

## Barcoding coffee grounds—Exploring pteropod gastropod biodiversity with dregs in collection jars

Christina Franziska Laibl<sup>1,2</sup>, Juan Lucas Cervera Currado<sup>3,4</sup>, Jérôme Morinière<sup>5</sup>,  
Michael Schrödl<sup>1,2,6</sup>

<sup>1</sup> SNSB-Bavarian State Collection of Zoology, Münchhausenstraße 21, 81247 München, Germany.  
(CFL) (Corresponding author) E-mail: [laibl@snsb.de](mailto:laibl@snsb.de). ORCID iD: <https://orcid.org/0000-0002-0306-3467>  
(MS) E-mail: [schroedl@snsb.de](mailto:schroedl@snsb.de). ORCID iD: <https://orcid.org/0000-0002-7629-1911>

<sup>2</sup> Ludwig-Maximilians-Universität München, Faculty of Biology, Großhadernerstraße 2, 82152 Planegg-Martinsried, Germany.

<sup>3</sup> Departamento de Biología, Facultad de Ciencias del Mar y Ambientales, Campus de Excelencia Internacional del Mar (CEIMAR), Universidad de Cádiz, Av. República Saharaui, s/n, P.O. Box 40, 11510 Puerto Real, Cádiz, Spain.  
(JLCC) E-mail: [lucas.cervera@uca.es](mailto:lucas.cervera@uca.es). ORCID iD: <https://orcid.org/0000-0002-8337-2867>

<sup>4</sup> Instituto Universitario de Investigación Marina (INMAR), Campus de Excelencia Internacional del Mar (CEIMAR), Universidad de Cádiz, Av. República Saharaui, s/n, P.O. Box 40, 11510 Puerto Real, Cádiz, Spain.

<sup>5</sup> AIM - Advanced Identification Methods GmbH, Niemeyerstr. 1, 04179 Leipzig, Germany.  
(JM) E-mail: [jerome.moriniere@aim.science](mailto:jerome.moriniere@aim.science). ORCID iD: <https://orcid.org/0000-0001-9167-6409>

<sup>6</sup> GeoBioCentrer LMU, Centre of Geobiology and Biodiversity Research at the Ludwig-Maximilians-University Munich, Germany.

**Summary:** Despite their cosmopolitan occurrence and massive plankton sampling during expeditions, the genetic diversity within Pteropoda Cuvier, 1804 is still largely unexplored. In this study we present a next-generation environmental barcoding approach to zooplankton bulk samples, which were collected during the circumglobal 2010 Malaspina expedition to evaluate pteropod diversity. We introduce a technique that avoids destructive procedures and leaves material intact for further morphological investigations. We extracted DNA out of the dregs (organic material such as mucus or body parts) of 27 sample containers for molecular barcoding (average 100-260 bp of COI). We were able to identify 7128 operational taxonomic units corresponding to the species composition contained in the examined samples. Among them were three species of thecosome pteropods, *Creseis acicula*, *Creseis virgula* and *Cavolinia inflexa*, which are discussed with respect to their taxonomy and their geographic distribution. Unidentified gymnosomes were also present in our samples from warmer regions in oceanic waters of the southern Indian Ocean. To facilitate identification of species, it is beneficial to create a better database of pteropod COI barcodes. Furthermore, gathering environmental barcoding data on a broad global scale will help to better understand species abundance and distribution of pteropods in the world's oceans, and potentially those of other planktonic organisms.

**Keywords:** Mollusca; Gastropoda; plankton; environmental DNA; circumglobally; pteropod diversity; Malaspina expedition.

**Barcoding a partir de posos de café - Explorando la biodiversidad de gasterópodos pterópodos a partir de posos de frascos de colección**

**Resumen:** A pesar de su presencia cosmopolita y las actividades de muestreo masivo de plancton durante las expediciones, la diversidad genética dentro de los Pteropoda Cuvier, 1804 está todavía inexplorada en gran medida. En este estudio se presenta una aproximación desde el barcoding ambiental aplicada a muestras generales de zooplancton recogidas durante la expedición circumglobal “Malaspina 2010”, con el fin de evaluar la diversidad de pterópodos. Se introduce una técnica que evita procedimientos destructivos de tal modo que el material permanece intacto para futuras investigaciones morfológicas. Extrajimos ADN de los posos (material orgánico como moco o partes del cuerpo) de 27 recipientes de muestras para el barcoding (promedio de 100- 260 bp de COI). Se pudieron identificar 7128 “OTUs” correspondientes a la composición de las especies contenidas en las muestras examinadas. Entre ellas se encontraron tres especies de pterópodos tecosomados, *Creseis acicula*, *Creseis virgula* y *Cavolinia inflexa*, cuya taxonomía y distribución geográfica son discutidas. Gimnosomados no identificados procedentes de regiones más templadas de aguas oceánicas del sur del Océano Índico también estaban presentes. Para facilitar la identificación de especies, es beneficioso crear una base de datos ampliada de códigos de barras COI de pterópodos. Además, la recopilación de datos de barcoding ambiental a una escala mundial amplia ayudará a comprender mejor la abundancia y distribución de especies de pterópodos en los océanos del mundo y de otros posibles organismos planctónicos.

**Palabras clave:** Mollusca; Gastropoda; plancton; ADN ambiental; circumglobal; diversidad de pterópodos; expedición Malaspina.

**Citation/Como citar este artículo:** Laibl C.F., Cervera Currado J.L., Morinière J., Schrödl M. 2023. Barcoding coffee grounds—Exploring pteropod gastropod biodiversity with dregs in collection jars. *Sci. Mar.* 87(2): e061. <https://doi.org/10.3989/scimar.05275.061>

**Editor:** J. Viñas.

**Received:** February 17, 2022. **Accepted:** January 10, 2023. **Published:** June 11, 2023.

**Copyright:** © 2023 CSIC. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License.

## INTRODUCTION

There are more than 230000 known metazoan species populating the world's oceans. However, as a result of climate change, ocean acidification and marine pollution, the increasing loss of biodiversity presents a daunting challenge to taxonomists, requiring the discovery and analysis of biodiversity at a greatly accelerated pace. In the face of growing extinction rates that are without much doubt outpacing the number of discoveries of new taxa, fast and accurate biodiversity analysis methods are urgently needed (Bucklin et al. 2011). Especially problematic is the taxonomic treatment of large-scale environmental bulk samples such as phyto- and zooplankton. Sorting and identifying the various organisms requires a lot of time before reliable diversity assessments are possible. Furthermore, traditional morphological approaches are limited and less efficient for analysing bulk samples or samples lacking distinguishing phenotypic features (for example immature or damaged specimens). It is well established that genetic markers, and especially COI barcoding, are a complementary tool to traditional morphology-based taxonomic research for the identification and delimitation of different lineages (Hajibabaei et al. 2007).

Applying barcoding methods to (environmental) organismic DNA material improves traditional biomonitoring activity: excluding uncertainties such as morphological identification, low detection probabilities and sampling methods (challenges of gear deployment) increases confidence in the monitoring results (e.g. Bucklin et al. 2021, Wang et al. 2021, Di Capua et al. 2022). DNA barcoding could thus accelerate the inventory analysis of biological diversity, especially of bulk samples such as those of sediments or plankton, and of older museum samples. A good example of such bulk sample collection is the worldwide multidisciplinary Malaspina expedition, in which over 70000 samples of water, air and plankton were gathered in different ocean regions from the surface down to 5000 m depth. This immense collection was sorted and divided into several sub-collections that were accessible for scientific research. Our focus is on the holopelagic group Pteropoda (thecosomes and gymnosomes), a group of gastropods with an important ecological role in the marine environment as microplankton grazers and as prey for fish and other zooplankton. Despite their cosmopolitan distribution, the genetic diversity within this group is still largely uncertain (Burrige et al. 2017a,b).

