Mitogenómica y filogenia de linajes de gasterópodos altamente diversificados (Vetigastropoda, Neritimorpha y Conoidea)

Juan Esteban Uribe Arboleda

2016

MITOGENÓMICA Y FILOGENIA DE LINAJES DE GASTERÓPODOS ALTAMENTE DIVERSIFICADOS (VETIGASTROPODA, NERITIMORPHA Y CONOIDEA)

Juan Esteban Uribe Arboleda

Tesis doctoral

2016

Mitogenómica y filogenia de linajes de gasterópodos altamente diversificados (Vetigastropoda, Neritimorpha y Conoidea)

PROGRAMA DE DOCTORADO EN BIOLOGÍA Y CIENCIAS DE LA ALIMENTACIÓN

MEMORIA PRESENTADA POR

JUAN ESTEBAN URIBE ARBOLEDA

PARA OPTAR AL GRADO DE DOCTOR POR LA UNIVERSIDAD AUTÓNOMA DE MADRID

DIRECTORES:

RAFAEL ZARDOYA SAN SEBASTIÁN

JOSÉ TEMPLADO GONZÁLEZ

TUTOR:

ÁNGEL ANTONIO LUQUE DEL VILLAR

Para mi madre y mis abuelos

AGRADECIMIENTOS

Muchas personas e instituciones han contribuido a esta tesis doctoral, en un sentido u otro, y quiero en estas líneas agradecerles:

Inicialmente agradezco a mis directores de tesis, que además de grandes personas son excelentes profesionales con un gran don para enseñar. Gracias de verdad, esto ha sido un experiencia enriquecedora, que seguro, recordaré toda mi vida.

A mi tutor de la Universidad Autónoma de Madrid Ángel Luque, que siempre estuvo dispuesto a ayudarme y a corregirme.

Al Museo Nacional de Ciencias Naturales de Madrid, por facilitar todo para que nosotros los becarios hagamos lo que hacemos. Incluyendo a todo el personal humano: investigadores, directivos, técnicos, servicios generales, a todos, gracias.

Al Ministerio de España por la financiación de esta tesis doctoral a partir de los proyectos (CGL2010-18216; CGL2013-45211-C2-2-P) y la beca FPI (BES-2011-051469). Como también, al Muséum National d'Histoire Naturelle de París, institución que me ha recibido en tres ocasiones para hacer mis estancias de investigación que tanto han enriquecido mi formación. Estancias que fueron patrocinadas por becas del Ministerio de España (estancias cortas) y beca de la comunidad europea (Synthesys).

Agradezco a los colaboradores expertos, que con sus sabios comentarios han mejorado cada uno de los escritos. Como también a los investigadores y amigos que tantos consejos me han dado:

Mario García-París, Diego San Mauro, Federico Abascal, Marta Barluenga, Borja Mila, Mónica Suarez, Ignacio Doadrio, Manuel J. Tenorio, Annie Machordom, Elena G. González, Cristina Grande, David Buckley, Iñigo Martínez, entre otros.

A toda mis amigos y compañeros del museo, que de una u otra forma han hecho parte de este proceso, y que sin ellos hubiera sido imposible terminar esta tesis doctoral: Jorge Gutiérrez, Silvia Perea, Iván Acevedo, Miriam Casal, Mario García, Samuel Abalde, David Osca, Iker Irisarri, Paula Rodríguez, Lourdes Alcaraz, Michel Domínguez, Guillermo Friis, Manuel Sánchez, Diushi Keri, Ernesto Recuero, Etienne Kornobis, Paloma Mas, Pilar Ochoa, Adelmo Martins (seguro he olvidado a alguien!). Y a todos con los que he compartido, chicos, gracias.

También, a la persona que me dio la oportunidad de iniciarme en el camino de la investigación, que tanto disfruto, gracias profe Lyda Castro.

A la familia Uribe, que a pesar de la distancia han estado presentes dando palabras de ánimo que me han fortalecido: mi abuela Josefina que es un pilar de vida; mis tíos (Doralba, Ildegar, Cesar, Chela, Carlos), que hicieron de mi infancia el mejor recuerdo; mis primas (Sandrita, Dianita, y Naty), que quiero mucho; y en especial a mi tío Libardo Uribe, que no volveré a ver, pero que me dejó la imagen de su sonrisa y positiva visión de la vida.

A la familia Arboleda: mis abuelos, mis alcahuetas incondicionales y comprensivos; las tías (Cristina, Mary Isabel,

Alejandra, Amanda), cada una con su punto especial; mi tío Humberto, que más que nadie dio ejemplo académico y su peculiar visión de la vida, que tanto me gusta; mis primos con los que crecí (Santy, Valen, Alejo, Susa, July, Laila, Mariana, Naty).

La familia Toro, que son parte fundamental del proceso: Alba, Mauricio, Jorge Moreno, Maricel. Y mis primos (Julián, Andrés, Caro, Caliche, Juan Camilo). En especial a mi hermano Jorge Esteban.

También, a muchas personas que han hecho parte de mi vida desde que llegué a España: Ferchito-Pilar, la familia de la Iglesia, Pablito, Isidro, a la gente de la "parcela", María José y Albert y Leire.

A todas esas personas que desde Colombia (o cualquier parte) han creído en este proceso y que incondicionalmente me han apoyado: Alexandra Granados, Jan Pájaro, Alberto Yepes, Lisset Gelves, Hadder Jiménez, Merlin, Ivania entre otros.

Y como último y más importante: a mi madre Luz Stella (luz de estrella) que siempre ha sido lo que me gustaría ser. A mi abuelo Fabio que es promotor de mi inquietud por los estudios y de terquedad; A Carlos Alberto que ha hecho de padre; y a mis dos hermanos Rosana y David que siempre están en mi mente. También a mi Chabelita, que con su sonrisa me alegra la vida.

Gracias, TOTALES.

ÍNDICE DE CONTENIDOS

ABREVIATURAS

ADN	Ácido desoxirribonucleico
ADNmt	Ácido desoxirribonucleico mitocondrial
ADNr	Ácido desoxirribonucleico ribosómico
AIC	Criterio de información de Akaike
ARN	Ácido ribonucleico
ARNr	Ácido ribonucleico ribosómico
ARNt	Ácido ribonucleico de transferencia
atp6-8	Genes mitocondriales que codifican para las subunidades 6 y 8 de la ATPasa mitocondrial
BI	Inferencia bayesiana
CAT	modelo mixto Bayesiano con diferentes categorias
cob	Gen mitocondrial que codifica para el citocromo *b*
cox1-3	Genes mitocondriales que codifican para las subunidades 1, 2 y 3 de la citocromo oxidasa
dNTP	Desoxirribonucleótidos Trifosfato
H3	Gen nuclear que codifica para la histona H3
Kb	Kilobases
LBA	Atracción de ramas largas
mya	Millones de años atrás
$MgCl_2$	Cloruro magnésico
$MgSO_4$	Sulfato magnésico
ML	Máxima verosimilitud
mM	Milimolar
mt	Mitocondrial
nad1-6-4L	Genes mitocondriales que codifican para las subunidades 1-6 y 4L de la NADH deshidrogenasa
PCR	Reacción en cadena de la polimerasa
pb	Pares de bases
rrnL	Gen mitocondrial que codifica para la subunidad mayor del ácido ribonucleico ribosómico
rrnS	Gen mitocondrial que codifica para la subunidad menor del ácido ribonucleico ribosómico
SNG	Secuenciación de nueva generación
trnA	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Alanina
trnC	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Cisteína

trnD	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Aspártico
trnE	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Glutámico
trnF	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Fenilalanina
trnG	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Glicina
trnH	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Histidina
trnI	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Isoleucina
trnK	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Lisina
trnL	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Leucina
trnM	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Metionina
trnN	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Asparagina
trnP	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Prolina
trnQ	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Glutamina
trnR	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Arginina
trnS	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Serina
trnT	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Treonina
trnV	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Valina
trnY	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Tirosina
trnW	Gen mitocondrial codificante del ácido ribonucleico de transferencia de Triptófano
U2	Gen nuclear corto no codificante espliceosoma subunidad 2
µM	Micromolar
µl	Microlitro

RESUMEN

A pesar de los importantes avances habidos en las últimas décadas en lo referente a la evolución y filogenia de los gasterópodos, las diferentes aproximaciones generales basadas en datos morfológicos, del registro fósil y moleculares han dado lugar a interpretaciones muy dispares sobre las relaciones filogenéticas dentro de esta clase de moluscos, que hoy día permanecen en buena parte sin resolver. Muchos de estos estudios coinciden en identificar algunos de los principales linajes, pero muestran discrepancias y no resuelven sus relaciones filogenéticas. Los cinco grandes clados que hoy día hay coincidencia en aceptar dentro de los gasterópodos son: Patellogastropoda, Vetigastropoda, Neritimorpha, Caenogastropoda y Heterobranchia, mientras que permanece sin resolver la identidad de Cocculinoidea y Neomphalina, así como las relaciones entre todos estos grupos.

En este sentido, y conociendo la demostrada utilidad de los genomas mitocondriales completos en la resolución de relaciones filogenéticas a diferentes niveles taxonómicos (subclases -en menor grado-, órdenes, superfamilias y familias), y que su catálogo es todavía muy insuficiente y desigual en lo referente a los gasterópodos, se incide en la presente tesis en la secuenciación de estos marcadores moleculares en algunos grupos dentro de Gastropoda aún poco representados y que, sin embargo, cuentan con un alto grado de diversificación. En concreto, la presente tesis doctoral se centra en las subclases Vetigastropoda, Neritimorpha y en la superfamilia Conoidea (dentro de la subclase

Caenogastropoda) y se plantea los siguientes objetivos principales: 1) incrementar el número de genomas mitocondriales completos secuenciados de Gastropoda, especialmente dentro de estas subclases para reconstruir con métodos probabilísticos sus relaciones filogenéticas internas; 2) reconstruir la filogenia de Gastropoda; 3) incrementar el número de genomas mitocondriales completos y datos *multilocus* dentro de un grupo altamente diversificado, la superfamilia Trochoidea (Vetigastropoda), para reconstruir sus relaciones filogenéticas internas; y 4) inferir la evolución de los reordenamientos génicos mitocondriales dentro de estos grupos objeto de estudio.

Como resultado, se han determinado las secuencias completas o casi completas de los genomas mitocondriales de 29 especies de gasterópodos utilizando técnicas de secuenciación masiva: 11 correspondientes a Vetigastropoda, 6 a Neritimorpha, y 12 a Conoidea (Caenogastropoda).

Dentro de Vetigastropoda, se han secuenciado por primera vez genomas mitocondriales de representantes de las superfamilias Phasianelloidea, Angarioidea, Lepetodriloidea y Seguenzioidea, y se ha ampliado el número de genomas mitocondriales completos de superfamilias ya representadas, Fissurelloidea y Trochoidea. Además, se ha incorporado a los análisis comparativos y filogenéticos el genoma mitocondrial completo de un representante de Neomphalina (*Chrysomallon squamiferum*), disponible en NCBI pero que no había sido estudiado.

Con los resultados obtenidos se comprueba que la ordenación de los genes mitocondriales se ajusta en la mayoría de los Vetigastropoda a la propuesta como ancestral para los Gastropoda, variando sólo en la posición relativa de algunos genes codificantes de ARNs de transferencia. Solo los genomas mitocondriales de las superfamilias Lepetodriloidea y Fissurelloidea muestran reordenamientos génicos más drásticos. Además, se ha comprobado que el genoma mitocondrial del representante de Neomphalina muestra un ordenamiento propio, no relacionado con ningún otro descrito para Gastropoda, por lo que se propone tentativamente su exclusión de los Vetigastropoda. Dentro del árbol de Vetigastropoda, se han distinguido cuatro linajes a nivel de superfamilia: Fissurelloidea, Lepetodriloidea, Seguenzioidea + Haliotoidea, y Trochoidea + Angarioidea + Phasianelloidea. La filogenia reconstruida permite inferir que la pérdida de la branquia paleal derecha ocurrió en múltiples ocasiones durante la evolución de este linaje.

Dentro de la superfamilia Trochoidea, se han secuenciado por primera vez los genomas mitocondriales de representantes de las familias Trochidae, Calliostomatidae y Margaritidae, y se ha ampliado el número de genomas mitocondriales de las familias ya representadas Tegulidae y Turbinidae. La filogenia reconstruida recuperó tres linajes principales: el primero formado por las familias Trochidae y Calliostomatidae; el segundo por Margaritidae; y un tercero que agrupa Angarioidea y Phasianelloidea, con un clado formado por los géneros *Tectus* y *Cittarium* (que formarían una nueva familia) más las familias Tegulidae y Turbinidae.

Se han generado también datos *multilocus* (secuencias parciales de los genes mitocondriales *cox1*, *cob*, *rrnS* y *rrnL* y de los genes nucleares *28S rRNA* e *histona H3*) de diversos representantes de la familia Trochidae, principalmente de las subfamilias Cantharidinae y Stomatellinae. La filogenia reconstruida a partir de estos datos muestra como todas las especies del Atlántico Noreste y del Mediterráneo se agrupan en un clado que recibe un alto apoyo estadístico. Los géneros *Phorcus* y *Jujubinus* forman sendos grupos monofiléticos con alto apoyo estadístico. Los géneros *Clelandella* y *Callumbonella* también forman un clado con alto apoyo. Sin embargo, el género *Gibbula* no ha resultado ser un grupo monofilético y ha sido dividido en cinco linajes independientes. El origen de los Cantharidinae del Atlántico Noreste y Mediterráneo se estimó (mediante un reloj molecular relajado) que ocurrió hace unos 47 millones de años (durante la denominada fase *Azolla*, Eoceno Medio), mientras que los eventos de diversificación a nivel genérico y específico coinciden con el cierre definitivo del mar de Tetis (hace 14 millones de años) y con la Crisis de Salinidad del Messiniense (hace 5,3 millones de años), respectivamente. Asimismo, se han reconstruido las relaciones filogenéticas de los géneros de la subfamilia Stomatellinae, que mostró que la diversidad en número de especies está actualmente infraestimada.

Dentro de Neritimorpha, se han secuenciado por primera vez genomas mitocondriales de representantes de las tres superfamilias, Neritopsoidea, Helicinoidea y Hydrocenoidea; así como de nuevos géneros de Neritoidea. En tres de ellas, el ordenamiento genómico

deducido coincidió con el considerado ancestral de Gastropoda. Solo la superfamilia Helicinoidea mostró reorganizaciones significativas en el orden de los genes mitocondriales. La filogenia de Neritimorpha basada en genomas mitocondriales completos recuperó con alto apoyo estadístico a la superfamilia Neritopsoidea como grupo hermano de un clado formado por Helicinoidea e Hydrocenoidea más Neritoidea.

Dentro de la familia Conidae, se han secuenciado por primera vez genomas mitocondriales de géneros de esta familia hasta ahora sin representación, en concreto de *Profundiconus*, *Californiconus*, *Conasprella* y *Lilliconus*, y se ha incrementado el número de genomas mitocondriales para el género *Conus*. También se secuenciaron genomas mitocondriales de familias de Conoidea relacionadas con Conidae (Conorbidae, Clathurellidae y Mangeliidae). La filogenia reconstruida a partir de los genomas mitocondriales completos recuperó la familia Conidae como un grupo monofilético. El género *Profundiconus* resultó grupo hermano de los demás géneros. Además, el género *Conus* se recuperó como grupo hermano de un clado que incluye el género *Conasprella* como grupo hermano de *Californiconus* y *Lilliconus* más *Pseudolilliconus*. La divergencia de los principales linajes dentro de la familia Conidae se estimó (mediante un reloj molecular relajado) que ocurrió entre el Paleoceno y el Eoceno (hace 56-30 millones de años) y la diversificación de especies en la transición del Oligoceno al Mioceno (hace 23 millones de años).

SUMMARY

Despite the important progress made in last decades regarding the evolution and phylogeny of gastropods, the different general approaches based on morphological, fossil record and molecular data have led to disparate interpretations of the phylogenetic relationships within this class of mollusks, which today remain largely unresolved. Many of these studies agree in identifying some of the major lineages, but show discrepancies and do not resolve their phylogenetic relationships. The five major clades that are currently accepted within gastropods are: Patellogastropoda, Vetigastropoda, Neritimorpha, Caenogastropoda and Heterobranchia, while the identity of Cocculinoidea and Neomphalina remains unresolved, as well as the relationships among all these groups.

In this regard, and knowing the proven usefulness of complete mitochondrial genomes in resolving phylogenetic relationships at different taxonomic levels (subclasses –to a lesser extent-, orders, superfamilies and families), and that their catalog is still insufficient and uneven for gastropods, the present thesis is focused on the sequencing of these molecular markers in some groups within Gastropoda still poorly represented and yet having a high degree of diversification. Specifically, this thesis focuses on subclasses Vetigastropoda and Neritimorpha, and the superfamily Conoidea (within subclass Caenogastropoda), and tackles the following main objectives: 1) to increase the number of sequenced complete mitochondrial genomes of Gastropoda, particularly within these

subclasses to reconstruct their internal phylogenetic relationships using probabilistic methods; 2) to reconstruct the phylogeny of Gastropoda; 3) to increase the number of complete mitochondrial genomes and *multilocus* data within a highly diversified group, the superfamily Trochoidea (Vetigastropoda), to reconstruct their internal phylogenetic relationships; and 4) to infer the evolution of mitochondrial gene rearrangements within these groups under study.

As a result, the complete or nearly complete mitochondrial genome sequences of 29 species of gastropods were obtained using massive sequencing techniques: 11 corresponding to Vetigastropoda, 6 to Neritimorpha, and 12 to Conoidea (Caenogastropoda).

Within Vetigastropoda, the mitochondrial genomes of representatives of superfamilies Phasianelloidea, Angarioidea, Lepetodriloidea and Seguenzioidea have been sequenced for the first time, and the number of complete mitochondrial genomes of already represented superfamilies, Fissurelloidea and Trochoidea has been expanded. In addition, the complete mitochondrial genome of a representative Neomphalina (*Chrysomallon squamiferum*), available at NCBI but not previously studied, was incorporated into the phylogenetic and comparative analyses.

The results obtained show that the arrangement of mitochondrial genes in most Vetigastropoda fits the proposed as ancestral for Gastropoda, varying only in the relative position of some genes encoding transfer RNAs. Only the mitochondrial

genomes of superfamilies Lepetodriloidea and Fissurelloidea show more drastic gene rearrangements. In addition, it was found that the mitochondrial genome of the representative of Neomphalina has its own arrangement, unrelated to any other described for Gastropoda, so its exclusion from Vetigastropoda is tentatively proposed. Within Vetigastropoda tree, four lineages were distinguished at the superfamily level: Fissurelloidea, Lepetodriloidea, Seguenzioidea + Haliotoidea, and Trochoidea + Angarioidea + Phasianelloidea. The reconstructed phylogeny allows inferring that the loss of the right pallial gill occurred on multiple occasions during the evolution of this lineage.

Within the superfamily Trochoidea, the mitochondrial genomes of representatives of the families Trochidae, Calliostomatidae and Margaritidae have been sequenced for the first time, and the number of mitochondrial genomes of already represented families, Tegulidae and Turbinidae has been expanded. The reconstructed phylogeny recovered three main lineages: the first formed by families Trochidae and Calliostomatidae; the second by Margaritidae, and a third one grouped Angarioidea and Phasianelloidea, with a clade formed by genera *Tectus* and *Cittarium* (which would form a new family) plus families Tegulidae and Turbinidae.

Multilocus data (partial sequences of mitochondrial genes *cox1*, *cob*, *rrnS* and *rrnL* and nuclear *histone H3* and *28S rRNA* genes) of various representatives of Trochidae, mainly from subfamilies Cantharidinae and Stomatellinae have been also

generated. The phylogeny reconstructed based on these data shows that all species from the Northeast Atlantic and Mediterranean are grouped in a clade that receives high statistical support. Genera *Phorcus* and *Jujubinus* form two monophyletic groups with high statistical support. Genera *Clelandella* and *Callumbonella*. Also form a clade with high support. However, the genus *Gibbula* was not monophyletic and it was divided into five independent lineages. The origin of Northeast Atlantic and Mediterranean Cantharidinae was estimated (using a relaxed molecular clock) to have occurred about 47 million years ago (during the so-called *Azolla* phase), while diversification events at the generic and specific levels match the final closure of the Tethys Sea (14 mya) and the Messinian Salinity Crisis (5.3 mya), respectively. In addition, phylogenetic relationships among genera of the subfamily Stomatellinae have been reconstructed, and showed that diversity in terms of number of species is underestimated at present.

Within Neritimorpha, the mitochondrial genomes of representatives of the three superfamilies, Neritopsoidea, Helicinoidea and Hydrocenoidea have been sequenced for the first time; as well as of new genera from Neritoidea. In three of these superfamilies, deduced genome organization coincided with that considered ancestral of Gastropoda. Only the superfamily Helicinoidea showed significant rearrangements in mitochondrial gene order. The phylogeny of Neritimorpha based on complete mitochondrial genomes recovered with high statistical support the superfamily Neritopsoidea as sister group to a clade formed by Helicinoidea and Hydrocenoidea plus Neritoidea.

Within family Conidae, the mitochondrial genomes of genera thus far not represented have been sequenced, specifically *Profundiconus*, *Californiconus*, *Conasprella* and *Lilliconus*, and the number of mitochondrial genomes for the genus *Conus* has been expanded. The mitochondrial genomes of conoidean families related to Conidae (Conorbidae, Clathurellidae and Mangeliidae) were also sequenced. The phylogeny reconstructed based on complete mitochondrial genomes recovered the family Conidae as a monophyletic group. The genus *Profundiconus* was the sister group to other genera. Moreover, the genus *Conus* was recovered as sister group to a clade that includes genus *Conasprella* as sister group to *Californiconus* and *Lilliconus* plus *Pseudolilliconus*. The divergence of major lineages within the family Conidae was estimated (using a relaxed molecular clock) to have occurred between the Paleocene and Eocene (56-30 million years ago) and species diversification in the transition from Oligocene to Miocene (23 mya).

1 INTRODUCCIÓN

1.1. Clase Gastropoda

1.1.1. Generalidades

Con un número estimado de especies vivas que oscila entre 40.000 y 150.000, la clase Gastropoda constituye el grupo más diverso dentro del filo Mollusca (Bieler, 1992; Aktipis *et al.*, 2008; Haszprunar y Wanninger, 2012; Rosenberg, 2014) y uno de los más exitosos dentro del reino animal en cuanto a sus adaptaciones y modos de vida (Aktipis *et al.*, 2008). Asimismo, cuenta con un extenso registro fósil que se remonta hasta el Cámbrico Inferior (hace 540-530 millones de años; Frýda *et al.*, 2008). Desde su origen, los gasterópodos han experimentado una extraordinaria radiación adaptativa, por lo que conocer cómo se ha llegado a las distintas formas y modos de vida actuales a partir de modificaciones de la morfología original ha sido objeto de un enorme interés y debate. Todo ello hace de los gasterópodos un grupo idóneo para estudios paleontológicos, evolutivos y filogenéticos, entre otros muchos aspectos. Los gasterópodos tienen asimismo importancia económica como fuente de alimento, de productos farmacéuticos, como objeto de coleccionismo y decoración, como plagas para la agricultura, como trasmisores de algunas enfermedades, y han sido también fuente de inspiración de artistas, ingenieros y arquitectos.

La clase Gastropoda se caracterizan por tener cefalización anterior con uno o dos pares de tentáculos, un pie reptante ventral y una concha calcárea que protege la masa visceral. La principal

sinapomorfía del grupo es la torsión de la masa visceral (una rotación de 180º en relación al pie) durante el desarrollo embrionario (Heller, 2015). La concha puede ser muy variable y ha sido el carácter morfológico más utilizado en la taxonomía del grupo aunque se ha demostrado que presenta altos niveles de homoplasia. Típicamente presenta un enrollamiento espiral dextrógiro, pero ha adoptado una forma aplastada tipo lapa de manera convergente en diversos grupos (Patellogastropoda, Cocculinoidea, Lepetelloidea, Fissurelloidea, Calyptraeoidea, Trimusculidae, entre otros). Puede presentar forma de tubo curvado (Caecidae) o irregular (Vermetoidea), puede llegar a ser bivalva en la familia Juliidae (Heterobranchia: Sacoglossa) o bien puede reducirse o desaparecer en diversos grupos, principalmente de Heterobranchia.

Los gasterópodos son uno de los pocos grupos de animales adaptados a vivir en todos los ambientes, tanto marinos (desde el intermareal a los fondos más profundos, incluyendo las fuentes hidrotermales profundas), dulciacuícolas y estuarios, como incluso los terrestres y arborícolas (Aktipis *et al.*, 2008; Haszprunar y Wanninger, 2012). Asimismo su plasticidad morfológica y ecológica es enorme, con adaptaciones, por ejemplo, a los más diferentes modos de alimentación, desde ramoneadores poco especializados (la mayor parte de los Patellogastropoda, Vetigastropoda y Littorinimorpha, dentro de los Caenogastropoda), hasta filtradores (como los Vermetoidea y Calyptroidea), detritívoros (como muchos Cerithioidea), macrohervívoros (como los Anaspidea), succionadores de jugos vegetales (los Sacoglossa) o

de fluidos corporales (como los Epitoniidae, Cancellaroidea o los Colubrariidae), carroñeros (como los Nassariidae), ramoneadores sobre animales sésiles (como los Caralliophilinae, Ovulidae o los Nudibranchia), depredadores (como buena parte de los Neogastropoda) o parásitos (como los Pyramidelloidea y Eulimidae). Asimismo, existen algunas especies que utilizan recursos quimiosintéticos derivados de bacterias simbiontes (como algunas especies de las fuentes hidrotermales pertenecientes a diversos grupos, como Neomphalina, Abyssochrysoidea u Orbitestellidae) o fotosintéticos mediante la incorporación de cloroplastos activos a sus tejidos procedentes de la algas que consumen (diversos Sacoglossa). El modo de alimentación ancestral se especula que fue micrófago, consistente es raspar con la rádula la película microbiana que recubre las superficies marinas, y debido a diferentes modificaciones en su aparato bucal y digestivo han desarrollado sus diversas estrategias alimenticias (Heller, 2015).

El origen de los gasterópodos y las relaciones entre sus principales grupos siguen siendo actualmente tema de debate entre las principales disciplinas que lo estudian (paleontología, morfología y biología molecular). Las extinciones masivas, seguidas de rápidas radiaciones, además de las convergencias de caracteres en los diferentes grupos, dificultan trazar la historia evolutiva del grupo. Por consiguiente, la mayoría de los linajes vivos representan sólo una idea más o menos aproximada de sus ancestros y de los diferentes procesos evolutivos ocurridos a lo largo de su historia.

En el ya clásico "*Treatise on Invertebrate Paleontology*" (Knight *et al.*, 1960), se establece una clasificación de los gasterópodos extintos y se postulan sus relaciones filogenéticas, basadas exclusivamente en los caracteres de la teloconcha (concha del adulto) en comparación con los taxones actuales. Sin embargo, se ha considerado posteriormente que los caracteres de la teloconcha no son informativos para establecer relaciones a niveles taxonómicos altos por los notables fenómenos de convergencia que ha sufrido, por lo que los postulados previos han sido muy cuestionados y modificados con posterioridad (Tracey *et al.*, 1993). En la actualidad, cada vez se concede más importancia a los caracteres de la protoconcha, pues son más estables y además reflejan aspectos del tipo de desarrollo larvario y, por tanto, de ciertas condiciones paleoecológicas (Nützel *et al.*, 2006). La protoconcha es la concha larvaria que se forma antes de la metamorfosis (es decir, antes de pasar al estado adulto). La fase larvaria puede desarrollarse en el plancton y alimentarse del mismo (desarrollo planctotrófico) o puede transcurrir dentro de una cápsula ovígera, eclosionando ya como un juvenil reptante o en una fase muy avanzada de desarrollo que no precisa de una fuente de alimentación externa (desarrollo lecitotrófico). En el primer caso se desarrolla una protoconcha denominada multiespiral, en la que pueden diferenciarse una protoconcha I (la desarrollada dentro de la cápsula ovígera) y una protoconcha II (la que se desarrolla durante la fase planctónica), existiendo una discontinuidad entre una y otra. En el caso del desarrollo lecitotrófico la protoconcha resultante se denomina pauciespiral y carece de protoconcha II.

Basándose en el estudio de las protoconchas, Nützel y Frýda (2003) y Nützel *et al.* (2006) concluyeron que no se conocen moluscos con protoconchas planctotróficas en el Cámbrico, pero ya eran abundantes en el Ordovícico. Por tanto, dichos autores asumen que el desarrollo planctotrófico se adquirió en los gasterópodos en la transición del Cámbrico al Ordovícico, coincidiendo con un aumento de nutrientes, y por tanto del plancton, en la columna de agua. Sin embargo, la ausencia de fósiles en la transición al Ordovícico (ver discusión de esta tesis) oscurece las relaciones de los primitivos linajes con los actuales (Parkhoev, 2008). Los representantes actuales de Patellogastropoda, Cocculinoidea, Vetigastropoda y Neomphalina no poseen larvas planctotróficas, característica que sólo presentan los representantes actuales de Neritimorpha, Caenogastropoda y Heterobranchia. Por tanto, la adquisición de la planctotrofia se considera una sinapomorfía de estos tres últimos clados.

1.1.2. Clasificación

La clasificación de los gasterópodos propuesta por Thiele (1929–35) es la que sienta las bases de la clasificación actual. Fue mayoritariamente adoptada por casi todos los malacólogos y libros de texto (en algunos todavía se mantiene hoy día), incluyendo los grandes tratados y monografías (Wenz, 1938-1944; Cox, 1960; Fretter y Graham, 1962; Boss, 1982, entre otros). Esta clasificación estaba basada en caracteres del sistema respiratorio, nervioso y radular y reconocía tres subclases (Figura 1): Opisthobranchia,

Pulmonata y Prosobranchia. Los Opisthobranchia fueron divididos a su vez en cuatro órdenes (Pleurocoela, Pteropoda, Sacoglossa y Acoela) y los Pulmonata en dos (Basommatophora y Stylommatophora). Los Prosobranchia, grupo en el que se centra la presente memoria, agrupaba los órdenes Archaeogastropoda, Mesogastropoda y Stenoglossa (este último equiparable a los Neogastropoda de Wenz 1938-1944). Estos órdenes fueron subdivididos a su vez en "estirpes", que incluían las diferentes familias. El orden Archaeogastropoda, se dividió en Zeugobranchia, Trochacea, Docoglossa, Neritacea y Cocculinacea (Figura 1), mientras que los Mesogastropoda fueron subdivididos en 15 linajes con 60 familias, y los Stenoglossa o Neogastropoda en cuatro linajes con 16 familias. Se trataba de un esquema de clasificación lineal y gradualista en el que cada uno de estos grandes grupos se correspondería con progresivos grados evolutivos, motivo que llevó a Cox (1960) a unir Mesogastropoda y Neogastropoda en un solo taxón (Caenogastropoda), al no encontrar una discontinuidad clara en la transición de un grupo a otro.

La clasificación de Thiele (1929–35) no fue modificada sustancialmente hasta la segunda mitad la década de 1980, con contadas excepciones, como la revolucionaria clasificación propuesta por Golikov y Starobogatov (1975), que no tuvo aceptación fuera de la "escuela rusa". A partir de 1980, los notables avances en los estudios anatómicos propiciados por las nuevas técnicas de microscopía electrónica, el descubrimiento de nuevos grupos asociados a las fuentes hidrotermales profundas y el establecimiento de la metodología cladista (que permitía el

establecimiento de filogenias más objetivas y comprobables, frente a los esquemas evolutivos y subjetivos que se habían propuesto con anterioridad) propiciaron una revolución en la interpretación y la clasificación de los gasterópodos. Cabe mencionar la gran influencia que autores como los británicos Vera Fretter y Alastair Graham, o los austriacos Luitfried von Salvini-Plawen y su discípulo Gerhard Haszprunar, ejercieron por entonces para auspiciar una nueva etapa de la malacología, que ha tenido su continuación hasta nuestros días con el auge de la biología molecular.

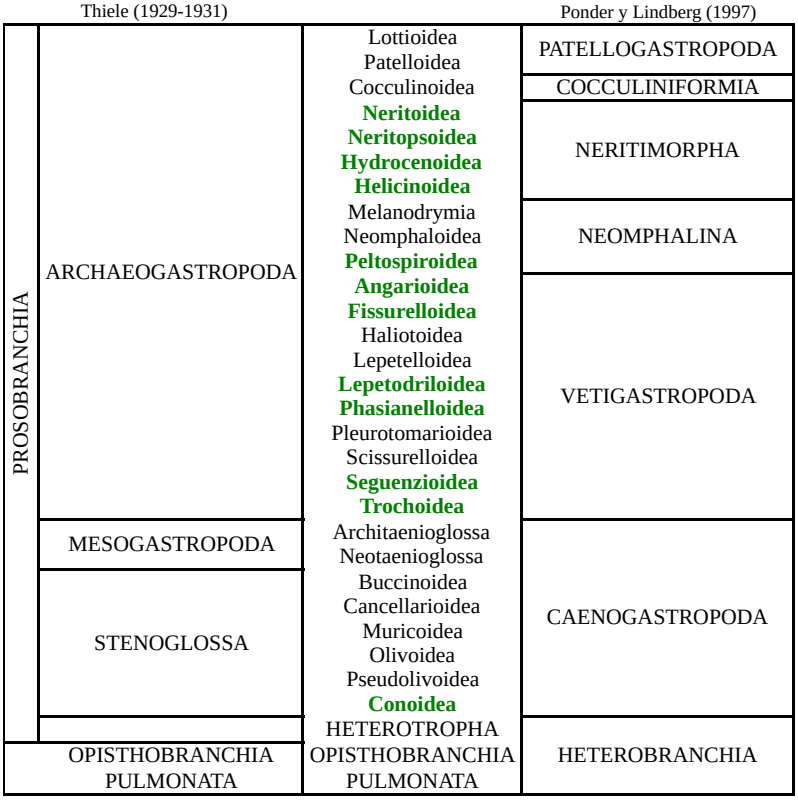


Figura 1. Clasificaciones más relevantes de Gastropoda.. En verde los grupos objeto de estudio en esta tesis.

Salvini ‐ Plawen (1980) propuso la unión de Zeugobranchia y Trochacea en el taxón Vetigastropoda y mantener Caenogastropoda. Años después, con los avances en el conocimiento de la anatomía de diferentes grupos se publica un estudio monográfico sobre la evolución y filogenia de los gasterópodos con una nueva propuesta de clasificación (Haszprunar, 1988). Dicho trabajo ha constituido sin duda uno de los hitos de mayor influencia y que dio mayor impulso a lo que se puede considerar como una nueva etapa en el estudio de la evolución y filogenia de los gasterópodos. Haszprunar (1988) propuso la división de los Gastropoda en dos subclases, Streptoneura y Euthyneura. La segunda comprendía los superórdenes Opisthobranchia y Pulmonata, mientras que dividió los Streptoneura en los órdenes Archaeogastropoda y Apogastropoda. Por un lado, subdividió los Archaeogastropoda en siete subórdenes: Docoglossa (verdaderas lapas), como el grupo más "ancestral", Cocculiniformia, "clado C de las fuentes hidrotermales" (denominado Lepetopsina por Fretter (1990), Neritimorpha, "clado A y *Neomphalus*" de las fuentes hidrotermales" (los actuales Neomphalina), Vetigastropoda con seis superfamilias (Lepetodriloidea, Fissurelloidea, Scissurelloidea, Haliotoidea, Pleurotomarioidea y Trochoidea), y Seguenziina (reconocidos como línea independiente a Vetigastropoda). Asimismo consideró a los Architaenioglossa sin un estatus taxonómico definido, como una línea divergente de los Caenogastropoda. Por otro lado, se utiliza el taxon Apogastropoda (propuesto por Salvini ‐ Plawen y Haszprunar 1987), para referirse

a los Caenogastropoda y Allogastropoda, estos últimos propuestos por Haszprunar (1985) para incluir a una serie de grupos con características intermedias entre Streptoneura y Euthyneura y denominados también “Heterostropha” por Ponder y Warén (1988) y posteriormente como “heterobranquios basales” (Bouchet *et al.*, 2005). La nueva clasificación propuesta por Haszprunar (1988) fue criticada por Bieler (1990) por la metodología utilizada, a medio camino entre el cladismo y el evolucionismo, y por la inclusión de grupos parafiléticos en la clasificación propuesta (ver también la respuesta de Haszprunar 1990).

Seguidamente, Bieler (1992) publicó una revisión histórica de la clasificación de los gasterópodos, haciendo hincapié en la necesidad de estabilizarla sustentando los cambios propuestos mediante estudios filogenéticos cladistas y que estuviera basada en grupos monofiléticos. Dicho autor consideró dentro de los Streptoneura a los Docoglossa (verdaderas lapas), Archaeogastropoda (incluyendo Vetigastropoda y los “taxa de fuentes hidrotermales”), Neritimorpha, Cocculiniformia y Caenogastropoda. Por otro lado, agrupó a los Allogastropoda (= Heterostropha) y Euthyneura en el gran taxa Heterobranchia, que había sido propuesto previamente por Haszprunar (1985). Asimismo, Bieler (1992) señaló a los Docoglossa como posible grupo hermano del resto de Gastropoda por sus simplesiomorfías compartidas con otros moluscos y por las autopomorfías propias que los diferencian del resto de gasterópodos. En cualquier caso insistió en la incertidumbre que todavía se mantenía sobre las

relaciones entre los diversos grupos y en la necesidad de nuevos estudios.

Años después, Ponder y Lindberg (1997) publicaron un nuevo y exhaustivo estudio filogenético de los gasterópodos usando estrictamente la metodología cladista y basándose en caracteres morfológicos, incluyendo los ultraestructurales, y del desarrollo embrionario. La nueva clasificación propuesta por Ponder y Lindberg (1997) rechaza los taxones parafiléticos de Haszprunar (1988) y establece una primera división de los gasterópodos en dos grandes grupos, los Eogastropoda (que comprendían a los Patellogastropoda y sus posibles ancestros de concha espiral) y los Orthogastropoda como grupo hermano que incluía al resto de gasterópodos (Figura 1). Dentro de estos últimos la primera línea divergente estaba formada por Neritimorpha + Cocculinoidea (parte de los Cocculiniformia de Haszprunar 1988), seguida del clado constituido por los Archaeogastropoda s. s., formado por los actuales Neomphalina (“taxa de fuentes hidrotermales”) y por los Vetigastropoda. Por último, en su filogenia diverge un gran clado formado por Caenogastropoda + Heterobranchia, como grupos hermanos, al que denominaron Apogastropoda. Dentro de los Caenogastropoda, la primera línea divergente la constituyeron los Architaenioglossa como grupo hermano del resto (Sorbeoconcha). Dentro de estos últimos, los Cerithioidea y Campaniloidea resultaron las primeras líneas divergentes y el resto de grupos fueron agrupados en el clado Hypsogastropoda.

Este esquema filogenético fue la base de la última clasificación general de los gasterópodos, que fue propuesta por Bouchet *et al.* (2005), aunque ésta incorpora algunos de los avances derivados de las filogenias moleculares que ya imperaban por aquellas fechas y que se comentan en el apartado siguiente de esta memoria. Esta clasificación divide a los gasterópodos en seis grandes clados: Patellogastropoda, Cocculiniformia s. s. (reducidos a los Cocculinoidea), Vetigastropoda (incluyendo a los Neomphalidae), Neritimorpha, Caenogastropoda y Heterobranchia. Hoy día, la base de datos WoRMS asigna la categoría de clase a estos seis clados y añade a los Neomphalina como clase independiente de los Vetigastropoda (Figura 1).

1.1.3. Filogenias moleculares

Las técnicas de secuenciación de ADN constituyeron un importante avance en la reconstrucción de la historia evolutiva de la vida y supusieron el comienzo de una nueva era. De hecho, hoy en día, en lo referente a los gasterópodos, así como a otros muchos grupos de organismos, las filogenias moleculares se han impuesto y han desplazado a las filogenias tradicionales basadas en caracteres morfológicos. En la década de 1990, las primeras filogenias dentro del filo Mollusca, y más específicamente dentro de Gastropoda, comenzaron a cuestionar algunas de las nociones establecidas sobre la historia evolutiva del grupo (Tillier *et al.*, 1992; Kenchington *et al.*, 1994; Rosenberg *et al.*, 1994). Estas primeras propuestas usaron un fragmento del gen nuclear *28S ARNr*. Sin embargo, la utilización

de un fragmento corto de este gen tenía una utilidad limitada para abordar una filogenia completa de los gasterópodos, sobre todo para dilucidar eventos de diversificación que tuvieron lugar en épocas tempranas (Cámbrico, Paleozoico o incluso en el Mesozoico) (McArthur y Koop, 1999). Además, estos primeros estudios estaban basados en una escasa representación de taxones. No fue hasta finales del siglo XX y principios del actual cuando se propusieron las primeras hipótesis filogenéticas que abarcaban una buena parte de los grupos reconocidos dentro de los gasterópodos, aunque presentaban un débil soporte estadístico (Winnepenninckx *et al.*, 1998; McArthur y Koop, 1999; McArthur y Harasewych, 2003). Estas filogenias utilizaban secuencias del gen *18S ARNr* con resultados dispares. Por su lado, Colgan *et al.* (2000) abordaron la filogenia de los gasterópodos utilizando dos segmentos del gen *28S ARNr* y del gen nuclear *histona H3*, e incluyeron una amplia representación del grupo. Posteriormente Colgan *et al.* (2003) añadieron al análisis secuencias de tres genes adicionales, un nuevo segmento del gen *28S ARNr*, el gen nuclear corto *U2 ARN*, y un segmento del gen mitocondrial *cox1*. Pero los resultados obtenidos no recuperaron algunos de los principales taxones resultantes de las filogenias morfológicas. Años más tarde Aktipis *et al.* (2008) analizaron de nuevo la filogenia de los gasterópodos con secuencias del gen *18S ARNr* completo, un fragmento del gen *28S ARNr*, el gen nuclear *histona H3*, y fragmentos de los genes mitocondriales *cox1* y *rrnL*. Estos autores realizaron además el primer análisis combinando caracteres morfológicos y moleculares, pues las filogenias obtenidas con una u otra fuente de datos

independientemente resultaban incongruentes. En definitiva, las topologías de los árboles en las diferentes hipótesis filogenéticas propuestas variaban en función de los marcadores utilizados, lo robusto de los análisis, los programas estadísticos utilizados y los taxones incluidos en los análisis.

Ya en la era de la filogenómica, tres estudios abordan las relaciones de Gastropoda, dos de ellos en el marco de la filogenia de Mollusca (Kocot *et al.*, 2011; Smith *et al.*, 2011) y el tercero centrado exclusivamente en resolver las relaciones dentro de Gastropoda (Zapata *et al.*, 2014). Cabe resaltar que, aun utilizando datos de secuenciación masiva, las relaciones dentro del grupo no resultaron consistentes y faltaba representación de algunos de los grupos principales. Lo más reseñable es que estos estudios coincidían en recuperar Apogastropoda, y en el caso de Kocot *et al.* (2011) y Zapata *et al.* (2014) a Neritimorpha como su grupo hermano.

En cualquier caso, a pesar de los importantes avances habidos en lo referente a la evolución y filogenia de los gasterópodos, las diferentes aproximaciones generales basadas en datos morfológicos, incluyendo los ultraestructurales (Golikov y Starobogatov, 1975; Bieler, 1992; Ponder y Lindberg, 1997; Aktipis *et al.*, 2008, entre otros), del registro fósil (p.e. Bandel, 1997; Frýda, 1999; Wagner, 2001), o moleculares (p.e. Colgan *et al.*, 2000; Colgan *et al.*, 2003; McArthur y Harasewych, 2003; Zapata *et al.*, 2014), han dado lugar a interpretaciones muy dispares sobre las relaciones filogenéticas de los gasterópodos, que hoy día permanecen sin resolver en buena

parte. Muchos de estos estudios coinciden en identificar algunos de los principales linajes, pero muestran discrepancias y no resuelven sus relaciones filogenéticas. Los cinco grandes clados que hoy día hay coincidencia en aceptar dentro de los gasterópodos son: Patellogastropoda, Vetigastropoda, Neritimorpha, Caenogastropoda y Heterobranchia, mientras que queda pendiente por resolver la identidad de los Cocculinoidea y Neomphalina, así como las relaciones entre todos estos grupos.

1.2. Mitogenomas de Gastropoda

1.2.1. Generalidades

El ADN o genoma mitocondrial (mitogenoma) tiene generalmente una herencia materna (una notable excepción son algunos bivalvos), baja recombinación y altas tasas evolutivas, por lo cual, es muy utilizado como marcador filogenético para resolver relaciones evolutivas a diferentes niveles taxonómicos (Irisarri *et al.*, 2012; San Mauro *et al.*, 2014; Timmermans *et al.*, 2014; Miya y Nishida, 2015; Shen *et al.*, 2015, entre otros), así como también en genética de poblaciones y filogeografía (Gissi *et al.*, 2008; Yu *et al.*, 2008). En los Metazoa, el genoma mitocondrial generalmente es una molécula circular de aproximadamente 16,000 pb (16kb), las cuales codifican para 37 genes (Boore, 1999): 13 codifican para proteínas, 2 para ARNs ribosómicos (ARNr); y 22 para ARNs de transferencia (ARNt; Figura 2).

Las reorganizaciones de los genes dentro del genoma mitocondrial (por translocaciones, inversiones, deleciones o inserciones), normalmente son ocasionados durante el proceso de replicación (BOORE 1999). De hecho, estas reorganizaciones ocurren principalmente alrededor de las regiones de control (Boore y Brown, 1998) y suelen implicar a los genes codificantes de ARNts (Gissi *et al.*, 2008). La probabilidad de que converjan estos reordenamientos en diferentes linajes es poco probable, y por ello son mapeados y comparados en filogenias, donde son interpretados como sinapomorfías moleculares que parten de un ancestro común, y pueden proporcionar cierta señal filogenética (Grande *et al.*, 2008). Normalmente, los reordenamientos están correlacionados con altas tasas evolutivas (Xu *et al.*, 2006), que dificultan en cierta medida las reconstrucciones filogenéticas con artefactos como el denominado de "Atracción de Ramas Largas" (LBA; Bergsten, 2005).

Desde el primer genoma mitocondrial secuenciado de la clase Gastropoda, el pulmonado *Albinaria coerulea* (Hatzoglou *et al.*, 1995), se han secuenciado previamente a esta tesis doctoral un total de 135 mitogenomas completos y parciales (>13000 pb, Tabla 1). A pesar de la enorme diversidad de la clase Gastropoda y del avance en las técnicas de secuenciación, como la Secuenciación de Nueva Generación (SNG), el número de mitogenomas secuenciados para este grupo es todavía muy exiguo, si lo comparamos sobre todo con los conocidos para diversos grupos de artrópodos y vertebrados. Aunque el número de mitogenomas secuenciados aumenta progresivamente (ya que sólo se conocían de 26 especies hasta

2010), quedan por representar numerosos linajes y existe una gran desproporción de unos grupos respecto a otros en cuanto al número de mitogenomas completos conocidos. Por ejemplo, la mayor parte

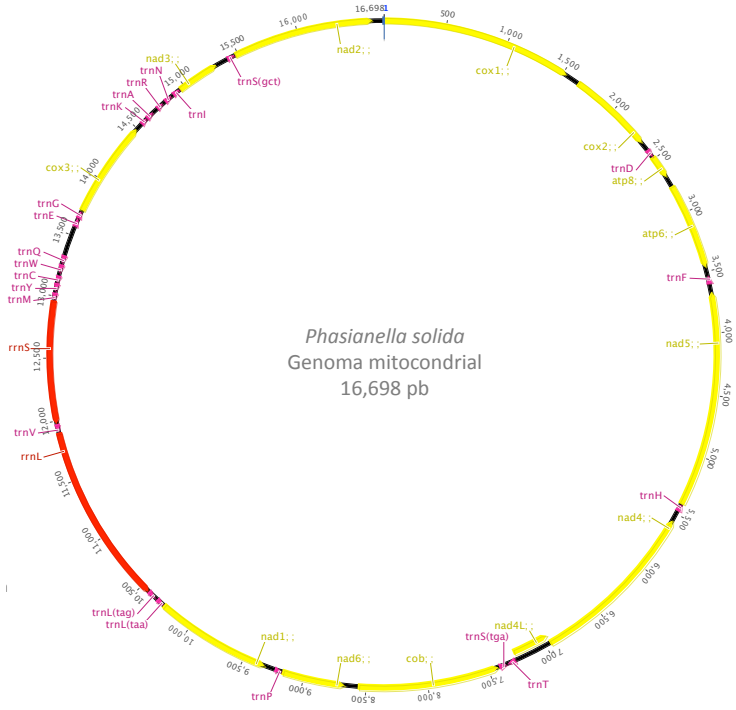


Figura 2. Representación circular de un genoma mitocondrial. En amarillo los genes codificantes de proteína; en rosa, los genes codificantes de ARNs de transferencia y en rojo los genes codificantes de ARNs ribosomales.

de las especies cuyo mitogenoma se conoce corresponden a las subclases Caenogastropoda (56 especies, de las cuales casi la mitad son Neogastropoda) y Heterobranchia (63 especies). Esta representación es todavía insuficiente, teniendo en cuenta que son los grupos de gasterópodos más diversificados y que abarcan

alrededor del 85% de las especies conocidas. Por otro lado, los mitogenomas conocidos correspondientes a las clases que divergieron primero dentro de los gasterópodos son llamativamente escasos: 1 especie de Patellogastropoda, 1 de Neomphalina, 4 de Neritimorpha (todas de especies del género *Nerita*) y 9 de Vetigastropoda (4 de ellas del género *Haliotis* y con representación de solo tres de las diez superfamilias reconocidas dentro del grupo). Por último, no se conocía al comienzo de este trabajo el mitogenoma de ninguna especie de Neomphalina ni Cocculiniformia.

Tabla 1. Genomas mitocondriales disponibles en la clase Gastropoda para cada subclase con su posición taxonómica (superfamilia: familia, respectivamente) y número de acceso a GenBank (No. Acc. NCBI). Genomas parciales (*). Genomas sin verificación (†).

Clado Especie	Posición Taxonómica	No. Acc. NCBI
Heterobranchia		
Aplysia vaccaria	Aplysioidea: Aplysiidae	DQ991928
Aplysia californica	Aplysioidea: Aplysiidae	NC_005827
Aplysia dactylomela	Aplysioidea: Aplysiidae	NC_015088
Aplysia kurodai	Aplysioidea: Aplysiidae	NC_024260
Bulla sp.*	Bulloidea: Bullidae	DQ991930
*Smaragdinella calyculata**	Bulloidea: Smaragdinellidae	DQ991938
*Odontoglaja guamensis**	Philinoidea: Aglajidae	DQ991935
*Sagaminopteron nigropunctatus**	Philinoidea: Gastropteridae	DQ991937
Melibe leonina	Dendronotoidea: Dendronotidae	NC_026987
Tritonia diomedea	Dendronotoidea: Tritoniidae	NC_026988
Notodoris gardineri	Anadoridoidea: Polyceridae	NC_015111
Roboastra europaea	Anadoridoidea: Polyceridae	NC_004321
Chromodoris magnifica	Eudoridoidea: Chromodorididae	NC_015096
Chromodoris quadricolor	Eudoridoidea: Chromodorididae	NC_030004
Phyllidia ocellata	Eudoridoidea: Phyllidiidae	NC_030039
Berthellina sp.	Pleurobranchoidea: Pleurobranchidae	NC_015091
Salinator rhamphidia	Amphiboloidea: Amphibolidae	NC_016185
Auriculinella bidentata	Ellobioidea: Ellobiidae	NC_016168
Myosotella myosotis	Ellobioidea: Ellobiidae	NC_012434
Ovatella vulcani	Ellobioidea: Ellobiidae	NC_016175
Pedipes pedipes	Ellobioidea: Ellobiidae	NC_016179

Tabla 1(cont.)

Achatinella mustelina	Achatinelloidea: Achatinellidae	NC_030190
Gastrocopta cristata	Pupilloidea: Pupillidae	NC_026043
Pupilla muscorum	Pupilloidea: Pupillidae	NC_026044
Vertigo pusilla	Pupilloidea: Vertiginidae	NC_026045
Achatina fulica	Achatinoidea: Achatinidae	NC_024601
Camaena cicatricosa	Camaenoidea: Camaenidae	NC_025511
Albinaria caerulea	Clausilioidea: Clausiliidae	NC_001761
Aegista diversifamilia	Helicoidea: Bradybaenidae	NC_027584
Aegista aubryana	Helicoidea: Bradybaenidae	NC_029419
Dolicheulota formosensis	Helicoidea: Bradybaenidae	NC_027493
Mastigeulota kiangsinensis	Helicoidea: Bradybaenidae	NC_024935
Cepaea nemoralis	Helicoidea: Helicidae	NC_001816
Cylindrus obtusus	Helicoidea: Helicidae	NC_017872
Helix aspersa	Helicoidea: Helicidae	NC_021747
Naesiotus nux	Orthalicoidea: Orthalicidae	NC_028553
Succinea putris	Succineoidea: Succineidae	NC_016190
Cerion incanum	Urocoptoidea: Cerionidae	NC_025645
Onchidella borealis	Onchidioidea: Onchidiidae	DQ991936
Onchidella celtica	Onchidioidea: Onchidiidae	NC_012376
Peronia peronii	Onchidioidea: Onchidiidae	NC_016181
Platevindex sp.†	Onchidioidea: Onchidiidae	KJ561352
Platevindex mortoni	Onchidioidea: Onchidiidae	NC_013934
Rhopalocaulis grandidieri	Rathouisioidea: Veronicellidae	NC_016183
Trimusculus reticulatus	Trimusculoidea: Trimusculidae	NC_016193
Galba pervia	Lymnaeoidea: Lymnaeidae	NC_018536
Radix swinhoei	Lymnaeoidea: Lymnaeidae	KP279638
Radix auricularia	Lymnaeoidea: Lymnaeidae	NC_026538
Radix balthica	Lymnaeoidea: Lymnaeidae	NC_026539
Physella acuta	Planorboidea: Physidae	NC_023253
Biomphalaria glabrata	Planorboidea: Planorbidae	NC_005439
Biomphalaria tenagophila	Planorboidea: Planorbidae	NC_010220
Planorbarius corneus	Planorboidea: Planorbidae	NC_026708
Pyramidella dolabrata	Pyramidelloidea: Pyramidellidae	NC_012435
Placida sp.	Limapontioidea: Limapontiidae	NC_020343
Ascobulla fragilis	Oxynooidea: Volvatellidae	NC_012428
Elysia chlorotica	Placobranchoidea: Placobranchidae	NC_010567
Thuridilla gracilis	Placobranchoidea: Placobranchidae	DQ991939
Siphonaria pectinata	Siphonarioidea: Siphonariidae	NC_012383
Siphonaria gigas	Siphonarioidea: Siphonariidae	NC_016188
Pupa strigosa	Acteonoidea: Acteonidae	NC_002176
Hydatina physis	Architectibranchia: Hydatinidae	DQ991932
Micromelo undata	Architectibranchia: Hydatinidae	NC_015106

Tabla 1(cont.)

Caenogastropoda		
Ifremeria nautilei	Abyssochrysoidea: Provannidae	NC_024642
Provanna sp.	Abyssochrysoidea: Provannida	KM675481
Marisa cornuarietis	Ampullarioidea: Ampullariidae	NC_025334
Pomacea canaliculata	Ampullarioidea: Ampullariidae	KJ739609
Pomacea aff. maculata	Ampullarioidea: Ampullariidae	KR350466
Pomacea canaliculata	Ampullarioidea: Ampullariidae	NC_024586
Pomacea maculata	Ampullarioidea: Ampullariidae	NC_027503
Cipangopaludina cathayensis	Ampullarioidea: Viviparidae	NC_025577
Obscurella hidalgoi	Cyclophoroidea: Cochlostomatidae	NC_028004
Strombus gigas	Stromboidea: Strombidae	NC_024932
*Calyptraea chinensis**	Calyptraeoidea: Calyptraeidae	EU827193
*Erosaria spurca**	Cypraeoidea: Cypraeidae	KP716636
Naticarius hebraeus	Naticoidea: Naticidae	NC_028002
Galeodea echinophora	Tonnoidea: Cassidae:	NC_028003
Cymatium parthenopeum	Tonnoidea: Ranellidae	NC_013247
Potamopyrgus antipodarum	Truncatelloidea: Hydrobiidae	NC_020790
Potamopyrgus estuarinus	Truncatelloidea: Hydrobiidae	NC_021595
Oncomelania hupensis	Truncatelloidea: Pomatiopsidae	NC_013073
Oncomelania hupensis hupensis	Truncatelloidea: Pomatiopsidae	NC_012899
Oncomelania hupensis robertsoni	Truncatelloidea: Pomatiopsidae	NC_013187
Tricula hortensis	Truncatelloidea: Pomatiopsidae	NC_013833
Ceraesignum maximum	Vermetoidea: Vermetidae	HM174253
Dendropoma gregarium	Vermetoidea: Vermetidae	NC_014580
Eualetes tulipa	Vermetoidea: Vermetidae	NC_014585
Thylacodes squamigerus	Vermetoidea: Vermetidae	NC_014588
Ceraesignum maximum	Vermetoidea: Vermetidae	NC_014583
Tylomelania sarasinorum	Cerithioidea: Pachychilidae	NC_030263
*Koreanomelania nodifila**	Cerithioidea: Pleuroceridae	KJ696780
Koreoleptoxis globus ovalis	Cerithioidea: Pleuroceridae	LC006055
Leptoxis ampla†	Cerithioidea: Pleuroceridae	KT153076
Semisulcospira libertina	Cerithioidea: Semisulcospiridae	NC_023364
Turritella bacillum	Cerithioidea: Turritellidae	NC_029717
*Columbella adansoni**	Buccinoidea: Columbellidae	KP716637
Babylonia lutosa	Buccinoidea: Buccinidae	NC_028628
Buccinum pemphigus	Buccinoidea: Buccinidae	NC_029373
Volutharpa perryi	Buccinoidea: Buccinidae	NC_028183
Babylonia areolata	Buccinoidea: Buccinidae	NC_023080
Ilyanassa obsoleta	Buccinoidea: Nassariidae	NC_007781
Nassarius reticulatus	Buccinoidea: Nassariidae	NC_013248
Varicinassa variciferus	Buccinoidea: Nassariidae	NC_029173
Concholepas concholepas	Muricoidea: Muricidae	NC_017886
Rapana venosa	Muricoidea: Muricidae	NC_011193

Tabla 1 (cont.)

Bolinus brandaris	Muricoidea: Muricidae	NC_013250
Reishia clavigera	Muricoidea: Muricidae	NC_010090
Amalda northlandica	Volutoidea: Olividae	NC_014403
*Cymbiola pulchra*B9*	Volutoidea: Volutidae	JN182216
Cymbium olla	Volutoidea: Volutidae	NC_013245
Cancellaria cancellata	Cancellarioidea: Cancellariidae	NC_013241
Conus consors	Conoidea: Conidae	NC_023460
Conus tulipa	Conoidea: Conidae	NC_027518
Conus tribblei	Conoidea: Conidae	NC_027957
Conus gloriamaris	Conoidea: Conidae	NC_030213
Africonus borgesi	Conoidea: Conidae	NC_013243
Conus textile	Conoidea: Conidae	NC_008797
Oxymeris dimidiata	Conoidea: Terebridae	NC_013239
Fusiturris similis	Conoidea: Turridae	NC_013242
Lophiotoma cerithiformis	Conoidea: Turridae	NC_008098
Vetigastropoda		
Chlorostoma brunnea	Trochoidea: Tegulidae	NC_016954
Lunella aff. *cinerea*	Trochoidea: Turbinidae	KF700096
Fissurella volcano	Fissurelloidea: Fissurellidae	NC_016953
Haliotis discus hannai	Haliotoidea: Haliotidae	KF724723
*Haliotis diversicolor supertexta**	Haliotoidea: Haliotidae	HQ832671
Haliotis laevigata	Haliotoidea: Haliotidae	NC_024562
Haliotis rubra	Haliotoidea: Haliotidae	NC_005940
Haliotis tuberculata tuberculata	Haliotoidea: Haliotidae	NC_013708
*Lepetodrilus nux**	Lepetodriloidea: Lepetodrilidae	LC107880
Neritimorpha		
*Nerita fulgurans**	Neritoidea: Neritidae	KF728888
*Nerita tessellata**	Neritoidea: Neritidae	KF728889
*Nerita versicolor**	Neritoidea: Neritidae	KF728890
*Nerita melanotragus**	Neritoidea: Neritidae	GU810158
Patellogastropoda		
Lottia digitalis	Lottioidea: Lottiidae	NC_007782
Neomphalina		
Chrysomallon squamiferum	Neomphaloidea: Peltospiridae	AP013032

1.2.2. Filogenias

Dentro de los moluscos, la utilización de los genomas mitocondriales para la reconstrucción de las relaciones evolutivas a altos niveles jerárquicos (a nivel de clase) no han sido muy útiles debido a la heterogeneidad de las tasa evolutivas y la composición nucleotídica del ADNmt en los diferentes linajes (Stöger y Schrödl,

2013; Osca *et al.*, 2014a; Schrödl y Stöger, 2014). Dentro de Gastropoda, las primeras filogenias con genomas mitocondriales completos fueron enfocadas a resolver linajes concretos, como Euthyneura (Grande *et al.*, 2002, 2004a) u Opisthobranchia (Grande *et al.*, 2004b). Posteriormente, Grande *et al.* (2008) abordaron la primera filogenia a nivel global, basada en los mitogenomas de las 16 especies de gasterópodos conocidos hasta la fecha, aunque por entonces todavía faltaban genomas mitocondriales de alguno de los principales grupos, como Neritimorpha. Dicha filogenia obtuvo a Heterobranchia + Patellogastropoda como grupo hermano de Vetigastropoda + Caenogastropoda (Grande *et al.*, 2008), aunque señalaron que el primero de estos grandes clados agrupaba a los linajes que habían sufrido notables recombinaciones, por lo que su unión podía deberse a un efecto de LBA (Figura 3). Más tarde, con la inclusión de los Neritimorpha en los análisis con genomas mt, se obtuvieron varias topologías, dependiendo de los métodos filogenéticos utilizados. En algunos análisis se recuperaba Patellogastropoda + Vetigastropoda como grupo hermano de Neritimorpha + Apogastropoda (Castro y Colgan, 2010) (Figura 3*a*), coincidiendo con las topologías filogenómicas (p.e. Zapata *et al.* 2014), pero en otros también se obtenía a Patellogastropoda + Heterobranchia como grupo hermano de Vetigastropoda y Neritimorpha más Caenogastropoda (Castro y Colgan, 2010; Arquez *et al.*, 2014; Osca *et al.*, 2014b; Figura 3*b*). En otros, se recupera Patellogastropoda + Heterobranchia como grupo hermano de Caenogastropoda y Neritimorpha más Vetigastropoda (Arquez *et al.*, 2014; Osca *et al.*, 2014b; Figura 3*c*). Simultáneamente, dos

análisis que no incluyeron a los Patellogastropoda recuperaban a Heterobranchia como grupo hermano de Vetigastropoda + Neritimorpha y Caenogastropoda (Williams *et al.*, 2014; Sevigny *et al.*, 2015), aunque en el primero de ellos Caenogastropoda resultó polifilético. Sin embargo, cabe destacar que ninguno de los análisis incluía representantes de Neomphalina y Cocculininoidea. También hay que mencionar algunos trabajos en los que se han abordado relaciones de grupos concretos basándose en genomas mitocondriales como los de Grande *et al.* (2004b) y Medina *et al.* (2011) sobre Opisthobranchia, Cunha *et al.* (2009) sobre Neogastropoda, White *et al.* (2011) sobre Pulmonata, y Osca *et al.* (2015) sobre Caenogastropoda.

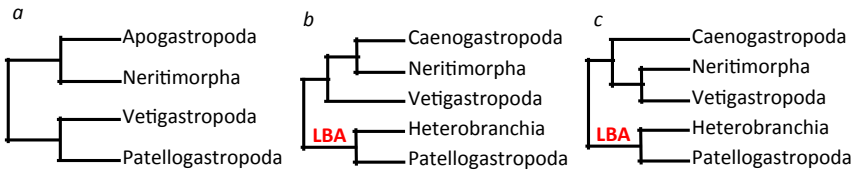


Figura 3. Hipótesis filogenéticas de gasterópodos basadas en genomas mitocondriales.

1.2.3. Reordenamientos

En ausencia de fenómenos de convergencia, los órdenes génicos dentro del genoma mitocondrial pueden reflejar las relaciones filogenéticas a diferentes niveles taxonómicos (Cameron *et al.*,

2007) y los reordenamientos compartidos por grupos diferentes pueden considerarse como sinapomorfías moleculares. Como se dijo anteriormente, los principales agentes móviles implicados en los reordenamientos son los genes codificantes de ARNts mientras que la posición relativa de los genes codificantes de proteínas y de ARNrs es más estable. Las comparaciones de reordenamientos genómicos han mostrado su utilidad para dilucidar las relaciones filogenéticas en diversas ramas del Reino Animal (Boore *et al.*, 2005; Cameron *et al.*, 2007; Lavrov, 2007; Havird y Santos, 2014; Miya y Nishida, 2015, entre otros). Mientras el ADNmt se mantiene muy estable en vertebrados, se han destacado numerosos cambios en el orden génico en el mitogenoma de diversos filos de invertebrados, y se ha demostrado que son particularmente frecuentes en algunos de ellos, como los nematodos o los moluscos (Boore *et al.*, 2005). A pesar de ello, y de que los moluscos son el grupo animal más diverso en número de especies después de los artrópodos, los reordenamientos de sus genomas mitocondriales no han sido apenas estudiados (Grande *et al.*, 2008; Osca *et al.*, 2015). Osca *et al.* (2014a) dedujeron en su estudio el supuesto orden mitocondrial ancestral de los moluscos y de los gasterópodos. Este último se mantiene conservado en diversos clados de gasterópodos, como en los Vetigastropoda, Neritimorpha y en buena parte de los Caenogastropoda (aunque aún es prematuro establecer generalizaciones con el escaso número de mitogenomas conocidos) pero no en Patellogastropoda y Heterobranchia. Precisamente, no hay que dejar de señalar que los reordenamientos drásticos que han sufrido algunos grupos, como los Patellogastropoda, los Vermetidae

o los Heterobranchia están asociados a altas tasas de evolución que se traducen en ramas largas en los árboles y que sufren el conocido efecto LBA, fuente de los principales problemas y artefactos en las reconstrucciones filogenéticas de gasterópodos y moluscos en general basadas en secuencias completas del genoma mitocondrial (Stöger y Schrödl, 2013).

1.3. Subclase Vetigastropoda Salvini-Plawen, 1980

Los Vetigastropoda aparecieron en la transición del Cámbrico al Ordovícico, junto a otros de los grandes linajes de gasterópodos, y se diversificaron durante la radiación temprana del Ordovícico, aproximadamente hace unos 490 millones de años (Frýda *et al.*, 2008). Comprenden varios miles de especies marinas actuales (y diversos grupos extintos) adaptadas a los más diversos ambientes, incluyendo algunos grupos exclusivos de las fuentes hidrotermales profundas.

El taxón Vetigastropoda fue propuesto inicialmente como grupo natural por Salvini - Plawen (1980) para incluir los Zeugobranchia (con órganos paleales pares), Trochoidea (que han perdido los órganos paleales del lado derecho) y Cocculinoidea. Posteriormente, sobre todo con el descubrimiento de distintos taxones asociados a las fuentes hidrotermales profundas (McLean, 1981; Hickman, 1984; McLean, 1988, 1989, 1990), este clado ha sido redefinido en varias ocasiones (incluyendo o excluyendo determinados grupos) y, asimismo, sus relaciones con los otros

grandes grupos de gasterópodos han sido y siguen siendo objeto de debate (Salvini‐Plawen y Haszprunar, 1987; Haszprunar, 1988; Bieler, 1992; Ponder y Lindberg, 1997; Sasaki, 1998; Geiger y Thacker, 2005; Geiger *et al.*, 2008; Kano, 2008; Aktipis y Giribet, 2010, 2012, entre otros).

Los Lepetelloidea, previamente considerados dentro de los Cocculiniformia (Haszprunar, 1987), fueron posteriormente incluidos dentro de Vetigastropoda en los análisis filogenéticos de Ponder y Lindberg (1997), Geiger y Thacker (2005) o Kano (2008). Bouchet *et al.* (2005), Geiger *et al.* (2008) y Kano (2008) incluyeron dentro de Vetigastropoda las siguientes superfamilias actuales (además de algunos grupos fósiles): Pleurotomarioidea, Fissurelloidea, Haliotioidea, Scissurelloidea, Lepetelloidea, Lepetodriloidea, Neomphaloidea, Seguenzioidea, Trochoidea y Turbinoidea. Por su parte, la posición de los Neomphalina (Neomphaloidea) hay sido muy controvertida y algunos autores los consideraron como un grupo independiente fuera de los Vetigastropoda (McArthur y Harasewych, 2003; Aktipis *et al.*, 2008; Heß *et al.*, 2008; Kano, 2008; Aktipis y Giribet, 2010, 2012). Asimismo, algunos análisis moleculares excluyen a los Pleurotomarioidea de los Vetigastropoda (Aktipis *et al.*, 2008; Aktipis y Giribet, 2010, 2012) o los recuperan considerándoles como grupo hermano del resto de Vetigastropoda (Kano, 2008). Por su lado, Williams y Ozawa (2006) elevaron a nivel de superfamilia dos grupos previamente incluidos en los Trochoidea: Angarioidea y Phasianelloidea, y Williams *et al.* (2008) y Williams (2012) redefinieron las superfamilias Trochoidea y Turbinoidea

introduciendo cambios significativos en lo que se refiere a sus respectivos componentes (a nivel de familia).

1.3.1. Superfamilia Trochoidea s. l. Rafinesque, 1815.

La superfamilia Trochoidea es la más diversa dentro de los Vetigastropoda y constituye uno de los grupos de gasterópodos más diversificado, con más de 2.000 especies actuales conocidas y cerca de 500 nombres genéricos (entre válidos y en desuso). Su registro fósil se extiende hasta el Triásico Medio (hace unos 225 millones de años), aunque su origen no está claro y puede ser mucho más antiguo (Hickman y McLean, 1990). Los trocoideos son, además, ecológicamente importantes, pues dominan en algunas comunidades intermareales, praderas de fanerógamas marinas o arrecifes de coral, aunque también están presentes en otros muchos hábitats. Aunque buena parte de sus especies son microherbívoros poco especializados, algunos grupos se han adaptado un régimen carnívoro sobre animales sésiles (p.e. Calliostomatidae) o son suspensívoros (p.e. Umboniinae). Todos los trocoideos presentan una concha espiralada y un opérculo que puede ser córneo o calcáreo, carácter que fue usado por Hickman y McLean (1990) para separar sus dos grandes grupos: tróquidos (opérculo córneo) y turbínidos (opérculo calcáreo).

