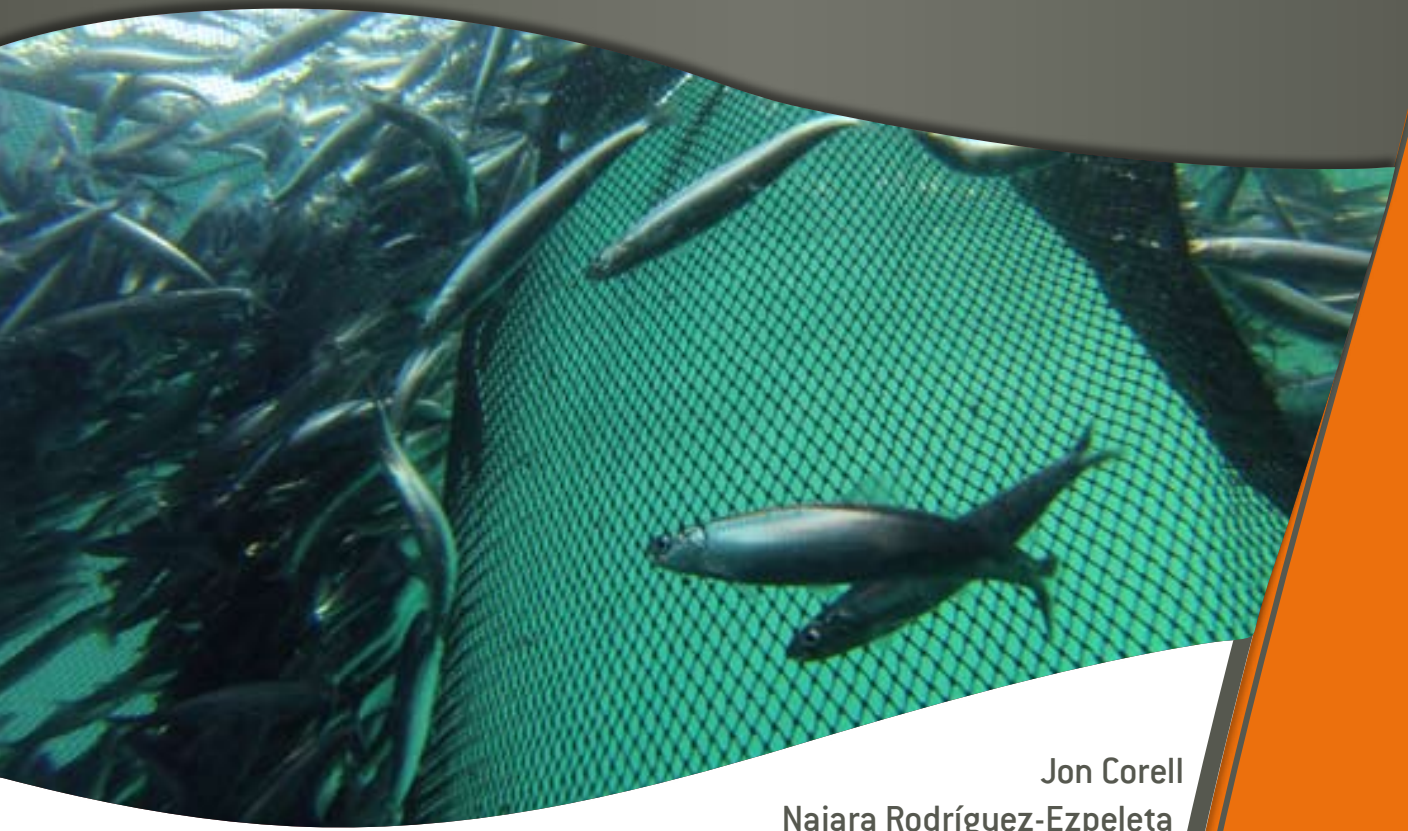
azti tecnalia

21 (2)

Tuning of protocols and marker selection to evaluate the diversity of zooplankton using metabarcoding



Jon Corell
Naiara Rodríguez-Ezpeleta

Revista de Investigación Marina
2013

Corell, J and Rodríguez-Ezpeleta, N, 2014. Tuning of protocols and marker selection to evaluate the diversity of zooplankton using metabarcoding . Revista de Investigación Marina, AZTI-Tecnalia, 21(2): 19-39

Edición: 1.ª Mayo 2014
© AZTI-Tecnalia
ISSN: 1988-818X
Unidad de Investigación Marina
Internet: www.azti.es
Edita: Unidad de Investigación Marina de Tecnalia
Herrera Kaia, Portualdea
20110 Pasaia
Foto portada: © Iñigo Onandia

Tuning of protocols and marker selection to evaluate the diversity of zooplankton using metabarcoding

Jon Corell[1]* and Naiara Rodríguez-Ezpeleta[1]

Abstract

DNA metabarcoding consists on taxonomic assignment of individuals from an environmental sample based on their DNA sequences. In recent years, this technique has enormously evolved due to the increasing effort in surveys and to the advent of new sequencing technologies. The success of the metabarcoding method largely depends on the efficiency of the protocol for extracting DNA from complex samples and on the genetic marker and primer pairs used for taxonomic identification. Here, we have experimentally evaluated alternative protocols for DNA extraction from zooplankton complex samples, and have theoretically tested the suitability of previously published primer pairs for barcoding of zooplankton. The experimental analyses show that the SDS-chloroform method is the most suitable for extracting DNA from mixed zooplankton samples. Additionally, the *in silico* analyses show that no primer pair is sufficient to amplify one of the most commonly used markers for barcoding, the gene codifying for the cytochrome c oxidase subunit I (COI), but that there are a set of primers that seem to capture the whole zooplankton diversity based on the small subunit ribosomal RNA (18S rRNA) gene. Our results have implications for future zooplankton metabarcoding projects, for which a prior knowledge of the best DNA extraction method and best performing primers is necessary.

Key words: 18S rRNA, COI, DNA extraction, metabarcoding, PCR, primers, zooplankton

Resumen

El *metabarcoding* consiste en asignar taxonómicamente individuos de una muestra medioambiental basándose en sus secuencias del ADN. El aumento del esfuerzo por descubrir la verdadera y oculta biodiversidad que se lleva a cabo en diferentes campañas, junto con la creación de nuevas tecnologías de secuenciación han hecho que el *metabarcoding* haya aumentado considerablemente en los últimos años. El éxito de dicha técnica depende sobre todo de la eficiencia del protocolo de extracción del ADN, así como también de la idoneidad de los marcadores genéticos y parejas de cebadores usados para la identificación taxonómica. En este estudio, hemos evaluado experimentalmente diferentes protocolos de extracción del ADN de muestras complejas de zooplancton y hemos testado teóricamente la idoneidad de varias parejas de cebadores previamente publicados para el *barcoding* de zooplancton. Nuestros análisis experimentales demuestran que el método más eficaz para extraer ADN de muestras complejas de zooplancton es el SDS-cloroformo. Además, nuestro análisis teórico muestra que ninguna pareja de cebadores es suficiente para amplificar uno de los marcadores más comúnmente utilizados para el *metabarcoding*, el gen de la subunidad I de la citocromo c oxidasa (COI). Sin embargo, hay varios cebadores basados en el gen de la pequeña subunidad ribosómica (18S rARN) que parecen capturar toda la diversidad. Nuestros resultados tienen implicaciones para futuros estudios de *metabarcoding* de zooplancton, para los cuales es necesario un previo conocimiento del mejor método de extracción del ADN y de los cebadores que permitan la mayor, o deseada, cobertura taxonómica posible.

