Postprint del artículo publicado en: Zoologica Scripta 45(2): 145-159 (2016)

Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry

JUAN E. URIBE, YASUNORI KANO, JOSÉ TEMPLADO & RAFAEL ZARDOYA

Uribe, J.E., kano, Y., Templado J., Zardoya, R. (2015) Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. — *Zoologica Scripta* 00. 000-000.

The nucleotide sequences of the complete or nearly complete mitochondrial (mt) genomes of seven vetigastropods were determined: *Angaria neglecta* (Angarioidea), *Phasianella solida* (Phasianelloidea), *Granata lyrata* (Seguenzioidea), *Tegula lividomaculata* and *Bolma rugosa* (Trochoidea), *Diodora graeca* (Fissurelloidea), and *Lepetodrilus schrolli* (Lepetodriloidea). While the mt genomes of the superfamilies Angarioidea, Phasianelloidea, Seguenzioidea, and Trochoidea conform generally to the ancestral gene order of Vetigastropoda and Gastropoda, those of the superfamilies Fissurelloidea and Lepetodriloidea have suffered important rearrangements. The gene order of the mtDNA of *Chrysomallon squamiferum*, a representative of Neomphalina, was also analyzed since it has been proposed to be closely related to Vetigastropoda, and showed a distinct arrangement. The reconstructed phylogenies recovered Neomphalina as a distinct gastropod lineage that is the sister group (only with moderate bootstrap support) of a clade including Vetigastropoda and Neritimorpha + Caenogastropoda while the relative position of Heteroranchia and Patellogastropoda in the gastropod tree could not be determined definitively due to their long branches. Within the monophyletic Vetigastropoda, the superfamily Fissurelloidea was recovered as the sister group of two lineages, one including Lepetodriloidea as the sister group of Seguenzioidea + Halitoidea, the other including Phasianelloidea, Angarioidea, and Trochoidea without resolved relationships. The long branches of Fissurelloidea were found to introduce significant tree instability in phylogenetic reconstruction. The new phylogeny supports that the loss of the right pallial gill occurred multiple times in vetigastropod evolution as previously suggested and that Phasianelloidea, Angarioidea, and Trochoidea radiated from a common asymmetric (single-gilled) ancestor that lived in the middle Paleozoic.

Juan E. Uribe, Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006, Madrid, Spain. Email: juanes@mncn.csic.es

Yasunori Kano, Atmosphere and Ocean Research Institute, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8564, Japan. Email: kano@aori.u-tokyo.ac.jp
José Templado, Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006, Madrid, Spain. Email: templado@mncn.csic.es
Rafael Zardoya, Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006, Madrid, Spain. Email: rafaz@mncn.csic.es

Introduction

Gastropods are the most diverse class of living molluscs. They have successfully adapted to marine as well as freshwater and terrestrial environments, have a rich fossil record, and constitute an excellent model system to study and understand the evolutionary mechanisms that are involved in the generation of biodiversity over long periods of time (Aktipis *et al.* 2008). At present, up to five main monophyletic groups are commonly recognized within gastropods: Patellogastropoda, Vetigastropoda, Neritimorpha, Caenogastropoda and Heterobranchia (Haszprunar 1988; Ponder & Lindberg 1997; Bouchet & Rocroi 2005). In addition, gastropods include other minor groups of uncertain taxonomic status, such as Cocculinoidea (also referred to as Cocculiniformia or Cocculinida) and the so-called 'hot-vent taxa' (Neomphalina). The Caenogastropoda and Heterobranchia (often grouped together as Apogastropoda; Ponder & Lindberg 1997) are considered the most derived and diversified living Orthogastropoda (all gastropods but Patellogastropoda). In contrast, the remaining less diverse orthogastropod groups (Cocculiniformia, Neomphalina, Vetigastropoda and Neritimorpha), most bearing a rhipidoglossan type radula, appear to be the intriguing living remnants of earlier gastropod radiations (Fryda *et al.* 2008; Bandel 2010), and their phylogenetic interrelationships are still a matter of hot debate.

Among these less-studied groups, the Vetigastropoda is the most species rich, comprising several thousands of living species and more extinct ones (Geiger *et al.* 2008; Kano 2008). This archaic clade originated in the Cambrian/Ordovician boundary, and was the most common gastropod group in the Paleozoic (Fryda *et al.* 2008). Vetigastropods are exclusively marine snails or limpets, and occur from the intertidal to deep sea, including hydrothermal vents, cold seeps and whale and wood falls (Geiger *et al.* 2008). Vetigastropoda was first recognized as a natural group by Salvini-Plawen (1980), but has been redefined several times ever since. The clade typically included the big slit shells (Pleurotomarioidea), little slit shells (Scissurelloidea), keyhole limpets (Fissurelloidea), abalones (Haliotoidea), and top and turban shells (Trochoidea). However, in recent times, other gastropod groups of uncertain phylogenetic position such as the Lepetelloidea, Seguenzioidea, and hot-vent Lepetodriloidea (initially ascribed to "Archaeogastropoda" by McLean 1988) were added to Vetigastropoda (Ponder & Lindberg 1997; Bouchet & Rocroi 2005). The Lepetelloidea were initially

included in Cocculiniformia, a group originally described as an assemblage of small white limpets that occur on a diversity of organic deposition mainly in the deep sea (Haszprunar 1987). However, more recent studies divided the Cocculiniformia into two independent lineages: Cocculinoidea (Cocculinidae + Bathysciadiidae) of uncertain phylogenetic relationships (fluctuating from being close to Patellogastropoda to being the sister taxa of Neomphalina), and Lepetelloidea, now included among vetigastropods (Ponder & Lindberg 1997; Sasaki 1998; Bouchet & Rocroi 2005; Geiger & Thacker 2005; Kano 2008; Kano *et al.* 2013). Likewise, the placement of Seguenziidae was uncertain in early studies. Initially ascribed to "Archaeogastropoda" (e.g., Thiele 1929–35), this taxonomic group was later placed either within the Caenogastropoda (Golikov & Starobogatov 1975) or considered as an independent order (Seguenziina) equally distant to Vetigastropoda and Caenogastropoda (Salvini-Plawen & Haszprunar 1987; Haszprunar 1988). However, nowadays it is generally accepted the placement of seguenzioids within the Vetigastropoda (Ponder & Lindberg 1997; Sasaki 1998; Bouchet & Rocroi 2005; Kano 2008). On the other hand, Bandel (2010) interpreted Seguenzioidea in a more restricted way than previously suggested (Bouchet & Rocroi 2005; Kano 2008; Kano *et al.* 2009) and regarded the plesiomorphic and paraphyletic Eucycloidea as a separate, valid superfamily.

Among the traditionally recognized vetigastropod superfamilies, Trochoidea, which is the most diverse, has a very confused taxonomic history. The traditional classification of Trochoidea recognized three families, namely Trochidae, Turbinidae, and Skeneidae (Hickman & McLean 1990). However, recent phylogenetic studies have revealed that Trochoidea as traditionally defined were polyphyletic (Williams & Ozawa 2006; Heß *et al.*, 2008; Kano 2008; Williams *et al.* 2008). Some of the taxa traditionally included in Trochoidea have been transferred to Seguenzioidea (Kano 2008; Kano *et al.* 2009), whereas others are placed in their own new superfamilies, Angarioidea and Phasianelloidea (Williams & Ozawa 2006; Williams *et al.* 2008). Trochoidea is currently restricted to the families Calliostomatidae, Liotiidae, Margaritidae, Skeneidae, Solariellidae, Tegulidae, Trochidae, and Turbinidae (Williams 2012), although its final composition is still under debate and for instance, some of the Skeneidae have recently been transferred to Seguenzioidea or Neomphalina (Kano 2008; Kunze 2011) or to the new family Crosseolidae (with only five species of which the radula is known for one)

of uncertain position (Hickman 2013).

Vetigastropoda (thus comprising the superfamilies Pleurotomarioidea, Scissurelloidea, Lepetodriloidea, Fissurelloidea, Haliotoidea, Lepetelloidea, Seguenzioidea, Trochoidea, Angarioidea, and Phasianelloidea) is accepted to be monophyletic by most authors (Ponder & Lindberg 1997; Sasaki 1998; Geiger & Thacker 2005; Kano 2008; Williams *et al.* 2008). However, in some molecular phylogenetic analyses based on mitochondrial (mt) and nuclear data and including a large outgroup sampling, Vetigastropoda not always turned out to be monophyletic: the Pleurotomarioidea were placed outside Vetigastropoda and the Lepetelloidea were the sister group to Patellogastropoda (Aktipis & Giribet 2010, 2012). Furthermore, although phylogenetic relationships among vetigastropod main lineages have been repeatedly studied using morphological and molecular data (Salvini-Plawen & Haszprunar 1987; Haszprunar 1988; Hedegaard 1997; Ponder & Lindberg 1997; Sasaki 1998; Geiger & Thacker 2005; Yoon & Kim 2005; Williams & Ozawa 2006; Kano 2008; Williams *et al.* 2008; Aktipis & Giribet 2010), the phylogeny of this diverse clade remains elusive (Aktipis & Giribet 2012) and discussion and changes continue at all its levels.

In addition, the related question on the relative phylogenetic position of Neomphalina is also a matter of a lively and yet unsolved debate. Some authors consider Neomphalina within the Vetigastropoda (Bouchet & Rocroi 2005; Geiger *et al.* 2008) whereas others consider Neomphalina as a separate lineage more closely related to other gastropod clades (e.g., Heß *et al.* 2008; Appeltans *et al.* 2012; Stöger *et al.* 2013).

The present study aims to address the open questions on the composition and phylogenetic relationships of Vetigastropoda. Over its evolutionary history, this clade has suffered rapid extinction/radiation events (Fryda *et al.* 2008), which challenge the recovery of a robust molecular phylogeny, and prompt for the use of multilocus data sets. Here, we based our phylogenetic reconstructions on mitochondrial (mt) genome sequence data, which have proven to recover well-resolved phylogenetic trees of gastropods when applied to moderately divergent lineages (White *et al.* 2011 and references therein). At present, there are only seven vetigastropod complete mt genomes available, including those of a fissurelloidean, *Fissurella volcano*; two trochoideans,

Lunella aff. cinerea (Williams *et al.* 2014) and *Tegula brunnea* (NC 016954, unpublished); and four haliotoideans, *Haliotis rubra* (Maynard *et al.* 2005), *H. tuberculata* (Van Wormhoudt *et al.* 2009), *H. diversicolor* (Xin *et al.* 2011), *H. laevigata* (Robinson *et al.* 2014), as well as the almost complete mt genome of *H. discus* (EU595789, unpublished). Here, we add the complete mt genomes of one angarioidean, one phasianelloidean, one fissurelloidean, two trochoidean, and one seguenzioidean species, as well as the nearly complete mt genome of one lepetodriloidean species. We reconstructed a phylogeny of Vetigastropoda including 12 mt genomes that represent seven of the ten monophyletic superfamilies nowadays recognized within the group, with the exception of Pleurotomarioidea, Scissurelloidea, and Lepetelloidea. We also included the mt genome of the scaly-foot gastropod *Chrysomallon squamiferum* (Chen *et al.* 2015), a member of the clade Neomphalina, available at GenBank (see Nakagawa *et al.* 2014), and some mt genomes of Neritimorpha, Caenogastropoda, Heterobranchia, and Patellogastropoda as outgroup taxa. A robust phylogeny of Vetigastropoda is crucial for understanding evolutionary trends within the group, and in particular the evolution of the symmetry/asymmetry of pallial organs including the gill, which is the subject of a long-standing debate (Haszprunar 1988; Sasaki 1998; Lindberg & Ponder 2001 and references therein).

Materials and methods

Samples and DNA extraction

One specimen of each *Angaria neglecta* (Angarioidea), *Phasianella solida* (Phasianelloidea), *Granata lyrata* (Seguenzioidea), *Bolma rugosa* and *Tegula lividomaculata* (Trochoidea), *Diodora graeca* (Fissurelloidea), and *Lepetodrilus schrolli* (Lepetodriloidea) was used for this study (See Table 1 for details on the locality and voucher ID of each sample). All samples were stored in 100% ethanol and total genomic DNA was isolated from up to 50-100 mg of foot tissue following a standard phenol-chloroform extraction.

PCR amplification and sequencing

We followed a three-step procedure to amplify the different mt genomes. First, fragments of the *cox1* (Folmer *et al.* 1994), *rrnL* (Palumbi *et al.* 1991), *rrnS* (Kocher *et*

al. 1989; Simon *et al.* 1994), and *cox3* (Boore & Brown 2000) genes were PCR amplified using universal primers. The standard PCR reactions contained 2.5 µl of 10x buffer, 1.5 µl of $MgCL_2$ (25 mM), 0.5 µl of dNTPs (2.5 mM each), 0.5 µl of each primer (10mM), 0.5-1 µl (20-100 ng) of template DNA, 0.2 µl of Taq DNA polymerase 5PRIME (Hamburg, Germany), and sterilized distilled water up to 25 µl. The following program was applied: a denaturalization step at 94°C for 60 s; 45 cycles of denaturalization at 94°C 30 s, annealing at different temperatures within the range of 44-52°C depending on the gene for 60 s and extension at 72°C for 90 s; a final extension step at 72°C for 5 m. Second, the amplified fragments were sequenced using Sanger sequencing, and new primers were designed in order to amplify long fragments outwards the short fragments (See Supplementary Material 1 for the long PCR primer sequences for each mt genome). Third, the remaining mtDNA was amplified in 2-3 overlapping fragments by long PCR. The long PCR reaction contained 2.5 µl of 10 × LA Buffer II (Mg^{+2} plus), 3 µl of dNTPs (2.5 mM each), 0.5 µl of each primer (10 mM), 0,5-1 µl (20-100 ng) of template DNA and 0.2 µl TaKaRa LA Taq DNA polymerase (5 units/µl), and sterilized distilled water up to 25 µl. The following PCR conditions were used: a denaturalization step at 94°C for 60 s; 45 cycles of denaturalization at 98°C for 10 s, annealing at 53°C for 30 s and extension at 68°C for 60 s per kb; and a final extension step at 68°C for 12 min.

The Long-PCR products were purified by ethanol precipitation. Overlapping fragments from the same mt genome were pooled together in equimolar concentrations and subjected to massive parallel sequencing. For each mt genome, an indexed library was constructed using the NEXTERA XT DNA library prep Kit (Illumina, San Diego, CA, USA) at AllGenetics (A Coruña, Spain). The constructed libraries were run in an Illumina HiSeq2000 (100 Pair-ended) at Macrogen (Seoul, Korea).

Genome assembly and annotation

The assembly of the mt genomes was performed in the TRUFA webserver (Kornobis *et al.* 2015). Briefly, reads corresponding to different mt genomes were sorted out using the indexes. Adapter sequences were removed using SeqPrep (St John 2011). The quality (randomness) of the sequencing was checked using FastQC v.0.10.1 (Andrews 2010). Reads were trimmed and filtered out according to their quality scores using

PRINSEQ v.0.20.3 (Schmieder & Edwards 2011). Filtered reads were used for *de novo* assembly of mt genomes, searching for contigs with a minimum length of 3kb. The complete circular sequence of each mt genome was finally assembled by overlapping the various contigs in Sequencher 5.0.1. The assembled sequence was used as reference to map the original (raw) reads with a minimum identity of 99% using Geneious® 8.0.3.