La clasificación de la superfamilia Trochoidea fue abordada de forma monográfica por Hickman y McLean (1990), que mantuvieron las tres familias que normalmente se reconocían dentro

del grupo (Trochidae, Turbinidae y Skeneidae), e introdujeron hasta 43 taxones supragenéricos (subfamilias y tribus) basándose en caracteres derivados compartidos. Desde entonces esta clasificación ha sufrido numerosos cambios con la introducción de los análisis filogenéticos moleculares. Estudios filogenéticos recientes desbarataban la monofilia de los Trochoidea, tal y como eran considerados hasta la fecha, sobre todo porque muchos de los géneros de pequeños skeneimorfos fueron trasferidos a los Seguenzioidea, a los Neomphalina (Williams y Ozawa, 2006; Kano, 2008; Kunze *et al.*, 2016) o a la nueva familia Crasseolidae de posición taxonómica incierta (Hickman, 2013). Esta autora restringe los Skeneidae sólo a un reducido número de géneros y (HASZPRUNAR *ET AL.*, 2016) proponen una nueva diagnosis de la familia e incluyen en ella provisionalmente sólo a siete géneros.

Por otro lado, son de destacar los trabajos de Williams y Ozawa (2006), Williams *et al*. (2008) y Williams (2012) en cuyos análisis filogenéticos moleculares la superfamilia Trochoidea y la familia Turbinidae, resultan polifiléticas. Basandose en estos estudios, la autora redescribe estos taxones y establece dos nuevas superfamilias: Angarioidea y Phasianelloidea. En el último de estos trabajos, Williams considera a la superfamilia Trochoidea comprendida por la familias Trochidae, Turbinidae, Solariellidae, Calliostomatidae, Liotiidae, Skeneidae, Margaritidae y Tegulidae. En cualquier caso son muchos los géneros que no han sido incluidos en los estudios moleculares y precisan ser reconsiderados. A su vez, algunos otros géneros, como *Cittarium* o *Tectus* quedan en una posición taxonómica incierta.

Hasta el inicio de esta tesis, dentro de la subclase Vetigastropoda sólo se conocían los genomas mitocondriales de tres superfamilias: dos genomas mitocondriales para la superfamilia Trochoidea correspondientes a las familias Tegulidae (género *Tegula*) y Turbinidae (género *Lunella*); varios genomas mitocondriales completos y casi completos para la superfamilia Haliotoidea, todos para el género tipo de la familia Haliotidae; un genoma mitocondrial de la superfamilia Fissurellidae para el género tipo de la familia Fissurellidae. En este sentido, este trabajo irá encaminado a incrementar los genomas mitocondriales completos disponibles a siete de 10 superfamilias dentro de la subclase y se enfatizará en la superfamilia Trochoidea, siendo esta la más diversa en número de especies dentro de la subclase.

1.3.1.1. Familia Trochidae Rafinesque, 1815.

Dentro de los Trochoidea, la familia Trochidae es, junto con Turbinidae, la más diversificada, con más de 600 especies actuales conocidas agrupadas en unos 60 géneros (Williams *et al.*, 2010; Williams, 2012). Aunque presentan una distribución cosmopolita y viven en muy diversos ambientes, se han diversificado sobre todo en el litoral rocoso de áreas tropicales del Indo-Pacífico (Williams *et al.*, 2010).

En su revisión monográfica basada en caracteres morfológicos, Hickman y McLean (1990) dividen los Trochidae en 13 subfamilias (subdivididas a su vez en tribus), algunas de las

cuales fueron elevadas a la categoría de familia (Bouchet *et al.*, 2005). Sin embargo, la clasificación de la familia ha sido objeto de una profunda revisión partiendo de análisis filogenéticos moleculares (Williams y Ozawa, 2006; Kano, 2008; Williams *et al.*, 2008; Williams *et al.*, 2010; Williams, 2012). Como consecuencia, algunos taxones previamente incluidos en esta familia se transfirieron a Turbinidae o a los Seguenzioidea, y algunas de las subfamilias fueron redefinidas. Finalmente, en la actualidad dentro de la familia Trochidae se reconocen 10 subfamilias de acuerdo con la nueva clasificación de los Trochoidea propuesta (Williams *et al.*, 2010): Trochinae, Umboniinae, Stomatellinae, Cantharidinae, Monodontinae, Halistylinae, Kaiparathininae, Fossarininae, Chrysostomatina y Alcyninae. Las redefiniciones de algunas de estas subfamilias han supuesto la transferencia de diversos géneros de unas a otras, pues la asignación genérica de muchas de las especies se ha basado exclusivamente en caracteres de la concha y se ha considerado tentativa. Dos de estas subfamilias, Trochinae y Cantharidinae, tienen representantes en las costas europeas, donde la segunda de ellas ha sufrido una notable diversificación.

La subfamilia Cantharidinae se considera en la actualidad que incluye 23 géneros y más de 200 especies (Gofas 2015 en WoRMS), pero la taxonomía del grupo a nivel genérico está sujeta a una continua revisión y redefinición (Williams *et al.*, 2010). Los análisis filogenéticos más recientes reconocen a la subfamilia Stomatellinae como grupo hermano de Cantharidinae (Williams *et al.*, 2008; Williams *et al.*, 2010; Williams, 2012). Las especies de esta subfamilia se distribuyen por las costas rocosas de áreas

tropicales y templadas del Indo-Pacífico. Su concha presenta una espira muy reducida (recuerdan a las de *Haliotis*, pero sin la hilera de orificios) y un pie muy desarrollado que no puede retraerse dentro de la concha y puede ser autotomizado. Se reconocen en la actualidad 5 géneros pertenecientes a los Stomatellinae con 36 especies (Gofas 2009b en WoRMS), aunque este número puede estar subestimado pues la taxonomía del grupo es muy confusa y precisa de una profunda revisión (Williams *et al.*, 2010).

1.4. Subclase Neritimorpha Golikov & Starobogatov, 1975.

Los Neritimorpha (o Neritopsina) constituyen un arcaico linaje de gasterópodos bien conocidos desde el Carbonífero (Kaim y Sztajner, 2005), pero su origen posiblemente se remonta incluso al Ordovícico, hace 500-450 millones de años (Bandel, 1999) y los miembros más antiguos que supuestamente dieron lugar a los Neritopsidae actuales aparecieron en el registro fósil en el Devónico Inferior (hace unos 400 millones de años).

Este grupo experimentó una gran radiación adaptativa y en la actualidad comprende unas 2.000 especies (Fukumori y Kano, 2014). A pesar de este relativo bajo número de especies actuales, este peculiar grupo de gasterópodos presenta una gran variedad de formas (desde las típicas formas de los caracoles a la morfología pateliforme o incluso la pérdida de la concha en estado adulto). Asimismo, se han adaptado a los hábitats más diversos, marinos (desde el nivel intermareal a los fondos profundos y a las fuentes

hidrotermales), dulceacuícolas y terrestre (incluyendo especies arborícolas) que reproducen a pequeña escala y en cierta manera la diversidad morfológica y ecológica que presentan los gasterópodos (Kano, 2008; Lindberg, 2008).

Los Neritimorpha fueron inicialmente incluidos dentro de los Archaeogastropoda como Neritacea (Thiele, 1929–35), y años más tarde fueron denominados como Neritopsina (Cox, 1960), o Neritimorpha (Golikov y Starobogatov, 1975), autores estos últimos que incluían dentro de ellos a la superfamilia Cocculinoidea. En la actualidad los Neritimorpha son considerados como una clase o un superorden dentro de los gasterópodos, que comprende dos órdenes Cyrtoneritimorpha (extintos) y Cycloneritimorpha (con representantes actuales) (Bouchet *et al.*, 2005). Las especies actuales se reparten en 4 superfamilias y en 6 familias: Neritopsidae y Titiscaniidae (Neritopsoidea), Hydrocenidae (Hydrocenoidea), Helicinidae Proserpinidae, Proserpinellidae y Neritiliidae (Helicinoidea), Phenacolepadidae y Neritidae (Neritoidea) (Gofas 2009a en WoRMS).

De todos los Neritimorpha, los Neritidae son el grupo más conocido y mejor estudiado, pues muchas especies de Nerita son los miembros de la malacofauna más conspicuos en el litoral rocoso intermareal y somero (o en manglares) de muchas áreas tropicales. Algunas especies también se han adaptado a vivir sobre las hojas de las fanerógamas marinas y a alimentarse de ellas (p.e. las del género *Smaragdia*). Algunos géneros de esta misma familia han invadido las aguas salobres y dulces (especies de los géneros *Theodoxus*,

Clithon , *Neritina*, *Septaria* y otros). La superfamilia Neritoidea también incluye a los Phenacolepadidae, cuyas especies presentan una concha mayoritariamente pateliforme y están adaptadas a ambientes muy diversos, incluyendo manglares, maderas sumergidas o feuntes hidrotermales profundas. Por otro lado, los miembros de las superfamilias Hydrocenoidea y Helicinoidea han invadido el medio terrestre y se han adaptado plenamente a vivir en él; dentro de Helicinoidea, las especies de la familia Neritiliidae viven en las aguas dulces subterráneas. Por su parte, las pocas especies actuales conocidas de Neritopsoidea pertenecen a los géneros *Neritopsis*, con especies exclusivas de cuevas y ambientes anquihalinos, y *Titiscania*, con dos especies actuales conocidas que han perdido enteramente la concha, se asemejan a un nudibranquio, se alimentan de ascidias compuestas y secretan sustancias defensivas (Templado y Ortea, 2001).

Mientras la monofilia de los Neritimorpha ha sido ampliamente aceptada, tanto con caracteres morfológicos como moleculares (Ponder y Lindberg, 1997; Aktipis *et al.*, 2008), en lo referente a su posición filogenética, puede decirse que los Neritimorpha son de los grupos más controvertidos, saltando de un lugar a otro en las diferentes propuestas filogenéticas. Tradicionalmente se habían incluido casi siempre dentro de los antiguos "arqueogasterópodos", aunque en ocasiones han sido relacionados con los Caenogastropoda por compartir con éstos algunas aparentes sinopomorfías (Bieler, 1992). Esta controversia se mantiene en la actualidad, siendo considerados unas veces como grupo hermano de los Vetigastropoda (Arquez *et al.*, 2014; Osca *et*

al., 2014a; Osca *et al.*, 2014b; Williams *et al.*, 2014; Osca *et al.*, 2015; Sevigny *et al.*, 2015), en otros de los Apogastropoda (Colgan *et al.*, 2000; McArthur y Harasewych, 2003; Aktipis *et al.*, 2008; Castro y Colgan, 2010; Zapata *et al.*, 2014), en otros de los Vetigastropoda + Apogastropoda (Haszprunar, 1988; Salvini‐Plawen y Steiner, 1996; Osca *et al.*, 2015) o de los Caenogastropoda (Castro y Colgan, 2010; Osca *et al.*, 2014b).

Las relaciones internas de los Neritimorpha se han abordado en pocas ocasiones. Holthuis (1995) analizó la filogenia del grupo sustentada en caracteres morfológicos, mientras que (Kano *et al.*, 2002) abordó el primer análisis molecular con secuencias parciales del gen *28S rADN*. Los resultados de este último análisis sitúan a los Neritopsoidea como grupo hermano de las otras tres superfamilias y a Neritoidea y Helicinoidea como grupos hermanos.

Hasta la realización de la presente tesis sólo se conocía el genoma mitocondrial de especies del género *Nerita* (Castro y Colgan, 2010; Arquez *et al.*, 2014) que fue considerado como representativo de los Neritimorpha y cuya organización resultó ser similar a la propuesta como ancestral para los gasterópodos (Osca *et al.*, 2014a). En la presente tesis se estudia el mitogenoma de representantes de las cuatro superfamilias que comprenden los Neritimorpha.

1.5. Subclase Caenogastropoda Cox, 1960

Durante el Devónico, hace unos 400 millones de años, surgió el grupo que se considera constituyen los “gasterópodos avanzados” (Caenogastropoda), que se convirtieró en el grupo dominante y más diversificado del medio marino desde el Jurásico (hace unos 170 millones de años) hasta la actualidad (Heller, 2015). Los Caenogastropoda comprenden alrededor del 60% de las especies de gasterópodos marinos (Ponder *et al.*, 2008). En estudios exhaustivos de los moluscos a nivel local, aproximadamente el 70% de las especies y casi el 80% de los ejemplares obtenidos eran caenogasterópodos (Bouchet *et al.*, 2002). Su diversidad morfológica es extraordinaria, se han adaptado a una gran variedad de formas de vida y ocupan una enorme cantidad de nichos ecológicos. Este grupo incluye a muchas de los de las familias de gasterópodos más conocidas y apreciadas por los aficionados y coleccionistas (Cypraeidae, Conidae, Volutidae, Olividae, Muricidae, etc).

El taxón Caenogastropoda fue propuesto como un orden de Prosobranchia que agrupaba a los Mesogastropoda y a los Stenoglossa (= Neogastropoda; Cox, 1960). Pero no fue casi utilizado hasta el trabajo de Salvini‐Plawen (1980), y fue Haszprunar (1985) el primer autor que proporcionó una definición del mismo basada en sinapomorfías. Dicho autor dividió los Caenogastropoda en cuatro subórdenes (Architaenioglossa, Neotaenioglossa, Neogastropoda y Heteroglossa), y poco después Haszprunar (1988) relaciona con ellos a los Seguenzioidea. Por su

lado, la mayor parte de las filogenias publicadas reconocen la monofilia de los Caenogastropoda, con contadas excepciones (Colgan *et al.*, 2003; Williams *et al.*, 2014), pero la identificación de su grupo hermano sigue sin estar resuelta.

La clasificación actual de los Caenogastropoda parte principalmente de la propuesta en la filogenia morfológica de Ponder y Lindberg (1997), en la que retoman el taxón Apogastropoda (inicialmente propuesto por Salvini‐Plawen y Haszprunar (1987) para incluir a algunos integrantes de Caenogastropoda y "Heterostropha"), pero en este caso para abarcar Caenogastropoda + Heterobranchia. Ponder y Lindberg (1997) dividen a su vez los Caenogastropoda en Architaenioglossa (Cyclophoridae y Ampullariidae) y Sorbeoconcha, que incluye el resto de Caenogastropoda, con Cerithioidea y Campanilidae como primeras ramas divergentes de la rama principal, constituida por los que denominaron Hypsogastropoda.

Uno de los grupos de Hypsogastropoda que se viene manteniendo desde la clasificación de Thiele (1929–35) son los Stenoglossa (bajo su denominación actual de Neogastropoda) que, según la clasificación de la base de datos WORMS, comprenden seis superfamilias con representantes actuales: Buccinoidea, Cancellarioidea, Conoidea, Muricoidea, Olivoidea y Pseudolivoidea. Los Neogastropoda son habitualmente considerados como el gran clado terminal (el más "avanzado") de los Caenogastropoda, cuyos miembros son casi exclusivamente marinos y carnívoros. A este grupo pertenecen las familias más

conocidas, como Muricidae, Volutidae, Buccinidae, Conidae, Mitridae, Cancellaridae o Terebridae.

Generalmente se ha considerado a los Neogastropoda como un grupo monofilético para el que se han descrito una serie de sinapomorfías morfológicas relacionadas con el aparato digestivo (Kantor, 1996; Ponder y Lindberg, 1996; Strong, 2003). Sin embargo, la monofilia de este gran clado no es apoyada por algunas de las filogenias moleculares, como la obtenida por Colgan *et al.* (2007) o las basadas en los genomas mitocondriales completos (Cunha *et al.*, 2009; Williams *et al.*, 2014; Osca *et al.*, 2015). En estas últimas, los Tonnoidea se recuperan en una posición intermedia entre los Cancellarioidea y el resto de los neogasterópodos, mientras que en la filogenia obtenida por Colgan *et al.* (2007) los Volutoidea se sitúan junto a los Tonnoidea.

1.5.1. Superfamilia Conoidea Fleming, 1822.

Los Neogastropoda se caracterizan por haber desarrollado la alimentación carnívora en sus distintas formas. Dentro de ellos, los Muricoidea, constituyen un hito en la evolución de los gasterópodos depredadores con la adquisición de un aparato bucal capaz de perforar las estructuras esqueléticas de sus presas. Sin embargo, posiblemente sean los Conoidea los que han alcanzado mayor sofisticación en este sentido, pues se caracterizan por poseer una rádula cuyos dientes se han convertido en algo así como un arpón hueco que contiene veneno y que inyectan a sus presas para

paralizarlas. Ello parece haber constituido una innovación clave que ha dotado a esta superfamilia de un enorme éxito evolutivo y es en la actualidad la que cuenta con un mayor número de especies dentro de los gasterópodos (Castelin *et al.*, 2012).

Tradicionalmente, Conoidea (=Toxoglossa) agrupaba a las familias Conidae, Terebridae y Turridae. Esta última familia, sin embargo, era considerada un "cajón de sastre" en el que se agrupaban a todas aquellas especies (generalmente diminutas) que no eran ni Conidae ni Terebridae. Aunque con los estudios recientes, los antiguos Turridae se han dividido y reordenado en diversos taxones, la denominación "túrridos" (en sentido amplio) se sigue utilizando para denominar a todos este vasto conjunto de especies y con este sentido se utiliza en esta tesis. Un punto de inflexión lo constituye la clasificación propuesta por Taylor *et al.* (1993) que, basada en caracteres morfológicos, dividió los Conoidea en seis familias (Conidae, Turridae, Terebridae, Drilliidae, Pseudomelatomidae y Strictispiridae), con la novedad de que ampliaron Conidae (familia previamente considerada monogenérica) incluyendo cinco subfamilias de los antiguos Turridae, mientras que éstos a su vez fueron subdivididos en cinco subfamilias. A partir de esta propuesta se han ido sucediendo nuevas reordenaciones y propuestas de clasificación de todo el grupo (Puillandre *et al.*, 2008; Tucker y Tenorio, 2009; Bouchet *et al.*, 2011; Puillandre *et al.*, 2011). Bouchet *et al.* (2011) propusieron una nueva clasificación de los hasta entonces polifiléticos "Turridae", reestructurándolos en 13 familias monofiléticas para contener hasta 358 géneros y subgéneros, e incluyendo una

diagnosis basada en la concha y rádula de todas ellas. De esta forma, la familia Turridae en sentido estricto quedaba reducida a 16 géneros. Esta es la clasificación que se sigue hoy día de forma general. Entre tanto, la monofilia dela familia Terebridae se ha confirmado en los diversos análisis filogenéticos (Holford *et al.*, 2009; Castelin *et al.*, 2012) en los cuales se propone una reorganización interna de los diferentes linajes.

1.5.1.1. Familia Conidae Fleming, 1822.

La familia Conidae incluye el tradicional (y bien conocido por los coleccionistas) género *Conus*, de origen relativamente reciente (Eoceno Inferior) y ampliamente distribuido por mares templados y cálidos, con más de 800 especies actuales conocidas, aunque su número aumenta continuamente con la descripción de nuevas especies (Puillandre *et al.*, 2014a).

La clasificación supraespecífica de los Conidae ha sufrido muchos cambios en las dos últimas décadas con la proliferación de subgéneros, inicialmente dentro de un único género *Conus*, pero sin establecerse unas relaciones filogenéticas claras. Según Puillandre *et al.* (2014b) se han propuesto se han llegado a proponer hasta la actualidad hasta cerca de 130 nombres genéricos válidos. Da Motta (1991) propuso seis géneros y 60 subgéneros dentro de la familia basándose sólo en caracteres de la concha, mientras que Tucker y Tenorio (2009, 2013) establecieron una clasificación completamente nueva basada también en caracteres de la rádula.

Estos autores propusieron subdivir la familia Conidae en cinco familias diferentes y 89 géneros. Sin embargo, la validez de esta clasificación ha sido cuestionada por Puillandre *et al.* (2014b).

Desde que se publicó la primera filogenia molecular del género *Conus* (Duda y Palumbi, 1999), se han ido sucediendo y solapando otras muchas, algunas de las cuales incluían también otros taxones relacionados (Puillandre *et al.*, 2008; Puillandre *et al.*, 2011; Puillandre *et al.*, 2014a) y otras centradas en algunos de los subgrupos (Duda y Palumbi, 1999; Espiritu *et al.*, 2001; Duda y Palumbi, 2004; Cunha *et al.*, 2005; Duda y Rolan, 2005; Duda *et al.*, 2008; Nam *et al.*, 2009; Kraus *et al.*, 2011; Kraus *et al.*, 2012, entre otros). En varios de estos trabajos se analiza también la evolución de los venenos y de la dieta, que principalmente puede ser vermívora, moluscivora o piscívora. La filogenia de Puillandre *et al.* (2014a) es la más completa y se basó en el análisis de tres genes mitocondriales de 330 especies mediante. Posteriormente esta filogenia fue utilizada para establecer una nueva clasificación a nivel de géneros y subgéneros de los Conidae (Puillandre *et al.*, 2014b), proponiendo cuatro géneros (*Californiconus*, *Profundiconus*, *Conasprella* y *Conus*), que representan los cuatro linajes que divergieron tempranamente hace unos 33 millones de años, y 71 subgéneros. Sin embargo, esta clasificación no es concluyente y requiere de un soporte muy robusto que permita entender como se ha generado la gran diversidad de la familia y otras cuestiones importantes, como el origen de la diferenciación de la dieta y la aparición de los venenos.

Hasta la fecha, sólo se conocía el genoma mitocondrial completo de cuatro especies del género *Conus*: *C.* (*Cylinder*) *textile* (Bandyopadhyay *et al.*, 2008), *C.* (*Gastridium*) *tulipa* (Chen *et al.*, 2015), *C.* (*Lautoconus*) *borgesi* (Cunha *et al.*, 2009), *C.* (*Splinoconus*) *tribblei* (Barghi *et al.*, 2015) y *C.* (*Pionoconus*) *consors* (Brauer *et al.*, 2012). No se conocen, por tanto, mitogenomas de especies de otros géneros de la familia, mientras que de otras familias de Conoidea se conocen los de *Xenuroturris cerithiformis* (Turridae; Bandyopadhyay *et al.* 2006), *Fusiturris similis* (Clavatulidae; Cunha *et al.* 2009) y *Oxymeris dimidiata* (Terebridae; Cunha *et al.* 2009). Con la representación de nuevos taxones dentro de la familia y superfamilia, en lo que se refiere a la secuenciación de sus mitogenomas, se pretende contribuir al establecimiento de una filogenia robusta de todo el grupo.

2 OBJETIVOS

Como se ha mencionado anteriormente, las relaciones entre los principales grupos de Gastropoda permanecen sin resolver y en constante controversia debido a las persistentes incongruencias que resultan de los diferentes análisis morfológicos y moleculares (tanto nucleares como mitocondriales). La falta de apoyo estadístico de las relaciones filogenéticas inferidas y la desigual representación de taxones se postulan como los principales problemas a resolver. En este sentido, y conociendo la demostrada utilidad de los genomas mitocondriales completos en la resolución de relaciones filogenéticas a diferentes niveles jerárquicos (subclases -en menor grado-, órdenes, superfamilias y familias), y que su catálogo es todavía muy insuficiente y desigual en los referente a los gasterópodos, se incide en la presente tesis en la secuenciación de estos marcadores moleculares en algunos grupos dentro de Gastropoda que están aún poco representados y, sin embargo, cuentan con un alto grado de diversificación.

Por lo tanto, esta tesis doctoral se plantea como objetivos principales:

1) Incrementar el número de genomas mitocondriales completos secuenciados de Gastropoda, especialmente dentro de las subclases Vetigastropoda y Neritimorpha para reconstruir con análisis probabilísticos sus relaciones filogenéticas internas.
2) Reconstruir la filogenia de Gastropoda.

3) Incrementar el número de genomas mitocondriales completos y datos *multilocus* dentro de dos grupos altamente diversificados, las superfamilias Conoidea (Caenogastropoda) y Trochoidea (Vetigastropoda), para reconstruir sus relaciones filogenéticas internas, e
4) Inferir la evolución de los reordenamientos génicos mitocondriales dentro de estos grupos objeto de estudio.

Para alcanzar estos objetivos generales, se han abordado los siguientes objetivos específicos:

2.1. Subclase Vetigastropoda

Esta subclase comprende diez superfamilias, pero sólo se conocen genomas mitocondriales completos de representantes de tres de ellas. En función de los taxones disponibles, se pretende secuenciar por primera vez genomas mitocondriales de representantes de las superfamilias Phasianelloidea, Angarioidea, Lepetodriloidea y Seguenzioidea, así como ampliar el número de genomas mitocondriales completos de superfamilias ya representadas como son Fissurelloidea y Trochoidea. Con los nuevos datos de secuencias se pretende:

1) Estudiar la evolución de reordenamientos génicos mitocondriales dentro de Vetigastropoda y en relación a otras subclases.

2) Reconstruir las relaciones filogenéticas dentro de Gastropoda con mejores apoyos estadísticos usando metodologías recientes que evitan la atracción de ramas largas, y

3) Reconstruir las relaciones filogenéticas dentro de Vetigastropoda.

Además, se pretende incorporar a los análisis comparativos y filogenéticos el genoma mitocondrial completo de un representante de Neomphalina (*Chrysomallon squamiferum*), que está disponible en NCBI, pero no ha sido estudiado. Neomphalina es un grupo de gasterópodos que en ocasiones se ha postulado como perteneciente a Vetigastropoda.

La superfamilia Trochoidea es el grupo más diverso dentro de Vetigastropoda y, sin embargo, sólo hay disponibles genomas mitocondriales completos para dos de sus ocho familias. En función de los taxones disponibles, se pretende secuenciar por primera vez los genomas mitocondriales de representantes de las familias Trochidae, Calliostomatidae y Margaritidae, así como ampliar el número de genomas mitocondriales de familias ya representadas, como Tegulidae y Turbinidae. Ello permitirá:

1) Estudiar la evolución de los reordenamientos génicos mitocondriales dentro de Trochoidea.

2) Establecer la posición relativa de Trochoidea dentro de Vetigastropoda, y

3) Reconstruir las relaciones filogenéticas dentro de Trochoidea.

Se pretende también generar datos *multilocus* (secuencias parciales de los genes mitocondriales *cox1*, *cob*, *rrnS* y *rrnL* y de los genes nucleares *28S ARNr* e *histona H3*) de la familia Trochidae y específicamente de las subfamilias hermanas Cantharidinae y Stomatellinae. Estos datos de secuencia permitirán:

1) Reconstruir las relaciones filogenéticas dentro de la subfamilia Cantharidinae, con especial atención a los géneros presentes en el Atlántico Noreste y Mediterráneo.

2) Inferir un cronograma con un reloj molecular relajado para determinar el tiempo de divergencia de los principales grupos dentro de esta subfamilia, y

3) Reconstruir las relaciones filogenéticas de los géneros de la subfamilia Stomatellinae.

2.2. Subclase Neritimorpha

La subclase Neritimorpha, a pesar de no ser tan numerosa en especies en comparación a otros grupos (Caenogastropoda y Vetigastropoda), representa una radiación peculiar e importante dentro de los gasterópodos en cuanto a formas de vida e invasión de los hábitats más diversos. Hasta la presente tesis, de los Neritimorpha solo había genomas mitocondriales completos disponibles para el género *Nerita* (superfamilia Neritoidea). Por lo tanto, en función de los taxones disponibles, se pretende secuenciar

por primera vez genomas mitocondriales de representantes de las demás superfamilias actuales (Neritopsoidea, Helicinoidea y Hydrocenoidea), así como ampliar el número de genomas mitocondriales a otros géneros de Neritoidea. Con ello se trata de:

1) Estudiar la evolución de reordenamientos génicos mitocondriales de las superfamilias incluidas en la subclase.

2) Reconstruir sus relaciones filogenéticas, y

3) Inferir un cronograma con un reloj molecular relajado para determinar el tiempo de divergencia de los diferentes linajes.

2.3. Subclase Caenogastropoda

La subclase Ceanogastropoda es la más diversificada dentro los gasterópodos y, dentro de ella, la superfamilia Conoidea es la que comprende un mayor número de especies actuales (se estima que más de 10.000). Actualmente sólo hay publicados ocho genomas mitocondriales que representan cuatro de las 16 familias de Conoidea. Además, su familia tipo (Conidae), solo cuenta con cinco genomas mitocondriales disponibles, todos ellos pertenecientes al mismo género (*Conus*).

En función de los taxones disponibles, se pretende secuenciar por primera vez genomas mitocondriales de géneros de la familia Conidae sin representación, como *Profundiconus*, *Californiconus*, *Conasprella* y *Lilliconus*, así como incrementar el número de

genomas mitocondriales para el género *Conus*. Además, se secuenciarán genomas mitocondriales de familias de Conoidea cercanas a Conidae. A partir de ello, se persigue:

1) Estudiar la evolución de reordenamientos génicos mitocondriales dentro de Conidae.

2) Reconstruir las relaciones filogenéticas entre los géneros de Conidae.

3) Reconstruir las relaciones filogenéticas dentro del género *Conus*, y

4) Producir un cronograma con un reloj molecular relajado para determinar el tiempo de divergencia de las especies del género *Conus* y de los diferentes géneros incluidos en la familia.

3 RESULTADOS

La metodología utilizada para la consecución de los objetivos y obtención de los resultados de esta tesis están detallados en cada uno de los capítulos correspondientes (publicaciones).

3.1. PUBLICACION 1:

Titulo: "Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry"
Autores: Juan E. Uribe, Yasunori Kano, José Templado y Rafael Zardoya.
Estado: Publicado
Año: 2016
Revista: *Zoologica Scripta* 42, 2:145-159

Resumen:

Se determinaron las secuencias de nucleótidos de los genomas mitocondriales (mt) completos o casi completos de siete vetigasterópodos: *Angaria neglecta* (Angarioidea), *Phasianella solida* (Phasianelloidea), *Granata lyrata* (Seguenzioidea), *Tegula lividomaculata* y *Bolma rugosa* (Trochoidea), *Diodora graeca* (Fissurelloidea) y *Lepetodrilus schrolli* (Lepetodriloidea). Mientras que los genomas mt de las superfamilias Angarioidea, Phasianelloidea, Seguenzioidea y Trochoidea se ajustan generalmente al orden ancestral de los genes de Vetigastropoda y Gastropoda, los de las superfamilias Fissurelloidea y Lepetodriloidea han sufrido importantes reordenamientos. El orden de los genes del ADNmt de *Chrysomallon squamiferum*, un representante de Neomphalina, también se analizó, debido a que se ha propuesto que este grupo está estrechamente relacionado con Vetigastropoda, y mostró un ordenamiento distinto. Las filogenias reconstruidas recuperaron Neomphalina como un linaje distinto de Gastropoda, que es el grupo hermano (sólo con un moderado apoyo de bootstrap) de un clado que incluye Vetigastropoda y Neritimorpha + Caenogastrpoda, mientras que la posición relativa

de Heterobranchia y Patellogastropoda en el árbol de Gastropoda no se pudo determinar definitivamente debido a sus ramas largas. Dentro de Vetigastropoda, la superfamilia Fissurelloidea se recuperó como el grupo hermano de dos linajes, uno incluyendo Lepetodriloidea como el grupo hermano de Seguenzioidea + Halitoidea, y otro incluyendo Phasianelloidea, Angarioidea y Trochoidea sin las relaciones resueltas. Se encontró que la rama larga de Fissurelloidea da inestabilidad significativa al árbol en la reconstrucción filogenética. La nueva filogenia apoya que la pérdida de la branquia del lado derecho ocurrió varias veces en la evolución de Vetigastropoda como se sugirió anteriormente y que Phasianelloidea, Angarioidea y Trochoidea radiaron de un antepasado común asimétrico (con branquia individual) que vivió en el Paleozoico medio.

Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry

Juan Esteban Uribe[1], Yasunori Kano[2], José Templado[1] and Rafael Zardoya[1*]

[1]Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006, Madrid, Spain

[2]Atmosphere and Ocean Research Institute, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8564, Japan.

* Corresponding author (rafaz@mncn.csic.es)

ABSTRACT

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), *Tegul*a *lividomaculata* and *Bolma rugosa* (Trochoidea), *Diodora graec*a (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 analysed 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 Heterobranchia 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 Palaeozoic.

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; Aktipis & Giribet, 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.

Table 1. New complete mitochondrial (mt) genomes analysed 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

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^{o}C for 60 s; 45 cycles of denaturalization at 94^{o}C 30 s, annealing at different temperatures within the range of 44-52^{o}C depending on the gene for 60 s and extension at 72^{o}C for 90 s; a final extension step at 72^{o}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^{o}C for 60 s;

45 cycles of denaturalization at 98^{o}C for 10 s, annealing at 53^{o}C for 30 s and extension at 68^{o}C for 60 s per kb; and a final extension step at 68^{o}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 (Table 1) available at NCBI (http://www.ncbi.nlm.nih.gov/). Eleven species of Gastropoda, one Cephalopoda, and one Caudofoveata were used as outgroups (Table 1).

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 1).

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 2). 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 3). The stop codons were variable depending on the gene and the species, and only *cox2* consistently ended with TAA (Supplementary Material 3). In *G. lyrata*, *nad1* and *atp8* genes were abnormally long (Supplementary Material 3). Each mt genome showed several intergenic regions, and those of *A. neglecta* were particularly long (up to 487 bp; see Supplementary Material 3). 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 3). 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 3). 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

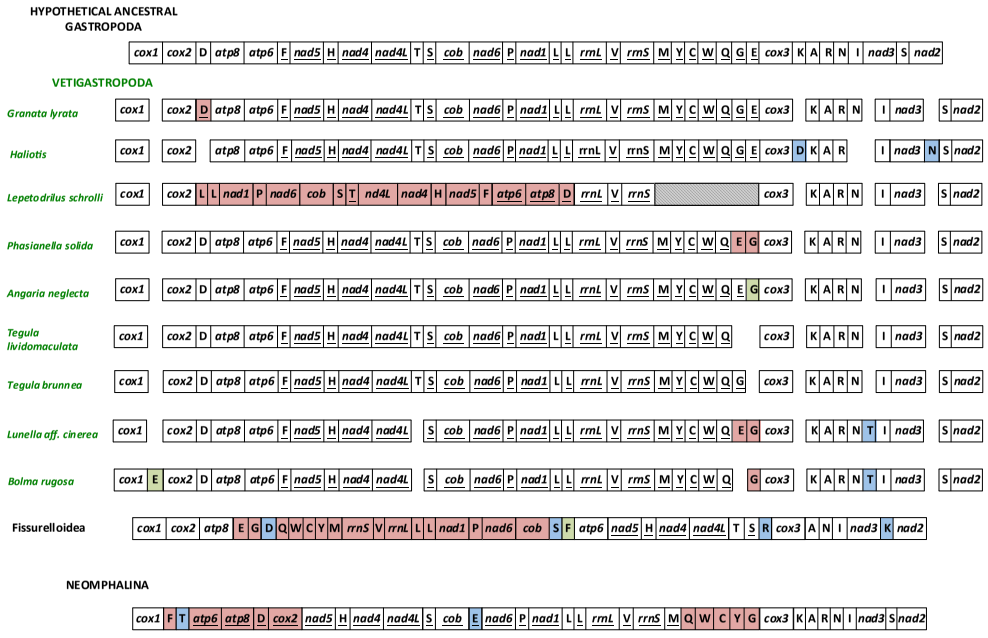


Figure 1. Mitochondrial gene orders of main lineages of Vetigastropoda. Gene orders in the hypothetical ancestral gastropod and Neomphalina are shown for comparison. Genes translocated are coloured in blue; inversions are in pink; genes translocated and inverted are in green. Genes encoded by the minor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

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).

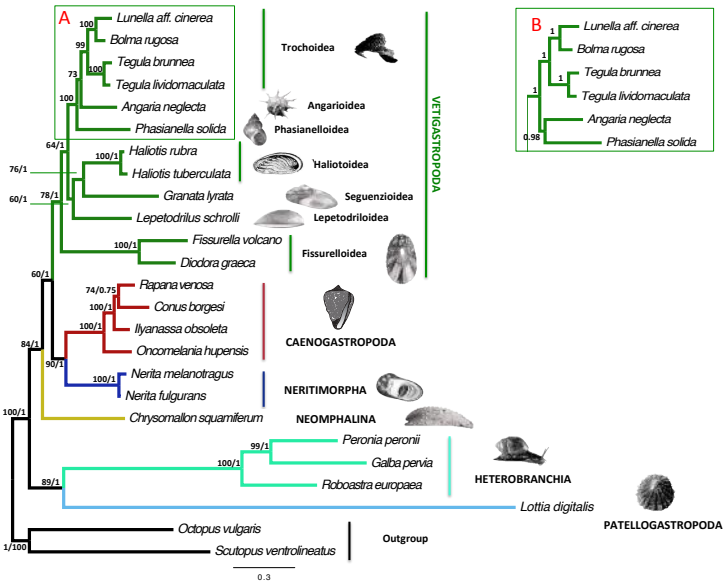


Figure 2. Phylogenetic relationships of Gastropoda based on mitochondrial sequence data. The maximum likelihood (ML) phylogram is shown (A). Topology differences in Bayesian inference (BI) are shown in the inset (B). Numbers at nodes are support values from ML (bootstrap proportions) and BI (Bayesian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

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).

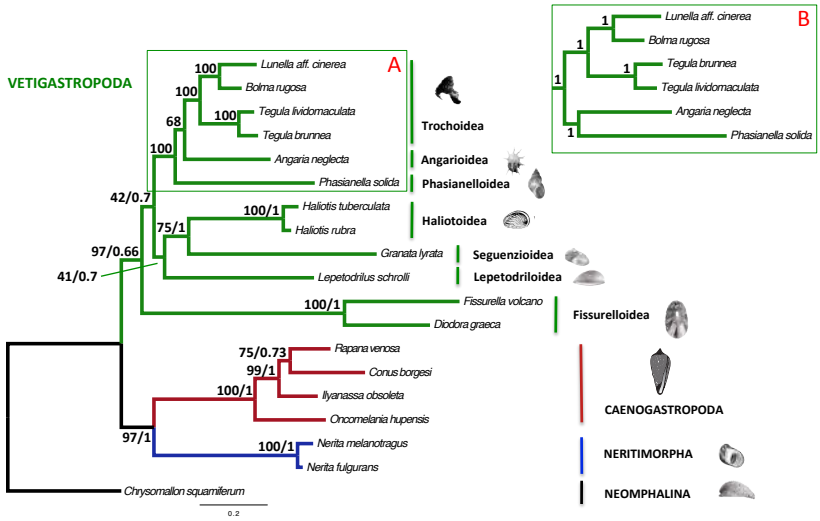


Figure 3. Phylogenetic relationships of Vetigastropoda based on mitochondrial sequence data. The maximum likelihood (ML) phylogram is shown (A). Topology differences in Bayesian inference (BI) are shown in the inset (B). Numbers at nodes are support values from ML (Bayesian bootstrap proportions) and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

Phylogenetic analyses using BI under the CAT-GTR+G model rendered a rather unresolved tree based on the gastropod data set (see Supplementary Material 4). 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 4). 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 4).

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; Aktipis & Giribet, 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; Aktipis & Giribet, 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; Aktipis & Giribet, 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; Aktipis & Giribet, 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; Williams, 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 *Halitis* 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.

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,.

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.

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*.

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.

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

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., 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.

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 Studies*, 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.

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.

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 Evolution*, 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.

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.

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 <italic>Galba pervia</italic> (Gastropoda: Mollusca), an Intermediate Host Snail of <italic>Fasciola</italic> 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.

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*, 0, 1-2.

Salvini-Plawen, L. v. (1980). A reconsideration of systematics in the Mollusca (phylogeny and higher classification). *Malacologia*, 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.

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.

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.

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.

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.

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.

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. *Proc. R. Soc. Lond. B*, 281.

SUPPORTING INFORMATION 1

Additional Supporting Information may be found in the online version of this article:

Data S1. Long PCR and primer walking primers.

Data S2. Complete mt genomes retrieved from Gen-Bank and analyzed in this study.

Data S3. Selected best -fit partitions and models.

Data S4. Annotation and main features of newly sequenced mt genomes.

Data S5. PhyloBayes of gastropod data set.

3.2. PUBLICACION 2:

Titulo: "Phylogenetic relationships within the superfamily Trochoidea (Gastropoda: Vetigastropoda) based on mitogenomes"
Autores: Juan E. Uribe, Suzanne Williams, José Templado, Samuel Abalde y Rafael Zardoya
Estado: En revisión
Año: 2016
Revista: *Journal of Molluscan Studies*

Resumen:

La gran diversidad morfológica y ecológica dentro de la superfamilia Trochoidea s.l. (Gastropoda: Vetigastropoda) ha impedido la reconstrucción de una filogenia robusta del grupo basada en morfología. Además, las filogenias moleculares recientes discrepan en la monofilia y relaciones internas de Trochoidea s.l., y en su posición relativa dentro de Vetigastropoda. Para mejorar la resolución del árbol filogenético de Trochoidea s.l. y Vetigastropoda se determinó la secuencia de nucleótidos del genoma mitocondrial (mt) completo de *Cittarium pica* (Tegulidae) y las secuencias de los genomas mt casi completos de *Tectus virgatus* (Tegulidae), *Gibbula umbilicaris* (Trochidae), y *Margarites vorticiferus* (Margaritidae). Además, se extrajeron de los transcriptómas de *Clanculus margaritarius* (Trochidae) y *Calliostoma zizyphinum* (Calliostomatidae) todos los genes codificantes de proteína y ARNr (excepto para las subunidades *atp* para *C. zizyphinum*). Los árboles filogenéticos reconstruidos con métodos probabilísticos y Neomphalina como grupo externo, no pudieron resolver las relaciones filogenéticas entre las superfamilias dentro de los vetigasterópodos, pero recuperaron con apoyo estadístico máximo el clado Trochoidea s.l., donde la familia Trochidae aparece como grupo hermano de la familia Calliostomatidae, formando un clado que era grupo hermano de los demás linajes: la familia Margaritidae fue hermana de un clado que incluye Phasianelloidea + Angarioidea y Turbinoidea + Tegulidae, siendo esta ultima familia parafíletica (*Cittarium* y *Tectus* deben ser adscritas a una nueva familia). El orden de los genes dentro de los genomas mt determinados, fue muy estable (solo con pocos reordenamientos restringidos a los genes ARNt) y conforme al genoma consenso para Vetigastropoda y Gastropoda.

Phylogenetic relationships within the superfamily Trochoidea (Gastropoda: Vetigastropoda) based on mitogenomes

Juan E. Uribe[1], Suzanne Williams[2], José Templado[1], Samuel Abalde[1], and Rafael Zardoya[1*]

[1]Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006, Madrid, Spain

[2]Department of Life Sciences, Natural History Museum, Cromwell Rd, London SW7 5BD, UK

*Corresponding author: rafaz@mncn.csic.es

ABSTRACT

The great morphological and ecological diversity within the superfamily Trochoidea s.l. (Gastropoda: Vetigastropoda) has hindered in the past the reconstruction of a robust phylogeny for the group based on morphology. Moreover, previous recent molecular phylogenies disagreed on the monophyly and internal relationships of Trochoidea s.l., as well as on its relative phylogenetic position within Vetigastropoda. In order to further resolve the trochoidean and vetigastropod phylogenetic trees, the nucleotide sequence of the complete mitochondrial (mt) genome of *Cittarium pica* (Tegulidae) and the nearly complete mt genomes of *Tectus virgatus* (Tegulidae), *Gibbula umbilicaris* (Trochidae), and *Margarites vorticiferus* (Margaritidae) were determined. In addition, the nucleotide sequences of all protein coding and rRNA genes of *Clanculus margaritarius* (Trochidae) and of *Calliostoma zizyphinum* (Calliostomatidae; except for *atp* subunits) were derived from transcriptomic sequence data. The reconstructed phylogenetic trees using probabilistic methods and Neomphalina as outgroup could not resolve with confidence phylogenetic relationships among vetigastropod superfamilies but recovered with maximal support a Trochoidea s.l. clade that included superfamilies Trochoidea, Angarioidea and Phasianelloidea. Within Trochoidea s.l., the family Trochidae was placed as sister group of the family Calliostomatidae, forming a clade that was sister group to the remaining lineages: the family Margaritidae was sister to a clade including Phasianelloidea + Angarioidea and Turbinidae + Tegulidae, this latter family being paraphyletic (*Cittarium* and *Tectus* need to be assigned to a new family). Gene order within newly determined mt genomes was very stable (with only few rearrangements restricted to tRNA genes) and conformed to the vetigastropod and gastropod consensus genome organizations.

INTRODUCTION

Trochoidea s.l. Rafinesque, 1815 (top shells, turban shells, and allies) is one of the most ecologically and morphologically diverse lineage of marine gastropods and, by far, the largest superfamily belonging to the subclass Vetigastropoda, with more than 2,000 living species grouped into about 500 recognized genera (Hickman, 1996; Geiger, Nützel & Sasaki, 2008). The clade is worldwide distributed and is present throughout all seas and oceans, at all latitudes and bathymetric ranges (Hickman & McLean, 1990; Williams, Karube & Ozawa, 2008). Trochoideans play an important ecological role as a predominant element in different marine communities such as intertidal rocky shores, seagrass beds, or coral reefs, and they are also found in many other marine habitats (Williams, *et al.*, 2008). They have a long fossil record that goes back to the Middle Triassic, 228-245 million years ago (Mya), but the time of the origin of the group is certainly much older (Hickman & McLean, 1990; Williams, *et al.*, 2008).

The taxonomic internal classification of Trochoidea has a long history of controversy and instability. In their comprehensive morphological monograph on trochacean gastropods, Hickman & McLean (1990) maintained the three families traditionally recognized within the superfamily i.e., Trochidae, Turbinidae and Skeneidae, but readjusted the different genera into various subfamilies and tribes based on suites of shared morphological characters. Later, in the taxonomic classification of gastropods proposed by (Bouchet *et al.*, 2005), the family Turbinidae

(including the subfamily Skeneinae) was classified within the superfamily Turbinoidea. However, major changes to the systematics of Trochoidea were based on recent molecular phylogenies (Geiger & Thacker, 2005; Williams & Ozawa, 2006; Kano, 2008; Williams, *et al.*, 2008; Williams, 2012), which challenged the monophyly of the superfamily as well as of several of the internal groups as defined by Hickman & McLean (1990), and prompted for important changes to the taxon composition and arrangement of families (Williams, 2012). For instance, some taxa were transferred to the new superfamily Seguenzioidea (Kano, 2008), and a number of minute skeneimorph genera were variously relocated either to Seguenzoidea (Kano, Chikyu & Warén, 2009; Haszprunar *et al.*, 2016), Neomphalina (Kunze *et al.*, 2008), or to the new family Crosseolidae of uncertain taxonomic position (Hickman, 2013). Furthermore, several molecular studies redefined the family Turbinidae (Williams & Ozawa, 2006), established the new superfamilies Angarioidea and Phasianelloidea (Williams, *et al.*, 2008), and restricted Trochoidea to the families Trochidae, Turbinidae, Solariellidae, Calliostomatidae, Liotiidae, Skeneidae, Margaritidae and Tegulidae (Williams, 2012).

None of these taxonomic changes was definitive and the debate over the final composition and internal phylogenetic relationships of Trochoidea remains more alive than ever. Moreover, this question is directly related to resolving phylogenetic relationships among the different superfamilies of Vetigastropoda. In this regard, some studies recovered Phasianelloidea and Angarioidea in early-branching positions of the Vetigastropoda tree

after the divergence of Pleurotomarioidea (Williams, *et al.*, 2008; Aktipis & Giribet, 2012) whereas more recent phylogenies grouped Phasianelloidea and/ or Angarioidea with Trochoidea (Zapata *et al.*, 2014; Uribe *et al.*, 2016). While earlier studies were based on few partial mitochondrial and nuclear genes and a rather extensive lineage representation, later ones were based on phylogenomic data but with reduced taxon sampling.

Phylogenetic analysis of complete mitochondrial (mt) genomes resulted in good resolution among vetigastropod superfamilies (Uribe, *et al.*, 2016) and therefore, they are good candidates to resolve phylogenetic relationships within Trochoidea. Thus far, there are available 13 complete or nearly complete mt genomes of Vetigastropoda, which represent the living superfamilies Fissurelloidea, Lepetodriloidea, Siguenzioidea, Haliotoidea, Angarioidea, Phasianelloidea, and Trochoidea (no mt genome has been sequenced for Pleurotomarioidea and Lepetelloidea). However, the great diversity of Trochoidea is clearly underrepresented, as mt genomes for only four species belonging to families Turbinidae and Tegulidae are published (Uribe, *et al.*, 2016). Here, we increased considerably the number of complete mt genomes representing different families within Trochoidea to test the monophyly and address internal phylogenetic relationships of the superfamily, as well as to resolve its relative phylogenetic position within Vetigastropoda. In addition, the reconstructed phylogeny was used to determine whether trochoidean mt genomes show rearrangements in their genes orders.

MATERIALS AND METHODS

Samples and DNA/ RNA extraction

One specimen each of *Cittarium pica* (Tegulidae), *Tectus virgatus* (Tegulidae), *Gibbula umbilicaris* (Trochidae), *Clanculus margaritarius* (Trochidae), *Calliostoma zizyphinum* (Calliostomatidae), and *Margarites vorticiferus* (Margaritidae) was used for this study (See Table 1, for details on the locality, collector, and voucher ID of each sample). Samples of *C. pica*, *T. virgatus*, *G. umbilicaris*, and *M. vorticiferus*, were stored in 100% ethanol at -20 ºC, and total genomic DNA was isolated from up to 30 mg of foot tissue following a standard phenol chloroform extraction.

Samples of *C. margaritarius*, and *C. zizyphinum* were stored in RNALater at -80 ºC, and total RNA was isolated from mantle tissue using the RNeasy Fibrous Tissue Mini Kit (Qiagen) according to the manufacturer's instructions. Total RNA was quantified and its integrity assessed using a Qubit® 2.0 Fluorometer RNA assay kit and an Agilent 2200 Tapestation using a high sensitivity R6K Screen Tape, respectively. Dynabeads® mRNA DIRECT™ Micro Kit (Ambion, Life Technologies) were used to isolate mRNA using the 100ng-1µg µg total RNA protocol.

PCR amplification and sequencing

Two alternative strategies were carried out to obtain mitogenomic sequence data. For *C. pica*, *T. virgatus*, *G. umbilicaris*, and *M.*

vorticiferus, complete or nearly complete mt genomes were PCR amplified and sequenced, whereas for *C. margaritarius*, and *C. zizyphinum*, transcriptomic sequence data was generated and mt protein-coding and rRNA genes were assembled.

For obtaining complete or nearly complete mt genomes a three-step strategy was used. First, fragments of *cox1*, *rrnL*, and *cox3* genes were amplified using the primers respectively detailed in (Folmer et al., 1994), (Palumbi et al., 1991), and (Boore and Brown, 2000). 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 (10 mM), 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 PCR temperature and cycle conditions used were: a denaturalization step at 94 °C for 60 s; 45 cycles of denaturalization at 94 °C 30 s, annealing at 44 (*cox1*) or 52 (*rrnL* and *cox3*) °C for 60 s and extension at 72 °C for 90 s; a final extension step at 72 °C for 5 m. Second, the amplified PCR fragments were sequenced using Sanger sequencing, and new primers were designed (see Suppl. Mat. for primer sequences) for amplifying outwards from the short fragments in the next step. Third, the remaining mtDNA was amplified in two-three overlapping fragments by long PCR using the newly designed primers. The long PCR reaction contained 2.5 µl of 10x 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, 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 denaturing

Table 1. Mitochondrial (mt) DNA data analyzed in this study. Length in bp, Genbank accesion number, museum voucher, sampling location, and name of collector are provided.

New mt data

Species	Family	Superfamily	bp	Acc. No.	Voucher	Location	Collector
Cittarium pica	Unassigned	Trochoidea	17,949	To be obtained	MNCN/ADN: XXX	Guanahacabibes, Bolivar, Cuba	José Templado
*Tectus virgatus**	Unassigned	Trochoidea	13,891	To be obtained	MNCN/ADN: XXX	Aqaba, Jordan	José Templado
*Margarites vorticiferus**	Margaritidae	Trochoidea	15,254	To be obtained	NHMUK: 20110451	Amchitka I., Cannikin, USA	Piotr Kuklinski
*Gibbula umbilicaris**	Trochidae	Trochoidea	12,885	To be obtained	MNCN/ADN: 86692	El Mohon, Murcia, SE Spain	José Templado
Clanculus margaritarius[†]	Trochidae	Trochoidea	—	To be obtained	NHMUK 20150502	[§]Kitahama, Shirahama, Nishimu	Tomo Nakano
Calliostoma zizyphinum[†]	Calliotomatidae	Trochoidea	—	To be obtained	NHMUK 20160315	Shetland Islands, 60° 14.9'N, 01[ε]	Piotr Kuklinski

GenBank mt data

Species	Family	Superfamily	bp	GenBank Acc. No.	Reference
Tegula brunnea	Tegulidae	Trochoidea	17690	NC_016954	Simison, 2011 (unpublished)
Tegula lividomaculata	Tegulidae	Trochoidea	17375	NC_029367	Uribe et al., 2016
Bolma rugosa	Turbinidae	Trochoidea	17432	NC_029366	Uribe et al., 2016
Lunella aff. cinerea	Turbinidae	Trochoidea	17670	KF700096	Williams et al., 2014
Phasianella solida	Phasianellidae	Phasianelloidea	16.698	KR297251	Uribe et al., 2016
Angaria neglecta	Angaridae	Angaroidea	1947	KR297248	Uribe et al., 2016
Haliotis rubra	Haliotidae	Haliotoidea	16907	NC_005940	Maynard et al., 2005
Haliotis tuberculata	Haliotidae	Haliotoidea	16521	NC_013708	VanWormhoudt et al., 2009
Granata lyrata	Chilofontidae	Seguenzioidea	17632	NC_028708	Uribe et al., 2016
Fissurella volcano	Fissurellidae	Fissurelloidea	17575	NC_016953	Simison, 2011 (unpublished)
Diodora graeca	Fissurellidae	Fissurelloidea	17209	KT207825	Uribe et al., 2016
Lepetodrilus nux	Lepetodrilidae	Lepetodriloidea	16353	LC107880	Nakajima et al., 2016
*Lepetodrilus schrolli**	Lepetodrilidae	Lepetodriloidea	15579	KR297250	Uribe et al., 2016
Chrysomallon squamiferum	Peltospiridae	Neomphaloidea	15388	AP013032	Nakagawa et al., 2014

*Nearly complete mt genomes

[†] The GenBank Acc. No. of each mt gene are showed consecutively in this order: *atp6, atp8, cob, cox1, cox2, cox3, nad1, nad2, nad3, nad4, nad4L, nad5, nad6, rrnL*, and *rrnS*.

[§]Locality shown for sequenced specimen; [ε]ote that voucher is a different specimen from another locality.

step at 94 °C for 60 s; 45 cycles of denaturation 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. 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, a separate indexed library was constructed using the NEXTERA XT DNA library prep Kit (Illumina, San Diego, CA, USA) and sequenced in a single lane of Illumina MiSeq at Sistemas Genómicos (Valencia, Spain).

Transcriptomes were sequenced and assembled for *Clanculus margaritarius* and *Calliostoma zizyphinum* using the methods outlined in Williams et al. (2016).

Assembly and annotation

The reads corresponding to the different PCR amplified mt genomes were sorted using the indexes. Adapter sequences were removed using SeqPrep (St John, 2011). Assembly was performed using the TRUFA webserver (Kornobis et al., 2015) 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 and Edwards, 2011). Filtered reads were used for *de novo* assembly of mt genomes, searching for contigs with a minimum length of 3 kb. The complete or nearly 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% and estimate coverage using Geneious® 8.0.3.

Genome annotation was performed by setting a limit of nucleotide identity of 75% to previously reported vetigatropods mt genomes (Uribe et al., 2016b) using Geneious® 8.0.3. The annotated 13 mt protein-coding genes were further corroborated by identifying the corresponding 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), which infers cloverleaf secondary structures. The ribosomal RNA (rRNA) genes were identified by sequence comparison with previously reported vetigastropod mt genomes, and assumed to extend to the boundaries of adjacent genes (Boore et al., 2005). GenBank accession numbers of each mt genome are provided in Table 1.

The mt protein coding and rRNA genes of *C. margaritarius*, and *C. zizyphinum* were extracted from the corresponding transcriptomes in Geneious by using published amino acid sequences for each mitochondrial gene from *Bolma rugosa* (GenBank KT207824) to identify matching sequences in the dataset of assembled contigs using the tBLASTx option, then the new contig was used as a reference sequence against the original reads to obtain full length genes.

Sequence alignment

The nucleotide sequences of the 13 protein coding and two rRNA genes encoded in the newly determined complete or nearly complete mt genomes were aligned each separately with the corresponding orthologous sequences of all vetigastropod complete or nearly complete mt genomes available at NCBI (www.ncbi.nlm.nih.gov/; see Table 1). The complete mt genome of *Chrysomallon squamiferum* (Neomphalina) was used as outgroup following (Uribe et al., 2016b). Each protein-coding gene was aligned with Translator X (Abascal et al., 2010) using the deduced amino acid sequence as guide whereas rRNA genes were aligned separately using MAFFT v7 (Katoh and Standley, 2013) with default parameters. Ambiguously aligned positions were removed using Gblocks v.0.91b (Castresana, 2000) with the following settings: minimum sequence for flanking positions: 85%; maximum contiguous non-conserved positions: 8; minimum block length: 10; gaps in final blocks: no. The generated single alignments were concatenated using Geneious® 8.0.3.

Phylogenetic analyses

Phylogenetic relationships were reconstructed using Bayesian inference (BI; (Huelsenbeck and Ronquist, 2001) and maximum likelihood (ML; (Felsenstein, 1981). BI analyses were conducted using MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003) and running four simultaneous Monte Carlo Markov chains (MCMC) for 10 million generations, sampling every 1,000 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 BI runs were performed to increase the chance of adequate mixing by the MCMC and to increase the chance of detecting failure to converge, as determined using Tracer v1.6 (Rambaut and Drummond, 2007). ML analyses were conducted with RAxML v7.3.1 (Stamatakis, 2006) and default parameters using the rapid hill-climbing algorithm and 10,000 bootstrap pseudoreplicates.

The program Partition Finder (Lanfear et al., 2012) was used to select best partition schemes and best-fit models of substitution (see Suppl. Mat.) according to the Bayesian information criterion (BIC; (Schwarz, 1978). For protein-coding genes, the partitions tested were: all genes combined, all genes separated except *atp6-atp8* and *nad4-nad4L*, and genes grouped by subunits (*atp*, *cox, cob* and *nad*). In addition, the three above partition schemes were tested considering first, second, and third codon positions separated. For the mt rRNA genes, the two genes combined or separated were tested.

RESULTS AND DISCUSSION

Sequencing and assembly

The nucleotide sequences and gene arrangement of the complete mt genome of *Cittarium pica* and the nearly complete mt genomes of *Tectus virgatus*, *Gibbula umbilicaris*, and *Margarites vorticiferus*

were determined (see annotation and main features in Suppl. Mat.). In *Tectus* and *Gibbula* a fragment of about 3kb between *rrnL* and *cox3* genes could not be PCR amplified. In the case of *Margarites*, a shorter fragment of about 2kb between *rrnS* and *cox3* genes was missing (Fig. 1). In addition, the nucleotide sequences of all protein coding and rRNA genes of *Clanculus margaritarius* and of *Calliostoma zizyphinum* (except for *atp* subunits) were derived from transcriptomic sequence data. The number of reads, mean coverage, and sequence length (bp) of each mt genome are: *C. pica* (165,292, 1,390x and 17,949); *T. virgatus* (205,498, 2,218x and 13,891); *G. umbilicaris* (142,074, 1,666x and 12,885); *M. vorticiferus* (290,484, 2,858x and 15,254).

Mitochondrial genome organization

Genome organization could only be determined for those mt genomes that were amplified by long PCR (all but C. *margaritarius* and *C. zizyphinum*). These mt genomes share the same gene order with regards to the relative position of protein-coding genes, and only minor changes affecting individual tRNA genes were observed (Fig. 1). The consensus gene order for Trochoidea s.l. (including Phasianelloidea and Anagraioidea) is the same observed in Haliotoidea and Siguenzoidea but not in Fissurelloidea and Lepetodriloidea (Uribe et al. 2016). Moreover, this consensus gene order is similar to that reported for Neritimorpha (except Helicinoidea; Uribe et al. 2016) and conforms to the genome arrangement of the hypothetical ancestor of gastropods (Fig. 1;

Uribe et al. 2016). With respect to this gastropod ancestral gene order, the mt genome of *T. virgatus* showed a translocation of the *trnQ* to a new relative position between *cob* and *nad6* genes in the minor strand (Fig. 1). The mt genome of *G. umbilicaris* had an inversion of the *trnF* and *trnT* genes from major to minor strand (Fig. 1). The mt genome of *M. vorticiferus,* showed a translocation of the *trnM* to a new relative position between *nad6* and *trnP* genes in the minor strand (Fig. 1). Finally, the mt genomes of *Tegula, Bolma*. *Lunella, Cittarium, Phasianella* and *Angaria* showed rearrangements affecting *trnG* and *trnE* genes, and in some instances, one or both genes were missing (Fig. 1). It is not possible to infer the exact evolution of these rearrangements given that this part of the mt genome could not be sequenced in *Tectus, Margarites*, and *Gibbula,* and is not available for *Clanculus* and *Calliostoma* (Fig. 1). However, it is important to note that these two genes are located at the end of the MCYWQGE tRNA gene cluster, and just before the hypothesized control region of gastropod mt genomes, which is known to act as hotspot of gene order rearrangements (Duarte, *et al*., 2008).

Phylogenetic relationships among vetigastropod superfamilies and within Trochoidea s.l.

A molecular phylogeny of Vetigastropoda was reconstructed using probabilistic methods. The final alignment was 12,088 positions long. The best partition scheme was the one having all protein-coding genes combined (but with each codon position analyzed separately) and the two rRNA genes combined. The best-fit model

for the different partitions was GTR+I+G. The ML (-lnL = 152731.92) and BI (-lnL = 15274,15 for run 1; -lnL = 15273,89 for run 2) phylogenetic analyses arrived at the same topology using Neomphalina as outgroup (Fig. 2). The superfamily Lepetodriloidea was recovered as sister group of the remaining vetigastropds, although only with moderate statistical support (56% BP, 0.94 BPP). The next lineage that branched off was Fissurelloidea, whose members exhibited relatively long branches (Fig. 2). The superfamilies Siguenzoidea and Haliotoidea formed a well-supported clade (74% BPP, 1 BPP), which was the sister group of Trochoidea s.l. (with Phasianelloidea and Angarioidea), although this relationship was only moderately supported (56% BP, 1 BPP). Resolving the relative phylogenetic position of vetigastropod lineages has been challenging, and previous molecular phylogenies have also rendered inconclusive results due to low statistical support at this part of the Vetigastropoda tree (e.g., (Kano, 2008; Aktipis and Giribet, 2012; Zapata et al., 2014; Uribe et al., 2016b). Moreover, in the present analysis, we could not incorporate a representative of the superfamily Pleurotomarioidea, which in other phylogenies is placed as sister group of the remaining vetigastropod lineages (Kano, 2008; Williams et al., 2008; Zapata et al., 2014) or even unrelated to Vetigastropoda (Aktipis and Giribet, 2012). Other missing superfamilies were Lepetelloidea and Scissurelloidea.

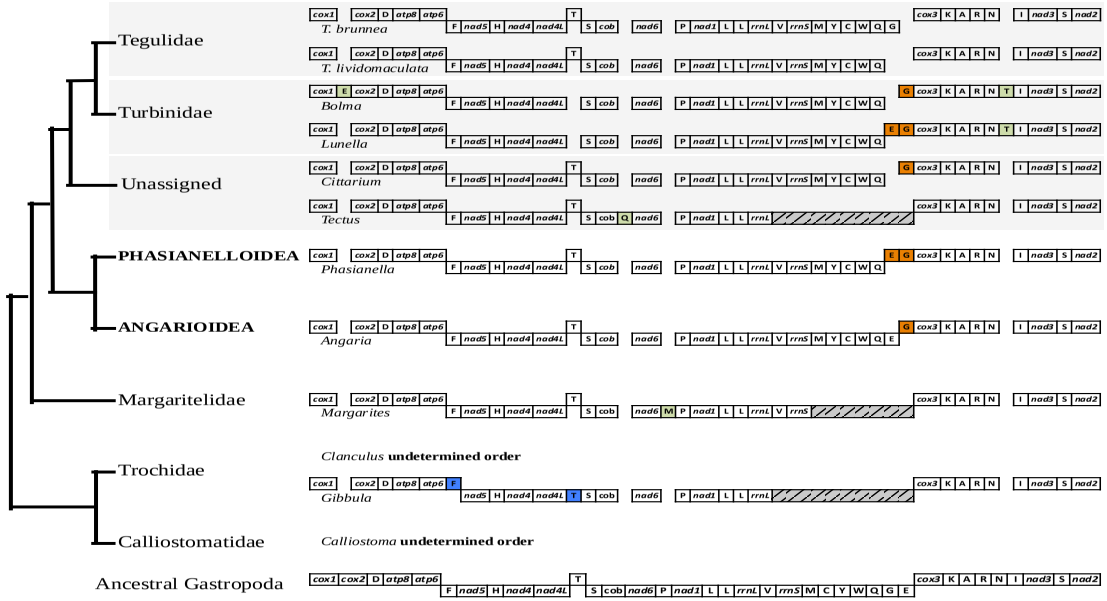


Figure 1. Gene orders of Trochoidea s.l. mitochondrial genomes. The consensus genome organization is shown for each lineage as well as the ancestral gene order for Gastropoda. The genes encoded in the major and minor strands are shown in the top and bottom lines, respectively. Gene rearrangements (restricted to tRNA genes) are indicated by colors. Translocated genes are in green. Inverted genes are in blue. Genes for which the exact rearrangement could not be inferred are in orange. Striped boxes indicate regions not sequenced. The gene order of *Clanculus* and *Calliostoma* could not be determined because sequence data was derived from RNA Seq.

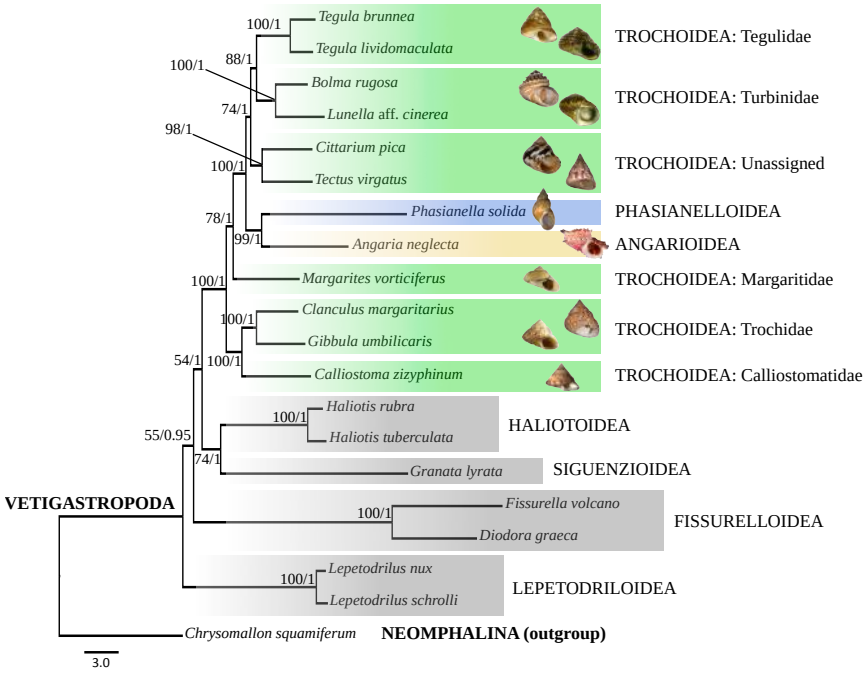


Figure 2. Phylogenetic relationships among vetigastropod superfamilies and within Trochoidea s.l. based on complete mt genomes. The reconstructed ML phylogram using Neomphalina as outgroup is shown. Numbers at nodes are statistical support values for ML (bootstrap proportions in percentage)/ BI (posterior probabilities). Vetigastropod superfamilies and trochoidean families are indicated.

The main focus of the present phylogenetic analysis was Trochoidea s.l. This clade received maximal support and included Trochoidea, Phasianelloidea and Angarioidea *sensu* Williams et al 2008 (Fig. 2). The recognition of Phasianelloidea and Angarioidea as new superfamilies different from Trochoidea was based on phylogenetic analyses of partial mt and nuclear genes that placed these two lineages in early diverging positions in the vetigastropod

tree (Kano, 2008; Williams et al., 2008; Aktipis and Giribet, 2012). However, our results are in agreement with more recent phylogenies based on mt (Uribe et al., 2016b) and nuclear (Zapata et al., 2014) genomic data sets, which also recovered a clade grouping Trochoidea together with Phasianelloidea and Angarioidea (this latter lineage was missing in (Zapata et al., 2014). Interestingly, Phasianelloidea and Angarioidea show relatively long branches, which in previous studies with different taxon sampling and molecular markers may have produced a long-branch attraction effect and the pulling of these two lineages to basal positions.

The representation of Trochoidea in recent phylogenomic analyses was rather limited, a situation that has been reverted in the present phylogenetic analysis. The reconstructed phylogenetic tree recovered the family Trochidae as sister group of the family Calliostomatidae with maximal statistical support, and this clade was the sister group of the remaining lineages, which formed a monophyletic group with high support (79% BP; 1 BPP). The family Margaritidae was sister of a maximally supported clade including Phasianelloidea + Angarioidea and Turbinidae + (paraphyletic) Tegulidae (Fig. 2). The recovered internal phylogenetic relationships are fully congruent with (Williams, 2012), who did not include Phasianelloidea and Angarioidea in her phylogenetic analysis. In particular, it is worth noting that *Cittarium* and *Tectus* (and possibly *Rochia*) need to be ascribed to a new family. Additionally, the recovered phylogeny prompts for a redefinition of Trochoidea, which awaits further mitogenomic

studies including missing families such as Skeneidae, Solariellidae, and Liotiidae.