Palabras clave: 18S rRNA, COI, extracción del DNA, metabarcoding, PCR, cebadores, zooplancton

[1] *AZTI-Tecnalia; Marine Research Division; Txatxarramendi irla z/g; 48395 Sukarrieta; Bizkaia, Spain*

** Corresponding author: jcorell@azti.es*

Introduction

Oceans are dominated in abundance and biomass by planktonic organisms (Machida *et al.*, 2009), which are small organisms that drift along the water masses. Among the most important ones, zooplankton are of particular interest, as they play a key role in ocean functioning (Carlotti *et al.*, 2000) for being downstream and upstream regulators, and also they are considered potential markers of climate change (Thomas *et al.*, 2004). The study of the distribution of zooplankton is therefore, crucial to better understand oceanic ecosystems and anticipate the effects of climate change.

Studying the distribution of zooplankton requires taxonomic identification of small organisms, traditionally performed by expert taxonomists employing optical techniques. This method is time consuming and in some cases there are limitations such as those that appear while studying cryptic species (morphologically indistinguishable species), early developmental stages (eggs and larvae), parts of specimen bodies (e.g. one leg) or semi-digested samples (e.g. gut contents) (Lindeque *et al.*, 2013). Genetic techniques have the potential of overcoming these limitations while providing cheaper, faster and more accurate taxonomic identification (Nowrousian, 2010). The metabarcoding is widely used in prokaryotes, for which standard protocols for DNA extraction and universal primers targeting the small subunit ribosomal RNA gene (16S rRNA) are well defined (see for example the Earth Microbiome Project: http://www.earthmicrobiome.org). The application of this technique to eukaryotic organisms is less straight forward as no standards yet exist. Several advances have been made in microbial eukaryotes (Creer & Sinniger, 2012) but this technique is still under experimental and analytical development for multicellular organisms such as zooplankton for assessing diversity studies. Yet, the potential of this technique to assist routine taxonomic identification of zooplankton samples collected during oceanographic campaigns has been acknowledged (Machida *et al.*, 2009).

The DNA barcoding method (Figure 1) consists of taxonomic assignment of a specimen (or part of it) by sequencing a standardized short DNA fragment (barcode) that is unique to each species and therefore used for species recognition and discrimination (Bucklin *et al.*, 2010b). Two short DNA sequences flanking the barcode (primers) are used to amplify it from the genomic DNA in a process named Polymerase Chain Reaction (PCR) (Green & Sambrook, 2012). If the process is performed, not on a single individual or a group of individuals from the same species, but on an environmental sample (filtered water, sediment, gut content…) containing a mix of individuals from different species, we name it metabarcoding (Bourlat *et al.*, 2013). In this case, DNA is extracted from the whole sample and the PCR does not amplify a single product, but a mixture of products corresponding to each of the species in the sample (Figure 1). The products obtained from DNA barcoding are generally sequenced using the Sanger method (Sanger *et al.*, 1977), whereas the mixed products obtained from metabarcoding are generally sequenced using high throughput sequencing technologies (Illumina, 454, etc) which allow to directly sequence mixed products (Nowrousian, 2010). Ultimately, both techniques rely on a previously built reference database that contains the correspondence between the barcodes and the species (Bourlat *et al.*, 2013; Hebert *et al.*, 2003).

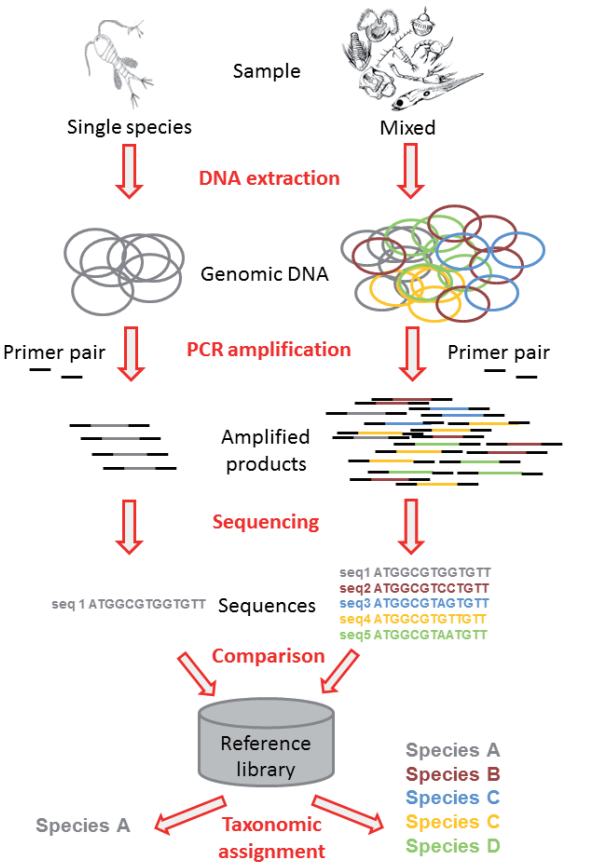


Figure 1. Schematic representation of the processes of DNA barcoding (left) and metabarcoding (right). Coloured circles represent extracted genomic DNA, which is composed of multiple copies of the same genome (barcoding) or of multiple copies of the genomes of the species composing the samples (metabarcoding). Amplified products are identical in barcoding, whereas a mixture of amplified products from the different genomes is obtained in metabarcoding. Once the amplified products have been sequenced, taxonomic assignment is preformed based on comparison of the obtained sequences with a reference database.

Being able to identify the species present in a sample by bulk sequencing the DNA, the metabarcoding approach is a promising avenue in biodiversity studies. Yet, several experimental and analytical issues need to be assessed in order to evaluate the suitability of this technique in zooplankton research. First, the identification of a suitable method to extract DNA from a complex sample is needed. The method of choice should account for the heterogeneity of organismal types (gelatinous organisms,

exoesqueletal structures, etc), yield enough DNA of good integrity and purity, and finally, remove compounds that could possibly inhibit subsequent reactions such as PCR (Wilson, 1997). Second, metabarcoding requires the identification of suitable barcodes which should be composed by an enough variable region (for species discrimination) flanked by conserved regions (for proper primer binding) (Bucklin *et al.*, 2011).

For eukaryotic organisms such as zooplankton, there are mainly two typically targeted genes from which barcode sequences are used: the mitochondrial cytochrome c oxidase subunit I (COI) (Ortman, 2008) and the nuclear small-subunit ribosomal RNA (18S rRNA) gene (Machida & Knowlton, 2012). COI is one of the most used genes for specifically barcoding metazoans, while 18S rRNA has been traditionally used for barcoding numerous eukaryotic taxonomic groups (Bucklin *et al.*, 2011). There have been more COI over 18S rRNA marine metazoan sequences generated, as COI sequences are significantly more distant between different species than between individuals from the same species (Bucklin *et al.*, 2010b; Bucklin *et al.*, 2011; Meyer & Paulay, 2005). This feature is the barcoding gap: a threshold defined as ten times the mean of the intraspecific variation of the group under study, compared with the interspecific variation (Hebert *et al.*, 2004a). It allows to distinguish between species through barcodes (Meyer & Paulay, 2005).

Within each barcode, a wide variety of primer options are possible depending on the targeted taxonomical group (Evans & Paulay, 2012), hence being extremely difficult to find a unique primer pair that allows amplification of a variable region in all species in the sample. This implies that the primers are mainly designed for specific groups, which makes results difficult to compare. To make them complementary to more than one sequence and therefore cover a higher range of taxonomic groups, primers can be degenerated (contain more than one possible base at a given position) (Meyer & Paulay, 2005) or include artificial nucleotides, such as the inosine, that bind to any of the four natural nucleotides (Geller *et al.*, 2013).

