Journal of
Molluscan Studies

The Malacological Society of London

Journal of Molluscan Studies (2022) **88**: eyac009. https://doi.org/10.1093/mollus/eyac009
Published online 27 April 2022

The mitogenome of the sunken wood limpet *Notocrater youngi*: insights into mitogenome evolution in Lepetellida (Gastropoda: Vetigastropoda)

Juan E. Uribe[1,2], Makiri Sei[2] and M. G. Harasewych[2]

[1]*Biodiversity and Evolutionary Biology Department (BEBD), Museo Nacional de Ciencias Naturales (MNCN-CSIC), Madrid, Spain; and*
[2]*Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, PO Box 37012, Washington, DC 20013-7012, USA*

Correspondence: J.E. Uribe; e-mail: juanuribearboleda@gmail.com

(Received 17 September 2021; editorial decision 24 November 2021)

ABSTRACT

The complete mitochondrial genome of the pseudococculinid limpet *Notocrater youngi* was determined using Illumina sequencing and found to be a double-stranded circular molecule 15,915 bp in length. It contains the usual 13 protein-coding genes, 2 ribosomal RNA genes and 22 transfer RNA genes, although with a gene order that differs from those of all other presently known Vetigastropoda, and that is derived within Lepetellida. Phylogenetic analyses based on amino acid sequences join the three pseudococculinid species in a clade and confirm the well-supported clade Scissurelloidea + (Lepetelloidea + Lepetodriloidea) but not a monophyletic Lepetellida as currently understood. Patterns of gene order rearrangements are concordant with these findings.

INTRODUCTION

Deep-sea limpets associated with biogenic substrates (e.g. wood, algal holdfasts, whale and fish bones, cephalopod beaks, crab carapaces and shark egg cases) had initially been grouped within the suborder Cocculiniformia, containing superfamilies Cocculinoidea (with two families) and Lepetelloidea (with eight families), based on morphological characters (Haszprunar, 1987, 1988a,b; McLean & Harasewych, 1995). Subsequent studies using anatomical and, increasingly, molecular characters (e.g. Ponder & Lindberg, 1997; McArthur & Harasewych, 2003; Lee *et al.*, 2019) have partitioned these limpets into two distinct lineages, the Cocculinoidea, a member of the subclass Neomphaliones, and the Lepetelloidea, one of five superfamilies within the order Lepetellida, a member of the subclass Vetigastropoda (Bouchet *et al.*, 2017).

There have been numerous investigations of phylogenetic relationships within Vetigastropoda, based on morphological characters (Ponder & Lindberg, 1997; Sasaki, 1998) and various molecular data (e.g. Geiger & Thacker, 2005; Aktipis *et al.*, 2008; Kano, 2008; Aktipis & Giribet, 2011), including mitogenomes (e.g. Lee *et al.*, 2016; Uribe *et al.*, 2016, 2017; Wort, Fenberg & Williams, 2017; Zhou *et al.*, 2019) and transcriptomic data (Cunha, Reimer & Giribet, 2021). However, phylogenetic hypotheses of relationships among the orders, superfamilies and families of Vetigastropoda remain dynamic, with tree topologies varying depending on taxon sampling, types of data and methods of data analysis. As currently understood, the subclass Vetigastropoda is comprised of the orders Pleurotomariida, Seguenziida, Trochida (each represented by a single superfamily in the Recent fauna) and Lepetellida (into which the superfamilies Lepetelloidea, Fissurelloidea, Haliotoidea, Lepetodriloidea and Scissurelloidea have been assigned) (Bouchet *et al.*, 2017). Mitogenomic information of the family Lepetelloidea is presently limited to a partial mitogenome (without entire gene order) from a single taxon (as Pseudococculinidae sp.; Lee *et al.*, 2019).

In the present study, we report the complete mitochondrial (mt) genome of *Notocrater youngi* McLean & Harasewych, 1995 (family Pseudococculinidae), the first complete mitogenome to be determined for any species within the superfamily Lepetelloidea. Features, such as genome size and gene order, are compared to those of representatives of other Vetigastropoda and Neomphaliones. The phylogenetic relationship of *Notocrater* within Vetigastropoda is inferred using mitogenome data from representatives of all lepetellid superfamilies and the main vetigastropod clades that are available in GenBank. Analyses include both maximum likelihood (ML) and Bayesian inference (BI) analyses of amino acid sequences of the protein-coding genes (PCGs) using probabilistic methods with two evolutionary replacement models (BI: site-heterogeneous model; ML: mixture model) in order to avoid a potential long branch attraction (LBA) bias generated by the accelerated mutation rates displayed in fissurelloid mitogenomes. Differences in gene order are compared, rearrangements among and within vetigastropod superfamilies are inferred, and patterns are evaluated against sequence-based phylogenetic hypotheses.

MATERIAL AND METHODS

DNA extraction

The foot and buccal muscles (*c.* 30 mg) were dissected from a single specimen of *Notocrater youngi* (National Museum of Natural History,

Table 1. List of taxa used in phylogenetic analyses, their GenBank accession numbers, source of sequence data and size of the mitogenome.

Order	Superfamily	Taxon	GenBank acc. no.	Source	Length (bp)
Subclass VETIGASTROPODA					
Lepetellida	Lepetelloidea	Pseudococculinidae sp.	MH837540	Lee *et al.* (2019)	15,580*
Lepetellida	Lepetelloidea	*Amphiplica gordensis*	SRR1505101	Zapata *et al.* (2014)	14,345†
Lepetellida	Lepetelloidea	*Notocrater youngi*	MT360647	This study	15,915
Lepetellida	Lepetodriloidea	*Lepetodrilus guaymasensis*	MH837537	Lee *et al.* (2019)	17,260*
Lepetellida	Lepetodriloidea	*Lepetodrilus nux*	LC107880	Nakajima *et al.* (2016)	16,353*
Lepetellida	Lepetodriloidea	*Lepetodrilus schrolli*	KR297250	Uribe *et al.* (2016)	15,579*
Lepetellida	Lepetodriloidea	*Pseudorimula* sp.	MK404176	Zhou *et al.* (2019)	16,682
Lepetellida	Scissurelloidea	*Anatoma* sp.	MH837531	Lee *et al.* (2019)	13,668*
Lepetellida	Fissurelloidea	*Diodora graeca*	KT207825	Uribe *et al.* (2016)	17,209*
Lepetellida	Fissurelloidea	*Fissurella volcano*	NC_016,953	Simison (unpubl.)	17,575
Lepetellida	Haliotoidea	*Haliotis discus hannai*	EU595789	Ren *et al.* (unpubl.)	15,784*
Lepetellida	Haliotoidea	*Haliotis rubra*	NC_005940	Maynard *et al.* (2005)	16,907
Trochida	Trochoidea	*Bolma rugosa*	NC_029366	Uribe *et al.* (2016)	17,432
Trochida	Trochoidea	*Gibbula umbilicaris*	KY205707	Uribe *et al.* (2017)	13,266*
Trochida	Trochoidea	*Margarites vorticiferus*	KY205708	Uribe *et al.* (2017)	15,253*
Seguenziida	Seguenzioidea	*Calliotropis micraulax*	MH837534	Lee *et al.* (2019)	17,140*
Seguenziida	Seguenzioidea	*Granata lyrata*	NC_028708	Uribe *et al.* (2016)	17,632
Pleurotomariida	Pleurotomarioidea	*Bayerotrochus teramachii*	MH837533	Lee *et al.* (2019)	13,473*
Pleurotomariida	Pleurotomarioidea	*Perotrochus caledonicus*	MH837539	Lee *et al.* (2019)	14,082*
Subclass NEOMPHALIONES					
Neomphalida	Neomphaloidea	*Chrysomallon squamiferum*	AP013032	Nakagawa *et al.* (2014)	15,388
Neomphalida	Neomphaloidea	*Peltospira smaragdina*	MH837538	Lee *et al.* (2019)	15,112*
Cocculinida	Cocculinoidea	*Coccocrater* sp.	MH837535	Lee *et al.* (2019)	13,983*
Cocculinida	Cocculinoidea	*Cocculina subcompressa*	MH837536	Lee *et al.* (2019)	18,167*