ACKNOWLEDGEMENTS

We are grateful to Jesus 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; BES-2014-069575 to SA) and by funding from the NHM Department of Life Sciences to STW.

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.

Aktipis, S.W., Giribet, G., 2012. Testing relationships among the vetigastropod taxa: a molecular approach. Journal of Molluscan Studies 78, 12-27.

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

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. Methods in Enzymology 395, 311 - 348.

Bouchet, P., Rocroi, J.P., Frýda, J., Hausdorf, B., Ponder , W.F., Valdés, Á., Warén, A., 2005. Classification and nomenclator of gastropod families. Malacologia 47, 1-397.

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

Duarte, G.T., de Azeredo-Espin, A.M.L., Junqueira, A.C.M. 2008. The mitochondrial control region of blowflies (Diptera: Calliphoridae): a hot spot for mitochondrial genome rearrangements. Journal of Medical Entomology 45, 667-676.

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.

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

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

Haszprunar, G., Kunze, T., Brückner, M., Heß, M., 2016. Towards a sound definition of Skeneidae (Mollusca, Vetigastropoda): 3D interactive anatomy of the type species, *Skenea serpuloides* (Montagu, 1808) and comments on related taxa. Organisms Diversity & Evolution, 1-19.

Hickman, C.S., 1996. Phylogeny and patterns of evolutionary radiation in trochoidean gastropods. In: Taylor, J.D. (Ed.), Origin and Evolutionary Radiation of the Mollusca. Oxford University Press, Oxford, pp. 177–198.

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.

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.

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

Kunze, T., Beck, F., Bruånckner, M., Heß, M., Haszprunar, G., 2008. Skeneimorph gastropods in Neomphalina and Vetigastropoda - A preliminary report. Zoosymposia 1.

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.

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.

Rambaut, A., Drummond, A.J., 2007. Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

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

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, W686-W689.

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

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics 6, 461-464.

St John, J., 2011. SeqPrep. . Available from https://github.com/jstjohn/SeqPrep.

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

Uribe, J.E., Kano, Y., Templado, J., Zardoya, R., 2016b. Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. Zoologica Scripta 45, 145-159.

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

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., Lockyer, A.E., Dyal, P., Hopkins, K., Nakano, N., Churchill, C.K.C., Speiser, D.I. 2016. Colourful seashells: identification of genes associated with the synthesis of shell colour in two marine snails. Molecular Ecology (in review).

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.

Zapata, F., Wilson, N.G., Howison, M., Andrade, S.C.S., Jörger, K.M., Schrödl, M., Goetz, F.E., Giribet, G., Dunn, C.W., 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. Proceeding of the Royal Society of London Biological series 281: 20141739.

SUPPORTING INFORMATION 2

Additional Supporting Information may be found in the online version of this article:

Data S1. Amplification strategy. Long PCR and primer walking primers.

Data S2. Mitochondrial genome features.

3.3. PUBLICACION 3:

Titulo: "Phylogenetic relationships of mediterranean and North-East Atlantic Cantharidinae and notes on Stomatellinae (Vetigastropoda: Trochidae)"
Autores: Juan E. Uribe, Suzanne T. Williams, José Templado, Barbara Buge y Rafael Zardoya
Estado: Aceptado
Año: 2016
Revista: *Molecular Phylogenetics and Evolution*

Resumen:

La subfamilia Cantharidinae Gray, 1857 (Trochoidea: Trochidae) incluye 23 géneros reconocidos y más de 200 especies vivas. Estos caracoles marinos son herbívoros microfagos que generalmente viven sobre rocas costeras, en macroalgas y en praderas de fanerógamas marinas de aguas subtropicales y templadas del la región biogeográfica Indo Pacíifico Central y Occidental hasta el Mar Mediterráneo y el Océano Atlántico Éste. Estudios filogenéticos moleculares recientes revisando la familia Trochidae apoyaron la monofília de la subfamilia Cantharidinae y a la subfamilia Stomatellinae como su grupo hermano. Estos estudios y otros más has estado mayoritariamente enfocados en los miembros del Indo Pacífico de la subfamilia Cantharidinae, por lo que aquí, investigamos las relaciones filogenéticas de su contraparte del Mar Mediterráneo y el Océano Atlántico Noreste (NE) incluyendo 33 especies de los géneros *Gibbula*, *Jujubinus*, *Phorcus*, *Clelandella*, y *Callumbonella*. Los grupos del Mediterráneo y NE Atlántico fueron complementados con 30 especies de Cantharidinae del Indo Pacífico más 19 especímenes del la familia hermana Stomatellinae. Se reconstruyeron árboles filogenéticos utilizando inferencia Bayesiana y máxima verosimilitud con dos sets de datos que incluían secuencias parciales de cuatro o seis genes mitocondriales (*cox1*, *rrnL*, *rrnS*, y *cob*) y nucleares (*28S rRNA* y *histona H3*). Se recuperó con un alto apoyo estadístico un clado que incluía todos los grupos del Mediterráneo y NE Atlántico., pero no se pudo determinar con confianza su grupo hermano entre los linajes del Indo Pacífico (aunque la asignación de *Trochus kotschyi* a *Priotrochus* pudo ser rechazada). Dentro del clado del Mediterráneo

y NE Atlántico, los géneros *Phorcus* y *Jujubinus* fueron recuperados como recíprocamente monofiléticos, y los géneros de profundidad *Clelandella* y *Callumbonella* se situaron como grupo hermano de *Jujubinus* con alto apoyo. Sin embargo, el género *Gibbula* como comvencionalmente es definido, no fue monofilético y las especies que lo conforman fueron divididas en tres grupos principales y dos linajes independientes. Las relaciones filogenéticas entre *Phorcus*, *Jujubinus* (más *Clelandella* y *Callumbonella*), y los diferentes clados de *Gibbula* se resolvieron pobremente en los análisis filogenéticos de cuatro genes, pero recibieron moderado apoyo en los basados en seis genes. Se realizó una primera aproximación a la resolución de las relaciones filogenéticas dentro de Stomatellinae que mostró como la diversidad de la subfamilia está actualmente muy imfraestimada, y que *Calliotrochus* es posiblemente un miembro de esta subfamilia. Se reconstruyó un cronográma usando un reloj molecular relajado lognormal, y el origen del clado del Mediterráneo y NE Atlántico fue datado justo después de la fase *Azolla* en el Eoceno Medio hace 48 millones de años, mientras que la diversificación de los principales clados (géneros) siguió al cierre oriental del Mar de Tetis en el Mioceno Medio hace 14 millones de años.

Phylogenetic relationships of Mediterranean and North-East Atlantic Cantharidinae and notes on Stomatellinae (Vetigastropoda: Trochidae)

Juan E. Uribe[1], Suzanne T. Williams[2], José Templado[1], Barbara Buge[3] and Rafael Zardoya[1*]

[1]Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006, Madrid, Spain

[2]Department of Life Sciences, Natural History Museum, Cromwell Rd, London SW7 5BD, UK

[3]Muséum National d'Histoire Naturelle, Direction des Collections, 55, rue de Buffon, 75005 Paris, France

*Corresponding author: rafaz@mncn.csic.es

ABSTRACT

The subfamily Cantharidinae Gray, 1857 (Trochoidea: Trochidae) includes 23 recognized genera and over 200 known living species. These marine top shell snails are microphagous grazers that generally live in shallow rocky shores and in macroalgae and seagrass beds of sub-tropical and temperate waters from the Central and Western Indo-Pacific biogeographic regions to the Mediterranean Sea and the Eastern Atlantic Ocean. Recent molecular phylogenetic studies revising the family Trochidae supported the monophyly of the subfamily Cantharidinae and its sister group relationship to the subfamily Stomatellinae. These studies and others has thus far mostly focused on Indo-Pacific members of the subfamily Cantharidinae whereas here, we investigated phylogenetic relationships among their counterparts from the Mediterranean Sea and the North-eastern (NE) Atlantic Ocean including 33 species of genera *Gibbula, Jujubinus, Phorcus, Clelandella,* and *Callumbonella*. The Mediterranean and NE Atlantic taxa were supplemented with 30 Indo-Pacific Cantharidinae species plus 19 members of the sister group subfamily Stomatellinae. Phylogenetic trees were constructed using Bayesian inference and maximum likelihood with two datasets comprised of partial sequences of four or six mitochondrial (*cox1*, *rrnL, rrnS,* and *cob*) and nuclear (*28S rRNA* and *histone H3*) genes. A clade comprised of all Mediterranean and NE Atlantic taxa was recovered with high support, but its sister group among the Indo-Pacific lineages could not be determined with confidence (although the ascription of *Trochus kotschyi* to *Priotrochus* could be rejected). Within the Mediterranean and NE Atlantic clade, genera *Phorcus* and *Jujubinus* were recovered as reciprocally monophyletic, and the deep-sea genera *Clelandella* and *Callumbonella* were placed with high support as sister to *Jujubinus*. However, the genus *Gibbula* as

currently defined was not monophyletic and constituent species were divided into three major clades and two independent lineages. Phylogenetic relationships among *Phorcus, Jujubinus* (plus *Clelandella* and *Callumbonella*), and the different clades of *Gibbula* were poorly resolved in phylogenetic analyses based on four genes but received moderate support in those based on six genes. A first approach to resolve phylogenetic relationships within Stomatellinae was conducted showing that the diversity of the subfamily is highly underestimated at present, and that *Calliotrochus* is possibly a member of this subfamily. A chronogram was reconstructed using an uncorrelated relaxed lognormal molecular clock and the origin of the Mediterranean and NE Atlantic clade was dated right after the *Azolla* phase in the Middle Eocene about 48 million years ago whereas diversification of major clades (genera) followed the eastern closure of the Tethys Ocean in the Middle Miocene about 14 million years ago.

INTRODUCTION

With more than 600 species grouped in about 60 genera (Gofas, 2015), the family Trochidae Rafinesque, 1815 is one of the largest within the highly diverse superfamily Trochoidea (Gastropoda: Vetigastropoda) (Hickman and McLean, 1990; Williams et al., 2008, 2010; Williams, 2012). Members of the family are ecologically and economically important, and commonly known as top shells due to their characteristic spiral, conical, and internally nacreous shell. Top shells are unspecialized grazers that typically feed by rasping algae and small organisms from rocky or vegetal surfaces (Hawkins et al., 1989). Although cosmopolitan in distribution and adapted to different marine environments, they radiated preferentially on tropical intertidal rocky shores and of the West Pacific and Indian Oceans (Williams et al., 2010). The phylogenetic relationships of the family (and superfamily) have traditionally been highly contentious due to high levels of homoplasy in anatomical, radular and shell characters (Hickman and McLean, 1990), which is also evident in recent molecular phylogenies (Kano, 2008; Williams et al., 2008, 2010; Williams, 2012). In their seminal work on the systematics of trochoidean gastropods, Hickman and McLean (1990) assigned the different genera within Trochidae to 13 subfamilies (some of which were further subdivided into tribes) based on morphology. Some of the subfamilies were later raised to the familial status (Bouchet et al., 2005). However, recently reconstructed molecular trees have prompted for a thorough taxonomic revision of the family. Some taxa traditionally thought to belong in Trochidae were excluded and

provisionally placed in Turbinidae or Seguenzioidea (Williams and Ozawa, 2006; Kano, 2008; Williams et al., 2008), whereas some subfamilies were redefined (Williams et al., 2010). Currently, Trochidae is divided into ten subfamilies according to the new classification for Trochoidea proposed by Williams (2012): Trochinae Rafinesque, 1815; Umboniinae H. Adams & A. Adams, 1854; Stomatellinae Gray, 1840; Cantharidinae Gray 1857; Monodontinae Gray, 1857; Halistylinae Keen, 1958 (pending molecular analysis); Kaiparathininae B. A. Marshall, 1993; Fossarininae Bandel, 2009; Chrysostomatinae Williams, Donald, Spencer and Nakano, 2010, and Alcyninae Williams, Donald, Spencer and Nakano, 2010.

In this study, we focused on the subfamily Cantharidinae, which is known from the Upper Cretaceous (Hickman and McLean, 1990), and is widely distributed from the Central and Western Indo-Pacific regions to the Mediterranean Sea and the Eastern Atlantic Ocean (Williams et al., 2010). Members of Cantharidinae are found predominantly in sub-tropical or temperate waters, and are mostly microphagous grazers mainly associated with macroalgae and seagrass beds (Hickman and McLean, 1990; Hickman, 2005; Williams et al., 2010; Donald and Spencer, 2016). The taxonomic history of Cantharidinae has been particularly controversial. The group was considered a tribe (Cantharidini) within Trochinae by Hickman and McLean (1990), and was only recently raised to subfamilial rank (Williams et al., 2008), sister to the subfamily Stomatellinae (Herbert, 1998; Williams et al., 2008, 2010; Donald and Spencer, 2016). New definitions of Cantharidinae and

Monodontinae (previously ‘Gibbulini’) resulted in the mutual transfer of many taxa between both subfamilies (Donald et al., 2005; Williams et al., 2008, 2010; Donald and Spencer, 2016). For instance, the genera *Oxystele*, *Osilinus*, *Gibbula* and *Diloma* traditionally included in Monodontinae now belong to Cantharidinae (Herbert, 1998; Donald et al., 2005; Williams et al., 2010). The subfamily Cantharidinae currently includes 23 genera and more than 200 species (Bouchet and Gofas, 2015) but the taxonomy of the group at the generic level is in constant revision and redefinition with numerous instances of polyphyly and synonymizations (e.g. *Osilinus* and *Phorcus*; Donald et al., 2012) due to rampant levels of anatomical convergence and lack of reliable morphological data (Williams et al., 2010; Donald et al., 2016).

The most extensive phylogenetic analysis of Cantharidinae to date (Williams et al., 2010) included representatives of 15 genera but was mostly focused on Indo-Pacific species. In that study (Williams et al., 2010), concatenated partial sequences from four genes (28S rRNA, *rrnL*, *rrnS*, and *cox1*) were analyzed using Bayesian Inference as implemented in MrBayes but rendered low resolution at the base of the Cantharidinae clade. Instead, BEAST analysis with incorporation of an uncorrelated relaxed, lognormal clock resulted in moderate support for *Kanekotrochus* as sister to all remaining Cantharidinae (Williams et al., 2010). *Oxystele* was recovered as sister to a clade including the New Zealand genera *Cantharidus* and *Micrelenchus* (plus the genus *Prothalotia* and two unidentified species from Vanuatu and the Solomon islands) and a

group of seven Mediterranean and North-eastern (NE) Atlantic species, although with poor support (Williams et al., 2010). Recently, Donald and Spencer (2016) have presented an extensive molecular study of 18 New Zealand and temperate Australian Cantharidinae species based on partial sequence data of three mitochondrial genes (*rrnL*, *rrnS* and *cox1*) and one nuclear gene (28S rRNA). The genera *Cantharidus*, *Micrelenchus* and *Prothalotia* were redefined, the new genus *Roseaplagis* was erected, and the Australian genus *Cantharidella* was removed from the Cantharidinae to the subfamily Trochinae (Donald and Spencer, 2016). In the reconstructed phylogeny based on the concatenated data set, New Zealand endemics were not monophyletic as *Roseaplagis* was recovered in a clade together with *Prothalotia*, *Oxystele*, and Mediterranean and NE Atlantic species, although with low support (Donald and Spencer, 2016). Due to limited taxon sampling, phylogenetic relationships within the Mediterranean and NE Atlantic clade have not been rigorously assessed, and remain virtually unstudied using molecular methods.

Three genera are currently recognized in the Mediterranean Sea and NE Atlantic Ocean: *Gibbula*, which was tentatively shown to be polyphyletic (Williams et al., 2010) and includes around 55 species (according to Gofas and Bouchet, 2015a), most of them from the European coasts; *Jujubinus*, which comprises about 34 species (Gofas and Bouchet, 2015b), most of them from the Mediterranean Sea, and *Phorcus*, which has 9 species (Donald et al., 2012), all of them from European and Macaronesian coasts. In addition, the deep-sea genus *Clelandella* was placed close to genus

Jujubinus by Cretella et al. (1990). This genus was considered monotypic until Gofas (2005) described four new species from different Macaronesian seamounts and Vilvens et al. (2011) described another one from off Western Sahara.

Top shells of the genus *Phorcus* are microherbivorous grazers that live in intertidal rocky shores, whereas species of the genus *Jujubinus* inhabit subtidal bottoms down to about 80 m, mainly associated with marine vegetation, where they feed on the periphyton. In contrast, species of the genus *Gibbula* are widespread, occurring in a variety of habitats, from intertidal and subtidal rocky shores to marine vegetation and detritic sublittoral bottoms, and form a guild of mainly microphagous herbivores (Templado, 2011). Finally, *Clelandella* species are adapted to life in deep detritic bottoms.

Here, we extend previous phylogenetic works on Cantharidinae (Williams et al., 2010; Donald and Spencer, 2016) by focusing on Mediterranean and NE Atlantic genera. We amplified and sequenced fragments of four mitochondrial (c*ox1*, *rrnL*, *rrnS*, and *cob*) and two nuclear (*28S rRNA* and *histone H3*) genes for 47 specimens representing 29 species (16 of them type species of their respective genera). We included in the phylogenetic analyses the deep-water genus *Callumbonella,* of uncertain taxonomic placement, in order to determine its evolutionary origin within the subfamily. The newly determined sequences were aligned to orthologous sequences available in GenBank from 13 previously studied specimens corresponding to 12 Mediterranean and NE

Atlantic species. We also included in the phylogenetic analyses, newly determined and previously published (Williams et al., 2010; Donald et al., 2012)

orthologous sequences from 14 and 23 specimens, respectively, corresponding to 30 Indo-Pacific Cantharidinae species. We aimed to test the monophyly of each of the studied genera, and whether the origin of the Mediterranean and NE Atlantic fauna was connected with the eastern closure of the Tethys Sea in the Late Miocene (14 million years ago, MYA).

In order to properly reconstruct phylogenetic relationships within subfamily Cantharidinae, we also included in the phylogenetic analyses an important representation of species from the subfamily Stomatellinae since most recent phylogenies suggest that this subfamily is the closest living sister group of Cantharidinae (Williams et al., 2008; Williams et al., 2010; Williams, 2012; Donald and Spencer, 2016). The Stomatellinae are distributed in the Indian and Pacific Oceans, in both temperate and tropical waters, and typically adapted to intertidal and shallow water hard substrata. Their shells are auriform, oblong and small (<40mm), and the foot has a highly extensible metapodium that cannot be withdrawn into the shell. At present, five genera with 36 species are recognized (Gofas, 2009), although both numbers may be underestimated since the group has been barely studied (Williams et al., 2010; Herbert, 2015).

MATERIALS AND METHODS

Samples and DNA extraction

A full list of species from the subfamilies Cantharidinae and Stomatellinae analyzed in this study along with sampling localities, voucher specimens, and GenBank accession numbers is provided in Table 1. Some specimens belonged to unknown genera and species (see Supplementary Material 1 for photos). Photos of most of the specimens used in Williams et al. (2010) are publicly available on MorphoBank at http://morphobank.geongrid.org/permalink/?P223. Type species of the following genera (original name in brackets; Gofas, 2015) were included: *Agagus* Jousseaume, 1894 (Type: *A. agagus* Jousseaume, 1894); *Calliotrochus* Fischer, 1879 (Type: *Margarita marmorea* Pease, 1861); *Callumbonella Thiele, 1924* (Type: *Trochus suturalis* Philippi, 1836); *Cantharidus* Montfort, 1810 (Type: *Limax opalus* Martyn, 1784); *Clelandella* Winckworth, 1932 (Type: *Trochus miliaris* Brocchi, 1814); *Gibbula* Risso, 1826 (Type: *Trochus magus* Linnaeus, 1758); *Jujubinus* Monterosato, 1884 (Type: *Trochus exasperatus* Pennant, 1777); *Kanekotrochus* Habe, 1958 (Type: *Zizyphinus infuscatus* Gould, 1861); *Oxystele* Philippi, 1847 (Type: *Trochus sinensis* Gmelin, 1791); *Phorcus* Risso, 1826 (Type: *Monodonta richardi* Payraudeau, 1826); *Pictodiloma* Habe, 1946 (Type: *Trochus suavis* Philippi, 1849); *Priotrochus* Fischer, 1879 (Type: *Trochus obscurus* Wood, 1828); *Roseaplagis* Donald & Spencer, 2016 (Type: *Cantharidis* [sic.] *rufozona* A. Adams); *Stomatia* Helbling, 1779 (Type: *Stomatia phymotis* Helbling, 1779);

Stomatolina Iredale, 1927 (Type: *Stomatolina rufescens* Gray, J.E., 1847); and *Thalotia* Gray, 1847 (Type: *Monodonta conica* Gray, 1827). Tissues were stored in 100% ethanol and total genomic DNA was isolated from up to 30-60 mg of foot following a standard phenol-chloroform extraction (Sambrook et al., 1989).

PCR amplification and sequencing

We carried out standard PCR reactions using universal primers to amplify partial regions of mitochondrial genes c*ox1* (Folmer et al., 1994), *rrnL* (Palumbi et al., 1991), *rrnS* (Oliverio and Mariottini, 2001), *cob* (F-151 5'-GTGGRGCNACYGTWATYACTAA-3' and R-270 5'-AANAGGAARTAYCAYTCNGGYTG-3' modified from Merritt et al. 1998) as well as two nuclear genes, *histone H3* (H3MF 5'-ATGGCTCGTACCAAGCAGACTGG-3 and H3MR 5'-TGGATGTCCTTGGGCATGATTGTTAC-3' modified from Colgan et al. 1998), and *28S rRNA* (see Williams et al., 2010 and references therein) (Table 1). Each PCR reaction contained 1x buffer, 1.5 mM $MgCL_2$, 0.05 mM dNTPs, 0.2 mM each primer, 20-100 ng of template DNA, 1 unit of Taq DNA polymerase 5PRIME (Hamburg, Germany), and sterilized distilled water to 25 µl. The following cycling scheme was applied: a denaturing step at 94 ºC for 60 s; 45 cycles of denaturation at 94 ºC for 30 s, annealing at different temperatures depending on the gene (44ºC for *cox1*, 50ºC for *rrnL*, 62ºC for *rrnS*, 45ºC for *cob*, 50ºC for *H3*, and 52ºC for *28S*) for 80 s, and extension at 72 ºC for 90 s; followed by a final extension step at 72 ºC for 5 min. The amplified fragments were

purified and sequenced using PCR primers with Sanger sequencing at Secugen, Spain.

Table 1. Species analyzed in this study. Specimens, locality and voucher are shown below. Type species in bold. Additionally, the expeditions, GenBank accesion numbers and availability of photos are shown in Supporting Information 3 (see Appendix).

Clade Species	Locality	Voucher No
Cantharidinae		
Gibbula umbilicaris (Linnaeus, 1758) 1	El Mohon, Murcia, SE Spain	MNCN:ADN:86692
Gibbula umbilicaris 2	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86693
Gibbula cineraria (Linnaeus, 1758)	A Guarda, Pontevedra, NW Spain	MNCN:ADN:86694
Gibbula pennanti (Philippi, 1846) 1	Roscoff, W France	NHMUK 20080944
Gibbula pennanti 2	A Guarda, Pontevedra, NW Spain	MNCN:ADN:86695
Gibbula racketti (Payraudeau, 1826) 1	Islas Chafarinas, S Spain	MNCN:ADN:86696
Gibbula racketti 2	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86697
Gibbula albida (Gmelin, 1791)	O Grove, Pontevedra, NW Spain	MNCN:ADN:86698
Gibbula adriatica (Philippi, 1844)	O Grove, Pontevedra, NW Spain	MNCN:ADN:86699
Gibbula varia (Linnaeus, 1758) 1	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86700
Gibbula varia 2	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86701
Gibbula varia 3	Islas Chafarinas, S Spain	MNCN:ADN:86702
Gibbula divaricata (Linnaeus, 1758)	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86703
Gibbula rarilineata (Michaud, 1829) 1	Dugi otok, Telašćica, Croatia	MNCN:ADN:86704
Gibbula rarilineata 2	Lumbarda, Korcula I., Croatia	NHMUK 20080375
Gibbula umbilicalis (da Costa, 1778) 1	O Grove, Pontevedra, NW Spain	MNCN:ADN:86705
Gibbula umbilicalis 2	A Guarda, Pontevedra, NW Spain	MNCN:ADN:86706
Gibbula umbilicalis 3	Wembury, Plymouth, UK	NHMUK 20080946
Gibbula umbilicalis 4	O Grove, Pontevedra, NW Spain	MNCN:ADN:86707
Phorcus richardi (Payraudeau, 1826) 1	Almuñecar, Granada, S Spain	See Donald et al. 2012
Phorcus richardi 2	Cala Rajá, Cabo de Gata, SE Spain	MNCN:ADN:86708
Phorcus articulatus (Lamarck, 1822) 1	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86709
Phorcus articulatus 2	Islas Chafarinas, S Spain	MNCN:ADN:86710
Phorcus mutabilis (Philippi, 1846)	Fano, Italy	See Donald et al. 2012
Phorcus atratus (Wood, 1828)	La Golfa, Lanzarote, Canary Islands	See Donald et al. 2012
Phorcus lineatus (da Costa, 1778) 1	A Guarda, Pontevedra, NW Spain	MNCN:ADN:86711
Phorcus lineatus 2	Pembrokeshire, Wales, UK	See Donald et al. 2012
Phorcus punctulatus (Lamarck, 1822)	Dakar, Senegal	See Donald et al. 2012
Phorcus mariae Templado & Rolán, 2012	Mordeira Bay, Sal, Cape Verde Islands	See Donald et al. 2012
Phorcus sauciatus (Koch, 1845) 1	Ría de Aldán, Pontevedra, NW Spain	MNCN:ADN:86712
Phorcus sauciatus 2	Funchal, Madeira	MNCN:ADN:86713
Phorcus sauciatus 3	Cabo Estai, Canido, Pontevedra, NW Spain	MNCN:ADN:86714
Phorcus turbinatus (Born, 1778) 1	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86715
Phorcus turbinatus 2	Coral Bay, Cyprus	See Donald et al. 2012
Jujubinus exasperatus (Pennant, 1777) 1	Cala San Esteve, Menorca, E Spain	MNCN:ADN:86716
Jujubinus exasperatus 2	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86717
Jujubinus exasperatus 3	Lumbarda, Korcula I., Croatia	NHMUK 20080387
Jujubinus exasperatus 4	Cabo Estai, Canido, Pontevedra, NW Spain	MNCN:ADN:86718
Jujubinus exasperatus 5	Ile Callot, France	NHMUK 20080943
Jujubinus pseudogravinae Nordsieck, 1973	San Miguel, Azores	MNCN:ADN:86719
Jujubinus vexationis Curini-Galletti, 1990	Funchal, Madeira	MNCN:ADN:86720
Jujubinus ruscurianus (Weinkauff, 1868) 1	Banalmádena, Málaga, S Spain	MNCN:ADN:86721
Jujubinus ruscurianus 2	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86722
Jujubinus gravinae (Dautzenberg, 1881)	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86723
Gibbula tingitana Pallary, 1901	Calahonda, Málaga, S Spain	MNCN:ADN:86724
Jujubinus striatus (Linnaeus, 1758) 1	Menorca, E Spain	MNCN:ADN:86725
Jujubinus striatus 2	Cabo Estai, Canido, Pontevedra, NW Spain	MNCN:ADN:86726
Jujubinus striatus 3	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86727
Callumbonella suturalis (Philippi, 1836)	La Herradura Seamount, Alboran Sea, SE Spain	MNCN:ADN:86728
Clelandella miliaris (Brocchi, 1814)	off Vélez-Málaga, Alborán Sea, SE Spain	MNCN:ADN:86729
Gibbula magus (Linnaeus, 1758) 1	Cabo Estai, Canido, Pontevedra, Spain	MNCN:ADN:86730
Gibbula magus 2	Cabo Pino, Málaga, Spain	MNCN:ADN:86731
Gibbula magus 3	Faro, Portugal	NHMUK 20030329
Gibbula fanulum (Gmelin, 1791)	Korcula I., Croatia	NHMUK 20080378
Gibbula ardens (Salis Marschlins, 1793) 1	Menorca, Spain	MNCN:ADN:86732
Gibbula ardens 2	Cala San Esteve, Menorca, E Spain	MNCN:ADN:86733
Gibbula ardens 3	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86734
Gibbula philberti (Récluz, 1843) 1	El Mohon, Murcia, SE Spain	MNCN:ADN:86735
Gibbula philberti 2	Banalmádena, Málaga, S Spain	MNCN:ADN:86736
Gibbula turbinoides (Deshayes, 1835)	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86737

Table 1 (cont.)

Cantharidus opalus (Martyn, 1784)	Stewart I., New Zealand	NMNZ M.287117
Cantharidus capillaceus (Philippi, 1849)	Enderby Island, New Zealand	—
Cantharidus dilatatus (G. B. Sowerby II, 1870)	Wellington, New Zealand	—
Cantharidus antipoda Hombront & Jackinot, 1854	Stewart Island, New Zealand	NMNZ M.287119
Micrelenchus tesselatus (A. Adams, 1853)	Warrington, New Zealand	NMNZ M.288193
Micrelenchus huttoni (E.A. Smith, 1876)	Purakaunui Bay, New Zealand	NMNZ M.288196
Micrelenchus sanguineus (Gray, 1843)	Warrington, New Zealand	NMNZ M.287121
Micrelenchus purpureus (Gmelin, 1791)	Bay of Islands, New Zealand	NMNZ M.287118
Roseaplagis rufozonus (A. Adams, 1853)	Tauranga, New Zealand	NMNZ M.284054
Prothalotia lehmanni (Menke, 1843)	Dunsborough, Australia	NHMUK 20070156
Cantharidinae unknown genus 1 sp. 1	Sta Isabel I., Solomon Islands	MNHN IM-2007-18160
"Cantharidus" sendersi Poppe, Tagaro & Dekker, 2006	Tutuba I., Vanuatu	MNHN IM-2007-18458
"Jujubinus" suarezensis (P. Fischer, 1878) 1	Sainte Luce, Madagascar	MNHN IM-2009-13642
"Jujubinus" suarezensis 2	Rocher de l'Albatros, Madagascar	MNHN IM-2009-13612
"Jujubinus" suarezensis 3	Ilha dos Portugueses, Maputo Bay, Mozambique	MNHN IM-2009-23230
"Jujubinus" suarezensis 4	Ilha dos Portugueses, Maputo Bay, Mozambique	MNHN IM-2009-23229
"Jujubinus" suarezensis 5	Xixuane, Maputo Bay, Mozambique	MNHN IM-2009-23217
"Jujubinus" suarezensis 6	Ponta Punduine, Maputo Bay, Mozambique	MNHN IM-2009-23257
Kanekotrochus infuscatus (Gould, 1861)	Katakuchihama, Kagoshima, Japan	NSMT Mo76821
Agagus agagus Jousseaume, 1894	Aliwal Shoal, KwaZulu-Natal, South Africa	W2743
Priotrochus kotschyi (Philippi, 1849)	Abu Dhabi, United Arab Emirates	OKOT.ABD.1
Thalotia conica (Gray, 1827)	Pagoda Point, Princess Royal Harbour, Australia	MNHN IM-2009-23270
"Jujubinus" gilberti (Montrouzier in Fischer, 1878) 1	Napaling, Panglao I., Philippines	MNHN IM-2007-18363
"Jujubinus" gilberti 2	Baclayon Takot, Bohol I., Philippines	MNHN IM-2007-18375
Cantharidinae unknown genus 2 sp. 1	Napaling, Panglao I., Philippines	MNHN IM-2007-18362
"Cantharidus" callichroa (Philippi, 1849)	Omaezaki, Shizuoka, Japan	NHMUK 20050419
"Cantharidus" jessoensis (Scherenck, 1863)	Otuchi, Iwate, Japan	NSMT Mo76819
"Jujubinus" geographicus Poppe, Tagaro & Dekker, 2006	E Aoré I., Aimbuei Bay, Vanuatu	MNHN IM-2007-18457
Tosatrochus attenuatus (Jonas, 1844)	Marukihama, Bounotsu, Kagoshima, Japan	—
Pictodiloma suavis (Phillippi, 1850)	Susaki, Ogasawara I., Japan	NSMT Mo76820
"Cantharidus" lepidus (Philippi, 1846)	Duke of Orleans Bay, Esperance, Australia	NHMUK 20080950
Priotrochus obscurus (W. Wood, 1828) 1	Ponta Punduine, Maputo Bay, Mozambique	MNHN IM-2009-23222
***Priotrochus obscurus* 2**	Ponta Punduine, Maputo Bay, Mozambique	MNHN IM-2009-23224
Oxystele impervia (Menke, 1843)	False Bay, South Africa	NMNZ M.287122
Oxystele variegata (Anton, 1838)	False Bay, South Africa	NMNZ M.287126
Oxystele tabularis (Krauss, 1848)	East London, South Africa	NMNZ M.287124
Oxystele sinensis (Gmelin, 1791)	East London, South Africa	NMNZ M.287123
Stomatellinae		
Stomatella planulata (Lamarck, 1816)	Teniya, Nago City, Okinawa, Japan	NHMUK 20080949
Stomatella sp. 1	Danao, Panglao I., Philippines	MNHN IM-2007-18370
Stomatella sp. 2	Napaling, Panglao I., Philippines	MNHN IM-2007-18365
Stomatella sp. 3	Baclayon Takot, Bohol I., Philippines	MNHN IM-2007-18355
Stomatella impertusa (Burrow, 1815) 1	Aqaba, Jordan	MNCN:ADN:86738
Stomatella impertusa 2	Esperance, Australia	NHMUK 20080948
Stomatella impertusa 3	Bruat Channel, Vanuatu	MNHN IM-2007-18154
Stomatolina rubra (Lamarck, 1822) 1	Napaling, Panglao I., Philippines	MNHN IM-2007-18366
Stomatolina rubra 2	Pontod Lagoon, Panglao I., Philippines	MNHN IM-2007-18156
Stomatolina rubra 3	Pontod Lagoon, Panglao I., Philippines	MNHN IM-2007-18359
Stomatolina rubra 4	Pontod Lagoon, Panglao I., Philippines	MNHN IM-2007-18360
Stomatolina angulata (A. Adams, 1850)	Marukihama, Bounotsu, Kagoshima, Japan	—
Pseudostomatella decolorata (Gould, 1848)	Benoki, Okinawa, Japan	NSMT Mo76824
Stomatellinae unknown genus 1 sp. 1	Pamilacan I., Philippines	MNHN IM-2007-18371
Stomatellinae unknown genus 1 sp. 2	Cortes Takot, Bohol I., Philippines	MNHN IM-2007-18372
Stomatellinae unknown genus 2 sp. 1	Baclayon Takot, Bohol I., Philippines	MNHN IM-2007-18356
Microtis sp. 1	Palikulo Bay, Vanuatu	MNHN IM-2007-18449
Microtis tuberculata H. & A. Adams, 1850	Napaling, Panglao I., Philippines	MNHN IM-2007-18364
Microtis? sp. 2	Santo, Malparavu I., Vanuatu	MNHN IM-2007-18463
Stomatia obscura Sowerby G.B. III, 1874	Marukihama, Bounotsu, Kagoshima, Japan	—
Stomatia phymotis Helbling, 1779	SE Malparavu Island, Vanuatu	MNHN IM-2007-18155
Calliotrochus marmoreus (Pease, 1861)	Pamilacan I., Philippines	MNHN IM-2007-18352
Umboniinae		
Isanda coronata A. Adams, 1854	Karratha Back Beach, Australia	NHMUK 20050421
Umbonium costatum (Kiener, 1839)	Pusan, South Korea	NHMUK 20050422

*indv2: the corresponding sequence comes from specimen 2
*SEQ=GenBank accesion Nos. available in for the final version

Editing and alignment of sequences

Sequences were edited with Sequencher 5.0.1 sequence analysis software (Gene Codes Corporation, Ann Arbor, MI, USA). The program TranslatorX (Abascal et al., 2010) was used to produce the alignments of each protein coding gene guided by the corresponding deduced open reading frames. The rRNA genes were aligned separately using MAFFT v. 7 (Katoh and Standley, 2013) with default parameters. No gaps were found in the protein-coding genes. Ambiguously aligned sites in the rRNA genes were removed with Gblocks, v.0.19b (Castresana, 2000) in the Phylogeny.fr server (Dereeper et al., 2008) with the following settings: minimum sequence for flanking positions: 85%; maximum contiguous non-conserved positions: 8; minimum block length: 10; gaps in final blocks: no. Trees were reconstructed for each of the individual genes to check for potential PCR contaminations, numts (nuclear pseudogenes), or taxon misidentifications. Phylogenetic analyses were based on two multi-locus datasets; one constructed by concatenating four genes (*rrnS, rrnL, cox1, and 28S rRNA*; see Figs 1, 2A and 3A) and a second with two extra genes (*cob* and histone *H3*). The 6-gene dataset included a reduced taxon set with representative species of the two less-studied groups, Stomatellinae, and Mediterranean and NE Atlantic Cantharidinae (see Figs. 2B and 3B).

Phylogenetic analyses

Phylogenetic relationships of Cantharidinae and Stomatellinae were reconstructed using maximum likelihood (ML; Felsenstein, 1981) and Bayesian inference (BI; Huelsenbeck and Ronquist, 2001). ML analyses were conducted with RAxML v7.3.1 (Stamatakis, 2006) using the rapid hill-climbing algorithm and 10,000 bootstrap pseudoreplicates (BP). Bootstrap values >70%, between 70-50%, and <50% were considered to indicate high, moderate and low statistical support, respectively (Hillis and Bull, 1993). BI analyses were conducted using MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003) and run with two independent analyses, each with four simultaneous Monte Carlo Markov chains (MCMC) run for 10 million generations. Trees were sampled every 1000 generations after a 25% burn-in when the analysis had reached stationarity (as judged by plots of ML scores and low SD of split frequencies). Statistical support of internal nodes was assessed using Bayesian posterior probabilities (BPP) with values >0.95, between 0.95-0.90, and <0.90 considered to indicate high, moderate and low statistical support, respectively.

The best partition schemes and best-fit models of substitution for the different analyzed data sets were identified using Partition Finder (Lanfear et al., 2012) according to the Bayesian Information Criterion (BIC; Schwarz, 1978); see Supplementary Material 2).

Estimation of divergence times

We used BEAST v.1.7 (Drummond and Rambaut, 2007) to perform a Bayesian estimation of divergence times among major Cantharidinae lineages based on the four-gene data set. This software was used to infer branch lengths and nodal ages on a fixed topology (the reconstructed ML tree topology in Fig. 1). Only one individual per species was included in the analysis, except in the case of *Gibbula racketti*, *Jujubinus exasperatus* and *Jujubinus striatus* that had two because these taxa showed internal highly divergent lineages that may represent cryptic species. For the clock model, we selected the lognormal relaxed-clock model, which allows rates to vary among branches without any *a priori* assumption of autocorrelation between adjacent branches. For the tree prior, we employed a Yule process of speciation. We employed the partitions and models selected by Partition Finder (see above). The final Markov chain was run twice, each for 50 million generations, sampling every 1,000 generations with the first 5 million discarded as the burn-in, after confirming convergence of chains with Tracer v.1.5. (Rambaut and Drummond, 2007). The effective sample size of all the parameters was > 200.

Estimated divergence times were obtained by specifying one calibration point based on fossil data using a lognormal distribution of prior probability. The origin of the New Zealand genus *Cantharidus* was constrained to have a minimum age corresponding to the Upper Eocene (35 MYA; 95% interval: 35.3–47.9 MYA) following Donald et al. (2012).

RESULTS

Our taxon sampling strategy was designed to maximize representation of Cantharidinae species from the Mediterranean Sea and NE Atlantic Ocean, which have been lacking in previous studies (Table 1). A total of 25 specimens of *Gibbula*, 11 of *Jujubinus*, eight of *Phorcus*, one of *Clelandella*, and one *Callumbonella* representing 29 different species (including the type species of these five genera) from these regions were sampled and analyzed in combination with 14 new specimens of Indo-Pacific Cantharidinae and 14 new specimens of Stomatellinae (Table 1). Altogether 346 new sequences were determined for four-six partial genes (*cox1* with ~600 bp; *rrnL* with ~550 bp; *rrnS* with ~630 bp; *cob* with ~340 bp; *28S rRNA* with ~1400 bp; *histone H3* with ~310 bp). Each of these genes was analyzed separately producing phylogenetic trees that essentially recover similar and congruent clades to those obtained in the phylogenetic analyses based on multi-locus matrices, but with poorer support (Supplementary Material 3).

Phylogenetic relationships of Cantharidinae

The newly determined sequences were combined with orthologous sequences available from GenBank for subfamilies Cantharidinae (36 specimens) and Stomatellinae (8 specimens), and phylogenetic analyses were based on four-gene (all Cantharidinae) and six-gene data sets (focused on Mediterranean and NE Atlantic Cantharidinae

and on Stomatellinae) using probabilistic methods of inference (Figs. 1-3). The nucleotide alignment of the four-gene data set (*rrnS, rrnL*, *cox1*, and *28S*) was 2,792 positions long. The ML (-lnL = 36651.18) and BI phylogenetic analyses arrived at the same topology using two species of the subfamily Umboniinae as outgroup (Fig. 1). Both reconstructed trees recovered with high statistical support the monophyly of the subfamilies Stomatellinae (100% BP, 1 BPP) and Cantharidinae (94% BP, 1 BPP), as well as of a Mediterranean and NE Atlantic Cantharidinae group (91% BP, 1 BPP) (Fig. 1). Statistical support at the genus level was generally high, but other internal nodes in the four-gene trees received only moderate to poor support (Fig. 1).

Within analyzed Cantharidinae, the genus *Oxystele* is recovered as sister to the remaining members of the subfamily, although with low support (Fig. 1). This genus is restricted to the South African region, and was represented in our analyses by four species (Table 1). Within the remaining Cantharidinae, four major clades and several independent lineages are recovered. Of these groups, only the one including the Mediterranean and NE Atlantic Ocean genera is highly supported (Fig. 1). One of the Indo-Pacific clades has high support (76% BP/ 1 BPP) and includes "*Jujubinus*" *geographicus* from Vanuatu as sister to "*Cantharidus*" from Japan and "*Jujubinus" gilberti* and Cantharidinae genus 2 from the Philippines (Fig. 1). Phylogenetic relationships within this clade are well resolved with nodes having high statistical support (Fig. 1).

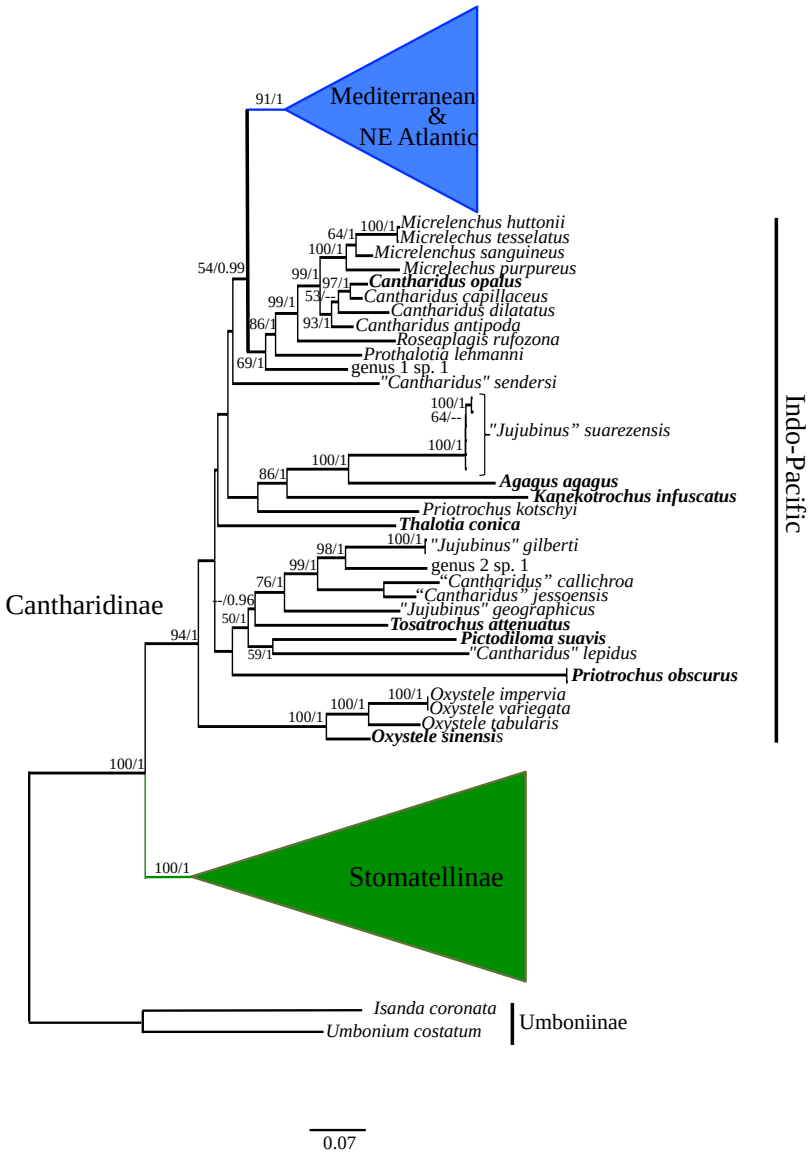


Figure 1. Phylogenetic relationships of Stomatellinae and Cantharidinae. The reconstructed ML tree based on the four-gene data set is shown. Numbers at nodes are statistical support values for ML (bootstrap proportions)/ BI (posterior probabilities). Only phylogenetic relationships among Indo-Pacific Cantharidinae are shown in full. The Mediterranean and NE Atlantic Cantharidinae and Stomatellinae clades are collapsed and shown separately in full in Figs. 2A and 3A, respectively. Type species are shown in bold.

A second highly supported Indo-Pacific group (86% BP/ 1 BPP) includes the genera *Kanekotrochus* from Japan and *Agagus* from South East Africa, as well as "*Jujubinus*" *suarezensis* from Mozambique and Madagascar. A third group is moderately well-supported in ML (69% BP) and highly supported in BI (1 BPP), and includes New Zealand species of the genera *Micrelenchus*, *Cantharidus* and *Roseaplagis* (clade with high statistical support, 99% BP/ 1 BPP), the genus *Prothalotia* from the west coast of Australia, and Cantharidinae genus 1 from the Solomon Islands (Fig. 1). The relative phylogenetic position of *Priotrochus obscurus* from Mozambique, genera *Pictodiloma* and *Tosatrochus* both from Japan, *"Cantharidus"* and *Thalotia* both from Western Australia, *Priotrochus kotschyi* from Abu Dhabi, and "*Cantharidus*" *sendersi* from Vanuatu could not be confidently resolved. Of the three major Indo-Pacific clades, the one including Cantharidinae genus 1 from Solomon Islands, Australian *Prothalotia*, and *Micrelenchus* + *Cantharidus* + *Roseaplagis* from New Zealand, was recovered as sister group to the Mediterranean and NE Atlantic clade, with moderate and high statistical support in ML (54% BP) and BI (0.99 BPP), respectively (Fig. 1). It is noteworthy that *Priotrochus obscurus* (type species of the genus) and *Priotrochus kotschyi* were recovered in different clades.

Phylogenetic relationships of Mediterranean and NE Atlantic Cantharidinae

Cantharidinae species living in the Mediterranean Sea and NE Atlantic Ocean form a highly supported (91% BP/ 1 BPP) monophyletic group in the phylogenetic analyses based on the four-gene data set (Fig. 1). Within this group, seven main lineages are recovered. The monophyly of both *Phorcus* and *Jujubinus* is maximally supported, as was a clade with the deep-sea genera *Clelandella* and *Callumbonella* sister to *Jujubinus* (99% BP/ 1BPP; Fig. 2A). Phylogenetic relationships within *Phorcus* are well resolved with *P. turbinatus* and *P. sauciatus* recovered as sister group to two distinct clades, one including *P. mariae*, *P. punctulatus*, and *P. lineatus*, and the other including *P. atratus*, *P. mutabilis*, *P. articulatus* and *P. richardi* (Fig. 2A). Within *Jujubinus*, phylogenetic relationships are relatively well resolved. Three individuals identified as *J. striatus* but showing high sequence divergences form a clade that is the sister to the remaining *Jujubinus* and *Gibbula tingitana*. Within this clade, two lineages are distinguished, one including *Gibbula tingitana* from Alborán sea (south Spain), *J. gravinae* and *J. ruscurianus*, and the other having Macaronesian endemics as sister to *J. exasperatus*, which itself consist of two divergent lineages (Fig. 2A). However, the most striking result of the phylogenetic analysis is the cryptic diversity and non-monophyly of *Gibbula*, which shows at least three distinct and highly supported clades (Fig. 2A). One clade includes *G. umbilicaris*, *G. cineraria*, *G. pennanti*, *G. racketti*, *G. albida*, and *G. adriatica*. Phylogenetic relationships within this clade are

unresolved (Fig. 2A). A second, closely related clade includes *G. umbilicalis* as sister to *G. rarilineata* plus *G. varia*, and *G. divaricata*. A third clade includes *G. ardens* as sister group of *G. magus* and *G. fanulum*. Two additional *Gibbula* species, *G. philberti* and *G. turbinoides* show relatively long branches and were recovered together as sister group of *Gibbula* clade 3 without support (Fig. 2A). Finally, as mentioned above, *G. tingitana* is nested deeply within *Jujubinus* in all phylogenetic analyses (Fig. 2A).

The nucleotide alignment of the six-gene (adding *cob* and *histone H3*) Mediterranean and NE Atlantic Cantharidinae data set was 3,507 positions long. The ML (-lnL = 18645.94) and BI phylogenetic analyses arrived at the same topology using *Agagus agagus* and "*Cantharidus*" *sendersi* as outgroups (Fig. 2B). The reconstructed trees confirmed the monophyly of *Jujubinus*, *Phorcus*, and of the three independent lineages of *Gibbula*, as well as the sister group relationship of *Clelandella* and *Callumbonella* to *Jujubinus* (Fig. 2B), providing higher statistical support for the corresponding tree nodes. Moreover, the addition of extra positions to the phylogenetic analyses produced higher resolution at deeper nodes connecting main lineages (Fig. 2B). A sister group relationship between *Gibbula* clades 1 and 2 was highly supported

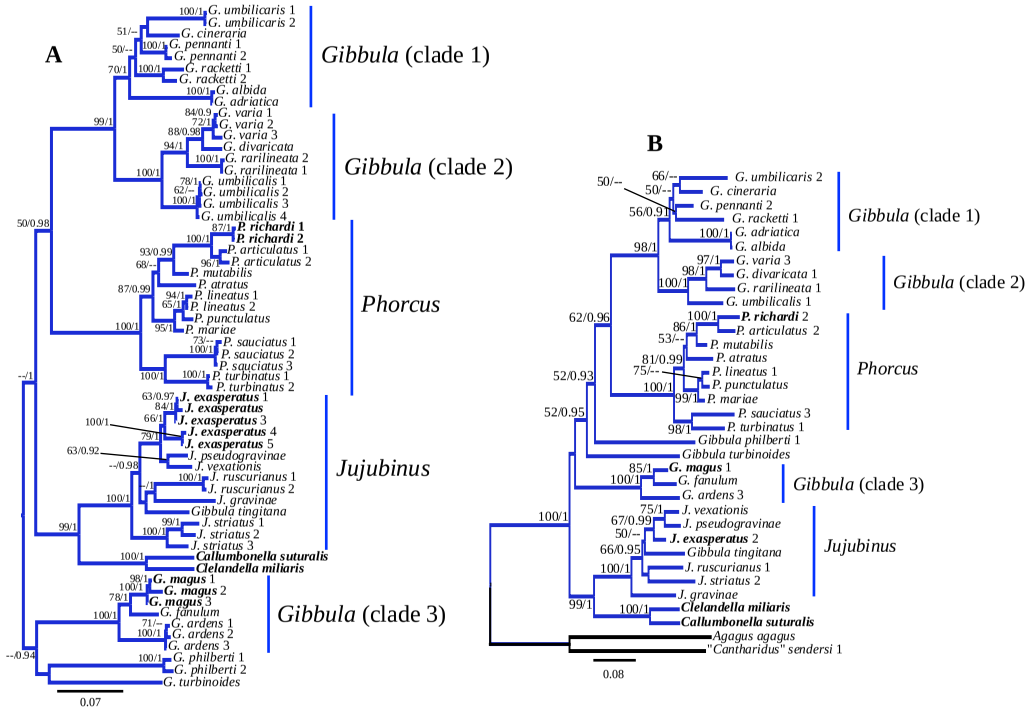


Figure 2. Phylogenetic relationships of Mediterranean and NE Atlantic Cantharidinae. The reconstructed ML trees based on the four-gene (A) and six-gene (B) data sets are shown. Numbers at nodes are statistical support values for ML (bootstrap proportions)/ BI (posterior probabilities). Type species are shown in bold.

(98% BP, 1 BPP), and this clade was sister to *Phorcus* with moderate support (62% BP, 0.96 BPP). This clade had *G. philberti* as sister to the exclusion of *G. turbinoides* with poor to moderate support (52% BP/ 0.93 BPP; Fig. 2B). The relative phylogenetic positions of *Gibbula* clade 3 and the clade with *Jujubinus* plus *Clelandella* and *Callumbonella* could not be resolved (Fig. 2B).

Phylogenetic relationships of Stomatellinae

The Indo-Pacific subfamily Stomatellinae has been poorly studied to date. Here, despite the fact that it was not the main focus of this study, we have significantly increased the number of species analyzed for this subfamily. In addition to the four-gene data set (see reconstructed trees in Fig. 3A) a six-gene (adding *cob* and *H3*) data set was also constructed for Stomatellinae. The nucleotide alignment of the six-gene Stomatellinae data set was 3,625 bp. The ML (-lnL = 13338.99) and BI phylogenetic analyses arrived at the same topology using *Agagus agagus* and "*Cantharidus*" *sendersi* as outgroups (Fig. 3B). Most internal nodes within Stomatellinae lacked resolution in both the four- and six-gene data sets (Fig. 3). However, phylogenetic analyses allowed delimitation of several independent lineages that indicate unrecognized diversity within the subfamily. One clade included species of the genera *Stomatia* and *Microtis* (Fig. 3). A second clade included four known and three unidentified species of the genus *Stomatella*. Individuals assigned to the species *S. impertusa* did not form a monophyletic group (Fig. 3). A third clade recovered the species *Pseudostomatella decolorata* as

sister to *Stomatolina rubra* with maximal support (Fig. 3). The relative phylogenetic position of *Stomatolina angulata* and three unidentified species (Stomatellinae unknown genus 1 sp 1and sp 2, and Stomatellinae unknown genus 2 sp 1) from the Philippines could not be confidently resolved. Lastly, *Calliotrochus marmoreus* was recovered as sister group of the Stomatellinae.

Estimation of divergence times

In order to date the major cladogenetic events within the Cantharidinae, a chronogram was estimated using a Bayesian uncorrelated relaxed molecular clock. One fossil calibration was used to constrain the origin of the New Zealand endemic clade to the Upper Eocene about 35 MYA (Donald et al., 2012). The resulting chronogram dated the split between the two sister subfamilies Stomatellinae and Cantharidinae about 77.5 MYA, although with a large confidence interval (106-56, 95%HPD; Fig. 4). The origin of the Mediterranean and NE Atlantic clade was estimated to have occurred 47 MYA (58-39, 95%HPD), whereas its diversification into major lineages (*Phorcus*, *Jujubinus*, and the different *Gibbula* clades) was dated between 38-28 MYA (Fig. 4) coinciding with the early diversification of the New Zealand endemic Cantharidinae. Within the Mediterranean NE Atlantic clade, most of the main diversification events within each genus started about 14 MYA (as in the Indo-Pacific Ocean) and a few more recently, about 5 MYA (Fig. 4).

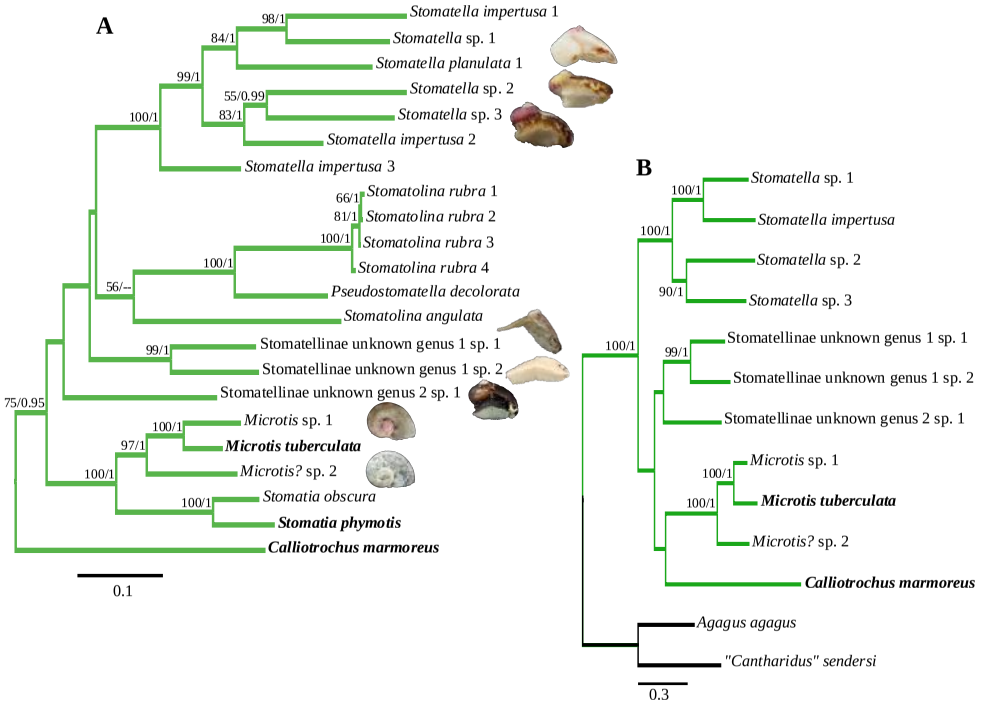


Figure 3. Phylogenetic relationships within Stomatellinae. The reconstructed ML trees based on the four-gene (A) and six-gene (B) data sets are shown. Numbers at nodes are statistical support values for ML (bootstrap proportions)/ BI (posterior probabilities). Photos of unidentified taxa are shown. Type species are shown in bold.

DISCUSSION

Phylogenetic relationships of NE Atlantic and Mediterranean Cantharidinae

In this study, we provide the first attempt to establish a robust phylogeny of Mediterranean and NE Atlantic Cantharidinae. All these taxa were recovered in a single, well-supported clade containing several well-supported subclades. Unfortunately, phylogenetic relationships among groups within the Mediterranean and NE Atlantic Cantharidinae clade lack resolution, possibly pointing to a rapid radiation event soon after the origin of these lineages. Reciprocal monophyly of both *Phorcus* and *Jujubinus* is strongly supported, but *Gibbula* as currently defined is not monophyletic. Interestingly, a sister group relationship between the deep-sea genera *Clelandella* and *Callumbonella* is recovered, and their sister relationship with *Jujubinus* is highly supported. The species *Clelandella miliaris* was originally included in the genus *Calliostoma*, and later ascribed to the genus *Jujubinus*. More recently, Cretella et al. (1990) proposed the genus *Clelandella* was valid, and phylogenetically closer to *Jujubinus* than to *Calliostoma* based on external head-foot and radular characters and allozyme data. The genus *Callumbonella* remained unassigned or variously assigned to either Umboniinae or Margaritinae (Rolán et al., 2009). Further work on deep sea Trochidae is on-going and will help determine the exact phylogenetic position of these enigmatic taxa (STW in prep).

The alpha-taxonomy of the genus *Jujubinus* and the exact distribution of many of its species remain unresolved despite considerable revisionary efforts (Mariottini et al., 2013; Smriglio et al., 2014; Oliverio pers. comm.). The reconstructed phylogenetic relationships within *Jujubinus* are relatively well resolved, and show a strong geographical component. The species from Madeira (*J. vexationis*) and Azores Islands (*J. pseudogravinae*) form a clade that is sister to the widespread *J. exasperatus*. This aptly named species is divided into two biogeographically distinct lineages: one found in the Mediterranean and one in the NE Atlantic. The divergence between these two lineages as well as that found among individuals assigned to *J. striatus* are similar to the one obtained between the Macaronesian endemic species, suggesting cryptic diversity within *J. exasperatus* and *J. striatus*. Hence, our results prompt for a revision of *Jujubinus* species with more thorough population sampling (currently being undertaken by M. Oliverio, pers. comm). One clear taxonomic outcome of the present phylogeny derives from the relative position of *Gibbula tingitana* from Alborán Sea, (south Spain), which is nested deeply within *Jujubinus* in all phylogenetic analyses next to species from the same region, *J. rucurianus* and *J. gravinae*. Thus, we propose officially reassigning this species to *Jujubinus*.

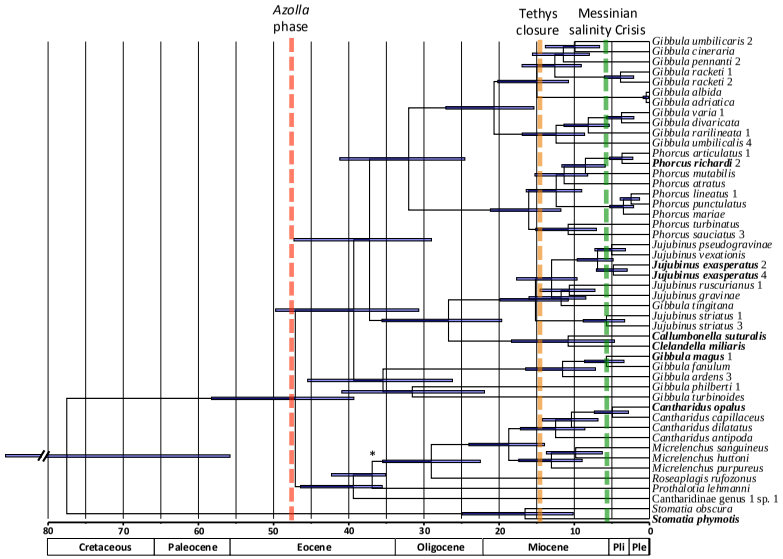


Figure 4. Chronogram with age estimates of major divergence events among Mediterranean and North-eastern Atlantic Cantharidinae based on the four-gene data set. A Bayesian uncorrelated relaxed lognormal clock was used in BEAST. Horizontal bars represent 95% credibility intervals for time estimates, and the calibration constraint is indicated with an asterisk on the corresponding node. Dates (and credibility intervals) are in millions of years. Paleogene and Neogene periods are shown. The *Azolla* phase, the eastern closure of Tethys Ocean, and the Messinian Salinity Crisis of the Mediterranean Sea are indicated with red, orange and green dashes lines, respectively. Type species are shown in bold.

We have included in our study all nine species currently recognized within genus *Phorcus* (Donald et al., 2012) and recovered two main lineages within the genus, which correspond to the subgenera *Phorcus* (*P. richardi*, *P. articulatus*, *P. mutabilis*, *P. atratus*, *P. lineatus*, *P. punctulatus*, and *P. mariae*) and *Osilinus* (*P. sauciatus* and *P. turbinatus*) (Donald et al., 2012). Within the subgenus *Phorcus*, the Atlantic species *P. punctulatus*, *P. mariae*, and *P. lineatus* were the sister to the Canary Island endemic and the Mediterranean species. Within the subgenus *Osilinus*, *P. sauciatus* is an Atlantic species and *P. turbinatus* is endemic to the Mediterranean Sea. Almost the same topology (varying only in the relative position of *P. mariae* as sister to *P. lineatus*) was recovered by Donald et al. (2012), who provided the first complete molecular phylogeny of *Phorcus* and clarified the evolution of this conspicuous group of intertidal grazers.

The reconstructed phylogeny divided species ascribed to the genus *Gibbula* into three major clades, with the exception of *G. turbinoides* and *G. philberti*, whose placement is uncertain. The relative position of the genus *Phorcus* (strongly supported in the six-gene phylogenetic analysis) makes *Gibbula* non-monophyletic. The clades obtained here largely coincide with those of a previous phylogeny of the genus based solely on *cox1* (Barco et al., 2013). All three *Gibbula* clades contain both Mediterranean and NE Atlantic species. In one clade, two eastern Mediterranean species, *G. albida* and *G. adriatica* are the sister group of a mix of Atlantic (*G. cineraria* and *G. pennanti*) and Mediterranean representatives (*G. umbilicaris, G. racketti*). According to Barco et al. (2013) the

species *G. spratti, G. nivosa, G. adansonii* and *G. nebulosa*, all from the Mediterranean Sea, could also belong to this clade. The high level of divergence found between two individuals assigned to *G. racketti* may indicate the presence of cryptic species and deserves further study. A second clade embraces the Atlantic *G. umbilicalis* and the Mediterranean species *G. rarilineata, G. varia,* and *G. divaricata.* Hence, the internal division of the clade follows the geographical distribution of taxa. This clade was also obtained by Barco et al. (2013), although with *G. rarilineata* as sister to *G. umbilicalis* plus *G. varia,* and *G. divaricata*. The Mediterranean species *G. rarilineata* and *G. divaricata* are known to form mixed populations in the same habitat, and morphologically intermediate forms can be found (López-Márquez, 2016).

Phylogenetic relationships among *Jujubinus*, *Phorcus* and the *Gibbula* clades are unresolved in phylogenetic trees reconstructed using the four-gene data set but some deeper nodes obtain statistical support in the six-gene phylogeny. In particular, the six-gene tree recovers high support for the sister group relationship of *Phorcus* and *Gibbula* clades 1 and 2 to the exclusion of *Gibbula* clade 3, *G. philberti*, and *G. turbinoides*, resulting in non-monophyly of the genus *Gibbula* as currently defined. Moreover, the mean sequence divergence among *Gibbula* clades is of the same magnitude as that estimated between *Phorcus* and *Jujubinus*, indicating that the three *Gibbula* clades plus the two independent species deserve treatment as separate genera. The clade 3 containing the type species, *G. magus,* plus *G. ardens* and *G. fanulus* should retain the genus name (*G. guttadauri*, not studied here, also belong to this clade according

to the results of Barco et al., 2013), whereas the remaining *Gibbula* clades could receive distinct generic names, as has been previously noted by Williams et al. (2010) and by Barco et al. (2013) employing *cox1* sequences. In fact, some junior synonyms are available, such as e.g., *Steromphalus* P. Fischer, 1875 (Type species *Trochus cinerarius* Linnaeus, 1758) for *Gibbula* clade 1, and *Gibbulastra* Monterosato, 1884 (Type species *Trochus divaricatus* Linnaeus, 1758) for *Gibbula* clade 2. However, we leave nomenclatural decisions for future studies with such specific aim and based on a more complete taxon sampling of European, Macaronesian and South African species.

Furthermore, some South African and Indo-Pacific species currently assigned to *Gibbula* and *Jujubinus* almost certainly belong to other genera. As a precedent, Hickman and McLean (1990) and Herbert (1991) assigned some western Indian Ocean species previously referred to *Gibbula* to the genera *Agagus* and *Calliotrochus*, and Beck (1995) based on morphological and radular characters erected the new genus *Rubritrochus* (which belongs to the subfamily Trochinae) for two species previously included in *Gibbula* from the Indian Ocean and Red Sea (*G. pulcherrima* and *G. declivis*). Similarly, Donald et al. (2012) showed that *Osilinus kotschyi*, from the Arabian and Red Seas, was not closely related to the other species of this NE Atlantic genus and tentatively referred it to the Indo-Pacific genus *Priotrochus* (see below). In particular, West African species currently attributed to *Gibbula* and *Jujubinus* should be studied molecularly to accurately determine the extent of both genera.

Phylogenetic relationships of Cantharidinae

The South African genus *Oxystele* was recovered as sister to the remaining members of the subfamily Cantharidinae here analyzed, but only with poor support. Previous phylogenetic analyses based on concatenated genes were also rather inconclusive for this part of the tree, and recovered either *Kanekotrochus* (Williams et al., 2010) or *Kanekotrochus* + *C. jessoensis* (Donald and Spencer, 2016) as sister to the remaining Cantharidinae with *Oxystele* in a more derived position. The different lineages within Cantharidinae are connected by short branches, suggesting a rapid radiation process soon after the origin of the subfamily. Genomic sequence data sets including relevant missing taxa may be needed to enhance resolution at these and other (e.g. sister group of Mediterranean and NE Atlantic taxa; see below) nodes in the phylogeny. A robust phylogeny is needed to shed light on the rather complex taxonomy of Indo-Pacific Cantharidinae. Most tropical species from this subfamily have been assigned to a specific genus based on shell or morphological characters, but the group requires extensive revision (Herbert, 1996; Vilvens, 2012), and many species will shift from one genus to another when robust phylogenies become available. For instance, our results further support the suggestion that the genus *Cantharidus* (the type species, *C. opalus*, is included) should be restricted to New Zealand species, whereas species from other areas traditionally included in this genus should be assigned to other genera (Williams et al., 2010). In this regard, our results mostly

matched those obtained by Donald and Spencer (2016), who recognized three New Zealand endemic genera *Cantharidus* s.s., *Micrelenchus* and *Roseaplagis*. The nomenclatural changes proposed by Donald and Spencer (2016) are accepted and followed here, and in fact, in our case the monophyly of New Zealand endemic taxa is highly supported.

A remarkable case is that of *Trochus kotschyi* from the Arabian and Red Seas, whose taxonomic placement has been controversial historically, as it was variously ascribed to genera *Cantharidus*, *Thalotia*, *Priotrochus* and *Gibbula* (Herbert, 1994). The species was once considered synonymous to *Priotrochus obscurus* (the type species of the genus) but Herbert (1988) confirmed its validity as endemic species from the Persian Gulf. Later, the same author (Herbert, 1994) included *T. kotschyi* in the European genus *Osilinus*. Nevertheless Donald et al. (2012) showed that this species is not sister to the European clade of *Osilinus* species, and suggested its tentative placement in *Priotrochus*. Our present results confirm that *T. kotschyi* is neither related to the NE Atlantic and Mediterranean clade nor to the clade that includes the type species of *Priotrochus*. Therefore, the generic assignment of this species remains unresolved.

Due to lack of statistical support, it is not possible to determine with confidence, the exact sister to the Mediterranean and NE Atlantic Ocean genera. However, our results agree with previous phylogenetic analyses (Williams et al., 2010) in considering that a putative clade including New Zealand

Cantharidus and related endemic genera (*Micrelenchus* and *Roseaplagis*; Donald and Spencer, 2016), *Prothalotia* from the west coast of Australia, and an undescribed species from the Solomon Islands could be the sister group to a clade containing Mediterranean and NE Atlantic genera. Donald and Spencer (2016) also add *Oxystele* to the list of potential genera with affinities to Mediterranean and NE Atlantic Cantharidinae. If further supported, this phylogenetic relationship between antipode taxa would be favor an ancient Tethyan origin for the Mediterranean and NE Atlantic Cantharidinae (see below).