The new vetigastropod mt genomes were annotated using the MITOS (Bernt *et al.* 2013) and DOGMA (Wyman *et al.* 2004) webservers. The 13 mt protein-coding genes were annotated by identifying their open reading frames using the invertebrate mitochondrial code. The transfer RNA (tRNA) genes were further identified with tRNAscan-SE 1.21 (Schattner *et al.* 2005) and ARWEN 1.2 (Laslett and Canbäck 2008), which infer cloverleaf secondary structures (almost all tRNAs were determined automatically but some had to be determined manually). The ribosomal RNA (rRNA) genes were identified by sequence comparison with other reported mollusc mt genomes, and assumed to extend to the boundaries of adjacent genes (Boore *et al.* 2005).

Sequence alignment

The complete sequences of the seven newly determined mt genomes were aligned to the orthologous sequences of five vetigastropod complete mt genomes (Supplementary Material 2) available at NCBI (http://www.ncbi.nlm.nih.gov/). Eleven species of Gastropoda, one Cephalopoda, and one Caudofoveata were used as outgroups (Supplementary Material 2).

Two different sequence data sets were constructed. The first data set (hereafter referred to as the gastropod data set) was aimed to test the monophyly of Vetigastropoda. It was rooted with one caudofoveate and one cephalopod, and included several species representing the following main lineages of gastropods as ingroup taxa: Patellogastropoda, Heterobranchia, Neomphalina, Neritimorpha, Caenogastropoda, and Vetigastropoda. The second data set (hereafter the vetigastropod data set) was aimed to test phylogenetic relationships within the Vetigastropoda, and was rooted with Neomphalina, Neritimorpha, and Caenogastropoda. Both data sets included the nucleotide sequence alignments of the two mt rRNA genes and the deduced amino acid sequences of the 13 mt protein coding genes. In order to construct these two data sets, the deduced amino acid sequences of the 13 mt protein-coding genes were aligned

separately using Translator X (Abascal *et al.* 2010) whereas the nucleotide sequences of the mt ribosomal RNA nuclear genes were aligned separately using MAFFT v7 (Katoh & Standley 2013) with default parameters. Ambiguously aligned positions were removed using Gblocks, v.0.91b (Castresana 2000) and allowing gap positions within the final blocks but not many contiguous non-conserved positions. Finally, the different single alignments were concatenated into the two data matrices using the ALTER webserver (Glez-Peña *et al.* 2010).

Phylogenetic analyses

Phylogenetic relationships were inferred using maximum likelihood (ML; Felsenstein, 1981) and Bayesian inference (BI; Huelsenbeck & Ronquist 2001). ML analyses were conducted with RAxML v7.3.1 (Stamatakis 2006) using the rapid hill-climbing algorithm and 10,000 bootstrap pseudoreplicates. BI analyses were conducted using MrBayes v3.1.2 (Ronquist & Huelsenbeck 2003) and running four simultaneous Markov chains for 10 million generations, sampling every 1000 generations, and discarding the first 25% generations as burn-in (as judged by plots of ML scores and low SD of split frequencies) to prevent sampling before reaching stationarity. Two independent Bayesian inference runs were performed to increase the chance of adequate mixing of the Markov chains and to increase the chance of detecting failure to converge.

The best partition schemes and best-fit models of substitution for the two data sets were identified using Partition Finder and Partition Finder Protein (Lanfear *et al.* 2012) with the Akaike information criterion (AIC; Akaike 1973). For the protein-coding gene alignments the partitions tested were: all genes combined; all genes separated except *atp6-atp8* and *nad4-nad4L;* genes grouped by subunits (*atp*, *cox*, *cytb* and *nad*; see Supplementary Material 3 for selected best fit partitions and models). For the rRNA genes, the two genes separated or combined were tested. In addition, following Williams *et al*. (2014), we tested manually whether the mtZoa model (Rota-Stabelli *et al.* 2009) could fit better than the selected models for each partition (see Supplementary Material 3).

Given the heterogeneity of evolutionary rates observed among the gastropod lineages included in the phylogenetic analyses, we also performed a BI using the site-heterogeneous mixture CAT model (Lartillot & Philippe 2004) as implemented in

PhyloBayes MPI v.1.5. (Lartillot *et al.* 2013). The CAT model assumes that the different sites of a protein evolve under distinct substitution processes and has proven to be less sensitive to (and alleviate) long-branch attraction biases in some instances (Lartillot *et al.* 2007). BI was performed without constant sites ('-dc' option), running two independent MCMC chains until convergence, sampling every cycle. The gastropod and vetigastropod data sets were analyzed only at the amino acid level (protein coding genes) under the best-fit CAT-GTR model, using the discrete gamma approximation to model among-site rate heterogeneity. The performance of the CAT-GTR+G model was assessed using a 10-fold cross-validation performed on subsamples of 6,000 non-constant positions randomly drawn from the original matrices. Convergence of analyses was checked a posteriori using the convergence tools implemented in PhyloBayes (maxdiff < 0.125, maximum discrepancy < 0.1 and effective size > 100; see Supplementary Material 4). Posterior probabilities provided branch support for BI analyses.

Results

Sequencing and assembly

The nucleotide sequences of the complete mt genomes of *A. neglecta, P. solida, B. rugosa, T. lividomaculata, D. graeca* and *G. lyrata* and the nearly complete mt genome of *L. schrolli* were determined. The Illumina sequencing produced a similar amount of sequences for *A. neglecta* (173,490 reads; 47 Mb), *P. solida* (158,008 reads; 43 Mb), *G. lyrata* (103,448 reads; 28 Mb), *D. graeca* (267,284 reads; 72 Mb), and *T. lividomaculata* (270,074 reads; 73 Mb). However, fewer data (34,300 reads; 36 Mb) were produced for *B. rugosa* because sequencing was based on a long PCR covering only a part of the mt genome. All these samples were run together with TruSeq RNA libraries (from other projects). Interestingly, *L. schrolli* produced one order of magnitude more data (6,592,262 reads; 1790 Mb) because it was run together with NEXTERA DNA libraries (from other projects). The average coverage was 857x, 280x, 715x, 974x, 984x, 771x, and 26,907x, respectively. However, due to local low coverage, it was not possible to assemble five fragments: *rrnL-cox3* in *A. neglecta*, *rrnS-cox1* in *P. solida* and *L. schrolli*, *rrnS-cox3* in *T. lividomaculata*, and *rrnL-cox1* in *B. rugosa*. These fragments were completed using Sanger sequencing and a primer walking

strategy (see Supplementary Material 1). In *L. schrolli*, primer walking through a cluster of RNA genes and the putative control region between *rrnS* and *cox3* failed.

Structural features and mitochondrial organization

The newly determined genomes contain 13 protein-coding, two ribosomal RNA and 22 transfer RNA genes. For the nearly complete mt genome of *L. schrolli*, only 15 of the 22 tRNAs were identified, and two tRNAs were missing from the *T. lividomaculata* genome). Five complete mt genomes (*A. neglecta*, *P. solida*, *B. rugosa*, *T. lividomaculata*, and *G. lyrata*) share the same gene order except for the relative position of the *trnG* and *trnE* genes (Fig. 1). The major strand encodes *cox1-3*, *atp6*, *atp8*, *nad2*, *nad3*, *trnD* (except in *G. lyrata*), *trnT*, *trnS* (gcu), and the KARNI (*trnK*, *trnA*, *trnR*, *trnN* and *trnI*) cluster (Fig. 1). The minus strand encodes the remaining protein-coding genes (*nad5*, *nad4*, *nad4L*, *cytb*, *nad6*, and *nad1*), the two rRNA genes (*rrnS* and *rrnL*), *trnF*, *trnH*, *trnS* (uga), *trnP*, *trnL* (uaa), *trnL* (uag) and the MYCWQ (*trnM*, *trnY*, *trnC*, *trnW*, and *trnQ*) cluster (Fig. 1). In *G. lyrata*, the cluster is extended with the *trnG* and *trnE*, also encoded by the minus strand. In *P. solida*, the cluster is prolonged with the *trnE* and *trnG* genes encoded by the major strand. In *A. neglecta*, the cluster is extended with the *trnE* and *trnG* genes encoded by the minus and major strands, respectively (Fig. 1). In *B. rugosa*, the cluster is prolonged with the *trnG* gene encoded by the major strand whereas the *trnE* gene is tentatively located (manually) between *cox1* and *cox2* genes, encoded by the major strand (Fig. 1). In this mt genome, the *trnT* gene is located between the *trnN* and *trnI* genes, as in *Lunella* (Fig. 1). In *T. lividomaculata*, we could not find the *trnE* and *trnG* genes (note that the former is also missing in *T. brunnea*; Fig. 1). The partial genome of *L. schrolli* shows a different gene arrangement in which *trnF* *nad5*, *trnH*, *nad4*, *nad4L*, *trnS* (uga), *cytb*, *nad6*, *trnP*, *nad1*, *trnL* (uaa), and *trnL* (uag) are encoded by the major strand whereas *trnD*, *atp8*, *atp6*, and *trnT* are encoded by the minus strand (Fig. 1). The mt genome organization of *D. graeca* is the same as that inferred automatically with MITOS for *Fissurella volcano* (i.e., the mt gene order reported in GenBank Accession No. NC_016953 is outdated). Both of the fissurellid mt genomes showed numerous rearrangements compared to other vetigastropod mt genomes. The genes *nad4*/*nad4L* overlapped in seven bp in all mt genomes (but those of Fissurelloidea). Almost all protein-coding genes start their open reading frame with the

codon ATG except *nad4* in *P. solida* that starts with ATT; *atp6* and *nad4* in *G. lyrata* that start with TTG and GTG, respectively; *nad1* and *nad4* that start with GTG in *D. graeca*; and *atp8* and *nad1* in *L. schrolli* that start with GTG (Supplementary Material 4). The stop codons were variable depending on the gene and the species, and only *cox2* consistently ended with TAA (Supplementary Material 4). In *G. lyrata*, *nad1* and *atp8* genes were abnormally long (Supplementary Material 4). Each mt genome showed several intergenic regions, and those of *A. neglecta* were particularly long (up to 487 bp; see Supplementary Material 4). Most intergenic regions of *A. neglecta*, *G. lyrata*, and *P. solida* showed an A-T% below 70% whereas most of these regions in *B. rugosa* and *T. lividomaculata* showed an A-T% above 70% (Supplementary Material 4). In *G. lyrata*, the intergenic region upstream *cox3* (putative control regions) was the longest (772 bp) but the A-T percentage was lower than 70% (62.7%) (Supplementary Material 4). The partial genome of *L. schrolli* was comparatively rather compact with short intergenic regions, and unfortunately the region upstream *cox3* could not be sequenced completely.

Phylogenetic relationships of Vetigastropoda

The molecular phylogeny of Gastropoda was reconstructed based on the deduced amino acid sequences of the 13 protein coding genes combined with the nucleotide sequences of the two rRNA genes (the gastropod data set) using probabilistic methods (Fig. 2). The final matrix was 4069 positions long. ML (*–lnL* = 72681.74) and BI (*–lnL* = 82710.11 for run1; *–lnL* = 82709.68 for run2) arrived at similar topologies (Fig. 2) that only differed in the relative position of *Phasianella* and *Angaria* (see below). The reconstructed trees recovered Heterobranchia + Patellogastropoda as the sister group to the remaining gastropods (Fig. 2). Within the latter, Neomphalina was the sister group of Vetigastropoda and Neritimorpha + Caenogastropoda. The vetigastropods were recovered as a monophyletic group with the maximal BPP and 78% bootstrap support (Fig. 2).

Phylogenetic relationships within the Vetigastropoda were also inferred based on another combined data set (the vetigastropod data set) of mitochondrial amino acid (13 protein coding gene) and nucleotide (two rRNA gene) sequences (Fig. 3). The final analyzed matrix was 4645 positions long. ML (*–lnL* = 59411.92) and BI (*–lnL* = 67558.08 for run1; *–lnL* = 67558.46 for run2) arrived at similar topologies (Fig. 3) only

differing on the relative position of *Phasianella* and *Angaria* (see below). Vetigastropods were recovered as a monophyletic group with 0.66 BPP and 97% bootstrap support (Fig. 3). Three main lineages were recovered within the Vetigastropoda (Fig. 3). The first lineage included *Fissurella* and *Diodora*, which were recovered as the sister group of the remaining vetigastropods (Fig. 3). The second lineage recovered Lepetodriloidea as the sister group of Seguenzioidea + Haliotoidea (Fig. 3). The third lineage included Phasinelloidea, Angarioidea, and Trochoidea. In ML, Phasinelloidea was recovered as the sister group of Angarioidea and Trochoidea whereas in BI, Phasinelloidea and Angarioidea are sister groups to the exclusion of Trochoidea (Fig. 3).

The two fissurelloidean representatives showed relatively long branches that produced significant tree instability as evidenced by only moderate statistical support in some particular nodes of the gastropod and vetigastropod trees (Figs. 2 and 3). When fissulleroideans were removed from phylogenetic analyses, all nodes in the trees had the maximal BPPs and above 70% bootstrap values and converged to a single topology in which Phasinelloidea was recovered as the sister group of Angarioidea and Trochoidea (not shown).

Phylogenetic analyses using BI under the CAT-GTR+G model rendered a rather unresolved tree based on the gastropod data set (see Supplementary Material 5). The best topology placed Heterobranchia together with Caneogastropoda and Neritimorpha in the same clade whereas Patellogastropoda was nested within the Vetigastropoda, and Neomphalina was recovered as the sister group of Vetigastropoda (including Patellogastropoda). Unfortunately, none of these relationships had meaningful statistical support (Supplementary Material 5). The reconstructed BI tree under the CAT-GTR+G model based on the vetigastropod data set had an identical topology and similar levels of nodal support with the ML tree shown in Figure 3 (Supplementary Material 5).

Discussion

Gene order evolution

As of May 2015, most of the complete mt genomes of gastropods sequenced thus far originate from the Heterobranchia (46 mtDNAs) and Caenogastropoda (31 mtDNAs)

whereas those of other main gastropod lineages are still underrepresented in sequence databases. Here, we provide six new complete (and one almost complete) mt genomes of Vetigastropoda to add to the six (and one almost complete) already available for this lineage. Several of the mtDNAs here sequenced represent vetigastropod superfamilies not previously sampled (Lepetodriloidea, Seguenzioidea, Phasianelloidea, and Angarioidea). In addition, we analyzed the mtDNA of one representative of Neomphalina (Peltospiridae) that was available in Genbank but thus far not properly analyzed since it was obtained as a by-product of the sequencing of the complete genome of a bacterial endosymbiont of the scaly-foot gastropod (Nakagawa *et al.* 2014). This latter mt genome has a striking genome organization that is different from those of other main lineages in Gastropoda. Compared to the hypothetical ancestral gene order of gastropods (Stöger & Schrödl 2013; Osca *et al.* 2014a), the mt genome of *Chrysomallon* has suffered two main inversions affecting a cluster including *cox2*, *trnD*, *atp8*, *atp6*, and *trnF* genes and a cluster including *trnY*, *trnC*, *trnW*, and *trnQ* genes (Fig. 1). In addition, two tRNA genes (*trnT* and *trnE*) have been translocated and one inverted (*trnG*).