Taxon names as well as their superfamily and family assignments have been updated according to WoRMS (2021).
*Only a partial mitogenome sequence was available.
†Partial mitogenome assembled from transcriptome files.

Smithsonian Institution, Washington, DC, USA; reg. no. USNM 1410123), taken from a sunken wooden branch collected in Prince Rupert Bay, Portsmouth, Dominica (15°33.558′N, 61°28.284′W) at a depth of 304 m by the submersible CURASUB, 8 March 2016, and preserved in 95% ethanol. Genomic DNA was extracted using the Qiagen DNeasy kit (Qiagen, Valencia, CA, USA) using the manufacturer's spin-column protocol for animal tissues, with the initial lysis step extended to 24 h at 56 °C with continuous agitation.

Sequencing, assembly and annotation

A portion of the cytochrome *c* oxidase subunit I (COX1) gene was PCR-amplified and Sanger-sequenced using the primers and protocols in Harasewych (2019) to serve as the initial reference sequence for assembly of the mitogenome. DNA libraries were constructed and sequenced on an Illumina MiSeq (Illumina, San Diego, CA, USA) with MiSeq Reagent Kit v2 at the Smithsonian National Museum of Natural History's Laboratories of Analytical Biology following previously published protocols (Harasewych *et al.*, 2019).

Low-quality reads were removed using TrimGalore v. 0.6.3 (https://github.com/FelixKrueger/TrimGalore), with thresholds for minimum Phred scores set to 20 and minimum read lengths set to 20 bp. The mitogenome was assembled using the "map to reference" feature of Geneious® v. 10.2.3 (http://www.geneious.com) with "minimum overlap identity" set to 99% and "minimum overlap" set to 40 bp, with the barcode region of the COX1 gene determined by Sanger sequencing as the initial reference sequence. Mitochondrial elements were annotated using MITOS (Bernt *et al.*, 2013), ARWEN v. 1.2 (Laslett & Canbäck, 2008) and the ORF finder in Geneious®. The size and sequence of the large intergenic region between the tRNA-M and COX3 genes were confirmed using standard PCR and Sanger sequencing with primers designed from the flanking regions within the 12S rRNA (ribosomal RNA) (NO-M-12S-R 5′-ACCAGAACCCATAAAGTACCATGTG-3′) and COX3 (NO-M-COX3-R 5′-GGGCTAAAGTCAACTAAATGATAGGG-3′) genes.

Sequence alignment and phylogenetic analyses

Phylogenetic analyses were conducted using data from the complete mitogenome of *N. youngi* as well as from complete and partial mitogenomes of 17 species representing all orders within the subclass Vetigastropoda and 4 species of the subclass Neomphaliones (selected to serve as the outgroup), which were available in GenBank (Table 1). Also included in the analyses was the partial mitogenome of *Amphiplica gordensis* McLean, 1991 that was reconstructed from transcriptome data (Zapata *et al.*, 2014) downloaded from the SRA repository. These data were assembled *de novo* using Trinity v. 2.8.5 (Grabherr *et al.*, 2011) with the Trimmomatic option (Bolger, Lohse & Usadel, 2014). Mitochondrial genes were identified following the protocol in Abalde *et al.*'s (2019) study by using a BLASTn search against a database of all gastropod mitogenomes available in GenBank assembled with BLAST v. 2.6.0 (Altschul *et al.*, 1997).

Homologous PCGs were processed at the amino acid level using PREQUEL v. 1.01 (Whelan, Irisarri & Burki, 2018) with default parameters in order to identify and mask regions with nonhomologous adjacent characters, and then aligned using MAFFT v. 7.407 (Katoh, Rosewicki & Yamada, 2019). Informative regions were then filtered from the raw alignments using BMGE (Criscuolo & Gribaldo, 2010) with default parameters, the matrixes concatenated