A comprehensive insight into the present diversity and distribution of these planktonic molluscs is an important prerequisite for stating possible future changes in species composition and also in species-specific responses to changing conditions in the marine environment. So far, studies focusing on pteropod species distribution patterns were geographically limited to certain marine regions (e.g. Jennings et al. 2010, Burrige et al. 2017b), and there are still areas where pteropod diversity remains unknown (Burrige et al. 2017b). Here, with the help of the 2010 Malaspina zooplankton samples, we generated new data using an environmental barcoding approach on the debris taken from 27 selected bulk samples from different oceans in order to gain a better understanding of species abundance and distribution of pteropods and to explore the potential of the method for broad-scale application.

## MATERIALS AND METHODS

### Sampling

Plankton samples were gathered globally at 154 locations. The hauls were conducted in the morning and in the evening. Out of 154 we selected 27 locations of interest in the Caribbean Sea, the Atlantic Ocean and the Indian Ocean (Table 1). The locations were chosen to coincide with those where frequent occurrence of certain pteropod species had been previously described (e.g. Burrige et al. 2017b), (Fig. 1).

### Sample preparation

Out of 27 bulk samples the preservation medium and the bottom content in the collection jars were extracted and filtered for organic material. Species identification of organic material was performed using DNA metabarcoding following the protocol published in Hausmann et al. (2020). Each single sample was dried in a 60°C oven for at least eight hours and subsequently homogenized in a FastPrep96 machine (MP Biomedicals) using sterile steel beads in order to generate a homogeneous mixture of faeces material before it was submitted for metabarcoding (conducted by AIM GmbH). Prior to DNA extraction, 1 mg of each homogenizate was weighed into sample vials and processed using adapted volumes of lysis buffer with the DNeasy 96 Blood and Tissue Kit (Qiagen) following the manufacturer's instructions. For

Table 1. – Sampling location, pteropod species detection. I, present; D, day catch; N, night catch; \* bin sharing, () sampling location [the first number refers to the collector, the second, separated by an underscore, to the collection point along the route. See Table 2 for more details].

Locality data:	North Atlantic 12 29.90 N, 25 59.17 W	North Atlantic 16 09.84 N, 26 01.53 W	North Atlantic 14 31.18 N, 26 00.02 W	Indian Ocean 28 07.65 S, 66 29.59 E	Indian Ocean 29 49.65 S, 79 36.66 E	Caribbean Sea 15 31.50 N, 67 00.86 W
Species						
Gymnosomata <i>sp.</i>				I/D (3_54)		
<i>Cavolinia inflexa</i>	I/D (1_11)		I/D (1_10)			I/N (7_130)
<i>Creseis acicula</i>				I/D (3_54)		
<i>Creseis virgula</i>		I/D * (1_9)		I/D (3_54)	I/N (3_58)	I/N (7_130) *
*	Creseis_conica   Creseis_virgula Atlantic_Ocean   Belize   Bermuda   Mexico			* Creseis_chierchiaie   Creseis_virgula Atlantic_Ocean   Belize   Mexico		



Fig. 1. – Course of the research vessel Hesperides during the 2010 Malaspina Expedition. Locations: Caribbean Sea (1), Atlantic Ocean (2) and Indian Ocean (3); red dot marks Cadiz; scale bar, 1000 km. Google (2021). Available at: <https://www.google.de/maps/> (Accessed: 06.02.2021).

amplification of the CO1-5P target region and preparation of the MiSeq libraries, a two-step PCR was performed. First, a 313-bp-long mini-barcode region was amplified by PCR (Leray et al. 2013, Morinière et al. 2016) using forward and reverse high-throughput sequencing (HTS) primers equipped with complementary sites for the Illumina sequencing tails. In a subsequent PCR reaction, index primers with unique i5 and i7 in-line tags and sequencing tails were used for amplification of indexed amplicons. Equimolar amplicon pools were then created and size-selected using preparative gel electrophoresis. Cleanup and concentration of amplicons were performed using the GeneJet Extraction Kit (Life Technologies). A bioanalyser (High Sensitivity DNA Kit, Agilent Technologies) was used for a final check of the bp distribution and concentration of the amplicons before the creation of the final library. All samples were pooled into one library, equimolar adjusted to 100 ng µL<sup>-1</sup>. Samples had DNA concentrations between 10.4 and 12.4 ng µL<sup>-1</sup>. HTS was performed on an Illumina MiSeq using v2 chemistry (2\*250 bp,

500 cycles, maximum of 20mio reads) (Illumina). All samples were analysed on a single MiSeq run. The bioinformatics processing of raw FASTQ files from Illumina was carried out using the VSEARCH suite v2.9.1 (Rognes et al. 2016) and Cutadapt v1.18 (Martin 2011). Forward and reverse reads in each sample were merged using the VSEARCH program “fastq\_mergepairs” with a minimum overlap of 10 bp, yielding approximately 313 bp sequences. Forward and reverse primers which were not reliably detected at >90% identity, were removed with Cutadapt using the “discard\_untrimmed” option. Quality filtering was done with the “fastq\_filter” in VSEARCH, and sequences with zero expected errors were kept (“fastq\_maxee” 1).

Sequences were dereplicated with “derep\_full-length,” first at the sample level and then concatenated into one FASTA file, which was subsequently dereplicated. Chimeric sequences were filtered out from the FASTA file using the “uchime\_denovo” VSEARCH program. The remaining sequences were then clustered into operational taxonomic units (OTUs) at 97%

identity with “cluster\_size”, a greedy centroid-based clustering program. OTUs were blasted against a custom Animalia database downloaded from BOLD on 28 November 2018, including taxonomy and barcode index number (BIN) information, by means of Geneious (v.10.2.5, Biomatters, Auckland, New Zealand) and following methods described in Morinière et al. (2016).

The resulting csv file, which included the OTU ID, BOLD Process ID, BIN, Hit-%-ID value (percentage of overlap similarity (identical basepairs) of an OTU query sequence with its closest counterpart in the database), length of the top BLAST hit sequence, phylum, class, order, family, genus, and species information for each detected out, was exported from Geneious and combined with the OTU table generated by the bioinformatic pipeline. The combined results table was then filtered by Hit-%-ID value and total read numbers per OTU. All entries with identifications below 97% and total read numbers below 0.01% of the summed reads per sample were removed from the analysis. OTUs were then assigned to the respective BIN. Additionally, the API provided by BOLD was used to retrieve BIN species and BIN countries for every OTU, and the Hit-%-IDs were aggregated over OTUs that found a hit in the same BIN and shown in the corresponding column as % range. To validate the BOLD BLAST results, a separate BLAST search was carried out in Geneious (using the same parameters) against a local copy of the NCBI nucleotide database downloaded from <ftp://ftp.ncbi.nlm.nih.gov/blast/db/>. Interactive Krona charts were produced from the taxonomic information using KronaTools v1.3 (Ondov et al. 2011). Species identification was based on the HTS of OTUs, before blasting and assignment to BINs (Ratnasingham and Hebert 2013) which are considered to be a good proxy for species numbers (Hausmann et al. 2013, Ratnasingham and Hebert 2013).

## Analyses

Generated sequence data were further analysed with a focus on three objectives: 1) assessment of species diversity using ABGD molecular species delineation (Puillandre et al. 2012); 2) discovery of new spe-

cies or potentially cryptic species assemblages; and 3) testing of distribution ranges of the respective lineages, especially in regions where pteropod diversity remains poorly understood so far.