Within the Vetigastropoda, the genera *Haliotis*, *Granata*, *Phasianella*, *Angaria*, *Bolma*, *Tegula*, and *Lunella* share almost the same genome organization, which is very similar to the hypothetical gastropod ancestral gene order (Fig. 1). Only rearrangements affecting the *trnE*, *trnG*, *trnT*, *trnN*, and *trnD* genes are detected (Fig. 1). The mt genome of *Lepetodrilus* shows one inversion event affecting a large fragment including the *trnD*, *atp8*, *atp6*, *trnF*, *nad5*, *trnH*, *nad4*, *trnT*, *trnS*, *cob*, *nad6*, *trnP*, *nad1*, *trnL* (uaa) and *trnL* (uag) genes; otherwise this mt genome shares the gastropod ancestral gene order (but note that the MYCWQGE cluster i.e, *trnM*, *trnY*, *trnC*, *trnW*, *trnQ*, *trnG*, and *trnE* genes could not be sequenced). Finally, the mt genomes of *Fissurella* (NC 016953, unpublished) and *Diodora* (this work) also show a large inverted fragment affecting the *cob*, *nad6*, *trnP*, *nad1*, *trnL* (uaa) and *trnL* (uag), *rrnL*, *trnV*, *rrnS* genes, and the MYCWQGE cluster (Fig. 1). In addition, the *trnF*, *trnD*, *trnS*, *trnR*, and *trnK* genes have also been rearranged independently (Fig. 1). The particularly high number of rearrangements of these mt genomes is correlated with the high evolutionary rates exhibited by these species (as evidenced by their long branches in the trees). This correlation between high rearrangement and evolutionary rates has been noticed in other

molluscs (Rawlings *et al.* 2010; Schrödl & Stöger 2014). In the overall context of gastropods, vetigastropods ancestrally retain the hypothetical ancestral gene order of gastropods as neritimorphs do (but note that only the genus *Nerita* has been sequenced thus far in this group; Castro & Colgan 2010; Arquez *et al.* 2014). In contrast, caenogastropods (Cunha *et al.*, 2009) and neomphalins (this work) show instances of discrete inversion events in their ancestors whereas Patellogastropoda (Simison *et al.* 2006) and Heterobranchia (Grande *et al.* 2008) had extensive rearrangements in their ancestors.

Phylogeny of Gastropoda

As in most previous phylogenetic analyses of gastropods based on the derived amino acid sequences of mt protein coding genes (Grande *et al.* 2008; Castro & Colgan 2010; Arquez *et al.* 2014; Osca *et al.* 2014b), the trees here reconstructed showed a strongly-supported sister group relationship of Patellogastropoda and Heterobranchia. This relationship is defined by the markedly long branches of both groups, and has been reported as spurious due to a long-branch attraction (LBA) artifact (Grande *et al.* 2008; Stöger & Schrödl 2013). In fact, phylogenetic analyses based on morphology supported a sister group relationship of Patellogastropoda to the remaining gastropods (Ponder & Lindberg 1997; Sasaki 1998). This result was also obtained by a phylogenetic analyses based on nuclear sequences (Osca *et al.* 2014b) but other phylogenies that used nuclear data (alone or combined with mt data) nested Patellogastropoda deeply within gastropods as the sister group of Vetigastropoda (Zapata *et al.* 2014) or even within the Vetigastropoda (Colgan *et al.* 2003; Aktipis & Giribet 2010, 2012). Interestingly, phylogenetic analyses performed at the nucleotide level based on the first and second codon positions of mt protein coding genes and rRNA genes have also recovered Patellogastropoda as the sister group of Vetigastropoda (Castro & Colgan 2010).

Morphology (Haszprunar 1988; Ponder & Lindberg 1997), nuclear sequences (McArthur & Harasewych 2003; Osca *et al.* 2014b; Zapata *et al.* 2014), first and second codon positions of mitochondrial protein coding genes and rRNA genes (Castro & Colgan 2010), and combined mt and nuclear sequence data (Aktipis & Giribet 2010, 2012) have recovered Heterobranchia as the sister group of Caenogastropoda, forming the clade Apogastropoda (Ponder & Lindberg 1997). In contrast, in our phylogenetic

analyses Caenogastropoda is placed as the sister group of Neritimorpha to the exclusion of Vetigastropoda. In previous phylogenetic analyses also based on mt amino acid sequences, these three groups always clustered together but in some instances Neritimorpha was recovered as the sister group of Caenogastropoda as here (Castro & Colgan 2010; Osca *et al.* 2014b) whereas in one case it was the sister group of Vetigastropoda (Arquez *et al.* 2014). Combined mt and nuclear data supported either Neritimorpha as the sister group of Caenogastropoda (Aktipis & Giribet 2010), of Vetigastropoda (Osca *et al.* 2014b) or of all other gastropods (Aktipis & Giribet 2012). The latest nuclear-based phylogeny supports a sister group relationship of Neritimorpha and Apogastropoda (Zapata *et al.* 2014). Altogether, this latter hypothesis seems to be the strongest after comparing the different studies and taking into account the above-mentioned biases introduced by the long branch of Heterobranchia in the mt-based phylogenetic analyses.

The BI phylogenetic analysis of the gastropod data set using the site-heterogeneous mixture CAT-GTR+G model was able to avoid the LBA artifact between Heterobranchia and Patellogastropoda, placing the former closer to Caenogastropoda (in support of the Apogastropoda hypothesis; Ponder & Lindberg 1997) and the latter within the Vetigastropoda as previously reported (Colgan *et al.* 2003; Aktipis & Giribet 2010, 2012). However, internal nodes in this tree had no meaningful statistical support.

The intriguing phylogenetic position of Neomphalina

The Neomphalina are enigmatic hydrothermal vent marine snails (McLean 1981; Warén *et al.* 2003) of an uncertain phylogenetic position ever since their discovery as they have been variously placed as the sister group of Vetigastropoda (Ponder & Lindberg 1997; Warén *et al.* 2003), within the Vetigastropoda (Bouchet & Rocroi 2005; Aktipis & Giribet 2012) or closest to Cocculinoidea (McArthur & Harasewych 2003; Aktipis & Giribet 2012; Stöger *et al.* 2013). Here, the phylogenetic analysis supports Neomphalina an independent lineage unrelated to Vetigastropoda and the sister group of a clade including Vetigastropoda and Neritimorpha + Caenogastropoda. However, it should be noted that (i) no Cocculinoidea was included in this analysis and (ii) the BI analysis under the CAT-GTR+G model, which was aimed to alleviate the above-mentioned long-branch attraction artifacts, recovered Neompahlina as the sister group of Vetigastropoda

and Patellogastropoda, although with insufficient statistical support. Also, (iii) the morphological resemblance between the Neomphalina and Vetigastropoda, including their similar radulae and shared ctenidial bursicles (Warén & Bouchet 2001; Heß *et al.* 2008), points to the inconclusiveness of the present topology.

Phylogeny of Vetigastropoda

The monophyly of Vetigastropoda (Fissurelloidea, Lepetodriloidea, Seguenzioidea, Haliotoidea, Phasianelloidea, Angarioidea, and Trochoidea in our analysis) is well supported in all but one (BI under CAT-GTR+G model based on the gastropod data set) of the present phylogenetic analyses, as is accepted by most authors (Ponder & Lindberg 1997; Geiger & Thacker 2005; Kano 2008; Williams *et al.* 2008; Zapata *et al.* 2014). However, note that members of Pleurotomarioidea, Lepetelloidea and Scissurelloidea were not included in the present study because their mt genomes are not yet available. Hence, we cannot discuss on the relative position neither of Pleurotomarioidea, which is commonly recognized as the sister group (earliest branch) to the remaining vetigastropods (Haszprunar 1988; Harasewych *et al.* 1997; Ponder & Lindberg 1997; Harasewych 2002; Geiger & Thacker 2005; Yoon & Kim 2005; Williams & Ozawa 2006; Kano 2008; Stöger *et al.* 2013; Zapata *et al.* 2014) nor of the deep sea Lepetelloidea, previously ascribed to the Cocculiniformia, and now included within the Vetigastropoda (Ponder & Lindberg 1997; Kano 2008; Lindberg 2008). Moreover, despite Fissurelloidea is placed as the sister group of the remaining vetigastropods (as in e.g., Kano 2008; but see e.g., Williams *et al.* 2008), we cannot reach any definitive conclusion regarding the relative phylogenetic position of this taxon due to the long branches of its representatives that caused significant instability of the tree. In fact, trees with either *Fissurella* or *Diodora* as the only representative of Fissurelloidea were even less stable. The addition of new representatives of Fissurelloidea will contribute to break down the long branch leading to this clade and improve the vetigastropod tree (Wägele & Mayer, 2007). Furthermore, when both taxa were removed from analyses, overall statistical support within the Vetigastropoda was stronger and all phylogenetic analyses converged to a single topology with regards to vetigastropod interrelationships. This topology was also recovered in the BI analysis with the CAT-GTR+G model, which has been proposed to be less sensitive to LBA

phenomena.

Vetigastropoda has been the subject of numerous morphological and molecular phylogenetic studies that agree on the monophyly of the different superfamilies, but conflict on the phylogenetic relationships among them (Salvini-Plawen & Haszprunar 1987; Haszprunar 1988; Hedegaard 1997; Ponder & Lindberg 1997; Sasaki 1998; Geiger & Thacker 2005; Yoon & Kim 2005; Williams & Ozawa 2006; Geiger *et al.* 2008; Kano 2008; Williams *et al.* 2008; Kano *et al.* 2009; Aktipis & Giribet 2010, 2012). Here, we recovered three distinct lineages within the Vetigastropoda that separate Fissurelloidea from the remaining vetigastropods, and Trochoidea + Angarioidea + Phasianelloidea from Haliotoidea + Seguenzioidea + Lepetodriloidea. The composition of the superfamily Trochoidea has been the source of taxonomic debate over the last few decades. In their seminal morphological monograph, (Hickman & McLean 1990) defined Trochoidea to comprise the families Turbinidae (including subfamilies Angariinae and Phasianellinae), Trochidae and Skeneidae. In recent years, changes to the systematics at the family level based on the comprehensive studies of (Williams & Ozawa 2006; Williams *et al.* 2008, 2012), led to corresponding changes at the superfamily level and the ultimate recognition of three superfamilies: Trochoidea, Angarioidea, Phasianelloidea. Interestingly, these three superfamilies form a monophyletic group in the reconstructed trees contrary to the results based on combined mt and nuclear sequences by Williams *et al.* (2008) and Aktipis & Giribet (2012), where Angarioidea and Phasianelloidea form the sister group of the remaining vetigastropods excluding pleurotomarioideans. Hence, our results emphasize the close affinity of Trochoidea, Angarioidea, and Phasianelloidea with the highest support values (see also Zapata *et al.* 2014) and prompt for further increasing the number of complete mt genomes of the highly diverse Turbinidae and Trochidae (Williams *et al.* 2014).

Among the non-trochoidean groups, our analyses recovered *Lepetodrilus* (Lepetodriloidea) as the sister group to *Granata* (Seguenzioidea) and *Haliotis* (Haliotoidea), although without statistical support in the vetigastropod tree (Fig. 3). This clade has been found in several previous studies, although internal phylogenetic relationships were different with Seguenzioidea as the sister group of Haliotoidea and Lepetodriloidea (Kano 2008) or Haliotoidea sister to Seguenzioidea and Lepetodriloidea (Williams *et al.* 2008). The close relationship between Haliotoidea and Seguenzioidea is

supported in another phylogenetic reconstruction based on combined mt and nuclear sequences (Aktipis & Giribet 2012), whereas neither this nor the above two previous phylogenies settled the position of Haliotoidea with meaningful support indices. The latest phylogenomic analysis recovered the three lineages branching off successively and paraphyletic with respect to Trochoidea, but again the position of Haliotoidea was ambiguous due to relatively poor gene sampling for this lineage (Zapata *et al.* 2014). Lepetodriloidea is recovered in recent studies as the sister group of Lepetelloidea (Kano *et al.* 2013; Zapata *et al.* 2014), a taxon not included in the present study.

Implications for the evolution of pallial asymmetry and paleontology

Our phylogenetic reconstruction of the Vetigastropoda sheds new light on the traditional debate on symmetry (or asymmetry) in gastropod pallial organs, including the gill (ctenidium), osphradium, hypobranchial gland, kidney and auricle (see Lindberg & Ponder 2001 for a review), and consequently the systematics and identification of Paleozoic and Mesozoic fossils. Many of vetigastropod taxa including the Trochoidea lack the gill on the right side, while others bear both left and right ones (Ponder & Lindberg 1997). The latter paired (zeugobranch) condition can usually be recognized in both extant and extinct taxa by the presence of a shell slit or a foramen, through which water is expelled after passing through the (more-or-less) symmetric mantle cavity (Haszprunar 1988; Ponder & Lindberg 1997; Sasaki 1998). The presence of such a structure contrasts with the simple, straight outer lip of the shell that characterizes trochoideans and other vetigastropods with the strongly asymmetric pallial cavity with the single left gill (Hickman & McLean 1990). Regarding the evolutionary polarity of single/paired conditions, recent molecular studies resolve the position of the zeugobranch Pleurotomarioidea as the basal-most Vetigastropoda (see above). The rich Paleozoic fossil record of zeugobranchs with shell slits agrees well with this topology (Knight *et al.* 1960; Lindberg & Ponder 2001; Fryda *et al.* 2008; Geiger *et al.* 2008).

The present mitochondrial phylogeny clusters Trochoidea, Angarioidea and Phasianelloidea (all asymmetric) on the one hand, and zeugobranch Haliotoidea and single-gilled Seguenzioidea on the other hand, both with high posterior and bootstrap indices (Fig. 3). This suggests not only the loss of the right gill occurred multiple times in vetigastropod evolution as proposed by previous authors (e.g. Ponder & Lindberg

1997; Lindberg & Ponder 2001; Kano 2008), but also that the clade containing Trochoidea, Angarioidea and Phasianelloidea might represent an ancient radiation from a common asymmetric ancestor that lived in the middle Paleozoic. The fossil history of 'trochomorphs' (trochoideans and other vetigastropod snails without slits or holes) undoubtedly goes back to the Devonian and probably to the Ordovician (Knight *et al.*, 1960; Geiger *et al.*, 2008). The monophyly of Trochoidea, Angarioidea and Phasianelloidea as a large, ancient clade thus appears to be in better agreement with the fossil record than previous phylogenetic hypotheses that regard the Trochoidea as an independent, more recent trochomorph radiation since the Mesozoic era (Kano 2008; Williams *et al.* 2008; Aktipis & Giribet 2012).

The Seguenzioidea represent the only other extant clade of trochomorphs with macroscopic (>2 mm) species (Kano 2008; Kano *et al.* 2009). Their abundant fossil record dates back to the Triassic (Hickman & McLean 1990; Bandel 2010). The present mtDNA phylogeny recovered a sister relationship between Seguenzioidea and Haliotoidea, the latter of which has a considerably younger record since the Late Cretaceous (Knight *et al.* 1960; Geiger *et al.* 2008). An apomorphic shift from the plesiomorphic slit shell, which is represented in Scissurelloidea and Fissurelloidea among extant taxa, would account for the apparent lack of pre-Cretaceous fossil evidence for the lineage leading to living haliotids.

Here it is interesting to note that the right pallial organs of *Haliotis* appear much later in post-metamorphic ontogeny than the left (Crofts 1937). One may infer a secondary evolutionary acquisition of the right gill from this asynchronous development (Sasaki 1998) as opposed to the traditional idea of the zeugobranch condition being plesiomorphic (see Lindberg & Ponder 2001). Crofts (1937) and Salvini-Plawen (1980) have explained in this regard that the juveniles of *Haliotis* and adults of single-gilled gastropods retain larval asymmetry caused by torsion, a unique synapomorphy for the entire Gastropoda (Haszprunar 1988; Ponder & Lindberg,1997). The retarded ontogeny therefore does not seem to carry a straightforward implication for assessing the evolutionary polarity of single/paired conditions in post-metamorphic pallial organs including the gill.