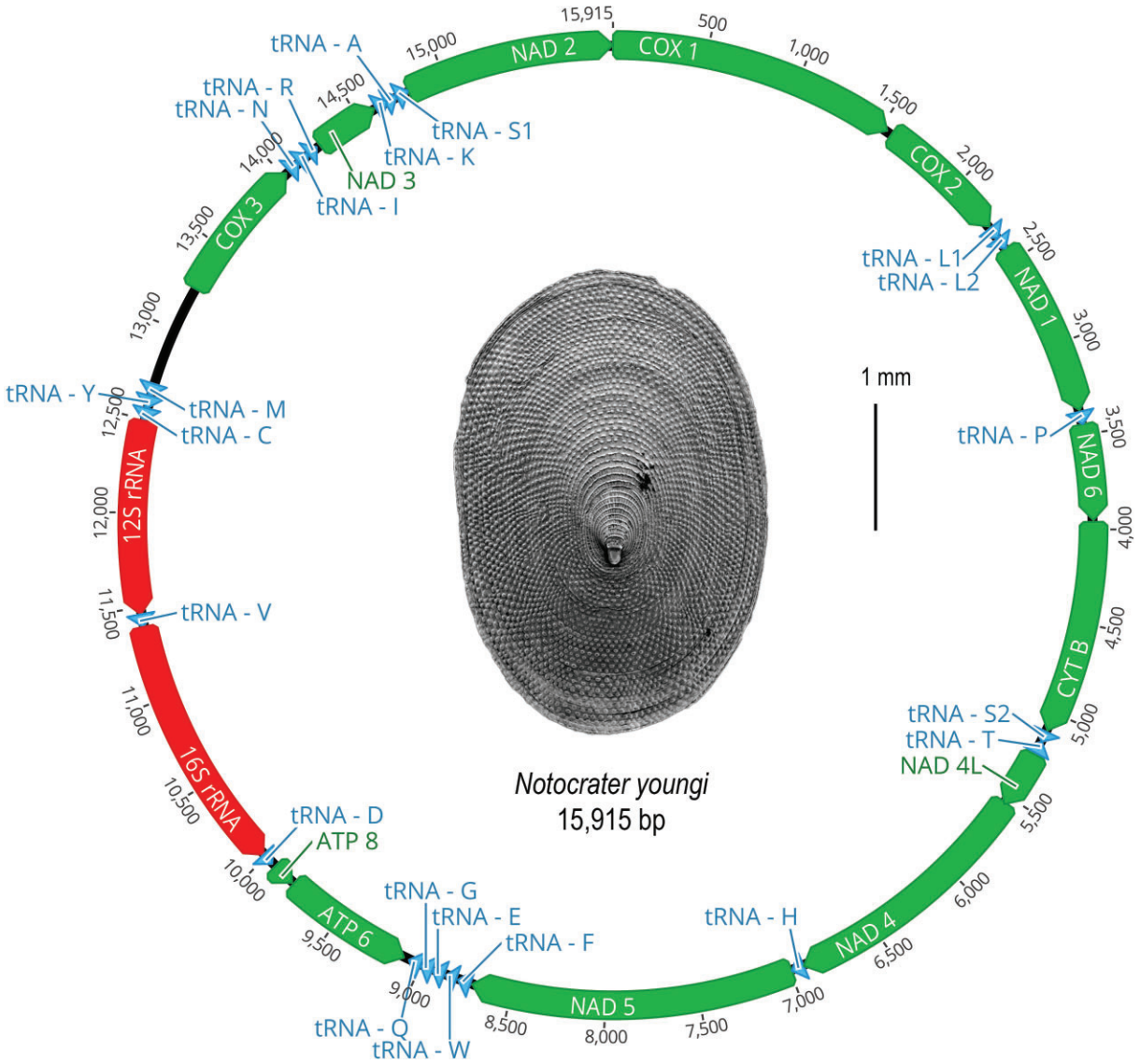

Figure 1. Map of the mitochondrial genome of *Notocrater youngi* (GenBank acc. no. MT360647). Arrows indicate the direction of transcription. PCGs are in green, rRNA genes in red and tRNA genes in blue.

at the amino acid level, and then used to infer the phylogenetic relationships within Vetigastropoda using probabilistic methods. Site-heterogeneous (Lartillot & Philippe, 2004) and mixture models (Le, Gascuel & Lartillot, 2008) have previously been used successfully to counteract LBA within Gastropoda (Uribe *et al.*, 2019) and were used, respectively, for BI- and ML-based tree reconstruction.

Unpartitioned matrices were analysed using PhyloBayes MPI v. 1.5a (Lartillot *et al.*, 2013), applying the GTR-CAT heterogeneity model on two independent MCMC (Markov chain Monte Carlo) chains, sampling every cycle, until convergence (checked *a posteriori* using the tracecomp tool: max-diff <0.1, maximum discrepancy <0.1 and effective sample size >100). ML analyses were run using IQ-TREE v. 1.6.1 (Nguyen *et al.*, 2014) using the best-fit mixture model (mtZOA + F + C50 + R5) selected using ModelFinder (Kalyaanamoorthy *et al.*, 2017) implemented with the option '-m TESTONLYNEW', under the Bayesian information criterion (Schwarz, 1978). A combination of rapid hill-climbing and stochastic perturbation methods was used and included a total of 1,000 pseudo-replicates of ultrafast bootstrap to assess robustness of the inferred tree. We considered branches with Bayesian posterior probabilities (PP) ≥0.9 or bootstrap support (BS) values ≥75 to be strongly supported.

Mitochondrial gene order

We used the program CREx (Bernt *et al.*, 2007) to infer the most parsimonious rearrangement events within the mitogenomes of Vetigastropoda concordant with our phylogenetic hypothesis.

RESULTS

Mitogenome content and organization

The Illumina sequencing run produced 3,203,700 paired reads ranging in length from 35 to 251 bp, with an average length of 247.3 (SD = 13.6). Of these, 3,648 reads mapped to the mt genome (coverage 55.7; SD = 12.6).

The mt genome of *Notocrater youngi* reconstructed from these data (GenBank acc. no. MT360627) is a double-stranded circular

Table 2. Position (start and stop), strand direction (+/−) and length (bp) of the genes in the mt genome of *Notocrater youngi* (15,915 bp; GenBank MT360647), with initiation and termination codons (init./term.) and amino acid sequence lengths (laa) for the PCGs.

Gene	Start	Stop	Strand direction	Length (bp)	Init./term.	Laa	Ign
COX1	1	1,536	+	1,536	ATG/TAA	511	34
COX2	1,571	2,278	+	708	ATG/TAA	235	9
tRNA-Leu (L1) (tag)	2,288	2,353	+	66			2
tRNA-Leu (L2) (taa)	2,356	2,420	+	65			0
NAD1	2,421	3,368	+	948	ATG/TAG	315	6
tRNA-Pro (P) (tgg)	3,375	3,440	+	66			1
NAD6	3,442	3,954	+	513	ATG/TAA	170	7
CYT B	3,962	5,098	+	1,137	ATG/TAA	378	5
tRNA-Ser (S2) (tga)	5,104	5,167	+	64			1
tRNA-Thr (T) (tgt)	5,236	5,169	–	68			8
NAD4L	5,245	5,547	+	303	ATG/TAA	100	−7
NAD4	5,541	6,925	+	1,385	ATG/TA(A)	461	0
tRNA-His (H) (gtg)	6,926	6,991	+	66			1
NAD5	6,993	8,724	+	1,732	ATG/T(AA)	577	0
tRNA-Phe (F) (gaa)	8,725	8,792	+	68			5
tRNA-Trp (W) (tca)	8,798	8,863	+	66			2
tRNA-Glu (E) (ttc)	8,866	8,931	+	66			2
tRNA-Gly (G) (tcc)	8,934	9,000	+	67			0
tRNA-Gln (Q) (ttg)	9,001	9,069	+	69			32
ATP6	9,797	9,102	–	696	ATG/TAG	231	16
ATP8	9,933	9,814	–	120	ATG/TAA	39	20
tRNA-Asp (D) (gtc)	10,021	9,954	–	68			0
16S rRNA	11,395	10,022	–	1,374			0
tRNA-Val (V) (tac)	11,461	11,396	–	66			0
12S rRNA	12,491	11,462	–	1,030			0
tRNA-Cys (C) (gca)	12,558	12,492	–	67			5
tRNA-Tyr (Y) (gta)	12,629	12,564	–	66			0
tRNA-Met (M) (cat)	12,693	12,630	–	64			553
COX3	13,247	14,026	+	780	ATG/TAA	259	18
tRNA-Asn (N) (gtt)	14,045	14,109	+	65			-2
tRNA-Ile (I) (gat)	14,108	14,174	+	67			0
tRNA-Arg (R) (tcg)	14,175	14,238	+	64			16
NAD3	14,255	14,605	+	351	ATG/TAG	116	3
tRNA-Lys (K) (ttt)	14,609	14,677	+	69			−5
tRNA-Ala (A) (tgc)	14,673	14,737	+	65			1
tRNA-Ser (S1) (gct)	14,739	14,805	+	67			2
NAD2	14,808	15,915	+	1,108	ATT/T(AA)	369	0