In the final analysis, publicly available sequence data from GenBank and BOLD were included for phylogenetic testing. To code the COI gene, the sequences were first aligned in protein and then converted into nucleotide using ClustalW implemented in the software package MEGA Version 3.0. This method allowed us to maximize the homology between nucleotide positions when amino acid deletion/insertion occurred.

## RESULTS

Barcode sequences were obtained with an average length of 100-268 bp. We obtained 69997 sequence clusters (coverage  $\geq$ CD-HIT-EST, min. 90.9% to 90.9%, max. 100% to 100%) that were blasted against 270000 DNA barcodes of identified specimens on the BOLD database BLAST (2019). This resulted in the detection of 206 BINs fitting the criterion of at least 90% sequence identity (supplementary material). This results in BINs that overlap with large groups at a phylum level: Arthropoda, Chaetognatha, Chordata, Cnidaria, Heterokontophytes, Mollusca, Nematoda, Porifera and Rotifera. For Mollusca we were able to identify 26 BINs (% identity: highest matches: 1.00, lowest 0.74; BOLD database BLAST) (Supplementary Table 1 a, b). Of these total sequence data, 6% were assigned to Gastropoda (Supplementary Table 2). Out of the gastropod sequence data, 46% matched with pteropod sequences (Table 2), 37% with thecosome and 9% with gymnosome data. Seven BINs were identified for Pteropoda (% identity: highest match, 1.00; lowest match, 0.909). Thirty-seven percent of snail sequence data matched with thecosome species. The sequence data matched with the euthecosome group Cavolinoidea, with the family of Creseidae (*Creseis acicula* (Rang, 1828) and *Creseis virgula* (Rang, 1828)) and Cavoliniidae (*Cavolinia inflexa* (Lesueur, 1813)). BIN sharing was observed for OTU\_3752 (*Creseis chierchiae*|*Creseis virgula*) and OTU\_13187 (*Creseis conica*|*Creseis virgula*) (Supplementary Table 1a).

Table 2. – Results of BLAST tool: BOLD (including BIN information) and GenBank.

BOLD database results of BLAST tool	BIN sharing	BIN location	GenBank/NCBI database results of BLAST tool
Gymnosomata	No	No location	Gymnosomata_sp.
Gymnosomata	No	No location	Gymnosomata_sp.
<i>Cavolinia inflexa</i>	No	Belize Mexico	<i>Cavolinia inflexa</i>
<i>Creseis acicula</i>	No	No location	<i>Creseis acicula</i>
<i>Creseis "clava"</i>	No	Bermuda	<i>Creseis acicula</i>
<i>Creseis virgula</i>	<i>Creseis chierchiae</i>   <i>Creseis virgula</i>	Atlantic_Ocean Belize Mexico	<i>Creseis virgula</i>
<i>Creseis virgula</i>	<i>Creseis conica</i>   <i>Creseis virgula</i>	Atlantic_Ocean Belize Bermuda Mexico	<i>Creseis virgula</i>
<i>Creseis virgula</i>	No	United_States	<i>Creseis virgula</i>

No correlation between OTU size and percentage species identity was observed (Fig. S1). ABGD analyses with cavoliniid sequence data grouped *Creseis clava* (note: outdated synonym) OTUs together with *Creseis virgula* OTUs. The *Cavolinia inflexa* OTU was distanced from the *Creseis acicula* OTU (initial partition (1-10) with prior maximal distance  $P=1.00e-03$  –  $P=1.00e-01$ ; barcode gap distance = 0.148 – barcode gap distance = 0.279; Jukes-Cantor JC69 distance MinSlope=1.500000) (Fig. S2).

Nine percent of total sequence data was assigned to gymnosome origin but not further assignable on family or species level, ABGD analyses revealed low genetic distance (range initial partition (1-10) with prior maximal distance  $P=1.00e-03$   $P=1.00e-01$ ; barcode gap distance = 0.095 Barcode gap distance = 0.299, Jukes-Cantor JC69 distance, MinSlope=1.500000), indicating intraspecific assignment of gymnosome specimens (Fig. S3).

## DISCUSSION

### Experimental approach

Standard DNA barcoding approaches for pteropod gastropods have already been successfully applied (e.g. Hunt et al. 2010, Jennings et al. 2010) and provide a suitable tool for assessing large-scale biodiversity (Makiola et al. 2020, Chimeno et al. 2022). With our approach we established a feasible protocol that overcomes current obstacles of having to sort the visible specimens or tissues. The sediment on the bottom of the provided plankton samples contained enough organic material (torn body parts, mucus and other secretions) to extract DNA. By using the dreg of the bottom, this method accelerates taxonomic procedures as the specimens in the sample jar remained unharmed and therefore stayed in suitable condition for further morphologic studies that might be of interest. This protocol might also be suitable in processing older collection material and therefore serves in the growing field of museomics, here enhancing comparative studies with modern and historical DNA material.

### Quality of data

This non-destructive method is based on the use of scarce material. The quantity of DNA material is already limited by the relatively small size of our targeted specimens as well as their restricted geographical and circadian presence in the field. DNA degradation, collection age, preparation treatment and storage conditions have a big impact on the quality and quantity of the already limited DNA material (Janik et al. 2020). Nevertheless, we were able to detect targeted molluscan and pteropod DNA. Many benthic marine gastropod species have pelagic larvae, so they may be encountered in traces in any plankton samples (e.g. Pulmonata, Caenogastropoda s.o and Stylommato-phora). Surprisingly, we found quite a large amount of DNA of terrestrial specimens in our samples, even though secure laboratory guidelines applied in order

to avoid cross contamination during extraction and amplification were reasonable. Cross contamination with land snail material in the field due to net storage ashore is possible. Contamination is ruled out regarding the amount and diversity of terrestrial gastropods in our data set. Scarce availability of DNA material in combination with too many PCR amplification cycles can lead to formation of chimeric products (Fonseca et al. 2012). Using a smaller number of PCR cycles served as a precaution of formation of such chimeric DNA fragments here, so we discard this as a main explanation. Terrestrial DNA matches with species identity percentages of 0.787% to 0.837%. This might be due to a lack of comparative data as it is well known that the likelihood of finding matches in public databases for invertebrates is lower than for vertebrates (Harris et al. 2016). This could lead to mis-assignment of a barcode to the wrong species with high confidence. Missing target taxa on reference databases leads to the risk of making both false positive and false negative taxonomic assignments. False positives occur due to mis-assignment of a barcode to the wrong species with high confidence because the target species is missing in the database, and false negatives occur because of gaps in the database resulting in low confidence assignments (Porter and Hajibabaei 2018). Matches to terrestrial data were not as high as those of our data assigned to marine snail DNA (0.93-1). As organisms in our samples varied in size, shape and anatomy (e.g. crustaceans vs. molluscs), we expected a disproportion in organismic material and therefore in the presence of usable DNA material. Length of the sequences is ruled out as they had a satisfying length of >221 bp. For further analyses the application of more sensitive and stricter quality filters seems a beneficial recommendation to avoid a trade-off between quality and quantity, low-input DNA material and bioinformatic obstacles (chimera), and most importantly, to eradicate statistical imbalances of species material in the dreg and preservation medium to avoid such statistic outbreaks in the future.