In this study, we used zooplankton samples collected during the Malaspina 2010 circumnavigation to (i) identify a suitable method to extract DNA from mixed samples of marine zooplankton and to (ii) assess the adequacy of published primer pairs for COI and 18S rRNA genes to study zooplankton diversity.

Material and Methods

Sample selection

29 mesozooplankton samples from the Malaspina expedition covering different depths and locations around the globe were used to tune up the DNA extraction.

DNA extraction

Aliquots of 5 or 25 ml of the total sample (50 ml) were used for DNA extraction. In order to remove the ethanol in which the samples are stored, samples were centrifuged at 3,500 *g* for 10 min and separated from ethanol either by inverting the tube, when the pellet was firmly attached to the bottom, or by filtration (pore diameter of 20 µm), when the pellet was unstable. Three different extraction methods were used (detailed below): the Sodium dodecyl sulphate (SDS)-chloroform method (Green & Sambrook, 2012; Manaffar *et al.*, 2010), the *Wizard Genomic DNA Purification Kit* (Promega Corporation, Wisconsin, USA) and the *QIAamp DNA Minikit* (QIAgen, Venlo, Netherlands). DNAs from all samples have been extracted with at least one of the protocols (Table S1).

For the SDS-chloroform method, samples were grinded with a mortar in 1 ml SDS buffer (Tris-HCl (10 mM), EDTA (100 mM, pH 8.0), NaCl (200 mM), SDS 1 %) until no integer organism could be appreciated. The buffer breaks the cells and the grinding facilitates the process. Proteins where digested through an overnight incubation with proteinase K (0.2 mg/ml, final concentration) at 55–56 ºC. After centrifugation at 3,500 *g* for 15 min, the supernatant, which contains the released nucleic acids (DNA and RNA), was transferred to 2 ml sterile Eppendorf tubes (1 ml per Eppendorf). Each aliquot was gently mixed with half volume of phenol-chloroform-isoamyl alcohol (25:24:1, vol:vol:vol) and centrifuged at 12,000 *g* for 10 min. From the three phases in which the sample was divided after centrifugation, the aqueous less dense superior phase containing the nucleic acids was transferred into a new tube. Remaining phenol traces were removed by mixing the sample with an equal volume of chloroform-isoamyl alcohol (24:1, vol:vol), centrifuging at 12,000 *g* for 10 min, and transferring the supernatant to a new tube. The final supernatant was mixed with two times its volume of ethanol 95 %, ammonium acetate 0.5 M, gently shooked and stored at - 20 ºC for at least 1 h. After centrifugation at 12,000 *g* for 20 min, the ethanol was removed with a pipette, leaving the white pellet containing the nucleic acids in the tube. Remaining salts were removed by adding 150 µl of ethanol 80 %, centrifuging 10 min at 12,000 *g* and removing the ethanol. Once all traces of ethanol were evaporated, the pellet was dissolved in 100 µl of Milli-Q water and stored at - 20 ºC until further analyses. After DNA extraction, some samples (Table S1) were incubated with RNAse (0.5%) at 37 ºC during 30 min or passed through the *PowerClean DNA Clean-Up Kit* (MO BIO Laboratories Inc., California, USA) following manufacturer's instructions, to eliminate or at least reduce the degraded RNA.

For the *Wizard Genomic DNA Purification kit* and the *QIAamp DNA Minikit* manufacturer´s instructions for the protocols "Isolating Genomic DNA from Tissue Culture Cells and Animal Tissue: Animal Tissue (Mouse Tail)" and "QIAamp DNA Minikit: DNA purification from tissues" were respectively followed, except for some modifications: samples were grinded with a mortar in the lysis buffer solution provided by the kit to ensure full digestion, the RNAse treatment was not applied, and final DNA pellet was dissolved in Milli-Q water instead of in the provided solutions. All exceptions were made to allow direct comparison with the SDS-chloroform method and avoid compounds (e.g. EDTA, which is present in most commercial elution buffers) that could inhibit further reactions.

DNA quantity, purity and integrity check

The DNA concentration and purity were measured with the ND-1000 Nanodrop (Thermo Scientific, Massachusetts, USA). DNA purity was inferred from 260/280 and 260/230 absorbance ratios as nucleic acids absorb light at 260 nm, proteins at 280 nm and organic compounds at 230 nm (Green & Sambrook, 2012). Good 260/280 ratios range between 1.8 and 2 and good 260/230 ratios, between 1.8 and 2.2 (Thermo-Scientific, 2010a; Thermo-Scientific, 2010b). Integrity of extracted genomic DNA and PCR amplification success were assessed by electrophoresis respectively in 0.7 % and 1.7 % agarose gels (using TAE buffer), stained with 20:1 of Red Safe (iNtRON Biotechnology Inc., Seoul, Korea) colorant. About 100 ng of DNA together with 1:5 loading buffer were charged. 2.5 µl of Perfect DNA™ 1 kbp Ladder (Merck KGaA, Darmstadt, Germany) and GeneRuler 100 bp (Thermo Scientific, Massachusetts, USA) were charged as ladder for DNA extractions and PCR products respectively. Based on the electrophoresis images, numerical values were assigned to each of the samples (0 = full degradation, 1 = medium integrity and 2 = full integrity) (modified from Salgado et al., 2007).

Zooplankton list assembly and genetic sequences retrieval

A list of 2993 zooplankton species was assembled based on the repositories *spp_barcodes*, *zoop_ANT-XXIV_1* and *zoop_RHB0603* of the Census of Marine Zooplankton database (www.cmarz.org/jg/dir/CMarZ/, checked on 13-06-2013), on the Marine Barcode of Life database (www.barcodingmarinelife.org/bc-bin/display_records.active, checked on 13-06-2013) and on Bucklin et al., (2010). The list was confronted against the World Register of Marine Species (WoRMS) (www.marinespecies.org/, checked on 13-06-2013), so that the species with synonyms or old names could be corrected. The species that were not in WoRMS or had ambiguous names in such database were not selected. In order to retrieve gene sequences from 18S rRNA and COI genes from the identified species, the GenBank database (http://www.ncbi.nlm.nih.gov/nuccores, checked on 18-06-2013) was interrogated. After the extraction of the sequences using a bash/shell script, some non-selected species appeared as they share some common characteristics with one of the list (e.g. genera…). Those species were confronted against WoRMS, and the ones that were zooplankton were added to the list. Finally, the species list was again reduced as there were some unverified species (according to the GenBank entry description). Sequences from the same species were clustered at 90% identity threshold with CD-HIT (Niu *et al.*, 2010) to reduce redundancy.

"In silico" primer analysis

67 primer pairs for COI and 31 primer pairs for 18S rRNA selected from the literature (Table S2 and S3) were tested against the retrieved zooplankton sequences. Most primers pairs were selected because they were developed for targeting zooplankton. In 2012 Evans and Paulay Showed that some pairs could also amplify other taxa apart from the groups for which they were originally designed. Moreover, in some cases those primer pairs have even higher amplification percentages in those "new" taxa, than in the original ones. Therefore, other primer pairs targeting alternative groups such as arachnids, fish or even birds were also tested for suitability in marine zooplankton. Some of the retrieved sequences do not cover complete gene length, implying that they cannot be tested for primers that fall outside the region they cover. Therefore, we tested, for each primer pair, only those sequences that cover the gene fragment covered by the corresponding primer pair. The complete COI and 18S rRNA gene sequences from *Mytilus galloprovincialis* (accession number DQ399833.1) and *Aplysia punctata* (accession number AJ224919.1) were used as reference to determine each primer´s start and end position in each gene using BLAST (Altschul *et al.*, 1990). All sequences were then compared to the references using BLAST to determine their start and end position. For each primer pair, only those sequences whose start position is smaller than the start position of the forward primer and whose end position is higher than the end position of the reverse primer were retained, so only those sequences including the primer sequence were maintained. The potential amplification of each primer pair per taxonomic group was tested using Primer Prospector (Walters *et al.*, 2011) with default parameters. When more than one sequence per species remained after clustering with CD-HIT, amplification was considered successful or unsuccessful when all sequences showed a positive or negative result in the amplification test respectively. When different sequences annotated as belonging to the same species provided contradictory results, amplification was considered in general positive as, in most cases, the non-amplified sequences were proven to be wrongly annotated. If obvious reason for the contradiction were observed, amplification was considered negative.