Standard abbreviations for PCGs are used. Both three- and one-letter abbreviations are listed for tRNA genes, along with the codon used. The numbers of intergenic nucleotides (ign) are shown. Negative numbers indicate gene overlaps.

molecule 15,915 bp in length. It is composed of 33.4% A, 33.8% T, 13.3% G and 19.5% C, and contains the usual 13 PCGs, 2 rRNA genes and 22 transfer RNA (tRNA) genes (Fig. 1, Table 2). Twenty-seven of these genes are coded on the heavy (+) strand and ten genes, including tRNA-Thr (T), as well as a segment containing ATP6, ATP8, tRNA-Asp (D), 16S rRNA, tRNA-Val (V), 12S rRNA and the tRNA genes for Cys (C), Tyr (Y) and Met (M) are coded on the light (−) strand.

PCGs span 11,317 bp (71.1%), rRNA genes span 2,404 bp (15.1%) and tRNA genes span 1,459 bp (9.2%) of the mitogenome. The most common start codon is ATG, occurring in all PCGs except NAD2, which uses the ATT start codon. Most PCGs end with the TAA stop codon, except for the NAD1, NAD3 and ATP6 genes, which end with the TAG codon, and the NAD2, NAD4 and NAD5 genes, which end with the truncated stop codon T that is completed by post-transcriptional polyadenylation (Donath *et al.*, 2019). The mitogenome includes 23 intergenic regions (749 bp in total; 4.7% of the genome), which range in size from 1 to 553 bp; the largest, between tRNA-M and COX3, includes the potential origin of replication due to its high A + T content (80.1%). Three overlapping regions (14 bp in total) are interspersed throughout the genome.

Phylogenetic analyses

Phylogenetic analyses of concatenated amino acid sequences of PCGs using both ML and BI each produced a single tree in which most branches were well supported. The two trees were congruent except for the relative positions of Trochida, Fissurelloidea and a clade consisting of Scissurelloidea + (Lepetodriloidea + Lepetelloidea) (Fig. 2). The BI tree indicated an unresolved trichotomy of these three lineages (Fig. 2A), while the ML analyses placed Fissurelloidea as sister to Trochida (Fig. 2B) with low support (BS = 32%), and Fissurelloidea + Trochida as sister to the clade Scissurelloidea + (Lepetodriloidea + Lepetelloidea) also with low support (BS = 55%).

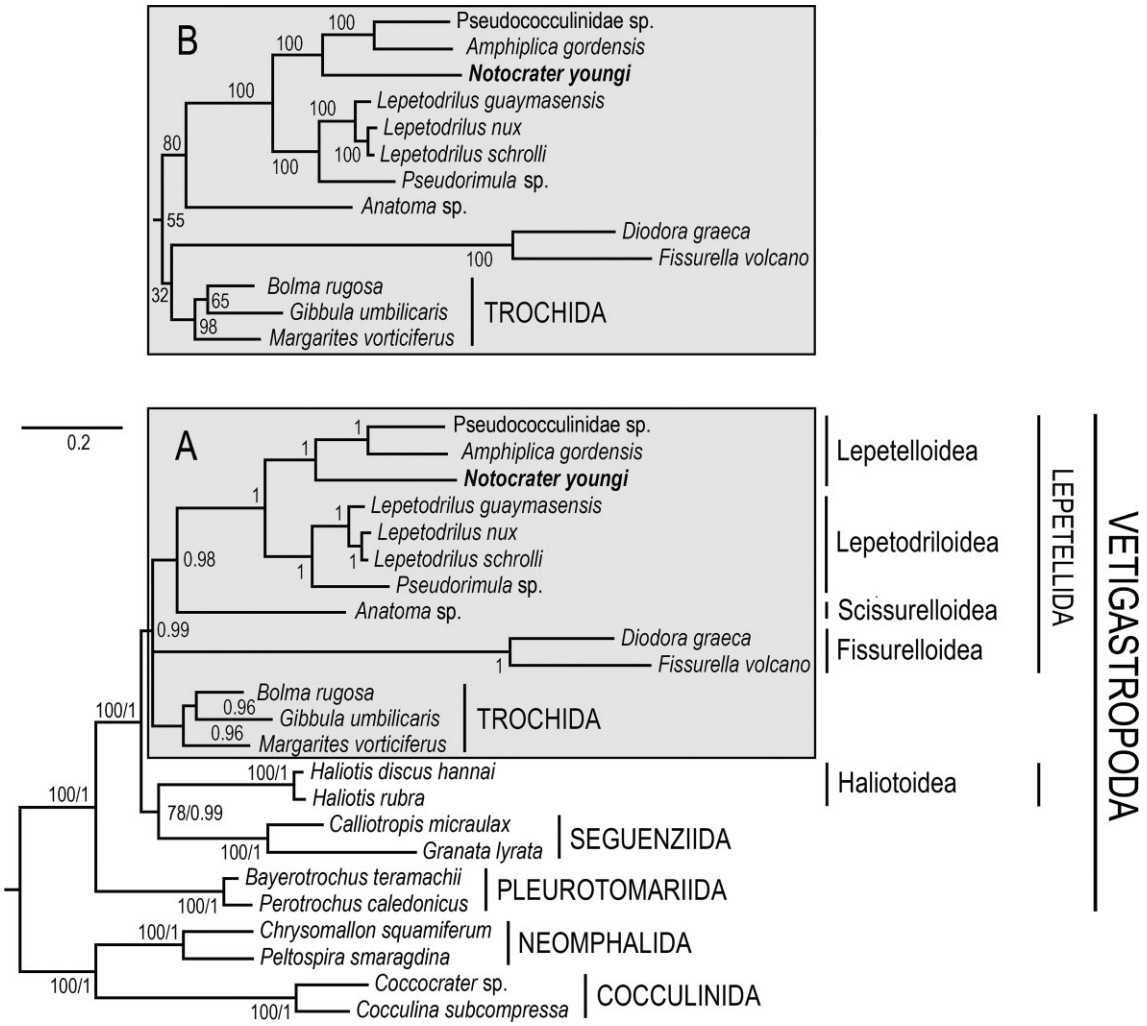

Figure 2. Phylogenetic relationships of *Notocrater youngi* based on BI (**A**) and ML (**B**) analyses of concatenated amino acid sequences of PCGs. Branch supports shown are ML bootstrap values followed by Bayesian posterior probabilities. The scale bar represents substitutions per site.