Phylogenetic relationships of Stomatellinae

The subfamily Stomatellinae has been long recognized as a natural group (Gray, 1840), but its phylogenetic placement was controversial until recently, when a robust sister group relationship with Cantharidinae was proposed based on morphology (Herbert, 1998), and further supported with molecules (Williams et al., 2010). The taxonomy of Stomatellinae has been neglected for many years and urgently demands monographic revisionary studies since nominal species were often inadequately described and type material has rarely been consulted by subsequent authors (Herbert, 2015). Therefore, it is not surprising that the phylogenetic relationships within Stomatellinae are largely unknown. Here, we provide an exploratory attempt to further resolve phylogenetic relationship within this subfamily, extending previous work (Williams et al., 2010) by adding new taxa, although several are

pending formal identification or description due to the problems mentioned above (Supplementary Material 1). The reconstructed phylogeny is congruent with that of Williams et al. (2010), and shows that some of the newly added taxa likely represent new genera, and that some species like *Stomatella impertusa* may have served as a catch-all taxon. Surprisingly, our results could prove that *Calliotrochus* is not a member of the Cantharidinae (cf. WoRMS; Bouchet and Gofas, 2010), but rather a basal and somewhat atypical (i.e., turbiniform rather than auriform) member of the Stomatellinae as suspected by Herbert (1998). Altogether, the phylogeny suggests that the diversity of the subfamily is highly underestimated at present.

Estimation of divergence times

The present-day distribution of the subfamily Cantharidinae extends from the Western Indo-Pacific region to the NE Atlantic-Mediterranean region and the coasts of Western Africa down to South Africa. This distribution strongly suggests a Tethyan origin for the group. The Tethys Ocean separated Laurasia and Gondwana supercontinents since the break-up of Pangea in the Triassic about 250 MYA. Continental drift produced collision of these two tectonic plates about 19 MYA and the final closure of the Eastern Tethys Seaway in the Middle Miocene around 14 MYA (Harzhauser et al., 2007; Okay et al., 2010). The closure definitively ended the previous exchange of water and tropical marine biota between the Indian Ocean and the proto-Mediterranean Basin

(Hamon et al., 2013). It is striking that the sister group of the Mediterranean and NE Atlantic Cantharidinae lineage may comprise antipode temperate genera from Australia and New Zealand, which would support an original widespread distribution in the Tethys Ocean, and likely extinction events in tropical intermediate regions. However, this hypothesis requires the postulated sister group relationship receiving higher statistical support in future studies.

According to our results, the origin of the Mediterranean and NE Atlantic lineage occurred in the boundary between the Early and Middle Eocene around 47 MYA (about 51 MYA in Donald et al. (2012), which greatly predates the closure of the Eastern Tethys seaway. Many Cantharidinae species are associated to seagrass beds and interestingly, the marine angiosperm genus *Posidonia* shows a similar geographic distribution with antipode (Australian and Mediterranean) clades. The divergence between these *Posidonia* clades is ancient, and was estimated based on ITS (internal transcribed spacer) sequence data to have occurred about 68 Myr (Aires et al., 2011), coinciding with the Cretaceous–Tertiary mass extinction (K–T boundary) event (Schulte et al., 2010). Other authors (Phillips and Meñez, 1988) estimated that the ancient *Posidonia* split could have occurred in the Late Eocene based on geological and tectonic history. This dating would be more in agreement with this study. While the Early Eocene was characterized by a warmer climate and sea level rise (which peaked in the so-called Eocene Climatic Optimum; 52 to 50 MYA), at the onset of Middle Eocene about 48 MYA, the *Azolla* phase started the

transition from greenhouse conditions towards modern icehouse conditions (Speelman et al., 2009). The *Azolla* phase is named for massive blooms of this floating aquatic fern that sank in the Arctic Ocean at this time, producing a drop in atmospheric carbon dioxide. The consequent steady cooling of the climate (a 17-My-long trend towards cooler conditions that persisted until the latter part of the Oligocene about 26 to 27 MYA, Zachos et al., 2001) over this time period is likely to have favored temperate taxa.

The origin of the major Mediterranean and NE Atlantic lineages (*Jujubinus*, *Phorcus*, and the different clades of *Gibbula*) was dated in this study to have occurred towards the end of the Eocene. The traffic of nearshore gastropods across the Eastern Tethys Seaway continued throughout the Oligocene until the Upper Miocene (23 MYA), when it dropped to zero despite the passage remaining open for sometime longer (Harzhauser et al., 2009). In fact, a re-establishment of migration pathways occurred about 16 MYA prior to the final closure (Harzhauser et al., 2009). Our estimates for the main diversification events within the different Cantharidinae lineages were dated to have occurred after the final closure of the Eastern Tethys Seaway due to the consequent isolation of the Mediterranean and Atlantic Ocean basin from the Indo-Pacific region. The Messinian Salinity Crisis (5.9-5.3 MYA) presumably led to important background extinction rates in the Mediterranean Sea, which would have affected subtidal taxa such as almost all *Jujubinus* species and many *Gibbula* species and less to *Phorcus* species and some *Gibbula* species that are well adapted to the intertidal zone and thus are likely capable of surviving stressful

conditions. The possibility of surviving the Messinian Salinity Crisis within the Mediterranean Sea has also been suggested for another intertidal group, the vermetid genus *Dendropoma* (Calvo et al., 2015). With the reopening of the Strait of Gibraltar (5.3 Ma), the refilling of the Mediterranean Sea by the Atlantic Ocean (the Zanclean flood; Garcia-Castellanos et al., 2009) allowed splitting among closely related species and population differentiation within species as suggested by our chronogram.

To conclude, this study builds on previous molecular phylogenetic studies of the hyperdiverse family Trochidae. It focuses on the subfamily Cantharidinae and in particular, Mediterranean and NE Atlantic species. The results of the study further corroborate the rampant levels of morphological homoplasy in top shell snails that have hindered their systematics, and are exemplified here by the non-monophyly of the genus *Gibbula*, which can be divided into at least three major clades. Moreover, it emphasizes the existence of unrecognized diversity both within the subfamily Cantharidinae and its sister subfamily Stomatellinae at the genus and species level, and the need for monographic works to thoroughly revise the taxonomy of both subfamilies. Future molecular phylogenies of the subfamily should add taxa from West Africa and use both mitochondrial and nuclear genomic data sets, which have proven to render highly resolved phylogenetic relationships within Vetigastropoda (Uribe et al., 2016). Additionally, the study highlights interesting biogeographical patterns as it supports a single origin of the Mediterranean and NE Atlantic clade from Indo-Pacific lineages, and shows important

divergences within each genus between Mediterranean and NE Atlantic species. The dating of the phylogeny using a fossil-calibrated molecular clock suggests that the *Azolla* phase, the eastern closure of the Tethys Ocean, and the Messinian Salinity Crisis of the Mediterranean Sea had major effects on the diversification of the group.

ACKNOWLEDGEMENTS

We are grateful to Serge Gofas for donating species from south Spain, and Emilio Rolán, who provided specimens from northwest Spain. The MNHN Indo-Pacific material originates from a series of expeditions and workshops, conducted in the context of the "Our Planet Reviewed" programme with Pro-Natura International (SANTO 2006 to Vanuatu, ATIMO VATAE 2010 to Madagascar "Deep South" with Université de Tuléar, and its spin-off INHACA 2011 at Estação de Biologia Maritima, Universidade Eduardo Mondlane), of the "Tropical Deep-Sea Benthos" programme with Institut de Recherche pour le Développement (SALOMON 2 to the Solomon Is), or stand-alone projects (PANGLAO 2004 to the Philippines, with University San Carlos; WESTERN AUSTRALIA 2011, with the Western Australian Museum). We thank the Total Foundation Total, Prince Albert II of Monaco Foundation, and Stavros Niarchos Foundation for their support to these expeditions. Jose Rosado and Hugh Morrison facilitated the field logistics in Mozambique and Australia, respectively. Alexander Fedosov, Yuri Kantor, Nicolas Puillandre and Ellen Strong processed the specimens in the field for molecular sequencing, and we are grateful to Philippe Bouchet for giving us access to this collection. All expeditions operated under the regulations then in force in the

countries in question and satisfy the conditions set by the Nagoya Protocol for access to genetic resources. We particularly thank Raquel Álvarez for her technical assistance in the molecular laboratory at early stages of this work. We are indebted 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. D.G. Herbert and an anonymous referee provided very useful comments that greatly improved a previous version of the manuscript. 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 by funding from the NHM Department of Life Sciences to STW.

REFERENCES

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

Aires, T., Marbà, N., Cunha, R.L., Kendrick, G.A., Walker, D.I., Serrão, E.A., Duarte, C.M., Arnaud-Haond, S., 2011. Evolutionary history of the seagrass genus *Posidonia*. Mar. Ecol. Prog. Ser. 421, 117-130.

Barco, A., Evans, J., Schembri, P.J., Taviani, M., Oliverio, M., 2013. Testing the applicability of DNA barcoding for Mediterranean species of top-shells (Gastropoda, Trochidae, *Gibbula* s.l.). Mar. Biol. Res. 9, 785-793.

Beck, L., 1995. *Rubritrochus*, a new genus name for *Gibbula pulcherrima* A. Adams 1855 and *Gibbula declivis* (Forskal 1775) (Gastropoda: Prosobranchia: Trochidae). Arch. Molluskenkd. 124, 65-85.

Bergsten, J., 2005. A review of long‐branch attraction. Cladistics 21, 163-193.

Bouchet, P.; Gofas, S. (2010). Cantharidinae Gray, 1857. In: MolluscaBase (2016). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p=taxdetails&id=382171 on 2016-06-02

Bouchet, P., Gofas, S., 2015. Cantharidinae Gray, 1857. In: MolluscaBase (2015). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p=taxdetails&id=382171 on 2016-04-30

Bouchet, P., Rocroi, J.P., Frýda, J., Hausdorf, B., Ponder , W.F., Valdés, Á., Warén, A. 2005. Classification and nomenclator of gastropod families. Malacologia 47, 1-397.

Calvo, M., Alda, F., Oliverio, M., Templado, J., Machordom, A., 2015. Surviving the Messinian Salinity Crisis? Divergence patterns in the genus *Dendropoma* (Gastropoda: Vermetidae) in the Mediterranean Sea. Mol. Phylogenet. Evol. 91, 17-26.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540-552.

Colgan, D.J., McLauchlan, A., Wilson, G.D.F., Livingston, S.P., Edgecombe, G.D., Macaranas, J., Cassis, G., Gray, M.R., 1998. Molecular phylogenetics of the Arthropoda: relationships based on histone H3 and U2 snRNA DNA sequences. Aust. J. Zool. 46, 419-437.

Cretella, M., Scillitani, G., Picariello, O., 1990. The systematic position of '*Trochus*' *miliaris* Brocchi, 1814 (Gastropoda: Trochidae); morphological and biological evidences. Lavori della Societa´ Italiana di Malacologia 23, 51–81.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M., Gascuel, O., 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36, W465-W469.

Donald, K.M., Kennedy, M., Spencer, H.G., 2005. The phylogeny and taxonomy of austral monodontine topshells (Mollusca: Gastropoda: Trochidae), inferred from DNA sequences. Mol. Phylogenet. Evol. 37, 474-483.

Donald, K.M., Preston, J., Williams, S.T., Reid, D.G., Winter, D., Alvarez, R., Buge, B., Hawkins, S.J., Templado, J., Spencer, H.G., 2012. Phylogenetic relationships elucidate colonization patterns in the intertidal grazers *Osilinus* Philippi, 1847 and *Phorcus* Risso, 1826 (Gastropoda: Trochidae) in the northeastern Atlantic Ocean and Mediterranean Sea. Mol. Phylogenet. Evol. 62, 35-45.

Donald, K.M., Spencer, H.G., 2016. Phylogeographic patterns in New Zealand and temperate Australian cantharidines (Mollusca: Gastropoda: Trochidae: Cantharidinae): Trans-Tasman divergences are ancient. Mol. Phylogenet. Evol. 100, 333-344.

Drummond, A., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 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. Mol. Mar. Biol. Biotechnol. 3, 294-299.

Garcia-Castellanos, D., Estrada, F., Jiménez-Munt, I., Gorini, C., Fernàndez, M., Vergés, J., De Vicente, R., 2009. Catastrophic flood of the Mediterranean after the Messinian salinity crisis. Nature 462, 778-781.

Gofas, S., 2005. Geographical Differentiation In *Clelandella* (Gastropoda: Trochidae) In The Northeastern Atlantic. J. Mollus. Stud. 71, 133-144.

Gofas, S., 2009. Stomatellinae Gray, 1840. In: MolluscaBase (2016). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p=taxdetails&id=382176 on 2016-05-16

Gofas, S., 2015. Trochidae. In: MolluscaBase (2015). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p=taxdetails&id=443 on 2016-04-30

Gofas, S., Bouchet, P., 2015a. *Gibbula* Risso, 1826. In: MolluscaBase (2016). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p

=taxdetails&id=138590 on 2016-05-16

Gofas, S., Bouchet, P., 2015b. Jujubinus Monterosato, 1884. In: MolluscaBase (2016). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p=taxdetails&id=138591 on 2016-05-16

Gray, J.E., 1840. Shells of molluscous animals. Synopsis of the contents of the British Museum. Woodfall and Son, London, pp. 106-156.

Hamon, N., Sepulchre, P., Lefebvre, V., Ramstein, G., 2013. The role of eastern Tethys seaway closure in the Middle Miocene Climatic Transition (ca. 14 Ma). Clim. Past 9, 2687-2702.

Harzhauser, M., Kroh, A., Mandic, O., Piller, W.E., Göhlich, U., Reuter, M., Berning, B., 2007. Biogeographic responses to geodynamics: A key study all around the Oligo–Miocene Tethyan Seaway. Zool. Anz. 246, 241-256.

Harzhauser, M., Reuter, M., Piller, W.E., Berning, B., Kroh, A., Mandic, O., 2009. Oligocene and Early Miocene gastropods from Kutch (NW India) document an early biogeographic switch from Western Tethys to Indo-Pacific. Paläontol. Z. 83, 333-372.

Hawkins, S.J., Watson, D.C., Hill, A.S., Harding, S.P., Kyriakides, M.A., Hutchinson, S., Norton, T.A., 1989. A comparison of feeding mechanisms in microphageous, herbivorous, intertidal prosobranchs in relation to resource partitioning. J. Mollus. Stud. 55, 151-165.

Herbert, D.G., 1988. Studies of *Priotrochus obscurus* and the systematic position of

Priotrochus (Mollusca: Gastropoda: Trochidae). J. Zool. Lond. 214, 261-268.

Herbert, D.G., 1991. A revision of the genus *Agagus* Jousseaume, 1894 (Mollusca: Gastropoda: Trochidae). J. Nat. Hist. 25, 883-900.

Herbert, D.G., 1994. *Trochus kotschyi*, the first Indian Ocean record of the genus

Osilinus (Mollusca: Gastropoda: Trochidae). J. Zool. Lond. 233, 345-357.

Herbert, D.G., 1996. A critical review of the trochoidean types in the Muséum d'Histoire Naturelle, Bordeaux (Mollusca: Gastropoda). Bull. Mus. Natl. Hist. Nat. Series 4, 18, Section A, 3-4, 409-445.

Herbert, D.G., 1998. Revision of the genus *Calliotrochus* Fischer, 1879 (Gastropoda : Trochoidea). Invertebr. Syst. 12, 545-565.

Herbert, D.G., 2015. An annotated catalogue and bibliography of the taxonomy, synonymy and distribution of the Recent Vetigastropoda of South Africa (Mollusca). Zootaxa 4049, 1-98.

Hickman, C.S., 2005. Seagrass fauna of the temperate southern coast of Australia I: the cantharidine trochid gastropods. In: Wells, F.E., Walker, D.I., Kendrick, G.A. (Eds.), The Marine Flora and Fauna of Esperance, Western Australia. Western Australia Museum, Perth.

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

Hillis, D.M., and Bull, J. J., 1993. An empirical testof bootstrapping as a method fo assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182-192.

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. Zool. Scr. 37, 1-21.

Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. 30, 772-780.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. Mol. Biol. Evol. 29, 1695-1701.

López-Márquez, V., García-Jiménez, R., Templado, J., Machordom, A. , 2016. Development and characterization of 26 novel microsatellite loci for the Trochidae *Gibbula divaricata* (Linnaeus, 1758), using Illumina MiSeq next generation sequencing technology. Peer J. 4:e1789

Mariottini, P., Di Giulio, A., Appolloni, M., Smriglio, C., 2013. Phenotypic diversity, taxonomic remarks and updated distribution of the Mediterranean *Jujubinus baudoni* (Monterosato,1891) (Gastropoda, Trochidae). Biodivers. J. 4, 343-354.

Merritt, T.J., Shi, L., Chase, M.C., Rex, M.A., Etter, R.J., Quattro, J.M., 1998. Universal cytochrome b primers facilitate intraspecific studies in molluscan taxa. Mol. Mar. Biol. Biotechnol. 7, 7-11.

Okay, A.I., Zattin, M., Cavazza, W., 2010. Apatite fission-track data for the Miocene Arabia-Eurasia collision. Geology 38, 35-38.

Oliverio, M., Mariottini, P., 2001. A molecular framework for the phylogeny of *Coralliophila* and related muricoids. J. Mollus. Stud. 67, 215-224.

Palumbi, S., Martin A, McMillan, W., Stice, L., Grabowski, G., 1991. The simple fool's guide to PCR. Available from http://palumbi.stanford.edu/SimpleFoolsMaster.pdf, 1-45.

Phillips, R.C., Meñez, E.G., 1988. Seagrasses. Smithsonian Institution Press, Washington, DC.

Rambaut, A., Drummond, A.J., 2007. Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Rolán, E., Gonzalez-Porto, M., de Matos-Pita, S., 2009. The genus *Callumbonella* (Gastropoda, Trochacea), with description of a new species from Namibia. J. Conchol. 39, 634-657.

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

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning. A Laboratory Manual. 2nd. Edition. Cold Spring Harbor Laboratory, New York.

Schulte, P., Alegret, L., Arenillas, I., Arz, J.A., Barton, P.J., Bown, P.R., Bralower, T.J., Christeson, G.L., Claeys, P., Cockell, C.S., Collins, G.S., Deutsch, A., Goldin, T.J., Goto, K., Grajales-Nishimura, J.M., Grieve, R.A.F., Gulick, S.P.S., Johnson, K.R., Kiessling, W., Koeberl, C., Kring, D.A., MacLeod, K.G., Matsui, T., Melosh, J., Montanari, A., Morgan, J.V., Neal, C.R., Nichols, D.J., Norris, R.D., Pierazzo, E., Ravizza, G., Rebolledo-Vieyra, M., Reimold,

W.U., Robin, E., Salge, T., Speijer, R.P., Sweet, A.R., Urrutia-Fucugauchi, J., Vajda, V., Whalen, M.T., Willumsen, P.S., 2010. The Chicxulub Asteroid Impact and Mass Extinction at the Cretaceous-Paleogene Boundary. Science 327, 1214-1218.

Schwarz, G., 1978. Estimating the Dimension of a Model. Ann. Statist. 6, 461-464.

Smriglio, C., Di Giulio, A., Mariottini, P., 2014. Description of two new *Jujubinus* species (Gastropoda: Trochidae) from the Sicily Channel, with notes on the *Jujubinus curinii* species complex. Zootaxa 381, 583–590.

Speelman, E.N., Van Kempen, M.M.L., Barke, J., Brinkhuis, H., Reichart, G.J., Smolders, A.J.P., Roelofs, J.G.M., Sangiorgi, F., De Leeuw, J.W., Lotter, A.F., Sinninghe DamstÉ, J.S., 2009. The Eocene Arctic *Azolla* bloom: environmental conditions, productivity and carbon drawdown. Geobiology 7, 155-170.

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

Templado, J., 2011. Familia Trochidae. In: Gofas, S., Moreno, D., Salas, C. (eds.). Moluscos marinos de Andalucía, vol 1. Servicio de Publicaciones e Intercambio Científico, Universidad de Málaga. pp 104-117.

Uribe, J.E., Kano, Y., Templado, J., Zardoya, R., 2016. Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. Zool. Scr. 45, 145-159.

Vilvens, C., 2012. New species and new records of Seguenzioidea and Trochoidea (Gastropoda) from French Polynesia. Novapex 13, 1-23.

Vilvens, C., Swinnen, F., Deniz Guerra, F. 2011. A new species of *Clelandella* (Gastropoda: Trochoidea: trochidae: Cantharidinae) from Western Sahara. Novapex 12(1-2): 49-55.

Williams, S.T., 2012. Advances in molecular systematics of the vetigastropod superfamily Trochoidea. Zool. Scr. 41, 571-595.

Williams, S.T., Donald, K.M., Spencer, H.G., Nakano, T., 2010. Molecular systematics of the marine gastropod families Trochidae and Calliostomatidae (Mollusca: Superfamily Trochoidea). Mol. Phylogenet. Evol. 54, 783-809.

Williams, S.T., Karube, S., Ozawa, T., 2008. Molecular systematics of Vetigastropoda: Trochidae, turbinidae and trochoidea redefined. Zool. Scr. 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). Mol. Phylogenet. Evol. 39, 33-51.

Zachos, J., Pagani, M., Sloan, L., Thomas E., Billups, K. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292, 686-693.

SUPPORTING INFORMATION 3

Data S1. Species analyzed in this study

Data S2. Photos of unknown genera and species

Data S3. Best fit partitions and evolutionary substitution models as selected by Partition Finder

Data S4. Individual gene trees

3.4. PUBLICACION 4:

Titulo: "Phylogenetic relationships among superfamilies of Neritimorpha (Mollusca: Gastropoda)"
Autores: Juan Esteban Uribe, Don Colgan, Lyda R. Castro, Yasunori Kano y Rafael Zardoya
Estado: Aceptado
Año: 2016
Revista: *Molecular Phylogenetics and Evolution*

Resumen:

A pesar de la extraordinaria diversidad morfológica y ecológica de Neritimorpha, pocos estudios se han enfocado sobre las relaciones filogenéticas de este linaje de gasterópodos, que incluye cuatro superfamilias vivas: Neritopsoidea, Hydrocenoidea, Helicinoidea y Neritoidea. Aquí, se determinaron las secuencias de nucleótidos de los genomas mitocondriales completos de *Georissa bangueyensis* (Hydrocenoidea), *Neritina usnea* (Neritoidea), y *Pleuropoma jana* (Helicinoidea) y los genomas mt casi completos de *Titiscania* sp. (Neritopsoidea), y *Theodoxus fluviatilis* (Neritoidea). Las reconstrucciones filogenéticas usando métodos probabilísticos se basaron en sets de datos mitocondriales (13 genes codificantes de proteína y dos genes de ARN), nucleares (genes parciales de *28S* rRNA, *18S* rRNA, *actin*, e *histona H3*) y combinados. Todos los análisis filogenéticos excepto uno convergieron sobre un árbol con alto apoyo, en el que se recuperó Neritoidea como grupo hermano de un clado que incluye Helicinoidea como grupo hermano de Hydrocenoidea y Neritoidea. Esta topología concuerda con el registro fósil y apoya al menos tres invasiones independientes del medio terrestre por los caracoles neritimorfos. Los genomas mitocondriales de *Titiscania* sp., *G. bangueyensis*, *N. usnea*, and *T. fluviatilis* comparten el mismo ordenamiento descrito previamente para el genoma mt de *Nerita*, mientras que el de *P. jana* ha sufrido importantes reorganizaciónes**.** Se secuenció alrededor de medio genoma mitocondrial de otra especie de Helicinoidea, *Viana regina* y se confirmel reordenamiento altamente derivado de *P. Jana.*

Phylogenetic relationships among superfamilies of Neritimorpha (Mollusca: Gastropoda)

Juan E. Uribe[1], Don Colgan[2], Lyda R. Castro[3], Yasunori Kano[4], and Rafael Zardoya[1*]

[1]Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006, Madrid, Spain

[2]AMRI, The Australian Museum, 1 William Street, Sydney 2010, Australia

[3]Grupo de Evolución, Sistemática y Ecología Molecular, Universidad del Magdalena, Santa Marta, Colombia

[4]Atmosphere and Ocean Research Institute, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8564, Japan.

* Corresponding author (rafaz@mncn.csic.es)

ABSTRACT

Despite the extraordinary morphological and ecological diversity of Neritimorpha, few studies have focused on the phylogenetic relationships of this lineage of gastropods, which includes four extant superfamilies: Neritopsoidea, Hydrocenoidea, Helicinoidea, and Neritoidea. Here, the nucleotide sequences of the complete mitochondrial genomes of *Georissa bangueyensis* (Hydrocenoidea), *Neritina usnea* (Neritoidea), and *Pleuropoma jana* (Helicinoidea) and the nearly complete mt genomes of *Titiscania* sp. (Neritopsoidea), and *Theodoxus fluviatilis* (Neritoidea) were determined. Phylogenetic reconstruction using probabilistic methods were based on mitochondrial (13 protein coding genes and two ribosomal rRNA genes), nuclear (partial 28S rRNA, 18S rRNA, actin, and histone H3 genes) and combined sequence data sets. All phylogenetic analyses except one converged on a single, highly supported tree in which Neritopsoidea was recovered as the sister group of a clade including Helicinoidea as the sister group of Hydrocenoidea and Neritoidea. This topology agrees with the fossil record and supports at least three independent invasions of land by neritimorph snails. The mitochondrial genomes of *Titiscania* sp., *G. bangueyensis*, *N. usnea*, and *T. fluviatilis* share the same gene organization previously described for *Nerita* mt genomes whereas that of *P. jana* has undergone major rearrangements. We sequenced about half of the mitochondrial genome of another species of Helicinoidea, *Viana regina* and confirmed that this species shares the highly derived gene order of *P. jana*.

INTRODUCTION

With approximately 2,000 living species (Fukumori and Kano, 2014; Richling, 2014), the Neritimorpha—also known as Neritopsina—forms a long-recognized and distinct lineage of Gastropoda that appears to be the remnant of an early radiation (Ponder and Lindberg, 1997; Lindberg, 2008). The earliest undisputed fossils of Neritimorpha date back to the Carboniferous (Kaim and Sztajner, 2005). However, its origin could be much older if platyceratid snails from the Ordovician to Devonian periods (Bandel and Frýda, 1999) are interpreted as stem lineages of the group (Frýda et al., 2008, 2009). Despite its current relatively low species richness, the group is characterized by an astonishing diversity of morphologies (besides spiral or conical shells, these snails have also evolved limpet- and slug-like forms) and the adaptation to very different habitats, thus resembling at a smaller scale the extraordinary morphological and ecological diversity achieved by all gastropods (Kano et al., 2002; Lindberg, 2008). Of marine evolutionary origin, they are currently found in deep-sea hydrothermal vents and seeps, as well as in submarine caves, but have particularly radiated in the intertidal and shallow subtidal zones of tropical regions (Warén and Bouchet, 2001; Frey and Vermeij, 2008; Kano and Kase, 2008). Additionally, several groups have independently invaded brackish, surface or underground freshwater and terrestrial (including arboreal) habitats (Kano et al., 2002; Lindberg, 2008; Schilthuizen et al., 2012; Fukumori and Kano, 2014).

Bouchet et al. (2005) classified the living species of Neritimoprha into four superfamilies: Neritopsoidea, Hydrocenoidea, Helicinoidea, and Neritoidea. The superfamily Neritopsoidea includes the marine families Neritopsidae and Titiscaniidae. The extant Neritopsidae are represented exclusively by a few species in the type genus *Neritopsis* from submarine caves and similar cryptic voids, but this family presents the oldest fossil record among living neritimorph superfamilies (Kano et al., 2002; Kaim and Sztajner, 2005). The Titiscaniidae, also represented by a few species in the single genus *Titiscania*, have adapted to life in even smaller interstices by evolving a slug-like phenotype with complete loss of the shell. Because of their highly divergent morphology, titiscaniids were initially placed in their own superfamily Titiscanoidea (Bergh, 1890), but more recent studies suggest that they originated in the early Cenozoic as a specialized lineage within Neritopsidae or even within *Neritopsis* (Kano et al., 2002). The superfamily Hydrocenoidea has one family, Hydrocenidae, which includes two widely-accepted genera (*Hydrocena* and *Georissa*) with presumably over a hundred terrestrial species in forest litter, on limestone crops or deep inside karstic caves (Schilthuizen et al., 2012). The superfamily Helicinoidea includes three terrestrial families, namely Helicinidae, Proserpinidae and Proserpinellidae (Thompson, 1980; Richling, 2004, 2014), as well as the aquatic family Neritiliidae, whose members live in freshwater streams, underground waters (e.g., *Neritilia* and *Platynerita*; Kano and Kase, 2004) and subtidal marine caves (e.g., *Pisulina* and *Siaesella*; Kano and Kase, 2002,

2008). Finally, the superfamily Neritoidea includes the Phenacolepadidae and Neritidae. The former family is mainly composed of limpet-like species that inhabit tidal flats and mangrove swamps (e.g., *Phenacolepas* and *Plesiothyreus*; Fretter, 1984), shipworm tunnels in sunken wood (*Phenacolepas*; Kano et al., 2013), or deep-sea hydrothermal vent fields (*Shinkailepas* and *Olgasoralis*; Warén and Bouchet, 2001). A monotypic genus *Bathynerita* from methane seeps in the Gulf of Mexico also belongs to Phenacolepadidae, regardless of its retention or re-acquisition of a coiled shell (Holthuis, 1995; Fukumori and Kano, 2014). The family Neritidae (the nerites) has radiated into brackish and freshwater habitats, and includes also the genus *Neritodryas*, which is partially terrestrial or even arboreal although some species retain planktotrophic early development (Little, 1990; Kano and Kase, 2002). Several genera within Neritidae such as *Nerita*, *Clithon* and *Neritina* comprise numerous species (Holthuis, 1995; Frey and Vermeij, 2008).

Only a few studies have specifically addressed relationships within Neritimorpha (Holthuis, 1995; Kano et al., 2002). The most complete phylogenetic studies on the internal relationships of the group were based on 57 morphological characters (Holthuis, 1995) and the partial sequences of the 28S rRNA gene (Kano et al., 2002). The monophyly of Neritimorpha is widely accepted and supported by numerous morphological synapomorphies (Ponder and Lindberg, 1997). However, only characters related to the reproductive system might be very useful for determining the phylogenetic relationships among the superfamiles (Kano et al., 2002). The 28S rRNA

molecular phylogeny recovered Neritopsoidea as the sister group of the remaining three superfamilies. Among these, Neritoidea and Helicinoidea were sister groups to the exclusion of Hydrocenoidea (Kano et al., 2002).

Thus far, the mitochondrial (mt) genomes available for Neritimorpha, all belonging to the genus *Nerita*, are only nearly complete, lacking the putative control region (Castro and Colgan, 2010; Arquez et al., 2014). These mt genomes have been used as representatives of Neritimorpha to determine its relative phylogenetic position within Gastropoda, and have allowed the description of the mt genome order of *Nerita*, which is similar to the inferred ancestral gastropod genome organization (Osca et al., 2014a) and nearly identical (except on a tRNA gene transposition) to the order found in some non-gastropod molluscs such as the cephalopod *Octopus vulgaris* Cuvier, 1797 (Yokobori et al., 2004). In this study, we aimed to (1) sequence the complete or almost complete mt genomes of representatives of all superfamilies of Neritimorpha, (2) determine whether the mt genome organization described in *Nerita* is common to the other neritimorph lineages, (3) reconstruct the phylogenetic relationships of the living superfamilies, (4) compare the mt-based phylogeny with the nuclear-based phylogeny of Kano et al. (2002), and (5) to infer several evolutionary trends within the group, which may be at the origin of its extraordinary morphological and ecological diversity.

MATERIAL AND METHODS

Samples and DNA extraction

In order to have representation of the four neritimorph superfamilies, one specimen each of *Georissa bangueyensis* Smith, 1895 (Hydrocenoidea), *Neritina usnea* (Röding, 1798) and *Theodoxus fluviatilis* (Linnaeus, 1758) (Neritoidea), *Pleuropoma jana* (Cox, 1872) and *Viana regina* (Morelet, 1849) (Helicinoidea), and *Titiscania* sp. (Neritopsoidea) was used for this study (See Table 1, for details on the locality, collector, and voucher ID of each sample). Except *P. jana*, which was held at -80° C after collection, all samples were stored in 100% ethanol and total genomic DNA was isolated from up to 30-60 mg of foot tissue following a standard phenol/ chloroform extraction.

Table 1. Complete mitochondrial (mt) genomes analyzed in this study. Length in bp, Genbank accesion number, museum voucher, sampling location, habitat, and name of collector are provided.

New mt genomes

Species	Superfamily	bp	GenBank	Voucher	Location	Habitat	Collector
Georissa bangueyensis	Hydrocenoidea	15.267	KU342664	MNCN/ ADN:86335	Sabah, Sandakan, Malaysia	Terrestrial	Menno Schilthuizen
Neritina usnea	Neritoidea	15.574	KU342665	MNCN/ ADN:86336	Cienaga de Mallorquín, Colombia	Estuarine	Moises Arquez
Pleuropoma jana	Helicinoidea	15.851	KU342666	AM/ EBU:65394	Yessabah Caves, New South Wales, Australia	Terrestrial	Michael Shea, Don Colgan, Sven Schreiter
*Theodoxus fluviatilis**	Neritoidea	14.122	KU342667	MNCN/ ADN:86338	La Fuente del Mal Nombre, Malaga, España	Freshwater	José Templado
*Viana regina**	Helicinoidea	7.557	KU342668	MNCN/ ADN:86339	Valle de Viñales, Pinar del Río, Cuba	Terrestrial	Jane Herrera
Titiscania sp.*	Neritopsoidea	15.046	KU342669	MNCN/ ADN:86337	Península de Santa Elena, Guanacaste, Costa Rica	Marine	Yolanda Camacho

GenBank mt genomes

Species	Superfamily	bp	GenBank	Reference
Haliotis rubra	Haliotoidea	16.907	NC_005940	Maynard et al., 2005
Lunella aff. cinerea	Trochoidea	17.670	KF700096	Williams et al., 2014
Angaria neglecta	Angaroidea	19.470	KR297248	Uribe et al., 2016
*Lepetodrilus schrolli**	Lepetodriloidea	15.579	KR297250	Uribe et al., 2016
Phasianella solida	Phasianelloidea	16.698	KR297251	Uribe et al., 2016
Ilyanassa obsoleta	Buccinoidea	15.263	NC_007781	Simison et al., 2006
Rapana venosa	Muricoidea	15.272	NC_011193	Chandler et al., 2008
Oncomelania hupensis	Truncatelloidea	15.182	NC_013073	Li and Zhou 2009
Africonus borgesi	Conoidea	15.536	NC_013243	Cunha et al., 2009
*Nerita melanotragus**	Neritoidea	15.261	GU810158	Castro and Colgan 2010
*Nerita fulgurans**	Neritoidea	15.261	KF728888	Arquez et al., 2014
*Nerita tessellata**	Neritoidea	15.741	KF728889	Arquez et al., 2014
*Nerita versicolor**	Neritoidea	15.866	KF728890	Arquez et al., 2014
Chrysomallon squamiferum	Neomphaloidea	15.388	AP013032	Nakagawa et al., 2014

*incomplete mt genomes

Table 2. Nuclear sequence data analyzed in this study. Length in bp, Genbank accesion number, museum voucher, and sampling location are provided.

New sequences		*18S*		*28S*		*actin*		*H3*			
Species	Superfamily	bp	Genbank	bp	Genbank	bp	Genbank	bp	Genbank	Voucher	Location
Pleuropoma jana	Helicinoidea	1768	KU342679	3714	KU342680	300	KU342681	314	KU342682	AM/ EBU: 65394	Yessabah Caves, New South Wales, Australia
Titiscania sp.	Neritopsoidea	988	KU342683	1437	KU342684	533	KU342685	—	—	MNCN/DNA: 86337	Península de Santa Elena, Guanacaste, Costa Rica
Georissa bangueyensis	Hydrocenoidea	971	KU342673	990	KU342674	488	KU342675	314	KU342676	MNCN/DNA: 86335	Sabah, Sandakan, Malaysia
Gibbula rarilineata†	Trochoidea	—	—	—	—	711	KU342677	314	KU342678	NHMUK/DNA: 20080375	Lumbarda, Korcula I., Croatia
Clanculus cruciatus	Trochoidea	—	—	1415	KU342672	720	KU342671	314	KU342670	MNCN/DNA: 86340	Cabo de Palos, Murcia, Spain

GenBank sequences		*18S*		*28S*		*actin*		*H3*	
Species	Superfamily	bp	Genbank	bp	Genbank	bp	Genbank	bp	Genbank
Titiscania limacina	Neritopsoidea	—	—	—	—	—	—	327	KF527280
*Nerita**	Neritoidea	957	AM048633	1479	AM048693	325	KM025036	—	—
Gibbula rarilineata†	Trochoidea	1182	AF534996	—	—	—	—	—	—
Clanculus cruciatus	Trochoidea	1733	AF120514	—	—	—	—	—	—
Lepetodrilus elevatus#	Lepetodriloidea	1856	AY145381	3408	AY145413	—	—	376	AY923959
Rapana venosa	Muricoidea	1814	HQ834011	1477	HE584197	1565	KF410817	354	HQ834144
Ilyanassa obsoleta	Buccinoidea	1826	AY145379	3415	AY145411	297	AY766115	—	—

* mix of sequences of different species from the same genus in the nuclear and combined data sets
† mixed with the *Lunella* aff. *cinerea* mt genome in the combined analyses (as Trochoidea; see Fig. 3).
# mixed with the *Lepetodrilus schrolli* mt genome in the combined analyses (as *Lepetodrilus*; see Fig. 3).

PCR amplification and sequencing

The mt genomes (except *P. jana*, which was sequenced at the Australian Museum) were PCR amplified following a three-step approach. First, fragments of the *cox1*, *rrnL* and *cox3* genes were PCR amplified using the primers respectively detailed in Folmer et al. (1994), Palumbi et al. (1991) and Boore and Brown (2000). 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 (10 mM), 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 denaturing step at 94°C for 60 s; 45 cycles of denaturation at 94°C for 30 s, annealing using 44ºC (*cox1*) and 52°C (*rrnL* and *cox3*) for 60 s and extension at 72°C for 90 s; a final extension step at 72°C for 5 min. Second, the amplified PCR fragments were sequenced using Sanger sequencing, and new primers were designed (see Appendix for primer sequences) for amplifying outwards from the short fragments in the next step. Third, the remaining mtDNA was amplified in two overlapping fragments by long PCR using the newly designed primers. The long PCR reaction contained 2.5 µl of 10x 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, 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 denaturing step at 94°C

for 60 s; 45 cycles of denaturation at 98°C for 10 s, annealing using a gradient of 50-65°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.

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 bp Paired-ends) at Macrogen (Seoul, Korea).

For the *P. jana* mt genome, next generation sequencing was performed on a CTAB DNA extraction (Saghai-Maroof et al., 1984) of one individual. RNAse A treatment was used after the pellet from the initial precipitation with ethanol was re-suspended. This solution was re-extracted once with phenol/chloroform/*iso*-amyl alcohol 25:24:1 and once with chloroform/*iso*-amyl alcohol 24:1) following which DNA was precipitated from the supernatant with an equal volume of *iso*-propanol. The pellet was washed with 70% ethanol and air dried.

In addition, the following partial nuclear genes were amplified from all species (except *P. jana*) with standard PCR reactions as described above for partial mitochondrial genes but using an annealing temperature of 50ºC in all cases: *28S* rRNA gene using primers LSU5 F (Littlewood et al., 2000) and LSU1600 R (Williams et al., 2003); *18S* rRNA gene using primers 1F

(Winnepenninckx et al., 1995) and 1100R (Williams et al., 2003); *actin* gene using primers FActin2 (5'-ATCTGGCATCACACCTTCTACAAC -3') and RActin2 (5'-ACAGTGTTRGGRTACAAGTCTTTACG -3'); and *histone H3* gene using primers H3MF (5'-ATGGCTCGTACCAAGCAGACTGG-3) and H3MR (5'-TGGATGTCCTTGGGCATGATTGTTAC-3') modified from Colgan et al. (1998). Each amplified fragment was sequenced using the Sanger technology (Table 2).

Contigs containing the individual nuclear genes were identified in *P. jana* by using known neritimorph sequences to search the CLC Genomic assembly (see below) using FASTA3, ver. 3.5 (Pearson, 1999).

Genome assembly and annotation

The reads corresponding to the different mt genomes were sorted using the indexes. Adapter sequences were removed using SeqPrep (StJohn, 2011). Assembly was performed using the TRUFA webserver (Kornobis et al., 2015). 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 and Edwards, 2011). Filtered reads were used for *de novo* assembly of mt genomes, searching for contigs with a minimum length of 3 kb. 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.

For *P. jana*, sequencing was performed at the Australian Genomics Research Facility. After quality-checking with FastQC version 0.10.1 (Andrews, 2010), the raw reads were assembled using the CLC Genomics Workbench. A variety of parameter values were examined and the assembly provided by the set (word = 20, bubble = 30, algorithm = fast) that gave the longest value for N50 was used. Contigs representing the mt genome were identified by FASTA3, ver. 3.5 searches on a local installation (Pearson, 1999) and confirmed by using them as test sequences in BLAST searches of GenBank.

The new neritimorph 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 mt 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. The ribosomal RNA (rRNA) genes were identified by sequence comparison with other reported gastropod mt genomes, and assumed to extend to the boundaries of adjacent genes (Boore et al., 2005).

Sequence alignment

Three different sequence data sets were constructed. The first one, hereafter designated the mt data set, was built aligning the newly determined mt genomes of Neritimorpha with the orthologous sequences from four *Nerita* species available in NCBI (www.ncbi.nlm.nih.gov/). Mt genomes of four Caenogastropoda, five Vetigastropoda, and one Neomphalina were used as gastropod outgroups (Table 1). No Heterobranchia were included in the present analyses because of their well-known high mitochondrial evolutionary rates that result in a long-branch attraction artefact (Grande et al., 2008; Stöger and Schrödl, 2013; Arquez et al., 2014; Uribe et al., 2016). This mt data set included the nucleotide sequence alignments of the two mt rRNA genes, which were combined either with the nucleotide (excluding third codon positions to avoid saturation) or the deduced amino acid sequences of the 13 mt protein coding genes. The second data set, hereafter the nuclear data set, included two ribosomal (*28S rRNA* and *18S rRNA*) and two protein-coding (*actin* and *histone H3*) genes. The nucleotide sequences for these genes produced in this work were aligned with orthologous sequences available in NCBI for Neritimorpha, Caenogastropoda and Vetigastropoda (Table 2). The third data set, named the combined data set, concatenated the mitogenomic and nuclear DNA sequence data. Protein-coding genes were aligned separately using Translator X (Abascal et al., 2010) whereas rRNA genes were aligned separately using MAFFT v7 (Katoh and Standley, 2013) with default parameters. Ambiguously aligned positions were removed using Gblocks v.0.91b (Castresana,

2000) with the following settings: minimum sequence for flanking positions: 85%; maximum contiguous non-conserved positions: 8; minimum block length: 10; gaps in final blocks: no. The generated single alignments were concatenated into the three data matrices using the ALTER webserver (Glez-Peña et al., 2010).

Phylogenetic analyses

Phylogenetic relationships were reconstructed using maximum likelihood (ML; Felsenstein, 1981) and Bayesian inference (BI; Huelsenbeck and Ronquist, 2001). ML analyses were conducted with RAxML v7.3.1 (Stamatakis, 2006) and default parameters using the rapid hill-climbing algorithm and 10,000 bootstrap pseudoreplicates. BI analyses were conducted using MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003) and default parameters, running four simultaneous Monte Carlo Markov chains (MCMC) for 10 million generations, sampling every 1,000 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 BI runs were performed to increase the chance of adequate mixing by the MCMC and to increase the chance of detecting failure to converge, as determined using Tracer v1.6 (Rambaut et al. 2014).

The best partition schemes and best-fit models of substitution for the three data sets were identified using Partition Finder and Protein Partition Finder (Lanfear et al., 2012) according to the

Bayesian information criterion (BIC; Schwarz, 1978). For mt protein-coding genes, the partitions tested were: all genes combined, all genes separated except *atp6-atp8* and *nad4-nad4L*, and genes grouped by subunits (*atp*, *cox*, *cob* and *nad*; see Appendix). In addition, at the nucleotide level the three above partition schemes were tested considering first and second codon positions separated. For the mt rRNA genes, the two genes combined or separated were tested. For the nuclear data set, the partitions tested were: all genes combined, all genes separated, and protein-coding genes grouped versus rRNA genes grouped. Partitions considering the three codon positions were also tested for the protein-coding nuclear genes. The best partitions selected for the mt and nuclear data sets were also used in the combined data set (see Appendix).

To check whether the presence of long branches was biasing the output of the BI analyses performed on the mt data set (with protein-coding genes analyzed at the amino acid level; see results), BI using the site-heterogeneous mixture CAT model (Lartillot and Philippe, 2004) was performed as implemented in PhyloBayes MPI v.1.5 (Lartillot et al., 2013). The site-heterogeneous mixture CAT model assumes that the different sites of a protein evolve under distinct substitution processes and has proven to be less sensitive to (and to 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. Due to restrictions in the program, the protein coding genes (at the amino acid level) and the rRNA genes were analyzed separately 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 Appendix). Posterior probabilities provided branch support for BI analyses.

Estimation of divergence times

The divergence dates between neritimorph clades were estimated using the mt data set at the nucleotide level and an uncorrelated relaxed molecular clock model in BEAST 1.7 (Drummond and Rambaut, 2007). This software was used to infer branch lengths and nodal ages on a fixed topology with regards to neritimorph internal phylogenetic relationships (the reconstructed ML/BI tree topology in Fig. 2). For the clock model, we selected the lognormal relaxed-clock model, which allows rates to vary among branches without any *a priori* assumption of correlation between adjacent branches. For the tree prior, we employed a Yule process of speciation. We employed the partitions selected by Partition Finder (see appendix 2; selected models were simplified to HKY85 to allow convergence). The final Markov chain was run twice, each for 120 million generations, sampling every 10,000 generations with the first 1,200 of the sampled generations discarded as the burn-in, after

confirming convergence of chains with Tracer v.1.5. (Rambaut and Drummond, 2007). The effective sample size of the majority of the parameters was > 200 (except the gamma distribution and the proportion of invariable site of *cox1*, *cox2* and *cox3* genes, first and second positions).

The tree was time-calibrated by setting the ages of (1) the first split within the Neritimorpha, i.e., between *Titiscania* sp. and other study taxa of the clade, and (2) the first split within *Nerita*, i.e., between *N. melanotragus* and three other congeners. The first calibration point was set at a minimum of 235 million years ago (Mya) with a 95% upper limit of 251 Mya (Gamma distribution, Shape: 1, Offset: 235, Scale: 5.3) based on the Triassic records of neritimorphs lacking the internal whorls of the teleoconch (Bandel and Frýda, 1999; Kaim and Sztajner, 2005). The retention of internal coiling is considered to be plesiomorphic and seen in *Neritopsis* (Neritopsoidea) alone among the extant Neritimorpha (Bandel and Frýda, 1999; Kano et al., 2002). The minimum bound was set by referring to the age of the Cassian Formation of the Italian Alps (Lower Carnian). This formation has yielded a variety of non-neritopsoid taxa that might have given rise to the Recent Neritoidea (Bandel, 2007).

The second calibration point, the divergence date of the lineage leading to *N. melanotragus*, was constrained at a minimum age of 49 Mya with a 95% upper limit of 56 Mya (Gamma distribution, Shape: 1, Offset: 49, Scale: 2.3) by referring to the fossil record and a previous molecular phylogeny of the genus

Nerita (Frey and Vermeij, 2008). The crown group representatives of the genus have appeared consistently throughout the record since the late Paleocene and many of the extant subgenera existed already in the Lutetian age (40–49 Mya; see Symonds, 2009 for figures of fossil specimens). *Nerita melanotragus* represents one of the earliest offshoots among the extant species of the genus (Frey and Vermeij, 2008).

RESULTS

Sequencing and assembly

The nucleotide sequences of the complete mt genomes of *Georissa bangueyensis, Neritina usnea*, and *Pleuropoma jana* and the nearly complete mt genomes of *Titiscania* sp., and *Theodoxus fluviatilis* were determined (see Table 1 for lengths and Genbank accession numbers). In addition, about half of the mtDNA of *Viana regina* (7,634 bp) was sequenced to determine whether other Helicinoidea share the derived gene order of *Pleuropoma*. The HiSeq2000 sequencer produced a similar amount of data for *Titiscania* sp. (580,418 reads; 79.2 Mb), *T. fluviatilis* (786,306 reads; 602.2 Mb), and *V. regina* (509,154 reads; 137.2 Mb). However, fewer data (7,584 reads; 2 Mb) were produced for *G. bangueyensis*. All these samples were run in lanes together with TruSeq RNA libraries (from other projects). Interestingly, *N. usnea* produced more data (10,113,794 reads; 2.74 Gb) and it was run in a lane together with Nextera DNA libraries (from other projects). After assembly, the

average coverage for each mt genome was: 42x for *G. bangueyensis*; 32,921x for *N. usnea*; 79x for *Titiscania* sp.; 602x for *T. fluviatilis*; 621x for *P. jana*; and 137x for *V. regina.* In the mt genomes of *Titiscania* sp. and *T. fluviatilis*, the sequencing strategy failed to read through a region of less than 1kb including a cluster of tRNA genes and the putative control region (see Appendix). The lengths and Genbank accession numbers for the nucleotide sequences of the partial nuclear genes of *Georissa bangueyensis, Titiscania* sp, *Pleuropoma jana, Gibbula rarilineata* (Michaud, 1829), and *Clanculus cruciatus* (Linnaeus, 1758) that were determined in this study are shown in Table 2.

Structural features and mitochondrial organization

The complete mt genomes of *G. bangueyensis*, *N. usnea,* and *P. jana* contain 13 protein coding, two ribosomal RNA, and 22 transfer RNA genes (Fig. 1). In addition, the mt genome of *P. jana* has an extra *trnM* gene. The gene arrangement of the mt genomes of *G. bangueyensis* and *N. usnea* is identical: the major strand encodes *cox1-3*, *atp6-8*, *nad3*, *nad2*, *trnD*, *trnT*, *trnS*(GCU), and the KARNI (*trnK*, *trnA*, *trnR*, *trnN*, and *trnI*) cluster whereas the minus strand encodes the remaining protein-coding genes (*nad5*, *nad4*, *nad4L*, *cob*, *nad6*, and *nad1*), the two ribosomal RNA genes (*rrnS* and *rrnL*), the *trnF*, *trnH*, *trnS*(UGA), *trnP*, *trnL*(UAA), *trnL*(UAG) genes and the MYCWQGE (*trnM*, *trnY*, *trnC*, *trnW*, *trnQ*, *trnG*, and *trnE*) cluster (Fig. 1). The nearly complete mt genomes of *Titiscania* sp. and *T. fluviatilis* generally have the same gene order

as *N. usnea* and *G. bangueyensis* (but note that we failed to sequence the MYCWQGE tRNA gene cluster and the control region in the two former species). The gene order of *P. jana* is markedly different from that shared by the above mentioned four species. All genes are encoded by the major strand except the *trnT* gene. This may have been due to a single inversion event that affected the following genes: *trnF*, *nad5*, *trnH*, *nad4*, *nad4L*, *trnY*, *trnS*, *cob*, *nad6*, *trnP*, *nad1*, *trnL*(UAA), *trnL*(UAG), *rrnL*, *trnV*, *rrnS*, *trnM*, and *trnY* (Fig. 1). In addition, this mt genome shows independent rearrangements (inversions and translocations) of the *trnL*(UAG), *trnW*, *trnQ*, *trnG*, and *trnE* genes. The sequenced half of the mt genome of *V. regina* shares the same order of *P. jana*, except in the cluster KARNI, where *trnN* and *trnI* genes were translocated (KARIN; see Fig. 1).

The genes *nad4*/*nad4L* overlapped by seven bp in all mt genomes where this region was sequenced (all but *V. regina*). All determined protein-coding genes start their open reading frame with the codon ATG, except *nad3* and *nad4* of *P. jana* that start with ATA and TTG, respectively (see Appendix). The most commonly used stop codon was TAA. The *atp8* and *cox3* genes of *G. bangueyensis*, the *cox2* and *nad4L* genes of *N. usnea*, the *atp8*, *cox3*, *nad1*, *nad4*, and *nad6* genes of *Titiscania* sp., the *cox2* and *nad6* genes of *P. jana*, the *cox3* and *nad3* genes of *T. fluviatilis*, and the *atp6* and *cox1* genes of *V. regina* end with TAG (see Appendix). Other genes end with incomplete stop codons either TA- as the *nad3* and *nad5* genes of *G. bangueyensis* or T-- as the *cox2*, *nad4* and *nad2* genes of *G. bangueyensis*, the *cob*, *nad3*, *nad5*, and *nad6*

genes of *P. jana*, the *nad4* and *nad2* genes of *N. usnea*, the *nad2* gene of *Titiscania* sp., the *nad2* and *nad5* genes of *T. fluviatilis*, and the *nad3* and *nad2* genes of *V. regina* (see Appendix). In the complete mt genomes, the longest intergenic regions were found between the *trnE* and *cox3* genes in *G. bangueyensis* (608 bp) and *N. usnea* (471 bp) and between the *trnF* and *trnL*(UAG) genes in *P. jana* (783 bp).

Phylogenetic relationships

The phylogenetic relationships of Neritimorpha were reconstructed based on the mt, nuclear, and combined data sets using probabilistic methods (Figs. 2 and 3). The alignment of the mt data set including protein-coding genes analyzed at the nucleotide level (only first and second codon positions) plus rRNA genes was 7,604 positions long. The ML (-*ln*L = 63,623.8) and BI (-*lnL* = 63,529.6 for run 1; -*ln*L = 63,528.9 for run 2) phylogenetic analyses arrived at the same topology (Fig. 2) by using Neomphalina to root the tree following Uribe et al. (2016). ML and BI reconstructed trees both showed high statistical support for all nodes but one referring to an internal relationship within the genus *Nerita* (Fig. 2). Neritimorpha was recovered monophyletic with maximal statistical support, and within Neritimorpha, the species *Titiscania* sp. representing the superfamily Neritopsoidea was recovered as the sister group of the remaining neritimorph superfamilies. The next clade that branched off within Neritimorpha included the species *P. jana* and *V. regina* as representatives of the superfamily Helicinoidea. The

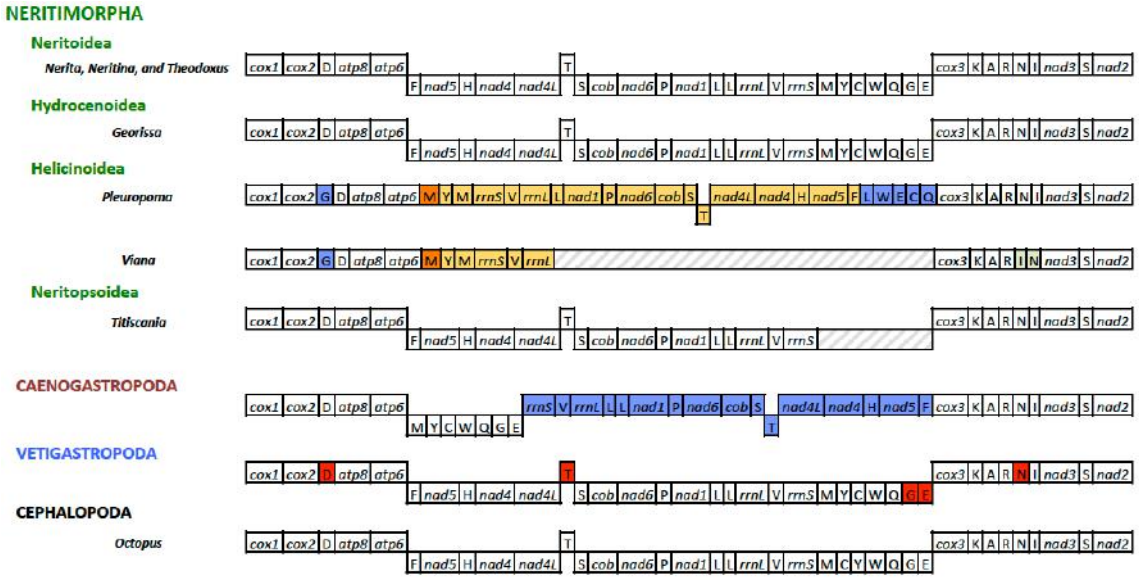


Figure 1. Mitochondrial gene orders of the four superfamilies of Neritimorpha. The gene organization of the mt genomes of Cephalopoda, Vetigastropoda, and Caenogastropoda is shown for comparison. The genes encoded in the major and minor strands are shown in the top and bottom lines, respectively. Inversions are colored in yellow. Translocated genes are in green. Genes translocated and inverted are in blue. Genes duplicated are in orange. The red boxes in Vetigastropoda show the tRNA genes that rearrange in some lineages of this group with respect to Neritimorpha. Striped boxes indicate regions not sequenced in Titiscania and Viana (the region not sequenced in Theodoxus is of similar position and length to that of Titiscania but is not shown).

superfamilies Hydrocenoidea (*G. bangueyensis*) and Neritoidea were recovered as sister groups. Within Neritoidea, *Neritina* and *Theodoxus* formed a clade that was the sister group of *Nerita*. Both Helicinoidea and Hydrocenoidea showed relatively long branches (Fig. 2). The protein-coding genes of the mt data set were also analyzed at the amino acid level combined with the rRNA genes. The final matrix was 4,484 positions long.

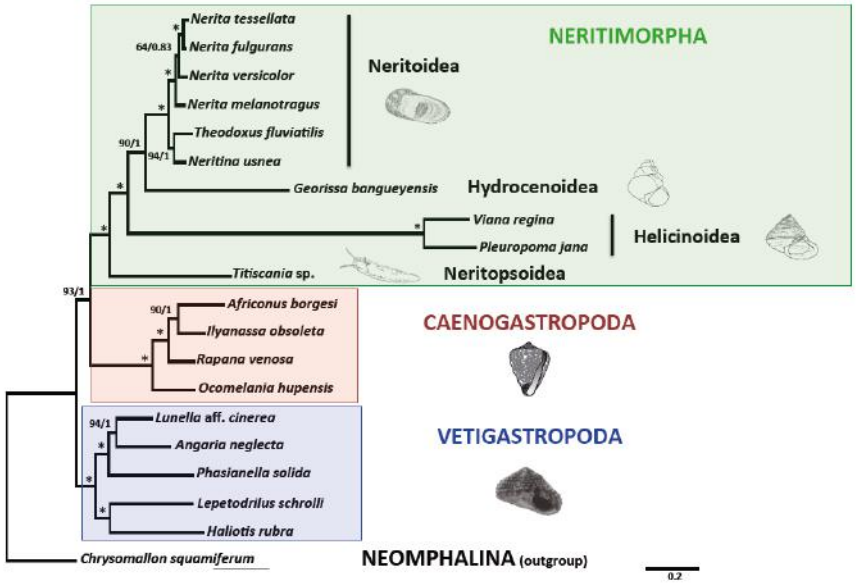


Figure 2. Phylogenetic relationships of Neritimorpha based on complete mt genomes (protein coding genes excluding third codon positions plus rRNA genes; 7,604 positions in total). The reconstructed ML phylogram using Neomphalina as outgroup is shown. Branch colors indicate main gastropod lineages. The names of the four neritimorph superfamilies are indicated. Numbers at nodes are statistical support values for ML (bootstrap proportions)/ BI (posterior probabilities). Red asterisks indicate nodes with maximal ML and BI support. Scale bar indicates substitutions/site.

The ML (-*ln*L = 57,096.7) and BI (-*ln*L = 67,830.4 for run 1; -*ln*L = 67,830.7 for run 2) trees recovered similar topologies that only differed in the relative phylogenetic positions of superfamilies Hydrocenoidea and Helicinoidea. While ML recovered the topology shown in Figure 2, BI grouped Hydrocenoidea and Helicinoidea together with maximal support (see Appendix). However, BI analyses using the site-heterogeneous mixture CAT model as implemented in PhyloBayes recovered the topology shown in Figure 2.

The alignment of the nuclear data set was 2,962 positions long. The ML (-lnL = 10,841.8) and BI (-lnL = 10,971.6 for run1; -lnL = 10,973.3 for run2) phylogenetic analyses arrived at the same topology (Fig. 3, inset). Vetigastropoda were used as outgroup since the analyzed nuclear genes were not available for Neomphalina. Within Neritimorpha, Neritopsoidea was recovered as the sister group of Helicinoidea plus Hydrocenoidea and Neritoidea (Fig. 3, inset). All nodes in this tree had maximal support in ML and BI, except that grouping Hydrocenoidea and Neritoidea (BP: 67%; PP: 0.94). Interestingly, *G. bangueyensis* but not *P. jana* showed a long branch mainly due to accelerated substitution rates in the 28S rRNA gene.

The matrix of the combined data set was 10,566 positions long. The ML (-lnL = 49,365.4) and BI (-lnL = 49,355.5 for run1; -lnL = 49,355.8 for run2) phylogenetic analyses arrived at the same topology, with all nodes receiving maximal support in both analyses (Fig. 3).

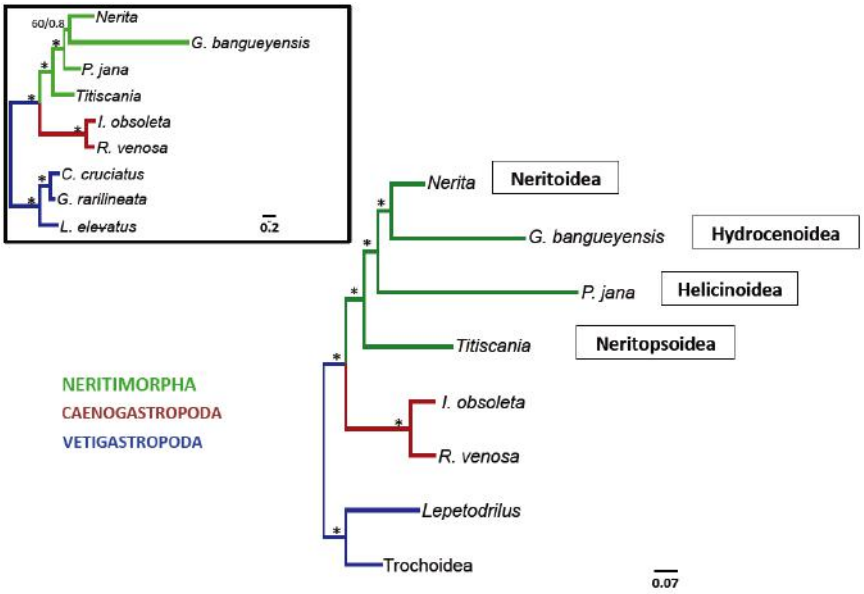


Figure 3. Phylogenetic relationships of Neritimorpha based on combined mt and nuclear data (10,566 positions long). The reconstructed ML phylogram using Vetigastropoda as outgroup is shown. Additionally, a ML phylogeny based solely on nuclear data (2,962 positions from *18S rRNA*, *28S rRNA*, *actin* and *histone H3* genes) is shown in the inset. Branch colors indicate main gastropod lineages. The names of the four neritimorph superfamilies are indicated. For some tree tips, sequences were concatenated from species for which there is strong evidence for the monophyly of the higher taxon rank, which is indicated. Numbers at nodes are statistical support values for ML (bootstrap proportions)/ BI (posterior probabilities). Red asterisks indicate nodes with maximal ML and BI support. Scale bar indicates substitutions/site.

According to the estimated divergence times (Fig. 4), the superfamily Helicinoidea branched off at 234 Mya (95% credible interval: 206–269 Mya) from the lineage leading to the Hydrocenoidea and Neritoidea. The split between the latter two superfamilies was estimated to have occurred at around 180 Mya

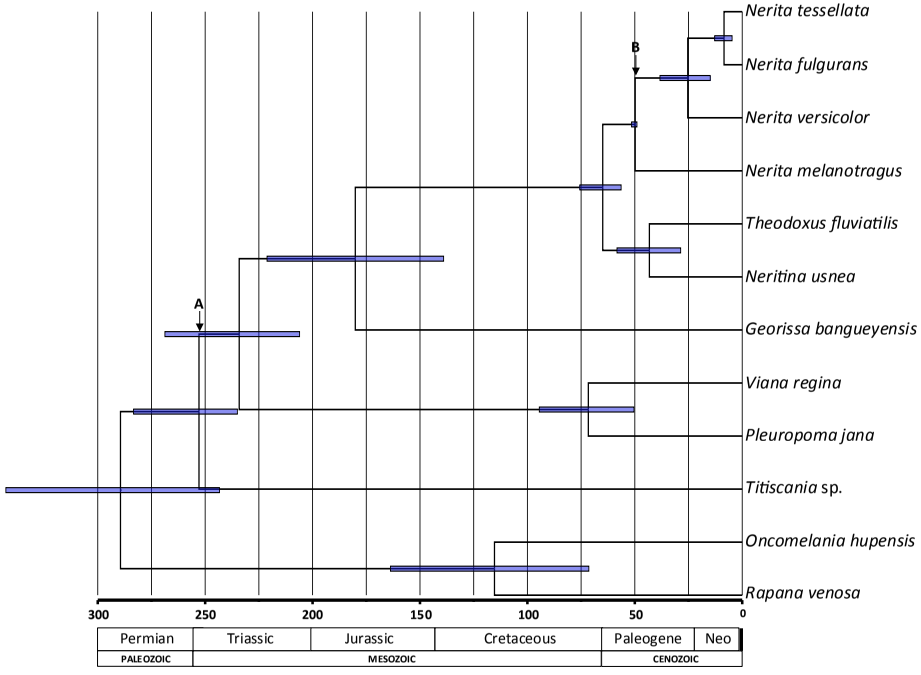


Figure 4. Chronogram of Neritimorpha based on mt data set (7,604 positions) and fossil-based calibration priors (A and B). A Bayesian uncorrelated relaxed lognormal clock was used in BEAST. Horizontal bars represent 95% credible intervals for time estimates; dates are in millions of years.

(139–221 Mya) in the Jurassic period. Within-family divergences of the studied taxa were suggested to have started in the Late Cretaceous and the boundary with the Paleocene time (Helicinidae: 72 Mya; Neritidae: 65 Mya).

DISCUSSION

Gene arrangement

Until the present work, no complete mt genome of any Neritimorpha was available, although up to four almost complete mt genomes lacking only the control region between the *trnE* and *cox3* genes had been determined for the genus *Nerita* (Castro and Colgan, 2010; Arquez et al., 2014). Here, we present three complete and two nearly complete mt genomes that together represent all known living superfamilies of Neritimorpha. The mt genomes of *Titiscania* sp., *G. bangueyensis*, *N. usnea*, and *T. fluviatilis* share the same gene organization previously described for *Nerita* mt genomes (Castro and Colgan, 2010; Arquez et al., 2014), indicating that three (Neritoidea, Hydrocenoidea, and Neritopsoidea) out of the four superfamilies have not undergone major gene rearrangements during their evolutionary history. Moreover, this genome organization is very similar to the one found in Vetigastropoda (Williams et al., 2014; Uribe et al., 2016) with only minor differences in the relative positions of several tRNA genes (marked in red in Fig. 1), nearly identical (except on a tRNA gene transposition) to the gene order reported for the cephalopod

Octopus vulgaris (Yokobori et al., 2004; Fig. 1), and only differing in a translocation of the *trnD* gene and an inversion of the *trnP* gene with respect to the mtDNA of *Katharina tunicata* (Wood, 1815) (Polyplacophora; Boore and Brown, 1994).

In contrast, the mt genome of *P. jana* (Helicinoidea) shows a markedly different gene order, which has been likely originated through a single inversion event that affected 17 genes. A second *trnM* gene in the upstream border of the inversion (Fig. 1) could represent a remnant of the rearrangement process. Moreover, in this mt genome the MYCWQGE cluster found in other gastropod mt genomes (Uribe et al., 2016) has been inverted, disrupted (e.g., the *trnG* gene has been inverted and translocated between the *cox2* and *trnD* genes) and rearranged. The sequencing of half the mt genome of *V. regina* confirmed that the peculiar gene order of *P. jana* including the large inversion, the inversion and translocation of the *trnG* gene, and the duplication of the *trnM* gene, is likely shared by all helicinids (Fig. 1). Interestingly, almost the same set of genes is involved in a gene rearrangement in the common ancestor of Caenogastropoda, where the integrity of the MYCWQGE cluster was nevertheless maintained (Cunha et al., 2009; Osca et al., 2014b, 2015). It is also noteworthy that the branch leading to the two sequenced species of Helicinoidea is the longest in the phylogenetic tree, indicating a significant acceleration of substitution rates. As in other reported cases (Shao et al., 2003; Stöger and Schrödl, 2013), this increase in substitution rate is correlated with an increase in rearrangement rates and a consequent major reorganization.

Phylogenetic relationships and radiation of Neritimorpha

All phylogenetic analyses performed in this study except one converged on the same topology, which received high statistical support. The only exception (BI based on the mt data set with protein-coding genes analyzed at the amino acid level) may be considered a spurious result related to a long-branch attraction artefact between Hydrocenoidea and Helicinoidea, which, in fact, was ameliorated when the data was analyzed using the site-heterogeneous mixture CAT model (Lartillot and Philippe, 2004; Lartillot et al., 2007). The monophyly of Neritimorpha is undisputed (Kano et al., 2002; Lindberg, 2008) and also supported by our phylogenetic analyses. Within Neritimorpha, the Neritopsoidea is recovered in our phylogenetic analyses as the sister group of the other three superfamilies. This result is in agreement with morphological evidence (Holthuis, 1995) as well as with the only available molecular phylogeny of Neritimorpha, which was based on the 28S rRNA gene (Kano et al., 2002). The Neritopsoidea have the oldest fossil record of all living neritimorph superfamilies, dating back at least to the Triassic and probably earlier (Bandel, 2000, 2007; Kaim and Sztajner, 2005), and show important conchological and anatomical characteristics. The shell of Neritopsoidea retains internal whorls as well as the central axis of coiling, which have been reabsorbed in the other superfamilies (Solem, 1983; Bandel and Frýda, 1999). Among many other anatomical discrepancies, the male pallial gonoduct of

Neritoposidea has open lamellae, which have been fused ventrally closing the duct in the other superfamilies (Holthuis, 1995; Kano et al., 2002; Kano and Kase, 2002).