Amplification by polymerase chain reaction

The PCR was done with the kit KAPA Taq DNA Polymerase (KAPABIOSYSTEM, Massachusetts, USA). Amplification was performed on DNA extractions (with all tested methods) from sample 118, except for the SDS-chloroform plus RNAse treatment, as it is assumed that this method performs similarly as the SDS-chloroform. For a 20 µl reaction 4 µl of Kapa buffer (1X), 2 µl of $MgCl_2$ (2.5 mM), 0.4 µl of dNTPs (0.2 mM), 0.4 µl per primer (0.2 µM), 0.1 µl of Taq polymerase (0.5 units/µl) and 100 ng of genomic DNA were added. Two COI primer pairs (LCO-1490/HCO-2198 and dgLCO1490/dgHCO2198) were tested under the following PCR thermal cycling conditions (Meyer, 2003): 3 min at 95 ºC, 35 cycles of 40 s at 95 ºC, 40 s at 45 ºC and 60 s at 72 ºC, and a final elongation of 7 min at 72 ºC.

Results and discussion

DNA extraction

A comparison of the different methods applied in the same samples reveals that good integrity DNA can be obtained with all of them (Table 1 and Figure 2). All protocols were not applied to all samples (Table S1) due to the high cost and the observed worse efficiency of the commercial kits. The SDS-chloroform method yielded a smear around 100 bp, which could correspond to degraded DNA or RNA. Both, RNAse treatment and purification by the cleaning column decreased this smear, suggesting that it corresponds to degraded RNA. Both, the *QIAamp DNA Minikit* and the *Wizard Genomic DNA Purification Kit* are based on a similar column-based system (Promega, 2010; QIAgen, 2010); therefore only results of the *QIAamp DNA Minikit* will be used as representatives for a commercial kit.

All samples extracted with the SDS-chloroform or *QIAamp DNA Minikit* methods yielded enough DNA quantity (> 57 ng/µl) of sufficient purity for subsequent steps (Figure 2 and 3). The purity of the DNA was inferred from the 260/280 and 260/230 ratios. The 260/280 and 230/260 ratios were statistically significantly lower and higher respectively (Student T test, *P*-value < 0.001) for the SDS-chloroform method when compared to the *QIAamp DNA Minikit* method (Figure 3). This suggested that the SDS-chloroform method produces less residual proteins but more salts. Importantly, DNA integrity differs between the two methods, yielding the majority of the extractions performed with the SDS-chloroform method DNA of perfect integrity while those performed with the *QIAamp DNA Minikit* method, yielded DNA of very low integrity.

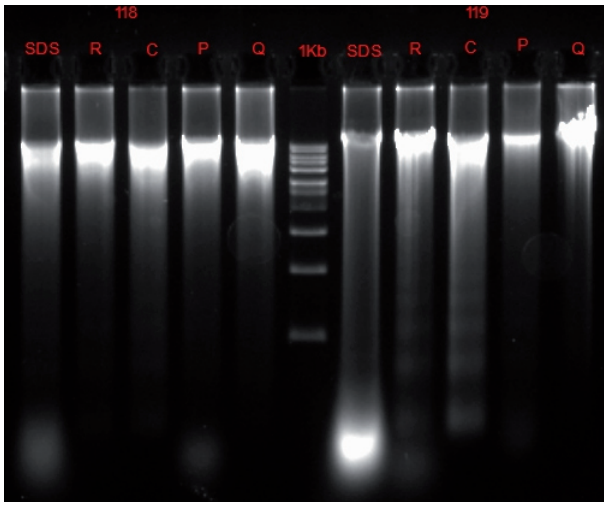


Figure 2. DNA integrity of the two samples (118 and 119) that have been processed with all the different protocols. 1 Kb ladder shows the followings bands: 500; 1,000; 1,500; 2,000; 2,500; 3,000; 4,000; 5,000; 6,000; 8,000 and 10,000 bp. Abbreviations: 1 Kb = Perfect DNA™ 1 kbp Ladder; SDS = SDS-chloroform; R = SDS-chloroform with RNAse treatment; C = SDS-chloroform with column treatment; P = *Wizard Genomic DNA Purification Kit*; Q = *QIAamp DNA Minikit*.

Table 1. Concentration (ng/ul) and purity (260/280 and 260/230 ratios) of DNA extracted from two samples (118 and 119) with different methods. Abbreviations of protocols: SDS: SDS-chloroform; SDS+RNAse: SDS-chloroform plus RNAse treatment; SDS+column: SDS-chloroform plus column treatment; Promega: *Wizard Genomic DNA Purification Kit;* QIAgen: *QIAamp DNA Minikit*.

Sample	Extraction met.	DNA Conc. (ng/µl)	260/280	260/230
118	SDS	88.2	1.86	1.72
	SDS+RNAse	92	1.91	1.75
	SDS+Column	49.4	1.66	2.51
	Promega	92.8	1.86	1.99
	QIAgen	82.8	1.79	2.39
119	SDS	100	1.81	1.92
	SDS+RNAse	102	1.80	2.81
	SDS+Column	84.6	1.88	2.04
	Promega	11.32	1.28	5.92
	QIAgen	63.4	1.74	2.35