Mitochondrial gene order

The gene order of the mt genome of *N. youngi*, the first complete mitogenome to be reported for the superfamily Lepetelloidea, is congruent with those of two other lepetelloideans for which only partial mitogenomes have been recovered (Fig. 3C) except for the transposition of the tRNA-Y and tRNA-M genes.

The gene orders of most vetigastropod mitogenomes correspond to the hypothetical ancestral gene order of gastropods (Stöger & Schrödl, 2013; Uribe *et al.*, 2016: fig. 1). This ancestral gene order is retained in members of Pleurotomariida, Seguenziida and Trochida (except for tRNA rearrangements in some trochid taxa; see Uribe *et al.*, 2017: fig. 1), but has undergone considerable rearrangements within Fissurelloidea and the clade containing Scissurelloidea, Lepetodrilloidea and Lepetelloidea (Fig. 3).

Rearrangements that occurred in the common ancestor of the clade Scissurelloidea + (Lepetodriloidea + Lepetelloidea) include reversal of a genome fragment containing 16 genes (tRNA-D to tRNA-L1), in terms of both gene order and strand orientation, as well as the transposition of three tRNA genes (tRNA-Y, tRNA-G and tRNA-E). This gene order is present in all members of the Scissurelloidea and Lepetodriloidea for which mitogenome data are available (Fig. 3A). The most parsimonious rearrangement events inferred by CREx to produce these changes in gene order are (1) the reversal of the 16-gene fragment (tRNA-D to tRNA-L1), (2) the reverse transposition of a fragment containing the tRNA-Y to tRNA-E, (3) the reversal of gene order and strand orientation of the gene fragment (tRNA-E to tRNA-C), (4) the transposition of fragment (tRNA-C to tRNA-W) with fragment (tRNA-Q to tRNA-G) and (5) transposition of tRNA-Q with fragment (tRNA-G to tRNA-E), as illustrated in Supplementary Material Figure S1.

Six additional transpositions involving eight tRNA genes have occurred between Lepetodriloidea and *N. youngi*. These include two (tRNA-K + tRNA-A and tRNA-R) in the region between COX3 and NAD2, and four (tRNA-M, tRNA-G, tRNA-W + tRNA-E and tRNA-Q) in the region between NAD5 and COX3 (Fig. 3B). CREx inferred that two transpositions account for the changes in the COX3 to NAD2 region (Supplementary Material Fig. S2): (1) the transposition of tRNA-R with the fragment tRNA-N + tRNA-I and (2) the transposition of fragment tRNA-K + tRNA-A with the fragment tRNA-N to NAD3.

Two alternative rearrangement scenarios were proposed by CREx to account for the gene order changes involving the five tRNA genes in the region between NAD5 and COX3. Following an initial transposition of tRNA-W and tRNA-E, one alternative scenario involved (1) the reversal in gene order and stand orientation for a fragment containing tRNA-G to tRNA-Q, (2) the reversal in strand orientation for tRNA-C, and (3) the transposition of the fragment tRNA-M + tRNA-Q and the fragment tRNA-W to tRNA-Y, followed by (4) a tandem-duplication random loss (TDRL) event