Other recent vetigastropod taxa with a single gill seem to have originated more recently than trochoids and seguenzioids, some probably even in the Cenozoic. Each of

the (originally zeugobranch) Scissurelloidea, Lepetodriloidea and Lepetelloidea contains one or more subclades with the strongly asymmetric pallial cavity and straight margin of the shell aperture (Kano 2008). Moreover, confamilial species with single or paired gills exist in Scissurellidae (Geiger 2012), Lepetodrilidae (Warén & Bouchet 2001) and Pseudococculinidae (Lepetelloidea; Kano *et al.* 2013). Most of these taxa with a single gill have small to minute body sizes, which may reduce respiratory demand or structurally constrain the complexity of the pallial organs on the narrower right side in a right-handed snail shell (Lindberg & Ponder 2001; Kano 2008). Summing up, the present phylogeny corroborates the multiple secondary losses of the pallial symmetry in the vetigastropod evolution, while it also proposes a possibility of longer geological histories for two extant clades of trochomorphs than previously calibrated using molecular data (Williams *et al.* 2008; Zapata *et al.* 2014).

Conclusions

The available complete mt genomes of Vetigastropoda were doubled. Several of the new mt genomes represent vetigastropod lineages not previously sampled and thus allowed reconstructing a vetigastropod tree based on complete mt genome sequence data. Neomphalina was tentatively recovered as a lineage independent of vetigastropods. The superfamily Fissurelloidea was recovered as the sister group of the remaining vetigastropods, although their representatives show high evolutionary and rearrangement rates that affect phylogenetic reconstruction and cause tree instability. The remaining analyzed vetigastropods are divided into two distinct groups: one including the superfamilies Trochoidea, Angarioidea and Phasianelloidea and the other including the superfamilies Lepetodriloidea, Haliotoidea and Seguenzioidea, suggesting that the former clade has descended from archaic trochomorphs that might have lost the pallial symmetry already in the Ordovician. Phylogenetic reconstruction based on complete mt genome sequence data seems to be particularly informative at the superfamily level and provides rather resolved vetigastropod trees. The addition of mt genomes from missing lineages (Pleurotomarioidea, Scissurelloidea and Lepetelloidea) as well as from controversial groups such as the polyphyletic skeneimorphs should help obtaining a robust phylogenetic framework to further understand the evolution of Vetigastropoda.

Acknowledgements

We are grateful to Iker Irisarri for his help in the molecular laboratory, David Osca for helping with phylogenetic analyses, and Etienne Kornobis for introducing us to the assembly programs and the Altamira environment, and Anders Warén and two anonymous reviewers for their invaluable comments to the manuscript. We are also grateful to Jesús Marco and Luis Cabellos who provided access to the supercomputer Altamira at the Institute of Physics of Cantabria (IFCA-CSIC), member of the Spanish Supercomputing Network, for performing phylogenetic analyses. This work was supported by the Spanish Ministry of Science and Innovation (CGL2010-18216 and CGL2013-45211-C2-2-P to RZ; BES-2011-051469 to JEU) and the Japan Society for the Promotion of Science (26291077 and 15H04412 to YK).

References

Abascal, F., Zardoya, R., & Telford, M. J. (2010). TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research, 38*, W7-13.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In; B. N. Petrov and F. Csaki (Eds). *2nd International Symposium on Information Theory* (pp 267-28). Budapest, Hungary: Akademiai Kiado.

Aktipis, S. W., & Giribet, G. (2010). A phylogeny of Vetigastropoda and other "archaeogastropods": re-organizing old gastropod clades. *Invertebrate Biology*, *129*, 220-240.

Aktipis, S. W., & Giribet, G. (2012). 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: W. F. Ponder & D. R. Lindberg (Eds) *Phylogeny and Evolution of the Mollusca* (pp. 201-237). Berkeley: University of California Press.

Andrews, S. (2010). FastQC. available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Appeltans, W., Ahyong, S. T., Anderson, G., Angel, M. V., Artois, T., Bailly, N., et al. (2012). The Magnitude of Global Marine Species Diversity. *Current Biology*, *22*, 2189-2202.

Arquez, M., Colgan, D., & Castro, L. R. (2014). Sequence and comparison of mitochondrial genomes in the genus *Nerita* (Gastropoda: Neritimorpha: Neritidae) and phylogenetic considerations among gastropods. *Marine Genomics*, *15*, 45-54.

Bandel, K. (2010). Relationships of the Triassic Eucycloidea Koken, 1897 (Mollusca, Gastropoda) to modern genera such as *Pagodatrochus, Calliotropis* and *Euchelus*, based on morphology of the early shell. *Bulletin of Geosciences*, *85*, 435-486.

Bernt, M., Donath, A., Jühling, F. Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M. & Stadler, P. F. (2013). MITOS: Improved *de novo* metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, *69*, 313-319.

Boore, J. L., & Brown, W. M. (2000). Mitochondrial Genomes of *Galathealinum, Helobdella,* and *Platynereis*: Sequence and Gene Arrangement Comparisons Indicate that Pogonophora Is Not a Phylum and Annelida and Arthropoda Are Not Sister Taxa. *Molecular Biology and Evolution*, *17*, 87-106.

Boore, J. L., Macey, J. R., & Medina, M. 2005. Sequencing and Comparing Whole Mitochondrial Genomes of Animals. In A. Z. Elizabeth & H. R. Eric (Eds) *Methods in Enzymology* pp. 311-348): Academic Press.

Bouchet, P., & Rocroi, J. P. (2005). Classification and Nomenclator of Gastropod Families. *Malacologia*, *47*, 397.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, *17*, 540-552.

Castro, L. R., & Colgan, D. J. (2010). The phylogenetic position of Neritimorpha based on the mitochondrial genome of *Nerita melanotragus* (Mollusca: Gastropoda). *Molecular Phylogenetics and Evolution*, *57*, 918-923.

Chen, C., Linse, K., Copley, J. T., & Rogers, A. D. (2015). The ‘scaly-foot gastropod’: a new genus and species of hydrothermal vent-endemic gastropod (Neomphalina: Peltospiridae) from the Indian Ocean. *Journal of Molluscan Studies*. doi: 10.1093/mollus/eyv013.

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

Crofts, D. R. (1937) The development of *Haliotis tuberculata*, with special reference to organogenesis during torsion. *Phil. Trans. R. Soc. Lond.* B, *228*, 219-268.

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

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution 17*, 368-376.

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, *3*, 294-299.

Fryda, J., Nützel, A., & Wagner, P. J. 2008. Paleozoic gastropoda. In: W. F. Ponder & D. R. Lindberg (Eds) *Phylogeny and Evolution of the Mollusca*. (pp. 239-270). Berkeley: University of California Press.

Geiger, D. L. (2012). *Monograph of the Little Slit Shells*. Santa Barbara: Santa Barbara Museum of Natural History.

Geiger, D. L., Nützel, A., & Sasaki, T. 2008. Vetigastropoda. In: W. F. Ponder & D. R. Lindberg (Eds) *Phylogeny and Evolution of the Mollusca*. (pp. 297-330). Berkeley: University of California Press.

Geiger, D. L., & Thacker, C. E. (2005). Molecular phylogeny of Vetigastropoda reveals non-monophyletic Scissurellidae, Trochoidea, and Fissurelloidea. *Molluscan Research, 25*, 47-55.

Glez-Peña, D., Gómez-Blanco, D., Reboiro-Jato, M., Fdez-Riverola, F., & Posada, D. (2010). ALTER: program-oriented conversion of DNA and protein alignments. *Nucleic Acids Research, 38*, W14-W18.

Golikov, A. N., & Starobogatov, Y. I. (1975). Systematics of prosobranch gastropods. *Malacologia, 15*, 185-232.

Grande, C., Templado, J., & Zardoya, R. (2008). Evolution of gastropod mitochondrial genome arrangements. *BMC Evolutionary Biology, 8*, 61.

Harasewych, M. G. (2002). Pleurotomarioidean gastropods. *Advances in Marine Biology*, *42*, 237-294.

Harasewych, M. G., Adamkewicz, S. L., Blake, J. A., Saudek, D., Spriggs, T., & Bult, C. J. (1997). Phylogeny and relationships of pleurotomariid gastropods (Mollusca: Gastropoda): an assessment based on partial *18S*

rDNA and cytochrome *c* oxidase I sequences. *Molecular Marine Biology and Biotechnology*, *6*, 1-20.

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

Haszprunar, G. (1988). On the origin and evolution of major gastropod groups, with special reference to the Streptoneura. *Journal of Molluscan Stu*dies, *54*, 367-441.

Hedegaard, C. (1997). Shell structures of the Recent Vetigastropoda. *Journal of Molluscan Studies*, *63*, 369-377.

Heß, M., Beck, F., Gensler, H., Kano, Y., Kiel, S., & Haszprunar, G. (2008). Microanatomy, shell structre and molecular phylogeny of *Leptogyra*, *Xyleptogyra* and *Leptogyropsis* (Gastropoda: Neomphalida: Melanodrymiidae) from sunken wood. *Journal of Molluscan Studies*, *74*, 383-401.

Hickman, C. S. (2013). Crosseolidae, a new family of skeneiform microgastropods and progress toward definition of monophyletic Skeneidae. *American Malacological Bulletin*, *31*, 1-16.

Hickman, C. S., & McLean, J. H. (1990). Systematic revision and suprageneric classification of trochacean gastropods. *Natural History Museum of Los Angeles County, Science Series*, *35*, 1-169.

Huelsenbeck, J. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics, 17*, 754-755.

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.

Kano, Y., Chikyu, E., & Warén, A. (2009). Morphological, ecological and molecular characterization of the enigmatic planispiral snail genus *Adeuomphalus* (Vetigastropoda: Seguenzioidea). *Journal of Molluscan Studies*, *75*, 397-418.

Kano, Y., Judge, J., Takano, T., Marshall, B., & Warèn, A. (2013). Illuminating relationships and habitat shifts in the lepetelloidean limpet radiation into deep-sea organic and chemosynthetic habitats: a molecular approach.

Book of Abstracts, World Congress of Malacology, Ponta Delgada, Portugal, 72–73.

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution, 30*, 772-780.

Knight, J. B., Cox, L. R., Keen, A. M., Batten, R. L., Yochelson, E. L., & Robertson, R. 1960. Archeogastropoda and some (mainly Paleozoic) Caenogastropoda and Opisthobranchia (Mollusca: Gastropoda). In: R. C. Moore (Ed) *Treatise on Invertebrate Paleontology* (pp. 1–351). Lawrence, Kansas: Geological Society of America and Kansas University Press.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., et al. (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of USA, 86*, 6196-6200.

Kornobis, E., Cabellos, L., Aguilar, F., Frías-López, C., Rozas, J., Marco, J., et al. (2015). TRUFA: A User-Friendly Web Server for de novo RNA-seq Analysis Using Cluster Computing. *Evolutionary Bioinformatics, 11*, 97-104.

Kunze, T. (2011). Dillwynella voightae new species, a new skeneimorph gastropod (Turbinidae) from the western Atlantic and a new record of *Dillwynella modesta* (Dall, 1889). *The Nautilus, 125*, 36-40.

Lanfear, R., Calcott, B., Ho, S. Y. W.& Guindon, S. (2012). PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Molecular Biology and Evolution, 29*, 1695-1701.

Lartillot, N., Brinkmann, H., & Philippe, H. (2007). Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. *BMC Evolutionary Biology, 7 Suppl 1*, S4.

Lartillot, N., & Philippe, H. (2004). A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Molecular Biology and Evolutio*n, *21*, 1095-1109.

Lartillot, N., Rodrigue, N., Stubbs, D., & Richer, 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.

Lindberg, D. R. 2008. Patellogastropoda, Neritimorpha and Cocculinoidea: The low-diversity gastropod clades. In: W. F. Ponder & D. R. Lindberg (Eds) *Phylogeny and Evolution of the Mollusca* (pp. 271-296). Berkeley: University of California Press.

Lindberg, D. R., & Ponder, W. F. (2001). The influence of classification on the evolutionary interpretation of structure a re-evaluation of the evolution of the pallial cavity of gastropod molluscs. *Organisms Diversity & Evolution*, *1*, 273-299.

Maynard, B., Kerr, L., McKiernan, J., Jansen, E., & Hanna, P. (2005). Mitochondrial DNA Sequence and Gene Organization in the Australian Blacklip Abalone *Haliotis rubra* (Leach). *Marine Biotechnology*, *7*, 645-658.

McArthur, A., & Harasewych, M. 2003. Molecular systematics of the major lineages of the Gastropoda. In: C. Lydeard & D. R. Lindberg (Eds) *Molecular Systematics and Phylogeography of Mollusks* (pp. 140-160). Washington DC: Smithsonian Books.

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

McLean, J. H. (1988). New Archaeogastropod Limpets from Hydrothermal Vents; Superfamily Lepetodrilacea I. Systematic Descriptions. *Phil. Trans. R. Soc. Lond.* (B), *319*, 1-32.

Nakagawa, S., Shimamura, S., Takaki, Y., Suzuki, Y., Murakami, S.-i., Watanabe, T., et al. (2014). Allying with armored snails: the complete genome of gammaproteobacterial endosymbiont. *ISME J*, *8*, 40-51.

Osca, D., Irisarri, I., Todt, C., Grande, C., & Zardoya, R. (2014a). The complete mitochondrial genome of *Scutopus ventrolineatus* (Mollusca:

Chaetodermomorpha) supports the Aculifera hypothesis. *BMC Evolutionary Biology*, *14*, 1-10.

Osca, D., Templado, J., & Zardoya, R. (2014b). The mitochondrial genome of *Ifremeria nautilei* and the phylogenetic position of the enigmatic deep-sea Abyssochrysoidea (Mollusca: Gastropoda). *Gene*, *547*, 257-266.

Palumbi, S., Martin A, McMillan, W., Stice, L., & Grabowski, G. (1991). The Simple Fool's Guide to PCR. http://palumbi.stanford.edu/SimpleFoolsMaster.pdf, 1-45.

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.

Robinson, N. A., Hall, N. E., Ross, E. M., Cooke, I. R., Shiel, B. P., Robinson, A. J., et al. (2014). The complete mitochondrial genome of *Haliotis laevigata* (Gastropoda: Haliotidae) using MiSeq and HiSeq sequencing. *Mitochondrial DNA*, doi:10.3109/19401736.2014.900611

Ronquist, F. & Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics, 19*, 1572-1574.

Rota-Stabelli, O., Yang, Z. & Telford, M. J. (2009). MtZoa: A general mitochondrial amino acid substitutions model for animal evolutionary studies. *Molecular Phylogenetics and Evolution*, *52*, 268-272.

Salvini-Plawen, L. V. (1980). A reconsideration of systematics in the Mollusca (phylogeny and higher classification). *Malacologi*a, 19, 249-278.

Salvini-Plawen, L. V., & Haszprunar, G. (1987). The Vetigastropoda and the systematics of streptoneurous Gastropoda (Mollusca). *Journal of Zoology*, *211*, 747-770.

Sasaki, T. (1998). Comparative anatomy and phylogeny of the recent Archaeogastropoda (Mollusca: Gastropoda). *The University Museum, The University of Tokyo, Bulletin*, *38*, 1-223.

Schattner, P., Brooks, A. N. & Lowe, T. M. (2005). The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Research, 33(suppl 2)*, W686-W689.

Schmieder, R. & Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics, 27*, 863-864.

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

Simison, W., Lindberg, D., & Boore, J. (2006). Rolling circle amplification of metazoan mitochondrial genomes. *Molecular Phylogenetics and Evolution*, *39*, 562-567.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., & Flook, P. (1994). Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers. *Annals of the Entomological Society of America*, 87, 651-701.

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

St John, J. SeqPrep. Available via https://github.com/jstjohn/SeqPrep.

Stöger, I., & Schrödl, M. (2013). Mitogenomics does not resolve deep molluscan relationships (yet?). *Molecular Phylogenetics and Evolution, 69*, 376-392.