### Notes on systematics of Pteropoda with a focus on euthecosome Cavolinioida

Three suborders divide the euthyneurian order Pteropoda Cuvier 1804 (WoRMs; <http://www.marinespecies.org>; 2022): Pseudothecosomata Meisenheimer, 1905, Euthecosomata Meisenheimer, 1905 and Gymnosomata Blainville, 1824. Peijinenburg et al. (2020) argued against the recent classification of Bouchet et al. (2017) and advocated two suborders, Gymnosomata and Thecosomata, where the latter is divided into Euthecosomata and Pseudothecosomata. Unfortunately, neither Pseudothecosomata nor Gymnosomata are represented in this study, so we further focus on Euthecosomata. Within Euthecosomata there are two superfamilies; Limacinoidea Gray, 1840 and Cavolinioida Gray, 1850 (Note: it is commented with “(1815)” in WoRMs). Cavolinioida comprises eight families, out of which two are represented in this study: Cavoliniidae Gray, 1850 (Note: it is also commented with

“(1815)” in WoRMs) and Creseidae Rampal, 1973. Creseidae comprise i.a. the genera: *Boasia* Dall, 1889 (*Boasia chierchiaie* (Boas, 1886)) and *Creseis* Rang, 1828, which will be further investigated here. Creseidae seems to be polyphyletic (Klussmann-Kolb and Dinapoli 2006, Corse et al. 2013, Burrridge et al. 2017a); current molecular studies place the genera *Styliola* Gray, 1847 and *Hyalocylis* Fol, 1875 apart from Creseidae (Corse et al. 2013, Burrridge et al. 2017a). The complicated history of *Creseis* nomenclature was discussed in Gasca and Janssen (2014) and Janssen (2018). Since their original descriptions, species of *Creseis* have been synonymized or separated into several formae (Janssen 2006, 2007, 2012). Especially the interpretations of *Creseis acicula* (Rang, 1828) and *C. clava* (Rang, 1828) have been confusing (Janssen 2018). The status of *C. clava* is unaccepted (Janssen 2018); it acts as synonym for *Creseis acicula* (Rang, 1828) (accepted, WoRMs) and will be referred to as such in the following.

We hope to contribute to and update the state of knowledge by proceeding with future investigations on interoceanic differences among the available sampling material on *Cavolinia inflexa*, *Creseis acicula* and *Creseis virgula*, as so far molecular backup is missing (e.g. Gasca and Janssen 2014). We plan to use more markers and longer barcodes in our future studies and to investigate the respective collection jars and pursue morphological studies if necessary.

### Gymnosomata

Shelled pteropod taxa have been in focus because of their usefulness in studying global climate change (i.e. ocean acidification), but unshelled pteropods have not yet been examined to the same extent. Some work has been done on an ecological and anatomic level, but little on genetics despite an increase in the recent years (e.g. Stromeck et al. 2015, Yamazaki et al. 2017, Kohnert et al. 2020). Gymnosomes are less abundant than thecosomes, but they are ecologically very important because of their feeding manners, primarily preying on thecosomes (Lalli and Gilmer 1989). Interestingly, some of our sequenced material from samples from a free water day catch in the Indian Ocean (roughly 28° S, 66° E) was assigned to Gymnosomata sp. According to Burrridge et al. (2017b), the presence of gymnosomes (and thecosomes) in their sampling material came from (sub)tropical free waters within a longitude gradient of ~28°N and ~28°S (Atlantic Ocean), and were most abundant in sub-Antarctic waters and rather less in warmer waters (e.g. Weldrick et al. 2019). Burrridge et al. (2017b) did not assign any species level. In our study, comparative data from online databanks were lacking at lower levels beyond order. To our knowledge, emergence of gymnosomes in our sampling locality (oceanic free water) is rather rare. Putatively, a connection between migration via the Agulhas current and influences by cold waters from the circumpolar current led to the appearance in our material.

### Thecosomata

Thecosome material indicates presence of the “usual suspects” in the sample jars: the cosmopolite species *Creseis acicula*, *Creseis virgula* and *Cavolinia inflexa* (Lesueur, 1813). Present in almost all the world’s oceans, these species are mostly found in warm water territories. *C. inflexa* was the most dominant we had in our samples, which we anticipated, as it is in general the most common representative of *Cavolinia* in the Atlantic Ocean. Our findings (appearance/detection of *C. inflexa* in spacious, scattered samples) agree with the accepted knowledge that *C. inflexa* is common and distributed widely and provide further input in the ongoing discussions about its taxonomic status, as it is a putative species complex (Rampal 2002, Janssen et al. 2019). The same is true for *Creseis acicula* and *Creseis virgula*. Morphological distinctions of *Creseis* species have always been subject to many controversies, as can be seen in Frontier (1965), Rampal (1985, 2002), Janssen (2007), Gasca and Janssen (2014). Genetic studies (e.g. Klussmann-Kolb and Dinapoli (2006) and Corse et al. (2013)) did help to build a more solid classification, but it is still open for taxonomic debates.

### E-DNA barcoding approach for Pteropoda

Using tools like BOLD and GenBank allows quick and easy barcode and phylogenetic analysis. But there are persisting issues. Identification errors found in already published sequence data are rarely re-evaluated. Using standard procedures there is a 95% probability of finding incorrectly described metazoan sequences in GenBank, ranging from 1% (Mollusca and Arthropoda) to 6.9% (Gastrotricha). Consequently, the increasing popularity of DNA barcoding and metabarcoding analysis may lead to overestimation of species diversity (e.g. Mioduchowska et al. 2018). Lack of species-specific comparative data and misidentifications/ incorrect sequence data in previously and newly published data are due to amplification of non-target taxa and insufficient analysis of the obtained sequences (Mioduchowska et al. 2018). The difficult taxonomic history of *Creseis* is still mirrored in our bioinformatic analyses (current study: Bold and Genbank). Unfortunately, outdated synonyms are still used for certain taxa, in our case *C. acicula*, which was wrongly assigned to *C. clava*. Furthermore, BIN sharing in two samples occurred. This is correctable on a small scale, but for a bigger approach it might lead to problematic consequences and therefore shows the need for updates.

### Barcoding as a tool for monitoring

We advocate using collection material from previous expeditions to monitor environmentally influenced changes in pteropod abundance/behaviour, using it as a benchmark for the occurrence of common traits or derivations. In addition to measuring biodiversity by monitoring species availability, putative changes in behaviour might also help to understand the effects of climate change on these marine organisms. Diel vertical migration is a known phenomenon

for most thecosome species, and in our samples we found DNA material in night and day hauls in anticipated quantities. Swimming and sinking behaviour by these pelagic snails is important in their ecology, predator-prey interaction, and vertical distribution (Karakas et al. 2020). Despite the costs, benefits like niche partitioning, metabolic advantage due to colder temperatures at depth, avoidance of light, high temperatures and predators seem to advocate this behaviour (Hays 2003, Antezana 2009). Factors such as ocean acidification might harmfully affect the migrating ability by altering shell condition (and thickness), leading to misbalancing in factors involved in locomotion and buoyancy processes, as already shown for other thecosomes (*Limacina retroversa*, here: Manno et al. 2012, Adhikari et al. 2016). This will have fatal consequences for the individuals and will potentially result in ecological cascades in the long run. This process will be mirrored by the catch success in future plankton hauls. Thus, if data are accessible for longer time periods and large geographic areas, comparisons and statements about harmful effects and future developments can be made for certain ecological key species, as in our case thecosome pteropods.

## CONCLUSION

In this pioneering study we applied a non-invasive, next-generation environmental barcoding approach to several selected (meso)zooplankton bulk samples collected during the 2010 Malaspina global circumnavigation. On a small scale we were able to support existing knowledge of the distribution of *Cavolinia inflexa*, *C. virgula* and *C. acicula*, and we made surprising findings about the putative broader distribution of gymnosomes hypothesized today. Environmental DNA approaches may streamline the search for new pteropod species in unsorted museum jars and streamline historic and future monitoring efforts. Limitations to our approach are related to the lack of comparative barcoding data from pteropods, especially of gymnosome data, and outdated use of synonyms and potential misidentifications in online sequence bases, which call for re-evaluation and up-dating of existing published data. With more and taxonomically broader barcoding sequences available in public databases, our environmental barcoding approach will improve our understanding of global species diversity and distribution patterns of Pteropoda and other planktonic organisms.