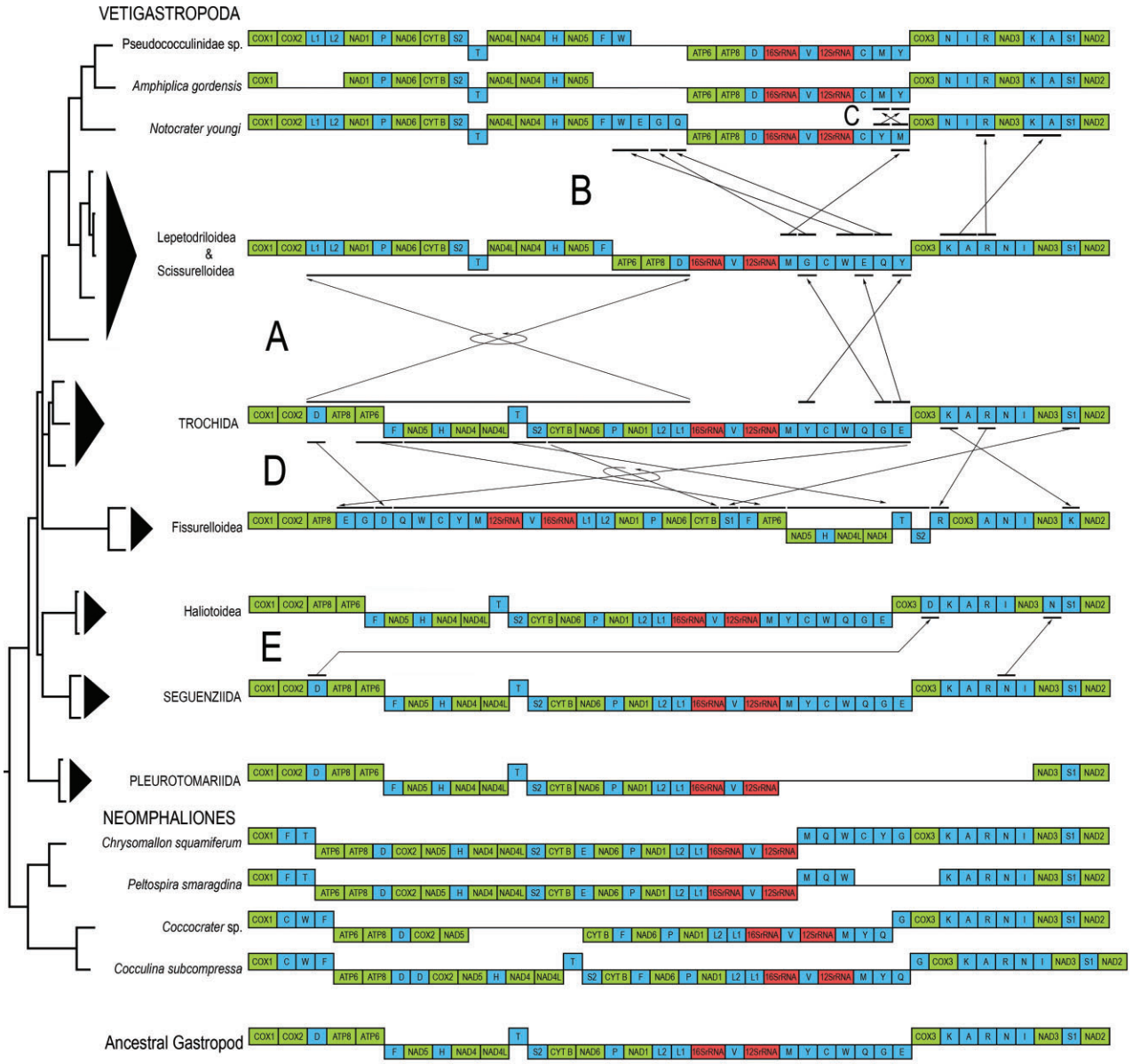

Figure 3. Linear representations of gene order in the 19 species of Vetigastropoda and 4 species of Neomphaliones used in this study shown according to their phylogenetic relationships (Fig. 2). Identical gene orders were combined to the highest taxonomic level. Rearrangements in gene order (diagonal lines with arrows) are shown in the lineages (**A**) leading to Lepetodriloidea, Scissurelloidea and Lepetelloidea; (**B**) leading to Lepetelloidea; (**C**) within Pseudococculinidae; (**D**) leading to Fissurelloidea; and (**E**) leading to Haliotoidea. The gene order of the hypothetical ancestral gastropod is shown for comparison.

within the region between ATP6 and tRNA-Q. The second alternative scenario consisted of (1) the reversal in strand orientation for tRNA-G and (2) the reversal in gene order and strand orientation for the fragment tRNA-E to tRNA-Q, followed by (3) two sequential TDRL events within the region between ATP6 and tRNA-Y (see Supplementary Material Fig. S2).

A reversal in the order of tRNA-Y and tRNA-M occurred between *N. youngi* and (*Amphiplica gordensis* + Pseudococculinidae sp.) (Fig. 3C), the result of a single transposition of these two tRNA genes recognized by CREx (Supplementary Material Fig. S3).

The mt gene order of Fissurelloidea differs substantially from the hypothetical ancestral gene order of Gastropoda (exemplified by the gene order of Trochida), and also from that of all other vetigastropods. This gene order (Fig. 3D) includes (1) the reversal of a genome fragment containing 16 genes (CYTB to tRNA-E) in terms of gene order and strand orientation (all from the light strand to the heavy strand), (2) the transposition of a genome fragment with five genes (NAD5 to tRNA-S2) retaining order and strand orientation, and (3) the reversal of a genome fragment with two genes (ATP6 + tRNA-F) in which the gene order and the strand orientation of tRNA-F are reversed. Four additional transpositions, each involving a single tRNA gene (tRNA-D, tRNA-K, tRNA-R and tRNA-S1) have occurred between the hypothetical ancestral gastropod gene order and Fissurelloidea. The transposition of tRNA-D occurred after the reversal of the fragment containing 16 genes (CYTB to tRNA-E).

CREx inferred these to be the result of a series of reversals, transpositions and TDRL events. These include (1) a reverse transposition of the ATP6-tRNA-F genes, (2) a TDRL event within the region between COX3 and NAD3, (3) a transposition of

tRNA-K within this region, (4) the reversal in direction and stand orientation of a 16-gene region between CYTB and tRNA-E, and (5) two TDRL events spanning the entire mitogenome as illustrated in Supplementary Material Figure S4.

The mt gene order of Haliotoidea differs from that of the hypothetical ancestral gene order of Gastropoda, exemplified by the gene order of Seguenziida (Fig. 3E), only in the transposition of tRNA-D from a position between COX2 and ATP8 to a position between COX3 and tRNA-K, and the transposition of tRNA-N from a position between tRNA-R and tRNA-I to a position between NAD3 and tRNA-S1. CREx shows these to be the result of two transpositions (Supplementary Material Fig. S5).

DISCUSSION

The mt genome of *Notocrater youngi* is 15,915 bp in length, the shortest among vetigastropods for which complete mitogenomes have thus far been reported (previously *Astralium haematragum*, 16,310 bp; Lee *et al.*, 2016: table 2). The genes are arranged predominantly on the heavy chain, with a smaller number of genes on the light chain (10) than in other Vetigastropoda studied, except for Fissurelloidea (5) (Fig. 3).

Phylogenetic analyses joined the three pseudococculinids in a clade and recovered a sister group relationship between Lepetelloidea and Lepetodriloidea, as well as between Scissurelloidea and Lepetelloidea + Lepetodriloidea, all with high levels of support. However, they did not recover a monophyletic Lepetellida. Both analyses showed Haliotoidea to be more closely related to Seguenziida (with high support) than to other Lepetellida. Vetigastropoda was monophyletic with high support.

The mt gene order of the Pleurotomariida, Seguenziida and Trochida is the same as the hypothetical ancestral gene order of Gastropoda (Stöger & Schrödl, 2013; Uribe *et al.*, 2016: fig. 1). Within Lepetellida, the order of the genes varies among the superfamilies. There appear to be three lineages within the Lepetellida: (1) Haliotoidea, (2) Fissurelloidea and (3) Scissurelloidea + (Lepetodriloidea + Lepetelloidea). Each of these lineages evolved independently from the hypothetical gastropod gene order through a different series of rearrangements.

The Scissurelloidea, Lepetelloidea and Lepetodriloidea all share a reversal of a gene fragment containing 16 genes (Fig. 3A) that is a synapomorphy for the clade containing these three superfamilies. Scissurelloidea and Lepetodriloidea also share transpositions of three tRNA genes, while *N. youngi* has undergone an additional six transpositions involving eight tRNA genes. As the superfamily Lepetelloidea includes eight families, it is not clear whether the changes within *Notocrater* are representative of Lepetelloidea, limited to the family Pseudococculinidae, or differently distributed among the remaining seven families. A single transposition involving two adjacent tRNA genes has occurred within the family Pseudococculinidae, between *N. youngi* and (*Amphiplica gordensis* + Pseudococculinidae sp.) (Fig. 3C). The distribution of this transposition among the 11 genera within Pseudococculinidae remains to be determined as mitogenomes of the remaining genera are analysed.

As previously reported (Uribe *et al.*, 2016), the mt gene order of Fissurelloidea differs extensively from those of all other vetigastropods. Rearrangements include the reversal of a different 16-gene fragment than occurred in the Scissurelloidea, Lepetelloidea and Lepetodriloidea, the transposition of a 5-gene fragment, a reversal of a 2-gene fragment and four transpositions of single tRNA genes. None of these rearrangements correspond to those that occur in the clade containing Scissurelloidea. The Fissurelloidea contains a single family Fissurellidae, which is subdivided into five subfamilies, of which only two, Fissurellinae and Diodorinae, are represented by mitogenomes with identical gene order.