Stöger, I., Sigwart, J. D., Kano, Y., Knebelsberger, T., Marshall, B. A., Schwabe, E., et al. (2013). The Continuing Debate on Deep Molluscan Phylogeny: Evidence for Serialia (Mollusca, Monoplacophora + Polyplacophora). *BioMed Research International, 2013*, Article ID 407072.

Thiele, J. (1929–35). *Handbuch der Systematischen Weichtierkunde*. Jena: Gustav Fischer.

Wägele, J., & Mayer, C. (2007). Visualizing differences in phylogenetic information content of alignments and distinction of three classes of long-branch effects. *BMC Evolutionary Biology, 7*, 147.

Warén, A., Bengtson, S., Goffredi, S. K., & Van Dover, C. L. (2003). A hot-vent gastropod with iron sulfide dermal sclerites. *Science, 302*, 1007.

Warén, A., & Bouchet, P. (2001). Gastropoda and Monoplacophora from hydrothermal vents and seeps; New taxa and records. *Veliger, 44*, 116-231.

White, T., Conrad, M., Tseng, R., Balayan, S., Golding, R., de Frias Martins, A., et al. (2011). Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. *BMC Evolutionary Biology, 11*, 295.

Williams, S. T. (2012). Advances in molecular systematics of the vetigastropod superfamily Trochoidea. *Zoologica Scripta*, *41*, 571-595.

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.

Williams, S. T., Karube, S., & Ozawa, T. (2008). Molecular systematics of Vetigastropoda: Trochidae, Turbinidae and Trochoidea redefined. *Zoologica Scripta*, *37*, 483-506.

Williams, S. T., & Ozawa, T. (2006). Molecular phylogeny suggests polyphyly of both the turban shells (family Turbinidae) and the superfamily Trochoidea (Mollusca: Vetigastropoda). *Molecular Phylogenetics and Evolution, 39*, 33-51.

Xin, Y., Ren, J., & Liu, X. (2011). Mitogenome of the small abalone *Haliotis diversicolor* Reeve and phylogenetic analysis within Gastropoda. *Marine Genomics, 4*, 253-262.

Yoon, S. H., & Kim, W. (2005). Phylogenetic relationships among six vetigastropod subgroups (Mollusca, Gastropoda) based on *18S* rDNA sequences. *Molecules and Cells*, *19*, 283-288.

Zapata, F., Wilson, N. G., Howison, M., Andrade, S. C. S., Jörger, K. M., Schrödl, M., et al. (2014). Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. Proceeding of the Royal Society of London series, *281*, 20141739.

Legends to figures

Fig.1 Mitochondrial gene orders of main lineages of Vetigastropoda. Gene orders in the hypothetical ancestral gastropod and Neomphalina are shown for comparison. Genes translocated are colored in blue; inversions are in pink; genes translocated and inverted are in green. Genes encoded by the minor strand are underlined. Shaded boxes indicate regions not sequenced in *L. schrolli*. Gaps are introduced to accommodate translocations except in Fissurelloidea and Neomphalina due to their high number of rearrangements.

Fig. 2 Phylogenetic relationships of Gastropoda based on mitochondrial sequence data. The ML phylogram is shown (A). Topology differences in BI are shown in the inset (B). Numbers at nodes are support values from BI (posterior probabilities) and ML (bootstrap proportions). Branch colors indicate main gastropod lineages. Scale bar indicates substitutions/site.

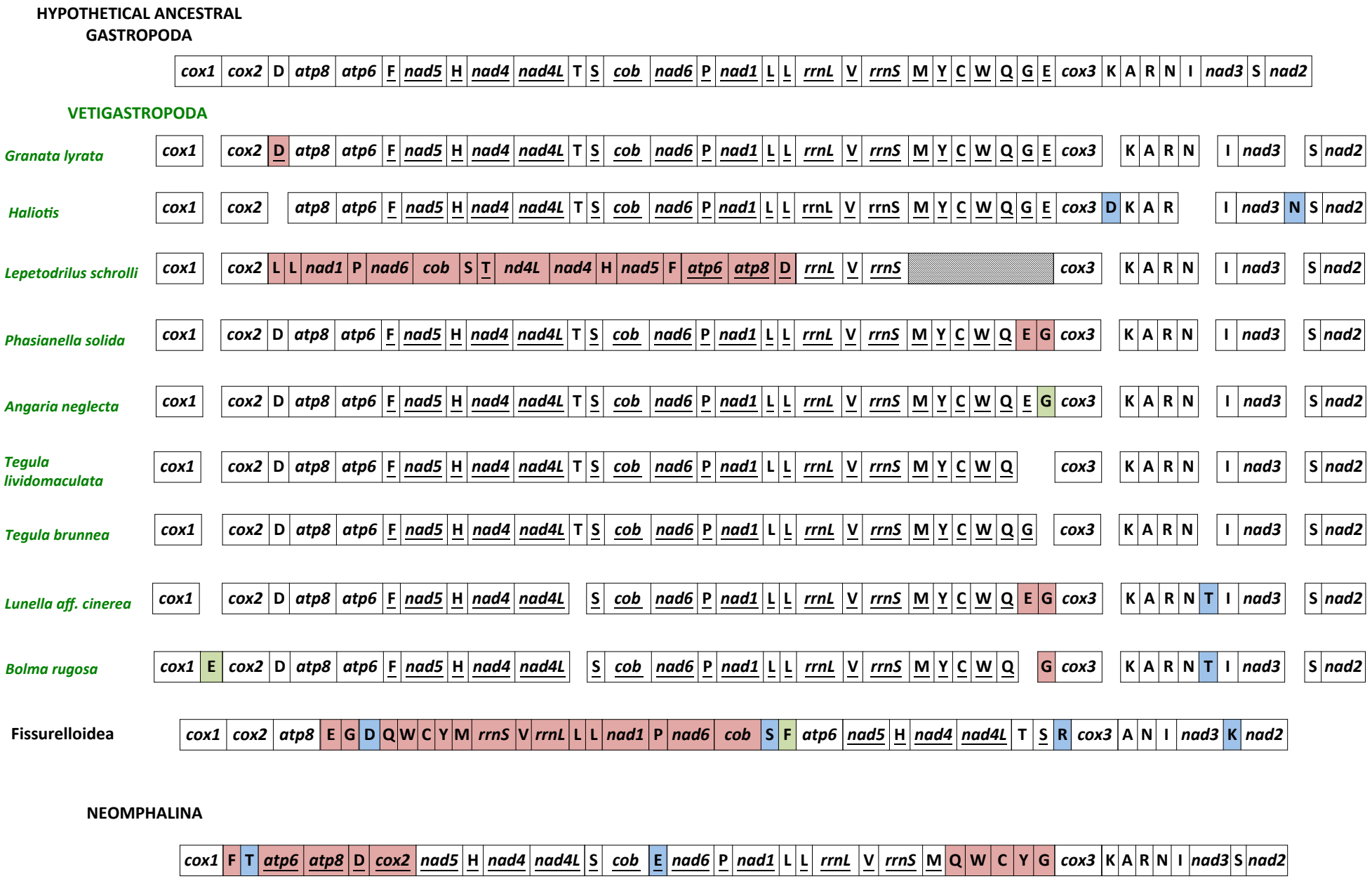
Fig.3 Phylogenetic relationships of Vetigastropoda based on mitochondrial sequence data. The ML phylogram is shown (A). Topology differences in BI are shown in the inset (B). Numbers at nodes are support values from BI (posterior probabilities) and ML (bootstrap proportions). Branch colors indicate main vetigastropod superfamilies. Scale bar indicates substitutions/site.

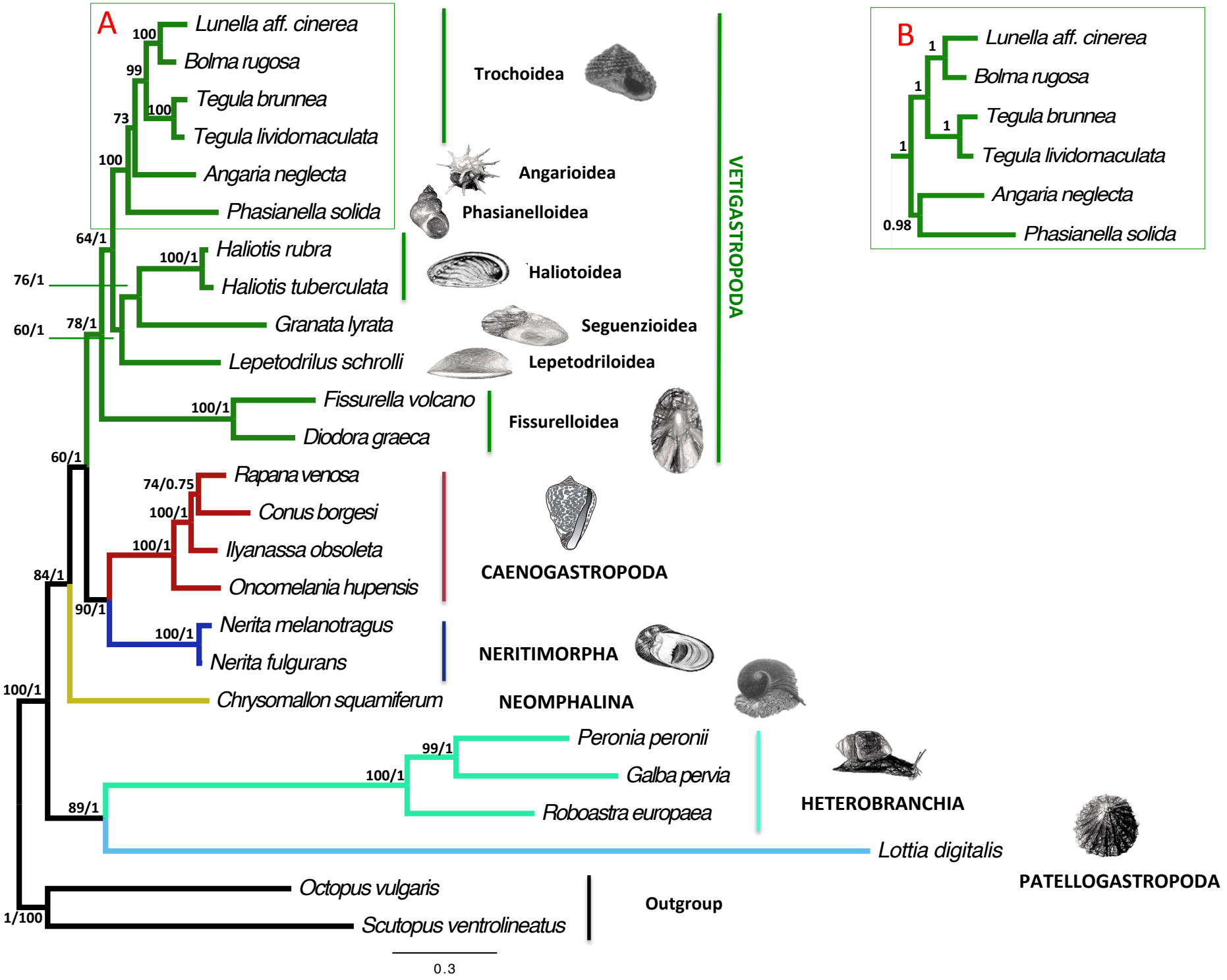
Table 1. Complete mitochondrial (mt) genomes analyzed in this study

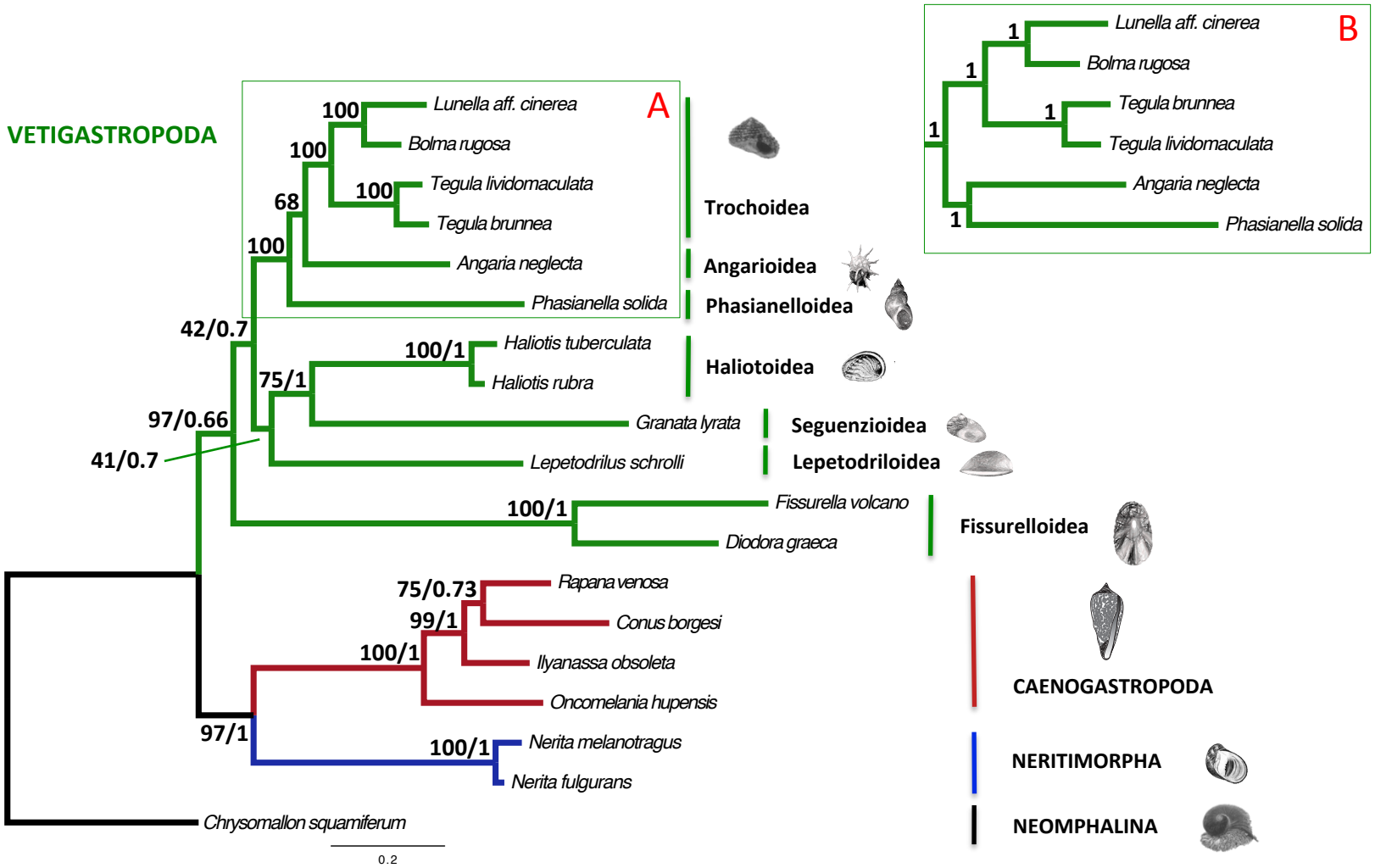
New mt genomes

Species	Superfamily	Length (bp)	GenBank Acc. No.	Location	Habitat	Voucher (MNCN/ADN)
Phasianella solida	Phasianelloidea	16698	KR297251	Bounotsu, Kagoshima, Kyushu, Japan	Rocky shore, intertidal	85259
Angaria neglecta	Angarioidea	19470	KR297248	Tsuji Is., Amakusa, Kumamoto, Kyushu, Japan	Rocky shore, intertidal	85258
*Lepetodrilus schrolli**	Lepetodriloidea	15579	KR297250	North Fiji Basin, South Pacific	Hydrothermal vent, 1990 m	85261
Granata lyrata	Seguenzioidea	17632	KR297249	Bounotsu, Kagoshima, Kyushu, Japan	Rocky shore, intertidal	85260
Bolma rugosa	Trochoidea	17432	KT207824	Islas Chafarinas, Spain	Rocky shore, intertidal	85637
Diodora graeca	Fissurelloidea	17209	KT207825	Cabo de Palos, Murcia, Spain	Rocky shore, intertidal	85530
Tegula lividomaculata	Trochoidea	17375	KT207826	Playa Girón, Bahía de Cochinos, Cuba	Rocky shore, intertidal	85638