With regards to the phylogenetic relationships among the remaining three superfamilies, our results conflict with those in the previous 28S rRNA phylogeny (Kano et al., 2002). While our present analyses place *Georissa* (as a representative of Hydrocenoidea) in a sister position to Neritoidea, the 28S rRNA phylogeny recovered the genus as the second earliest offshoot of the Recent Neritimorpha (Kano et al., 2002). However, the latter relationship was only moderately supported in their ML analysis (Kano et al., 2002), due to a long branch leading to *Georissa* that biased the phylogenetic reconstruction. Actually, the addition of the other extant genus of the superfamily, *Hydrocena*, renders the 28S rRNA topology quite the same as the present topology based on mt genomes (Y. Kano and I. Richling, unpublished data).

Our divergence time estimates based the molecular data and two fossil-based calibration points suggested a Jurassic origin of the Hydrocenoidea, in accordance with the disjunct geographic distribution of its living taxa; the species of *Hydrocena* occur in southern Europe and on eastern Atlantic islands while *Georissa* has a wide Indo-West Pacific distribution, hence possibly suggesting a Pangaean origin for the family (Bandel, 2000). The estimated date of the divergence does not contradict with the assumption that *Schwardtina cretacea* (Tausch, 1856) from fresh- or brackish-water deposits of the Late Cretaceous (Bandel and Riedel, 1994)

constitutes the earliest representative of Hydrocenoidea (Kano et al., 2002). This species shows an intriguing resemblance to the extant Hydrocenidae in the protoconch and teleoconch shapes, irrespective of its aquatic habitat (Bandel, 2000; Fukumori and Kano, 2014). The earliest Neritoidea that can be unambiguously classified into extant genera also appeared in the Late Cretaceous or in the Paleocene (Bandel and Kiel, 2003; Frey and Vermeij, 2008). Although a number of older Mesozoic taxa have been assigned to Neritidae based solely on the plesiomorphic hemispherical shell (Kano et al., 2002), we suggest based on the time-calibrated phylogeny that neritids diverged in the Cretaceous period and subsequently colonized freshwater habitats.

The present topology further emphasizes the phenotypic plasticity in neritimorph evolution. The astonishing diversity seen in the shell morphology and anatomy of the group is evidently the outcome of their major ecological radiation (Kano et al., 2002; Lindberg, 2008). The comparison of the morphological traits and even the identification of character homology are particularly difficult between the aquatic and terrestrial taxa, due to drastic changes in the body plan to adapt to a life in the air (Little, 1990; Kano and Kase, 2002). Based on their 28S rRNA phylogeny, Kano et al. (2002) proposed that the evolutionary history of the four superfamilies could only be traced morphologically by different conditions of the female reproductive system: the Neritopsoidea and Hydrocenoidea exhibit a monaulic condition with a single gonopore, which might represent a plesiomorphic character state, while the others have independent vaginal and ootype openings,

which are separately located in the Helicinoidea and adjoin in the Neritoidea (e.g. Bourne, 1911; Fretter, 1984; Holthuis, 1995; Kano and Kase, 2002; Richling, 2004). The presently recovered sister relationship between the monaulic Hydrocenoidea and diaulic Neritoidea clearly rejects such a simple, gradual increase of complexity in the female reproductive system. To our knowledge, none of distribution patterns of morphological traits adequately explains the relationships among the extant superfamilies of Neritimorpha as reconstructed in our phylogeny.

In any case, our results agree with those of Kano et al. (2002) and with morphological studies (Bourne, 1911; Haszprunar, 1988; Little, 1990) in supporting that the transition from marine to terrestrial mode of life, despite being complex processes involving numerous adaptations at morphological, physiological and ecological levels, has occurred more than twice convergently within the evolutionary history of Neritimorpha. Besides the independent terrestrialization of Hydrocenoidea, Helicinoidea and *Neritodryas*, the Paleozoic *Dawsonella meeki* most probably represents an early offshoot of Neritimorpha as one of the first terrestrial gastropods (Bourne, 1911; Solem, 1983). The Carboniferous occurrence of the species, however, long precedes the presumed Triassic split between the Neritopsoidea and the other extant superfamilies (Bandel and Frýda, 1999; Kaim and Sztajner, 2005; Bandel, 2007), hence exemplifying another line of land invasion (Bourne, 1911; Kano et al., 2002).

CONCLUSION

Phylogenetic information contained in complete mt genome sequences is useful in resolving phylogenetic relationships within major lineages of gastropods. Here we sequenced complete mtDNAs of representatives of the four superfamilies of Neritimorpha and recovered a highly supported phylogenetic tree at this taxonomic level that is congruent with nuclear-based inferences. Hence, the reconstructed phylogeny provides a statistically robust framework for a variety of comparative studies within the group as exemplified here by the inference of multiple invasions of land. Future mt genome phylogenetic studies incorporating a denser taxon sampling at the family level will additionally contribute to a more refined understanding of the evolutionary processes that generated the astonishing morphological diversity and habitat adaptation observed in Neritimorpha.

ACKNOWLEDGEMENTS

We want to express our gratitude to Iker Irisarri (University of Konstanz) for his seminal work in this project. We are indebted to Yolanda Camacho (Universidad of Costa Rica) for her gift of several specimens of *Titiscania* sp.; José Templado (MNCN-CSIC) for the samples of *Theodoxus fluviatilis*; Jane Herrera Uria (Museo Nacional de Historia Natural de Cuba) and Thierry Backeljau (RBINS) for the individuals of *Viana regina*; Moises Arquez (Universidad Simón Bolivar) for the samples of *Neritina usnea,* and

Menno Schilthuizen (Naturalis Biodiversity Center) for the specimens of *Georissa bangueyensis*. 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). We are grateful for the critical reading of the original version of the manuscript by two anonymous reviewers.

REFERENCES

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

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

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. Mar. Genomics 15, 45-54.

Bandel, K., 2000. The new family Cortinellidae (Gastropoda, Mollusca) connected to a review of the evolutionary history of the subclass Neritimorpha. Neues Jahrb Geol P-A 217, 111-129.

Bandel, K., 2007. Description and classification of Late Triassic Neritimorpha (Gastropoda, Mollusca) from the St Cassian Formation, Italian Alps. Bull. Geosci. 82, 215-274.

Bandel, K., Frýda, J., 1999. Notes on the evolution and higher classification of the subclass Neritimorpha (Gastropoda) with the description of some new taxa. Geol. Palaeontol. 33, 219-235.

Bandel, K., Kiel, S., 2003. Relationships of Cretaceous Neritimorpha (Gastropoda, Mollusca), with the description of seven new species. Bull. Czech Geol. Surv. 78, 53-65.

Bandel, K., Riedel, F., 1994. The late Cretaceous gastropod fauna from Ajka (Bakony Mountains, Hungary): a revision. Ann. Nat. Hist. Mus. Wien A 96, 1–65.

Bergh, R., 1890. Die Titiscanien, eine Familie der rhipidoglossen Gastropoden. Morphol. Jahrb. 16, 1–26.

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. Mol. Phylogen. Evol. 69, 313-319.

Boore, J.L., Brown, W.M., 1994. Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. Genetics 138, 423-443.

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. Mol. Biol. Evol. 17, 87-106.

Boore, J.L., Macey, J.R., Medina, M., 2005. Sequencing and comparing whole mitochondrial genomes of animals. Methods Enzymol. 395, 311 - 348.

Bouchet, P., Rocroi, J.P., Frýda, J., Hausdorf, B., Ponder , W.F., Valdés, Á., A., W., 2005. Classification and nomenclator of gastropod families. Malacologia 47, 1-397.

Bourne, G.C., 1911. Contributions to the morphology of the group Neritacea of the aspidobranch gastropods. Part II. The Helicinidae. Proc. Zool. Soc. Lond. 1911, 759-809.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 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). Mol. Phylogen. Evol. 57, 918-923.

Colgan, D.J., McLauchlan, A., Wilson, G.D.F., Livingston, S.P., Edgecombe, G.D., Macaranas, J., Cassis, G., Gray, M.R., 1998. Molecular phylogenetics of the Arthropoda: relationships based on histone H3 and U2 snRNA DNA sequences. Aust. J. Zool. 46, 419-437.

Cunha, R., Grande, C., Zardoya, R., 2009. Neogastropod phylogenetic relationships based on entire mitochondrial genomes. BMC Evol. Biol. 9, 210.

Drummond, A., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 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. Mol. Mar. Biol. Biotechnol. 3, 294-299.

Fretter, V., 1984. The functional anatomy of the neritacean limpet *Phenacolepas omanensis* Biggs and some comparison with *Septaria*. J. Molluscan Stud. 50, 8-18.

Frey, M.A., Vermeij, G.J., 2008. Molecular phylogenies and historical biogeography of a circumtropical group of gastropods (Genus: *Nerita*): Implications for regional diversity patterns in the marine tropics. Mol. Phylogen. Evol. 48, 1067-1086.

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

Frýda, J., Racheboeuf, P.R., Frýdová, B., Ferrová, L., Mergl, M., Berkyová, S., 2009. Platyceratid gastropods—stem group of patellogastropods, neritimorphs or something else? . Bull. Geosci. 84, 107-120.

Fukumori, H., Kano, Y., 2014. Evolutionary ecology of settlement size in planktotrophic neritimorph gastropods. Mar. Biol. 161, 213–227.

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 Res. 38, W14-W18.

Grande, C., Templado, J., Zardoya, R., 2008. Evolution of gastropod mitochondrial genome arrangements. BMC Evol. Biol. 8, 61.

Haszprunar, G., 1988. On the origin and evolution of major gastropods group, with special reference to the streptoneura. J. Molluscan Stud. 54, 367 - 441.

Holthuis, B.V., 1995. Evolution between marine and freshwater habitats: a case study of the gastropod Neritopsina PhD thesis. University of Washington, Seattle.

Huelsenbeck, J., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754-755.

Kaim, A., Sztajner, P., 2005. The opercula of neritopsid gastropods and their phylogenetic importance. J. Molluscan Stud. 71, 211–219.

Kano, Y., Chiba, S., Kase, T., 2002. Major adaptive radiation in neritopsine gastropods estimated from 28S rRNA sequences and fossil records. Proc. R. Soc. Biol. Sci. Ser. B 269, 2457-2465.

Kano, Y., Fukumori, H., Brenzinger, B., Warén, A., 2013. Driftwood as a vector for the oceanic dispersal of estuarine gastropods

(Neritidae) and an evolutionary pathway to the sunken-wood community. J. Molluscan Stud. 79, 378-382.

Kano, Y., Kase, T., 2002. Anatomy and systematics of the submarine-cave gastropod *Pisulina* (Neritopsina, Neritiliidae). J. Molluscan Stud. 68, 365–384.

Kano, Y., Kase, T., 2004. Genetic exchange between anchialine cave populations by means of larval dispersal: the case of a new gastropod species *Neritilia cavernicola*. Zool. Scr. 33, 423-437.

Kano, Y., Kase, T., 2008. Diversity and distributions of the submarine-cave Neritiliidae in the Indo-Pacific (Gastropoda: Neritimorpha). Org Divers Evol 8, 22-43.

Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. 30, 772-780.

Kornobis, E., Cabellos, L., Aguilar, F., Frías-López, C., Rozas, J., Marco, J., Zardoya, R., 2015. TRUFA: A user-friendly web server for de novo RNA-seq analysis using cluster computing. Evol. Bioinform. 11, 97-104.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 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 Evol. Biol. 7 Suppl. 1, S4.

Lartillot, N., Philippe, H., 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol. Biol. Evol. 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. Syst. Biol. 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: Ponder, W.F., Lindberg, D.R. (Eds.), Phylogeny and Evolution of the Mollusca. University of California Press, Berkeley, pp. 271-296

Little, C., 1990. The terrestrial invasion. An ecophysiological approach to the origin of land animals. Cambridge University Press, Cambridge.

Littlewood, D.T.J., Curini-Galletti, M., Herniou, E.A., 2000. The Interrelationships of Proseriata (Platyhelminthes: Seriata) tested with molecules and morphology. Mol. Phylogen. Evol. 16, 449-466.

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 Evol. Biol. 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.

Osca, D., Templado, J., Zardoya, R., 2015. Caenogastropod mitogenomics. Mol. Phylogen. Evol. 93, 118-128.

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.

Pearson, W.R., 1999. Flexible sequence similarity searching with the FASTA3 program package. In: Misener, S., Krawetz, S.A. (Eds.), Methods in Molecular Biology. Bioinformatics methods and protocols. Humana Press, Totowa, New Jersey, pp. 185–219.

Ponder, W.F., Lindberg, D.R., 1997. Towards a phylogeny of gastropod molluscs: an analysis using morphological characters. Zool. J. Linn. Soc. 119, 83-265.

Rambaut, A., Drummond, A.J., 2007. Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Richling, R., 2004. Classification of the Helicinidae: Review of morphological characteristics based on a revision of the Costa Rican species and application to the arrangement of the Central American mainland taxa (Mollusca: Gastropoda: Neritopsina). Malacologia 45, 195-440.

Richling, R., 2014. Poorly explored jewels of the tropics: Estimating diversity in non-pulmonate land snails of the family Helicinidae (Gastropoda: Neritopsina). Am. Malacol. Bull. 32, 246-258.

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

Saghai-Maroof, M.A., K.M., S., Jorgensen, R.A., Allard, R.W., 1984. Ribosomal DNA spacer-length polymorphism in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. USA 81, 8014-8019.

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 Res. 33, W686-W689.

Schilthuizen, M., Rutten, E.M.J., Haase, M., 2012. Small-scale genetic structuring in a tropical cave snail and admixture with its above-ground sister species. Biol. J. Linn. Soc. 105, 727–740.

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

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461-464.

Sevigny, J.L., Kirouac, L.E., Thomas, W.K., Ramsdell, J.S., Lawlor, K.E., Sharifi, O., Grewal, S., Baysdorfer, C., Curr, K., Naimie, A.A., Okamoto, K., Murray, J.A., Newcomb, J.M., 2015. The mitochondrial genomes of the nudibranch mollusks, *Melibe leonina* and *Tritonia diomedea*, and their Impact on gastropod phylogeny. PLoS ONE 10, e0127519.

Shao, R., Dowton, M., Murrell, A., Barker, S.C., 2003. Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. Mol. Biol. Evol. 20, 1612 - 1619.

Solem, A., 1983. Lost or kept internal whorls: ordinal differences in land snails. J. Molluscan Stud. Suppl. A 12, 172–178.

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

StJohn, J., 2011. SeqPrep. https://github.com/jstjohn/SeqPrep.

Stöger, I., Schrödl, M., 2013. Mitogenomics does not resolve deep molluscan relationships (yet?). Mol. Phylogenet. Evol. 69, 376-392.

Symonds, M. F. 2009. The Neritopsidae and the Neritidae of the Bracklesham Group (Early and Middle Eocene) of the Hampshire Basin. Cainozoic Res. 6, 37–51.

Thompson, F.G., 1980. Proserpinoid land snails and their relationships within the Archaeogastropoda. Malacologia 20, 1–33.

Uribe, J.E., Kano, Y., Templado, J., Zardoya, R., 2016. Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. Zool. Scr. 45, 145-159.

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

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., Reid, D.G., Littlewood, D.T.J., 2003. A molecular phylogeny of the Littorininae (Gastropoda: Littorinidae): unequal evolutionary rates, morphological parallelism, and biogeography of the Southern Ocean. Mol. Phylogen. Evol. 28, 60-86.

Winnepenninckx, B., Backeljau, T., De Wachter, R., 1995. Phylogeny of protostome worms derived from 18S rRNA sequences. Mol. Biol. Evol. 12, 641-649.

Wyman, S.K., Jansen, R.K., Boore, J.L., 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20, 3252-3255.

Yokobori, S.-i., Fukuda, N., Nakamura, M., Aoyama, T., Oshima, T., 2004. Long-term conservation of six duplicated structural genes in cephalopod mitochondrial genomes. Mol. Biol. Evol. 21, 2034-2046.

SUPPORTING INFORMATION 4

Additional Supporting Information may be found in the online version of this article:

Data S1. Long PCR and primer walking primers.

Data S2. Evolutionary sustitution models selected by Partition Finder

Data S3. Mitochondrial genome features

Data S4. Phylobayes tree

3.5. PUBLICACION 5:

Titulo: "Beyond *Conus*: phylogenetic relationships of Conidae based on complete mitochondrial genomes"

Autores: Juan Esteban Uribe, Nicollas Puillandre y Rafael Zardoya

Estado: en revisión

Año: 2016

Revista: *Molecular Phylogenetics and Evolution*

Resumen

La comprensión de cómo evolucionó la extraordinaria diversidad taxonómica y ecológica de los caracoles cono (Caenogastropoda: Conidae), requiere un enfoque filogenético estadísticamente robusto, que hasta ahora no está disponible. Aunque las filogenias moleculares recientes han sido capaces de distinguir varios linajes profundos dentro de la familia Conidae, incluyendo *Profundiconus*, *Californiconus*, *Conasprella*, y *Conus* (y dentro de este, varios subgéneros), las relaciones filogenéticas entre estos géneros permanecen sin esclarecer. Además, la posibilidad de que puedan existir linajes profundos dentro de la familia permanece abierta. Aquí, recontruímos con métodos probabilísticos una filogenia molecular usando nuevas secuencias de genomas mitocondriales (mt) completos y casi completos de las siguientes nueve especies que representan todos los linajes principales de Conidae y potencialmente uno más: *Profundiconus teramachii*, *Californiconus californicus*, *Conasprella wakayamaensis*, *Lilliconus sagei*, *Pseudolilliconus traillii*, *Conus* (*Kalloconus*) *venulatus*, *Conus* (*Lautoconus*) *ventricosus*, *Conus* (*Lautoconus*) *hybridus*, y *Conus* (*Eugeniconus*) *nobilis*. Para testar la monofília de la familia, también secuenciamos los genomas casi completos de las siguientes tres especies representandes de familias de conoideos con relación cercana: *Benthomangelia* sp. (Mangeliidae), *Tomopleura* sp. (Borsoniidae) y *Glyphostoma* sp. (Clathurellidae). Todos los genomas mt secuenciados de conoideos comparten un orden de genes relativamente constante con los reordenamientos limitados a los genes de ARNts. La reconstrucción filogenética recuperó con

alto apoyo estadístico la monofilia de Conidae y las relaciones dentro de la familia. El género *Profundiconus* se posicionó como hermano del resto de géneros. Dentro de estos, un clado incluyendo *Californiconus* y *Lilliconus* + *Pseudolilliconus* fue grupo hermano de *Conasprella* excluyendo a *Conus*. La filogenia incluyó un nuevo linaje cuya relativa posición filogenética era desconocida (*Lilliconus*) y destapó diversidad hasta ahora oculta dentro de la familia (*Pseudolilliconus*). Además, las relaciones filogenéticas reconstruidas permitieron inferir que la peculiar dieta de *Californiconus* básada en gusanos, moluscos, crustáceos y peces es derivada, y reforzar la hipótesis de que el ancestro de Conidae fue un cazador de gusanos. Se reconstruyó un cronograma basado en un reloj molecular no correlacionado, que dató el origen de la familia poco después del limite del Cretácico-Terciario (hace 59 millones de años) y la divergencia entre los principales linajes durante el Paleoceno y el Eoceno (56-30 millones de años atrás).

Beyond *Conus*: phylogenetic relationships of Conidae based on complete mitochondrial genomes

Juan E. Uribe[1], Nicolas Puillandre[2] & Rafael Zardoya[1*]

[1]Museo Nacional de Ciencias Naturales (MNCN-CSIC), José Gutiérrez Abascal 2, 28006, Madrid, Spain

[2] Institut de Systématique, Évolution, Biodiversité ISYEB – UMR 7205 – CNRS, MNHN, UPMC, EPHE, Muséum National d'Histoire Naturelle, Sorbonne Universités, 43 rue Cuvier, CP26, F-75005, Paris, France.

*Corresponding author (rafaz@mncn.csic.es)

ABSTRACT

Understanding how the extraordinary taxonomic and ecological diversity of cone snails (Caenogastropoda: Conidae) evolved requires a statistically robust phylogenetic framework, which thus far is not available. While recent molecular phylogenies have been able to distinguish several deep lineages within the family Conidae, including the genera *Profundiconus*, *Californiconus*, *Conasprella*, and *Conus* (and within this one, several subgenera), phylogenetic relationships among these genera remain elusive. Moreover, the possibility that additional deep lineages may exist within the family is open. Here, we reconstructed with probabilistic methods a molecular phylogeny of Conidae using the newly sequenced complete or nearly complete (mt) mitochondrial genomes of the following nine species that represent all main Conidae lineages and potentially new ones: *Profundiconus teramachii, Californiconus californicus, Conasprella wakayamaensis, Lilliconus sagei, Pseudolilliconus traillii, Conus* (*Kalloconus*) *venulatus*, *Conus* (*Lautoconus*) *ventricosus*, *Conus* (*Lautoconus*) *hybridus*, and *Conus* (*Eugeniconus*) *nobilis*. To test the monophyly of the family, we also sequenced the nearly complete mt genomes of the following three species representing closely related conoidean families: *Benthomangelia* sp. (Mangeliidae), *Tomopleura* sp. (Borsoniidae), and *Glyphostoma* sp. (Clathurellidae). All newly sequenced conoidean mt genomes shared a relatively constant gene order with rearrangements limited to tRNA genes. The reconstructed phylogeny recovered with high statistical support the monophyly of Conidae and phylogenetic relationships within the family. The

genus *Profundiconus* was placed as sister to the remaining genera. Within these, a clade including *Californiconus* and *Lilliconus* + *Pseudolilliconus* was the sister group of *Conasprella* to the exclusion of *Conus*. The phylogeny included a new lineage whose relative phylogenetic position was unknown (*Lilliconus*) and uncovered thus far hidden diversity within the family (*Pseudolilliconus*). Moreover, reconstructed phylogenetic relationships allowed inferring that the peculiar diet of *Californiconus* based on worms, mollusks, crustaceans and fish is derived, and reinforce the hypothesis that the ancestor of Conidae was a worm hunter. A chronogram was reconstructed under an uncorrelated relaxed molecular clock, which dated the origin of the family shortly after the Cretaceous-Tertiary boundary (about 59 million years ago) and the divergence among main lineages during the Paleocene and the Eocene (56-30 million years ago).

INTRODUCTION

With more than 830 described species (WoRMS, 2016), cone snails (Family Conidae, Fleming, 1822 *sensu lato*) constitute a major component of the biodiversity of tropical and subtropical oceans (Tucker and Tenorio, 2013). The species diversity of cones is highest in the Indo-West Pacific region (Röckel et al., 1995) but notably about 10% of the species radiated in the Cape Verde archipelago (Cunha et al., 2005; Duda and Rolan, 2005). While some species show widespread distributions (e.g., *Conus litteratus* Linnaeus, 1758 throughout the Indo-West Pacific region), others are narrowly restricted to an island or a bay (e.g., *Conus verdensis* Trovão, 1979 from Santiago Island in Cape Verde). Cones are found from deep waters to the intertidal zone, associated to rocky shores, coral reefs, and sandy bottoms (Kohn, 1959). These marine gastropods are predatory carnivores feeding mostly on marine worms, snails and fishes (Duda et al., 2001), and have evolved a sophisticated mechanism to capture preys, which are paralyzed thanks to harpoon-like radular teeth coated with a cocktail of toxins produced in a venom gland (Olivera, 2002). Interestingly, recent transcriptomic studies have shown that predation- and defense-evoked venoms are produced in the distal and proximal regions of the venom duct, respectively (Dutertre et al., 2014; Prashanth et al., 2016). Moreover, conotoxins are of important medical and pharmaceutical interest since they are potent and have very specific inhibitors of ion channels in the human brain (Terlau and Olivera, 2004). Reconstructing a statistically robust phylogeny of Conidae is mandatory for understanding how the great species diversity of

the family was generated and addressing other important evolutionary open questions in the group such as the origin of the different diet specializations or how did predation and defense venoms appeared and evolved (Duda et al., 2001; Puillandre et al., 2014a). Moreover, current discovery of pharmacologically important conotoxins could be enhanced and improved by using a concerted discovery strategy that takes into account robustly inferred phylogenetic relationships to target most divergent and poorly studied groups (Holford et al., 2009; Puillandre and Holford, 2010).

All cones share a typical conical shell of different sizes (mm to about 20 cm), often brightly colored, and with diverse banding patterns that is highly appreciated by collectors (Tucker and Tenorio, 2013). The inner walls of the shell are re-absorbed during growth, and this is considered a synapomorphy of the family Conidae (Tucker and Tenorio, 2009). In general, the shell is helpful for species identification but has limited utility for discrimination of higher taxonomic levels, at which other characters such as the shape of the radula and DNA sequences are used (Tucker and Tenorio, 2009; Puillandre et al., 2014a). The family Conidae belongs to the superfamily Conoidea (Caenogastropoda: Neogastropoda) together with closely related families such as e.g., Conorbidae, Raphitomidae, Mangeliidae, Borsoniidae, Clathurellidae, and Mitromorphidae (See Puillandre et al. 2011, and references therein). Traditionally, most authors assumed that the family Conidae

contained only the genus *Conus* (e.g., Röckel et al., 1995; but see Cotton, 1945; Walls, 1978; Da Motta, 1991; Taylor et al., 1993). However, two recent studies have proposed considerable changes to the classification of the family. One study (Tucker and Tenorio, 2009) was based on cladistic analysis of radular teeth and shell characters and proposed to recognize some previously introduced genera in addition to *Conus*, to raise some previously known subgenera to the genus level, and to erect completely new genera. The proposed classification distinguished up to four living families (including Conidae, Conilithidae, Conorbidae, and Taranteconidae) and 86 extant genera. Later, Conorbidae was tentatively maintained as a separate family (Bouchet et al., 2011) and species within Taranteconidae were found to be closely related to *Conus* (*Stephanoconus*) (Watkins et al., 2010). The other study (Puillandre et al., 2014a) was based on probabilistic analyses of three partial mitochondrial (mt) genes and included 330 species belonging to Conidae, Conilithidae and Taranteconidae *sensu* Tucker and Tenorio (2009). The presence of several deep lineages within the analyzed taxa prompted for a new taxonomic classification (that we follow here naming the subgenus only the first time) with a single family Conidae, which included four genera, namely *Californiconus*, *Profundiconus*, *Conasprella*, and *Conus* (Puillandre et al., 2014b). The latter two genera were further subdivided into 11 and 60 subgenera, respectively (Puillandre et al., 2014b). The reconstructed phylogeny showed that *Profundiconus* was the sister group of the remaining Conidae, although without support (and thus questioning the limits of the family; Puillandre et al., 2014a).

Within the remaining taxa, *Californiconus* was the sister group of *Conasprella* and *Conus* (Puillandre et al., 2014a). Therefore, the genera *Conus* and *Conasprella sensu* Puillandre et al. (2014b) more or less corresponded to the families Conidae and Conilithidae *sensu* Tucker and Tenorio (2009), respectively. However, the genera *Profundiconus* and *Californiconus* were excluded from other Conilithidae (*Conasprella*) (Puillandre et al., 2014b). Besides that, major lineages within *Conasprella* and *Conus* were highly congruent between both studies, only differing in their subgeneric (Puillandre et al., 2014b) or generic (Tucker and Tenorio, 2009) status, and on the placement of some species for which there was no radula and/or DNA data available and that were ascribed based on shell characters only (Tucker and Tenorio, 2009). Moreover, the new molecular phylogeny was confirming previous ones (Duda and Kohn, 2005; Biggs et al., 2010) that had already distinguished *Conus californicus*, a "Small Major Clade" (*Conasprella*) and a "Large Major Clade (*Conus*). In addition, different lineages within *Conasprella* (Kraus et al., 2012) and *Conus* (Espiritu et al., 2001; Nam et al., 2009; Kraus et al., 2011) were also recovered in several previous molecular phylogenies.

Here, we aimed to confirm the main deep lineages reported within Conidae (*sensu* Puillandre et al., 2014b) and in particular to define the phylogenetic relationships between these main deep lineages, which were mostly unresolved in published phylogenies (e.g., Puillandre et al., 2014a). To achieve these goals, we used complete or almost complete (without control region) mt genome sequence data, which have proven useful in recovering internal

nodes with high support at this level of divergence or higher in other gastropods (Grande et al., 2008; White et al., 2011; Uribe et al., 2016). Thus far, the only complete mt genomes available for Conidae are those of *Conus* (*Cylinder*) *textile* (Bandyopadhyay et al., 2008); *Conus* (*Gastridium*) *tulipa* (Chen et al., 2015); *Conus* (*Lautoconus*) *borgesi* (Cunha et al., 2009); *Conus* (*Splinoconus*) *tribblei* (Barghi et al., 2015); and *Conus* (*Pionoconus*) *consors* (Brauer et al., 2012). No complete mt genomes are available for other cone snails genera and for related families within Conoidea, and the closest conoideans available are *Xenuroturris cerithiformis* (Turridae; Bandyopadhyay et al., 2006), *Fusiturris similis* (Clavatulidae; (Cunha et al., 2009), and *Oxymeris dimidiata* (Terebridae; Cunha et al., 2009), which some authors place in a different superfamily, Turroidea (Tucker and Tenorio, 2009). Therefore, we sequenced mtDNAs of several species representing the main lineages of Conidae (*Profundiconus*, *Californiconus*, *Conasprella*, and *Conus*), as well as closely related conoidean families (Mangeliidae, Clathurellidae, and Borsoniidae). In addition, we sequenced the mt genomes of two highly divergent species of cones that may represent additional genera (*Lilliconus* and *Pseudolilliconus*). Our aims were: (1) to confirm the previously identified main lineages within cone snails and eventually identify new ones; (2) to reconstruct a robust phylogeny of Conidae that could be used as framework for further evolutionary studies; (3) to assess whether there have been major rearrangements of the mtDNA genome organization among the analyzed conoidean families, and (4) to date main cladogenetic events within Conidae.

MATERIALS AND METHODS

Samples and DNA extraction

The complete list of species analyzed in this study corresponding to families Conidae, Borsoniidae, Mangeliidae and Clathurellidae, is shown in Table 1, along with their respective sampling localities and museum vouchers. Specimens from the MNHN were either found in old collections or newly collected during several recent expeditions (Atimo Vatae in Madagascar, Papua Niugini and Kavieng in Papua New-Guinea). All samples were stored in ethanol 100% and total genomic DNA was isolated from up to 30-50 mg of foot tissue following a standard phenol-chloroform extraction.

PCR amplification and sequencing

Complete or nearly complete (without the control region; see results and discussion) mt genomes were amplified through long PCR using different combinations of conserved primers newly designed in mt *cox1*, *cox3*, *rrnL* and *trnF* genes (Supplementary material 1). The long PCR reactions 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 (10-40 ng) of template DNA, 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: initial denaturing step

at 94°C for 60 s; 45 cycles of denaturing at 98°C for 10 s, annealing at 53°C for 30 s and extending at 68°C for 60 s per kb; final extending step at 68°C for 12 min. In addition, two standard PCR reactions were performed (Supplementary material 1). One used the *rrnL* gene universal primers (Palumbi et al., 1991) to close the gap between long PCR *rrnL* primers, and the other used *cox1*gene universal primers (Folmer et al., 1994) to amplify a fragment, which after Sanger sequencing at the MNHN was used to check that final assemblies corresponded to the correct species. 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 (10-40 ng/µl) 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: initial denaturing step at 94°C for 60 s; 45 cycles of denaturalization at 94°C for 30 s, annealing at 44°C for 60 s and extending at 72°C for 90 s; final extending step at 72°C for 5 m.

Table 1. Mitochondrial (mt) genomes analyzed in this study. Numero (#) of reads and mean (M) coverage of each species.

New mt genomes

Species	Family	Length (bp)	GenBank Acc. No.	Location	Voucher Museum	# reads	M. Coverage
*Californiconus californicus**	Conidae	15444	KX263249	Aqua Hedionda lagoon, Carlsbad, California, USA	MNCN:ADN:86740	15542	693
Conus (*Kalloconus*) *venulatus**	Conidae	15524	KX263250	Boavista Island, Ilheu Sal Rei, Cape Verde	MNCN:ADN:86741	155050	1009
Conus (*Lautoconus*) *ventricosus**	Conidae	15534	KX263251	Faro, Portugal	MNCN:ADN:86742	249121	1619
Conus (*Lautoconus*) *hybridus*	Conidae	15276	KX263252	Dakar, 0-2 m, 14° 45 ' N; 17°32' W	MNHN-IM-2009-18301	54697	355
Conus (*Eugeniconus*) *nobilis*	Conidae	15379	KX263253	Indonesia, NE Flores	MNHN-IM-2009-29800	80526	521
Conasprella wakayamaensis	Conidae	15927	KX263254	Papua Niugini expedition (Papua New-Guinea), st. CP4059, 335 m, 02°38'S; 141°18'E	MNHN-IM-2013-19091	128645	801
Lilliconus sagei	Conidae	15485	KX263255	Atimo Vatae expedition (Madagascar), BS03, 14-18 m, 25°26.4'S; 44°56.1'E	MNHN-IM-2009-31328	27906	177
Profundiconus teramachii	Conidae	15279	KX263256	Papua Niugini expedition (Papua New-Guinea), st. CP3979, 540-580 m, 04°44'S; 146°11'E	MNHN-IM-2013-19686	99314	645
Pygmaeconus traillii	Conidae	14963	KX263257	Kavieng expedition (Papua New-Guinea), st. KB8, 13m, 02°33,2'S; 150°48,2'E	MNHN-IM-2013-47771	34311	636
Benthomangelia sp.	Mangeliidae	15034	KX263258	Papua Niugini expedition (Papua New-Guinea), st. CP4024, 420-490 m, 05°22'S; 145°48'E	MNHN-IM-2013-09858	18107	119
Tomopleura sp.	Borsoniidae	15182	KX263259	Papua Niugini expedition (Papua New-Guinea), st. CP4023, 340-385 m, 05°22'S; 145°48'E	MNHN-IM-2013-09849	30684	271
Glyphostoma sp.	Clathurellidae	13370	KX263260	Papua Niugini expedition (Papua New-Guinea), st. CP4065, 380, 03°19'S; 143°01'E	MNHN-IM-2013-19173	37382	277

GenBank mt genomes

Species	Family	Length (bp)	GenBank Acc. No.	Reference
Conus (*Lautoconus*) *borgesi*	Conidae	15536	NC_013243	Cunha et al., 2009
Conus (*Pionoconus*) *consors*	Conidae	16112	NC_023460	Brauer et al., 2012
Conus (*Cylinder*) *textile*	Conidae	15562	NC_008797	Bandyopadhyay et al., 2008
Conus (*Splinoconus*) *tribblei*	Conidae	15570	KT199301	Barghi et al., 2015
Conus (*Gastridium*) *tulipa*	Conidae	16599	NC_027518	Chen et al., 2015
Fusiturris similis	Clavatulidae	15595	NC_013242	Cunha et al., 2009
Xenuroturris cerithiformis	Turridae	15380	NC_008098	Bandyopadhyay et al., 2006
Oxymeris dimidiata	Terebridae	16513	NC_013239	Cunha et al., 2009

**Complete genomes*

Long-PCR products were purified by ethanol precipitation. Amplified fragments from the same mt genome were pooled together in equimolar concentrations and subjected to massive parallel sequencing. For each conoidean mt genome a separate indexed library was constructed using the NEXTERA XT DNA library prep Kit (Illumina, San Diego, CA, USA) at AllGenetics (A Coruña, Spain). Each of the libraries contained in addition mt genomes of unrelated animals (e.g., snakes, spiders) from different projects. The indexed libraries were run in a single lane in an Illumina HiSeq2000 (100 Pair-ended) at Macrogen (Seoul, Korea).

Genome assembly and annotation

Reads were sorted according to their indexes, and the assembly of the different mt genomes was performed in the TRUFA webserver (Kornobis et al., 2015). Briefly, adapters were removed using SeqPrep (StJohn, 2011), quality of the reads was checked using FastQC v.0.10.1 (Andrews, 2010), and raw sequences were trimmed and filtered out according to their quality scores using PRINSEQ v.0.20.3 (Schmieder and Edwards, 2011). Filtered reads were used for *de novo* assembly of each mt genome using TRUFA default settings (minimum contig length; 200; sequence identity threshold: 0.95) only retaining contigs with a minimum length of 3kb. These contigs were finally overlapped in Sequencher 5.0.1 to render the different complete or nearly complete mt genomes included within

each index (the one belonging to a conoidean species and those belonging to a snake or a spider). In order to estimate mean coverage, each assembled conoidean mt genome was used as reference to map the original (raw) reads with a minimum identity of 100% using Geneious® 8.0.3.

The newly determined mt genomes were annotated using Geneious® 8.0.3 by setting a limit of nucleotide identity of 75% to previously reported conoidean mt genomes (i.e., *C. textile, C. borgesi, C. consors, F. similis, X. cerithiformis,* and *O. dimidiata*). Annotations of the 13 mt protein-coding genes were corroborated manually identifying the corresponding 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), which infer cloverleaf secondary structures (with a few exceptions that were determined manually). The ribosomal RNA (rRNA) genes were identified by sequence comparison with previously reported conoidean mt genomes, and assumed to extend to the boundaries of adjacent genes (Boore et al., 2005). GenBank accession numbers of each mt genome are provided in Table 1.

Sequence alignment

The newly sequenced complete or nearly complete mt genomes were aligned with all orthologous conoidean mt genomes available in NCBI (Table 1). Two sequence data sets were constructed and analyzed: the first data set (hereafter referred to as the Conidae data

set) was aimed to test the monophyly of Conidae and included main lineages within the family as well as closely related conoidean families. Three species of less related conoideans were selected as outgroup taxa following Puillandre et al., (2011): *F. similaris* (Clavatulidae); *X. cerithiformis* (Turridae), and *O. dimidiata* (Terebridae). This data set included the deduced amino acid sequences of the 13 mt protein coding genes and the nucleotide sequences of the two rRNA genes. The second data set (hereafter referred to as the *Conus* data set) was aimed to test the internal phylogenetic relationships of *Conus*. This data set included newly determined and previously published *Conus* species and it was rooted with *Conasprella wakayamaensis* and *Californiconus californicus*. The data set included 13 mt protein-coding genes and two rRNA genes, both analyzed at the nucleotide level. Phylogenetic analyses of the protein-coding genes at the amino acid and nucleotide levels in the Conidae and *Conus* data sets, respectively, was aimed at maximizing phylogenetic information (by selecting the appropriate levels of sequence variation) as each data set was addressing different taxonomic questions (see discussion). In order to construct these two data sets, the deduced amino acid sequences of the 13 mt protein-coding genes were aligned separately and used to guide the alignment of the corresponding nucleotide sequences with Translator X (Abascal et al., 2010). Nucleotide sequences of the mt rRNA genes were aligned separately using MAFFT v7 (Katoh and Standley, 2013) with default parameters. Ambiguously aligned positions were removed using Gblocks, v.0.91b (Castresana, 2000) with the

following settings: minimum sequence for flanking positions: 85%; maximum contiguous non-conserved positions: 8; minimum block length: 10; gaps in final blocks: no. Finally, the different single alignments were concatenated into the two data matrices using Geneious® 8.0.3.

Phylogenetic analyses

Phylogenetic relationships of family Conidae and genus *Conus* were inferred using maximum likelihood (ML; Felsenstein, 1981) and Bayesian inference (BI; Huelsenbeck and Ronquist, 2001). ML analyses were conducted with RAxML v7.3.1 (Stamatakis, 2006) using the rapid hill-climbing algorithm and 10,000 bootstrap pseudoreplicates (BP). BI analyses were conducted using MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003), 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, as determined using Tracer v1.6 (Rambaut and Drummond, 2007). Node support was assessed based on Bayesian Posterior Probabilities (BPP).

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 Bayesian Information Criterion (BIC; Schwarz, 1978). For the protein-coding genes of the Conidae data set (analyzed at the amino acid level) the partitions tested were: all genes grouped; all genes separated (except nad4/ 4L and atp6/8); genes grouped by enzymatic complexes (*nad*, *cox*, *atp*, *cob*; see Supplementary Material 2 for selected best fit partitions and models). For the protein-coding genes of the *Conus* data set, which were analyzed at the nucleotide level, the partitions tested were: all genes grouped; all genes separated (except nad4/ 4L and atp6/8); genes grouped by subunits (see Supplementary Material 2). In addition, these three partitions schemes were tested taking into account separately the three codon positions). The rRNA genes (analyzed at the nucleotide level) in both data sets were tested separately with two different schemes, as genes separated or combined.

Estimation of divergence times

The program BEAST v.1.7 (Drummond and Rambaut, 2007) was used to perform a Bayesian estimation of divergence times among major conoidean lineages based on the mt amino acid data set. An uncorrelated relaxed molecular clock was used to infer branch lengths and nodal ages. The tree topology was set based on a combination of the Conidae and *Conus* trees. For the clock model, the lognormal relaxed-clock model was selected, which allows rates to vary among branches without any a priori assumption of autocorrelation between adjacent branches. For the tree prior, a

Yule process of speciation was employed. The partitions selected by Partition Finder Protein (see above) were applied. The final Markov chain was run twice for 100 million generations, sampling every 10,000 generations and the first 10 million was discarded as part of the burn-in process, according to the convergence of chains checked with Tracer v.1.5. (Rambaut and Drummond, 2007). The effective sample size of all the parameters was above 200.

The posterior distribution of the estimated divergence times was obtained by specifying two calibration points as priors for divergence times of the corresponding splits. Fossils provided hard minimum bounds (offset) and mean and standard deviations were chosen so that the 95% probability limit corresponds to a soft maximum bound. For the divergence of Conidae, a calibration point was set at a minimum of 55 million years ago (Mya) with a 95% upper limit of 58.1 MYA (lognormal distribution, offset: 55; mean: 1; standard deviation: 1) based on the oldest known fossils of *Hemiconus rouaulti* (France) and *Hemiconus concinnus* (England) that likely belong to the stem group of the family Conidae (Tucker and Tenorio, 2009) and were documented from the Lower Eocene (Kohn, 1990). A second calibration point was set at the divergence time between *C. ventricosus* and *C. borgesi*. Fossils of *C. (Lautoconus) ventricosus* become recognized in the Middle-Lower Miocene (16.4 to 20.5 Mya) of Cuenca de Piemonte (Italy) (Sacco, 1893). This interval coincides with the inferred origin of Cape Verde cone snails and the age of the archipelago (Cunha et al., 2005). Therefore, a normal distribution (recommended for inferred a secondary calibrations and biogeographical datings; Ho and

Phillips, 2009) was applied. The 95% upper and lower limits were set to 21 and 16 MYA, respectively (mean: 18.5; standard deviation: 1.5).

RESULTS

Sequencing and assembly

Within Conidae, the mt genomes of *C. californicus*, *Conus* (*Kalloconus*) *venulatus*, and C. *ventricosus* were determined complete whereas those of *Conus* (*Lautoconus*) *hybridus*, *Conus* (*Eugeniconus*) *nobilis* (subspecies *victor*), *C. wakayamaensis*, *Lilliconus sagei*, *Profundiconus teramachii*, and *Pseudolilliconus traillii* lacked the control region because it could not be amplified. In addition, the nearly complete (without control region) mt genomes of *Benthomangelia* sp. (Mangeliidae), *Tomopleura* sp. (Borsoniidae), and *Glyphostoma* sp. (Clathurellidae) were also amplified and sequenced. The number of reads, mean coverage, and length of each mt genome are provided in Table 1. The mt genomes of *C. californicus* and *C. ventricosus* received the minimum (15,542) and maximum (249,121) reads, respectively. The minimum (119x) and maximum (1,619x) coverage corresponded to *Benthomangelia* sp. and *C. ventricosus*, respectively.

Structural features and mitochondrial organization

The newly determined mt genomes had the usual 13 protein coding, 2 rRNA, and 22 tRNA genes reported in other animal mt genomes (see annotation and main features of each of these mt genomes in Supplementary Material 3). In few instances, the control region between *trnF* and *cox3* genes was also amplified allowing the completion of the mt genome. All but two of the analyzed conoidean mt genomes conformed to the consensus genome organization described for Caenogastropoda (Osca et al., 2015) with most genes encoded by the major strand and only a cluster of tRNA genes (MYCWQGE) and the *trnT* gene encoded by the minor strand (Fig. 1). The only exceptions were the mt genome of *L. sagei*, which showed the translocation of the *trnL*(uag) and *trnL* (uaa) genes and the inversion and translocation of the *trnT* gene, as well as the mt genomes of *P. traillii* and *Tomopleura* sp., which showed the translocation of the *trnT* gene (Fig.1). In addition, we were not able to find the *trnR* gene of the mt genome of *L. sagei* in its usual position (within the cluster KARNI), but we cannot discard that it might have been translocated near to the control region, which could not be amplified in this mt genome (Fig. 1).

Phylogenetic relationships of Conidae

The molecular phylogeny of Conidae was reconstructed based on the Conidae data set using probabilistic methods (Fig. 2). The final matrix was 5870 positions long. Both, ML (-lnL = 57997.47) and BI

(-lnL = 59051.42 for run1; -lnL = 59051.06 for run2) arrived at almost identical topologies (Fig. 2) only differing in the internal relationships within *Conus.* The reconstructed phylogeny recovered Borsoniidae + Clathurellidae as sister group to Mangeliidae + Conidae, although both groupings received moderate and low statistical support, respectively (Fig. 2). The monophyly of Conidae received strong statistical support (1 BPP, 90% BP; Fig. 2). Within Conidae, *Profundiconus* was recovered as sister group of the remaining members of the family. Within the latter, a clade including *Californiconus* and *Lilliconus* + *Pseudolilliconus* was the sister group of *Conasprella* to the exclusion of *Conus* (Fig. 2). All recovered phylogenetic relationships within Conidae received strong support (Fig. 2) except those within *Conus*.

***Phylogenetic relationships of* Conus**

In order to further determine phylogenetic relationships within the genus *Conus*, a second alignment named Conus data set was analyzed with probabilistic methods (Fig. 3). The final matrix was 13473 positions long. Both ML (-lnL = 69568.03) and BI (-lnL = 69594.80 for run1; -lnL = 69592.68 for run2) arrived at fully resolved phylogenetic trees with all nodes strongly supported (above 70% BP and 0.99 BPP; Fig. 3). Among *Conus* studied species, *C. tribblei* was recovered as sister group of the remaining, which were organized into two sister clades. One clade included *C. consors* + *C. tulipa* as sister group of *C. textile* + *C. nobilis*. The other clade included *C. venulatus* as sister group of a clade including *C. hybridus* and *C. ventricosus* + *C. borgesi* (Fig. 3).

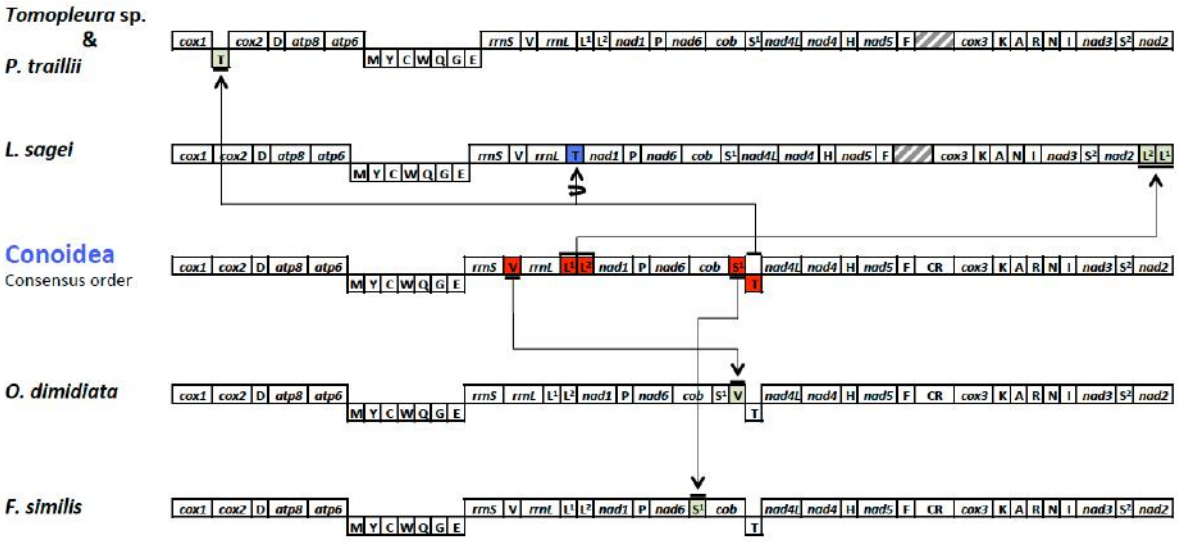


Figure 1. Mitochondrial gene orders of conoidean mitochondrial genomes. The consensus genome organization is shown as well as known exceptions. The genes encoded in the major and minor strands are shown in the top and bottom lines, respectively. Gene rearrangements (restricted to tRNA genes) are indicated by arrows. Translocated genes are in green. A gene both translocated and inverted is in blue. Striped boxes indicate regions not sequenced (note that the *trnR* gene is missing in *L. sagei*).

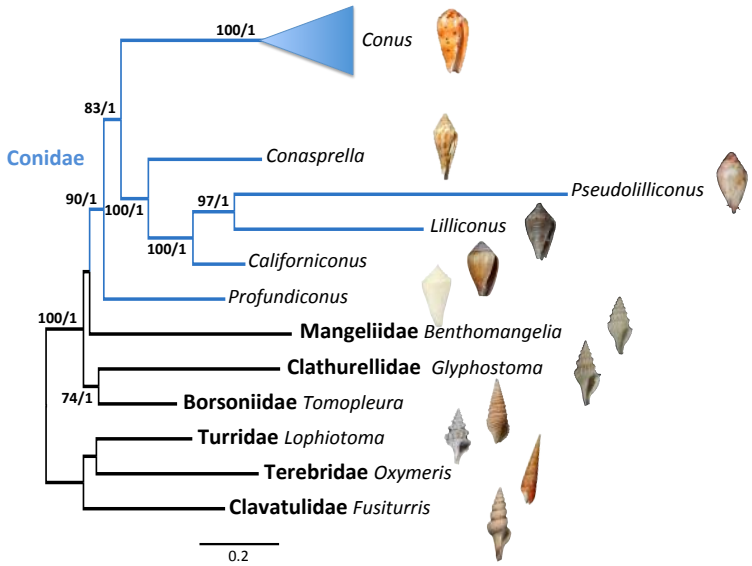
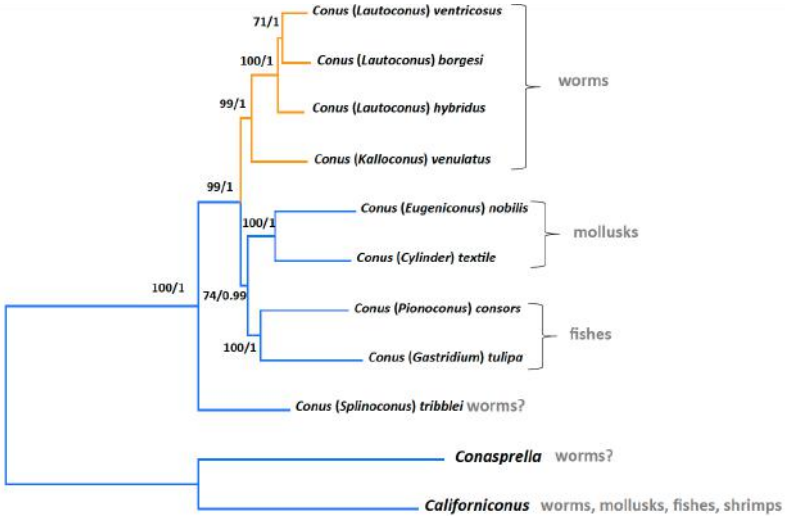


Figure 2. Phylogenetic relationships of Conoidea based on complete mt genomes. The reconstructed ML phylogram using Terebridae, Turridae and Clavtulidae as outgroup is shown. The family Conidae is indicated in blue. Numbers at nodes are statistical support values for ML (bootstrap proportions in percentage)/ BI (posterior probabilities). Drawings are taken from (Puillandre et al., 2014a).



←Figure 3. Phylogenetic relationships within *Conus* based on complete mt genomes. The reconstructed ML phylogram using *Californiconus* and *Conasprella* as outgroup is shown. Numbers at nodes are statistical support values for ML (bootstrap proportions in percentage)/ BI (posterior probabilities). The distributions of the taxa (Indo-Pacific region in blue; Western Atlantic Ocean and Mediterranean in orange) and their diet are indicated.

Divergence times

Major cladogenetic events within Conoidea were dated using an uncorrelated relaxed molecular clock model, which was calibrated with several European fossils belonging to the stem and crown groups of Conidae. The origin of the conoidean families closely related to Conidae is dated at a mean of 67 (84-57, credible intervals) Mya, quite close in geological times to the origin of the family Conidae itself about 59 (73-55) Mya (Fig. 4). The branching of *Profundiconus* is estimated to have occurred around 56 (70-49) Mya and the split between the lineage leading to extant *Conus* and the clade containing *Californicus*, *Conasprella*, *Lilliconus*, and *Pseudolilliconus* was dated at 51 (64-44) Mya (Fig. 4). Divergence among these latter four genera occurred successively between 46-30 (59-22) Mya. The radiation of the analyzed *Conus* species was estimated to have occurred between 24-15 (30-12) Mya (Fig. 4).

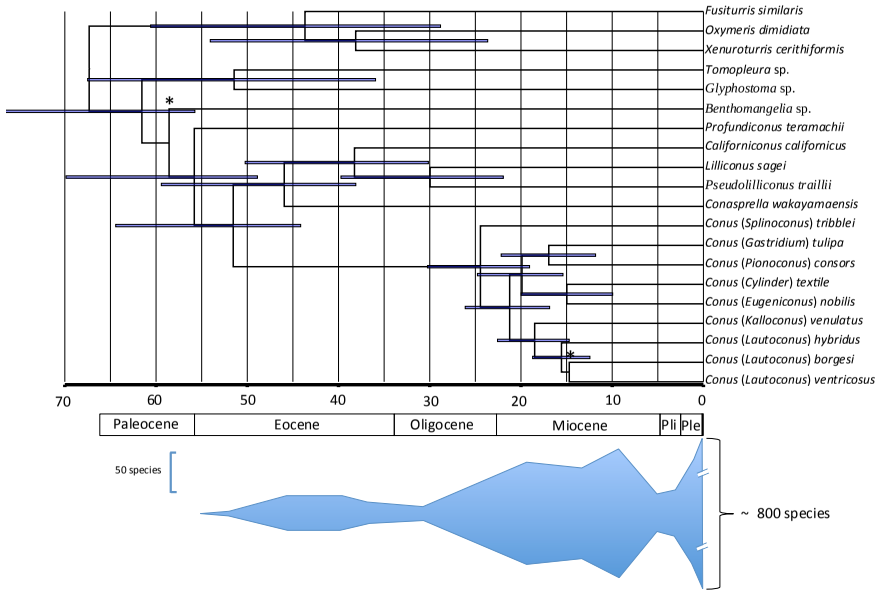


Figure 4. Chronogram with age estimates of major divergence events among conoideans, based on the Conidae data set, and using Bayesian relaxed dating methods (BEAST). Horizontal bars represent 95% credibility intervals of relevant nodes, and calibration constraints are indicated with an asterisk on the corresponding nodes. Dates (and credibility intervals) are in millions of years. A geological table with periods is shown as well as species diversity of cone snails in the fossil record (modified from (Kohn, 1990).

DISCUSSION

Thanks to the combination of long PCR and massive sequencing techniques, we were able to add in the present work up to12 new mt genomes to the catalogue of conoidean mt genomes. Not only we more than double the number of available mt genomes for this superfamily of Caenogastropoda, but also we provide a better representation of the diversity of the superfamily by adding the first representatives of five genera within Conidae and three closely related families. Both, number of reads per mt genome and final coverage were high, with a direct relationship between both parameters except in the case of the mt genomes of *Californiconus* and *Pseudolilliconus*, which showed higher coverage than expected. The presence of reads with the same index and in the same lane corresponding to distantly related animal species did not interfere in the correct assembly of each conoidean mt genome as assessed by empirical PCR amplification and sequencing of the *cox1* gene of each analyzed species. We were able to complete only three out of the 12 mt genomes. Completed mt genomes showed short control regions and interestingly, the coverage in these regions was much lower than average (despite being part of a longer PCR fragment in equimolar concentration). It is likely that longer and more complex (with secondary structures) control regions in the remaining mt genomes prevented *Taq* polymerase for completing the PCR reactions in some species. In those cases, outward primers were designed in the *trnF* and *cox3* genes at the boundaries of the control region (see Supplementary Material 1).

Gene order evolution

The mt genomes of mollusks, and of gastropods in particular, are known for having relatively high rates of gene rearrangement (Grande et al., 2008; Stöger and Schrödl, 2013). Major changes in mt genome organization including translocations and inversions of protein coding and/or rRNA genes normally occur between main lineages of gastropods (e.g., Patellogastropoda or Heterobranchia versus other gastropod lineages; Grande et al., 2008) or in particular groups within main lineages (e.g., superfamily Vermetoidea witihin Caenogastropoda; Rawlings et al., 2010). Interestingly, these high rates of rearrangement are normally associated with high mutational rates, leading to long branches in phylogenetic trees (Stöger and Schrödl, 2013; Osca et al., 2015; Uribe et al., 2016). Nevertheless, for the majority of groups and species within a main gastropod lineage, gene order is relatively stable and rearrangements are restricted to tRNA genes, if any (Grande et al., 2008). Hence, it is possible to reconstruct a consensus gene order for the hypothetical ancestor of the different main gastropod lineages (Grande et al., 2008; Stöger and Schrödl, 2013; Osca et al., 2014; Osca et al., 2015; Uribe et al., 2016). The gene order of the 12 mt genomes here sequenced generally conforms to the consensus genome organization for Caenogastropoda and is identical to the one inferred for Neogastropoda (Cunha et al., 2009; Osca et al., 2015). Among previously published conoidean mt genomes, it was reported the translocation of the *trnV* and *trnS* in *O. dimidiata* and

F. similis, respectively (Cunha et al., 2009). Here, the *Tomopleura* sp. mt genome shows a translocation of the *trnT*, which is normally found next to the *trnS* (uga) gene and encoded by the minor strand, to a location between the *cox1* and *cox2* genes. The rearrangement of this tRNA gene is relatively frequent among caenogastropods (Osca et al., 2015) and occupies the same position in *P. traillii*. In addition, the mt genome of *L. sagei* presents a translocation and inversion of the *trnT* gene, which is found next to the *rrnL* gene and encoded by the major strand, in a position were normally the *trnL* (uag) and *trnL* (uaa) genes are found. In this mt genome, however, the two *trnL* genes have moved next to the *cox2* gene. Interestingly, both events seem to be connected because at the same position where the *trnT* is found, the minor strand could putatively encode for a *trnL* (uaa) gene (see Supplementary Material 4), indicating that the *trnT* and the reverse complementary *trnL (uaa)* gene sequences are very similar. Moreover, between the two *trnL* genes there is space for the coding of a *trnT* gene in the major strand (see Supplementary Material 4), which could be the remnant of an ancient duplication. In addition, we were not able to detect the *trnR* gene, in the otherwise highly conserved KARNI cluster. This missing gene might have moved next to the control region, which could not be sequenced. Finally, it is worth mentioning that in many gastropod mt genomes high rates of rearrangement and of substitution rates are normally correlated (Rawlings et al., 2010; Stöger and Schrödl, 2013; Osca et al., 2014; Osca et al., 2015). However, here this correlation does not hold. The mt genomes of *Tomopleura* and *Pseudolilliconus* have the same gene order, but

only the latter genus has a very long branch in the phylogenetic tree, much longer than that of *Lilliconus*, whose mt genome has more rearrangements than any other (and even in this case only associated to minor tRNA gene rearrangements).

Phylogenetic relationships of Conidae

The hyperdiverse superfamily Conoidea has been the subject of recent molecular phylogenetic studies (Puillandre et al., 2008; Puillandre et al., 2011; Puillandre et al., 2014a) that supported some morphology-based classifications (Taylor et al., 1993) and allowed discerning the closest families to Conidae, i.e., Conorbidae, Raphitomidae, Mangeliidae, Borsoniidae, Clathurellidae, and Mitromorphidae. These molecular phylogenies were based on the concatenation of partial mt genes and were unable to resolve phylogenetic relationships among these families, and thus determining the sister group of Conidae. A clade including Conorbidae and Borsoniidae was tentatively recovered as sister group of Conidae but without statistical support (Puillandre et al., 2011). In the phylogeny here reconstructed based on complete mt genomes, the Mangeliidae were recovered as sister group of Conidae but this relationship showed low statistical support impeding the resolution of this long-standing question. Here, we added a considerable amount of sequence data (mt genomes) in trying to gain further resolution in this part of the Conoidea tree but without success. However, our data set was biased towards representatives of the family Conidae. Hence, in future studies, it

would be important to increase taxon representation within closely related families, as well as include missing important families such as Conorbidae, Raphitomidae, and Mitromorphidae. Moreover, the possibility nowadays of obtaining a considerable number of nuclear loci using next-generation sequencing techniques opens a potent approach to increase phylogenetic resolution. In any case, it is clear from the reconstructed phylogenetic trees that the lengths of internal nodes connecting these families are rather short, which may indicate and ancient radiation, and therefore that achieving high statistical support and final resolution of these phylogenetic relationships will be challenging.

The monophyly of the family Conidae was highly supported in the reconstructed phylogeny, as were relationships among its main deep lineages. In this case, we had a complete representation of main lineages and even new ones, allowing us to reach stronger conclusions. The genus *Profundiconus* was recovered as sister group to the remaining members of Conidae in agreement with previous molecular phylogenies (Puillandre et al., 2011; Puillandre et al., 2014a) but here showing high statistical support. Phylogenetic relationships among the remaining Conidae differed with respect to previous studies. Here, *Conus* was recovered as the sister group of a clade containing *Conasprella* as sister group of *Californiconus* and *Lilliconus* + *Pseudolilliconus*. In previous molecular phylogenies (Puillandre et al., 2011; Puillandre et al., 2014a), *Californiconus* was recovered as sister group of *Conasprella* + *Conus*, with low BP support (50-63 %) in ML and relatively high BPP support in BI (0.96-0.98). The differences

between the present study and previous ones are the increased number of analyzed positions, the use of amino acids, which show a better phylogenetic information/ noise ratio at deeper nodes due to lower saturation levels, and the inclusion of new lineages of Conidae that proved to be highly divergent. Interestingly, a close relationship between *Californiconus* and *Conasprella* (*Lilliconus* and *Pseudolilliconus* were not included in the study) was already suggested by Tucker and Tenorio (2009). The reconstructed phylogeny is statistically robust within Conidae and serves as a framework for studying evolutionary processes associated with the diversification of the family. All members of Conidae are presumed vermivorous except *C. californicus,* which has a wide diet based on worms, mollusks, crustaceans and fishes (Biggs et al., 2010), and certain derived groups of *Conus* that feed on fishes or mollusks. The strongly supported phylogenetic position of *C. californicus* deeply nested within the Conidae tree reinforces the hypothesis that the ancestor of Conidae was a hunter of polychaete worms (Puillandre et al., 2014a). The 16 extant species belonging to *Profundiconus* live in the deep sea in the Indo-Pacific region (Tenorio and Castelin, 2016). The relative phylogenetic position of this genus within the family Conidae suggests that the group represents an early offshoot that has survived since the middle Eocene.

Phylogenetic relationships of Conus

The reconstructed phylogeny based on the Conidae data set lacked resolution within *Conus*. This was likely due to low levels of

variation in the amino acids at this hierarchical taxonomic level. In order to maximize phylogenetic information, a second data set was constructed with protein coding genes analyzed at the nucleotide level. The *Conus* data set rendered a fully resolved phylogeny with high statistical support in all nodes. The reconstructed phylogenetic relationships are fully congruent with those recovered in a previous molecular phylogeny with an extended taxon sampling (Puillandre et al., 2014a). The presumed vermivorous species *C. tribblei* from the West Pacific and Indian Oceans was recovered as sister group of the remaining taxa. Within this, two main groups were recovered, corresponding to Indo-Pacific and Western Atlantic- Mediterranean species, respectively. Within the Indo-Pacific clade, two species, *C. consors* and *C. tulipa*, feeding on fish clustered together as sister group of two species, *C. textile* and *C. nobilis*, feeding on mollusks (Tucker and Tenorio, 2009; Puillandre et al., 2014a). All species in the Western Atlantic- Mediterranean clade are worm hunters. Obviously, the present phylogenetic tree has only a minor representation of the species diversity of the genus (nine out of 800 species) and is biased in terms of taxonomy (three of the species belong to the same subgenus), distribution (half of the species are from the Atlantic Ocean), and life history traits (about half of the species are fish- or mollusk-hunters) when compared with the genus as a whole. Therefore, we limited our interpretation of the results to character states at the tips of the tree and refrained from performing proper ancestral character-state reconstructions, which would be meaningless at this moment. Nevertheless, the present work emphasizes that complete mt genomes are a very promising tool for

achieving important levels of resolution within *Conus*, and that a more complete data set will certainly help a better understanding of the evolutionary processes (diet and conotoxin evolution, biogeography) that led to the extraordinary diversity encompassed by the genus.

Divergence times and taxonomic levels within family Conidae

The reconstructed time tree using a relaxed molecular clock model dated the origin of the family Conidae in the Paleocene, shortly after the Cretaceous-Tertiary boundary (about 59 Mya, at the Danian/ Selanian transition), which is right before the earliest fossils of cone snails are documented. According to the chronogram, the first burst of cladogenetic events within the family Conidae occurred successively during the Paleocene and Eocene and corresponded to the origin of the major lineages (genera) in parallel to the appearance of closely related conoidean families (Fig. 4; Kohn, 1990). The fossil record would suggest that some of these conoidean families may have appeared before (Powell, 1942), but this would need to be confirmed with a full revision of the different reported fossils and confirmation of their ascriptions to the different families. The diversity of Conidae increased steadily during the Oligocene until a major radiation in the Indo-Pacific region occurred in the Miocene corresponding to the appearance of subgenera within *Conus* (Fig. 4; Kohn, 1990). However, the analyzed *Conus* species correspond to clades that appeared relatively late during the evolution of the genus (Puillandre et al.,

2014a): the inclusion of species belonging to the subgenera *Fraterconus*, *Stephanoconus*, *Strategoconus*, *Klemaeconus*, and *Turriconus*, which supposedly diverged before (Puillandre et al., 2014a), will likely push back our estimates for the original radiation of *Conus*. Given our taxon sampling we could not date the most recent radiation in the family corresponding to the appearance of extant species in the Pleistocene (Kohn, 1990; the magnitude of this radiation, >800 living species versus 100-150 fossils species at the maximum diversity 10 Mya, directly depends on how complete and unbiased is the fossil record). Another important radiation of *Conus* occurred locally during the middle-lower Miocene in the Cape Verde archipelago shortly after the emergence of these volcanic islands (Cunha et al., 2005).

In an ultrametric tree, the distances from the root to every branch tip are equal, the length of the branches is proportional to the time of divergence, and hence, branch length of the different lineages can be roughly compared in order to provide a criterion for taxonomic level delimitation above species (Johns and Avise, 1998). The hierarchical level of the main clades within Conidae has been an important source of conflict between morphological- (Tucker and Tenorio, 2009) and molecular-based (Puillandre et al., 2014b) classifications. By comparing the branch lengths of the different accepted families within Conoidea (Puillandre et al., 2011), it seems that earlier lineages within Conidae could have appeared before some other related families of Conoidea. However, the fossil record suggests that some families of “turrids” (i.e., Conoidea except Conidae and Terebridae) likely appeared before the cone snails

(Powell, 1942). Furthermore, while cone snails are represented by multiple lineages, and several species within *Conus*, in our phylogenetic analyses, the closely related families are represented by only one species each, and even some families, also suggested as closely related to cone snails (Puillandre et al. 2011), e.g., Clathurellidae, Mitromorphidae, Raphitomidae, are absent. Therefore, more data, and in particular a better coverage of the Conoidea diversity, together with calibration points for non-cone-snails Conoidea, are needed to provide a time-calibrated phylogeny that could be used to discuss taxonomic ranks.

CONCLUSIONS

The ancient radiation at the origin of conoidean families combined with the extraordinary species diversity within Conidae have hindered past attempts of resolving the phylogeny of the family based on concatenated partial mt genes. Here, up to 12 complete or nearly complete (without control region) mt genomes of all main lineages of Conidae and certain selected closely related conoidean families were sequenced and used for phylogenetic analyses. The monophyly of the family including genus *Profundiconus* was recovered with high support, and high resolution of phylogenetic relationships was achieved not only among all genera, but also among an abridged representation of species within the most diverse genus (*Conus*). Our results indicate that complete mt genomes are a very promising phylogenetic tool to reconstruct a statistically robust phylogeny of the family. This approach could be complemented

with the development (using next generation sequencing techniques) of nuclear markers, which could be particularly useful for resolving deeper phylogenetic relationships i.e., those among conoidean families. Altogether, these robust molecular phylogenies would allow setting the needed framework to further our understanding of the evolutionary processes that generated and maintain the remarkable taxonomic and ecological diversity of cone snails.

ACKNOWLEDGEMENTS

The MNHN Indo-Pacific material originates from a series of expeditions and workshops, conducted in the context of the "Our Planet Reviewed" programme with Pro-Natura International (ATIMO VATAE to South Madagascar in partnership with Université de Tuléar; PAPUA NIUGINI 2012 and KAVIENG 2014 to Papua New Guinea, in partnership with University of Papua New Guinea and the National Fisheries College). The organizers thank the Total Foundation, Prince Albert II of Monaco Foundation, Stavros Niarchos Foundation, Fondation EDF, Vinci Entrepose Contracting and Richard Lounsbery Foundation for their support to these expeditions. We are indebted to José Utge, Agnès Dettaï and Laetitia Aznar (MNHN, Paris), and Ana Echeverry (MNCN) for technical assistance in the laboratory. Several specimens of *C. californicus* were kindly donated by Rick Negus. We are 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), by the "Service de Systématique Moléculaire" (UMS 2700 CNRS-MNHN), by the ATM ("Action transversal du Muséum") Emergences of the MNHN and by the project CONOTAX, funded by the French Agence Nationale de la Recherche (ANR-13-JSV7-0013-01 to NP). Access of JEU to the MNHN collections and facilities was funded by the EU Synthesys Project.