Of the superfamilies currently included in the order Lepetellida, the mt gene order of Haliotoidea is by far the closest to that of the hypothetical ancestral gene order of Gastropoda, differing only in the transposition of two tRNA genes. As in prior studies using mitogenome data (e.g. Lee *et al.*, 2016; Uribe *et al.*, 2016; Wort *et al.*, 2017) Haliotoidea emerged as a sister taxon to the Seguenziida, rather than to members of Lepetellida. Other phylogenetic studies based on transcriptome data (Zapata *et al.*, 2014; Cunha *et al.*, 2021) reported an uncertain position of *Haliotis* within Vetigastropoda.

The use of mt gene sequences in phylogenetic studies of gastropods has increased in the past decade, and these data have performed well in some but not all groups. Their utility in reconstructing deep nodes has been questioned due to errors, such as LBA artefacts (Stöger & Schrödl, 2013). Rearrangements of mt gene order have been reported to occur at varying rates within taxa of differing ranks within Mollusca, ranging from among orders to between congeneric species (e.g. Gissi, Iannelli & Pesole, 2008; Grande, Templado & Zardoya, 2008; Rawlings *et al.*, 2010; Wu *et al.*, 2010). Xu *et al.* (2006) reported a correlation between levels of mitogenome rearrangement and rates of sequence evolution. Thus, phylogenetic analyses based on mitogenome sequences would likely be most robust when analysing samples with identical or similar gene orders, while shared gene order characters can serve to support sequence-based phylogenetic hypotheses of more heterogeneous samples.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of Molluscan Studies* online.

ACKNOWLEDGEMENTS

Field work was conducted while participating in the Smithsonian Institution's Deep Reef Observation Project. Support from this program is gratefully acknowledged. Thanks also to Adriaan Schrier and the staff and crew of Substation Curaçao and the Seaquarium for their gracious hospitality and support. This is Ocean Heritage Foundation/Curaçao Sea Aquarium/Substation Curaçao contribution no. 50.

Many thanks to Laboratories of Analytical Biology facilities of the National Museum of Natural History, which supported all laboratory work, sequencing and analyses. Also, we are grateful to Smithsonian High-Performance Cluster for their facilities to perform all computational analyses. JEU was supported by the Peter Buck Postdoctoral Fellowship Program from the Smithsonian Institution (2017–2019) and the Atracción Talento de la Comunidad de Madrid Fellowship Program (REFF 2019-T2/AMB-13166).

Authors' contributions: MGH collected the material and extracted the DNA, JEU and MS generated the molecular data, JEU analysed the data, MGH and JEU wrote the first draft of the manuscript and all authors contributed to writing the final version. All authors read and approved the final manuscript.

REFERENCES

ABALDE, S., TENORIO, M.J., URIBE, J.E. & ZARDOYA, R. 2019. Conidae phylogenomics and evolution. *Zoologica Scripta*, **48**: 194–214.

AKTIPIS, S.W. & GIRIBET, G. 2011. Testing relationships among the vetigastropod taxa: a molecular approach. *Journal of Molluscan Studies*, **78**: 12–27.

AKTIPIS, S.W., GIRIBET, G., LINDBERG, D.R. & PONDER, W.F. 2008. Gastropoda: an overview and analysis. In: *Phylogeny and evolution of the Mollusca* (W.F. Ponder & D.R. Lindberg, eds), pp. 201–237. University of California Press, Berkeley, CA.

ALTSCHUL, S.F., MADDEN, T.L., SCHÄFFER, A.A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**: 3389–3402.

BERNT, M., DONATH, A., JUHLING, F., EXTERNBRINK, F., FLORENTZ, C., FRITSCH, G., PÜTZ, J., MIDDENDORF, M. & STADLER, P.F. 2013. MITOS: improved *de novo* metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, **69**: 313–319.

BERNT, M., MERKLE, D., RAMSCH, K., FRITSCH, G., PERSEKE, M., BERNHARD, D., SCHLEGEL, M., STADLER, P.F. & MIDDENDORF, M. 2007. CREx: inferring genomic rearrangements based on common intervals. *Bioinformatics*, **23**: 2957–2958.

BOLGER, A.M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**: 2114–2120.

BOUCHET, P., ROCROI, J.P., HAUSDORF, B., KAIM, A., KANO, Y., NÜTZEL, A., PARKHAEV, P., SCHRÖDL, M. & STRONG, E.E. 2017. Revised classification, nomenclator and typification of gastropod and monoplacophoran families. *Malacologia*, **61**: 1–527.

CRISCUOLO, A. & GRIBALDO, S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evolutionary Biology*, **10**: 210.

CUNHA, T.J., REIMER, J.D. & GIRIBET, G. 2021. Investigating sources of conflict in deep phylogenomics of vetigastropod snails. *Systematic Biology*. DOI: https://doi.org/10.1093/sysbio/syab071.

DONATH, A., JÜHLING, F., AL-ARAB, M., BERNHART, S.H., REINHARDT, F., STADLER, P.F., MIDDENDORF, M. & BERNT, M. 2019. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. *Nucleic Acids Research*, **47**: 10543–10552.

GEIGER, D.L. & THACKER, C.E. 2005. Molecular phylogeny of Vetigastropoda reveals non-monophyletic Scissurellidae, Trochoidea, and Fissurelloidea. *Molluscan Research*, **25**: 47–55.

GISSI, C., IANNELLI, F. & PESOLE, G. 2008. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity*, **101**: 301–320.

GRABHERR, M.G., HAAS, B.J., YASSOUR, M., LEVIN, J.Z., THOMPSON, D.A., AMIT, I., ADICONIS, X., FAN, L., RAYCHOWDHURY, R., ZENG, Q., CHEN, Z., MAUCELI, E., HACOHEN, N., GNIRKE, A., RHIND, N., di PALMA, F., BIRREN, B.W., NUSBAUM, C., LINDBLAD-TOH, K., FRIEDMAN, N. & REGEV, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**: 644–652.

GRANDE, C., TEMPLADO, J. & ZARDOYA, R. 2008. Evolution of gastropod mitochondrial genome arrangements. *BMC Evolutionary Biology*, **8**: 61.

HARASEWYCH, M.G. 2019. *Tropidofusus ypotethys*: a new genus and species of Columbariidae (Mollusca: Gastropoda: Turbinelloidea). *Molluscan Research*, **39**: 148–158.

HARASEWYCH, M.G., SEI, M., WIRSHING, H., GONZÁLEZ, V.L. & URIBE, J.E. 2019. The complete mitochondrial genome of *Neptuneopsis gilchristi* G.B. Sowerby III, 1898 (Neogastropoda: Volutidae: Calliotectinae). *Nautilus*, **133**: 67–73.

HASZPRUNAR, G. 1987. Anatomy and affinities of cocculinid limpets (Mollusca, Archaeogastropoda). *Zoologica Scripta*, **16**: 305–324.

HASZPRUNAR, G. 1988a. Comparative anatomy of cocculiniform gastropods and its bearing on archaeogastropod systematics. *Malacological Review*, **Supplement 4**: 64–84.

