Ma and 140 Ma when the main superfamilies of derived Caenogastropoda originated. Time-calibrated phylogenies showed that Caenogastropoda diverged throughout the Triassic epoch. Following the disappearance of Paleozoic biota, the Triassic period was the transitional period for the formation of modern biota (Carter et al., 2001). Thus, there have been several evolutionary changes in marine invertebrate groups. Although the rock traces of this period are distinct, the precise beginning and end time cannot be determined, as with other estimations of ancient geological times. Therefore, our research does not sufficiently address all aspects.

5 Conclusions

In this study, the mitochondrial genome sequences of *V. ampullacea*, *M. albuginosa*, *Mauritia arabica asiatica* and *E. errones* were obtained by next-generation sequencing. Their lengths

were 16 177 bp, 16 244 bp, 16 926 bp and 15 422 bp respectively. Each mitogenome consisted of 2 rRNAs, 13 PCGs, and 22 tRNAs, which were highly conserved. It was inaccurate to assume that species classification of Mollusca was judged by appearance alone, so further molecular phylogenetic studies were urgently needed. The phylogenetic tree contributed to the scientific classification of Caenogastropoda. This study provided an effective basis for the genetic characteristics and evolution of this subclass. These four species emerged at different times and their evolution may be linked to geological events that altered their habitat.

References

Bankevich A, Nurk S, Antipov D, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology, 19(5): 455–477, doi: 10.1089/cmb.2012.0021

Bernt M, Donath A, Jühling F, et al. 2013. MITOS: improved *de novo* metazoan mitochondrial genome annotation. Molecular Phylogenetics and Evolution, 69(2): 313–319, doi: 10.1016/j.ympev.2012.08.023

Bittner A. 1912. Brachiopods from the triassic of the bakony forest. Resultate der Wissenschaftlichen Erforschung des Balatonsees, II Band: Paläontologie der Umgebung des Balatonsees, 1(1): 1–60

Bridge D, Cunningham C W, Schierwater B, et al. 1992. Class-level relationships in the phylum Cnidaria: evidence from mitochondrial genome structure. Proceedings of the National Academy of Sciences of the United States of America, 89(18): 8750–8753, doi: 10.1073/pnas.89.18.8750

Carter A, Roques D, Bristow C, et al. 2001. Understanding Mesozoic accretion in Southeast Asia: significance of Triassic thermotectonism (Indosinian orogeny) in Vietnam. Geology, 29(3): 211–214, doi: 10.1130/0091-7613(2001)029<0211:UMAISA>2.0.CO;2

Colgan D J, Ponder W F, Beacham E, et al. 2007. Molecular phylogenetics of Caenogastropoda (Gastropoda: Mollusca). Molecular Phylogenetics and Evolution, 42(3): 717–737, doi: 10.1016/j.ympev.2006.10.009

Cunha R L, Grande C, Zardoya R. 2009. Neogastropod phylogenetic relationships based on entire mitochondrial genomes. BMC Evolutionary Biology, 9: 210, doi: 10.1186/1471-2148-9-210

Darriba D, Taboada G L, Doallo R, et al. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics, 27(8): 1164–1165, doi: 10.1093/bioinformatics/btr088

Feng Jiantong, Xia Liping, Yan Chengrui, et al. 2021. Characterization of four mitochondrial genomes of family Neritidae (Gastropoda: Neritimorpha) and insight into its phylogenetic relationships. Scientific Reports, 11(1): 11748, doi: 10.1038/s41598-021-91313-0

Foster W J, Garvie C L, Weiss A M, et al. 2020. Resilience of marine invertebrate communities during the early Cenozoic hyperthermals. Scientific Reports, 10(1): 2176, doi: 10.1038/s41598-020-58986-5

Gomes-dos-Santos A, Lopes-Lima M, Castro L F C, et al. 2020. Molluscan genomics: the road so far and the way forward. Hydrobiologia, 847(7): 1705–1726, doi: 10.1007/s10750-019-04111-1