*nearly complete mt genome


HYPOTHETICAL ANCESTRAL GASTROPODA
cox1 cox2 D atp8 atp6 F nad5 H nad4 nad4L T S cob nad6 P nad1 L L rrnL V rrnS M Y C W Q G E cox3 K A R N I nad3 S nad2
VETIGASTROPODA
Granata lyrata
Haliotis
Lepetodrilus schrolli
Phasianella solida
Angaria neglecta
Tegula lividomaculata
Tegula brunnea
Lunella aff. cinerea
Bolma rugosa
Fissurelloidea
cox1 cox2 atp8 E G D Q W C Y M rrnS V rrnL L L nad1 P nad6 cob S F atp6 nad5 H nad4 nad4L T S R cox3 A N I nad3 K nad2
NEOMPHALINA
cox1 F T atp6 atp8 D cox2 nad5 H nad4 nad4L S cob E nad6 P nad1 L L rrnL V rrnS M Q W C Y G cox3 K A R N I nad3 S nad2


A
Lunella aff. cinerea
Bolma rugosa
Tegula brunnea
Tegula lividomaculata
Angaria neglecta
Phasianella solida
100
99
100
73
100
64/1
100/1
76/1
60/1
78/1
100/1
60/1
74/0.75
100/1
100/1
84/1
90/1
100/1
100/1
89/1
99/1
100/1
1/100
Haliotis rubra
Haliotis tuberculata
Granata lyrata
Lepetodrilus schrolli
Fissurella volcano
Diodora graeca
Rapana venosa
Conus borgesi
Ilyanassa obsoleta
Oncomelania hupensis
Nerita melanotragus
Nerita fulgurans
Chrysomallon squamiferum
Peronia peronii
Galba pervia
Roboastra europaea
Lottia digitalis
Octopus vulgaris
Scutopus ventrolineatus
Trochoidea
Angarioidea
Phasianelloidea
Haliotoidea
Seguenzioidea
Lepetodriloidea
Fissurelloidea
VETIGASTROPODA
CAENOGASTROPODA
NERITIMORPHA
NEOMPHALINA
HETEROBRANCHIA
PATELLOGASTROPODA
Outgroup
0.3
B
1
1
1
1
0.98
Lunella aff. cinerea
Bolma rugosa
Tegula brunnea
Tegula lividomaculata
Angaria neglecta
Phasianella solida


A
B
Lunella aff. cinerea
Bolma rugosa
Tegula lividomaculata
Tegula brunnea
Angaria neglecta
Phasianella solida
100
100
100
68
100
1
1
1
1
1
VETIGASTROPODA
Trochoidea
Angarioidea
Phasianelloidea
42/0.7
100/1
75/1
Haliotis tuberculata
Haliotis rubra
Haliotoidea
97/0.66
41/0.7
Granata lyrata
Lepetodrilus schrolli
Seguenzioidea
Lepetodriloidea
100/1
Fissurella volcano
Diodora graeca
Fissurelloidea
75/0.73
99/1
100/1
Rapana venosa
Conus borgesi
Ilyanassa obsoleta
Oncomelania hupensis
CAENOGASTROPODA
97/1
100/1
Nerita melanotragus
Nerita fulgurans
NERITIMORPHA
Chrysomallon squamiferum
NEOMPHALINA
0.2

Supplementary Material 1. Long PCR and primer walking primers

Angaria neglecta

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
ANcox1HF	GCTACTATCTTTACCGGTTTTGGCTGGGGC	*cox1-rrnL* (12028)
AN16sHF	CGGTTAAACGAGGGCCATGCTGTCTCCTC	
AN16sHR	ATCTTAGTCCAACATCGAGGTCGTAAACC	*rrnL-cox3* (4079)
ANcox3R	AACAGCAGTATTCAATAGCGGAACCTG	
ANcox3HF	TAAGGTTTCTAGTGGGTTGCGTTGAGG	*cox3-cox1* (3480)
ANcox1HR	TGACCTAACTCAGCCCGAATCAAAAGTCT	

Primer walking

Primer	Sequence 5'-3'
AN12sF	AAGGTGAGGTTGATCGTGGACTATCG
ANTyrF	AGATCTACAGTCTTTCGCTTCCTTGC
ANcox3RIn	CAAGGACTAAACTCAACTAAATGAAACGG

Granata lyrata

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
GLcox1HF	GGCACCAGACATAGCCTTTCCTCGGCTC	*cox1-rrnL* (10951)
GL16sHF	GGGACAAGAAGACCCTATCGAGCTTTAGTGGC	
GL16sHR	ATCTTAGTCCAACATCGAGGTCGCAAAC	*rrnL-cox1* (5994)
GLcox1HR	ACTAGAGACGACCTCAGTAATAGGGCTA	

Phasianella solida

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
PHscox1F_1	GTTGCTGTCTTTGCCTGTGTTAGCTGGGGC	*cox1-rrnS* (11731)
PHs12SF_1	CCAGCCTGTATACCGTCGTCACCAGATCAC	
PHs12SR_1	CATTAGCTGCACCTTGATCTGACATGGA	*rrnS-cox1* (5085)
PHscox1R_1	TGCACCCAAAATAGAAGAAATACCTGCCAAG	

Primer walking

Primer	Sequence 5'-3'
Pha_12SRW1	ATTCGTCCAAATACTGTAGTTTAAGGGC
Phacox3RW1	AATTTAAGTGATAGAACCGGAAGCCACC
PhaNAD2R	AACAACAAAGACAGGTAATAATACAGCC
Phacox3FW1	TACTCTTAGGTGTATACTTTACGGTGC
Phanad2R2W	CTTCCTACTATAAGTAACCCAGAACCC
Phanad3FW	AAATATGGGAGAACGATACCCCTTTGC

Diodora graeca

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
40DGcox1F	TTTCTTGTGATGCCTATAATGATTGGGGG	*cox1-rrnL* (5573)
40DG16SR	TGTTATCCCCACGGTAACTTATTCTTCC	
40DGcox1R	ACAGCACCCAAAATAGAMGACACACC	*rrnL-cox1* (12105)
40DG16sR	ACCCCATCGAGCTTTAGTGGAATTTTGG	

Bolma rugosa

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
BRcox1F	GCTCCAGATATAGCATTTCCTCGTCTTAAT	*cox1-rrnL* (10837)
BR16SF	CGACCTCGATGTTGGACTAAGATATC	

Primer walking

Primer	Sequence 5'-3'
BRLeuR	GCTTAAACCTAATGCACTAATCTGCC
BR16sRW1	CACTAAAGCTCAACGGGGTCTTCTTGTCCCT
BR16sRW2	TCTTCTTGTCCCTCAGTTAAATGTTAGGC
BR16sRW3	AAAGTTTCGGAAGGCATTTTACCCCT
BRTrpF	GCAAGTTTAAAGGTGTATAGTTTGTACC
BRQF	TACTTGGAGTTTTGATCTCTGCGGG
BRcox3R	CTGTTGCCGTGAGTCCTTGAAGTCCACC
BRcox3F	GGGTTCTGGGGTAACAGTAACTTGAGCTC
BRAla_F	GTACTAGGAAGTGAGAAAATTACATGCG
BRnad3F	CCTGTAATTAAGATTTCTGGTGGAATGG
BRcox12R	TCCCGAGAATAAGGTATAATGTCCC
BRcox11R	ACAGCCCCTAGAATAGATGAAATACCTGCA

Tegula lividomaculata

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
86TLcox1F	GCTGCTGTAGAAAGAGGGGCCGGTACTG	*cox1-rrnS* (12850)
86TL12SF	GGCGGTGTCTTAAGTCCTTCTAGGGGAACC	
86TL86cox1R	TCCCGCTAATACAGGAAGAGACAACAAC	*cox3-cox1* (3161)
86TL86cox3F	CTTCTTTTGCCATTTCGGACGGAGC	

Primer walking

Primer	Sequence 5'-3'
86TLcox3RW	AGCCTGGAGTCGAAATAAGCAAACCC
86TLcox3RW2	CTCCGTCCGAAATGGCAAAAGAAGC
86TL12S-MRW	CTTGCTTTTAACAGAGGATACATCCG
86TL12SR2W	TGGACTATCGATTATAGGACAGGTTCCC
86TL12SR1W	CCATCTCTACCTTTTCATTAGCTGCACCT

Lepetodrilus schroli

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
LScox1F	TGACATCTGCCGCTGTAGAAAGAGGTGCTGG	*cox1-rrnS* (11602)
LS12SF	AACCTGCCCCATAAACTGATGATCCAC	
LScox1R	CCACCTCCTGCCGGGTCGAAGAAAGAG	*rrnS-cox1*(5100)*
LS12SR	CCCACCTTTCCGCCTTATTATAAGCTGCACC	

Primer walking

Primer	Sequence 5'-3'
LScox3RW	ATCCTAATTCTGGAGTTGGGGCAAGTC
LS12SR	TTATAAGCTGCACCTCGATCTGACGTC
LS12sRW2	TTCCTGCCTATACTCACCAGATCCC
LSContRw	ACTTTGCAAAGTTGCGAATGAGCTCAG

*Approximate based on the agarose gel

Supplementary Material 2. Complete mitochondrial (mt) genomes retrieved from GenBank and analyzed in this study

Species	Superfamily	Length (bp)	GenBank Acc. No.	Reference
Tegula brunnea	Trochoidea	17690	NC_016954	Simison, 2011 (unpublished)
Lunella aff. cinerea	Trochoidea	17670	KF700096	Williams et al., 2014
Haliotis rubra	Haliotoidea	16907	NC_005940	Maynard et al., 2005
Haliotis tuberculata	Haliotoidea	16521	NC_013708	VanWormhoudt et al., 2009
Fissurella volcano	Fissurelloidea	17575	NC_016953	Simison, 2011 (unpublished)
Chrysomallon squamiferum	Neomphaloidea	15388	AP013032	Nakagawa et al., 2014
Lottia digitalis	Lottioidea	26835	NC_007782	Simison et al., 2006
*Nerita fulgurans**	Neritoidea	15261	KF728888	Arquez et al., 2014
*Nerita melanotragus**	Neritoidea	15261	GU810158	Castro and Colgan 2010
Oncomelania hupensis	Truncatelloidea	15182	NC_013073	Li and Zhou, 2009 (unpublished)
Ilyanassa obsoleta	Buccinoidea	15263	NC_007781	Simison et al., 2006
Rapana venosa	Muricoidea	15272	NC_011193	Chandler et al., 2008 (unpublished)
Conus borgesi	Conoideea	15536	NC_013243	Cunha et al., 2009
Galba pervia	Lymnaeoidea	13768	NC_018536	Liu et al., 2012
Peronia peronii	Onchidioidea	13968	NC_016181	White et al., 2011
Roboastra europaea	Anadoridoidea	14472	NC_004321	Grande et al., 2002
Octopus vulgaris	Neocoleoidea	15744	NC_006353	Yokobori et al., 2004
Scutopus ventrolineatus	Scutopodidae**	14662	NC_025284	Osca et al., 2014

*nearly complete mt genomes

**unassigned to a superfamily

References

Arquez, M., Colgan, D., & Castro, L. R. (2014). Sequence and comparison of mitochondrial genomes in the genus *Nerita* (Gastropoda: Neritimorpha: Neritidae) and phylogenetic considerations among gastropods. *Marine Genomics*, 15, 45-54.

Castro, L. R., & Colgan, D. J. (2010). The phylogenetic position of Neritimorpha based on the mitochondrial genome of *Nerita melanotragus* (Mollusca: Gastropoda). *Molecular Phylogenetics and Evolution*, 57, 918-923.

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

Grande, C., Templado, J., Lucas Cervera, J., & Zardoya, R. (2002). The Complete Mitochondrial Genome of the Nudibranch *Roboastra europaea* (Mollusca: Gastropoda) Supports the Monophyly of Opisthobranchs. *Molecular Biology and Evolution*, 19, 1672-1685.

Liu, G.-H., Wang, S.-Y., Huang, W.-Y., Zhao, G.-H., Wei, S.-J., Song, H.-Q., et al. (2012). The Complete Mitochondrial Genome of *Galba pervia* (Gastropoda: Mollusca), an Intermediate Host Snail of *Fasciola* spp. PLoS ONE, 7, e42172.

Maynard, B., Kerr, L., McKiernan, J., Jansen, E., & Hanna, P. (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., et al. (2014). Allying with armored snails: the complete genome of gammaproteobacterial endosymbiont. ISME J, 8, 40-51.

Osca, D., Irisarri, I., Todt, C., Grande, C., & Zardoya, R. (2014). The complete mitochondrial genome of *Scutopus ventrolineatus* (Mollusca: Chaetodermomorpha) supports the Aculifera hypothesis. *BMC Evolutionary Biology*, 14, 1-10.

Simison, W., Lindberg, D., & Boore, J. (2006). Rolling circle amplification of metazoan mitochondrial genomes. Molecular Phylogenetics and Evolution, 39, 562-567.

Yokobori, S.-i., Fukuda, N., Nakamura, M., Aoyama, T., & Oshima, T. (2004). Long-Term Conservation of Six Duplicated Structural Genes in Cephalopod Mitochondrial Genomes. *Molecular Biology and Evolution*, 21, 2034-2046.

Van Wormhoudt, A., Le Bras, Y., & Huchette, S. (2009). *Haliotis marmorata* from Senegal; a sister species of Haliotis tuberculata: Morphological and molecular evidence. Biochemical Systematics and Ecology, 37, 747-755.

White, T., Conrad, M., Tseng, R., Balayan, S., Golding, R., de Frias Martins, A., et al. (2011). Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. *BMC Evolutionary Biology*, 11, 295.

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.