HASZPRUNAR, G. 1988b. On the origin and evolution of major gastropod groups, with special reference to the Streptoneura. *Journal of Molluscan Studies*, **54**: 367–441.

KALYAANAMOORTHY, S., MINH, B.Q., WONG, T.K.F., VON HAESELER, A. & JERNIIN, L.S. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**: 587.

KANO, Y. 2008. Vetigastropod phylogeny and a new concept of Seguenzioidea: independent evolution of copulatory organs in the deep-sea habitats. *Zoologica Scripta*, **37**: 1–21.

KATOH, K., ROSEWICKI, J. & YAMADA, K.D. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, **20**: 1160–1166.

LARTILLOT, N. & PHILIPPE, H. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Molecular Biology and Evolution*, **21**: 1095–1109.

LARTILLOT, N., RODRIGUE, N., STUBBS, D. & RICHTER, J. 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Systematic Biology*, **62**: 611–615.

LASLETT, D. & CANBÄCK, B. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics*, **24**: 172–175.

LE, S.Q., GASCUEL, O. & LARTILLOT, N. 2008. Empirical profile mixture models for phylogenetic reconstruction. *Bioinformatics*, **24**: 2317–2323.

LEE, H., CHEN, W.J., PUILLANDRE, N., AZAR-CORMANO, L., TSAI, M.H. & SAMADI, S. 2019. Incorporation of deep-sea and small-sized species provides new insights into gastropods phylogeny. *Molecular Phylogenetics and Evolution*, **135**: 136–147.

LEE, H., SAMADI, S., PUILLANDRE, N., TSAI, M.H., DAI, C.F. & CHEN, W.J. 2016. Eight new mitogenomes for exploring the phylogeny and classification of Vetigastropoda. *Journal of Molluscan Studies*, **82**: 534–541.

McARTHUR, A.G. & HARASEWYCH, M.G. 2003. Molecular systematics of the major lineages of the Gastropoda. In: *Molecular systematics and phylogeography of mollusks* (C. Lydeard & D.R. Lindberg, eds), pp. 140–160. Smithsonian Books, Washington, DC.

McLEAN, J.H. & HARASEWYCH, M.G. 1995. Review of western Atlantic species of cocculinid and pseudococculinid limpets, with descriptions of new species (Gastropoda: Cocculiniformia). *Contributions in Science, Natural History Museum of Los Angeles County*, **453**: 1–33.

MAYNARD, B.T., KERR, L.J., McKIERNAN, J.M., JANSEN, E.S. & HANNA, P.J. 2005. Mitochondrial DNA sequence and gene organization in the Australian blacklip abalone *Haliotis rubra* (Leach). *Marine Biotechnology*, **7**: 645–658.

NAKAGAWA, S., SHIMAMURA, S., TAKAKI, Y., SUZUKI, Y., MURAKAMI, S.I., WATANABE, T., FUJIYOSHI, S., MINO, S., SAWABE, T., MAEDA, T. & MAKITA, H. 2014. Allying with armored snails: the complete genome of gammaproteobacterial endosymbiont. *ISME Journal*, **8**: 40.

NAKAJIMA, Y., SHINZATO, C., KHALTURINA, M., NAKAMURA, M., WATANABE, H., SATOH, N. & MITARAI, S. 2016. The mitochondrial genome sequence of a deep-sea, hydrothermal vent limpet, *Lepetodrilus nux*, presents a novel vetigastropod gene arrangement. *Marine Genomics*, **28**: 121–126.

NGUYEN, L.T., SCHMIDT, H.A., VON HAESELER, A. & MINH, B.Q. 2014. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**: 268–274.

PONDER, W.F. & LINDBERG, D.R. 1997. Towards a phylogeny of gastropod molluscs: an analysis using morphological characters. *Zoological Journal of the Linnean Society*, **119**: 83–265.

RAWLINGS, T., MacINNIS, M., BIELER, R., BOORE, J. & COLLINS, T. 2010. Sessile snails, dynamic genomes: gene rearrangements within the mitochondrial genome of a family of caenogastropod molluscs. *BMC Genomics*, **11**: 440.

SASAKI, T. 1998. Comparative anatomy and phylogeny of the Recent Archaeogastropoda (Mollusca: Gastropoda). *University Museum, University of Tokyo, Bulletin*, **38**: 1–223.

SCHWARZ, G. 1978. Estimating the dimension of a model. *Annals of Statistics*, **6**: 461–464.

STÖGER, I. & SCHRÖDL, M. 2013. Mitogenomics does not resolve deep molluscan relationships (yet?). *Molecular Phylogenetics and Evolution*, **69**: 376–392.

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., KANO, Y., TEMPLADO, J. & ZARDOYA, R. 2016. Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. *Zoologica Scripta*, **45**: 145–159.

URIBE, J.E., WILLIAMS, S.T., TEMPLADO, J., ABALDE, S. & ZARDOYA, R. 2017. Denser mitogenomic sampling improves resolution of the phylogeny of the superfamily Trochoidea (Gastropoda: Vetigastropoda). *Journal of Molluscan Studies*, **83**: 111–118.

WHELAN, S., IRISARRI, I. & BURKI, F. 2018. PREQUAL: detecting non-homologous characters in sets of unaligned homologous sequences. *Bioinformatics*, **34**: 3929–3930.

WORMS EDITORIAL BOARD. 2021. World Register of Marine Species. Available at: https://www.marinespecies.org/index.php. Accessed 25 September 2021.

WORT, E.J., FENBERG, P.B. & WILLIAMS, S.T. 2017. Testing the contribution of individual genes in mitochondrial genomes for assessing phylogenetic relationships in Vetigastropoda. *Journal of Molluscan Studies*, **83**: 123–128.

WU, X., XU, X., YU, Z., WIE, Z. & XIA, J. 2010. Comparison of seven *Crassostrea* mitogenomes and phylogenetic analyses. *Molecular Phylogenetics and Evolution*, **57**: 448–454.

XU, W., JAMESON, D., TANG, B. & HIGGS, P.G. 2006. The relationship between the rate of molecular evolution and the rate of genome rearrangement in animal mitochondrial genomes. *Journal of Molecular Evolution*, **63**: 375–392.

ZAPATA, F., WILSON, N.G., HOWISON, M., ANDRADE, S.C., JÖRGER, K.M., SCHRÖDL, M., GOETZ, F.E., GIRIBET, G. & DUNN, C.W. 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. *Proceedings of the Royal Society B*, **281**: 20141739.

ZHOU, Y., CHENG, H., ZHANG, D. & WANG, C. 2019. The mitochondrial genome of a slit limpet *Pseudorimula* sp. (Vetigastropoda: Lepetodrilidae) from hydrothermal vent on the Southwest Indian Ridge. *Mitochondrial DNA Part B*, **4**: 1189–1190.