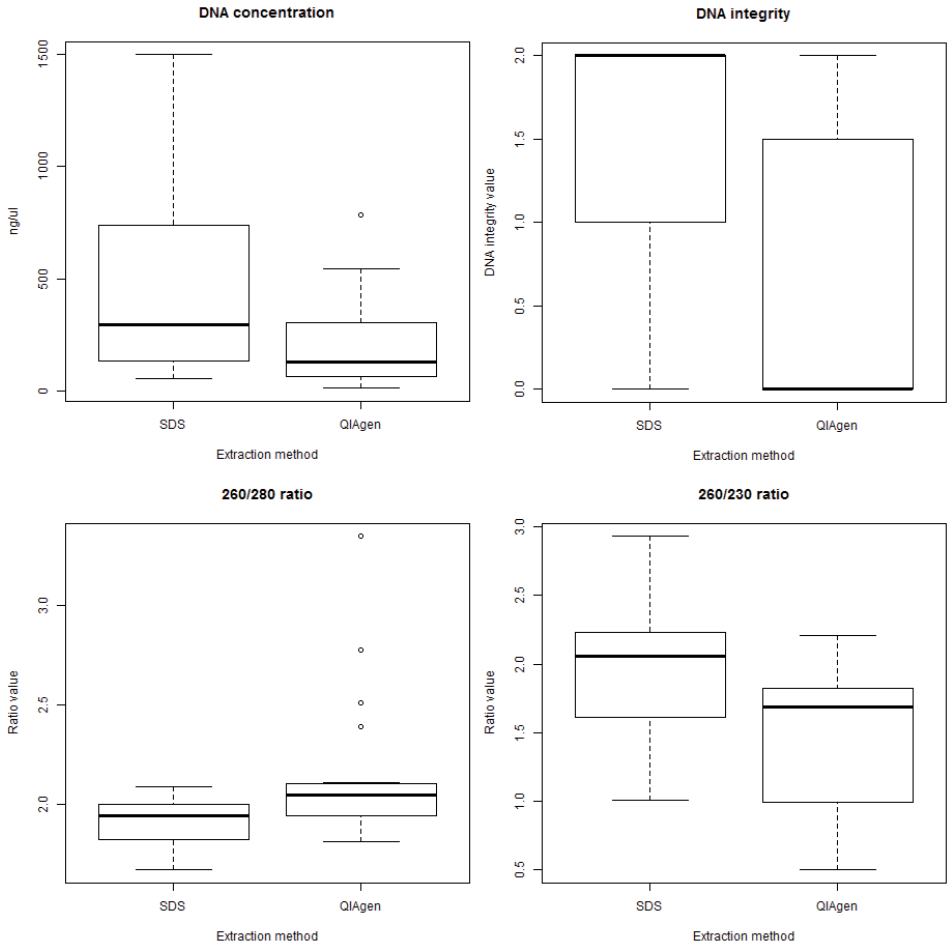


Figure 3. Bloxplots for all DNA concentration, integrity and purity (measured as 230/260 and 280/260 ratios) of the SDS-chloroform (alone and with post-treatment) method (SDS) and the *QIAamp DNA Minikit* (QIAgen). Median (second quartile) is represented as a black line creating the upper third quartile and below first quartile. Standard deviation is represented as whiskers, while outliers as individual points. The four parameters were statistically different between SDS and QIAgen according to Student T test (*P*-value < 0.001).

Unlike home-made protocols, commercial kits have fixed reagent volumes and are optimized for a limited range of input sample volume. It is therefore possible that in our case, the SDS-chloroform method was outperforming the the *QIAamp DNA Minikit* method because of the large volumes of input sample we were using. Indeed, from the 22 samples of 5 ml tested with the *QIAamp DNA Minikit* method, 8 yielded good integrity DNA, whereas none of the 25 ml samples tested yielded good DNA integrity. One possible explanation could be that with so much sample volume the filter clogs avoiding reagents to pass through. Since the aim of this study was to obtain the full representation of the zooplankton species present in our samples, the subsample used for DNA extraction should be large enough so that it represents the content of the total sample. Therefore, in our case the SDS-chloroform method was the most suitable one. Moreover, this method was the least expensive as it costs less than 1 € per sample compared to the 4 € per sample of the *QIAamp DNA Minikit*. In terms of time consumption, the SDS-chloroform was the only one that required more than 6 h, whereas the commercial kits usually take around 2 h (the overnight lysis step is not taken into account in any cases). In summary, for now, the SDS-chloroform plus RNAse treatment is selected as the best performing DNA extraction method for zooplankton samples, although testing commercial kits that are suitable for larger volumes could help reducing the per sample processing time.

Zooplankton species list and sequences

The compiled zooplankton list contains 2993 species, being the great majority chordates (1304 sp, 44 %) and arthropods (1279 sp, 43 %) (Figure 4A). Yet, this list does not represent the real marine zooplankton diversity (Figure 4B) where arthropods, especially calanoid copepods (Bucklin *et al.*, 2010b), followed by molluscs are the main zooplankton groups (Brusca & Brusca, 2003; Lenz, 2000; Postel *et al.*, 2000). This disagreement is due to the fact that our lists were compiled from zooplankton databases including mostly sequenced species.

Molecular databases are largely biased towards economically important species such as fish, which have largely contributed to increase the chordate fraction of our dataset.

From the 2993 species, sequences from 1,975 and 654 were retrieved from GeneBank for COI and 18S rRNA respectively. In both cases, more than one sequence was retrieved for some of the species, being the total number of sequences retrieved 17,201 and 3,133 respectively. After clustering the sequences for each species at 90% similarity with CD-HIT, the number of sequences per species was reduced, but there were still 312 and 328 species with more than one sequence for COI and 18S rRNA respectively. In the former, the maximum number of sequences per species was 28 and in the later, 4. The presence of quite different sequences (> 90% similarity) per species may be explained by the fact that they corresponded to different isolates coming from parts of the world (Evans *et al.*, 2009) or to the poor curation of taxonomic databases which can include sequences from two different species labelled as one.

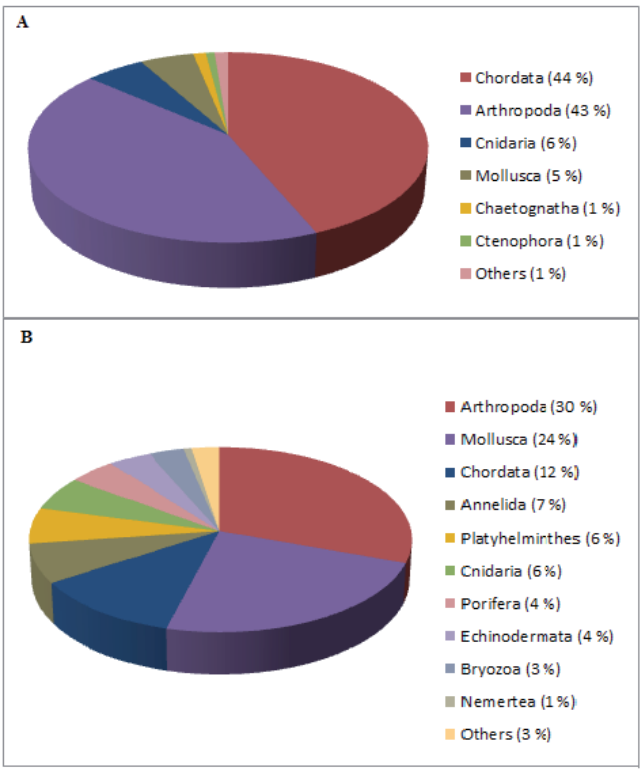


Figure 4. A. Percentage of each phylum in the created zooplankton list. Category *others* includes Annelida (14 sp), Foraminifera (6 sp), Rotifera (6 sp), Nemertea (5 sp), Ciliophora (4 sp), Bryozoa (1 sp) and Myzozoa (1 sp). B. Expected zooplankton diversity according to actual estimations (Brusca & Brusca, 2003; Lenz, 2000)

Primer analysis

Primer selection is a key step in organism identification by metabarcoding. According to the taxonomic groups of interest, different primer combinations are possible. Primer Prospector determines if a given primer pair would successfully amplify a given sequence, that is, if the two primers would bind to the sequence to be amplified. We evaluated the amplification potential of 67 and 31 primer pairs for COI and 18S rRNA respectively on the retrieved sequences. From those only 23 and 24 respectively were able to amplify at least one species (Figure 5 and 6).

Because most of the sequences in the database for COI do not include the binding site for many of the primers under study, the number of species on which each primer could be tested for this gene was limited. Despite this limitation, we can conclude that the best performing primer pairs are Leray_Meyer1 and Leray_Geller (Leray *et al.*, 2013a) as they have the highest rate of species tested versus amplified (Figure 5). Both of these primers pairs combine an internal forward primer (Leray) with a modified widely used reverse primer, Folmer (Folmer *et al.*, 1994). The first primer pair uses a degenerated version of the Folmer primer (Meyer *et al.*, 2005) whilst the second replaces degenerated bases by inosines (Geller *et al.*, 2013) (Table S2).