Supplementary Material 3. Selected best fit partitions and models

	Partition selected (PartitionFinder)	Model selected (PartitionFinder)	RaxML (-lnL) (PartitionFinder)	Model (MtZoa)	RaxML (-lnL) (MtZoa)	AIC (PartitionFinder)	AIC (MtZoa)
Gastropod	*atp6 + atp8*	MtArt+I+G	-5226.587607	MtZoa+I+G	-5199.492818	10453.175312	10398.985734
	cob	MtArt+I+G+F	-7017.424662	MtZoa+I+G+F	-6980.311661	14034.849460	13960.623458
	cox1	LG+I+G+F	-6108.531815	MtZoa+I+G+F	-6094.228767	12217.063766	12188.457670
	cox2	LG+I+G+F	-3651.186035	MtZoa+I+G+F	-3626.893971	7302.372206	7253.788078
	cox3	MtArt+G+F	-4317.787699	MtZoa+G+F	-4302.668418	8635.575532	8605.336970
	nad1	MtArt+I+G+F	-4286.383740	MtZoa+I+G+F	-4276.359446	8572.767616	8552.719028
	nad2	MtArt+I+G	-4103.552493	MtZoa+I+G	-4101.632823	8207.105084	8203.265744
	nad3	LG+G	-826.241576	MtZoa+G	-808.305197	16524.83248	16166.10490
	nad4 + nad4L	MtArt+G+F	-7614.716066	MtZoa+G+F	-7590.436652	15229.432266	15180.873438
	nad5	LG+I+G+F	-9484.075013	MtZoa+I+G+F	-9451.556845	18968.150162	18903.113826
	nad6	MtArt+G	-2604.453604	MtZoa+G	-2598.201473	5208.907304	5196.403042
	rrnL + rrnS	GTR+I+G	-16350.245	—	—	32814.489	—
Vetigastropod	*atp6 + atp8*	MtArt+I+G+F	-3870.311856	MtZoa+I+G+F	-3857.873546	7740.623824	7715.747204
	cob	MtArt+I+G+F	-4724.556803	MtZoa+I+G+F	-4702.951708	9449.113718	9405.903528
	cox1	LG+I+G+F	-3969.442087	MtZoa+I+G+F	-3957.914211	7938.884286	7915.828534
	cox2	LG+I+G+F	-2973.174396	MtZoa+I+G+F	-2951.638563	5946.348902	5903.277198
	cox3	MtArt+I+G+F	-2737.626254	MtZoa+I+G+F	-2727.576836	5475.252618	5455.153782
	nad1	MtArt+I+G+F	-3606.786211	MtZoa+I+G+F	-3600.434217	7213.572496	7200.868546
	nad2	MtArt+G+F	-3444.551420	MtZoa+G+F	-3427.286657	6889.102912	6854.573424
	nad3	LG+G+F	-1148.509381	MtZoa+G+F	-1136.311663	2297.018872	2272.623436
	nad4 + nad4L	MtArt+I+G+F	-6661.174569	MtZoa+I+G+F	-6641.039689	13322.349210	13282.079488
	nad5	LG+I+G+F	-8244.068129	MtZoa+I+G+F	-8209.163132	16488.136370	16418.326376
	nad6	MtArt+G	-2010.77533	MtZoa+G	-1996.776360	4021.55138	3993.552792
	rrnL + rrnS	GTR+I+G	-15269.621	—	—	30629.242	—

Supplementary Material 4. Annotation and main features of newly sequenced mt genomes

Angaria neglecta

Name Gene	Type	Start	Stop	Length	Codon start	Codon stop	Strand	%A-T
cox1	CDS	1	1536	1,536	ATG	TAG	forward	
	Intergenic			314				68,4
cox2	CDS	1851	2543	693	ATG	TAA	forward	
	Intergenic			353				66,6
trnD	tRNA	2897	2971	75			forward	
atp8	CDS	2972	3157	186	ATG	TAG	forward	
	Intergenic			213				68
atp6	CDS	3371	4063	693	ATG	TAG	forward	
	Intergenic			50				68
trnF	tRNA	4114	4183	70			reverse	
	Intergenic			41				70,7
nad5	CDS	4225	5967	1,743	ATG	TAA	reverse	
trnH	tRNA	5968	6032	65			reverse	
	Intergenic			218				78,4
nad4	CDS	6251	7642	1,392	ATG	TAA	reverse	
nad4L	CDS	7636	7935	300	ATG	TAG	reverse	
	Intergenic			43				69,8
trnT	tRNA	7979	8050	72			forward	
	Intergenic			21				76,2
trnS(uga)	tRNA	8072	8138	67			reverse	
	Intergenic			21				47,6
cob	CDS	8160	9299	1,14	ATG	TAA	reverse	
	Intergenic			257				62,3
nad6	CDS	9557	10063	507	ATG	TAA	reverse	
	Intergenic			4				
trnP	tRNA	10068	10138	71			reverse	
	Intergenic			487				74
nad1	CDS	10626	11570	945	ATG	TAG	reverse	
	Intergenic			3				
trnL(uaa)	tRNA	11574	11641	68			reverse	
	Intergenic			278				68
trn(uag)	tRNA	11920	11987	68			reverse	
rrnL	rRNA	11988	13666	1,679			reverse	
trnV	tRNA	13667	13735	69			reverse	
rrnS	rRNA	13736	14857	1,122			reverse	
trnM	tRNA	14858	14926	69			reverse	
	Intergenic			8				
trnY	tRNA	14935	15002	68			reverse	
	Intergenic			27				66,6
trnC	tRNA	15030	15096	67			reverse	
	Intergenic			48				66,7
trnW	tRNA	15145	15213	69			reverse	
	Intergenic			8				
trnQ	tRNA	15222	15290	69			reverse	
	Intergenic			51				64,7
trnE	tRNA	15342	15406	65			reverse	
	Intergenic			323				77,4
trnG	tRNA	15730	15797	68			forward	
	Intergenic			99				68,7
cox3	CDS	15897	16676	780	ATG	TAA	forward	
	Intergenic			195				59
trnK	tRNA	16872	16937	66			forward	
trnA	tRNA	16938	17008	71			forward	
	Intergenic			116				69
trnR	tRNA	17125	17193	69			forward	
	Intergenic			194				62,9
trnN	tRNA	17388	17459	72			forward	
	Intergenic			97				54,6
trnI	tRNA	17557	17625	69			forward	
	Intergenic			3				
nad3	CDS	17629	17982	354	ATG	TAG	forward	
	Intergenic			46				52,1
trnS(gcu)	tRNA	18029	18097	69			forward	
	Intergenic			3				
nad2	CDS	18101	19276	1,176	ATG	TAG	forward	
	Intergenic			194				67,5

Granata lyrata

Name Gene	Type	Start	Stop	Length	Codon start	Codon stop	Strand	%A-T
cox1	CDS	1	1587	1,587	ATG	TAA	forward	
	Intergenic			42				50
cox2	CDS	1630	2334	705	ATG	TAA	forward	
	Intergenic			68				57,4
trnD	tRNA	2565	2639	79			reverse	
	Intergenic			159				58,5
atp8	CDS	2641	3,051	411	ATG	TAA	forward	
	Intergenic			155				52,3
atp6	CDS	3207	3926	720	TTG	TAA	forward	
	Intergenic			27				62,9
TrnF	tRNA	3954	4019	66			reverse	
	Intergenic			6				
nad5	CDS	4026	5768	1,743	ATG	TAG	reverse	
	Intergenic			1				
trnH	tRNA	5770	5836	67			reverse	
	Intergenic			24				62,5
nad4	CDS	5861	7246	1,386	GTG	TAG	reverse	
nad4L	CDS	7240	7542	303	ATG	TAG	reverse	
	Intergenic			7				
trnT	tRNA	7550	7619	70			forward	
	Intergenic			2				
trnS(uga)	tRNA	7622	7687	66			reverse	
	Intergenic			7				
cob	CDS	7695	8831	1,137	ATG	TAA	reverse	
	Intergenic			38				56,4
nad6	CDS	8870	9373	504	ATG	TAG	reverse	
	Intergenic			3				
trnP	tRNA	9377	9446	70			reverse	
	Intergenic			58				53,2
nad1	CDS	9505	10626	1,122	ATG	TAG	reverse	
	Intergenic			1				
trnL(uaa)	tRNA	10628	10695	68			reverse	
	Intergenic			2				
trn(tag)	tRNA	10698	10766	69			reverse	
rrnL	rRNA	10767	12259	1,493			reverse	
trnV	tRNA	12260	12328	69			reverse	
rrnS	rRNA	12329	13506	1,178			reverse	
trnM	tRNA	13507	13576	70			reverse	
	Intergenic			10				
trnY	tRNA	13587	13654	68			reverse	
	Intergenic			13				
trnC	tRNA	13668	13734	67			reverse	
	Intergenic			11				
trnW	tRNA	13746	13824	79			reverse	
	Intergenic			7				
trnQ	tRNA	13832	13904	73			reverse	
	Intergenic			3				
trnG	tRNA	13908	13974	67			reverse	
	Intergenic			20				
trnE	tRNA	13995	14069	75			reverse	
	Intergenic			772				62,7
cox3	CDS	14842	15627	786	ATG	TAA	forward	
	Intergenic			63				39,7
trnK	tRNA	15691	15759	69			forward	
	Intergenic			24				62,5
trnA	tRNA	15784	15850	67			forward	
trnR	tRNA	15851	15907	57			forward	
	Intergenic			15				
trnN	tRNA	15923	15989	67			forward	
	Intergenic			2				
trnI	tRNA	15992	16068	77			forward	
	Intergenic			1				
nad3	CDS	16070	16423	354	ATG	TAA	forward	
	Intergenic			13				
trnS(gctu)	tRNA	16437	16504	68			forward	
	Intergenic			4				
nad2	CDS	16509	17597	1,089	ATG	TAA	forward	
	Intergenic			35				60

Phasianella solida

Name Gene	Type	Start	Stop	Length	Codon start	Codon stop	Strand	%A-T
cox1	CDS	1	1545	1,545	ATG		forward	
	Intergenic			110				62,7
cox2	CDS	1656	2354	699	ATG	TAA	forward	
	Intergenic			73				53,4
trnD	tRNA	2428	2495	68			forward	
atp8	CDS	2496	2687	192	ATG	TAA	forward	
	Intergenic			57				64,9
atp6	CDS	2745	3479	735	ATG	TAG	forward	
	Intergenic			67				67,2
trnF	tRNA	3547	3614	68			reverse	
	Intergenic			62				66,1
nad5	CDS	3677	5413	1,737	ATG	TAA	reverse	
	Intergenic			1				
trnH	tRNA	5415	5479	65			reverse	
	Intergenic			85				52,9
nad4	CDS	5565	6956	1,392	ATT	TAG	reverse	
nad4L	CDS	6950	7258	309	ATG	TAG	reverse	
	Intergenic			7				
trnT	tRNA	7266	7334	69			forward	
	Intergenic			7				
trnS(uga)	tRNA	7342	7408	67			reverse	
	Intergenic			6				
cob	CDS	7415	8554	1,14	ATG	TAA	reverse	
	Intergenic			106				65,1
nad6	CDS	8661	9164	504	ATG	TAG	reverse	
	Intergenic			1				
trnP	tRNA	9166	9231	66			reverse	
	Intergenic			82				51,2
nad1	CDS	9314	10258	945	ATG	TAG	reverse	
	Intergenic			1				
trnL(uaa)	tRNA	10260	10327	68			reverse	
	Intergenic			24				70,8
trnL(uag)	tRNA	10352	10420	69			reverse	
rrnL	rRNA	10421	11893	1,473			reverse	
trnV	tRNA	11894	11964	71			reverse	
rrnS	rRNA	11965	12951	987			reverse	
trnM	tRNA	12952	13018	67			reverse	
	Intergenic			7				
trnY	tRNA	13026	13091	66			reverse	
trnC	tRNA	13092	13156	65			reverse	
	Intergenic			23				60,9
trnW	tRNA	13180	13246	67			reverse	
	Intergenic			7				
trnQ	tRNA	13254	13322	69			reverse	
	Intergenic			250				79,2
trnE	tRNA	13573	13637	65			forward	
trnG	tRNA	13638	13704	67			forward	
	Intergenic			5				
cox3	CDS	13710	14489	780	ATG	TAA	forward	
	Intergenic			59				61
trnK	tRNA	14549	14607	59			forward	
trnA	tRNA	14608	14677	70			forward	
	Intergenic			42				59,5
trnR	tRNA	14720	14786	67			forward	
	Intergenic			19				
trnN	tRNA	14806	14879	74			forward	
	Intergenic			10				
trnI	tRNA	14890	14957	68			forward	
nad3	CDS	14958	15311	354	ATG	TAG	forward	
	Intergenic			97				62,9
trnS(gcu)	tRNA	15409	15476	68			forward	
nad2	CDS	15477	16616	1,14	ATG	TAA	forward	
	Intergenic			82				68,3

Diodora graeca

Name Gene	Type	Start	Stop	Length	Codon start	Codon stop	Strand	%A-T
cox1	CDS	1	1533	1,533	ATG	TAA	forward	
	Intergenic			38				71,1
cox2	CDS	1572	2264	693	ATG	TAA	forward	
	Intergenic			92				60,9
atp8	CDS	2357	2563	207	ATG	TAG	forward	
	Intergenic			153				50,3
trnE	tRNA	2717	2784	68			forward	
	Intergenic			17				
trnG	tRNA	2802	2871	70			forward	
	Intergenic			28				64,3
trnD	tRNA	2900	2968	69			forward	
	Intergenic			21				71,4
trnQ	tRNA	2990	3060	71			forward	
	Intergenic			8				
trnW	tRNA	3069	3137	69			forward	
	Intergenic			48				50
trnC	tRNA	3186	3256	71			forward	
	Intergenic			3				
trnY	tRNA	3260	3330	71			forward	
	Intergenic			29				65,5
trnM	tRNA	3360	3427	68			forward	
rrnS	rRNA	3428	4476	1049			forward	
trnV	tRNA	4477	4545	69			forward	
rrnL	rRNA	4546	6022	1477			forward	
trnL(uag)	tRNA	6023	6091	69			forward	
	Intergenic			6				
trnL(uaa)	tRNA	6098	6166	69			forward	
	Intergenic			3				
nad1	CDS	6170	7144	975	GTG	TAA	forward	
	Intergenic			34				52,9
trnP	tRNA	7179	7248	70			forward	
	Intergenic			3				
nad6	CDS	7252	7776	525	ATG	TAG	forward	
	Intergenic			18				
cob	CDS	7795	8934	1140	ATG	TAG	forward	
	Intergenic			26				42,3
trnS(ucu)	tRNA	8961	9033	73			forward	
	Intergenic			16				
trnF	tRNA	9050	9116	67			forward	
	Intergenic			692				60,5
atp6	CDS	9809	10510	702	ATG	TAG	forward	
	Intergenic			70				74,3
nad5	CDS	10581	12299	1719	ATG	TAA	reverse	
trnH	tRNA	12300	12363	64			reverse	
	Intergenic			16				
nad4	CDS	12380	13786	1407	GTG	TAA	reverse	
nad4L	CDS	13780	14076	297	ATG	TAA	reverse	
	Intergenic			11				
trnT	tRNA	14088	14152	65			forward	
	Intergenic			4				
trnS(uga)	tRNA	14157	14224	68			reverse	
	Intergenic			49				81,6
trnR	tRNA	14274	14344	71			forward	
	Intergenic			38				63,2
cox3	CDS	14383	15162	780	ATG	TAG	forward	
	Intergenic			39				43,6
trnA	tRNA	15202	15270	69			forward	
	Intergenic			13				
trnN	tRNA	15284	15354	71			forward	
	Intergenic			15				
trnI	tRNA	15370	15437	68			forward	
	Intergenic			1				
nad3	CDS	15439	15792	354	ATG	TAG	forward	
	Intergenic			9				
trnK	tRNA	15802	15871	70			forward	
	Intergenic			111				57,7
nad2	CDS	15983	17088	1106	ATG	TAA	forward	
	Intergenic			121				60,3

%A-T was calculated only for intergenic regions >20 pb
red cell indicates %A-T above 70%