Grande C, Templado J, Zardoya R. 2008. Evolution of gastropod mitochondrial genome arrangements. BMC Evolutionary Biology, 8: 61, doi: 10.1186/1471-2148-8-61

Greiner S, Lehwark P, Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3. 1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Research, 47(W1): W59–W64, doi: 10.1093/nar/gkz238

He Chongbo, Wang Jian, Gao Xianggang, et al. 2011. The complete mitochondrial genome of the hard clam *Meretrix meretrix*. Molecular Biology Reports, 38(5): 3401–3409, doi: 10.1007/s11033-010-0449-8

Huelsenbeck J P, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics, 17(8): 754–755, doi: 10.1093/bioinformatics/17.8.754

Kearse M, Moir R, Wilson A, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12): 1647–1649, doi: 10.1093/bioinformatics/bts199

Kosnik M A. 2005. Changes in Late Cretaceous-early Tertiary benthic marine assemblages: analyses from the North American coastal plain shallow shelf. Paleobiology, 31(3): 459–479, doi: 10.1666/0094-8373(2005)031[0459:CILCTB]2.0.CO;2

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, doi: 10.1093/molbev/msw054

Lee S Y, Lee H J, Kim Y K. 2019. Comparative analysis of complete mitochondrial genomes with Cerithioidea and molecular phylogeny of the freshwater snail, *Semisulcospira gottschei* (Caenogastropoda, Cerithioidea). International Journal of Biological Macromolecules, 135: 1193–1201, doi: 10.1016/j.ijbiomac.2019.06.036

Linnaeus C. 1758. Naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Stockholm: Holmiae

Meyer C P. 2003. Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. Biological Journal of the Linnean Society, 79(3): 401–459, doi: 10.1046/j.1095-8312.2003.00197.x

Middendorff A T. 1848. Vorläufige Anzeige einiger neuen Konchylien aus den Geschlechtern: littorina, Tritonium, Bullia, Natica und Margarita. Bulletin de la Classe Physico-Mathématique de l'Académie Impériale des Sciences de Saint-Pétersbourg, 7(16): 241–246

Navajas M, Le Conte Y, Solignac M, et al. 2002. The complete sequence of the mitochondrial genome of the honeybee ectoparasite mite *Varroa destructor* (Acari: Mesostigmata). Molecular Biology and Evolution, 19(12): 2313–2317, doi: 10.1093/oxfordjournals.molbev.a004055

Osca D, Templado J, Zardoya R. 2014. The mitochondrial genome of *Ifremeria nautilei* and the phylogenetic position of the enigmatic deep-sea Abyssochrysoidea (Mollusca: Gastropoda). Gene, 547(2): 257–266, doi: 10.1016/j.gene.2014.06.040

Osca D, Templado J, Zardoya R. 2015. Caenogastropod mitogenomics. Molecular Phylogenetics and Evolution, 93: 118–128, doi: 10.1016/j.ympev.2015.07.011

Peretolchina T E, Sitnikova T Y, Sherbakov D Y. 2020. The complete mitochondrial genomes of four Baikal molluscs from the endemic family Baicaliidae (Caenogastropoda: Truncatelloida). Journal of Molluscan Studies, 86(3): 201–209, doi: 10.1093/mollus/eyaa004

Perna N T, Kocher T D. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. Journal of Molecular Evolution, 41(3): 353–358, doi: 10.1007/BF00186547

Ponder W F. 1973. The origin and evolution of the Neogastropoda. Malacologia, 12(2): 295–338

Ponder W F, Colgan D J, Healy J M, et al. 2008. Caenogastropoda. In: Ponder W, ed. Phylogeny and Evolution of the Mollusca. Berkeley: University of California Press, 331–383

Pu Deqiang, Liu Hongling, Gong Yiyun, et al. 2017. Mitochondrial genomes of the hoverflies *Episyrphus balteatus* and *Eupeodes corollae* (Diptera: Syrphidae), with a phylogenetic analysis of Muscomorpha. Scientific Reports, 7: 44300, doi: 10.1038/srep44300