REFERENCES

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

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

Bandyopadhyay, P.K., Stevenson, B.J., Cady, M.T., Olivera, B.M., Wolstenholme, D.R., 2006. Complete mitochondrial DNA sequence of a Conoidean gastropod, *Lophiotoma* (*Xenuroturris*) *cerithiformis*: Gene order and gastropod phylogeny. Toxicon 48, 29-43.

Bandyopadhyay, P.K., Stevenson, B.J., Ownby, J.-P., Cady, M.T., Watkins, M., Olivera, B.M., 2008. The mitochondrial genome of *Conus textile*, *coxI–coxII* intergenic sequences and Conoidean evolution. Mol. Phylogenet. Evol. 46, 215-223.

Barghi, N., Concepcion, G.P., Olivera, B.M., Lluisma, A.O., 2015. Characterization of the complete mitochondrial genome of *Conus tribblei* Walls, 1977. Mitochondrial DNA, 1-2.

Biggs, J.S., Watkins, M., Puillandre, N., Ownby, J.-P., Lopez-Vera, E., Christensen, S., Moreno, K.J., Bernaldez, J., Licea-Navarro, A., Corneli, P.S., 2010. Evolution of *Conus* peptide toxins: analysis of *Conus californicus* Reeve, 1844. Mol. Phylogenet. Evol. 56, 1-12.

Boore, J.L., Macey, J.R., Medina, M., 2005. Sequencing and comparing whole mitochondrial genomes of animals. Methods Enzymol. 395, 311 - 348.

Bouchet, P., Kantor, Y.I., Sysoev, A., Puillandre, N., 2011. A new operational classification of the Conoidea (Gastropoda). J. Mollus. Stud. 77, 273-308.

Brauer, A., Kurz, A., Stockwell, T., Baden-Tillson, H., Heidler, J., Wittig, I., Kauferstein, S., Mebs, D., Stöcklin, R., Remm, M., 2012. The mitochondrial genome of the venomous cone snail *Conus consors*. PLoS. ONE 7, e51528.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540-552.

Chen, P.-W., Hsiao, S.-T., Huang, C.-W., Chen, K.-S., Tseng, C.-T., Wu, W.-L., Hwang, D.-F., 2015. The complete mitochondrial genome of *Conus tulipa* (Neogastropoda: Conidae). Mitochondrial DNA, 1-2.

Cotton, B.C., 1945. A catalogue of the cone snails (Conidae) in the South Australian Museum. Rec. South Aust. Mus 8, 229–280.

Cunha, R., Grande, C., Zardoya, R., 2009. Neogastropod phylogenetic relationships based on entire mitochondrial genomes. BMC Evol. Biol. 9, 210.

Cunha, R.L., Castilho, R., Rüber, L., Zardoya, R., 2005. Patterns of cladogenesis in the venomous marine gastropod genus *Conus* from the Cape Verde Islands. Syst. Biol 54, 634-650.

Da Motta, A.J., 1991. A systematic classification of the gastropod family Conidae at the generic level. La Conchiglia, Roma.

Drummond, A., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7, 214.

Duda, T.F., Kohn, A.J., 2005. Species-level phylogeography and evolutionary history of the hyperdiverse marine gastropod genus *Conus*. Mol. Phylogenet. Evol. 34, 257-272.

Duda, T.F., Kohn, A.J., Palumbi, S.R., 2001. Origins of diverse feeding ecologies within *Conus*, a genus of venomous marine gastropods. Biol. J. Linn. Soc. 73, 391-409.

Duda, T.F., Rolan, E., 2005. Explosive radiation of Cape Verde *Conus*, a marine species flock. Mol. Ecol. 14, 267-272.

Dutertre, S., Jin, A.-H., Vetter, I., Hamilton, B., Sunagar, K., Lavergne, V., Dutertre, V., Fry, B.G., Antunes, A., Venter, D.J., Alewood, P.F., Lewis, R.J., 2014. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. Nat. Comm. 5, 3521.

Espiritu, D.J.D., Watkins, M., Dia-Monje, V., Cartier, G.E., Cruz, L.J., Olivera, B.M., 2001. Venomous cone snails: molecular

phylogeny and the generation of toxin diversity. Toxicon 39, 1899-1916.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 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. Mol. Mar. Biol. Biotechnol. 3, 294-299.

Grande, C., Templado, J., Zardoya, R., 2008. Evolution of gastropod mitochondrial genome arrangements. BMC Evol. Biol. 8, 61.

Ho, S.Y.W., Phillips, M.J., 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst. Biol 58, 367-380.

Holford, M., Zhang, M.-M., Gowd, K.H., Azam, L., Green, B.R., Watkins, M., Ownby, J.-P., Yoshikami, D., Bulaj, G., Olivera, B.M., 2009. Pruning nature: Biodiversity-derived discovery of novel sodium channel blocking conotoxins from *Conus bullatus*. Toxicon 53, 90-98.

Huelsenbeck, J., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754-755.

Johns, G.C., Avise, J.C., 1998. A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. Mol. Biol. Evol. 15, 1481-1490.

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

Kohn, A.J., 1959. The Ecology of *Conus* in Hawaii. Ecol. Monogr. 29, 47-90.

Kohn, A.J., 1990. Tempo and mode of evolution in Conidae. Malacologia 32, 55-67.

Kornobis, E., Cabellos, L., Aguilar, F., Frías-López, C., Rozas, J., Marco, J., Zardoya, R., 2015. TRUFA: A user-friendly web server for *de novo* RNA-seq analysis using cluster computing. Evol. Bioinform. 11, 97-104.

Kraus, N.J., Corneli, P.S., Watkins, M., Bandyopadhyay, P.K., Seger, J., Olivera, B.M., 2011. Against expectation: a short sequence with high signal elucidates cone snail phylogeny. Mol. Phylogenet. Evol. 58, 383-389.

Kraus, N.J., Watkins, M., Bandyopadhyay, P.K., Seger, J., Olivera, B.M., Corneli, P.S., 2012. A very short, functionally constrained sequence diagnoses cone snails in several *Conasprella* clades. Mol. Phylogenet. Evol. 65, 335-338.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. Mol. Biol. Evol. 29, 1695-1701.

Nam, H.H., Corneli, P.S., Watkins, M., Olivera, B., Bandyopadhyay, P., 2009. Multiple genes elucidate the evolution of venomous snail-hunting *Conus* species. Mol. Phylogenet. Evol. 53, 645-652.

Olivera, B.M., 2002. *Conus* venom peptides: reflections from the biology of clades and species. Annu. Rev. Ecol. Syst. 33, 25-47.

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 Evol. Biol. 14, 1-10.

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

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.

Powell, A.W.B., 1942. The New-Zealand Recent and Fossil Mollusca of the Family Turridae. With General Notes on Turrid Nomenclature and Systematics. Bull. Auckl. Inst. Mus. 2, 1–192.

Prashanth, J.R., Dutertre, S., Jin, A.H., Lavergne, V., Hamilton, B., Cardoso, F.C., Griffin, J., Venter, D.J., Alewood, P.F., Lewis, R.J., 2016. The role of defensive ecological interactions in the evolution of conotoxins. Mol. Ecol. 25, 598-615.

Puillandre, N., Bouchet, P., Duda Jr, T.F., Kauferstein, S., Kohn, A.J., Olivera, B.M., Watkins, M., Meyer, C., 2014a. Molecular phylogeny and evolution of the cone snails (Gastropoda, Conoidea). Mol. Phylogenet. Evol. 78, 290-303.

Puillandre, N., Duda, T.F., Meyer, C., Olivera, B.M., Bouchet, P., 2014b. One, four or 100 genera? A new classification of the cone snails. J. Mollus. Stud.

Puillandre, N., Holford, M., 2010. The Terebridae and teretoxins: Combining phylogeny and anatomy for concerted discovery of bioactive compounds. BMC Chem. Biol. 10, 1-12.

Puillandre, N., Kantor, Y.I., Sysoev, A., Couloux, A., Meyer, C., Rawlings, T., Todd, J.A., Bouchet, P., 2011. The dragon tamed? A molecular phylogeny of the Conoidea (Gastropoda). J. Mollus. Stud. 77, 259-272.

Puillandre, N., Samadi, S., Boisselier, M.C., Sysoev, A.V., Kantor, Y.I., Cruaud, C., Couloux, A., Bouchet, P., 2008. Starting to unravel the toxoglossan knot: molecular phylogeny of the “turrids”(Neogastropoda: Conoidea). Mol. Phylogenet. Evol. 47, 1122-1134.

Rambaut, A., Drummond, A.J., 2007. Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

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.

Röckel, D., Korn, W., Kohn, A.J., 1995. Manual of the Living Conidae. Vol. I, Indo-Pacific. Christa Hemmen Verlag, Wiesbaden.

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

Sacco, F., 1893. I molluschi dei terreni terziarii del Piemonte e della Liguria. Conidae e Conorbidae. Stamperia Reale, Torino. Italy.

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 Res. 33, W686-W689.

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

Schwarz, G., 1978. Estimating the Dimension of a Model. Annals of Statistics, 461-464.

StJohn, J., 2011. SeqPrep. https://github.com/jstjohn/SeqPrep.

Stöger, I., Schrödl, M., 2013. Mitogenomics does not resolve deep molluscan relationships (yet?). Mol. Phylogenet. Evol. 69, 376-392.

Taylor, J.D., Kantor, Y.I., Sysoev, A.V., 1993. Foregut anatomy, feeding mechanisms, relationships and classification of the Conoidea (= Toxoglossa) (Gastropoda). Bull. Nat. Hist. Mus. Lond. Zool. 59, 125–170.

Tenorio, M.J., Castelin, M., 2016. Genus *Profundiconus* Kuroda, 1956 (Gastropoda, Conoidea): Morphological and molecular studies, with the description of five new species from the Solomon Islands and New Caledonia. Eur. J. Taxon. 173, 1-42.

Terlau, H., Olivera, B.M., 2004. *Conus* venoms: a rich source of novel ion channel-targeted peptides. Physiol. Rev. 84, 41-68.

Tucker, J.K., Tenorio, M.J., 2009. Systematic classification of recent and fossil conoidean gastropods: with keys to the genera of cone shells. Conchbooks.

Tucker, J.K., Tenorio, M.J., 2013. Illustrated catalog of the living cone shells. MDM Publishing, Wellington, FL, USA.

Uribe, J.E., Kano, Y., Templado, J., Zardoya, R., 2016. Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. Zool. Scr. 45, 145-159.

Walls, J.G., 1978. Supraspecific groupings in living cones (Mollusca: Conidae). The Pariah 3, 1-13.

Watkins, M., Corneli, P.S., Hillyard, D., Olivera, B.M., 2010. Molecular phylogeny of *Conus chiangi* (Azuma, 1972) (Gastropods:Conidae). The Nautilus 124, 129-136.

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

WoRMS, E.B., 2016. World Register of Marine Species. Available from http://www.marinespecies.org at VLIZ. Accessed 2016-05-07.

SUPPORTING INFORMATION 5

Additional Supporting Information may be found in the online version of this article:

Data S1. Amplification strategy. Long PCR and primer walking primers.

Data S2. Best fit partitions and evolutionary substitution models as selected by Partition Finder

Data S3. Mitochondrial genome features

Data S4. Secundary structure of tRNAs from manual search

4 DISCUSION

A pesar de los grandes avances en tecnologías de la secuenciación, el número de genomas mitocondriales completos conocidos sigue siendo muy escaso en los moluscos y, en particular, se considera muy insuficiente en los gasterópodos, a pesar de ser la clase con más especies dentro del filo y uno de los grupos más diversificado del Reino Animal. Ello es debido principalmente a que muy pocas de las especies de gasterópodos son consideradas sistemas modelo y sigue existiendo un importante trabajo taxonómico de identificación y descripción por realizar. Por este motivo, en la presente tesis se ha realizado un gran esfuerzo para incrementar la información de secuencias del genoma mitocondrial y de genes parciales mitocondriales y nucleares en grupos de gasterópodos que han experimentado una importante diversificación, pero que apenas han sido estudiados desde este punto de vista. Los resultados obtenidos en el presente trabajo de tesis doctoral han sido ampliamente discutidos en sus correspondientes capítulos (publicaciones), por lo que aquí sólo se expone una breve síntesis de los mismos.

En algunos grupos de animales, como los moluscos, los genomas mitocondriales presentan altas tasas de reordenamiento (translocación, inversión, deleción o inserción) de sus genes. En un contexto evolutivo, las reorganizaciones de los genes mitocondriales pueden ser interpretadas como sinapomorfías (caracteres derivados compartidos a partir de un ancestro común) moleculares portadoras de señal filogenética (Grande *et al.*, 2008).

En general, las reorganizaciones afectan principalmente a genes codificantes de ARNs de transferencia y en menor medida a genes codificantes de proteínas y de ARNs ribosómico (Gissi *et al.*, 2008). Así, la organización del genoma mitocondrial se ha mantenido casi invariable con respecto a los genes codificantes de proteínas en la mayoría de los linajes del filo Mollusca a lo largo de su historia evolutiva. Ello ha permitido mediante estudios comparativos proponer el orden ancestral para Mollusca y Gastropoda en particular (Osca *et al.*, 2014a).

Dentro de la subclase Vetigastropoda, la mayoría de los genomas secuenciados en la presente tesis doctoral se ajustan en su organización a la de los pocos publicados previamente (del género *Haliotis* por Maynard *et al.* (2005) y por Robinson *et al.* (2016), de *Tegula* (NC_016954, sin publicar), y de *Lunella* por Williams *et al.* (2014), y comparten el orden ancestral propuesto para Gastropoda. Como en otros grupos, los cambios mínimos deducidos de la comparación de los genomas mitocondriales completos o casi completos aquí determinados, afectan principalmente a los genes codificantes de ARNs de transferencia. No obstante, el orden del genoma mitocondrial obtenido para *Diodora graeca* confirma el antes descrito para la superfamilia Fissurelloidea (*Fissurella volcano*: NC_016953, sin publicar) y, al igual que el reordenamiento del genoma mitocondrial parcial aquí obtenido para *Lepetodrilus schrolli* (superfamilia Lepetodriloidea), constituyen los cambios más drásticos dentro de Vetigastropoda. El genoma mitocondrial de *L. schrolli* muestra una inversión de varios de sus genes codificantes de proteína, lo que supone un cambio notable y

nuevo para el grupo (ver publicación 1). Asimismo, la organización del genoma mitocondrial de Neomphalina (Chen *et al.*, 2015) mostró ser muy diferente a todos los genomas mitocondriales publicados de Gastropoda y, en particular, de Vetigastropoda, grupo dentro del cual se ha clasificado en ocasiones (Bouchet *et al.*, 2005; Aktipis y Giribet, 2012). Los resultados de esta tesis, en cambio, apoyan su exclusion de los Vetigastropoda y favorecen otras hipótesis como las que proponen a Neomphalina como grupo hermano de los Vetigastropoda (Ponder y Lindberg, 1997; Warén *et al.*, 2003; Geiger y Thacker, 2005; Geiger *et al.*, 2008), o próximo a Cocculinoidea (McArthur y Harasewych, 2003; Aktipis y Giribet, 2012).

Por otro lado, dentro la subclase Neritimorpha, sólo se había publicado hasta la fecha el ordenamiento mitocondrial correspondiente al género *Nerita* (Castro y Colgan, 2010; Arquez *et al.*, 2014), que precisamente se ajusta al orden ancestral propuesto para Gastropoda. En el presente trabajo, se ha observado que dicho ordenamiento mitocondrial es compartido por los otros géneros de la superfamilia Neritoidea estudiados (*Theodoxus* y *Neritina*), así como por las superfamilias Hydrocenoidea (*Georissa*) y Neritopsoidea (*Titiscania*). Sin embargo, en el genoma mitocondrial de Helicinoidea (*Pleuropoma*) se encontró un importante reordenamiento, con inversiones en varios de los genes codificantes de proteínas. Dichos cambios se verificaron además al secuenciar el genoma mitocondrial parcial de otro género (*Viana*) de la superfamilia.

Los reordenamientos encontrados para los 12 genomas mitocondriales secuenciados de la superfamilia Conoidea coinciden con el orden encontrado mayoritariamente en Caenogastropoda (Osca *et al.*, 2015). Este orden, a diferencia del orden ancestral para Gastropoda, presenta una inversión de la cadena menor a mayor que deja todos los genes codificantes para proteína y ARNs ribosómico en la misma disposición, y solo el "cluster" MYCWQGE y el gen *trnT* en la cadena menor. Curiosamente, los nuevos reordenamientos encontrados en Vetigastropoda (*Lepetodrilus schrolli*) y Neritimorpha (*Pleuropoma jana*) presentan una inversión muy similar a la encontrada en Caenogastropoda, y que sólo difiere en la disposición relativa de los ARNs de trasferencia incluidos en el "cluster" MYCWQGE de Caenogastropoda y los genes *atp6-8* y *rrnL-S* de *L. schrolli*.

Las secuencias del genoma mitocondrial completo han sido utilizadas con gran éxito y capacidad de resolución en diversos grupos de animales para determinar las relaciones filogenéticas a diferentes niveles taxonómicos (Irisarri *et al.*, 2012; San Mauro *et al.*, 2014; Crampton-Platt *et al.*, 2015; Yang *et al.*, 2015, entre otros). Sin embargo, en los gasterópodos, los frecuentes reordenamientos mitocondriales están correlacionados con altas tasas evolutivas (Xu *et al.*, 2006), lo que dificulta en cierta medida las reconstrucciones filogenéticas debido a la posibilidad de recuperar artefactos como el denominado de "Atracción de Ramas Largas" (Long Branch Attraction, LBA; Bergsten 2005).

Las relaciones filogenéticas internas de Gastropoda han sido objeto de numerosos estudios, tanto mediante análisis morfológicos como moleculares (Ponder y Lindberg, 1997; Sasaki, 1998; Colgan *et al.*, 2000; Colgan *et al.*, 2003; McArthur y Harasewych, 2003; Grande *et al.*, 2008; Castro y Colgan, 2010; Kocot *et al.*, 2011; Smith *et al.*, 2011; Williams *et al.*, 2014; Zapata *et al.*, 2014, entre otros). A pesar de ello, siguen siendo hoy día objeto de controversia. La filogenia y la clasificación morfológica de Ponder y Lindberg (1997) sigue siendo la más completa y aceptada para Gastropoda. Según esta propuesta, Gastropoda está dividido en dos clados: Eogastropoda (Patellogastropoda) y Orthogastropoda, este último compuesto por Vetigastropoda, Neritimorpha y Apogastropoda (Caenogastropoda + Heterobranchia), quedando sin resolver la posición de Cocculinoidea y Neomphalina. Los diversos análisis moleculares que han intentado recuperar y apoyar esta hipótesis filogenética han tenido resultados dispares y en muchas ocasiones incongruentes. Así, algunos trabajos (Kocot *et al.*, 2011) recuperan topologías acordes a las relaciones filogenéticas propuestas por Ponder y Lindberg (1997), mientras que otros recuperan Patellogastropoda como grupo hermano de Vetigastropoda, y Neritimorpha como grupo hermano de Apogastropoda (Smith *et al.*, 2011; Zapata *et al.*, 2014). Por su parte, los análisis basados en genomas mitocondriales completos presentan un sesgo, debido a la atracción de grupos (Heterobranchia y Patellogastropoda) que presentan drásticos reordenamientos y tasas evolutivas aceleradas (Grande *et al.*, 2008; Castro y Colgan, 2010).

Estos resultados incongruentes de las filogenias moleculares son debidos en parte a la escasa representación en los estudios filogenéticos de algunos de sus principales linajes, como Patellogastropoda, Cocculinoidea, Neomphalina, Neritimorpha y Vetigastropoda. Incluso en los linajes mejor representados en los análisis, como Heterobranchia y Caenogastropoda, tampoco se puede decir que están exentos de cierto sesgo, dada su gran diversidad y la dificultad de conseguir muestras de todos los linajes que componen estos grupos para poder incluirlas en los estudios. En el presente trabajo de tesis doctoral se ha avanzado precisamente en la incorporación de linajes menos representados para paliar esta situación. Así, en la reconstrucción de las relaciones filogenéticas de Gastropoda (Publicación 1), se incorporó por primera vez en los análisis basados en genomas mitocondriales completos a la subclase Neomphalina. Previamente, se había considerado que Neomphalina podía ser un grupo hermano de Vetigastropoda (Ponder y Lindberg, 1997; Warén *et al.*, 2003; Geiger y Thacker, 2005; Geiger *et al.*, 2008), pertenecer a Vetigastropoda (Bouchet *et al.*, 2005; Aktipis y Giribet, 2012) o estar próximo a Cocculinoidea (McArthur y Harasewych, 2003; Aktipis y Giribet, 2012). Los resultados de los análisis filogenéticos realizados en esta tesis doctoral apoyan con alto soporte estadístico que Neomphalina es una línea evolutiva diferente a Vetigastropoda y posiblemente su grupo hermano (a falta de incorporar Cocculinoidea en los análisis).

Por otra parte, como mencionábamos anteriormente, la heterogeneidad de las tasas evolutivas entre linajes sesga fuertemente la reconstrucción filogenética a partir de los datos de

secuencias de genomas mitocondriales completos. Aquí, se ha tratado de solucionar en parte estos sesgos mediante diferentes aproximaciones. Por un lado, se ha investigado el efecto sobre la topología final de realizar los análisis utilizando aminoácidos, nucleótidos e incorporar o no las terceras posiciones de codón de los genes codificantes de proteínas. Asimismo, se han realizado análisis utilizando las particiones óptimas (seleccionadas de acuerdo a criterios de información Bayesianos) con los modelos evolutivos que mejor se ajustan a los datos. Una novedad ha sido la aplicación de modelos evolutivos (CAT; Lartillot *et al.*, 2007), que tienen en cuenta de forma independiente las tasas evolutivas para cada posición del alineamiento (Lartillot y Philippe, 2004) y han sido específicamente diseñados para evitar la atracción de ramas largas. Los análisis realizados con nucleótidos y excluyendo las terceras posiciones de codón (Castro y Colgan, 2010), así como aquellos basados en el modelo CAT (Publicación 1), consiguen contrarrestar en parte los sesgos conocidos asociados a los genomas mitocondriales y recuperar el linaje Apogastropoda, lo cual hasta ahora no era fácil. En cualquier caso, los apoyos estadísticos siguen siendo bajos, por lo que será necesario continuar mejorando está línea de investigación de perfeccionamiento de los análisis filogenéticos para la resolución de la filogenia de Gastropoda basada en genomas mitocondriales.

Pero el gran potencial y la capacidad resolutiva de los genomas mitocondriales se centran en niveles taxonómicos más bajos (por debajo de subclase). En los análisis filogenéticos de Vetigastropoda se han obtenido cuatro linajes cuyas relaciones

quedan por resolver: Fissurelloidea, Lepetodriloidea, Seguenzioidea + Haliotoidea, y Trochoidea + Angarioidea + Phasianelloidea. En gran parte, muchos de los problemas asociados a la filogenia de Vetigastropoda se centran en la definición de la superfamilia Trochoidea. En el presente trabajo se han incluido tres grupos que tradicionalmente se habían incluido dentro de Trochoidea, pero recientemente fueron considerados como superfamilias independientes: Seguenzioidea (Kano, 2008), Angarioidea y Phasianelloidea (Williams *et al.*, 2008). Los resultados obtenidos en esta tesis apoyan claramente que Seguenzioidea es una superfamilia independiente cercana a Haliotoidea y cuyos límites están todavía por resolver (Bandel, 2010), para lo cual sería necesario secuenciar el genoma mitocondrial de más representantes, pues hasta la fecha solo se conoce el del género *Granata* (Publicación 1). En cambio, las superfamilias Angarioidea y Phasianelloidea ya aparecían estrechamente relacionadas con Trochoidea en un primer trabajo (Publicación 1) y, de hecho, en un segundo análisis en el que se amplía la representación de familias de Trochoidea, se observa como se recupera con un apoyo estadístico alto a Trochoidea s. l. (en la concepción tradicional mantenida por Hickman y McLean 1990), incluyendo a Angarioidea y Phasianelloidea, y en contra de Williams *et al.* (2008). Las relaciones entre las familias de Trochoidea s. l. coinciden con las publicadas que se basan en análisis realizados con genes parciales (Williams, 2012), pero que no incluyeron a Angarioidea y Phasianelloidea, tal vez siguiendo la consideración previa de excluir a ambos grupos de Trochoidea, según Williams *et al.* (2008). Cabe resaltar la estrecha relación entre

las familias Trochidae y Calliostomatidae, así como que *Cittarium* y *Tectus* son géneros que deben formar una nueva familia que debe ser formalmente nombrada y descrita (al no estar directamente relacionados con los miembros de la familia Tegulidae).

Los resultados obtenidos en esta tesis doctoral avanzan en la delimitación de linajes dentro de Vetigastropoda y permiten inferir que la evolución de la pérdida de la branquia derecha (condición zeugobranquial) ha ocurrido múltiples veces dentro del grupo. La filogenia que se presenta aquí requiere, no obstante, ser completada en el futuro con la incorporación de genomas mitocondriales de las superfamilias Lepetelloidea, Pleurotomarioidea y Scissurelloidea, así como con la ampliación de la representación de familias de Trochoidea s. l., la superfamilia más diversa dentro del grupo.

Continuando con la resolución de relaciones filogenéticas a diferentes niveles taxonómicos, uno de los trabajos de la presente tesis (Publicación 3) se ha centrado en dos subfamilias de Trochidae (Vetigastropoda: Trochoidea), Cantharidinae y Stomatellinae. La subfamilia Cantharidinae es la que está mejor representada en las costas del Atlántico, mientras que la subfamilia Stomatellinae, que se considera grupo hermano de la anterior, presenta una distribución Indo-Pacífica. La subfamilia Cantharidinae fue redefinida recientemente basándose en análisis moleculares que incluían mayoritariamente especies del Indo-Pacífico (Williams *et al.*, 2010). El presente trabajo (Publicación 3) se ha centrado, por tanto, en las especies del Atlántico Noreste y Mediterráneo, que representan a todos los géneros de Cantharidinae presentes en esta área

geográfica, así como la especie tipo del género *Callumbonella* (*C. suturalis*), de posición taxonómica incierta. Las sospechas previas de que los linajes del Atlántico Este y Mediterráneo parecían constituir un grupo monofilético (Williams *et al.*, 2010; Williams, 2012) y que el género *Gibbula* fuera polifilético (Williams *et al.*, 2010; Barco *et al.*, 2013), han sido corroboradas con un alto apoyo estadístico en los análisis realizados para esta tesis (Publicación 3). También, se recuperaron con un apoyo robusto los géneros *Phorcus* y *Jujubinus* (quedando incluida en este último la especie considerada hasta ahora como *Gibbula tingitana*), y *Jujubinus* quedó como grupo hermano de un clado formado por *Clellandella* y *Callumbonella*. Sin embargo, las relaciones entre los grupos restantes dentro de Cantharidinae permanecen aún inciertas. El cronograma obtenido utilizando como calibración un fósil que determina el origen de los grupos endémicos de Nueva Zelanda (Donald *et al.*, 2012), determinó que los principales eventos cladogenéticos sucedieron en periodos geológicos concretos, habiéndose estimado la separación de la fauna Atlanto-Mediterránea hace 47 millones de años (durante la denominada fase *Azolla*). La diversificación a nivel de género se estimó en 14 millones de años (coincidiendo con el cierre definitivo del mar de Tetis). Finalmente, la diversificación específica se data a partir de 5,3 millones de años (Crisis de Salinidad del Messiniense). En la presente tesis también se ha abordado una primera exploración filogenética de la diversidad de la subfamilia Stomatellinae, que previamente se había determinado como el grupo hermano de Cantharidinae (Williams *et al.*, 2010). La diversidad de Stomatellinae aparece claramente

infraestimada a la luz de los resultados obtenidos en este tesis y las relaciones entre sus múltiples géneros quedan abiertas, a la espera de incluir un mayor número de ellos en los análisis, así como de secuenciar genomas mitocondriales completos de los principales linajes identificados que permitan obtener nodos con apoyos estadísticos robustos.

Otro de los grupos objetivo de la presente tesis doctoral fue la subclase Neritimorpha. A pesar de su gran diversificación de hábitats y modos de vida, del importante número de especies que comprende (Fukumori y Kano, 2014) y de su completo registro fósil (Bandel, 1999; Frýda, 1999), las numerosas sinapomorfías que presentan no ofrecen duda sobre su monofilia (Ponder y Lindberg, 1997; Kano *et al.*, 2002). Sin embargo, su historia evolutiva solo ha sido objeto de un único trabajo basado en un gen parcial ribosómico (*28S*) (Kano *et al.*, 2002). En este trabajo, las relaciones filogenéticas inferidas carecían de suficiente apoyo estadístico para ser concluyentes, además de observarse un sesgo debido a tasas aceleradas en el gen *28S rDNA* de Hydrocenoidea. En esta tesis (Publicación 4), se han determinado las secuencias de genomas mitocondriales de representantes de las cuatro superfamilias de Neritimopha, lo cual ha permitido reconstruir las relaciones internas del grupo de manera robusta. Además se han obtenido secuencias parciales de genes nucleares (*h3*, *actin 18S* y *28S*) con el fin de dar mayor consistencia al análisis previo realizado con los genomas mitocondriales. Tanto los análisis con genes nucleares, como mitocondriales y combinados convergieron en la misma topología, con un apoyo máximo en todos los nodos en los análisis

combinados. Neritopsoidea resultó ser el primer linaje divergente y se recuperó como grupo hermano de un clado formado por Helicinoidea e Hydrocenoidea más Neritoidea. La posición relativa de Helicinoidea y Hydrocenoidea difiere de la hipótesis previa (KANO *et al*., 2002), y en los resultados obtenidos en esta tesis es más coincidente con los datos morfológicos, que muestran una pérdida gradual de los órganos de la cavidad paleal, siendo Hydrocenoidea + Neritoidea los grupos asimétricos (Barker, 2001). Sin embargo, el número de taxones representados es aún insuficiente y serían necesarios datos de todas las familias incluidas dentro de la subclase para establecer una reconstrucción más precisa de los estados de carácter.

En el cronograma elaborado para estimar los principales eventos de divergencia dentro de la subclase Neritimorpha, la separación de la superfamilia Helicinoidea de Hydrocenoidea y Neritoidea se ha estimado en 234 millones de años, mientras que la separación de Hydrocenoidea y Neritoidea se ha estimado en alrededor de 180 millones de años, lo que es coincidente con la distribución del grupo y la separación de los continentes actuales hace alrededor de 175 millones de años (Bandel, 2000). Las separaciones a nivel de familia dentro de Helicinoidea y Neritoidea se han estimado al final de Cretácico e inicios del Paleoceno (hace entre 72 y 65 millones de años), acorde con lo señalado para el grupo por varios autores (Kiel, 2003; Frey y Vermeij, 2008).

En el caso de Caenogastropoda, la presente tesis se ha centrado en la superfamilia Conoidea, por ser ésta la más extensa en

número de especies dentro de la subclase (Olivera *et al.*, 2014). Tradicionalmente, y como sucede con la mayoría de las grandes radiaciones, el grupo presenta una enorme complejidad taxonómica y ha resultado muy difícil ordenar en el tiempo los diferentes eventos cladogenéticos que ha experimentado. Tradicionalmente, los Conoidea eran divididos en tan solo tres familias: Turridae, Terebridae y Conidae (Vaught, 1989). Sin embargo, el grupo ha sido redefinido recientemente de forma sucesiva (Bouchet *et al.*, 2005; Tucker y Tenorio, 2009; Bouchet *et al.*, 2011), aumentando considerablemente el número de taxones supraespecíficos (con hasta 16 familias), pero sin que se alcance un consenso. En estos últimos años se ha intensificado el esfuerzo para delimitar los principales linajes de la superfamilia con herramientas moleculares. Como consecuencia, se ha demostrado que esta superfamilia efectivamente presenta un elevado número de linajes independientes (Puillandre *et al.*, 2014a; Puillandre *et al.*, 2014b). En estos trabajos también se abordan las relaciones filogenéticas de los principales linajes dentro de la familia tipo (Conidae) (Puillandre *et al.*, 2014a) y se definen varios clados con buenos apoyos estadísticos, pero cuyas relaciones quedan sin resolver. La presente tesis ha centrado el esfuerzo en la familia Conidae (Publicación 5). Los doce genomas mitocondriales aquí secuenciados abarcan cinco clados incluidos en la familia Conidae y tres familias relacionadas (Conorbidae, Clathurellidae y Mangeliidae). La reconstrucción de los árboles filogenéticos obtenidos ha mostrado un fuerte apoyo estadístico en la monofilia de Conidae, así como en sus relaciones internas. En particular cabe destacar la adscripción de

Profundiconus a la familia Conidae y el establecimiento de un clado que incluye a *Conasprella* como grupo hermano de *Californiconus* y dos géneros (*Lilliconus* + *Pseudolilliconus*) que hasta ahora no habían sido incluidos en los análisis filogenéticos (de hecho, el segundo es un género nuevo, descubierto recientemente y formado por especies con ejemplares diminutos, que permite avanzar que la diversidad real del grupo está aún por descubrir). La reconstrucción filogenética obtenida dentro del género *Conus* ha resultado ser congruente con resultados previos (Puillandre *et al.*, 2014a), con la ventaja de dotarlos de apoyo estadístico adicional. En el caso del género *Conus*, sin embargo, la filogenia dista mucho de tener una buena representación de la diversidad del grupo y, por lo tanto, será necesario secuenciar más genomas mitocondriales completos de numerosos linajes con el fin de poder establecer un marco filogenético robusto en el que interpretar los patrones evolutivos referentes a la dieta (gusanos, caracoles, peces), a la biogeografía o a las familias de conotoxinas.

Las estimaciones de los tiempos de divergencia dentro de la familia Conidae fueron datadas con fósiles ya usados previamente (Cunha *et al.*, 2005). Como se ha visto en los demás grupos estudiados en esta tesis doctoral, los principales eventos de divergencia se asocian a eras geológicas concretas. Así, el origen de la familia Conidae se ha estimado en la transición entre el Paleoceno y el Eoceno, acorde a lo planteado previamente según el registro fósil (Kohn, 1990). Las divisiones de los principales géneros se sucedieron después del Paleoceno, y la gran diversificación, con el incremento abrupto de número de especies,

se ha estimado que se produjo en la transición entre el Paleoceno y el Mioceno.

A la vista de los resultados obtenidos para los diferentes grupos de gasterópodos estudiados en el presente trabajo, se puede concluir que los genomas mitocondriales aportan información filogenética de calidad que permite resolver las relaciones filogenéticas dentro de los grandes linajes (subclases) de Gastropoda con altos apoyos estadísticos. La posibilidad de obtener genomas mitocondriales completos (o casi completos si no se puede secuenciar la región de control) de forma relativamente sencilla y económica con las técnicas de secuenciación masiva, permite imaginar que en un futuro cercano se podrá ir completando la filogenia de cada grupo con los linajes peor representados hasta el momento. El principal reto actual es, por lo tanto, la obtención de muestras de dichos linajes y contar con el conocimiento taxonómico necesario para poder identificarlos correctamente (lamentablemente, cada vez más escaso). No cabe duda de que las filogenias basadas en genomas mitocondriales deben ser contrastadas con análisis filogenómicos basados en marcadores nucleares, aunque la obtención de estos sigue siendo aún complicada y costosa. Asimismo es necesaria la integración de metodologías novedosas de inferencia filogenética probabilística, como se ha visto en el presente trabajo, para poder resolver los posibles sesgos derivados de procesos moleculares relacionados con la aceleración en las tasas mutacionales y que tantos problemas generan en grupos hiperdiversos como los que componen Gastropoda.

5 CONCLUSIONES

De los resultados de la presente tesis doctoral pueden obtenerse las siguientes conclusiones generales:

1) Se han determinado las secuencias completas o casi completas de los genomas mitocondriales de 29 especies de gasterópodos utilizando técnicas de secuenciación masiva: 11 corresponden a Vetigastropoda; seis a Neritimorpha; y 12 a Conoidea (Caenogastropoda).

2) La ordenación de los genes mitocondriales se ajusta en la mayoría de los Vetigastropoda a la propuesta como ancestral para los Gastropoda, variando sólo en la posición relativa de algunos genes codificantes de ARNs de transferencia. Solo los genomas mitocondriales de las superfamilias Lepetodriloidea y Fissurelloidea muestran reordenamientos génicos más drásticos. Además, se comprobó que el genoma mitocondrial del representante de Neomphalina muestra un ordenamiento propio, no relacionado con ningún otro descrito para Gastropoda.

3) En la filogenia reconstruida de Gastropoda usando genomas mitocondriales completos, Neomphalina está solo lejanamente emparentado con Vetigastropoda. Dentro de Vetigastropoda, se distinguen cuatro linajes a nivel de superfamilia: Fissurelloidea, Lepetodriloidea, Seguenzioidea + Haliotoidea, y Trochoidea + Angarioidea

+ Phasianelloidea. Las relaciones filogenéticas entre ellos no pudieron resolverse completamente. No obstante, la filogenia permite inferir que la pérdida de la branquia paleal derecha ocurrió en múltiples ocasiones durante la evolución de Vetigastropoda.

4) La filogenia centrada en la superfamilia Trochoidea s. l. recuperó tres linajes principales: el primero formado por las familias Trochidae y Calliostomatidae; el segundo por Margaritidae; y un tercero que agrupa Angarioidea y Phasianelloidea con un clado formado por los géneros *Tectus* y *Cittarium* (que formarían una nueva familia), más las familias Tegulidae y Turbinidae.

5) Se aportan 350 secuencias parciales de seis genes (*rrnL*, *rrnS*, *cox1*, *cob*, *28S ARNr*, e *Histona H3*) correspondientes a especies de las subfamilias Cantharidinae y Stomatellinae (Trochoidea: Trochidae). Dentro de la subfamilia Cantharidinae, la filogenia muestra como las especies del Atlántico Noreste y Mediterráneo se agrupan en un clado que recibe alto apoyo estadístico.

6) Dentro de los Cantharidinae del Atlántico Noreste y Mediterráneo, los géneros *Phorcus* y *Jujubinus* forman sendos grupos monofiléticos con alto apoyo estadístico, al igual que un clado compuesto por los géneros *Clelandella* y *Callumbonella*. Sin embargo, el género *Gibbula* no es

monofilético y queda dividido en cinco linajes independientes.

7) El origen de la divergencia de los Cantharidinae del Atlántico Noreste y Mediterráneo se ha estimado en hace unos 47 millones de años (durante la denominada fase Azolla, en el Eoceno Medio), mientras que los eventos de diversificación a nivel genérico y específico coinciden con el cierre definitivo del mar de Tetis (hace 14 millones de años) y la Crisis de Salinidad del Messiniense (hace 5,3 millones de años), respectivamente.

8) Los seis genomas mitocondriales correspondientes a la subclase Neritimorpha comprenden las cuatro superfamilias del grupo. En tres de ellas, el ordenamiento genómico coincide con el considerado ancestral de Gastropoda. Sólo la superfamilia Helicinoidea mostró reorganizaciones significativas en el orden de los genes mitocondriales. La filogenia de Neritimorpha basada en genomas mitocondriales completos recuperó con alto apoyo estadístico a la superfamilia Neritopsoidea como grupo hermano de un clado formado por Helicinoidea e Hydrocenoidea más Neritoidea.

9) De los 12 genomas mitocondriales secuenciados para Conoidea, nueve representan todos los géneros reconocidos de la familia Conidae según la clasificación más actual, y uno más, un nuevo género aún sin describir; el resto pertenecen a familias estrechamente relacionadas

con Conidae: Conorbidae, Clathurellidae y Mangeliidae. Todos los genomas mitocondriales de Conoidea comparten esencialmente el ordenamiento genómico de Caenogastropoda, y sólo difieren en cambios en la posición relativa de genes codificantes para ARNs de transferencia.

10) La filogenia reconstruida a partir de los genomas mitocondriales completos recuperó la familia Conidae como un grupo monofilético. El género *Profundiconus* se obtiene como grupo hermano de los demás géneros, siendo el género *Conus* el grupo hermano de un clado que incluye el género *Conasprella* como grupo hermano del género *Californiconus*, y *Lilliconus* más *Pseudolilliconus*. La divergencia de los principales linajes dentro de la familia Conidae se ha estimado que tuvo lugar entre el Paleoceno y el Eoceno (hace 56-30 millones de años) y la diversificación de especies en la transición del Oligoceno al Mioceno (hace 23 millones de años).

5 CONCLUSIONS

The results obtained in the present PhD Thesis allow deriving the following general conclusions:

1) The complete or nearly complete sequences of the mitochondrial genomes of 29 gastropods were determined using next generation sequencing: 11 belong to Vetigastropoda, six to Neritimorpha, and 12 to Conoidea (Caenogastropoda).

2) The mitochondrial gene order of most Vetigastropoda conforms to the ancestral one of Gastropoda, and only varies in the relative position of some genes encoding tRNAs. Only the mitochondrial genomes of superfamilies Lepetodriloidea and Fissurelloidea showed more drastic gene rearrangements. Furthermore, it was found that the mitochondrial genome of Neomphalina shows its own arrangement, which is not related to any other reported for Gastropoda.

3) In the reconstructed phylogeny of Gastropoda based on complete mitochondrial genomes, Neomphalina is only distantly related to Vetigastropoda. Within Vetigastropoda, up to four distinct linages at the superfamily level are distinguished: Fissurelloidea, Lepetodriloidea, Seguenzioidea + Haliotoidea, and Trochoidea + Angarioidea + Phasianelloidea. Phylogenetic relationships among them could not be fully resolved.

However, the phylogeny allowed inferring that the loss of the right pallial gill occurred multiple times in vetigastropod evolution.

4) The phylogeny focused on the superfamily Trochoidea s. l. recovered three main lineages: the first included families Trochidae and Calliostomatidae; the second included Margaritidae; and a third one grouped Angarioidea and Phasianelloidea to a clade formed by genera *Tectus* and *Cittarium* (that would constitute a new family) plus the families Tegulidae and Turbinidae.

5) A total of 350 partial sequences of six genes (*rrnL, rrnS, cox1, cob, 28S rR*NA and *histone H3*) were determined for species of the subfamilies Cantharidinae and Stomatellinae (Trochoidea: Trochidae). Within the subfamily Cantharidinae, the phylogeny shows that species from the North-eastern Atlantic and Mediterranean Sea are grouped in a clade that receives high statistical support.

6) Within Cantharidinae from North-eastern Atlantic and Mediterranean Sea, *Phorcus* and *Jujubinus* formed monophyletic groups with high statistical support, as had the clade composed by *Clelandella* and *Callumbonella*. However, the genus *Gibbula* is not monophyletic and was divided into five independent lineages.

7) The origin of the divergence of the Cantharidinae from North-eastern Atlantic and Mediterranean was estimated to

have occurred about 47 mya (during the so-called Azolla event, Middle Eocene), whereas diversification events at the generic and specific levels coincided with the Tethys closure (14 mya) and the Messinian Salinity Crisis (5.3 mya), respectively.

8) The six mitochondrial genomes of Neritimorpha cover the four superfamilies included in the group. The genome arrangement of three of the superfamilies coincides with the ancestral one of Gastropoda. Only the superfamily Helicinoidea showed significant gene order rearrangements. The phylogeny of Neritimorpha based on complete mitochondrial genomes recovered the superfamily Neritopsoidea as sister group to a clade formed by Helicinoidea and Hydrocenoidea plus Neritoidea.

9) Of the 12 mitochondrial genomes sequenced for Conoidea, nine represent all genera known to the family Conidae, and another one a new undescribed genus; the remaining belong to families closely related to Conidae: Conorbidae, Clathurellidae and Mangeliidae. All conoidean mitochondrial genomes essentially share the common gene order described for Caenogastropoda, and only differ in changes in the relative position of some genes encoding tRNAs.

10) The reconstructed phylogeny based on complete mitochondrial genomes recovered the family Conidae as a

monophyletic group. The genus *Profundiconus* was the sister group of the remaining genera, and the genus *Conus* was the sister group of a clade that included the genus *Conasprella* as sister group of the genus *Californiconus* and *Lilliconus* plus *Pseudolilliconus*. The divergence time of the main lineages within the family Conidae was estimated between the Palaeocene and the Eocene (56-30 mya), and the diversification of the species occurred in the transition from Oligocene to Miocene (23 mya).

6 BIBLIOGRAFÍA

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: Ponder, W.F., Lindberg, D.R. (Eds.), Phylogeny and Evolution of the Mollusca. University of California Press,, Berkeley, pp. 201-237.

Appeltans, W., Bouchet, P., Boxshall, G.A., Fauchald, K., Gordon, D.P., Hoeksema, B.W., Poore, G.C.B., Van Soest, R.W.M., Stöhr, S., Walter, T.C., 2012. World register of marine species. Accessed online: http://www. marinespecies. org (accessed on 28 February 2014).

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., 1997. Higher classification and pattern of evolution of the Gastropoda. Courier Forschungsinstitut Senckenberg 201, 57-81.

Bandel, K., 1999. Notes on the evolution and higher classification of the subclass Neritimorpha (Gastropoda) with the description of some new taxa. Geologica et Palaeontologica 33, 219-235.

Bandel, K., 2000. The new family Cortinellidae (Gastropoda, Mollusca) connected to a review of the evolutionary history of the subclass Neritimorpha. Neues Jahrbuch fur Geologie und Palaontologie-Abhandlungen 217, 111-130.

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.

Bandyopadhyay, P.K., Stevenson, B.J., Cady, M.T., Olivera, B.M., Wolstenholme, D.R., 2006. Complete mitochondrial DNA

sequence of a Conoidean gastropod, *Lophiotoma* (*Xenuroturris*) *cerithiformis*: gene order and gastropod phylogeny. Toxicon 48, 29-43.
Bandyopadhyay, P.K., Stevenson, B.J., Ownby, J.-P., Cady, M.T., Watkins, M., Olivera, B.M., 2008. The mitochondrial genome of *Conus textile*, *coxI–coxII* intergenic sequences and Conoidean evolution. Molecular phylogenetics and evolution 46, 215-223.
Barco, A., Evans, J., Schembri, P.J., Taviani, M., Oliverio, M., 2013. Testing the applicability of DNA barcoding for Mediterranean species of top-shells (Gastropoda, Trochidae, *Gibbula* s.l.). Marine Biology Research 9, 785-793.
Barghi, N., Concepcion, G.P., Olivera, B.M., Lluisma, A.O., 2015. Characterization of the complete mitochondrial genome of *Conus tribblei* Walls, 1977. Mitochondrial DNA, 1-2.
Barker, G.M., 2001. The biology of terrestrial molluscs. CABI publishing.
Bergsten, J., 2005. A review of long‐branch attraction. Cladistics 21, 163-193.
Bieler, R., 1990. Haszprunar's' clado-evolutionary'classification of the Gastropoda: a critique. Malacologia 31, 371-380.
Bieler, R., 1992. Gastropod phylogeny and systematics. Annual Review of Ecology and Systematics 23, 311-338.
Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic acids research 27, 1767-1780.
Boore, J.L., Brown, W.M., 1998. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. Current opinion in genetics & development 8, 668-674.
Boore, J.L., Macey, J.R., Medina, M., 2005. Sequencing and comparing whole mitochondrial genomes of animals. Methods in enzymology 395, 311-348.
Boss, K.J., 1982. Mollusca, p. 945-1166. Synopsis and Classification of Living Organisms 2.
Bouchet, P., Kantor, Y.I., Sysoev, A., Puillandre, N., 2011. A new operational classification of the Conoidea (Gastropoda). Journal of Molluscan Studies 77, 273-308.
Bouchet, P., Lozouet, P., Maestrati, P., Heros, V., 2002. Assessing the magnitude of species richness in tropical marine environments: exceptionally high numbers of molluscs at a

New Caledonia site. Biological Journal of the Linnean Society 75, 421-436.
Bouchet, P., Rocroi, J.-P., Frýda, J., Hausdorf, B., Ponder, W., Valdés, Á., Warén, A., 2005. Classification and nomenclator of gastropod families.
Brauer, A., Kurz, A., Stockwell, T., Baden-Tillson, H., Heidler, J., Wittig, I., Kauferstein, S., Mebs, D., Stöcklin, R., Remm, M., 2012. The mitochondrial genome of the venomous cone snail *Conus consors*. PloS one 7, e51528.
Cameron, S.L., Lambkin, C.L., Barker, S.C., Whiting, M.F., 2007. A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high precision. Systematic Entomology 32, 40-59.
Castelin, M., Puillandre, N., Kantor, Y.I., Modica, M.V., Terryn, Y., Cruaud, C., Bouchet, P., Holford, M., 2012. Macroevolution of venom apparatus innovations in auger snails (Gastropoda; Conoidea; Terebridae). Molecular phylogenetics and evolution 64, 21-44.
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, P.-W., Hsiao, S.-T., Huang, C.-W., Chen, K.-S., Tseng, C.-T., Wu, W.-L., Hwang, D.-F., 2015. The complete mitochondrial genome of *Conus tulipa* (Neogastropoda: Conidae). Mitochondrial DNA, 1-2.
Colgan, D.J., Ponder, W.F., Beacham, E., Macaranas, J., 2007. Molecular phylogenetics of Caenogastropoda (Gastropoda: Mollusca). Molecular phylogenetics and evolution 42, 717-737.
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.
Colgan, D.J., Ponder, W.F., Eggler, P.E., 2000. Gastropod evolutionary rates and phylogenetic relationships assessed using partial *28S rDNA* and *histone H3* sequences. Zoologica Scripta 29, 29-63.
Cox, L.R., 1960. Thoughts on the classification of the Gastropoda. Journal of Molluscan Studies 33, 239-261.

Crampton-Platt, A., Timmermans, M.J.T.N., Gimmel, M.L., Kutty, S.N., Cockerill, T.D., Khen, C.V., Vogler, A.P., 2015. Soup to tree: the phylogeny of beetles inferred by mitochondrial metagenomics of a Bornean rainforest sample. Molecular biology and evolution 32, 2302-2316.

Cunha, R.L., Castilho, R., Rüber, L., Zardoya, R., 2005. Patterns of cladogenesis in the venomous marine gastropod genus *Conus* from the Cape Verde Islands. Systematic biology 54, 634-650.

Cunha, R.L., Grande, C., Zardoya, R., 2009. Neogastropod phylogenetic relationships based on entire mitochondrial genomes. BMC Evolutionary Biology 9, 1.

Da Motta, A.J., 1991. A systematic classification of the gastropod family Conidae at the generic level. La Conchiglia.

Donald, K.M., Preston, J., Williams, S.T., Reid, D.G., Winter, D., Alvarez, R., Buge, B., Hawkins, S.J., Templado, J., Spencer, H.G., 2012. Phylogenetic relationships elucidate colonization patterns in the intertidal grazers Osilinus Philippi, 1847 and Phorcus Risso, 1826 (Gastropoda: Trochidae) in the northeastern Atlantic Ocean and Mediterranean Sea. Molecular phylogenetics and evolution 62, 35-45.

Duda, T.F., Bolin, M.B., Meyer, C.P., Kohn, A.J., 2008. Hidden diversity in a hyperdiverse gastropod genus: discovery of previously unidentified members of a *Conus* species complex. Molecular Phylogenetics and Evolution 49, 867-876.

Duda, T.F., Palumbi, S.R., 1999. Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. Proceedings of the National Academy of Sciences 96, 6820-6823.

Duda, T.F., Palumbi, S.R., 2004. Gene expression and feeding ecology: evolution of piscivory in the venomous gastropod genus *Conus*. Proceedings of the Royal Society of London B: Biological Sciences 271, 1165-1174.

Duda, T.F., Rolan, E., 2005. Explosive radiation of Cape Verde *Conus*, a marine species flock. Molecular Ecology 14, 267-272.

Espiritu, D.J.D., Watkins, M., Dia-Monje, V., Cartier, G.E., Cruz, L.J., Olivera, B.M., 2001. Venomous cone snails: molecular phylogeny and the generation of toxin diversity. Toxicon 39, 1899-1916.

Fretter, V., 1990. The anatomy of some new archaeogastropod limpets (Order Patellogastropoda, Suborder Lepetopsina) from hydrothermal vents. Journal of Zoology 222, 529-555.

Fretter, V., Graham, A., 1962. British prosobranch molluscs. Their functional anatomy and ecology. British prosobranch molluscs. Their functional anatomy and ecology.

Frey, M.A., Vermeij, G.J., 2008. Molecular phylogenies and historical biogeography of a circumtropical group of gastropods (Genus: *Nerita*): implications for regional diversity patterns in the marine tropics. Molecular Phylogenetics and Evolution 48, 1067-1086.

Frýda, J., 1999. Higher classification of Paleozoic gastropods inferred from their early shell ontogeny. Journal of GEOsciences 44, 137-154.

Frýda, J., Nützel, A., Wagner, P.J., 2008. Paleozoic gastropoda. Phylogeny and Evolution of the Mollusca, 239-270.

Fukumori, H., Kano, Y., 2014. Evolutionary ecology of settlement size in planktotrophic neritimorph gastropods. Marine biology 161, 213-227.

Geiger, D.L., Nützel, A., Sasaki, T., 2008. Vetigastropoda. Phylogeny and evolution of the mollusca. University of California Press, Berkeley, 297-330.

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

Gissi, C., Iannelli, F., Pesole, G., 2008. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity 101, 301-320.

Gofas, S., 2009a. Neritimorpha. In: MolluscaBase (2016). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p=taxdetails&id=156484 on 2016-06-09

Gofas, S., 2009b. Stomatellinae Gray, 1840. In: MolluscaBase (2016). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p=taxdetails&id=382176 on 2016-05-16

Gofas, S., 2015. Trochidae. In: MolluscaBase (2015). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p=taxdetails&id=443 on 2016-04-30

Gofas, S., Bouchet, P., 2015a. Gibbula Risso, 1826. In: MolluscaBase (2016). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p=taxdetails&id=138590 on 2016-05-16

Gofas, S., Bouchet, P., 2015b. Jujubinus Monterosato, 1884. In: MolluscaBase (2016). Accessed through: World Register of Marine Species at http://www.marinespecies.org/aphia.php?p=taxdetails&id=138591 on 2016-05-16

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

Grande, C., Templado, J., Cervera, J.L., Zardoya, R., 2002. The complete mitochondrial genome of the nudibranch *Roboastra europaea* (Mollusca: Gastropoda) supports the monophyly of opisthobranchs. Mol Biol Evol 19, 1672 - 1685.

Grande, C., Templado, J., Cervera, J.L., Zardoya, R., 2004a. Molecular phylogeny of Euthyneura (Mollusca: Gastropoda). Molecular Biology and Evolution 21, 303 - 313.

Grande, C., Templado, J., Cervera, J.L., Zardoya, R., 2004b. Phylogenetic relationships among Opisthobranchia (Mollusca: Gastropoda) based on mitochondrial *cox1*, *trnV*, and *rrnL* genes. Molecular phylogenetics and evolution 33, 378-388.

Grande, C., Templado, J., Zardoya, R., 2008. Evolution of gastropod mitochondrial genome arrangements. BMC Evol. Biol. 8, 61.

Haszprunar, G., 1985. The Heterobranchia–a new concept of the phylogeny of the higher Gastropoda. Journal of Zoological Systematics and Evolutionary Research 23, 15-37.

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 Studies 54, 367-441.

Haszprunar, G., 1990. Towards a phylogenetic system of Gastropoda. 1. Traditional methodology-a reply. Malacologia, pp. 195-202.

Haszprunar, G., Kunze, T., Brückner, M., Heß, M., 2016. Towards a sound definition of Skeneidae (Mollusca, Vetigastropoda): 3D

interactive anatomy of the type species, *Skenea serpuloides* (Montagu, 1808) and comments on related taxa. Organisms Diversity & Evolution, 1-19.
Haszprunar, G., Wanninger, A., 2012. Molluscs. Current Biology 22, R510-R514.
Hatzoglou, E., Rodakis, G.C., Lecanidou, R., 1995. Complete sequence and gene organization of the mitochondrial genome of the land snail *Albinaria coerulea*. Genetics 140, 1353-1366.
Havird, J.C., Santos, S.R., 2014. Performance of single and concatenated sets of mitochondrial genes at inferring metazoan relationships relative to full mitogenome data. PloS one 9, e84080.
Heller, J., 2015. Sea Snails: A natural history. Springer.
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., 1984. A new archaeogastropod (Rhipidoglossa, Trochacea) from hydrothermal vents on the East Pacific Rise. Zoologica Scripta 13, 19-25.
Hickman, C.S., 2013. Intaglicollonia, a new name for the Eocene gastropod *Nehalemia* Hickman, 1974 (Gastropoda: Coloniidae). Bailey-Matthews shell museum c/o Dr. Jose H Leal, Associate/Managing Editor, 3075 Sanibel-Captiva Rd, Sanibel, fl 33957
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.
Holford, M., Puillandre, N., Terryn, Y., Cruaud, C., Olivera, B., Bouchet, P., 2009. Evolution of the Toxoglossa venom apparatus as inferred by molecular phylogeny of the Terebridae. Molecular Biology and Evolution 26, 15-25.
Holthuis, B.V., 1995. Evolution between marine and freshwater habitats: a case study of the gastropod suborder Neritopsina. University of Washington.
Irisarri, I., San Mauro, D., Abascal, F., Ohler, A., Vences, M., Zardoya, R., 2012. The origin of modern frogs (Neobatrachia) was accompanied by acceleration in mitochondrial and nuclear substitution rates. BMC genomics 13, 1.

Kaim, A., Sztajner, P., 2005. The opercula of neritopsid gastropods and their phylogenetic importance. Journal of Molluscan Studies 71, 211-219.
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., Chiba, S., Kase, T., 2002. Major adaptive radiation in neritopsine gastropods estimated from *28S rRNA* sequences and fossil records.
Kantor, Y.I., 1996. Phylogeny and relationships of Neogastropoda. Origin and evolutionary radiation of the mollusca. Oxford University Press, Oxford, 221-230.
Kenchington, E.L., Roddick, D.L., Singh, R.K., Bird, C.J., 1994. Analysis of small-subunit rRNA gene sequences from six families of molluscs. Journal of Marine Biotechnology 1, 215-217.
Kiel, K.B.S., 2003. Relationships of Cretaceous Neritimorpha (Gastropoda, Mollusca), with the description of seven new species. Bulletin of the Czech Geological survey 78, 53-65.
Knight, J.B., Cox, L.R., Keen, A.M., Smith, A.G., Batten, R.L., Yochelson, E.L., Ludbrook, N.H., Robertson, R., Yonge, C.M., Moore, R.C., 1960. Treatise on Invertebrate Paleontology. Pt. I, Mollusca I. Geological Society of America and University of Kansas Press, Lawrence.
Kocot, K.M., Cannon, J.T., Todt, C., Citarella, M.R., Kohn, A.B., Meyer, A., Santos, S.R., Schander, C., Moroz, L.L., Lieb, B., 2011. Phylogenomics reveals deep molluscan relationships. Nature 477, 452-456.
Kohn, A.J., 1990. Tempo and mode of evolution in Conidae. Malacologia 32, 55-67.
Kraus, N.J., Corneli, P.S., Watkins, M., Bandyopadhyay, P.K., Seger, J., Olivera, B.M., 2011. Against expectation: a short sequence with high signal elucidates cone snail phylogeny. Molecular phylogenetics and evolution 58, 383-389.
Kraus, N.J., Watkins, M., Bandyopadhyay, P.K., Seger, J., Olivera, B.M., Corneli, P.S., 2012. A very short, functionally constrained sequence diagnoses cone snails in several *Conasprella* clades. Molecular phylogenetics and evolution 65, 335-338.

Kunze, T., Heß, M., Haszprunar, G., 2016. 3D-interactive microanatomy of *Ventsia tricarinata* Warén & Bouchet, 1993 (Vetigastropoda: Seguenzioidea) from Pacific hydrothermal vents. Journal of Molluscan Studies, eyw002.
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, S4.
Lartillot, N., Philippe, H., 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Molecular biology and evolution 21, 1095-1109.
Lavrov, D.V., 2007. Key transitions in animal evolution: a mitochondrial DNA perspective. Integrative and Comparative Biology 47, 734-743.
Lindberg, D.R., 2008. Patellogastropoda, Neritimorpha and Cocculinoidea: the low-diversity gastropod clades. Phylogeny and evolution of the mollusca. University of California Press, Berkeley, CA.
Maynard, B.T., Kerr, L.J., McKiernan, J.M., Jansen, E.S., Hanna, P.J., 2005. Mitochondrial DNA sequence and gene organization in the Australian blacklip abalone Haliotis rubra (Leach). Marine Biotechnology 7, 645-658.
McArthur, A., Harasewych, M., 2003. Molecular systematics of the major lineages of the Gastropoda. In: Lydeard, C., Lindberg, D.R. (Eds.), Molecular Systematics and Phylogeography of Mollusks. Smithsonian Books, Washington DC, pp. 140-160.
McArthur, A.G., Koop, B.F., 1999. Partial *28S rDNA* sequences and the antiquity of hydrothermal vent endemic gastropods. Mol Phylogenet Evol 13, 255-274.
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. Philosophical Transactions of the Royal Society of London B: Biological Sciences 319, 1-32.
McLean, J.H., 1989. New archaeogastropod limpets from hydrothermal vents: new family Peltospiridae, new superfamily Peltospiracea. Zoologica Scripta 18, 49-66.
McLean, J.H., 1990. Neolepetopsidae, a new docoglossate limpet family from hydrothermal vents and its relevance to patellogastropod evolution. Journal of Zoology 222, 485-528.

Medina, M., Lal, S., Vallès, Y., Takaoka, T.L., Dayrat, B.A., Boore, J.L., Gosliner, T., 2011. Crawling through time: transition of snails to slugs dating back to the Paleozoic, based on mitochondrial phylogenomics. Marine genomics 4, 51-59.
Miya, M., Nishida, M., 2015. The mitogenomic contributions to molecular phylogenetics and evolution of fishes: a 15-year retrospect. Ichthyological Research 62, 29-71.
Nam, H.H., Corneli, P.S., Watkins, M., Olivera, B., Bandyopadhyay, P., 2009. Multiple genes elucidate the evolution of venomous snail-hunting *Conus* species. Molecular phylogenetics and evolution 53, 645-652.
Nützel, A., Frýda, J., 2003. Paleozoic plankton revolution: evidence from early gastropod ontogeny. Geology 31, 829-831.
Nützel, A., Lehnert, O., Frýda, J.í., 2006. Origin of planktotrophy—evidence from early molluscs. Evolution & development 8, 325-330.
Olivera, B.M., Showers Corneli, P., Watkins, M., Fedosov, A., 2014. Biodiversity of cone snails and other venomous marine gastropods: Evolutionary success through neuropharmacology. Annu. Rev. Anim. Biosci. 2, 487-513.
Osca, D., 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, 197.
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.
Osca, D., Templado, J., Zardoya, R., 2015. Caenogastropod mitogenomics. Molecular phylogenetics and evolution 93, 118-128.
Parkhoev, P.Y., 2008. The early Cambrian radiation of Mollusca. Phylogeny and Evolution of the Mollusca 33.
Ponder, W.F., Colgan, D.J., Healy, J.M., Nützel, A., Simone, L.R.L., Strong, E.E., 2008. Caenogastropoda. Phylogeny and evolution of the Mollusca, 331-383.
Ponder, W.F., Lindberg, D.R., 1996. Gastropod phylogeny—challenges for the 90s. Origin and evolutionary radiation of the Mollusca, 135-154.
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.
Ponder, W.F., Warén, A., 1988. ClassiWcation of the Caenogastropoda and Heterostropha – a list of the family-group names and higher taxa. In: Ponder, W.F. (Ed.), Prosobranch Phylogeny, Malacological Review Supplement 4, pp. 288-328.
Puillandre, N., Bouchet, P., Duda Jr, T.F., Kauferstein, S., Kohn, A.J., Olivera, B.M., Watkins, M., Meyer, C., 2014a. Molecular phylogeny and evolution of the cone snails (Gastropoda, Conoidea). Molecular Phylogenetics and Evolution 78, 290-303.
Puillandre, N., Duda, T.F., Meyer, C., Olivera, B.M., Bouchet, P., 2014b. One, four or 100 genera? A new classification of the cone snails. Journal of Molluscan Studies.
Puillandre, N., Kantor, Y.I., Sysoev, A., Couloux, A., Meyer, C., Rawlings, T., Todd, J.A., Bouchet, P., 2011. The dragon tamed? A molecular phylogeny of the Conoidea (Gastropoda). Journal of Molluscan Studies 77, 259-272.
Puillandre, N., Samadi, S., Boisselier, M.C., Sysoev, A.V., Kantor, Y.I., Cruaud, C., Couloux, A., Bouchet, P., 2008. Starting to unravel the toxoglossan knot: molecular phylogeny of the “turrids”(Neogastropoda: Conoidea). Molecular phylogenetics and evolution 47, 1122-1134.
Robinson, N.A., Hall, N.E., Ross, E.M., Cooke, I.R., Shiel, B.P., Robinson, A.J., Strugnell, J.M., 2016. The complete mitochondrial genome of *Haliotis laevigata* (Gastropoda: Haliotidae) using MiSeq and HiSeq sequencing. Mitochondrial DNA 27, 437-438.
Rosenberg, G., 2014. A new critical estimate of named species-level diversity of the recent Mollusca. American Malacological Bulletin 32, 308-322.
Rosenberg, G., Kuncio, R.S., Davis, G.M., Harasewych, M.G., Harasewych, M.G., Tillier, S., 1994. Ribosomal RNA phylogeny of selected gastropod and unionacean bivalve mollusks. Nautilus 108, 111-121.
Salvini‐Plawen, L.V., 1980. A reconsideration of systematics in the Mollusca (phylogeny and higher classification). Malacologia: International Journal of Malacology.

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

Salvini‐Plawen, L.V., Steiner, G., 1996. Synapomorphies and plesiomorphies in higher calssification of Mollusc. In: Taylor, J.D. (Ed.), Origins and evolutionary radiation of the Mollusca. Oxford University Press, Oxford, pp. 29-51.

San Mauro, D., Gower, D.J., Müller, H., Loader, S.P., Zardoya, R., Nussbaum, R.A., Wilkinson, M., 2014. Life-history evolution and mitogenomic phylogeny of caecilian amphibians. Molecular phylogenetics and evolution 73, 177-189.

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

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

Sevigny, J.L., Kirouac, L.E., Thomas, W.K., Ramsdell, J.S., Lawlor, K.E., Sharifi, O., Grewal, S., Baysdorfer, C., Curr, K., Naimie, A.A., Okamoto, K., Murray, J.A., Newcomb, J.M., 2015. The mitochondrial genomes of the nudibranch mollusks, *Melibe leonina* and *Tritonia diomedea*, and their impact on gastropod phylogeny. PLoS ONE 10, e0127519.

Shen, X., Tian, M., Yan, B., Chu, K., 2015. Phylomitogenomics of Malacostraca (Arthropoda: Crustacea). Acta Oceanologica Sinica 34, 84-92.

Smith, S.A., Wilson, N.G., Goetz, F.E., Feehery, C., Andrade, S.C.S., Rouse, G.W., Giribet, G., Dunn, C.W., 2011. Resolving the evolutionary relationships of molluscs with phylogenomic tools. Nature 480, 364-367.

Stöger, I., Schrödl, M., 2013. Mitogenomics does not resolve deep molluscan relationships (yet?). Mol. Phylogenet. Evol. 69, 376-392.

Strong, E.E., 2003. Refining molluscan characters: morphology, character coding and a phylogeny of the Caenogastropoda. Zoological Journal of the Linnean Society 137, 447-554.

Taylor, J.D., Kantor, Y.I., Sysoev, A.V., 1993. Foregut anatomy, feeding mechanisms, relationships and classification of the Conoidea. Bulletin of the Natural History Museum. Zoology series 59, 125-170.

Templado, J., Ortea, J., 2001. The occurrence of the shell-less neritacean gastropod *Titiscania limacina* in the Galapagos Islands. Veliger 44, 404-406.
Thiele, J., 1929–35. Handbuch der Systematischen Weichtierkunde. Gustav Fischer, Jena.
Tillier, S, Masselot, M, Philippe, H, Tillier, A, 1992. Phylogénie moléculaire des Gastropoda (Mollusca) fondée sur le séquençage partiel de l'ARN ribosomique 28S. Elsevier, Paris, FRANCE.
Timmermans, M.J.T.N., Lees, D.C., Simonsen, T.J., 2014. Towards a mitogenomic phylogeny of Lepidoptera. Molecular phylogenetics and evolution 79, 169-178.
Tracey, S., Todd, J.A., Erwin, D.H., 1993. Mollusca: gastropoda. The fossil record 2, 131-167.
Tucker, J.K., Tenorio, M.J., 2009. Systematic classification of recent and fossil conoidean gastropods: with keys to the genera of cone shells. Conchbooks.
Tucker, J.K., Tenorio, M.J., 2013. Illustrated catalog of the living cone shells. MdM Publishing.
Vaught, K.C., 1989. A classification of the living Mollusca. Amer Malacologists.
Wagner, P.J., 2001. Gastropod phylogenetics: progress, problems, and implications. Journal of Paleontology 75, 1128-1140.
Warén, A., Bengtson, S., Goffredi, S.K., Van Dover, C.L., 2003. A hot-vent gastropod with iron sulfide dermal sclerites. Science 302, 1007-1007.
Wenz., W., 1938-1944. Gastropoda, Teil 1: Allgemeiner Teil und Prosobranchia. In: Schindewolf (Ed.), Handbuch der Palaozoologie. Gebruder Borntrager, Berlin, pp. 1-240.
White, T.R., Conrad, M.M., Tseng, R., Balayan, S., Golding, R., de Frias Martins, A., Dayrat, B.A., 2011. Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. BMC evolutionary biology 11, 1.
Williams, S.T., 2012. Advances in molecular systematics of the vetigastropod superfamily Trochoidea. Zoologica Scripta 41, 571-595.
Williams, S.T., Donald, K.M., Spencer, H.G., Nakano, T., 2010. Molecular systematics of the marine gastropod families Trochidae and Calliostomatidae (Mollusca: Superfamily

Trochoidea). Molecular Phylogenetics and Evolution 54, 783-809.
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). Mol Phylogenet Evol 39, 33-51.
Winnepenninckx, B., Steiner, G., Backeljau, T., De Wachter, R., 1998. Details of gastropod phylogeny inferred from *18S rRNA* sequences. Molecular phylogenetics and evolution 9, 55-63.
Xu, W., Jameson, D., Tang, B., Higgs, P.G., 2006. The relationship between the rate of molecular evolution and the rate of genome rearrangement in animal mitochondrial genomes. Journal of Molecular Evolution 63, 375-392.
Yang, X., Cameron, S.L., Lees, D.C., Xue, D., Han, H., 2015. A mitochondrial genome phylogeny of owlet moths (Lepidoptera: Noctuoidea), and examination of the utility of mitochondrial genomes for lepidopteran phylogenetics. Molecular phylogenetics and evolution 85, 230-237.
Yu, Z., Wei, Z., Kong, X., Shi, W., 2008. Complete mitochondrial DNA sequence of oyster *Crassostrea hongkongensis-a* case of " Tandem duplication-random loss" for genome rearrangement in *Crassostrea*? Bmc Genomics 9, 1.
Zapata, F., Wilson, N.G., Howison, M., Andrade, S.C.S., Jörger, K.M., Schrödl, M., Goetz, F.E., Giribet, G., Dunn, C.W., 2014. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. Proc. R. Soc. Lond. B 281.