## ACKNOWLEDGEMENTS

We wish to thank the Malacological Society of London (YRA) for funding this project. Special thanks go to Juan Ignacio González-Gordillo (University of Cádiz) for his support and help in choosing and gathering suitable samples. Stefan Filser and Andrew Brodie are thanked for their constructive input. Special thanks are due for providing the samples to the project “Expedición de circumnavegación Malaspina 2010: Cambio global y exploración de la biodiversidad del Océano global (CSD2008-00077)”, funded by the Spanish Ministry of Economy and Competitiveness.

## REFERENCES

- Adhikari D., Webster D. R., Yen J. 2016. Portable tomographic PIV measurements of swimming shelled Antarctic pteropods. *Exp. Fluids*. 57: 1-17.  
<https://doi.org/10.1007/s00348-016-2269-7>
- Antezana T. 2009. Species-specific patterns of diel migration into the Oxygen Minimum Zone by euphausiids in the Humboldt Current Ecosystem. *Prog. Oceanogr.* 83: 228-236.  
<https://doi.org/10.1016/j.pocean.2009.07.039>
- Bucklin A., Steinke D., Blanco-Bercial L. 2011. DNA barcoding of marine metazoa. *Annu. Rev. Mar. Science* 3: 471-508.  
<https://doi.org/10.1146/annurev-marine-120308-080950>
- Bucklin A., Peijnenburg K.T., Kosobokova K.N., et al. 2021. Toward a global reference database of COI barcodes for marine zooplankton. *Mar. Biol.* 168: 1-26.  
<https://doi.org/10.1007/s00227-021-03887-y>
- Burridge A.K., Hörnlein C., Janssen A.W., et al. 2017a. Time-calibrated molecular phylogeny of pteropods. *PLoS ONE* 12: e0177325.  
<https://doi.org/10.1371/journal.pone.0177325>
- Burridge A.K., Goetze E., Wall-Palmer D., et al. 2017b. Diversity and abundance of pteropods and heteropods along a latitudinal gradient across the Atlantic Ocean. *Prog. Oceanogr.* 158: 213-223.  
<https://doi.org/10.1016/j.pocean.2016.10.001>
- Bouchet P., Rocroi J.-P., Hausdorf B., et al. 2017. Revised classification, nomenclator and typification of gastropod and monoplacophoran families. *Malacologia* 61: 1-526.
- Chimeno C., Hübner J., Seifert L., et al. 2022. Depicting environmental gradients from Malaise trap samples: Is ethanol-based DNA metabarcoding enough? *Insect Conservation and Diversity*.  
<https://doi.org/10.1111/icad.12609>
- Corse É., Rampal J., Cuoc C., et al. 2013. Phylogenetic analysis of Thecosomata Blainville, 1824 (Holoplanktonic Opisthobranchia) using morphological and molecular data. *PLoS ONE* 8: e59439.  
<https://doi.org/10.1371/journal.pone.0059439>
- Di Capua I., D'Angiolo R., Piredda R., et al. 2022. From Phenotypes to Genotypes and Back: Toward an Integrated Evaluation of Biodiversity in Calanoid Copepods. *Front. Mar. Sci.* 75.  
<https://doi.org/10.3389/fmars.2022.833089>
- Fonseca V. G., Nichols B., Lallias D., et al. 2012. Sample richness and genetic diversity as drivers of chimera formation in nSSU metagenetic analyses. *Nucleic Acids Res.* 40: e66.  
<https://doi.org/10.1093/nar/gks002>
- Frontier S. 1965. Le problème des *Creseis*. *Océanographie (Nosy-Bé), Cah. ORSTOM. Sér. Sci. Hum.* 3:11-17.
- Gasca R., Janssen A. W. 2014. Taxonomic review, molecular data and key to the species of Creseidae from the Atlantic Ocean. *J. Molluscan Stud.* 80: 35-42.  
<https://doi.org/10.1093/mollus/eyt038>
- Harris D. J., Rosado D., Xavier R. 2016. DNA barcoding reveals extensive mislabeling in seafood sold in Portuguese supermarkets. *J. Aquat. Food Prod. Technol.* 25: 1375-1380.  
<https://doi.org/10.1080/10498850.2015.1067267>
- Hausmann A., Godfray H. C.J., Huemer P., et al. 2013. Genetic patterns in European geometrid moths revealed by the Barcode Index Number (BIN) system. *PLoS ONE* 8: e84518.  
<https://doi.org/10.1371/journal.pone.0084518>
- Hausmann A., Segerer A.H., Greifenstein T., et al. 2020. Toward a standardized quantitative and qualitative insect monitoring scheme. *Ecol. Evol.* 10: 4009-4020.  
<https://doi.org/10.1002/ece3.6166>
- Hays G. C. 2003. A review of the adaptive significance and ecosystem consequences of zooplankton diel vertical migrations. *Migrations and dispersal of marine organisms. Hydrobiologia* 503: 163-170.  
[https://doi.org/10.1007/978-94-017-2276-6\\_18](https://doi.org/10.1007/978-94-017-2276-6_18)
- Hajibabaei M., Singer G.A., Hebert P.D., et al. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends Genet.* 23: 167-172.  
<https://doi.org/10.1016/j.tig.2007.02.001>
- Janik P., Ronikier M., Ronikier A. 2020. New protocol for successful isolation and amplification of DNA from exiguous fractions of specimens: a tool to overcome the basic obstacle in molecular analyses of myxomycetes. *PeerJ* 8: e8406.  
<https://doi.org/10.7717/peerj.8406>