For 18S rRNA, there are more sequences that cover the primer binding sites (Figure 6) and therefore more sequences can be tested per primer pair. All Machida (Machida & Knowlton, 2012) and Stoeck (Stoeck *et al.*, 2010) primer pairs are the ones that could be tested in more species and also the ones that have the highest amplification proportions. The Machida primers amplified regions of different length of the 18S rRNA gene and have been designed for 24 different phyla (Machida & Knowlton, 2012), and the Stoeck primer pairs amplified two different regions of 18S and have been designed to amplify all kind of organisms (wide range primers) (Dunthorn *et al.*, 2012; Guillou *et al.*, 2012; Stoeck *et al.*, 2010) (Table S3). Other primer pairs, such as Hillis_Halanych (Halanych *et al.*, 1995), also showed relative high proportion of amplified species through many phyla, even though they have been designed for specific taxonomic groups.

In general, when focusing on one specific group, the primer pair to use should be selected based on the amplification percentage of that group. On the other hand, if focussing on the whole community, primers that are not the best performing on a specific group but that capture the whole phyla spectrum might be desirable. Yet, ideally, the combination of wide-range and group-specific primers should be used in order to cover the whole biodiversity of the target groups.

DNA amplification by the Polymerase Chain Reaction

LCO-1490/HCO-2198 (Folmer *et al.*, 1994) and dgLCO1490/dgHCO2198 (Meyer *et al.*, 2005) (Folmer and Meyer respectively in the Primer Prospector analysis) were chosen for their wide coverage and common usage (Geller *et al.*, 2013; Leray *et al.*, 2011; Leray *et al.*, 2013a) to test amplification of COI on the DNA obtained from sample 118 with different protocols (Figure 7). The amplified product appeared as a clear band around 600 bp, which confirms that amplification worked well for both primers (the expected length of the products amplified with these primer pairs is 658 bp).

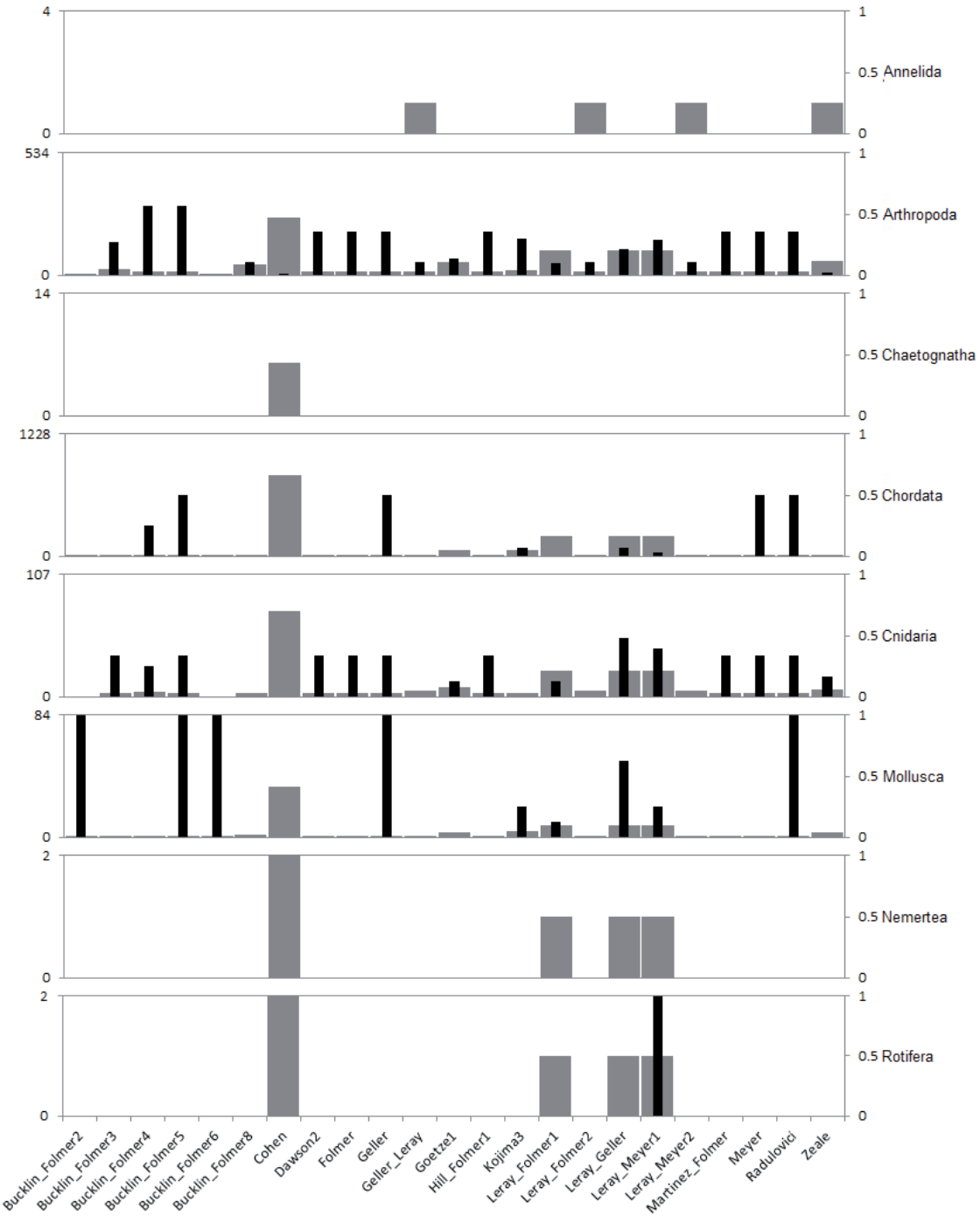


Figure 5. Number of species tested (left y axis, grey bars) and, within them, proportion of potentially amplified (right y axis, black bars) for COI using different primer pairs (x axis) for the most representative phyla. Only primers that showed amplification for at least one species are shown.