7 APÉNDICES

A continuación se muestran los apéndices correspondientes a los diferentes capítulos (publicaciones) de la sección de resultados.

SUPPORTING INFORMATION 1

Data S1. 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
ANcox3Rln	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

Data S2. Complete mt genomes retrieved from Gen-Bank 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

Arquez, M., Colgan, D., and 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., and 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., and Zardoya, R. (2009). Neogastropod phylogenetic relationships based on entire mitochondrial genomes. BMC Evolutionary Biology 9, 210.

Grande, C., Templado, J., Lucas Cervera, J., and 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., and 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., and 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., and 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., and 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., and 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 mitocondrial genomes of pulmonates (Mollusca: Gastropoda) and their impact on phylogenetic relationships. BMC Evolutionary Biology 11, 295.

Williams, S. T., Foster, P. G., and 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.

Data S3. 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	—

Data S4. 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

%A-T was calculated only for intergenic regions >20 pb

Data S4 (cont.)

Granata lyrata								
Name Gene	**Type**	**Start**	**Stop**	**Length**	**Codon**		**Strand**	**%A-T**
					start	**stop**		
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

Data S4 (cont.)

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

Data S4 (cont.)

Diodora graeca								
Name Gene	**Type**	**Start**	**Stop**	**Length**	**Codon**		**Strand**	**%A-T**
					start	**stop**		
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

Data S4 (cont.)

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

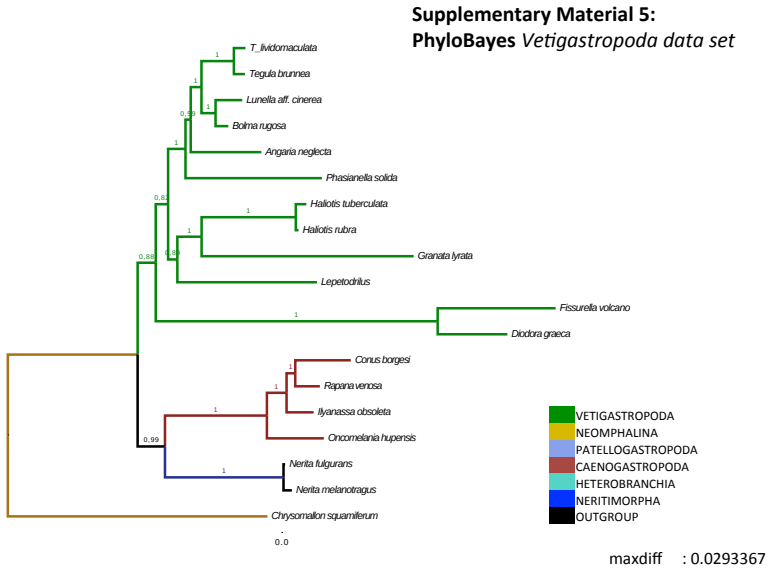
Data S4 (cont.)

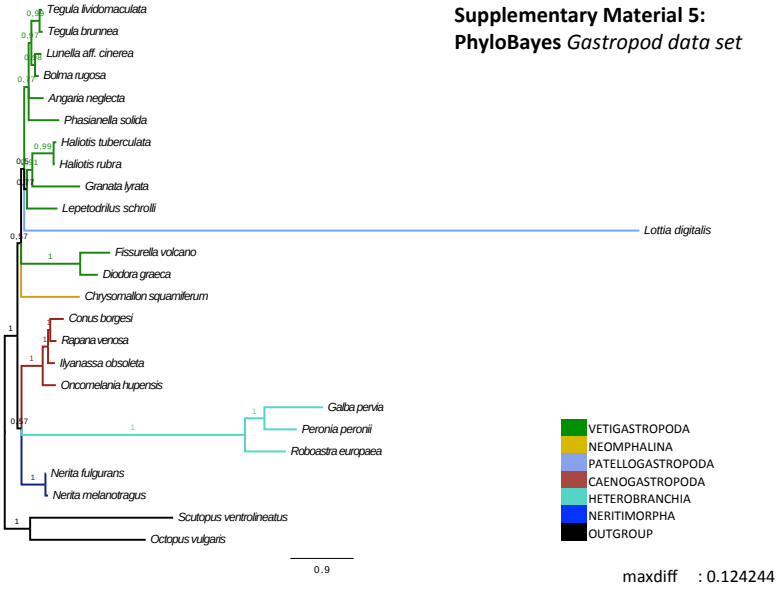
Tegula lividomaculata								
Name Gene	**Type**	**Start**	**Stop**	**Length**	**Codon**		**Strand**	**%A-T**
					start	**stop**		
cox1	CDS	1	1536	1.536	ATG	TAA	forward	
	Intergenic			72				83.3
cox2	CDS	1609	2304	696	ATG	TAA	forward	
	Intergenic			158				80.4
trnD	tRNA	2463	2537	75			forward	
atp8	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

Data S4 (cont.)

Lepetodrilus schroli								
Name Gene	**Type**	**Start**	**Stop**	**Length**	**Codon**		**Strand**	**%A-T**
					start	**stop**		
	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

Data S5. PhyloBayes of gastropod data set.





SUPPORTING INFORMATION 2

Data S1. Amplification strategy. Long PCR and primer walking primers.

Cittarium pica

Long PCR		
Primer	Sequence 5'-3'	Fragment (bp)
Citt-cox1-F	TGGTTAATTCCTCTGATATTGGGAGCTCC	*cox1-rrnL* (11436)
TROmt16sF	GATAACAGCGTAATCTTTCTGGAGAGATC	
TROmt16sR	AAGCTCAACAGGGTCTTCTTGTCCC	*rrnL-cox3* (3439)
85CPcox3R	CATAGACACCATCTGAGATAGTTAACGG	
85CPcox3F	GAGCTTATTTTCATAGAAGTCTCGCTTC	*cox3-cox1* (3580)
Citt-cox1-R	GCAGGATCAAAGAAGGATGTGTTAAAATTTC	

Margarites vorticiferus

Long PCR		
Primer	Sequence 5'-3'	Fragment (bp)
Alecox3F_UJ	CTGAGCATATTTCCATAGAAGCCTGGC	*cox3-cox1* (3074)
Alecox1R	CTGATCAAGTGAATAGTGGTAGGCGTTC	
Alecox1F	CTTAGTTTTCGGGATTTGAGCAGGCC	*cox1-rrnS* (12686)
Ale12SF	TTTAAATCCTTCCAGGGGAACCTGTCC	

Gibbula umbilicaris

Long PCR		
Primer	Sequence 5'-3'	Fragment (bp)
GVcox3F	TTTCCACAGAAGACTTGCTCCTACTCC	*cox3-cox1* (2867)
G2-cox1-r	AATAGAAGAAACACCYGCTAAGTGAAGGGA	
G2-cox1-F	CCGGTGCTATTACTATGCTGCTCACTGA	*cox1-rrnL* (9870)
TROmt16sF	GATAACAGCGTAATCTTTCTGGAGAGATC	

Tectus virgatus

Long PCR		
Primer	Sequence 5'-3'	Fragment (bp)
TVcox3F	GTATTTCCACAGAAGGTTGGCTTCTGC	*cox3-cox1* (3567)
TVcox1R	GAAGAGATAGCAGCAACAAAATGGCCGT	
TVcox1F	GCATTTCCGCGACTTAATAACATGAGATT	*cox1-rrnL* (10646)
TROmt16sF	GATAACAGCGTAATCTTTCTGGAGAGATC	

Data S2. Mitochondrial genome features.

Cittarium pica

		Gene			Codon		
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Strand**
cox1	CDS	1	1539	1539	ATG	TAA	forward
cox2	CDS	1669	2364	696	ATG	TAG	forward
trnD	tRNA	2535	2605	71			forward
atp8	CDS	2608	2785	178	—	TAA	forward
atp6	CDS	2975	3670	696	ATG	TAA	forward
trnF	tRNA	3730	3799	70			reverse
nad5	CDS	3937	5683	1747	ATG	TAA	reverse
trnH	tRNA	5684	5754	71			reverse
nad4	CDS	5805	7202	1398	ATG	TAA	reverse
nad4L	CDS	7196	7495	300	ATG	TAG	reverse
trnT	tRNA	7600	7669	70			forward
trnS (tga)	trna	7803	7870	68			reverse
cob	CDS	7902	9041	114	ATG	TAA	reverse
nad6	CDS	9231	9737	507	ATG	TAA	reverse
trnP	tRNA	9741	9810	70			reverse
nad1	CDS	9953	10903	951	ATG	TAA	reverse
trnL (taa)	tRNA	10905	10972	68			reverse
trnL (tag)	tRNA	11281	11348	68			reverse
rrnL	rRNA	11349	13012	1664			reverse
trnV	tRNA	13013	13086	74			reverse
rrnS	rRNA	13087	14153	1067			reverse
trnM	tRNA	14154	14222	69			reverse
trnY	tRNA	14312	14379	68			reverse
trnC	tRNA	14381	14447	67			reverse
trnW	tRNA	14455	14521	67			reverse
trnQ	tRNA	14539	14607	69			reverse
trnG	tRNA	14618	14686	69			forward
cox3	CDS	14740	15519	780	ATG	TAA	forward
trnK	tRNA	15718	15777	60			forward
trnA	tRNA	15778	15845	68			forward
trnR	tRNA	15906	15974	69			forward
trnN	tRNA	15989	16061	73			forward
trnI	tRNA	16109	16179	71			forward
nad3	CDS	16184	16537	354	ATG	TAA	forward
trnS (cgt)	tRNA	16626	16692	67			forward
nad2	CDS	16696	17949	1254	ATG	T—	forward

Margarites vorticiferus

		Gene			Codon		
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Strand**
cox1	CDS	2522	4066	1545	ATG	TAA	forward
cox2	CDS	4117	4844	728	ATA	TAA	forward
trnD	tRNA	4875	4941	67			forward
atp8	CDS	4943	5122	180	ATG	TAG	forward
atp6	CDS	5200	5895	696	ATG	TAA	forward
trnF	tRNA	5928	5993	66			reverse
nad5	CDS	6012	7765	1754	ATG	TAG	reverse
trnH	tRNA	7766	7831	66			reverse
nad4	CDS	7894	9282	1389	GTG	TAA	reverse
nad4L	CDS	9276	9575	300	ATG	TAG	reverse
trnT	tRNA	9635	9705	71			forward
trnS (tga)	tRNA	9710	9776	67			reverse
cob	CDS	9786	10925	114	ATG	TAA	reverse
nad6	CDS	10997	11500	504	ATG	TAA	reverse
trnM	tRNA	11500	11563	64			reverse
trnP	tRNA	11641	11709	69			reverse
nad1	CDS	11769	12716	948	ATG	TAA	reverse
trnL (taa)	tRNA	12718	12785	68			reverse
trnL (tag)	tRNA	12811	12878	68			reverse
rrnL	rRNA	12879	14458	158			reverse
trnV	tRNA	14459	14527	69			reverse
rrnS	rRNA	14528	15254	727			reverse
cox3	CDS	1	484	484	—	TAA	forward
trnK	tRNA	520	578	59			forward
trnA	tRNA	579	650	72			forward
trnR	tRNA	690	751	62			forward
trnN	tRNA	752	821	70			forward
trnI	tRNA	828	894	67			forward
nad3	CDS	897	125	354	ATG	TAG	forward
trnS (cgt)	tRNA	1271	1338	68			forward
nad2	CDS	1343	2499	1157	TGT	TAA	forward

Gibbula umbilicaris

		Gene			Codon		
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Strand**
cox1	CDS	2419	3954	1536	ATG	TAA	forward
cox2	CDS	3983	4675	693	ATG	TAA	forward
trnD	tRNA	4706	4770	65			forward
atp8	CDS	4772	4936	165	ATG	TAA	forward
atp6	CDS	4981	5679	699	ATG	TAG	forward
trnF	tRNA	5712	5776	65			forward
nad5	CDS	5801	7555	1755	ATG	TAA	reverse
trnH	tRNA	7556	7623	68			reverse
nad4	CDS	7700	9088	1389	ATG	TAG	reverse
nad4L	CDS	9082	9381	300	ATG	TAG	reverse
trnT	tRNA	9406	9474	69			reverse
trnS (tga)	tRNA	9503	9569	67			reverse
cob	CDS	9601	1074	114	ATG	TAA	reverse
nad6	CDS	10856	11362	507	ATG	TAA	reverse
trnP	tRNA	11366	11431	66			reverse
nad1	CDS	11494	12435	942	ATG	TAG	reverse
trnL (taa)	tRNA	12437	12504	68			reverse
trnL (tag)	tRNA	12530	12597	68			reverse
rrnL	rRNA	12598	12886	289			reverse
cox3	CDS	1	479	479	—	TAA	forward
trnK	tRNA	513	570	58			forward
trnA	tRNA	571	638	68			forward
trnR	tRNA	641	709	69			forward
trnN	tRNA	716	782	67			forward
trnI	tRNA	784	850	67			forward
nad3	CDS	855	1208	354	ATG	TAA	forward
trnS (cgt)	tRNA	1218	1285	68			forward
nad2	CDS	1289	2418	113	ATG	T—	forward

Tectus virgatus

		Gene			Codon		
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Strand**
cox1	CDS	2979	4514	1536	ATG	TAA	forward
cox2	CDS	4586	5276	691	ATG	TAA	forward
trnD	tRNA	5453	5520	68			forward
atp8	CDS	5521	5706	186	ATG	TAG	forward
atp6	CDS	5866	6561	696	ATG	TAA	forward
trnF	tRNA	6596	6665	70			reverse
nad5	CDS	6782	8515	1734	ATG	TAA	reverse
trnH	tRNA	8516	8584	69			reverse
nad4	CDS	8682	10079	1398	ATG	TAA	reverse
nad4L	CDS	10073	10372	300	ATG	TAA	reverse
trnT	tRNA	10429	10497	69			forward
trnS (tga)	tRNA	10553	10618	66			reverse
cob	CDS	10629	11768	114	ATG	TAA	reverse
trnQ	tRNA	11767	11831	65			reverse
nad6	CDS	11834	12340	507	ATG	TAA	reverse
trnP	tRNA	12345	12415	71			reverse
nad1	CDS	12497	13447	951	ATG	TAG	reverse
trnL (taa)	tRNA	13449	13516	68			reverse
trnL (tag)	tRNA	13561	13628	68			reverse
rrnL	rRNA	13629	13891	263			reverse
cox3	CDS	1	547	547	—	TAA	forward
trnK	tRNA	734	791	58			forward
trnA	tRNA	792	859	68			forward
trnR	tRNA	916	985	70			forward
trnN	tRNA	103	1098	69			forward
trnI	tRNA	1146	1215	70			forward
nad3	CDS	1219	1572	354	ATG	TAG	forward
trnS (cgt)	tRNA	1703	1769	67			forward
nad2	CDS	1774	2978	1205	ATG	T—	forward

SUPPORTING INFORMATION 3

Data S1. Species analyzed in this study

Subfamily	Species	Expedition	Locality	Voucher No	28S	COI	16S	12S	H3	COB	PHOTO
Cantharidinae											
	Gibbula umbilicaris (Linnaeus, 1758) 1	Mediterranean 2005	El Mohon, Murcia, SE Spain	MNCN:ADN:86692	KX277459	KX277553	KX277387	KX277313	–	–	
	Gibbula umbilicaris 2	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86693	KX277460	KX277554	KX277388	KX277314	KX277612	KX277506	
	Gibbula cineraria (Linnaeus, 1758)	Galicia 2005	A Guarda, Pontevedra, NW Spain	MNCN:ADN:86694	KX277461	KX277555	KX277389	KX277315	KX277613	–	
	Gibbula pennanti (Philippi, 1846) 1	—	Roscoff, W France	NHMUK 20080944	GQ232393	GQ232365	GQ232295	GQ232330	–	–	MorphoBank
	Gibbula pennanti 2	Galicia 2005	A Guarda, Pontevedra, NW Spain	MNCN:ADN:86695	KX277462	KX277556	KX277390	KX277316	KX277614	KX277507	
	Gibbula racketti (Payraudeau, 1826) 1	Mediterranean 2005	Islas Chafarinas, S Spain	MNCN:ADN:86696	KX277463	KX277557	KX277391	KX277317	KX277615	KX277508	
	Gibbula racketti 2	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86697	KX277464	KX277558	KX277392	KX277318	–	–	
	Gibbula albida (Gmelin, 1791)	Galicia 2005	O Grove, Pontevedra, NW Spain	MNCN:ADN:86698	KX277465	KX277559	KX277393	KX277319	KX277616	–	
	Gibbula adriatica (Philippi, 1844)	Galicia 2005	O Grove, Pontevedra, NW Spain	MNCN:ADN:86699	KX277466	KX277560	KX277394	KX277320	KX277617	–	
	Gibbula varia (Linnaeus, 1758) 1	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86700	KX277467	KX277561	KX277395	KX277321	–	–	
	Gibbula varia 2	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86701	KX277468	KX277562	KX277396	KX277322	–	–	
	Gibbula varia 3	Mediterranean 2005	Islas Chafarinas, S Spain	MNCN:ADN:86702	KX277469	KX277563	KX277397	KX277323	KX277618	KX277509	
	Gibbula divaricata (Linnaeus, 1758)	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86703	KX277470	KX277564	KX277398	KX277324	KX277619	–	
	Gibbula rarilineata (Michaud, 1829) 1	Croatia 2011	Dugi otok, Telaščica, Croatia	MNCN:ADN:86704	KX277471	KX277565	KX277399	KX277325	indv 2*	indv 2*	
	Gibbula rarilineata 2	—	Lumbarda, Korcula I., Croatia	NHMUK 20080375	GQ232394	GQ232366	GQ232296	–	KX277610	KX277504	MorphoBank
	Gibbula umbilicalis (da Costa, 1778) 1	Galicia 2005	O Grove, Pontevedra, NW Spain	MNCN:ADN:86705	indv 2*	KX277566	KX277400	KX277326	KX277620	KX277510	
	Gibbula umbilicalis 2	Galicia 2005	A Guarda, Pontevedra, NW Spain	MNCN:ADN:86706	KX277472	KX277567	KX277401	KX277327	–	–	
	Gibbula umbilicalis 3	—	Wembury, Plymouth, UK	NHMUK 20080946	GQ232395	GQ232367	GQ232297	–	–	–	–
	Gibbula umbilicalis 4	Galicia 2005	O Grove, Pontevedra, NW Spain	MNCN:ADN:86707	KX277473	KX277568	KX277402	KX277328	–	–	
	Phorcus richardi (Payraudeau, 1826) 1	—	Almuñecar, Granada, S Spain	See Donald et al. 2012	JN686264	JN686361	JN686181	–	–	–	–
	Phorcus richardi 2	Mediterranean 2005	Cala Rajá, Cabo de Gata, SE Spain	MNCN:ADN:86708	KX277474	KX277569	KX277403	KX277329	KX277621	KX277511	
	Phorcus articulatus (Lamarck, 1822) 1	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86709	KX277475	KX277570	KX277404	KX277330	–	–	
	Phorcus articulatus 2	Mediterranean 2005	Islas Chafarinas, S Spain	MNCN:ADN:86710	KX277476	KX277571	KX277405	KX277331	KX277622	KX277512	
	Phorcus mutabilis (Philippi, 1846)	—	Fano, Italy	See Donald et al. 2012	JN686268	JN686364	JN686179	–	–	–	–
	Phorcus atratus (Wood, 1828)	—	La Golfa, Lanzarote, Canary Islands	See Donald et al. 2012	JN686209	JN686295	JN686113	–	–	–	–
	Phorcus lineatus (da Costa, 1778) 1	Galicia 2005	A Guarda, Pontevedra, NW Spain	MNCN:ADN:86711	JN686232	KX277572	KX277406	KX277332	KX277623	KX277513	–
	Phorcus lineatus 2	—	Pembrokeshire, Wales, UK	See Donald et al. 2012	JN686241	JN686328	JN686125	–	–	–	–
	Phorcus punctulatus (Lamarck, 1822)	—	Dakar, Senegal	See Donald et al. 2012	JN686259	JN686350	JN686158	–	–	–	–
	Phorcus mariae Templado & Rolán, 2012	—	Mordeira Bay, Sal, Cape Verde Islands	See Donald et al. 2012	JN686249	JN686337	JN686156	–	–	–	–
	Phorcus sauciatus (Koch, 1845) 1	Galicia 2005	Ría de Aldán, Pontevedra, NW Spain	MNCN:ADN:86712	KX277477	KX277573	KX277407	KX277333	–	–	
	Phorcus sauciatus 2	Azores & Madeira 2008	Funchal, Madeira	MNCN:ADN:86713	JN686227	KX277574	KX277408	KX277334	–	–	
	Phorcus sauciatus 3	Galicia 2005	Cabo Estai, Canido, Pontevedra, NW Spain	MNCN:ADN:86714	KX277478	KX277575	KX277409	KX277335	KX277624	KX277514	
	Phorcus turbinatus (Born, 1778) 1	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86715	JN686167	JN686340	KX277410	KX277336	KX277625	KX277515	
	Phorcus turbinatus 2	—	Coral Bay, Cyprus	See Donald et al. 2012	GQ434030	GQ434018	GQ434024	GQ434013	–	–	–
	Jujubinus exasperatus (Pennant, 1777) 1	Balearic Islands 2007	Cala San Esteve, Menorca, E Spain	MNCN:ADN:86716	KX277479	KX277576	KX277411	KX277337	–	–	
	Jujubinus exasperatus 2	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86717	KX277480	KX277577	KX277412	KX277338	KX277626	KX277516	
	Jujubinus exasperatus 3	—	Lumbarda, Korcula I., Croatia	NHMUK 20080387	GQ232396	GQ232368	GQ232298	GQ232333	–	–	MorphoBank
	Jujubinus exasperatus 4	Galicia 2005	Cabo Estai, Canido, Pontevedra, NW Spain	MNCN:ADN:86718	KX277481	KX277578	KX277413	KX277339	–	–	
	Jujubinus exasperatus 5	—	Ile Callot, France	NHMUK 20080943	GQ232387	GQ232361	GQ232288	GQ232323	–	–	–

Data S1 (cont.)

Jujubinus pseudogravinae Nordsieck, 1973	Azores & Madeira 2008	San Miguel, Azores	MNCN:ADN:86719	KX277482	KX277579	KX277414	KX277340	KX277627	KX277517	
Jujubinus vexationis Curini-Galletti, 1990	Azores & Madeira 2008	Funchal, Madeira	MNCN:ADN:86720	KX277483	KX277580	KX277415	KX277341	KX277628	KX277518	
Jujubinus ruscurianus (Weinkauff, 1868) 1	Mediterranean 2005	Banalmádena, Málaga, S Spain	MNCN:ADN:86721	KX277484	KX277581	KX277416	KX277342	KX277629	KX277519	
Jujubinus ruscurianus 2	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86722	–	KX277582	KX277417	KX277343	–	–	
Jujubinus gravinae (Dautzenberg, 1881)	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86723	KX277485	KX277583	KX277418	KX277344	KX277630	KX277520	
Gibbula tingitana Pallary, 1901	Mediterranean 2005	Calahonda, Málaga, S Spain	MNCN:ADN:86724	–	KX277584	KX277419	KX277345	KX277631	KX277521	
Jujubinus striatus (Linnaeus, 1758) 1	Balearic Islands 2008	Menorca, E Spain	MNCN:ADN:86725	–	KX277585	KX277420	KX277346	–	–	
Jujubinus striatus 2	Galicia 2005	Cabo Estai, Canido, Pontevedra, NW Spain	MNCN:ADN:86726	KX277486	KX277586	KX277421	KX277347	KX277632	KX277522	
Jujubinus striatus 3	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86727	–	KX277587	KX277422	KX277348	–	–	
Callumbonella suturalis (Philippi, 1836)	Fauna I, stn 26A	La Herradura Seamount, Alboran Sea, SE Spain	MNCN:ADN:86728	KX277487	–	KX277423	KX277349	–	–	
Clelandella miliaris (Brocchi, 1814)	Fauna I, stn 5A	off Vélez-Málaga, Alborán Sea, SE Spain	MNCN:ADN:86729	KX277488	KX277588	KX277424	KX277350	–	–	
Gibbula magus (Linnaeus, 1758) 1	Galicia 2005	Cabo Estai, Canido, Pontevedra, Spain	MNCN:ADN:86730	KX277489	KX277589	KX277425	KX277351	KX277633	–	
Gibbula magus 2	Mediterranean 2005	Cabo Pino, Málaga, Spain	MNCN:ADN:86731	KX277490	KX277590	KX277426	KX277352	–	–	
Gibbula magus 3	—	Faro, Portugal	NHMUK 20030329	GQ232392	GQ232364	GQ232294	GQ232329	–	–	–
Gibbula fanulum (Gmelin, 1791)	—	Korcula I., Croatia	NHMUK 20080378	GQ232391	GQ232363	–	GQ232328	KX277611	KX277505	MorphoBank
Gibbula ardens (Salis Marschlins, 1793) 1	Balearic Islands 2008	Menorca, Spain	MNCN:ADN:86732	KX277491	KX277591	KX277427	KX277353	–	–	
Gibbula ardens 2	Balearic Islands 2007	Cala San Esteve, Menorca, E Spain	MNCN:ADN:86733	KX277492	KX277592	KX277428	KX277354	–	–	
Gibbula ardens 3	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86734	KX277493	KX277593	KX277429	KX277355	KX277634	KX277523	
Gibbula philberti (Récluz, 1843) 1	Mediterranean 2005	El Mohon, Murcia, SE Spain	MNCN:ADN:86735	KX277494	KX277594	KX277430	KX277356	KX277635	–	
Gibbula philberti 2	Mediterranean 2005	Banalmádena, Málaga, S Spain	MNCN:ADN:86736	KX277495	KX277595	KX277431	KX277357	–	–	
Gibbula turbinoides (Deshayes, 1835)	Mediterranean 2005	Cabo de Palos, Murcia, SE Spain	MNCN:ADN:86737	KX277496	KX277596	KX277432	KX277358	KX277636	KX277524	
Cantharidus opalus (Martyn, 1784)	—	Stewart I., New Zealand	NMNZ M.287117	GQ249739	GQ249696	GQ249830	GQ249799	–	–	MorphoBank
Cantharidus capillaceus (Philippi, 1849)	—	Enderby Island, New Zealand	—	GQ249742	GQ249700	GQ249834	GQ249804	–	–	–
Cantharidus dilatatus (G. B. Sowerby II, 1870)	—	Wellington, New Zealand	—	GQ249727	GQ249689	GQ249809	GQ249771	–	–	–
Cantharidus antipoda Hombront & Jackinot, 1854	—	Stewart Island, New Zealand	NMNZ M.287119	GQ249726	GQ249682	GQ249807	GQ249767	–	–	MorphoBank
Micrelenchus tesselatus (A. Adams, 1853)	—	Warrington, New Zealand	NMNZ M.288193	GQ434028	GQ434017	GQ434022	GQ434011	–	–	MorphoBank
Micrelenchus huttoni (E.A. Smith, 1876)	—	Purakaunui Bay, New Zealand	NMNZ M.288196	GQ434027	GQ434016	GQ434021	GQ434010	–	–	MorphoBank
Micrelenchus sanguineus (Gray, 1843)	—	Warrington, New Zealand	NMNZ M.287121	GQ249728	GQ249685	GQ249812	GQ249773	–	–	MorphoBank
Micrelenchus purpureus (Gmelin, 1791)	—	Bay of Islands, New Zealand	NMNZ M.287118	GQ249744	GQ249702	GQ249835	GQ249805	–	–	–
Roseaplagis rufozonus (A. Adams, 1853)	—	Tauranga, New Zealand	NMNZ M.284054	GQ249740	GQ249697	GQ249831	GQ249800	–	–	MorphoBank
Prothalotia lehmanni (Menke, 1843)	—	Dunsborough, Australia	NHMUK 20070156	EU530021	EU530123	GQ232302	GQ232338	–	–	MorphoBank
Cantharidinae unknown genus 1 sp. 1	Salomon 2, st. CP2203	Sta Isabel I., Solomon Islands	MNHN IM-2007-18160	GQ232384	GQ232358	GQ232286	FN435321	–	–	–
"Cantharidus" sendersi Poppe, Tagaro & Dekker, 2006	Santo 2006, st. EP39	Tutuba I., Vanuatu	MNHN IM-2007-18458	—	KX277545	KX277379	KX277305	–	–	
"Jujubinus" suarezensis (P. Fischer, 1878) 1	Atimo Vatae, st. TM19	Sainte Luce, Madagascar	MNHN IM-2009-13642	KX277436	KX277528	KX277362	KX277290	–	–	
"Jujubinus" suarezensis 2	Atimo Vatae, st. BS08	Rocher de l'Albatros, Madagascar	MNHN IM-2009-13612	KX277435	KX277527	KX277361	KX277289	–	–	
"Jujubinus" suarezensis 3	Inhaca 2011, st. MR06	Ilha dos Portugueses, Maputo Bay, Mozambique	MNHN IM-2009-23230	KX277457	KX277551	KX277385	KX277311	–	–	–
"Jujubinus" suarezensis 4	Inhaca 2011, st. MR06	Ilha dos Portugueses, Maputo Bay, Mozambique	MNHN IM-2009-23229	KX277456	KX277550	KX277384	KX277310	–	–	–
"Jujubinus" suarezensis 5	Inhaca 2011, st. MM11	Xixuane, Maputo Bay, Mozambique	MNHN IM-2009-23217	KX277453	KX277547	KX277381	KX277307	KX277609	KX277503	–
"Jujubinus" suarezensis 6	Inhaca 2011, st. MM04	Ponta Punduine, Maputo Bay, Mozambique	MNHN IM-2009-23257	KX277458	KX277552	KX277386	KX277312		–	–
Kanekotrochus infuscatus (Gould, 1861)	—	Katakuchihama, Kagoshima, Japan	NSMT Mo76821	AB505235	AB505282	AB505328	AB505371	–	–	MorphoBank
Agagus agagus Jousseaume, 1894	—	Aliwal Shoal, KwaZulu-Natal, South Africa	NMSA W2743	KX277498	KX277598	KX277434	KX277360	KX277638	KX277526	–
Priotrochus kotschyi (Philippi, 1849)	—	Abu Dhabi, United Arab Emirates	OKOT.ABD.1	JN686228	JN686315	JN686118	–	–	–	–

Data S1 (cont.)

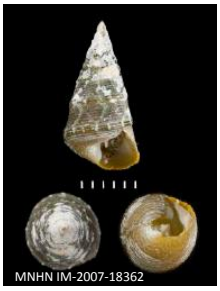
	Thalotia conica (Gray, 1827)	Western Australia 2011, st. WA10	Pagoda Point, Princess Royal Harbour, Australia	MNHN IM-2009-23270	EU530022	FN435322	GQ232309	GQ232345	–	–	–
	"Jujubinus" gilberti (Montrouzier in Fischer, 1878) 1	Panglao 2004, st. B8	Napaling, Panglao I., Philippines	MNHN IM-2007-18363	KX277443	KX277535	KX277369	KX277297	–	–	
	"Jujubinus" gilberti 2	Panglao 2004, st. S2	Baclayon Takot, Bohol I., Philippines	MNHN IM-2007-18375	—	KX277542	KX277376	KX277302	–	–	
	Cantharidinae unknown genus 2 sp. 1	Panglao 2004, st. B8	Napaling, Panglao I., Philippines	MNHN IM-2007-18362	KX277442	KX277534	KX277368	KX277296	–	–	Supp. Data 1
	"Cantharidus" callichroa (Philippi, 1849)	—	Omaezaki, Shizuoka, Japan	NHMUK 20050419	AM048703	AM049338	AM048892	GQ232314	–	–	MorphoBank
	"Cantharidus" jessoensis (Scherenck, 1863)	—	Otuchi, Iwate, Japan	NSMT Mo76819	AB505405	AB505280	AB505326	AB505369	–	–	MorphoBank
	"Jujubinus" geographicus Poppe, Tagaro & Dekker, 2006	Santo 2006, st. EP36	E Aoré I., Aimbuei Bay, Vanuatu	MNHN IM-2007-18457	KX277451	KX277544	KX277378	KX277304	–	–	
	Tosatrochus attenuatus (Jonas, 1844)	—	Marukihama, Bounotsu, Kagoshima, Japan	—	AB505236	AB505283	AB505329	AB505372	–	–	MorphoBank
	Pictodiloma suavis (Phillippi, 1850)	—	Susaki, Ogasawara I., Japan	NSMT Mo76820	AB505234	AB505281	AB505327	AB505370	–	–	MorphoBank
	"Cantharidus" lepidus (Philippi, 1846)	—	Duke of Orleans Bay, Esperance, Australia	NHMUK 20080950	GQ232402	GQ232371	GQ232305	GQ232341	–	–	MorphoBank
	Priotrochus obscurus (W. Wood, 1828) 1	Inhaca 2011, st. MM07	Ponta Punduine, Maputo Bay, Mozambique	MNHN IM-2009-23222	KX277454	KX277548	KX277382	KX277308	–	–	–
	***Priotrochus obscurus* 2**	Inhaca 2011, st. MM02	Ponta Punduine, Maputo Bay, Mozambique	MNHN IM-2009-23224	KX277455	KX277549	KX277383	KX277309	–	–	–
	Oxystele impervia (Menke, 1843)	—	False Bay, South Africa	NMNZ M.287122	GQ249748	DQ061093	DQ061084	GQ249769	–	–	MorphoBank
	Oxystele variegata (Anton, 1838)	—	False Bay, South Africa	NMNZ M.287126	GQ249751	DQ061092	DQ061083	GQ249770	–	–	MorphoBank
	Oxystele tabularis (Krauss, 1848)	—	East London, South Africa	NMNZ M.287124	GQ249750	DQ061090	DQ061081	GQ249766	–	–	MorphoBank
	Oxystele sinensis (Gmelin, 1791)	—	East London, South Africa	NMNZ M.287123	GQ249749	DQ061089	DQ061080	GQ249765	–	–	MorphoBank
Stomatellinae											
	Stomatella planulata (Lamarck, 1816)	—	Teniya, Nago City, Okinawa, Japan	NHMUK 20080949	EU530029	EU530130	GQ232307	GQ232343	–	–	MorphoBank
	Stomatella sp. 1	Panglao 2004, st. M3	Danao, Panglao I., Philippines	MNHN IM-2007-18370	KX277447	KX277539	KX277373	KX277300	KX277604	–	Supp. Data 1
	Stomatella sp. 2	Panglao 2004, st. B9	Napaling, Panglao I., Philippines	MNHN IM-2007-18365	KX277445	KX277537	KX277371	–	KX277603	KX277501	Supp. Data 1
	Stomatella sp. 3	Panglao 2004, st. B13	Baclayon Takot, Bohol I., Philippines	MNHN IM-2007-18355	KX277438	KX277530	KX277364	KX277292	KX277600	–	Supp. Data 1
	Stomatella impertusa (Burrow, 1815) 1	Jordan 2005	Aqaba, Jordan	MNCN:ADN:86738	KX277497	KX277597	KX277433	KX277359	KX277637	KX277525	–
	Stomatella impertusa 2	—	Esperance, Australia	NHMUK 20080948	GQ232400	GQ232369	GQ232303	GQ232339	–	–	MorphoBank
	Stomatella impertusa 3	Santo 2006, st. DB8	Bruat Channel, Vanuatu	MNHN IM-2007-18154	GQ232401	GQ232370	GQ232304	GQ232340	–	–	MorphoBank
	Stomatolina rubra (Lamarck, 1822) 1	Panglao 2004, st. B9	Napaling, Panglao I., Philippines	MNHN IM-2007-18366	KX277446	KX277538	KX277372	KX277299			Supp. Data 1
	Stomatolina rubra 2	Panglao 2004, st. B39	Pontod Lagoon, Panglao I., Philippines	MNHN IM-2007-18156	GQ232404	GQ232373	GQ232308	GQ232344			MorphoBank
	Stomatolina rubra 3	Panglao 2004, st. B39	Pontod Lagoon, Panglao I., Philippines	MNHN IM-2007-18359	KX277440	KX277532	KX277366	KX277294			Supp. Data 1
	Stomatolina rubra 4	Panglao 2004, st. B39	Pontod Lagoon, Panglao I., Philippines	MNHN IM-2007-18360	KX277441	KX277533	KX277367	KX277295			Supp. Data 1
	Stomatolina angulata (A. Adams, 1850)	—	Marukihama, Bounotsu, Kagoshima, Japan	—	AB505248	AB505295	AB505339	AB505382	–	–	MorphoBank
	Pseudostomatella decolorata (Gould, 1848)	—	Benoki, Okinawa, Japan	NSMT Mo76824	AB505244	AB505291	AB505335	AB505379	–	–	MorphoBank
	Stomatellinae unknown genus 1 sp. 1	Panglao 2004, st. R31	Pamilacan I., Philippines	MNHN IM-2007-18371	KX277448	KX277540	KX277374	KX277301	KX277605	–	
	Stomatellinae unknown genus 1 sp. 2	Panglao 2004, st. R43	Cortes Takot, Bohol I., Philippines	MNHN IM-2007-18372	KX277449	KX277541	KX277375	–	KX277606	–	Supp. Data 1
	Stomatellinae unknown genus 2 sp. 1	Panglao 2004, st. B13	Baclayon Takot, Bohol I., Philippines	MNHN IM-2007-18356	KX277439	KX277531	KX277365	KX277293	KX277601	KX277499	Supp. Data 1
	Microtis sp. 1	Santo 2006, st. DB40	Palikulo Bay, Vanuatu	MNHN IM-2007-18449	KX277450	KX277543	KX277377	KX277303	KX277607	–	Supp. Data 1
	Microtis tuberculata H. & A. Adams, 1850	Panglao 2004, st. B9	Napaling, Panglao I., Philippines	MNHN IM-2007-18364	KX277444	KX277536	KX277370	KX277298	KX277602	KX277500	
	Microtis? sp. 2	Santo 2006, st. LD8	Santo, Malparavu I., Vanuatu	MNHN IM-2007-18463	KX277452	KX277546	KX277380	KX277306	KX277608	KX277502	Supp. Data 1
	Stomatia obscura Sowerby G.B. III, 1874	—	Marukihama, Bounotsu, Kagoshima, Japan	—	AB505247	AB505294	AB505338	AB505381	–	–	MorphoBank
	Stomatia phymotis Helbling, 1779	Santo 2006, st. VM59	SE Malparavu Island, Vanuatu	MNHN IM-2007-18155	GQ232403	GQ232372	GQ232306	GQ232342	–	–	MorphoBank
	Calliotrochus marmoreus (Pease, 1861)	Panglao 2004, st. B11	Pamilacan I., Philippines	MNHN IM-2007-18352	KX277437	KX277529	KX277363	KX277291	KX277599	–	
Umboniinae											
	Isanda coronata A. Adams, 1854	—	Karratha Back Beach, Australia	NHMUK 20050421	AM048705	AM049340	AM048894	GQ232332	–	–	MorphoBank
	Umbonium costatum (Kiener, 1839)	—	Pusan, South Korea	NHMUK 20050422	AM048706	AM049341	AM048895	GQ232348	–	–	MorphoBank

*indv2: the corresponding sequence comes from specimen 2

Data S2. Photos of unknown genera and species

No photo

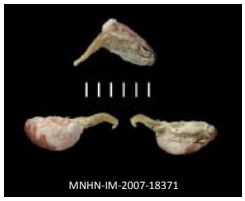
Cantharidinae unknown genus 1 sp. 1



Cantharidinae unknown genus 2 sp. 1

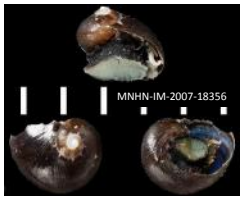
Microtis sp. 1

Microtis? sp. 2

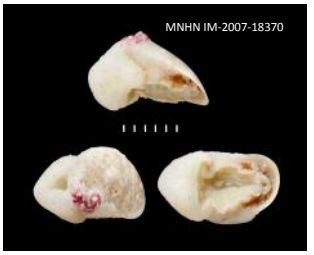


Stomatellinae unknown genus 1 sp. 1

Stomatellinae unknown genus 1 sp. 2

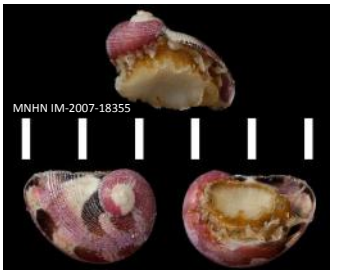


Stomatellinae unknown genus 2 sp. 1



Stomatella sp. 1

Stomatella sp. 2

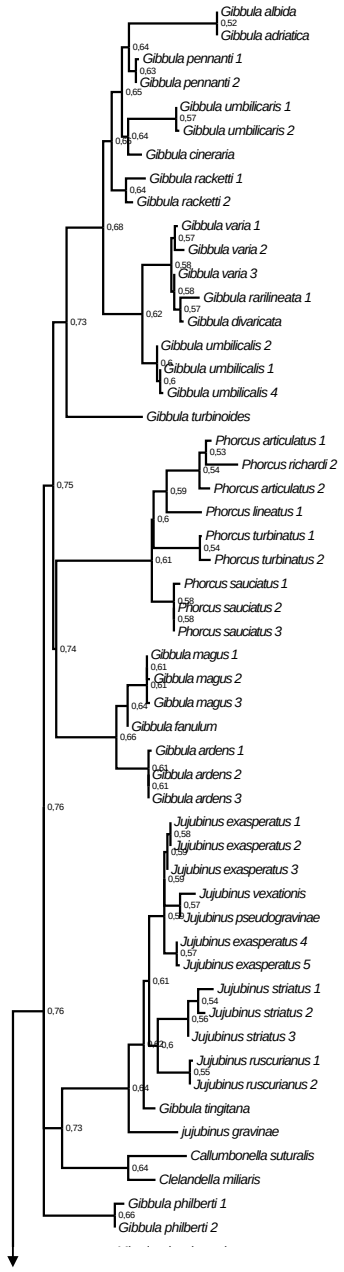


Stomatella sp. 3

Data S3. Selected best -fit partitions and models.

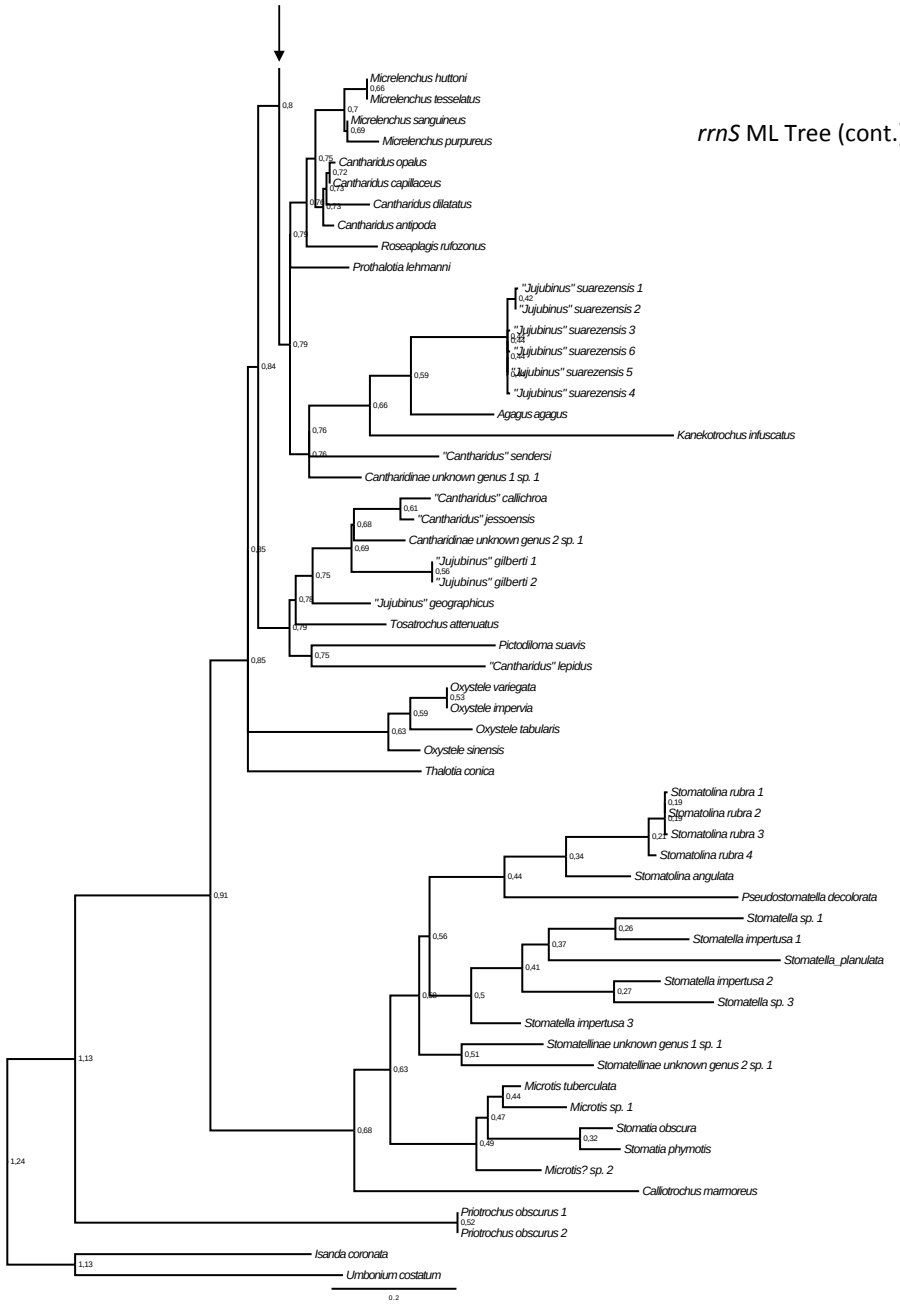
	Partition selected (PartitionFinder)	Model selected (PartitionFinder)	RaxML (-lnL) (PartitionFinder)	Model (MtZoa)	RaxML (-lnL) (MtZoa)	AIC (PartitionFinder)	AIC (MtZoa)
Gastropod	*atp6 + atp8*	MtArt+I+G	-5226.587607	MtZoa+I+G	-5199.492818	10453.175312	10398.985734
	cob	MtArt+I+G+F	-7017.424662	MtZoa+I+G+F	-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	—

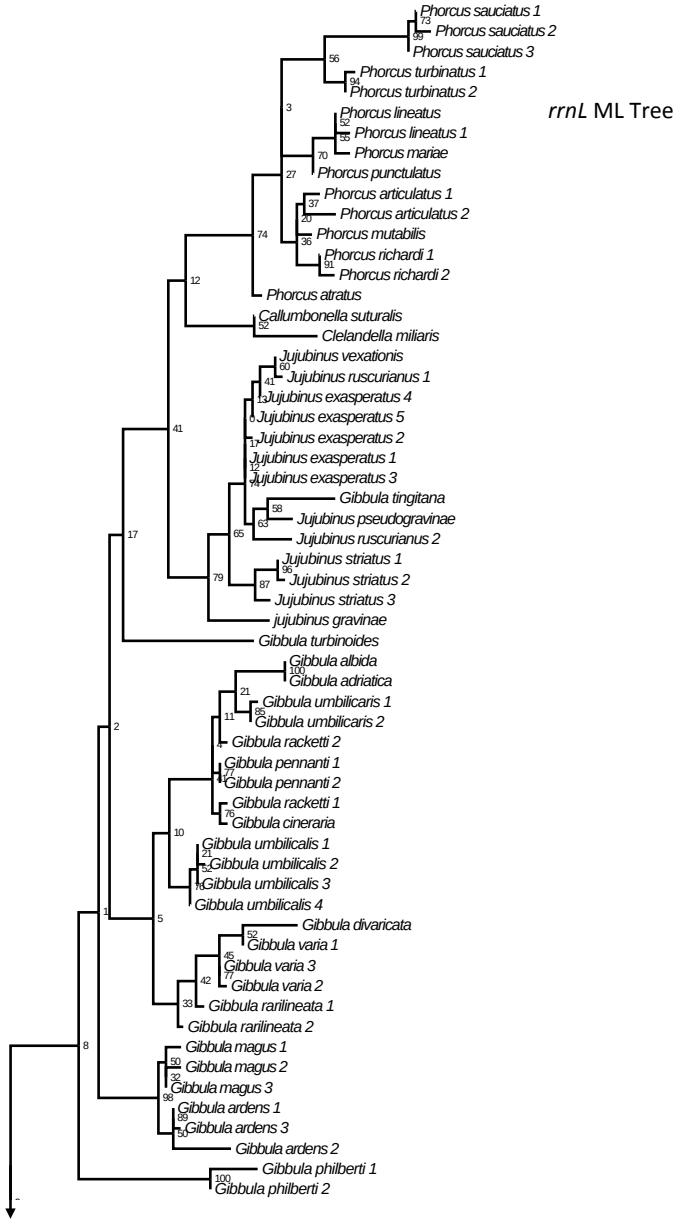
Data S4. Individual gene trees



rrnS ML Tree

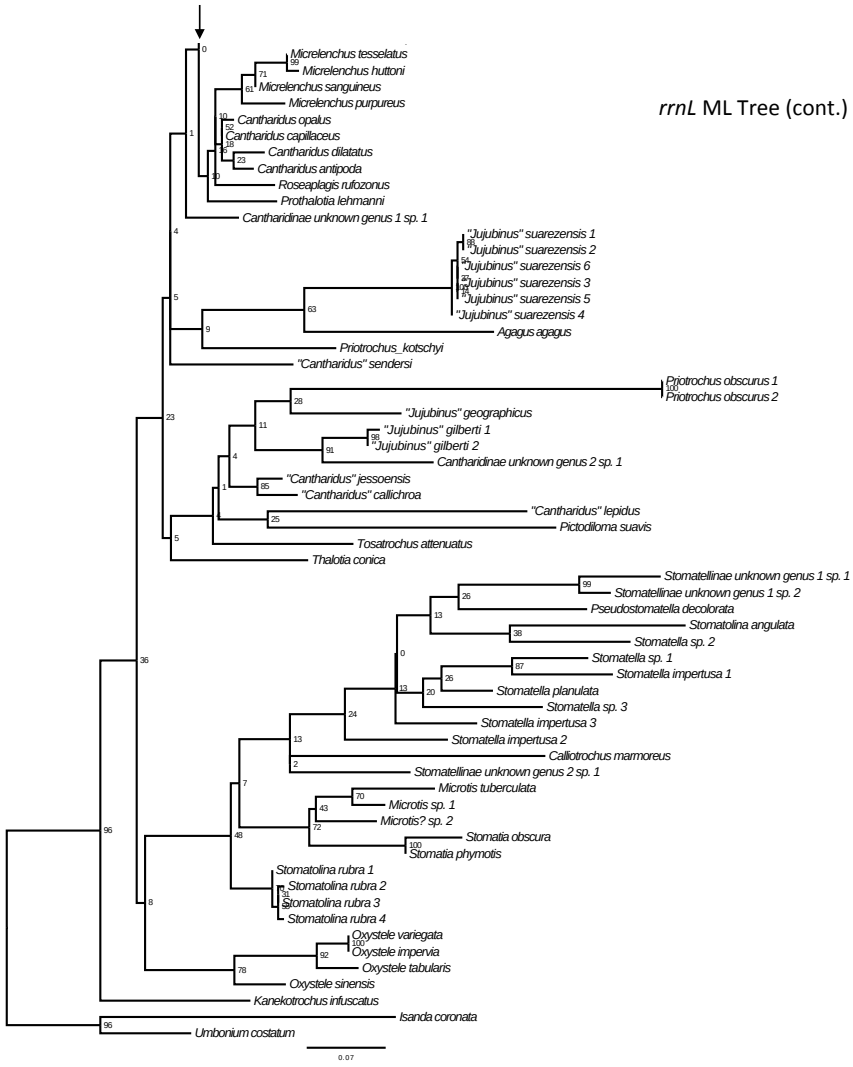
rrnS ML Tree (cont.)

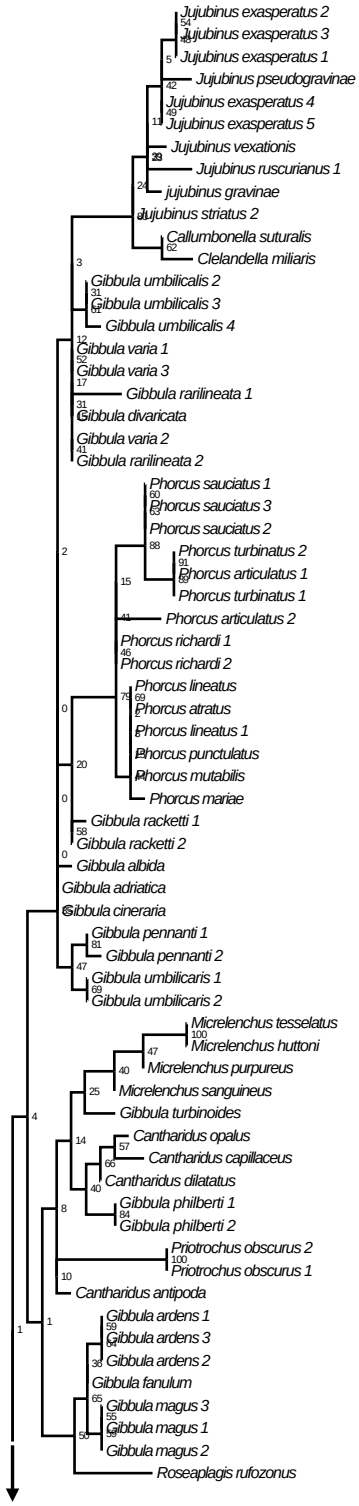



Phorcus sauciatus 1
Phorcus sauciatus 2
Phorcus sauciatus 3
Phorcus turbinatus 1
Phorcus turbinatus 2
Phorcus lineatus
Phorcus lineatus 1
Phorcus mariae
Phorcus punctulatus
Phorcus articulatus 1
Phorcus articulatus 2
Phorcus mutabilis
Phorcus richardi 1
Phorcus richardi 2
Phorcus atratus
Callumbonella suturalis
Clelandella miliaris
Jujubinus vexationis
Jujubinus ruscurianus 1
Jujubinus exasperatus 4
Jujubinus exasperatus 5
Jujubinus exasperatus 2
Jujubinus exasperatus 1
Jujubinus exasperatus 3
Gibbula tingitana
Jujubinus pseudogravinae
Jujubinus ruscurianus 2
Jujubinus striatus 1
Jujubinus striatus 2
Jujubinus striatus 3
jujubinus gravinae
Gibbula turbinoides
Gibbula albida
Gibbula adriatica
Gibbula umbilicaris 1
Gibbula umbilicaris 2
Gibbula racketti 2
Gibbula pennanti 1
Gibbula pennanti 2
Gibbula racketti 1
Gibbula cineraria
Gibbula umbilicalis 1
Gibbula umbilicalis 2
Gibbula umbilicalis 3
Gibbula umbilicalis 4
Gibbula divaricata
Gibbula varia 1
Gibbula varia 3
Gibbula varia 2
Gibbula rarilineata 1
Gibbula rarilineata 2
Gibbula magus 1
Gibbula magus 2
Gibbula magus 3
Gibbula ardens 1
Gibbula ardens 3
Gibbula ardens 2
Gibbula philberti 1
Gibbula philberti 2

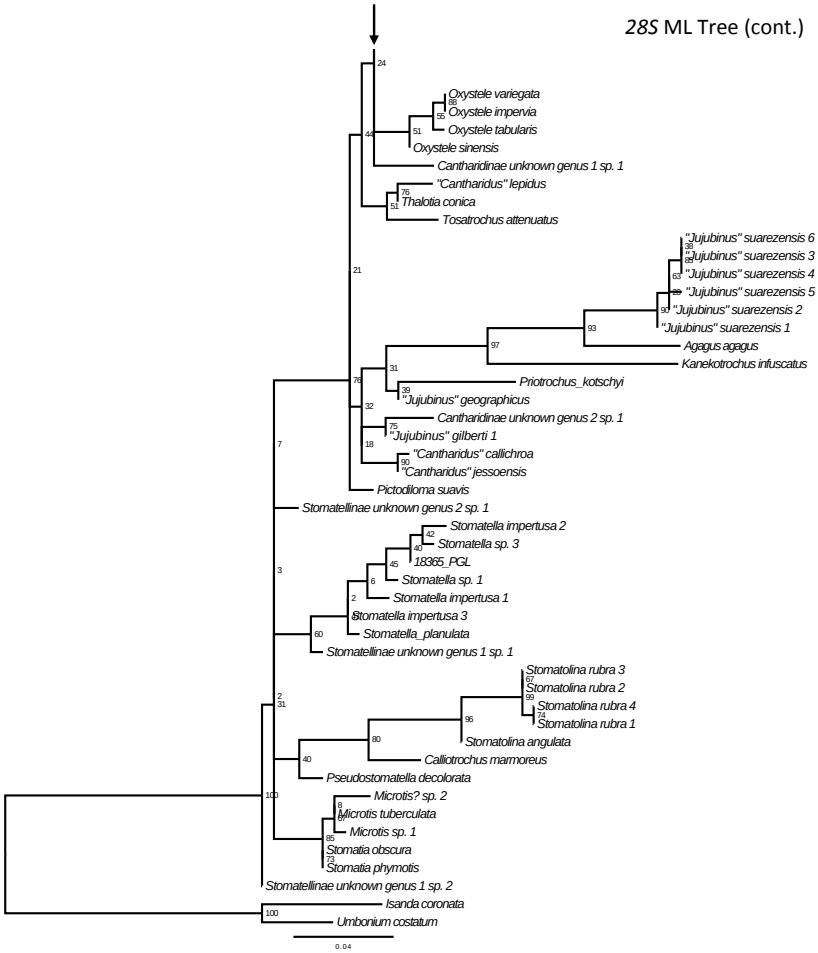
rrnL ML Tree

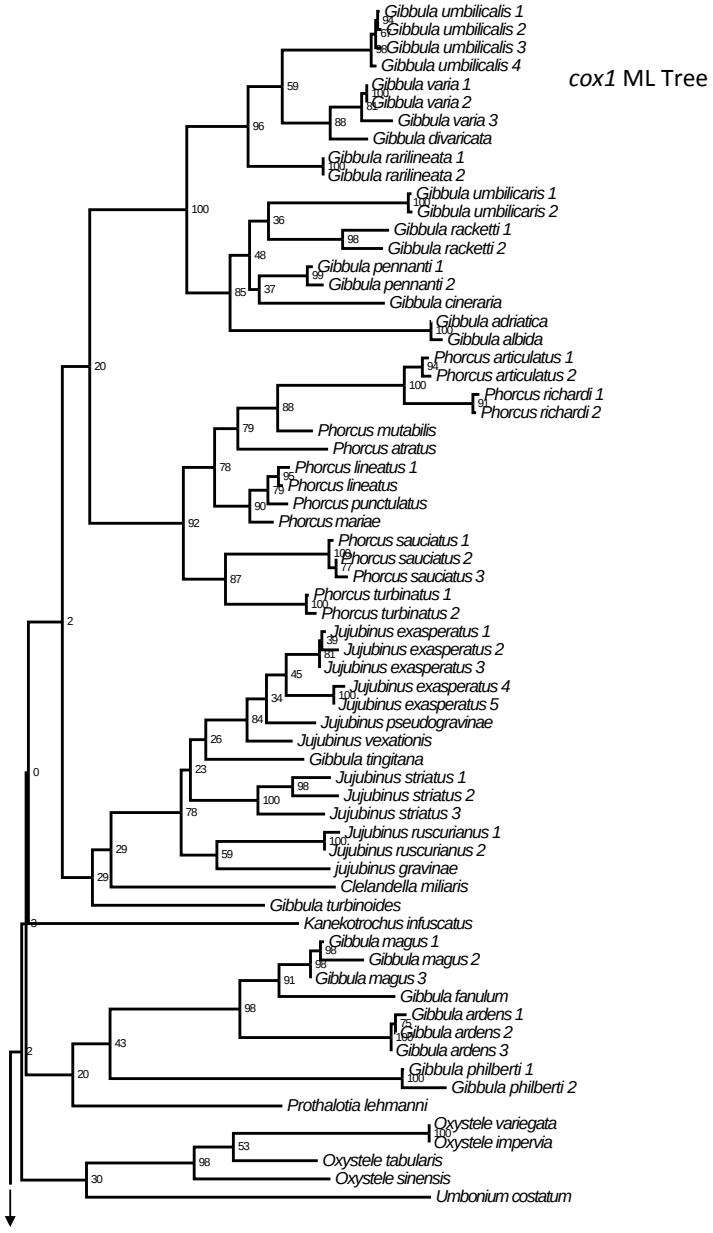
rrnL ML Tree (cont.)



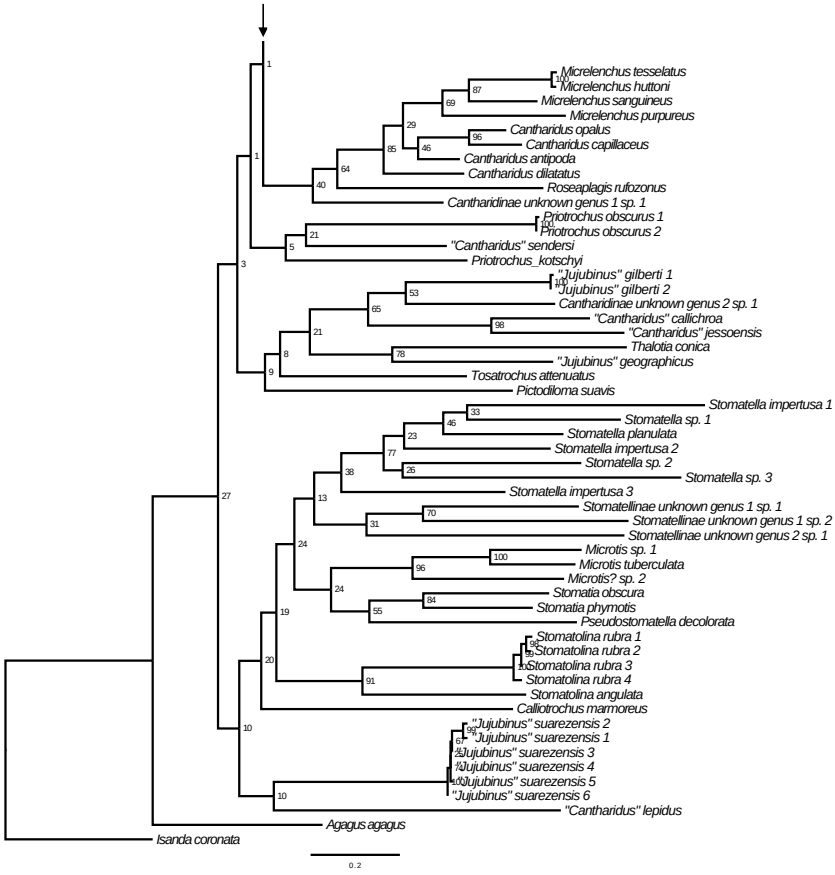

Jujubinus exasperatus 2
Jujubinus exasperatus 3
Jujubinus exasperatus 1
Jujubinus pseudogravinae
Jujubinus exasperatus 4
Jujubinus exasperatus 5
Jujubinus vexationis
Jujubinus ruscurianus 1
jujubinus gravinae
Jujubinus striatus 2
Callumbonella suturalis
Clelandella miliaris
Gibbula umbilicalis 2
Gibbula umbilicalis 3
Gibbula umbilicalis 4
Gibbula varia 1
Gibbula varia 3
Gibbula rarilineata 1
Gibbula divaricata
Gibbula varia 2
Gibbula rarilineata 2
Phorcus sauciatus 1
Phorcus sauciatus 3
Phorcus sauciatus 2
Phorcus turbinatus 2
Phorcus articulatus 1
Phorcus turbinatus 1
Phorcus articulatus 2
Phorcus richardi 1
Phorcus richardi 2
Phorcus lineatus
Phorcus atratus
Phorcus lineatus 1
Phorcus punctulatus
Phorcus mutabilis
Phorcus mariae
Gibbula racketti 1
Gibbula racketti 2
Gibbula albida
Gibbula adriatica
Gibbula cineraria
Gibbula pennanti 1
Gibbula pennanti 2
Gibbula umbilicaris 1
Gibbula umbilicaris 2
Micrelenchus tesselatus
Micrelenchus huttoni
Micrelenchus purpureus
Micrelenchus sanguineus
Gibbula turbinoides
Cantharidus opalus
Cantharidus capillaceus
Cantharidus dilatatus
Gibbula philberti 1
Gibbula philberti 2
Priotrochus obscurus 2
Priotrochus obscurus 1
Cantharidus antipoda
Gibbula ardens 1
Gibbula ardens 3
Gibbula ardens 2
Gibbula fanulum
Gibbula magus 3
Gibbula magus 1
Gibbula magus 2
Roseaplagis rufozonus

28S ML Tree

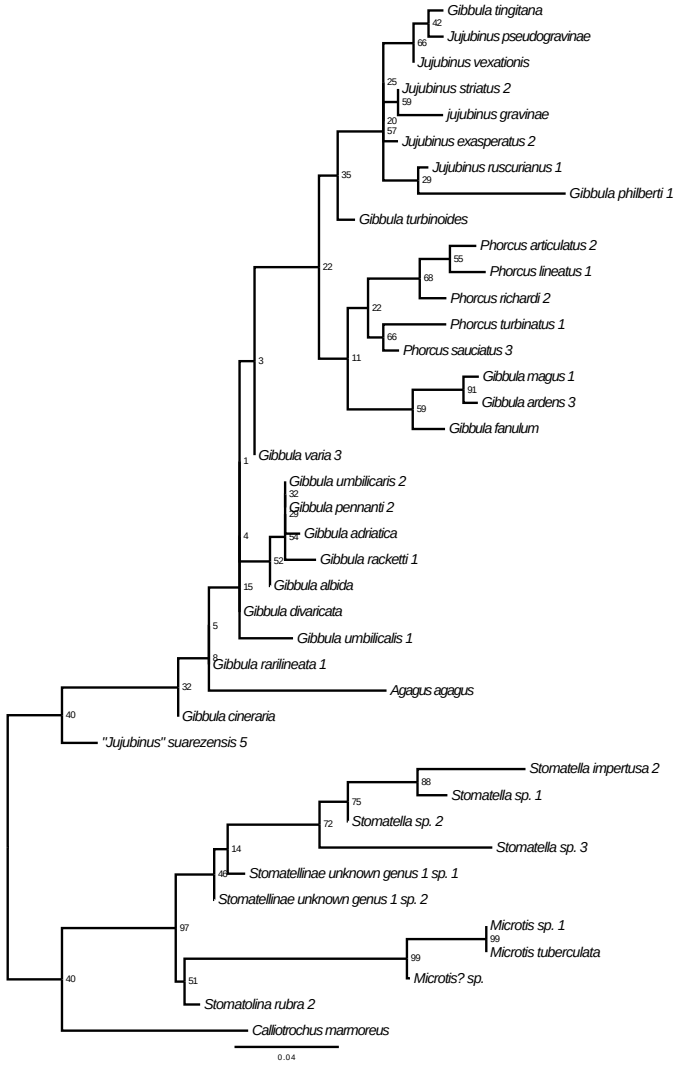

28S ML Tree (cont.)
Oxystele variegata
Oxystele impervia
Oxystele tabularis
Oxystele sinensis
Cantharidinae unknown genus 1 sp. 1
"Cantharidus" lepidus
Thalotia conica
Tosatrochus attenuatus
"Jujubinus" suarezensis 6
"Jujubinus" suarezensis 3
"Jujubinus" suarezensis 4
"Jujubinus" suarezensis 5
"Jujubinus" suarezensis 2
"Jujubinus" suarezensis 1
Agagus agagus
Kanekotrochus infuscatus
Priotrochus_kotschyi
"Jujubinus" geographicus
Cantharidinae unknown genus 2 sp. 1
"Jujubinus" gilberti 1
"Cantharidus" callichroa
"Cantharidus" jessoensis
Pictodiloma suavis
Stomatellinae unknown genus 2 sp. 1
Stomatella impertusa 2
Stomatella sp. 3
18365_PGL
Stomatella sp. 1
Stomatella impertusa 1
Stomatella impertusa 3
Stomatella_planulata
Stomatellinae unknown genus 1 sp. 1
Stomatolina rubra 3
Stomatolina rubra 2
Stomatolina rubra 4
Stomatolina rubra 1
Stomatolina angulata
Calliotrochus marmoreus
Pseudostomatella decolorata
Microtis? sp. 2
Microtis tuberculata
Microtis sp. 1
Stomatia obscura
Stomatia phymotis
Stomatellinae unknown genus 1 sp. 2
Isanda coronata
Umbonium costatum
0.04


Gibbula umbilicalis 1
Gibbula umbilicalis 2
Gibbula umbilicalis 3
Gibbula umbilicalis 4
Gibbula varia 1
Gibbula varia 2
Gibbula varia 3
Gibbula divaricata
Gibbula rarilineata 1
Gibbula rarilineata 2
Gibbula umbilicaris 1
Gibbula umbilicaris 2
Gibbula racketti 1
Gibbula racketti 2
Gibbula pennanti 1
Gibbula pennanti 2
Gibbula cineraria
Gibbula adriatica
Gibbula albida
Phorcus articulatus 1
Phorcus articulatus 2
Phorcus richardi 1
Phorcus richardi 2
Phorcus mutabilis
Phorcus atratus
Phorcus lineatus 1
Phorcus lineatus
Phorcus punctulatus
Phorcus mariae
Phorcus sauciatus 1
Phorcus sauciatus 2
Phorcus sauciatus 3
Phorcus turbinatus 1
Phorcus turbinatus 2
Jujubinus exasperatus 1
Jujubinus exasperatus 2
Jujubinus exasperatus 3
Jujubinus exasperatus 4
Jujubinus exasperatus 5
Jujubinus pseudogravinae
Jujubinus vexationis
Gibbula tingitana
Jujubinus striatus 1
Jujubinus striatus 2
Jujubinus striatus 3
Jujubinus ruscurianus 1
Jujubinus ruscurianus 2
jujubinus gravinae
Clelandella miliaris
Gibbula turbinoides
Kanekotrochus infuscatus
Gibbula magus 1
Gibbula magus 2
Gibbula magus 3
Gibbula fanulum
Gibbula ardens 1
Gibbula ardens 2
Gibbula ardens 3
Gibbula philberti 1
Gibbula philberti 2
Prothalotia lehmanni
Oxystele variegata
Oxystele impervia
Oxystele tabularis
Oxystele sinensis
Umbonium costatum

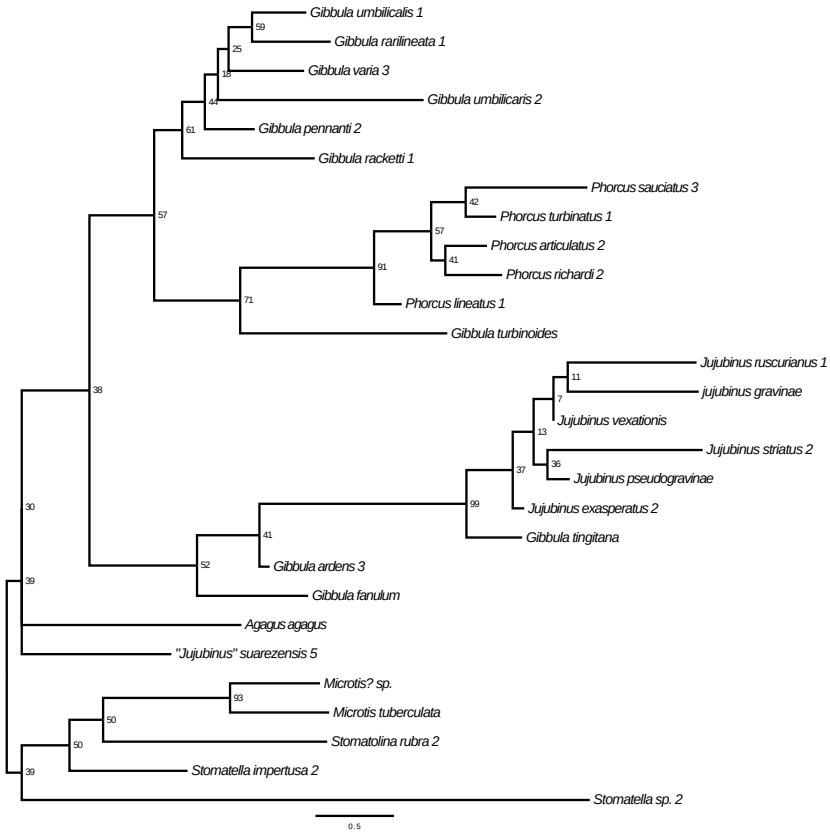
cox1 ML Tree


Micrelenchus tesselatus
Micrelenchus huttoni
Micrelenchus sanguineus
Micrelenchus purpureus
Cantharidus opalus
Cantharidus capillaceus
Cantharidus antipoda
Cantharidus dilatatus
Roseaplagis rufozonus
Cantharidinae unknown genus 1 sp. 1
Priotrochus obscurus 1
Priotrochus obscurus 2
"Cantharidus" sendersi
Priotrochus_kotschyi
"Jujubinus" gilberti 1
"Jujubinus" gilberti 2
Cantharidinae unknown genus 2 sp. 1
"Cantharidus" callichroa
"Cantharidus" jessoensis
Thalotia conica
"Jujubinus" geographicus
Tosatrochus attenuatus
Pictodiloma suavis
Stomatella impertusa 1
Stomatella sp. 1
Stomatella planulata
Stomatella impertusa 2
Stomatella sp. 2
Stomatella sp. 3
Stomatella impertusa 3
Stomatellinae unknown genus 1 sp. 1
Stomatellinae unknown genus 1 sp. 2
Stomatellinae unknown genus 2 sp. 1
Microtis sp. 1
Microtis tuberculata
Microtis? sp. 2
Stomatia obscura
Stomatia phymotis
Pseudostomatella decolorata
Stomatolina rubra 1
Stomatolina rubra 2
Stomatolina rubra 3
Stomatolina rubra 4
Stomatolina angulata
Calliotrochus marmoreus
"Jujubinus" suarezensis 2
"Jujubinus" suarezensis 1
"Jujubinus" suarezensis 3
"Jujubinus" suarezensis 4
"Jujubinus" suarezensis 5
"Jujubinus" suarezensis 6
"Cantharidus" lepidus
Agagus agagus
Isanda coronata
0.2

H3 ML Tree


Gibbula tingitana
Jujubinus pseudogravinae
Jujubinus vexationis
Jujubinus striatus 2
jujubinus gravinae
Jujubinus exasperatus 2
Jujubinus ruscurianus 1
Gibbula philberti 1
Gibbula turbinoides
Phorcus articulatus 2
Phorcus lineatus 1
Phorcus richardi 2
Phorcus turbinatus 1
Phorcus sauciatus 3
Gibbula magus 1
Gibbula ardens 3
Gibbula fanulum
Gibbula varia 3
Gibbula umbilicaris 2
Gibbula pennanti 2
Gibbula adriatica
Gibbula racketti 1
Gibbula albida
Gibbula divaricata
Gibbula umbilicalis 1
Gibbula rarilineata 1
Agagus agagus
Gibbula cineraria
"Jujubinus" suarezensis 5
Stomatella impertusa 2
Stomatella sp. 1
Stomatella sp. 2
Stomatella sp. 3
Stomatellinae unknown genus 1 sp. 1
Stomatellinae unknown genus 1 sp. 2
Microtis sp. 1
Microtis tuberculata
Microtis? sp.
Stomatolina rubra 2
Calliotrochus marmoreus
0.04

cob ML Tree


Gibbula umbilicalis 1
Gibbula rarilineata 1
Gibbula varia 3
Gibbula umbilicaris 2
Gibbula pennanti 2
Gibbula racketti 1
Phorcus sauciatus 3
Phorcus turbinatus 1
Phorcus articulatus 2
Phorcus richardi 2
Phorcus lineatus 1
Gibbula turbinoides
Jujubinus ruscurianus 1
jujubinus gravinae
Jujubinus vexationis
Jujubinus striatus 2
Jujubinus pseudogravinae
Jujubinus exasperatus 2
Gibbula tingitana
Gibbula ardens 3
Gibbula fanulum
Agagus agagus
"Jujubinus" suarezensis 5
Microtis? sp.
Microtis tuberculata
Stomatolina rubra 2
Stomatella impertusa 2
Stomatella sp. 2
59
25
18
44
61
57
42
57
41
91
71
11
7
13
37
36
99
41
52
38
30
39
93
50
50
39
0.5

SUPPORTING INFORMATION 4

Data S1. Long PCR and primer walking primers.