- Janssen A. 2006. Notes on the systematics, morphology and biostratigraphy of fossil holoplanktonic Mollusca. On the status of some pteropods (Gastropoda, Euthecosomata) from the Miocene of New Zealand, referred to as species of *Vaginella*. *Basteria* 70: 71-83.
- Janssen A.W. 2007. Holoplanktonic Mollusca (Gastropoda: Pterotracheoidea, Janthinoidea, Thecosomata and Gymnosomata) from the Pliocene of Pangasinan (Luzon, Philippines). *Scr. Geol.* 135: 29-177.
- Janssen A. 2012. Early Pliocene heteropods and pteropods (Mollusca, Gastropoda) from Le Puget-sur-Argens (Var), France. *Cainozoic Res.* 9: 145-166.
- Janssen A. W. 2018. Notes on the systematics, morphology and biostratigraphy of holoplanktonic Mollusca, 25 (1). Once more: the correct name for the type species of the genus *Creseis* Rang, 1828 (Pteropoda, Euthecosomata, Creseidae). *Basteria* 82: 110-112.
- Janssen A.W., Bush S.L., Bednaršek N. 2019. The shelled pteropods of the northeast Pacific Ocean (Mollusca: Heterobranchia, Pteropoda). *Zoosymposia* 13: 305-346. <https://doi.org/10.11646/zoosymposia.13.1.22>
- Karakas F., Wingate J., Blanco-Bercial L. et al. 2020. Swimming and Sinking Behavior of Warm Water Pelagic Snails. *Front. Mar. Sci.* 7:749. <https://doi.org/10.3389/fmars.2020.556239>
- Klussmann-Kolb A., Dinapoli A. 2006. Systematic position of the pelagic Thecosomata and Gymnosomata within Opisthobranchia (Mollusca, Gastropoda)-revival of the Pteropoda. *J. Zool. Syst. Evol. Res.* 44: 118-129. <https://doi.org/10.1111/j.1439-0469.2006.00351.x>
- Kohnert P.C., Cerwenka A.F., Brandt A., Schrödl M. 2020. Pteropods from the Kuril-Kamchatka Trench and the sea of Okhotsk (Euopisthobranchia; Gastropoda). *Prog. Oceanogr.* 181:102259. <https://doi.org/10.1016/j.pocan.2019.102259>
- Lalli C.M., Gilmer R.W. 1989. Pelagic snails: the biology of holoplanktonic gastropod mollusks. Palo Alto, Stanford Univ. Press. <https://doi.org/10.1515/9781503623088>
- Leray M., Yang J.Y., Meyer C.P., et al. 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* 10: 1-14. <https://doi.org/10.1186/1742-9994-10-34>
- Makiola A., Compson Z.G., Baird D.J., et al. 2020. Key questions for next-generation biomonitoring. *Front. Environ. Sci.* 7: 197. <https://doi.org/10.3389/fenvs.2019.00197>
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17:10-12. <https://doi.org/10.14806/ej.17.1.200>
- Manno C., Morata N., Primicerio R. 2012. *Limacina retroversa*'s response to combined effects of ocean acidification and sea water freshening. *Estuar. Coast. Shelf Sci.* 113: 163-171. <https://doi.org/10.1016/j.ecss.2012.07.019>
- Morinière J., Cancian de Araujo B., Lam A. W., et al. 2016. Species identification in malaise trap samples by DNA barcoding based on NGS technologies and a scoring matrix. *PLoS ONE* 11: e0155497. <https://doi.org/10.1371/journal.pone.0155497>
- Mioduchowska M., Czyż M.J., Goldyn B., et al. 2018. Instances of erroneous DNA barcoding of metazoan invertebrates: Are universal cox1 gene primers too "universal"? *PLoS ONE.* 13: e0199609. <https://doi.org/10.1371/journal.pone.0199609>
- Ondov B. D., Bergman N. H., Phillippy A. M. 2011. Interactive metagenomic visualization in a Web browser. *BMC Bioinform.* 12: 1-10. <https://doi.org/10.1186/1471-2105-12-385>
- Porter T. M., Hajibabaei M. 2018. Over 2.5 million COI sequences in GenBank and growing. *PLoS ONE* 13: e0200177. <https://doi.org/10.1371/journal.pone.0200177>
- Puillandre N., Lambert A., Brouillet S., et al. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol. Resour.* 21: 1864-1877. <https://doi.org/10.1111/j.1365-294X.2011.05239.x>
- Rampal J. 1985. Systématique du genre *Creseis* (Mollusques, Thécosomes), Rapport de la Commission Internationale pour l'Exploration Scientifique de la Mer Méditerranée. *Bull. Comm. Int. Explor. Sci. Mer Mediter.* 29: 259-263.
- Rampal J. 2002. Biodiversité et biogéographie chez les Cavoliniidae (Mollusca, Gastropoda, Opisthobranchia, Euthecosomata). *Régions faunistiques marines, Zoosystema* 24 :209-258.
- Ratnasingham S., Hebert P. D. 2013. A DNA-based registry for all animal species: The Barcode Index Number (BIN) system. *PLoS ONE* 8: e66213. <https://doi.org/10.1371/journal.pone.0066213>
- Rognes T., Flouri T., Nichols B., et al. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4: e2584. <https://doi.org/10.7717/peerj.2584>
- Stromek L., Lasota R., Szymelfenig M., Wolowicz M. 2015. Genetic evidence for the existence of two species of the "bipolar" pelagic mollusk *Clione limacinae*. *Am. Malacol. Bull.* 33: 118-120. <https://doi.org/10.4003/006.033.0108>
- Wang S., Yan Z., Hänfling B., et al. 2021. Methodology of fish eDNA and its applications in ecology and environment. *Sci. Total Environ.* 755: 142622. <https://doi.org/10.1016/j.scitotenv.2020.142622>
- Weldrick C. K., Trebilco R., Davies D. M., Swadling K. M. 2019. Trophodynamics of Southern Ocean pteropods on the southern Kerguelen Plateau. *Ecol. Evol.* 9: 8119-8132. <https://doi.org/10.1002/ece3.5380>
- Yamazaki T., Kuwahara T. 2017. A new species of *Clione* distinguished from sympatric *Clione limacina* (Gastropoda: Gymnosomata) in the southern Okhotsk Sea, Japan, with remarks on the taxonomy of the genus. *J. Molluscan Stud.* 83: 19-26. <https://doi.org/10.1093/mollus/eyw032>

SUPPLEMENTARY MATERIAL

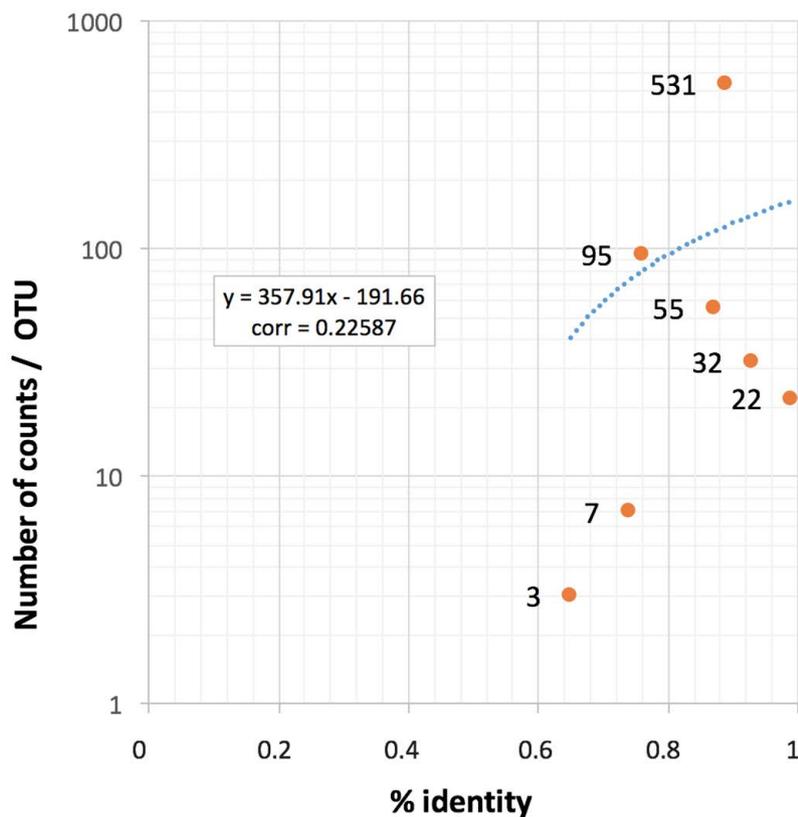


Fig. S1. – Correlation estimation between number of counts per OTU and percentage of species assignment; here, *C. virgula*.