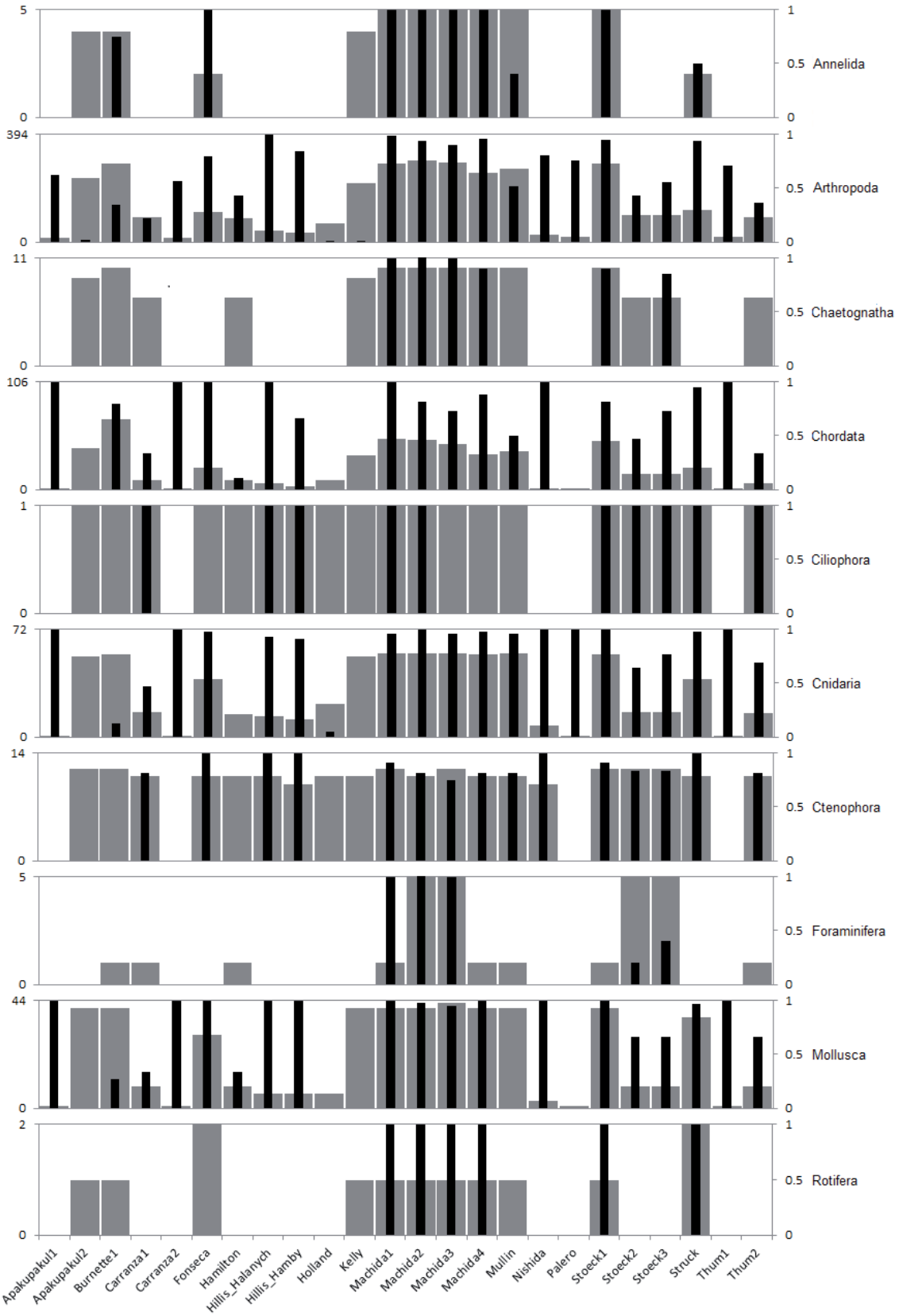


Figure 6. Number of species tested (left y axis, grey bars) and, within them, proportion of potentially amplified (right y axis, black bars) for 18S rRNA using different primer pairs (x axis) for the most representative phyla. Only primers that showed amplification for at least one sequences are shown.

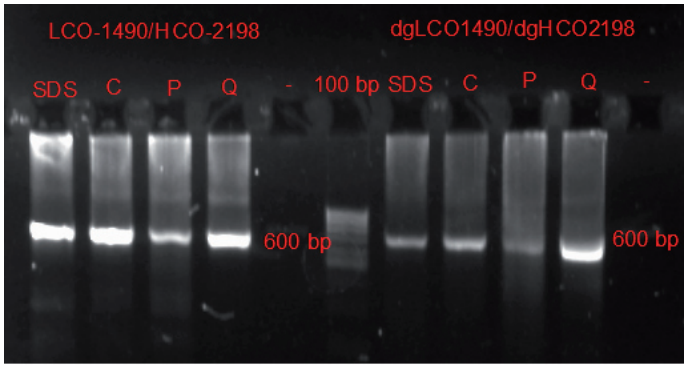


Figure 7. PCR products obtained with 2 different primer pairs (in the left side LCO-1490/HCO-2198 and in the right dgLCO1490/dgHCO2198) for sample 118 extracted using different DNA extraction protocols. The 100 bp ladder shows bands from 100 bp to 1,000 bp every 100 bp. Abbreviations: 100 bp = GeneRuler 100bp ladder; SDS = SDS-chloroform; C = SDS-chloroform with column treatment; P = Wizard Genomic DNA Purification Kit; Q = QIAamp DNA Minikit; - = negative control.

Conclusions

Here, we have addressed two of the requirements needed for a metabarcoding analysis of zooplankton samples: the definition of an efficient protocol for DNA extraction and the identification of the best suitable primer pairs for barcode amplification. Due to the large volumes of sample needed, the commercial kits tested have proven to be inefficient, being the SDS-chloroform based method the most suitable protocol to extract DNA from zooplankton complex samples. From the primer pairs tested, the ones presenting a certain degree of degeneracy are the ones that most successfully amplify a wide-range of species. These findings allow us to set the bases for applying metabarcoding to different zooplankton samples all around the globe, as well as to other researchers an easy access to identify the best approach for their studies. In the last years high throughput sequencing technologies have had an incredible evolution. Adding to this evolution, the selection of appropriate barcodes and the construction of well-populated DNA libraries, the metabarcoding has a bright future regarding molecular taxonomic identification.

Acknowledgements

We are grateful to Fernando Villate (UPV-EHU) and Ann Bucklin (University of Conneticut) for their help in gathering a comprehensive zooplankton list, to Xabier Irigoien (KAUST) for fruitful discussions and support, to Iñaki Mendibil (AZTI-Tecnalia) for excellent technical assistance and to Eva Aylagas (AZTI-Tecnalia) for useful comments on the manuscript. We are also grateful to Unai Cotano (AZTI-Tecnalia) and an anonymous reviewer for useful comments on the manuscript. This work has been supported by the project Malaspina (Consolider-Ingenio 2010, CSD2008-00077), by Basque Government (Department of Economic Development and Competitiveness) and by a "Fundación Centros Tecnológicos – Iñaki Goenaga" grant to Jon Corell. This is contribution number 670 from the Marine Research Division (AZTI-Tecnalia).

Bibliography

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipmanl, D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.*, 215: 403-410

Bourlat, S. J., Borja, A., Gilbert, J., Taylor, M. I., Davies, N., Weisberg, S. B., Griffith, J. F., Lettieri, T., Field, D., Benzie, J., Glockner, F. O., Rodriguez-Ezpeleta, N., Faith, D. P., Bean, T. P. & Obst, M. 2013. Genomics in marine monitoring: New opportunities for assessing marine health status. *Marine pollution bulletin*, 74: 19-31

Brusca, R. C. & Brusca, G. J. 2003. *Invertebrates*. 2nd. Sinauer Associates, Inc., United States of America

Bucklin, A., Ortman, B. D., Jennings, R. M., Nigro, L. M., Sweetman, C. J., Copley, N. J., Sutton, T. & Wiebe, P. H. 2010. A "rosetta stone" for metazoan zooplankton: DNA barcode analysis of species diversity of the sargasso sea (northwest atlantic ocean). *Deep Sea Research Part II: Topical Studies in Oceanography*, 57 (24-26): 2234-2247

Bucklin, A., Steinke, D. & Blanco-Bercial, L. 2011. DNA barcoding of marine metazoa. *Annual Review of Marine Science*, 3 (1): 471-508

Carlotti, F., Giske, J. & Werner, F. 2000. Modeling zooplankton dynamics. 571-668. *In*: Harris, Wiebe, Lenz, Skjoldal&Huntley. *Ices zooplankton methodology manual*. Academic Press, United Kingdom

Creer, S. & Sinniger, F. 2012. Cosmopolitanism of microbial eukaryotes in the global deep seas. *Molecular ecology*, 21 (5): 1033-1035

Dunthorn, M., Klier, J., Bunge, J. & Stoeck, T. 2012. Comparing the hyper-variable v4 and v9 regions of the small subunit rdna for assessment of ciliate environmental diversity. *The Journal of eukaryotic microbiology*, 59 (2): 185-187

Evans, K. M., Chepurnov, V. A., Sluiman, H. J., Thomas, S. J., Spears, B. M. & Mann, D. G. 2009. Highly differentiated populations of the freshwater diatom *sellaphora capitata* suggest limited dispersal and opportunities for allopatric speciation. *Protist*, 160 (3): 386-396

Evans, N. & Paulay, G. 2012. DNA barcoding methods for invertebrates. 47-77. *In*: *DNA barcodes: Methods and protocols in molecular biology*. 2012/06/12. Methods Mol Biol