Georissa bangueyensis

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
GBcox1F	GCTCCTGATATGGCTTTTCCTCGGTTAA	cox1-rrnL (9527)
NS16SF	GAGAGATTGTGACCTCGATGTTGGACTAGG	
Geocox1R	AGAACCCCCCCCATGAGCCAAATTACT	rrnL-cox1 (5917)
NERmt16SR	AGTCCAACATCGAGGTCACAAWCTCTT	

Neritina usnea

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
NUcos1FH	GAGCTTGGACAGCCTGGTGCTTTATTAGGGG	cox1-rrnL (9773)
SPN16sF	CGGTYACCCTAGTCCAACATCGAGGTCACA	
Nas16sR	AGGCTGGTATGAACGGTTTGACGAGA	rrnL-cox1 (5835)
SV-NU-cox1R	CCAAGCTCAGCTCGAATTAACAAACTAAGAG	

Primer walking

Primer	Sequence 5'-3'
NULeuR	TAAACCTAATGCACTTCTTCTGCCACC
NU16SF	CAAAAAGACGAGAAGACCCTGTTGAGC

Titiscania sp.

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
TLmtcox1F	GGTTCGGTGGACTTAGCTATTTTTTCACT	cox1-rrnL (10507)
NERmt16SF	CTGACYGTGCAAAGGTAGCATAATCAT	

Primer walking

Primer	Sequence 5'-3'
COIRbruTit	CGTCAACAACATAGTAATGGCCCCAGC
TLcox3F	AGTTACAGTTACTTGGGCGCATCATGC
TLLys3F	TATAGGCAGAAGCATGAAGAGGAGGC
TLnad2R	ATCTCCAAAAAGAGCACCAGGTTGCC
EctremFTit	AGAAGATAGAAACCAACCTGGCTCGC
12Sb (Kocher et al., 1989)	GAGGGTGACGGGCGGTGTGT
TL16s1F_w	ACCTCGTAATATGAAGTGTCAGCCC
TL12sR_w	TTAGATGTTTCAGTTCACTGCGTTCGC

Theodoxus fluviatilis

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
GastCOX1F	CTGGWACAGGATGRACTGTWTAYC	cox1-rrnL (9844)
GAS116SF	TGCCTTTTAATTGAAGGCTGGWATGAATGGTTTRACG	

Primer walking

Primer	Sequence 5'-3'
Nercox3F	GAGGGAACWTTTCAAGGMTTTCATACTGG
Gas2cox1R	ATTAAGCCGAGGAAAAGCCATATCAGGAGC
TXPrimer2R	CGTATTCATCTCAATACATAACAAAAGCCC
Txnad2R	GCAGAAACTGAAAATATAGTTCCAAACCC
ExtremFNer	AAAGATAGAAACTGACCTGGCTTACGCC
12Sb (Kocher et al., 1989)	GAGGGTGACGGGCGGTGTGT
Tx12SR_w	CGTTGCTAATAGCAGGGTGTCTAATCC

Viana regina

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Vivcox1F	CTTTGATGTTGAAGGTGCCAGATATGGC	cox1-rrnL (5071)
Viv16SRUJ	ATAAAGCTCAACAGGGTCTCATCGCC	
Vivcox1R	GGTGGATAAACAGTCCAACCAGTACC	cox3-cox1 (2676)
Vivcox3F	CTCATCATGCTTTAGTTGTTGGTTGTCG	

Data S2. Evolutionary sustitution models selected by Partition Finder

mt data set

nucleotides	Partition (BIC = 99518.19)	Best Model	Alpha	pinvar	A <-> C	A <-> G	A <-> T	C <-> G	C <-> T	pi(A)	pi(C)	pi(G)	pi(T)
1	All 1st codon positions of mt protein-coding genes	GTR+I+G	0.94	0.33	2.16	6.25	3.15	0.75	22.00	0.25	0.17	0.26	0.31
2	All 2nd codon positions of mt protein-coding genes	GTR+G	0.73	—	0.68	2.60	0.67	5.61	2.01	0.17	0.21	0.16	0.43
3	*rrnS-rrnL* genes	GTR+I+G	0.60	0.16	1.18	6.34	2.96	0.33	11.57	0.33	0.15	0.21	0.29

amino acids	Partition (BIC=86576.14)	Best Model	Alpha	pinvar	A <-> C	A <-> G	A <-> T	C <-> G	C <-> T	pi(A)	pi(C)	pi(G)	pi(T)
1	All mt protein-coding genes in aminoacids level combined	MtArt+I+G+F	0.56	0.15									
2	*rrnS-rrnL* genes	GTR+I+G	0.47	0.17	0.45	2.87	0.93	0.07	5.51	0.33	0.15	0.21	0.29

nuclear data set	Partition (BIC=22139.97)	Best Model	Alpha	pinvar	A <-> C	A <-> G	A <-> T	C <-> G	C <-> T	pi(A)	pi(C)	pi(G)	pi(T)
1	*18S* gene	TrNef+G	1000	—	0.91	1.88	0.48	0.60	3.94	0.25	0.21	0.26	0.25
2	*28S* gene	GTR+I+G	0.31	0.00	0.57	1.80	1.41	1.62	6.62	0.22	0.25	0.32	0.19
4	1st codon positions of *H3* and *actin* genes	GTR+G	0.58	—	4.41	0.54	6.11	0.18	17.63	0.27	0.23	0.31	0.16
5	2nd codon positions of *H3* and *actin* genes	GTR+G	1.60	—	2.03	1.69	0.40	4.34	1.77	0.30	0.25	0.17	0.26
6	3rd codon positions of *H3* and *actin* genes	GTR+G	0.96	—	8.75	49.06	39.18	0.12	28.36	0.10	0.41	0.28	0.19

combined data set	Partition (mt + nuclear)	Best Model	Alpha	pinvar	A <-> C	A <-> G	A <-> T	C <-> G	C <-> T	pi(A)	pi(C)	pi(G)	pi(T)
1	All 1st codon positions of mt protein-coding genes	GTR+I+G	1.34	0.35	3.04	7.40	4.96	0.70	18.49	0.25	0.16	0.25	0.31
2	All 2nd codon positions of mt protein-coding genes	GTR+G	0.68	—	1.08	2.43	0.75	5.62	2.27	0.17	0.21	0.16	0.44
3	*rrnS-rrnL* genes	GTR+I+G	0.73	0.18	1.05	7.39	4.33	0.25	11.63	0.32	0.15	0.21	0.30
4	*18S* gene	TrNef+G	1000	—	0.66	1.42	0.36	0.46	2.89	0.25	0.21	0.26	0.25
5	*28S* gene	GTR+I+G	0.36	0.04	0.34	1.19	1.02	1.04	4.36	0.22	0.25	0.32	0.19
6	1st codon positions of *H3* and *actin* genes	GTR+G	0.63	—	3.53	0.33	2.32	0.00	9.39	0.27	0.24	0.32	0.15
7	2nd codon positions of *H3* and *actin* genes	GTR+G	1.69	—	3.08	0.19	0.45	8.16	1.38	0.30	0.25	0.17	0.26
8	3rd codon positions of *H3* and *actin* genes	GTR+G	1.33	–	4.06	19.21	24.70	0.23	15.19	0.09	0.42	0.29	0.18

BEAST data set	Partition (BIC= 70900.79)	Best Model	Alpha	pinvar
1	1st codon positions of *atp6-8* genes	HKY+G	0.51	—
2	2nd codon positions of *atp6-8* genes	GTR+G	0.41	—
3	1st codon positions of *cob* gene	GTR+I+G	0.46	0.27
4	2nd codon positions of *cob* gene	GTR+G	0.22	—
5	1st codon positions of *cox1-3* genes	GTR+I+G	0.4	0.23
6	2nd codon positions of *cox1-3* genes	GTR+I+G	0.6	0.41
7	1st codon positions of *nad1-6* genes	GTR+I+G	1.11	0.25
8	2nd codon positions of *nad1-6* genes	GTR+G	0.4	—
9	*rrnS-rrnL* genes	GTR+I+G	0.46	3.45

Data S3. Mitochondrial genome features

Georissa bangueyensis								
		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	1	1539	1539	TAG	TAA	4	forward
cox2	CDS	1544	2228	687	TAG	T--	0	forward
trnD	tRNA	2229	2291	63			0	forward
atp8	CDS	2292	2447	156	TAG	TAG	9	forward
atp6	CDS	2457	3155	699	TAG	TAA	29	forward
trnF	tRNA	3185	3251	67			0	reverse
nad5	CDS	3252	4954	1704	TAG	T--	0	reverse
trnH	tRNA	4955	5018	64			0	reverse
nad4	CDS	5019	6381	1363	TAG	T--	-7	reverse
nad4L	CDS	6375	6665	291	TAG	TAA	5	reverse
trnT	tRNA	6671	6733	63			-1	forward
trnS(tga)	tRNA	6733	6798	66			25	reverse
cob	CDS	6824	7969	1146	TAG	TAA	1	reverse
nad6	CDS	7971	8468	498	TAG	TAA	3	reverse
trnP	tRNA	8472	8534	63			0	reverse
nad1	CDS	8535	9454	924	TAG	T--	0	reverse
trnL(taa)	tRNA	9455	9518	64			0	reverse
trnL(tag)	tRNA	9519	9581	63			0	reverse
rrnL	rRNA	9582	10810	1229			0	reverse
trnV	tRNA	10811	10873	63			0	reverse
rrnS	rRNA	10874	11620	747			0	reverse
trnM	tRNA	11621	11684	64			0	reverse
trnY	tRNA	11685	11747	63			1	reverse
trnC	tRNA	11749	11810	62			2	reverse
trnW	tRNA	11813	11880	68			-3	reverse
trnQ	tRNA	11878	11943	66			13	reverse
trnG	tRNA	11957	12019	63			0	reverse
trnE	tRNA	12020	12081	62			608	reverse
cox3	CDS	12690	13469	780	TAG	TAG	17	forward
trnK	tRNA	13487	13551	65			1	forward
trnA	tRNA	13553	13617	65			0	forward
trnR	tRNA	13618	13683	66			-1	forward
trnN	tRNA	13683	13748	66			0	forward
trnI	tRNA	13749	13816	68			0	forward
nad3	CDS	13817	14169	353	TAG	TA-	0	forward
trnS(gct)	tRNA	14170	14237	68			0	forward
nad2†	CDS	14238	15267	1030	TAG	T--	0	forward

Neritina usnea

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	1	1548	1548	TAG	TAA	11	forward
cox2	CDS	1560	2249	690	TAG	TAG	1	forward
trnD	tRNA	2251	2316	66			0	forward
atp8	CDS	2317	2481	165	TAG	TAA	6	forward
atp6	CDS	2488	3189	702	TAG	TAA	22	forward
trnF	tRNA	3212	3279	68			6	reverse
nad5	CDS	3286	4993	1708	TAG	TAA	0	reverse
trnH	tRNA	4994	5059	66			0	reverse
nad4	CDS	5060	6425	1366	TAG	T--	-7	reverse
nad4L	CDS	6419	6712	294	TAG	TAA	5	reverse
trnT	tRNA	6718	6785	68			6	forward
trnS(tga)	tRNA	6792	6856	65			5	reverse
cob	CDS	6862	7998	1137	TAG	TAA	12	reverse
nad6	CDS	8011	8517	507	TAG	TAA	1	reverse
trnP	tRNA	8519	8584	66			1	reverse
nad1	CDS	8586	9518	933	TAG	TAA	0	reverse
trnL(taa)	tRNA	9519	9586	68			12	reverse
trnL(tag)	tRNA	9599	9668	70			0	reverse
rrnL	rRNA	9669	10972	1304			0	reverse
trnV	tRNA	10973	11039	67			0	reverse
rrnS	rRNA	11040	11901	862			0	reverse
trnM	tRNA	11902	11968	67			5	reverse
trnY	tRNA	11974	12041	68			5	reverse
trnC	tRNA	12047	12110	64			0	reverse
trnW	tRNA	12111	12176	66			0	reverse
trnQ	tRNA	12177	12245	69			0	reverse
trnG	tRNA	12246	12311	66			2	reverse
trnE	tRNA	12314	12379	66			471	reverse
cox3	CDS	12851	13630	780	TAG	TAA	28	forward
trnK	tRNA	13659	13727	69			5	forward
trnA	tRNA	13733	13800	68			16	forward
trnR	tRNA	13817	13885	69			6	forward
trnN	tRNA	13892	13963	72			13	forward
trnI	tRNA	13977	14045	69			0	forward
nad3	CDS	14046	14399	354	TAG	TAA	5	forward
trnS(gct)	tRNA	14405	14472	68			0	forward
nad2†	CDS	14473	15574	1102	TAG	T--	-29	forward

Pleuropoma jana

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	1	1539	1539	TAG	TAA	13	forward
cox2	CDS	1553	2239	687	TAG	TAG	8	forward
trnG	tRNA	2248	2310	63			5	forward
trnD	tRNA	2316	2378	63			0	forward
atp8	CDS	2379	2537	159	TAG	TAA	11	forward
atp6	CDS	2549	3241	693	TAG	TAA	28	forward
trnM	tRNA	3270	3332	63			6	forward
trnY	tRNA	3339	3404	66			0	forward
trnM	tRNA	3405	3470	66			0	forward
rrnS	rRNA	3471	4299	829			0	forward
trnV	tRNA	4300	4365	66			0	forward
rrnL	rRNA	4366	5626	1261			0	forward
trnL(taa)	tRNA	5627	5690	64			0	forward
nad1	CDS	5691	6614	924	TAG	TAA	76	forward
trnP	tRNA	6691	6756	66			1	forward
nad6	CDS	6758	7247	490	TAG	TAG	-8	forward
cob	CDS	7255	8388	1141	TAG	T--	0	forward
trnS(tga)	tRNA	8389	8454	66			14	forward
trnT	tRNA	8469	8538	70			43	reverse
nad4L	CDS	8582	8881	300	TAG	TAA	-7	forward
nad4	CDS	8875	10242	1368	TTG	TAA	0	forward
trnH	tRNA	10243	10303	61			0	forward
nad5	CDS	10304	11996	1693	TAG	T--	0	forward
trnF	tRNA	11997	12063	67			783	forward
trnL(tag)	tRNA	12847	12910	64			3	forward
trnW	tRNA	12914	12980	67			26	forward
trnE	tRNA	13007	13068	62			2	forward
trnC	tRNA	13071	13134	64			7	forward
trnQ	tRNA	13142	13209	68			29	forward
cox3	CDS	13239	14018	780	ATG	TAA	11	forward
trnK	tRNA	14030	14087	58			0	forward
trnA	tRNA	14088	14154	67			0	forward
trnR	tRNA	14155	14216	62			2	forward
trnN	tRNA	14219	14285	67			8	forward
trnI	tRNA	14294	14359	66			3	forward
nad3	CDS	14363	14708	346	ATA	T--	0	forward
trnS(gct)	tRNA	14709	14776	68			1	forward
nad2	CDS	14778	15833	1056	TAG	TAA	18	forward

Titiscania sp.

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2307	3842	1536	ATG	TAA	9	forward
cox2	CDS	3852	4547	696	ATG	TAA	21	forward
trnD	tRNA	4569	4636	68			1	forward
atp8	CDS	4638	4817	180	ATG	TAG	53	forward
atp6	CDS	4871	5569	699	ATG	TAA	39	forward
trnF	tRNA	5609	5676	68			393	reverse
nad5	CDS	6070	7812	1743	ATG	TAA	2	reverse
trnH	tRNA	7815	7883	69			38	reverse
nad4	CDS	7922	9304	1383	ATG	TAG	-7	reverse
nad4L	CDS	9298	9594	297	ATG	TAA	10	reverse
trnT	tRNA	9605	9673	69			6	forward
trnS(tga)	tRNA	9680	9746	67			5	reverse
cob	CDS	9752	10888	1137	ATG	TAA	29	reverse
nad6	CDS	10918	11424	507	ATG	TAG	2	reverse
trnP	tRNA	11427	11496	70			45	reverse
nad1	CDS	11542	12477	936	ATG	TAG	1	reverse
trnL(taa)	tRNA	12479	12547	69			12	reverse
trnL(tag)	tRNA	12560	12627	68			0	reverse
rrnL	rRNA	12628	14042	1415			0	reverse
trnV	tRNA	14043	14110	68			0	reverse
rrnS	rRNA	14111	15046	936				reverse
cox3	CDS	1	288	288	partial	TAG	66	forward
trnK	tRNA	355	427	73			9	forward
trnA	tRNA	437	505	69			18	forward
trnR	tRNA	524	592	69			6	forward
trnN	tRNA	599	670	72			1	forward
trnI	tRNA	672	739	68			2	forward
nad3	CDS	742	1101	360	ATG	TAA	47	forward
trnS(gct)	tRNA	1149	1216	68			0	forward
nad2†	CDS	1217	2306	109	ATG	T--	0	forward

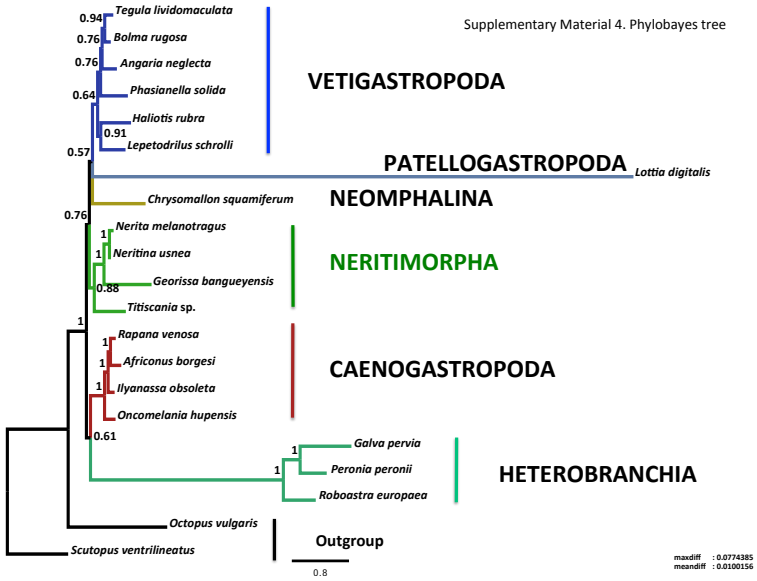
Theodoxus fluviatilis

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2479	4026	1548	ATG	TAA	10	forward
cox2	CDS	4037	4726	690	ATG	TAA	6	forward
trnD	tRNA	4733	4798	66			0	forward
atp8	CDS	4799	4963	165	ATG	TAA	8	forward
atp6	CDS	4972	5673	702	ATG	TAA	22	forward
trnF	tRNA	5696	5762	67			0	reverse
nad5	CDS	5763	7476	1714	ATG	T--	0	reverse
trnH	tRNA	7477	7541	65			0	reverse
nad4	CDS	7542	8907	1366	ATG	T--	-7	reverse
nad4L	CDS	8901	9194	294	ATG	TAA	4	reverse
trnT	tRNA	9199	9266	68			7	forward
trnS(tga)	tRNA	9274	9338	65			2	reverse
cob	CDS	9341	1048	114	ATG	TAA	9436	reverse
nad6	CDS	10485	10991	507	ATG	TAA	1	reverse
trnP	tRNA	10993	11056	64			1	reverse
nad1	CDS	11058	1199	933	ATG	TAA	10791	reverse
trnL(taa)	tRNA	11991	12058	68			0	reverse
trnL(tag)	tRNA	12059	12128	70			0	reverse
rrnL	rRNA	12129	13424	1296			0	reverse
trnV	tRNA	13425	13492	68			0	reverse
rrnS	rRNA	13493	14122	630			0	reverse
cox3	CDS	1	537	537	partial	TAG	26	forward
trnK	tRNA	564	631	68			5	forward
trnA	tRNA	637	704	68			19	forward
trnR	tRNA	724	792	69			12	forward
trnN	tRNA	805	879	75			4	forward
trnI	tRNA	884	951	68			0	forward
nad3	CDS	952	1305	354	ATG	TAG	3	forward
trnS(gct)	tRNA	1309	1376	68			0	forward
nad2†	CDS	1377	2478	1102	ATG	T--	-29	forward

Viana regina

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2246	3781	1536	ATG	TAG	80	forward
cox2	CDS	3862	4548	687	ATG	TAA	10	forward
trnG	tRNA	4559	4622	64			2	forward
trnD	tRNA	4625	4687	63			0	forward
atp8	CDS	4688	4846	159	ATG	TAA	18	forward
atp6	CDS	4865	5554	690	ATG	TAG	19	forward
trnM(cat)	tRNA	5574	5638	65			9	forward
trnY	tRNA	5648	5712	65			0	forward
trnM(cat)	tRNA	5713	5780	68			0	forward
rrnS	rRNA	5781	6617	837			0	forward
trnV	tRNA	6618	6682	65			0	forward
rrnL	rRNA	6683	7557	875			0	forward
cox3	CDS	1	339	>339	partial	TAA	1	forward
trnK	tRNA	341	403	63			3	forward
trnA	tRNA	407	475	69			1	forward
trnR	tRNA	477	541	65			15	forward
trnI	tRNA	557	622	66			6	forward
trnN	tRNA	629	697	69			20	forward
nad3	CDS	718	1121	403	ATG	T--	0	forward
trnS(gct)	tRNA	1121	1188	68			0	forward
nad2†	CDS	1189	2245	1057	ATG	T--	-29	forward

Data S4. Phylobayes tree



Additional taxa included in the PhyloBayes analysis

Species	Order	Length (bp)	GenBank Acc. No.	Publication
Tegula lividomaculata	Trochoidea	17,375	KT207826	Uribe et al., 2015
Bolma rugosa	Trochoidea	17,432	KT207824	Uribe et al., 2015
Lottia digitalis	Lottioidea	26,835	NC_007782	Simison et al., 2006
Peronia peronii	Onchidioidea	13,968	NC_016181	White et al., 2011
Galba pervia	Lymnaeoidea	13,768	NC_018536	Liu et al., 2012
Roboastra europaea	Anadoridoidea	14,472	NC_004321	Grande et al., 2002
Octopus vulgaris	Neocoleoidea	15,744	NC_006353	Yokobori et al., 2004
Scutopus ventrolineatus	Scutopodidae*	14,662	NC_025284	Osca et al., 2014

*unassigned to a superfamily

Grande, C., Templado, J., Cervera, J.L. and 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., Xu, M.-J., Lin, R.-Q., Zhou, D.-H. and Zhu, X.-Q., 2012. The complete mitochondrial genome of *Galba pervia* (Gastropoda: Mollusca), an intermediate host snail of Fasciola spp. PLoS ONE 7, e42172.

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

Simison, W.B., Lindberg, D.R. and Boore, J.L., 2006. Rolling circle amplification of metazoan mitochondrial genomes. Moleculaer Phylogenetic and Evolutions 39, 562 - 567.

Uribe, J.E., Kano, Y., Templado, J., Zardoya, R., 2015. Mitogenomics of Vetigastropoda: insights into the evolution of pallial symmetry. Zoologica Scripta doi: 10.1111/zsc.12146.

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

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

SUPPORTING INFORMATION 5

Data S1. Amplification strategy. Long PCR and primer walking primers.

Conus californicus

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Forcox1F	AGCTTTTGACTTTTACCCCCTGCTTTG	*cox1-rrnL* (5346)
Cdea16sR	CTACCTTTGCACGGTCAGAGTACC	
Forcox1R	GTCTACCGAACCTCCTGCATGAGCTAGG	*rrnL-cox3* (10219)
Cdea16sF	GCCTTATAATTGAAGGCTRGWATGAATGG	

Primer Link *rrnL*

Primer	Sequence 5'-3'	
16sinicioF2	TTCTGCCTGTTTAKCAAAAACATGGCTTC	*rrnL-rrnL* (425)
16sfinR	AAAGATAATGCTGTTATCCCTRCGG	

Conus (*Lautoconus*) *venulatus*

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Conus_cox1_F	AGTTTYTGRCTTCTTCCTCCTGCGCTT	*cox1-rrnL* (5371)
Conus_16S_R	GATTATGCTACCTTTGCACGGTCAGAG	
Conus_12S_F	GAGATAAGTCGTAACAYAGTAGGGGTAATG	*rrnL-cox1* (6024)
Conus_nd4_R	GAATTTAGGACTACCTCCGTGATGAATAG	
Conus_nd4_F	GTTTATTAAGCGTACTCGTCTTTGCAGCAT	*rrnL-cox1* (5926)
Conus_cox1_R	CCTAAAATAGAAGAHACMCCAGCAAGATG	

Conus (*Lautoconus*) *ventricosus*

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Conus_cox1_F	AGTTTYTGRCTTCTTCCTCCTGCGCTT	*cox1-rrnL* (5358)
Conus_16S_R	GATTATGCTACCTTTGCACGGTCAGAG	
Conus_12S_F	GAGATAAGTCGTAACAYAGTAGGGGTAATG	*rrnL-cox1* (6037)
Conus_nd4_R	GAATTTAGGACTACCTCCGTGATGAATAG	
Conus_nd4_F	GTTTATTAAGCGTACTCGTCTTTGCAGCAT	*rrnL-cox1* (5935)
Conus_cox1_R	CCTAAAATAGAAGAHACMCCAGCAAGATG	

Conus (*Lautoconus*) *hybridus*

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Cdeacox3F	ATGGCACGAAATCCATTTCATTTRGTTGA	*cox3-cox1* (3339)
COIbfol_R	TATAAAATDGGATCHCCACCTCCTGC	
COIfol_F	TATTTTCTACHAATCATAAAGATATTGG	*cox1-rrnL* (5618)
Cdea16sR	CTACCTTTGCACGGTCAGAGTACC	
Cdea16sF	GCCTTATAATTGAAGGCTRGWATGAATGG	*rrnL-trnF* (7011)
CdeaPheR	TACYTTAGCATCTTCAGCGCTAYGCTCT	

Primer Link *rrnL*

Primer	Sequence 5'-3'	
16sinicioF2	TTCTGCCTGTTTAKCAAAAACATGGCTTC	*rrnL-rrnL* (426)
16sfinR	AAAGATAATGCTGTTATCCCTRCGG	

Conus (*Eugeniconus*) *nobilis victor*

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Cdeacox3F	ATGGCACGAAATCCATTTCATTTRGTTGA	*cox3-cox1* (3397)
COIbfol_R	TATAAAATDGGATCHCCACCTCCTGC	
COIfol_F	TATTTTCTACHAATCATAAAGATATTGG	*cox1-rrnL* (5694)
Cdea16sR	CTACCTTTGCACGGTCAGAGTACC	
Cdea16sF	GCCTTATAATTGAAGGCTRGWATGAATGG	*rrnL-trnF* (7005)
CdeaPheR	TACYTTAGCATCTTCAGCGCTAYGCTCT	

Data S1 (cont.)

Primer Link *rrnL*

Primer	Sequence 5'-3'	
16sinicioF2	TTCTGCCTGTTTAKCAAAAACATGGCTTC	*rrnL-rrnL* (427)
16sfinR	AAAGATAATGCTGTTATCCCTRCGG	

Conasprella wakayamaensis

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Cdeacox3F	ATGGCACGAAATCCATTTCATTTRGTTGA	*cox3-cox1* (3324)
COIbfol_R	TATAAAATDGGATCHCCACCTCCTGC	
COIfol_F	TATTTTCTACHAATCATAAAGATATTGG	*cox1-rrnL* (6282)
Cdea16sR	CTACCTTTGCACGGTCAGAGTACC	
Cdea16sF	GCCTTATAATTGAAGGCTRGWATGAATGG	*rrnL-trnF* (7037)
CdeaPheR	TACYTTAGCATCTTCAGCGCTAYGCTCT	

Primer Link *rrnL*

Primer	Sequence 5'-3'	
16sinicioF2	TTCTGCCTGTTTAKCAAAAACATGGCTTC	*rrnL-rrnL* (445)
16sfinR	AAAGATAATGCTGTTATCCCTRCGG	

Liliconus sagei

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Cdeacox3F	ATGGCACGAAATCCATTTCATTTRGTTGA	*cox3-cox1* (3863)
Lilicox1R	CTGCACCTAAAATTGATGAAGCACCAGC	
Lilicox1F	TTAAGTCAACCTGGGGCTCTGTTAGG	*cox1-rrnL* (5553)
Cdea16sR	CTACCTTTGCACGGTCAGAGTACC	
Cdea16sF	GCCTTATAATTGAAGGCTRGWATGAATGG	*rrnL-trnF* (6964)
CdeaPheR	TACYTTAGCATCTTCAGCGCTAYGCTCT	

Primer Link *rrnL*

Primer	Sequence 5'-3'	
16sinicioF2	TTCTGCCTGTTTAKCAAAAACATGGCTTC	*rrnL-rrnL* (427)
16sfinR	AAAGATAATGCTGTTATCCCTRCGG	

Profundiconus terimachi

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Cdeacox3F	ATGGCACGAAATCCATTTCATTTRGTTGA	*cox3-rrnL* (8325)
Cdea16sR	CTACCTTTGCACGGTCAGAGTACC	
Cdea16sF	GCCTTATAATTGAAGGCTRGWATGAATGG	*rrnL-trnF* (700)
CdeaPheR	TACYTTAGCATCTTCAGCGCTAYGCTCT	

Primer Link *rrnL*

Primer	Sequence 5'-3'	
16sinicioF2	TTCTGCCTGTTTAKCAAAAACATGGCTTC	*rrnL-rrnL* (426)
16sfinR	AAAGATAATGCTGTTATCCCTRCGG	

Pseudolilliconus traillii

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Cdeacox3F	ATGGCACGAAATCCATTTCATTTRGTTGA	*cox3-rrnS* (8124)
Cdea16sR	CTACCTTTGCACGGTCAGAGTACC	
Pseu12sF	AATCTGTGAAAGTTTTGAGGGAAACCGGG	*rrnS-trnF* (8175)
CdeaPheR	TACYTTAGCATCTTCAGCGCTAYGCTCT	

Benthomangelia sp.

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)

Data S1 (cont.)

Cdeacox3F	ATGGCACGAAATCCATTTCATTTRGTTGA	*cox3-cox1* (3065)
Mangcox1R	ACAGCHCCTAAAATAGAAGAAACACC	
Mangcox1F	GGAGCTCCHGATATAGTWTTTCCTCG	*cox1-rrnL* (5261)
Cdea16sR	CTACCTTTGCACGGTCAGAGTACC	
Cdea16sF	GCCTTATAATTGAAGGCTRGWATGAATGG	*rrnL-trnF* (7019)
CdeaPheR	TACYTTAGCATCTTCAGCGCTAYGCTCT	

Primer Link *rrnL*

Primer	Sequence 5'-3'	
16sinicioF2	TTCTGCCTGTTTAKCAAAAACATGGCTTC	*rrnL-rrnL* (422)
16sfinR	AAAGATAATGCTGTTATCCCTRCGG	

Tomopleura sp.

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Cdeacox3F	ATGGCACGAAATCCATTTCATTTRGTTGA	*cox3-cox1* (3299)
Borcox1R	GATATAARATAGGATCWCCRCCTCCTGC	
Borcox1F	ATTGGAGGATTTGGRAATTGRTTRGTTCC	*cox1-rrnL* (5380)
Cdea16sR	CTACCTTTGCACGGTCAGAGTACC	
Cdea16sF	GCCTTATAATTGAAGGCTRGWATGAATGG	*rrnL-cob* (3249)
BorcobR	GGGCACCTTCCAATCCAAGTTAAAAC	
Borcob4F	GAAGTCCTATTCGAAAAGTTCATCCGG	*cobtrnF* (4737)
CdeaPheR	TACYTTAGCATCTTCAGCGCTAYGCTCT	

Primer Link *rrnL*

Primer	Sequence 5'-3'	
16sinicioF2	TTCTGCCTGTTTAKCAAAAACATGGCTTC	*rrnL-rrnL* (423)
16sfinR	AAAGATAATGCTGTTATCCCTRCGG	

Glyphostoma sp.

Long PCR

Primer	Sequence 5'-3'	Fragment (bp)
Cdeacox3F	ATGGCACGAAATCCATTTCATTTRGTTGA	*cox3-cox1* (3145)
Clathcox1R	AGCACCTAAAATAGAAGAAACACCNGCAAG	
Clathcox1F	GGGGCTCCYGATATGGTYTTTCCTCG	*cox1-rrnL* (5351)
Cdea16sR	CTACCTTTGCACGGTCAGAGTACC	
Cdea16sF	GCCTTATAATTGAAGGCTRGWATGAATGG	*rrnL-nad4* (5147)
Conus_nd4_R	GAATTTAGGACTACCTCCGTGATGAATAG	

Primer Link *rrnL*

Primer	Sequence 5'-3'	
16sinicioF2	TTCTGCCTGTTTAKCAAAAACATGGCTTC	*rrnL-rrnL* (422)
16sfinR	AAAGATAATGCTGTTATCCCTRCGG	

Data S2. Best fit partitions and evolutionary substitution models as selected by Partition Finder

Conidae matrix

		Set Partition	Best Model	Alpha	Pinvar	A<->C	A<->G	A<->T	C<->G	C<->T	pi(A)	pi(C)	pi(G)	pi(T)
Best Partition to CDS genes (BIC =72340.71)	1	*atp6-8*	MtMam+I+G+F	0.34	0.09									
	2	*cob*	MtArt+I+G	0.98	0.54									
	3	*cox1-2-3*	MtMam+I+G+F	0.61	0.62									
	4	*nad1-2-3-4s-5-6*	MtMam+I+G+F	0.37	0.10									
Best Partition to rARNs genes (BIC =46628.80)	5	*rrnL-S*	GTR+I+G	0.61	0.27	0.29	5.14	1.04	0.08	5.80	0.36	0.12	0.18	0.33

Conus matrix

		Set Partition	Best Model	Alpha	Pinvar	A<->C	A<->G	A<->T	C<->G	C<->T	pi(A)	pi(C)	pi(G)	pi(T)
Best Partition CDS genes (BIC =117487.53)	1	*atps* 1st	HKY+G	1.57	0.56	1.79	35.88	3.42	1.57	43.41	0.28	0.15	0.20	0.34
	2	*atps* 2nd	GTR+I+G	1.12	0.70	0.73	10.38	0.97	8.95	6.46	0.17	0.22	0.13	0.47
	3	*atps* 3rd	HKY+I+G	1.11	0.02	27.02	226.17	4.16	12.15	367.55	0.28	0.08	0.13	0.49
	4	*cob* 1st	GTR+G	1.45	0.68	2.09	13.92	1.22	0.00	47.87	0.22	0.20	0.26	0.30
	5	*cob* 2nd	GTR+I+G	0.77	0.81	0.00	2.88	0.69	4.58	6.03	0.21	0.21	0.15	0.41
	6	*cob* 3rd	HKY+I+G	0.57	0.03	15.09	1745.51	4.62	54.55	1796.65	0.32	0.10	0.09	0.47
	7	*cox1-2-3* 1st	GTR+I+G	0.68	0.65	0.96	5.93	0.00	0.00	33.81	0.23	0.16	0.31	0.28
	8	*cox1-2-3* 2nd	GTR+I+G	1.20	0.89	1.90	6.12	0.46	9.23	3.83	0.19	0.21	0.18	0.39
	9	*cox1-2-3* 3rd	HKY+G	0.96	0.00	4.13	295.70	6.55	32.23	342.50	0.30	0.06	0.16	0.47
	10	*nad1-2-3-4s-5-6* 1st	GTR+I+G	0.47	0.27	3.16	15.82	1.71	2.01	40.21	0.28	0.15	0.21	0.33
	11	*nad1-2-3-4s-5-6* 2nd	GTR+I+G	0.69	0.62	1.09	12.35	1.11	9.74	8.34	0.16	0.20	0.16	0.46
	12	*nad1-2-3-4s-5-6* 3rd	GTR+I+G	1.65	0.02	13.47	256.02	6.36	25.46	197.03	0.33	0.10	0.12	0.42
Best Partition to rARNs genes (BIC =25260.64)	13	*rrnL-S*	GTR+I+G	0.28	0.06	0.98	22.03	3.82	1.24	23.43	0.35	0.13	0.19	0.32

Data S3. Mitochondrial genome features

Californiconus californicus

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	1	1548	1548	ATG	TAA	119	forward
cox2	CDS	1668	2354	687	ATG	TAA	3	forward
trnD	tRNA	2358	2427	70			0	forward
atp8	CDS	2428	2586	159	ATG	TAA	2	forward
atp6	CDS	2589	3284	696	ATG	TAA	38	forward
trnM	tRNA	3323	3389	67			10	reverse
trnY	tRNA	3400	3467	68			0	reverse
trnC	tRNA	3468	3534	67			0	reverse
trnW	tRNA	3535	3600	66			-3	reverse
trnQ	tRNA	3598	3664	67			20	reverse
trnG	tRNA	3685	3750	66			-1	reverse
trnE	tRNA	3750	3817	68			1	reverse
rrnS	rRNA	3819	4770	952			1	forward
trnV	tRNA	4772	4838	67			0	forward
rrnL	rRNA	4839	6198	1360			1	forward
trnL (tag)	tRNA	6200	6268	69			8	forward
trnL (taa)	tRNA	6277	6345	69			0	forward
nad1	CDS	6346	7287	942	ATG	TAA	6	forward
trnP	tRNA	7294	7361	68			0	forward
nad6	CDS	7362	7859	498	ATG	TAA	18	forward
cob	CDS	7878	9017	1140	ATG	TAA	13	forward
trnS (tga)	tRNA	9031	9095	65			0	forward
trnT	tRNA	9096	9161	66			20	reverse
nad4L	CDS	9182	9478	297	ATG	TAG	-7	forward
nad4	CDS	9472	10843	1372	ATG	T--	0	forward
trnH	tRNA	10844	10907	64			0	forward
nad5	CDS	10908	12626	1719	ATG	TAA	0	forward
trnF	tRNA	12627	12694	68			97	forward
cox3	CDS	12792	13571	780	ATG	TAA	24	forward
trnK	tRNA	13596	13662	67			8	forward
trnA	tRNA	13671	13737	67			4	forward
trnR	tRNA	13742	13811	70			7	forward
trnN	tRNA	13819	13886	68			8	forward
trnI	tRNA	13895	13964	70			0	forward
nad3	CDS	13965	14318	354	ATG	TAA	1	forward
trnS (cgt)	tRNA	14320	14387	68			0	forward
nad2	CDS	14388	15443	1056	ATG	TAG	1	forward

Conus (*Lautoconus*) *venulatus*

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	1	1548	1548	ATG	TAA	130	forward
cox2	CDS	1679	2365	687	ATG	TAA	0	forward
trnD	tRNA	2366	2432	67			0	forward
atp8	CDS	2433	2594	162	ATG	TAG	6	forward
atp6	CDS	2601	3296	696	ATG	TAA	40	forward
trnM	tRNA	3337	3404	68			13	reverse
trnY	tRNA	3418	3483	66			4	reverse
trnC	tRNA	3488	3552	65			0	reverse
trnW	tRNA	3553	3618	66			1	reverse
trnQ	tRNA	3620	3677	58			14	reverse
trnG	tRNA	3692	3757	66			1	reverse
trnE	tRNA	3759	3826	68			0	reverse
rrnS	rRNA	3827	4781	955			0	forward
trnV	tRNA	4782	4848	67			0	forward
rrnL	rRNA	4849	6212	1364			0	forward
trnL (tag)	tRNA	6213	6282	70			6	forward
trnL (taa)	tRNA	6289	6357	69			0	forward
nad1	CDS	6358	7299	942	ATG	TAA	3	forward
trnP	tRNA	7303	7369	67			0	forward
nad6	CDS	7370	7870	501	ATG	TAA	11	forward
cob	CDS	7882	9021	1140	ATG	TAA	12	forward
trnS (tga)	tRNA	9034	9098	65			9	forward
trnT	tRNA	9108	9174	67			21	reverse
nad4L	CDS	9196	9492	297	ATG	TAG	-7	forward
nad4	CDS	9486	10867	1382	ATG	TA-	0	forward
trnH	tRNA	10868	10934	67			0	forward
nad5	CDS	10935	12649	1715	ATG	TA-	0	forward
trnF	tRNA	12650	12715	66			126	forward
cox3	CDS	12842	13621	780	ATG	TAA	25	forward
trnK	tRNA	13647	13716	70			4	forward
trnA	tRNA	13721	13787	67			16	forward
trnR	tRNA	13804	13872	69			10	forward
trnN	tRNA	13883	13951	69			14	forward
trnI	tRNA	13966	14035	70			5	forward
nad3	CDS	14041	14394	354	ATG	TAA	8	forward
trnS (cgt)	tRNA	14403	14470	68			0	forward
nad2	CDS	14471	15524	1054	ATG	T--	0	forward

Conus (*Lautoconus*) *ventricosus*

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	1	1548	1548	ATG	TAA	130	forward
cox2	CDS	1679	2365	687	ATG	TAA	0	forward
trnD	tRNA	2366	2432	67			0	forward
atp8	CDS	2433	2594	162	ATG	TAA	6	forward
atp6	CDS	2601	3296	696	ATG	TAA	35	forward
trnM	tRNA	3332	3399	68			12	reverse
trnY	tRNA	3412	3477	66			1	reverse
trnC	tRNA	3479	3543	65			0	reverse
trnW	tRNA	3544	3610	67			1	reverse
trnQ	tRNA	3612	3669	58			15	reverse
trnG	tRNA	3685	3750	66			2	reverse
trnE	tRNA	3753	3818	66			0	reverse
rrnS	rRNA	3819	4769	951			0	forward
trnV	tRNA	4770	4837	68			0	forward
rrnL	rRNA	4838	6201	1364			0	forward
trnL (tag)	tRNA	6202	6271	70			18	forward
trnL (taa)	tRNA	6290	6358	69			0	forward
nad1	CDS	6359	7300	942	ATG	TAG	4	forward
trnP	tRNA	7305	7372	68			0	forward
nad6	CDS	7373	7873	501	ATG	TAG	11	forward
cob	CDS	7885	9024	1140	ATG	TAG	11	forward
trnS (tga)	tRNA	9036	9100	65			9	forward
trnT	tRNA	9110	9176	67			20	reverse
nad4L	CDS	9197	9493	297	ATG	TAG	-7	forward
nad4	CDS	9487	10868	1382	ATG	TA-	0	forward
trnH	tRNA	10869	10935	67			0	forward
nad5	CDS	10936	12651	1716	ATG	TAA	10	forward
trnF	tRNA	12662	12726	65			126	forward
cox3	CDS	12853	13632	780	ATG	TAA	23	forward
trnK	tRNA	13656	13725	70			6	forward
trnA	tRNA	13732	13797	66			18	forward
trnR	tRNA	13816	13884	69			10	forward
trnN	tRNA	13895	13964	70			9	forward
trnI	tRNA	13974	14043	70			5	forward
nad3	CDS	14049	14402	354	ATG	TAA	8	forward
trnS (cgt)	tRNA	14411	14478	68			0	forward
nad2	CDS	14479	15534	1056	ATG	TAA	0	forward

Conus (*Lautoconus*) *hybridus*

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2655	4202	1548	ATG	TAA	120	forward
cox2	CDS	4323	5009	687	ATG	TAG	0	forward
trnD	tRNA	5010	5076	67			0	forward
atp8	CDS	5077	5238	162	ATG	TAA	6	forward
atp6	CDS	5245	5940	696	ATG	TAG	34	forward
trnM	tRNA	5975	6042	68			12	reverse
trnY	tRNA	6055	6121	67			1	reverse
trnC	tRNA	6123	6187	65			0	reverse
trnW	tRNA	6188	6253	66			-3	reverse
trnQ	tRNA	6251	6317	67			10	reverse
trnG	tRNA	6328	6393	66			2	reverse
trnE	tRNA	6396	6461	66			0	reverse
rrnS	rRNA	6462	7409	948			0	forward
trnV	tRNA	7410	7476	67			0	forward
rrnL	rRNA	7477	8840	1364			0	forward
trnL (tag)	tRNA	8841	8910	70			2	forward
trnL (taa)	tRNA	8913	8981	69			0	forward
nad1	CDS	8982	9923	942	ATG	TAA	4	forward
trnP	tRNA	9928	9994	67			0	forward
nad6	CDS	9995	10495	501	ATG	TAA	12	forward
cob	CDS	10508	11647	1140	ATG	TAG	12	forward
trnS (tga)	tRNA	11660	11724	65			10	forward
trnT	tRNA	11735	11801	67			20	reverse
nad4L	CDS	11822	12118	297	ATG	TAG	-7	forward
nad4	CDS	12112	13493	1382	ATG	TA-	0	forward
trnH	tRNA	13494	13560	67			0	forward
nad5	CDS	13561	15276	1716	ATG	TAA	0	forward
cox3	CDS	1	753	753	---	TAA	22	forward
trnK	tRNA	776	847	72			5	forward
trnA	tRNA	853	919	67			18	forward
trnR	tRNA	938	1006	69			10	forward
trnN	tRNA	1017	1085	69			9	forward
trnI	tRNA	1095	1164	70			5	forward
nad3	CDS	1170	1523	354	ATG	TAA	8	forward
trnS (cgt)	tRNA	1532	1599	68			0	forward
nad2	CDS	1600	2654	1055	ATG	TA-	0	forward

Conus (*Eugeniconus*) *nobilis victor*

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2701	4248	1548	ATG	TAA	159	forward
cox2	CDS	4408	5094	687	ATG	TAA	0	forward
trnD	tRNA	5095	5161	67			0	forward
atp8	CDS	5162	5323	162	ATG	TAG	6	forward
atp6	CDS	5330	6025	696	ATG	TAG	36	forward
trnM	tRNA	6062	6129	68			8	reverse
trnY	tRNA	6138	6206	69			0	reverse
trnC	tRNA	6207	6270	64			0	reverse
trnW	tRNA	6271	6336	66			-3	reverse
trnQ	tRNA	6334	6412	79			-4	reverse
trnG	tRNA	6409	6474	66			1	reverse
trnE	tRNA	6476	6540	65			0	reverse
rrnS	rRNA	6541	7502	962			0	forward
trnV	tRNA	7503	7570	68			0	forward
rrnL	rRNA	7571	8942	1372			0	forward
trnL (tag)	tRNA	8943	9012	70			6	forward
trnL (taa)	tRNA	9019	9087	69			0	forward
nad1	CDS	9088	10029	942	ATG	TAA	4	forward
trnP	tRNA	10034	10100	67			0	forward
nad6	CDS	10101	10601	501	ATG	TAA	9	forward
cob	CDS	10611	11750	1140	ATG	TAA	12	forward
trnS (tga)	tRNA	11763	11827	65			9	forward
trnT	tRNA	11837	11903	67			20	reverse
nad4L	CDS	11924	12220	297	ATG	TAG	-7	forward
nad4	CDS	12214	13595	1382	ATG	TA-	0	forward
trnH	tRNA	13596	13663	68			0	forward
nad5	CDS	13664	15379	1716	ATG	TAA	0	forward
cox3	CDS	1	780	780	---	TAG	31	forward
trnK	tRNA	812	881	70			20	forward
trnA	tRNA	902	968	67			15	forward
trnR	tRNA	984	1054	71			9	forward
trnN	tRNA	1064	1131	68			7	forward
trnI	tRNA	1139	1211	73			4	forward
nad3	CDS	1216	1569	354	ATG	TAG	8	forward
trnS (cgt)	tRNA	1578	1645	68			0	forward
nad2	CDS	1646	2700	1055	ATG	TA-	0	forward

Conasprella wakayamaensis

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2639	4210	1572	ATG	TAA	734	forward
cox2	CDS	4945	5630	686	ATG	TA-	0	forward
trnD	tRNA	5631	5699	69			0	forward
atp8	CDS	5700	5858	159	ATG	TAA	3	forward
atp6	CDS	5862	6557	696	ATG	TAA	46	forward
trnM	tRNA	6604	6670	67			1	reverse
trnY	tRNA	6672	6738	67			0	reverse
trnC	tRNA	6739	6804	66			0	reverse
trnW	tRNA	6805	6870	66			-2	reverse
trnQ	tRNA	6869	6945	77			-2	reverse
trnG	tRNA	6944	7010	67			0	reverse
trnE	tRNA	7011	7077	67			0	reverse
rrnS	rRNA	7078	8029	952			0	forward
trnV	tRNA	8030	8097	68			0	forward
rrnL	rRNA	8098	9479	1382			0	forward
trnL (tag)	tRNA	9480	9548	69			5	forward
trnL (taa)	tRNA	9554	9621	68			0	forward
nad1	CDS	9622	10563	942	ATG	TAA	5	forward
trnP	tRNA	10569	10638	70			0	forward
nad6	CDS	10639	11139	501	ATG	TAG	21	forward
cob	CDS	11161	12300	1140	ATG	TAA	11	forward
trnS (tga)	tRNA	12312	12376	65			1	forward
trnT	tRNA	12378	12445	68			36	reverse
nad4L	CDS	12482	12778	297	ATG	TAG	-7	forward
nad4	CDS	12772	14143	1372	ATG	T--	0	forward
trnH	tRNA	14144	14208	65			0	forward
nad5	CDS	14209	15927	1719	ATG	TAG	0	forward
cox3	CDS	1	759	759	---	TAA	24	forward
trnK	tRNA	784	852	69			6	forward
trnA	tRNA	859	925	67			0	forward
trnR	tRNA	926	995	70			11	forward
trnN	tRNA	1007	1077	71			10	forward
trnI	tRNA	1088	1158	71			2	forward
nad3	CDS	1161	1512	352	ATG	T--	0	forward
trnS (cgt)	tRNA	1513	1580	68			0	forward
nad2	CDS	1581	2636	1056	ATG	TAA	0	forward

Lilliconus sagei

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2964	4511	1548	ATG	TAA		forward
cox2	CDS	4584	5270	687	ATG	TAA		forward
trnD	tRNA	5277	5346	70				forward
atp8	CDS	5347	5508	162	ATG	TAA		forward
atp6	CDS	5518	6213	696	ATG	TAA		forward
trnM	tRNA	6249	6317	69				reverse
trnY	tRNA	6318	6387	70				reverse
trnC	tRNA	6386	6451	66				reverse
trnW	tRNA	6454	6520	67				reverse
trnQ	tRNA	6518	6584	67				reverse
trnG	tRNA	6592	6658	67				reverse
trnE	tRNA	6657	6724	68				reverse
rrnS	rRNA	6799	7762	964				forward
trnV	tRNA	7763	7831	69				forward
rrnL	rRNA	7832	9174	1343				forward
trnT	tRNA	9181	9250	70				forward
nad1	CDS	9317	10258	942	ATG	TAG		forward
trnP	tRNA	10262	10328	67				forward
nad6	CDS	10329	10826	498	ATG	TAA		forward
cob	CDS	10837	11976	114	ATG	TAA		forward
trnS (tga)	tRNA	11977	12041	65				reverse
nad4L	CDS	12043	12339	297	ATG	TAG		forward
nad4	CDS	12333	13701	1369	GTG	T--		forward
trnH	tRNA	13702	13766	65				forward
nad5	CDS	13767	15485	1719	ATG	TAA		forward
cox3	CDS	1	756	756	---	TAA		forward
trnK	tRNA	767	833	67				forward
trnA	tRNA	844	912	69				forward
trnN	tRNA	951	1017	67				forward
trnI	tRNA	1021	1090	70				forward
nad3	CDS	1094	1447	354	ATG	TAA		forward
trnS (cgt)	tRNA	1449	1516	68				forward
nad2	CDS	1517	2590	1074	ATG	TAA		forward
trnL (taa)	tRNA	2619	2691	73				forward
trnL (tag)	tRNA	2836	2903	68				forward

Profundiconus terimachi

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2636	4183	1548	ATG	TAA	160	forward
cox2	CDS	4344	5028	685	ATG	T--	0	forward
trnD	tRNA	5029	5097	69			0	forward
atp8	CDS	5098	5256	159	ATG	TAA	4	forward
atp6	CDS	5261	5956	696	ATG	TAA	38	forward
trnM	tRNA	5995	6062	68			4	reverse
trnY	tRNA	6067	6133	67			0	reverse
trnC	tRNA	6134	6199	66			0	reverse
trnW	tRNA	6200	6265	66			-3	reverse
trnQ	tRNA	6263	6341	79			-10	reverse
trnG	tRNA	6332	6398	67			0	reverse
trnE	tRNA	6399	6466	68			0	reverse
rrnS	rRNA	6467	7422	956			0	forward
trnV	tRNA	7423	7490	68			0	forward
rrnL	rRNA	7491	8877	1387			0	forward
trnL (tag)	tRNA	8878	8946	69			2	forward
trnL (taa)	tRNA	8949	9017	69			0	forward
nad1	CDS	9018	9959	942	ATG	TAG	9	forward
trnP	tRNA	9969	10035	67			0	forward
nad6	CDS	10036	10536	501	ATG	TAG	6	forward
cob	CDS	10543	11682	114	ATG	TAA	8	forward
trnS (tga)	tRNA	11691	11755	65			2	forward
trnT	tRNA	11758	11824	67			15	reverse
nad4L	CDS	11840	12136	297	ATG	TAG	-7	forward
nad4	CDS	12130	13498	1369	ATG	T--	0	forward
trnH	tRNA	13499	13563	65			0	forward
nad5	CDS	13564	15279	1716	ATG	TAA	0	forward
cox3	CDS	1	771	771	---	TAA	18	forward
trnK	tRNA	790	856	67			8	forward
trnA	tRNA	865	931	67			0	forward
trnR	tRNA	932	1001	70			14	forward
trnN	tRNA	1016	1081	66			-5	forward
trnI	tRNA	1077	1153	77			2	forward
nad3	CDS	1156	1509	354	ATG	TAA	0	forward
trnS (cgt)	tRNA	1510	1577	68			0	forward
nad2	CDS	1578	2635	1058	ATG	TA-	0	forward

Pseudolilliconus traillii

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2611	4157	1547	ATG	TAA	-12	forward
trnT	tRNA	4146	4207	62			10	reverse
cox2	CDS	4218	4904	687	ATG	TAG	10	forward
trnD	tRNA	4915	4983	69			5	forward
atp8	CDS	4989	5144	156	ATG	TAA	16	forward
atp6	CDS	5161	5888	728	ATG	TA-	0	forward
trnM	tRNA	5889	5954	66			8	reverse
trnY	tRNA	5963	6029	67			1	reverse
trnC	tRNA	6031	6093	63			2	reverse
trnW	tRNA	6096	6159	64			-3	reverse
trnQ	tRNA	6157	6229	73			-4	reverse
trnG	tRNA	6226	6288	63			-1	reverse
trnE	tRNA	6288	6350	63			0	reverse
rrnS	rRNA	6351	7278	928			0	forward
trnV	tRNA	7279	7349	71			0	forward
rrnL	rRNA	7350	8657	1308			0	forward
trnL (tag)	tRNA	8658	8724	67			3	forward
trnL (taa)	tRNA	8728	8792	65			0	forward
nad1	CDS	8793	9737	945	GTG	TAA	2	forward
trnP	tRNA	9740	9806	67			0	forward
nad6	CDS	9807	10298	492	GTG	TAA	2	forward
cob	CDS	10301	11439	1139	ATG	TAG	15	forward
trnS (cgt)	tRNA	11455	11519	65			8	forward
nad4L	CDS	11528	11824	297	ATG	TAG	-7	forward
nad4	CDS	11818	13183	1366	ATG	T--	0	forward
trnH	tRNA	13184	13250	67			0	forward
nad5	CDS	13251	14963	1713	ATA	TAA	0	forward
cox3	CDS	1	771	771	---	TAA	23	forward
trnK	tRNA	794	860	67			3	forward
trnA	tRNA	863	929	67			7	forward
trnR	tRNA	936	999	64			5	forward
trnN	tRNA	1004	1074	71			2	forward
trnI	tRNA	1076	1144	69			1	forward
nad3	CDS	1145	1497	353	ATG	TA-	1	forward
trnS (cgt)	tRNA	1498	1565	68			1	forward
nad2	CDS	1566	2610	1045	ATG	T--	0	forward

Benthomangelia sp.

		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2585	4135	1551	ATG	TAA	63	forward
cox2	CDS	4199	4883	685	ATG	T--	0	forward
trnD	tRNA	4884	4949	66			0	forward
atp8	CDS	4950	5108	159	ATG	TAA	2	forward
atp6	CDS	5111	5806	696	ATG	TAG	36	forward
trnM	tRNA	5843	5909	67			-1	reverse
trnY	tRNA	5909	5974	66			1	reverse
trnC	tRNA	5976	6038	63			-6	reverse
trnW	tRNA	6033	6104	72			-3	reverse
trnQ	tRNA	6102	6169	68			3	reverse
trnG	tRNA	6173	6237	65			-1	reverse
trnE	tRNA	6237	6301	65			0	reverse
rrnS	rRNA	6302	7252	951			0	forward
trnV	tRNA	7253	7315	63			0	forward
rrnL	rRNA	7316	8649	1334			0	forward
trnL (tag)	tRNA	8650	8715	66			0	forward
trnL (taa)	tRNA	8716	8784	69			0	forward
nad1	CDS	8785	9726	942	ATG	TAA	8	forward
trnP	tRNA	9735	9802	68			0	forward
nad6	CDS	9803	10304	502	ATG	T--	0	forward
cob	CDS	10305	11444	114	ATG	TAA	11	forward
trnS (tga)	tRNA	11456	11520	65			0	forward
trnT	tRNA	11521	11584	64			6	reverse
nad4L	CDS	11591	11887	297	ATG	TAG	-7	forward
nad4	CDS	11881	13252	1372	ATG	T--	0	forward
trnH	tRNA	13253	13317	65			1	forward
nad5	CDS	13319	15034	1716	ATG	TAG	0	forward
cox3	CDS	1	756	756	---	TAA	11	forward
trnK	tRNA	768	832	65			0	forward
trnA	tRNA	833	899	67			1	forward
trnR	tRNA	901	967	67			5	forward
trnN	tRNA	973	1041	69			4	forward
trnI	tRNA	1046	1112	67			1	forward
nad3	CDS	1114	1465	352	ATG	T--	0	forward
trnS (cgt)	tRNA	1466	1533	68			0	forward
nad2	CDS	1534	2584	1051	ATG	T--	0	forward

Tomopleura sp.

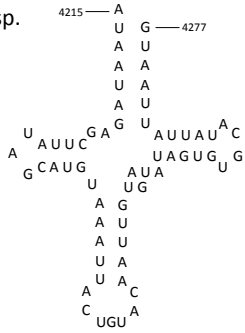
		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2615	4162	1548	ATG	TAG	44	forward
trnT	tRNA	4207	4276	70			-2	reverse
cox2	CDS	4275	4959	685	ATG	T--	0	forward
trnD	tRNA	4960	5028	69			0	forward
atp8	CDS	5029	5187	159	ATG	TAA	7	forward
atp6	CDS	5195	5890	696	ATG	TAA	36	forward
trnM	tRNA	5927	5993	67			3	reverse
trnY	tRNA	5997	6064	68			0	reverse
trnC	tRNA	6065	6129	65			0	reverse
trnW	tRNA	6130	6195	66			-3	reverse
trnQ	tRNA	6193	6263	71			-2	reverse
trnG	tRNA	6262	6326	65			0	reverse
trnE	tRNA	6327	6396	70			0	reverse
rrnS	rRNA	6397	7347	951			0	forward
renV	tRNA	7348	7414	67			0	forward
rrnL	rRNA	7415	8772	1358			0	forward
trnL (tag)	tRNA	8773	8841	69			5	forward
trnL (taa)	tRNA	8847	8915	69			0	forward
nad1	CDS	8916	9857	942	ATG	TAA	3	forward
trnP	tRNA	9861	9929	69			0	forward
nad6	CDS	9930	10430	501	ATG	TAA	11	forward
cob	CDS	10442	11581	114	ATG	TAG	6	forward
trnS (tga)	tRNA	11588	11652	65			84	forward
nad4L	CDS	11737	12033	297	ATG	TAG	-7	forward
nad4	CDS	12027	13365	1339	ATG	TAA	33	forward
trnH	tRNA	13399	13462	64			1	forward
nad5	CDS	13464	15182	1719	ATG	TAA	0	forward
cox3	CDS	1	756	756	---	TAG	29	forward
trnK	tRNA	786	855	70			5	forward
trnA	tRNA	861	928	68			0	forward
trnR	tRNA	929	994	66			0	forward
trnN	tRNA	995	1063	69			6	forward
trnI	tRNA	1070	1139	70			0	forward
nad3	CDS	1140	1493	354	ATG	TAA	0	forward
trnS (cgt)	tRNA	1494	1561	68			0	forward
nad2	CDS	1562	2614	1053	ATG	TAA	0	forward

Glyphostoma sp.

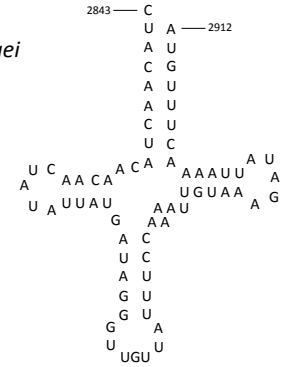
		Gene			Codon			
Gene	**Type**	**Start**	**Stop**	**Length**	**Start**	**Stop**	**Intergenic**	**Strand**
cox1	CDS	2672	4219	1548	ATG	TAG	93	forward
cox2	CDS	4313	4997	685	ATG	T--	0	forward
trnD	tRNA	4998	5066	69			0	forward
atp8	CDS	5067	5225	159	ATG	TAG	20	forward
atp6	CDS	5246	5941	696	ATG	TAA	32	forward
trnM	tRNA	5974	6040	67			1	reverse
trnY	tRNA	6042	6109	68			12	reverse
trnC	tRNA	6122	6186	65			-6	reverse
trnW	tRNA	6181	6254	74			-3	reverse
trnQ	tRNA	6252	6318	67			3	reverse
trnG	tRNA	6322	6388	67			0	reverse
trnE	tRNA	6389	6454	66			0	reverse
rrnS	rRNA	6455	7399	945			0	forward
trnV	tRNA	7400	7467	68			0	forward
rrnL	rRNA	7468	8845	1378			0	forward
trnL (tag)	tRNA	8846	8914	69			10	forward
trnL (taa)	tRNA	8925	8993	69			0	forward
nad1	CDS	8994	9935	942	ATG	TAA	5	forward
trnP	tRNA	9941	10009	69			0	forward
nad6	CDS	10010	10519	510	ATG	TAA	19	forward
cob	CDS	10539	11678	114	ATG	TAG	5	forward
trnS (cgt)	tRNA	11684	11748	65			0	forward
trnT	tRNA	11749	11815	67			9	reverse
nad4L	CDS	11825	12121	297	ATG	TAG	-7	forward
nad4	CDS	12115	13370	132	ATG	---	0	forward
cox3	CDS	1	741	741	---	TAA	14	forward
trnK	tRNA	756	820	65			34	forward
trnA	tRNA	855	921	67			1	forward
trnR	tRNA	923	988	66			17	forward
trnN	tRNA	1006	1071	66			24	forward
trnI	tRNA	1096	1163	68			4	forward
nad3	CDS	1168	1521	354	ATG	TAG	0	forward
trnS (cgt)	tRNA	1522	1589	68			0	forward
nad2	CDS	1590	2642	1053	ATG	TAA	29	forward

Data S4. Secundary structure of tRNAs from manual search

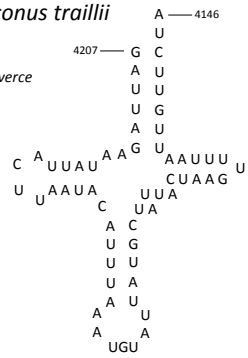
Tomopleura sp.
tRNA^{Thr} *reverce*



Liciconus sagei
tRNA^{Thr}



Pseudolilliconus traillii
tRNA^{Thr} *reverce*



Liciconus sagei
tRNA^{Leu} *reverce*