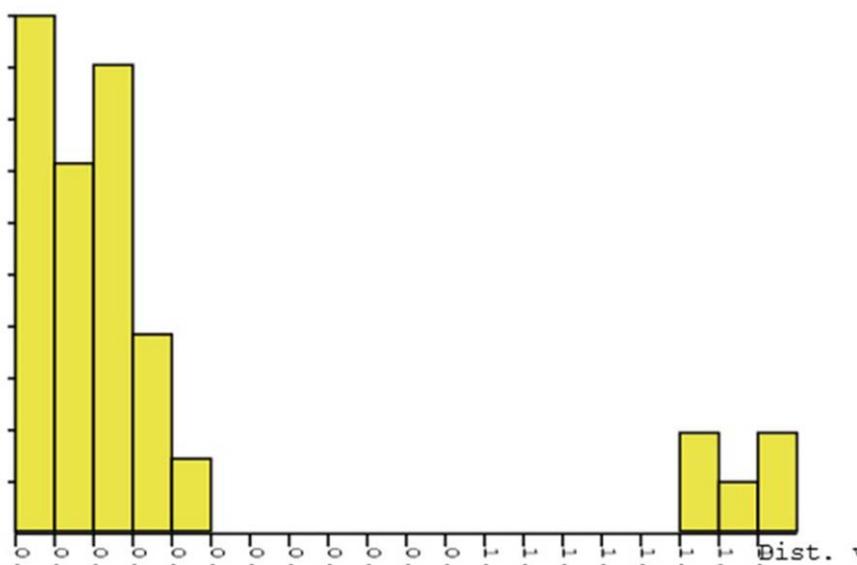


Fig. S2. – ABGD Histogram Cavolinidae (Cavolinid sequences): Group [ 1] n, 1; id, OTU\_1636size27. Group [ 2] n, 1; id, OTU\_12743size2. Group [ 3] n, 11; id, OTU\_12626size14 OTU\_12747size2 OTU\_502size124 OTU\_1314size133 OTU\_3752size6 OTU\_13187size2 OTU\_12776size2 OTU\_2234size12 OTU\_3131size15 OTU\_5297size15 OTU\_91size326.

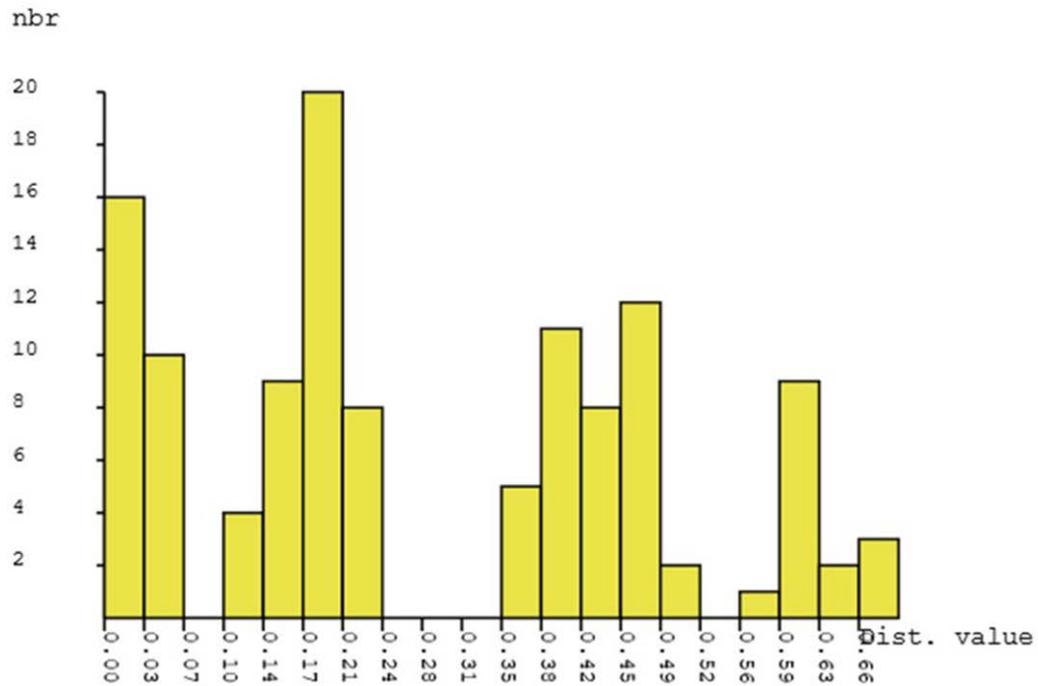


Fig. S3. – ABGD histogram gymnosomes (pteropod sequences): Group [4] n, 2; id, OTU\_12723size6 OTU\_440size167.

Table S1a. – Pteropod BOLD BLAST (BIN sharing) / NCBI GenBank BLAST [sum raw reads in sample (after filtering for OTUs with less than 0.01% reads per sample)]

	NCBI_nt	Species	%_identity2	BIN sharing?	HIT%IDrange
OTU_12723; size=6	KC774091	<i>Gymnosomata_sp.</i>	0.932	No	93.2%_to_93.9%
OTU_440; size=167	KC774091	<i>Gymnosomata_sp.</i>	0.939	No	93.2%_to_93.9%
OTU_1636; size=27	MF048913	<i>Cavolinia_inflexa</i>	1	No	100%_to_100%
OTU_12743; size=2	KC774054	<i>Creseis_acicula</i>	0.961	No	96.1%_to_96.1%
OTU_12626; size=14	KC774054	<i>Creseis_acicula</i>	0.97	No	97%_to_98.5%
OTU_12747; size=2	KC774054	<i>Creseis_acicula</i>	0.97	No	97%_to_98.5%
OTU_502; size=124	KC774054	<i>Creseis_acicula</i>	0.974	No	97%_to_98.5%
OTU_1314; size=133	KC774054	<i>Creseis_acicula</i>	0.985	No	97%_to_98.5%
OTU_3752; size=6	KC774047	<i>Creseis_virgula</i>	0.985	Yes	98.5%_to_99.2%
OTU_13187; size=2	FJ876889	<i>Creseis_virgula</i>	0.909	Yes	90.9%_to_90.9%
OTU_12776; size=2	HM385051	<i>Creseis_virgula</i>	0.936	No	95.5%_to_97%
OTU_2234; size=12	KC774047	<i>Creseis_virgula</i>	0.951	No	95.5%_to_97%
OTU_3131; size=15	HM385051	<i>Creseis_virgula</i>	0.951	No	95.5%_to_97%
OTU_5297; size=15	KC774047	<i>Creseis_virgula</i>	0.947	No	95.5%_to_97%
OTU_91; size=326	KC774047	<i>Creseis_virgula</i>	0.951	No	95.5%_to_97%

Table S1b. – Abnormalities observed: Stylommatophora (terrestrial slugs and snails).

Pulmonata	Bradybaenidae	<i>Trishoplita_cretacea</i>
Sorbeoconcha (Caenogastropoda s.o)	Potamididae	<i>Cerithidea_anticipata</i>
Sorbeoconcha (Caenogastropoda s.o)	Thiaridae	<i>Brotia_episcopalis</i>
Stylommatophora	Arionidae	<i>Arion_ater</i>
Stylommatophora	Arionidae	<i>Arion_rufus</i>
Stylommatophora	Arionidae	<i>Geomalacus_maculosus</i>
Stylommatophora	Arionidae	<i>Geomalacus_maculosus</i>
Stylommatophora	Clausiliidae	<i>Cochlodina_laminata</i>
Stylommatophora	Limacidae	<i>Lehmannia_marginata</i>
Stylommatophora	Vitrinidae	<i>Vitrina_angelicae</i>
Vetigastropoda	Haliotidae	<i>Discus_rotundatus</i>

Table S2. – Collection jars. Sampling locations. Sampling gear.