Reeve LA. 1859. Monograph of the Genus Columbella. In: Conchologia Iconica, or, illustrations of the shells of molluscous animals. 11th ed. London: Lovell Reeve and Company, 1-37

Schilder F A, Schilder M. 1939. Prodrome of a monograph on living Cypraeidae. Journal of Molluscan Studies, 23(4): 181–231, doi:

10.1093/oxfordjournals.mollus.a064355
Sigwart J D, Lindberg D R, Chen C, et al. 2021. Molluscan phylogenomics requires strategically selected genomes. Philosophical Transactions of the Royal Society B:Biological Sciences, 376(1825): 20200161, doi: 10.1098/rstb.2020.0161
Simone L R L. 2004. Comparative morphology and phylogeny of representatives of the superfamilies of architaenioglossans and the Annulariidae (Mollusca, Caenogastropoda). Arquivos do Museu National, Rio de Janeiro, 62(4): 387–504
Simone L R L. 2005. Comparative morphological study of representatives of the three families of Stromboidea and the Xenophoroidea (Mollusca, Caenogastropoda), with an assessment of their phylogeny. Arquivos de Zoologia, 37(2): 141–267, doi: 10.11606/issn.2176-7793.v37i2p141-267
Sohl N F. 1969. The fossil record of shell boring by snails. American Zoologist, 9(3): 725–734, doi: 10.1093/icb/9.3.725
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22(21): 2688–2690, doi: 10.1093/bioinformatics/btl446
Sun Yan, Li Qi, Kong Lingfeng, et al. 2012. DNA barcoding of Caenogastropoda along coast of China based on the COI gene. Molecular Ecology Resources, 12(2): 209–218, doi: 10.1111/j.1755-0998.2011.03085.x
Vaux F, Trewick S A, Crampton J S, et al. 2018. Evolutionary lineages of marine snails identified using molecular phylogenetics and geometric morphometric analysis of shells. Molecular Phylogenetics and Evolution, 127: 626–637, doi: 10.1016/j.ympev.2018.06.009
Wilke T, Falniowski A. 2001. The genus *Adriohydrobia* (Hydrobiidae: Gastropoda): polytypic species or polymorphic populations? Journal of Zoological Systematics and Evolutionary Research, 39(4): 227–234,doi: 10.1046/j.1439-0469.2001.00171.x
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(1): 38–47, doi: 10.1016/j.gene.2013.10.005
Wolstenholme D R. 1992. Animal mitochondrial DNA: structure and evolution. International Review of Cytology, 141: 173–216, doi: 10.1016/S0074-7696(08)62066-5
Xu Xiaodong, Wu Xiangyun, Yu Ziniu. 2012. Comparative studies of the complete mitochondrial genomes of four *Paphia* clams and reconsideration of subgenus *Neotapes* (Bivalvia: Veneridae). Gene, 494(1): 17–23, doi: 10.1016/j.gene.2011.12.002
Yang Huirong, Zhang Jiaen, Xia Jun, et al. 2018. Comparative characterization of the complete mitochondrial genomes of the three Apple Snails (Gastropoda: Ampullariidae) and the phylogenetic analyses. International Journal of Molecular Sciences, 19(11): 3646, doi: 10.3390/ijms19113646
Yang Yi, Li Qi, Kong Lingfeng, et al. 2019. Mitogenomic phylogeny of *Nassarius* (Gastropoda: Neogastropoda). Zoologica Scripta, 48(3): 302–312, doi: 10.1111/zsc.12343
Zinsmeister W J, Feldmann R M, Woodburne M O, et al. 1989. Latest Cretaceous/earliest Tertiary transition on Seymour Island, Antarctica. Journal of Paleontology, 63(6): 731–738, doi: 10.1017/S0022336000036453

Supplementary information:

Table S1. Results of alignment of protein-coding sequences for all species in the phylogenetic tree of this study.

The supplementary information is available online at https://doi.org/10.1007/s13131-023-2258-7 and http://www.aosocean.com/. The supplementary information is published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.