Bolma rugosa

Name Gene	Type	Start	Stop	Length	Codon start	Codon stop	Strand	%A-T
cox1	CDS	1	1536	1536	ATG	TAA	forward	
	Intergenic			1				
trnE	tRNA	1538	1606	69			forward	
cox2	CDS	1607	2296	690	ATG	TAA	forward	
	Intergenic			136				86,8
trnD	tRNA	2433	2504	72			forward	
atp8	CDS	2505	2684	180	ATG	TAA	forward	
	Intergenic			85				85,9
atp6	CDS	2770	3465	696	ATG	TAA	forward	
	Intergenic			33				93,9
trnF	tRNA	3498	3566	69			reverse	
	Intergenic			256				78,1
nad5	CDS	3823	5568	1746	ATG	TAA	reverse	
trnH	tRNA	5569	5635	67			reverse	
	Intergenic			93				82,8
nad4	CDS	5729	7120	1392	ATG	TAA	reverse	
nad4L	CDS	7114	7413	300	ATG	TAA	reverse	
	Intergenic			69				84,1
trnS(uga)	tRNA	7483	7549	67			reverse	
	Intergenic			14				
cob	CDS	7564	8703	1140	ATG	TAG	reverse	
	Intergenic			139				83,5
nad6	CDS	8843	9349	507	ATG	TAA	reverse	
	Intergenic			5				
trnP	tRNA	9355	9428	74			reverse	
	Intergenic			240				79,6
nad1	CDS	9669	10616	948	ATG	TAG	reverse	
	Intergenic			4				
trnL(uaa)	tRNA	10621	10688	68			reverse	
	Intergenic			87				80,5
trnL(uag)	tRNA	10776	10843	68			reverse	
rrnL	rRNA	10844	12466	1623			reverse	
trnV	tRNA	12467	12535	69			reverse	
rrnS	rRNA	12536	13582	1047			reverse	
trnM	tRNA	13583	13651	69			reverse	
	Intergenic			80				82,5
trnY	tRNA	13732	13801	70			reverse	
trnC	tRNA	13801	13867	67			reverse	
	Intergenic			5				
trnW	tRNA	13873	13943	71			reverse	
	Intergenic			1				
trnQ	tRNA	13945	14013	69			reverse	
	Intergenic			11				
trnG	tRNA	14025	14092	68			forward	
	Intergenic			38				76,3
cox3	CDS	14131	14910	780	ATG	TAA	forward	
	Intergenic			107				72
trnK	tRNA	15018	15081	64			forward	
trnA	tRNA	15082	15150	69			forward	
	Intergenic			21				81
trnR	tRNA	15172	15240	69			forward	
	Intergenic			132				74,2
trnN	tRNA	15373	15443	71			forward	
	Intergenic			20				65
trnT	tRNA	15464	15535	72			forward	
	Intergenic			69				81,2
trnI	tRNA	15605	15672	68			forward	
	Intergenic			3				
nad3	CDS	15676	16029	354	ATG	TAA	forward	
	Intergenic			141				79,4
trnS(gcu)	tRNA	16171	16238	68			forward	
	Intergenic			3				
nad2	CDS	16242	17396	1155	ATG	TAA	forward	
	Intergenic			36				77,8

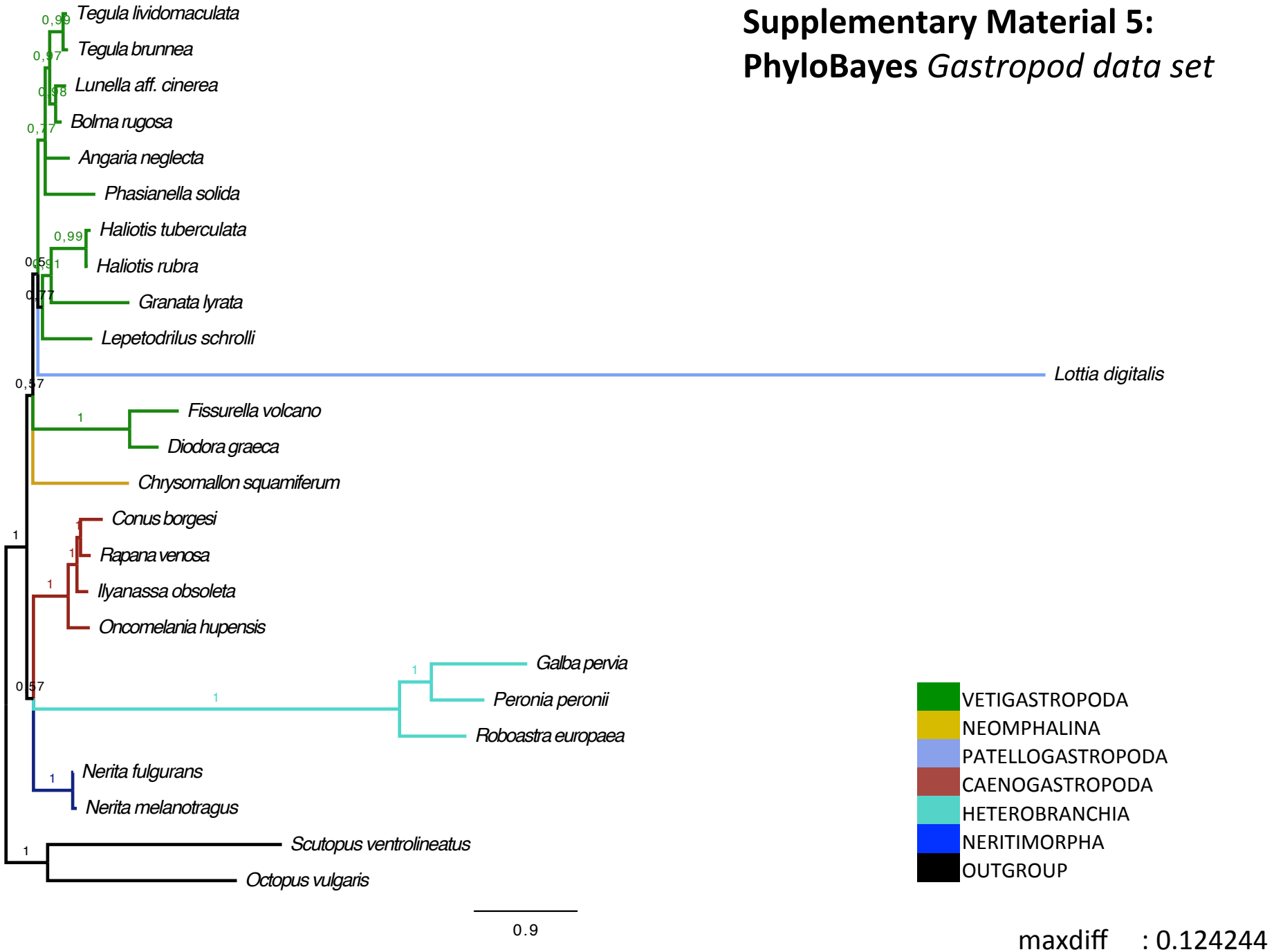
Tegula lividomaculata

Name Gene	Type	Start	Stop	Length	Codon start	Codon stop	Strand	%A-T
cox1	CDS	1	1536	1,536	ATG	TAA	forward	
	Intergenic			72				83,3
cox2	CDS	1609	2304	696	ATG	TAA	forward	
	Intergenic			158				80,4
trnD	tRNA	2463	2537	75			forward	
atp8	CDS	2538	2714	177	ATG	TAG	forward	
	Intergenic			245				80
atp6	CDS	2960	3658	699	ATG	TAA	forward	
	Intergenic			49				87,8
trnF	tRNA	3708	3776	69			reverse	
	Intergenic			357				76,5
nad5	CDS	4134	5873	1,74	ATG	TAG	reverse	
trnH	tRNA	5874	5940	67			reverse	
	Intergenic			69				76,8
nad4	CDS	6010	7404	1,395	ATG	TAA	reverse	
nad4L	CDS	7398	7697	300	ATG	TAA	reverse	
	Intergenic			83				75,9
trnT	tRNA	7781	7855	75			forward	
	Intergenic			91				80,2
trnS(uga)	tRNA	7947	8012	66			reverse	
	Intergenic			13				
cob	CDS	8026	9165	1,14	ATG	TAA	reverse	
	Intergenic			96				81,2
nad6	CDS	9262	9768	507	ATG	TAA	reverse	
	Intergenic			4				
trnP	tRNA	9773	9841	69			reverse	
	Intergenic			25				68
nad1	CDS	9867	1082	954	ATG	TAA	reverse	
	Intergenic			1				
trnL(uaa)	tRNA	10822	10889	68			reverse	
	Intergenic			91				67
trnL(uag)	tRNA	10981	11048	68			reverse	
rrnL	rRNA	11049	12625	1,577			reverse	
trnV	tRNA	12626	12695	70			reverse	
rrnS	rRNA	12696	13759	1,064			reverse	
trnM	tRNA	13760	13829	70			reverse	
	Intergenic			31				64,5
trnY	tRNA	13861	13927	67			reverse	
	Intergenic			6				
trnC	tRNA	13934	14009	76			reverse	
	Intergenic			3				
trnW	tRNA	14013	14079	67			reverse	
	Intergenic			2				
trnQ	tRNA	14082	14150	69			reverse	
	Intergenic			123				74,8
cox3	CDS	14274	15053	780	ATG	TAA	forward	
	Intergenic			70				82,9
trnK	tRNA	15124	15183	60			forward	
trnA	tRNA	15184	15252	69			forward	
	Intergenic			87				75,9
trnR	tRNA	15340	15408	69			forward	
	Intergenic			12				
trnN	tRNA	15421	15488	68			forward	
	Intergenic			42				83,3
trnI	tRNA	15531	15599	69			forward	
	Intergenic			4				
nad3	CDS	15604	15957	354	ATG	TAA	forward	
	Intergenic			158				75,9
trnS(gcu)	tRNA	16116	16183	68			forward	
	Intergenic			3				
nad2	CDS	16187	17350	1,164	ATG	TAA	forward	
	Intergenic			25				68

Lepetodrilus schroli

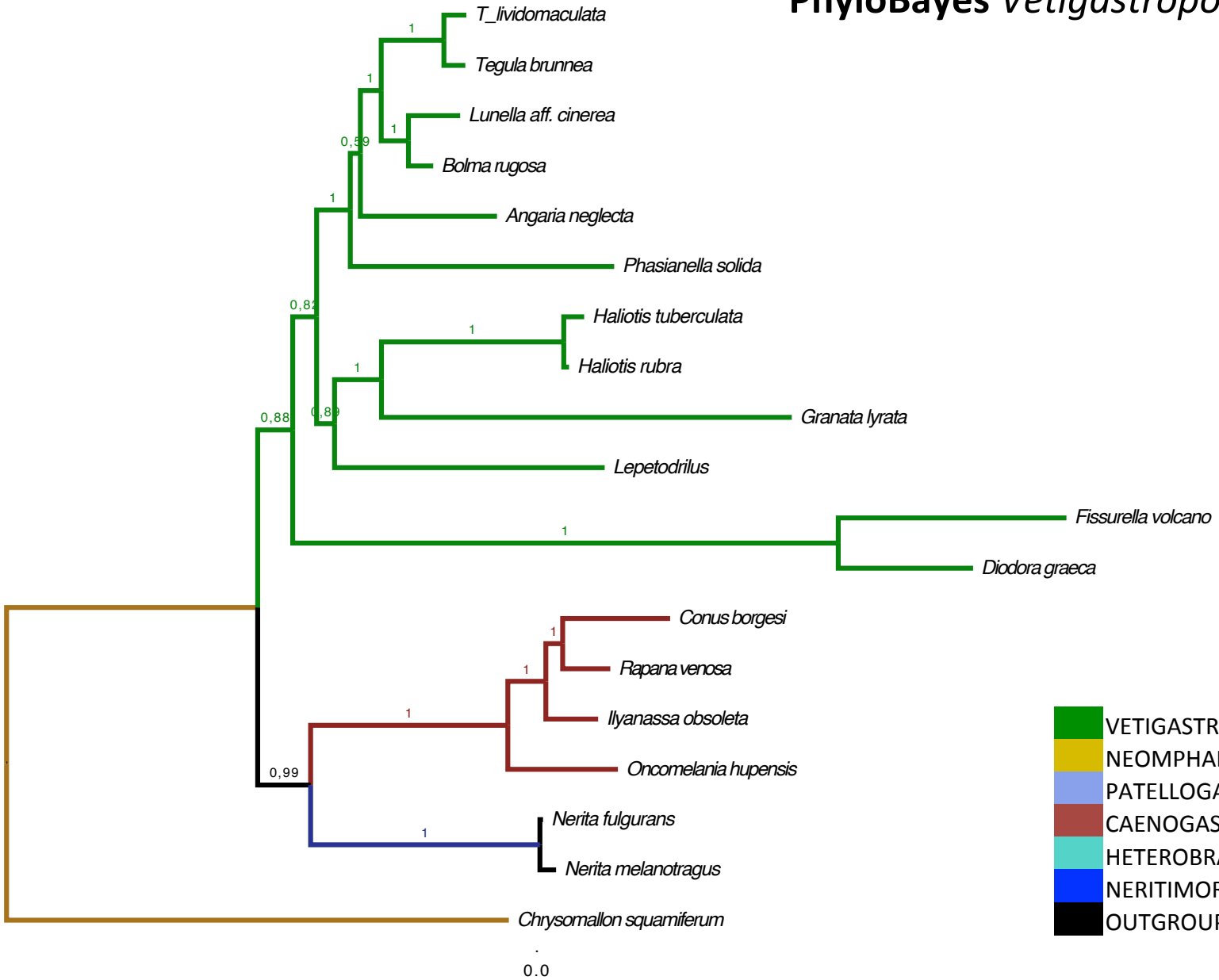
Name Gene	Type	Start	Stop	Length	Codon start	Codon stop	Strand	%A-T
	Intergenic			378				70
cox3	CDS	379	1158	780	ATG	TAA	forward	
	Intergenic			21				77,8
trnK	tRNA	1180	1240	61			forward	
trnA	tRNA	1241	1305	65			forward	
trnR	tRNA	1305	1373	69			forward	
	Intergenic			1				
trnN	tRNA	1375	1440	66			forward	
trnI	tRNA	1441	1507	67			forward	
nad3	CDS	1508	1858	351	ATG	TAG	forward	
trnS(gcu)	tRNA	1859	1924	66			forward	
	Intergenic			2				
nad2	CDS	1927	3018	1092	ATG	TAA	forward	
	Intergenic			4				
cox1	CDS	3023	4567	1545	ATG	TAA	forward	
	Intergenic			22				86,4
cox2	CDS	4590	5297	708	ATG	TAA	forward	
	Intergenic			22				72,7
trnL(uag)	tRNA	5320	5386	67			forward	
	Intergenic			1				
trnL(uaa)	tRNA	5388	5452	65			forward	
	Intergenic			1				
nad1	CDS	5454	6392	939	GTG	TAA	forward	
	Intergenic			12				
trnP	tRNA	6405	6473	69			forward	
	Intergenic			6				
nad6	CDS	6477	6989	513	ATG	TAA	forward	
	Intergenic			5				
cob	CDS	6995	8131	1,137	ATG	TAA	forward	
	Intergenic			5				
trnS(uga)	tRNA	8137	8203	67			forward	
trnT	tRNA	8204	8268	65			reverse	
	Intergenic			11				
nad4L	CDS	8280	8579	300	ATG	TAG	forward	
nad4	CDS	8573	9958	1,386	ATG	TAA	forward	
	Intergenic			12				
trnH	tRNA	9971	10035	65			forward	
nad5	CDS	10036	11766	1,731	ATG	TAG	forward	
	Intergenic			2				
trnF	tRNA	11769	11835	67			forward	
	Intergenic			19				
atp6	CDS	11855	12559	705	ATG	TAA	reverse	
	Intergenic			52				76,9
atp8	CDS	12612	12836	225	GTG	TAG	reverse	
	Intergenic			68				67,6
trnD	tRNA	12905	12971	67			reverse	
rrnL	rRNA	12972	14456	1,485			reverse	
trnV	tRNA	14457	14524	68			reverse	
rrnS	rRNA	14525	15522	998			reverse	
	Intergenic			57				63,8

Supplementary Material 5:
PhyloBayes *Gastropod data set*



maxdiff : 0.124244

Supplementary Material 5:
PhyloBayes *Vetigastropoda data set*



maxdiff : 0.0293367