CODE	LEG	STAT	DATE	LATITUDE	LONGITUDE	SAMPLER
MH005N003S011GGE1	1	3	19/12/2010	29°41'00.0"N	017°17'28.0"W	Neuston net
MH010N008S011GGE1	1	8	24/12/2010	20°15'67.0"N	024°15'07.0"W	Neuston net
MH011N009S011GGE1	1	9	25/12/2010	16°09'84.0"N	026°01'53.0"W	Neuston net
MH012M010C041DGE1	1	10	26/12/2010	14°31'18.0"N	026°00'02.0"W	Multinet
MH012M010C101DGE1	1	10	26/12/2010	14°31'18.0"N	026°00'02.0"W	Multinet
MH012M010C101DGE2	1	10	26/12/2010	14°31'18.0"N	026°00'02.0"W	Multinet
MH012M010C211DGE1	1	10	26/12/2010	14°31'18.0"N	026°00'02.0"W	Multinet
MH012N010S011GGE1	1	10	26/12/2010	14°31'18.0"N	026°00'02.0"W	Neuston net
MH013N011S011GGE1	1	11	27/12/2010	12°29'90.0"N	025°59'17.0"W	Neuston net
MH013N011S012GGE1	1	11	27/12/2010	12°29'90.0"N	025°59'17.0"W	Neuston net
MH014N012S011GGE1	1	12	28/12/2010	09°33'82.0"N	025°59'60.0"W	Neuston net
MH014N012S012GGE1	1	12	28/12/2010	09°33'82.0"N	025°59'60.0"W	Neuston net
MH065N049S011GGE1	3	49	17/02/2011	33°54'43.0"S	037°02'53.0"E	Neuston net
MH065N049S012GGE1	3	49	17/02/2011	33°54'43.0"S	037°02'53.0"E	Neuston net
MH072N052S011GGE1	3	52	24/02/2011	30°03'30.0"S	061°25'84.0"E	Neuston net
MH072N052S012GGE1	3	52	24/02/2011	30°03'30.0"S	061°25'84.0"E	Neuston net
MH074N054S011GGE1	3	54	26/02/2011	28°07'65.0"S	066°29'59.0"E	Neuston net
MH074N054S012GGE1	3	54	26/02/2011	28°07'65.0"S	066°29'59.0"E	Neuston net
MH076N056S011GGE1	3	56	28/02/2011	29°33'63.0"S	072°26'65.0"E	Neuston net
MH076N056S012GGE1	3	56	28/02/2011	29°33'63.0"S	072°26'65.0"E	Neuston net
MH078N058S012GGE1	3	58	02/03/2011	29°49'65.0"S	079°36'66.0"E	Neuston net
MH080M060C101DGE1	3	60	04/03/2011	29°44'93.0"S	086°15'39.0"E	Multinet
MH080M060C221DGE1	3	60	04/03/2011	29°44'93.0"S	086°15'39.0"E	Multinet
MH080N060S011GGE1	3	60	04/03/2011	29°44'93.0"S	086°15'39.0"E	Neuston net
MH080N060S012GGE1	3	60	04/03/2011	29°44'93.0"S	086°15'39.0"E	Neuston net
MH082N062S011GGE1	1	62	06/03/2011	29°37'61.0"S	092°59'05.0"E	Neuston net
MH084N064S012GGE1	3	64	08/03/2011	30°19'96.0"S	103°18'45.0"E	Neuston net
MH190N129S011GGE1	7	129	22/06/2011	15°04'11.0"N	069°17'72.0"W	Neuston net
MH191N130S012GGE1	7	130	23/06/2011	15°31'50.0"N	067°00'86.0"W	Neuston net

Table S3. – Gastropod bins identified in dreg of processed bulk samples with %identity and OTUs with cluster size

bin_uri	%_identity	Seq_length	order_name	species_name	OTU;cluster_size
BOLD:AAD2596	0.759	228	Basommatophora	<i>Bulinus truncatus</i>	OTU_5390; size=6
BOLD:ACQ2738	0.832	119	Basommatophora	<i>Gyraulus</i> sp. _15911	OTU_313; size=106
BOLD:ADR7487	0.74	192	Caenogastropoda	<i>Bittium reticulatum</i>	OTU_12183; size=4
BOLD:ACQ5099	0.83	100	Cephalaspidea	<i>Chelidonura sandrana</i>	OTU_12754; size=2
BOLD:ACI0947	0.932	265	Gymnosomata	Not determined	OTU_12723; size=6
BOLD:ACI0947	0.939	264	Gymnosomata	Not determined	OTU_440; size=167
BOLD:ADR2784	0.81	232	Hygrophila	<i>Galba truncatula</i>	OTU_13162; size=2
BOLD:ACT8286	0.996	266	Lepetellida	Not determined	OTU_2458; size=17
BOLD:ADH1065	0.992	266	Littorinimorpha	<i>Atlanta helicinoidea</i>	OTU_2089; size=13
BOLD:ACQ5602	0.985	264	Littorinimorpha	<i>Atlanta meteori</i>	OTU_2088; size=21
BOLD:ACZ0738	0.936	264	Littorinimorpha	<i>Atlanta selvagensis</i>	OTU_1833; size=13
BOLD:ADK8167	0.802	268	Littorinimorpha	<i>Pseudamnicola moussoni</i>	OTU_1937; size=17
BOLD:AAM3343	1	266	Pteropoda	<i>Cavolinia inflexa</i>	OTU_1636; size=27
BOLD:ACH6682	0.961	181	Pteropoda	<i>Creseis acicula</i>	OTU_12743; size=2
BOLD:ACZ1440	0.97	266	Pteropoda	<i>Creseis clava</i>	OTU_12626; size=14
BOLD:ACZ1440	0.974	266	Pteropoda	<i>Creseis clava</i>	OTU_12747; size=2
BOLD:ACZ1440	0.977	266	Pteropoda	<i>Creseis clava</i>	OTU_502; size=124
BOLD:ACZ1440	0.985	266	Pteropoda	<i>Creseis clava</i>	OTU_1314; size=133
BOLD:AAE9544	0.985	264	Pteropoda	<i>Creseis virgula</i>	OTU_3752; size=6
BOLD:AAC6583	0.909	264	Pteropoda	<i>Creseis virgula</i>	OTU_13187; size=2
BOLD:ACV0071	0.955	264	Pteropoda	Not determined	OTU_12776; size=2
BOLD:ACV0071	0.958	264	Pteropoda	Not determined	OTU_2234; size=12
BOLD:ACV0071	0.97	264	Pteropoda	Not determined	OTU_3131; size=15
BOLD:ACV0071	0.97	264	Pteropoda	Not determined	OTU_5297; size=15
BOLD:ACV0071	0.97	264	Pteropoda	Not determined	OTU_91; size=326
BOLD:ACV8715	0.787	197	Pulmonata	<i>Trishoplita cretacea</i>	OTU_5200; size=15
BOLD:AAF0822	0.769	221	Sorbeoconcha	<i>Cerithidea anticipata</i>	OTU_13236; size=2
BOLD:AAF6997	0.821	224	Sorbeoconcha	<i>Brotia episcopalis</i>	OTU_2842; size=5
BOLD:AAD2027	1	266	Stylommatophora	<i>Arion ater</i>	OTU_506; size=214
BOLD:AAE6652	1	266	Stylommatophora	<i>Arion rufus</i>	OTU_1508; size=17
BOLD:ACW0753	0.837	104	Stylommatophora	<i>Geomalacus maculosus</i>	OTU_13177; size=6
BOLD:ACW0753	0.837	104	Stylommatophora	<i>Geomalacus maculosus</i>	OTU_1729; size=40
	0.996	266	Stylommatophora	<i>Cochlodina laminata</i>	OTU_3587; size=6
BOLD:AAF1156	1	266	Stylommatophora	<i>Lehmannia marginata</i>	OTU_478; size=265
BOLD:AAN0223	0.996	266	Stylommatophora	<i>Vitrina angelicae</i>	OTU_3969; size=7
BOLD:AAI9791	0.989	266	Vetigastropoda	<i>Discus rotundatus</i>	OTU_1429; size=25