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 (5): 294-299

Geller, J., Meyer, C., Parker, M. & Hawk, H. 2013. Redesign of pcr primers for mitochondrial cytochrome c oxidase subunit i for marine invertebrates and application in all-taxa biotic surveys. *Molecular ecology resources*: doi: 10.1111

Green, M. R. & Sambrook, J. 2012. *Molecular cloning: A laboratory manual*. 4th edition. Cold spring Harbor Laboratory press, United States of America

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., Del Campo, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W. H., Lara, E., Le Bescot, N., Logares, R., Mahe, F., Massana, R., Montresor, M., Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A.-L., Siano, R., Stoeck, T., Vaulot, D., Zimmermann, P. & Christen, R. 2012. The protist ribosomal reference database (pr2): A catalog of unicellular eukaryote small sub-unit rrna sequences with curated taxonomy. *Nucleic acids research*: 1-8

Halanych, K. M., Bacheller, J. D., Aguinaldo, A. M. A., Liva, S. M., Hillis, D. M. & Lake, J. A. 1995. Evidence from 18s ribosomal DNA that the lophophorates are protostome animals. *Science*, 267: 1641-1643

Hebert, P. D., Ratnasingham, S. & deWaard, J. R. 2003. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings. Biological sciences / The Royal Society*, 270 Suppl 1: S96-99

Hebert, P. D., Stoeckle, M. Y. S., Zemlak, T. S. & Francis, C. M. 2004. Identification of birds through DNA barcodes. *PLoS Biology*, 2 (10): 1657-1663

Lenz, J. 2000. Introduction. 1-32. *In*: Harris, Wiebe, Lenz, Skjoldal&Huntley. *Ices zooplankton methodology manual*. Academic Press, United Kingdom

Leray, M., Boehm, J. T., Mills, S. C. & Meyer, C. P. 2011. Moorea biocode barcode library as a tool for understanding predator–prey interactions: Insights into the diet of common predatory coral reef fishes. *Coral Reefs*, 31 (2): 383-388

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T. & Machida, R. J. 2013. A new versatile primer set targeting a short fragment of the mitochondrial coi region for metabarcoding metazoan diversity: Applicationfor characterizing coral reef fish gut contents. *Frontiers in zoology*, 10 (34): 1-14

Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J. & Atkinson, A. 2013. Next generation sequencing reveals the hidden diversity of zooplankton assemblages. *PloS one*, 8 (11): e81327

Machida, R. J., Hashiguchi, Y., Nishida, M. & Nishida, S. 2009. Zooplankton diversity analysis through single-gene sequencing of a community sample. *BMC genomics*, 10: 438

Machida, R. J. & Knowlton, N. 2012. Pcr primers for metazoan nuclear 18s and 28s ribosomal DNA sequences. *PloS one*, 7 (9): e46180

Manaffar, R., Maleki, R., Zare, S., Agh, N., Soltanian, S., Sehatnia, B., Sorgeloo, P., Bossier, P. & Van Stappen, G. 2010. A new method for rapid DNA extraction from *artemia* (branchipoda, crustacea). *International Journal of Biological and Life Sciences*, 6 (2): 123-127

Meyer, C. 2003. Molecular systematics of cowries (gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biological Journal of the Linnean Society*, 79: 401-459

Meyer, C. P., Geller, J. B. & Paulay, G. 2005. Fine scale endemism on coral reefs: Archipelagic differentiation in turbinid gastropods. *Evolution*, 59 (1): 113-125

Meyer, C. P. & Paulay, G. 2005. DNA barcoding: Error rates based on comprehensive sampling. *PLoS Biology*, 3 (12): 2229-2238

Niu, B., Fu, L., Sun, S. & Li, W. 2010. Artificial and natural duplicates in pyrosequencing reads of metagenomic data. *BMC bioinformatics*, 11: 187

Nowrousian, M. 2010. Next-generation sequencing techniques for eukaryotic microorganisms: Sequencing-based solutions to biological problems. *Eukaryotic cell*, 9 (9): 1300-1310

Ortman, B. D. 2008. *DNA barcoding the medusozoa and ctenophora*. UMI Dissertation Publishing, United States of America

Postel, L., Fock, H. & Hagen, W. 2000. Biomass and abundence. 83-192. *In*: Harris, Wiebe, Lenz, Skjoldal&Huntley. *Ices zooplankton methodology manual*. Academic Press, United Kingdom

Promega. 2010. Wizard genomic DNA purification kit protocol. 1-21

QIAgen. 2010. Qiaamp DNA mini and blood mini handbook. 72

Salgado, A., Vieiralves, T., Lamarão, F. R. M., Assumpção, L. L. M., Gomes, D., Jascone, L., Valadão, A. L., Albano, R. M. & Lôbo-Hajdu, G. 2007. Field preservation and optimization of a DNA extraction method for porifera. *Porifera Research: Biodiversity, Innovation and Sustainability*: 555-560

Sanger, F., Nicklen, S. & Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.*, 74 (12): 5463-5467

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D., Breiner, H.-W. & Richards, T. A. 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular ecology*, 19 Suppl 1: 21-31

Thermo-Scientific. 2010a. *Nanodrop spectophotometers: 260/280 and 260/230 ratios*. Thermo Scientific. United States of America

Thermo-Scientific. 2010b. *Nanodrop spectrophotometers: Nucleic acid purity ratios*. Thermo Scientific
United States of America

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F. N., Ferreira de Siquiera, M., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A. S., Midgley, G. F., Miles, L., Ortega-Huerta, A. A., Peterson, T. A., Phillips, O. L. & Williams, S. E. 2004. Extinction risk from climate change. *Nature*, 427: 145-148

Walters, W. A., Caporaso, J. G., Lauber, C. L., Berg-Lyons, D., Fierer, N. & Knight, R. 2011. Primerprospector: De novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics*, 27 (8): 1159-1161

Wilson, I. G. 1997. Inhibition and facilitation of nucleic acid amplification. *Applied and environmental microbiology*, 63 (10): 3741-3751

SUPPLEMENTARY MATERIAL

Tuning of protocols and marker selection to evaluate the diversity of zooplankton using metabarcoding

Jon Corell and Naiara Rodríguez-Ezpeleta

Table 1. Different DNA extraction protocols used in this study. The number below the protocols represents the number of times that sample has been extracted with that protocol. Abbreviations of protocols: SDS = SDS-chloroform; SDS+RNAse = SDS-choroform plus RNAse treatment; SDS+column = SDS-choroform plus column treatment; Promega = *Wizard Genomic DNA Purification Kit*; QIAgen = *QIAamp DNA Minikit*.

	DNA extraction protocols				
Sample	**SDS**	**SDS+RNAse**	**SDS+column**	**Promega**	**Quiagen**
17	2				
18	3				
19	2				
20					1
21	2				
50					2
51					2
53					2
54					1
57					1
58					1
59					1
76	2				
77					2
96					2
102					1
104					2
106					2
107		1			2
108		1			2
118	1	1	1	1	1
119	1	1	1	1	1
127	2	1	1		
128	2				
129	1				
130	1				
131	1	1			
132	1	1	1		
133	2	1			

Table 2. Details of the primers tested for COI amplification. For each primer pair, primer names, forward and reverse sequences, taxonomic group for which it was designed and reference are indicated.

Primer pair name (forward name; reverse name)	Primer pair sequences in sense 5'-3' (top: forward and bottom: reverse)	Taxa for which it was designed	Reference
Barret1 (chelicerateforward1;chelicerateverse1)	TACTCTACTAATCATAAAGACATTGG CCTCCTCCTGAAGGGTCAAAAAATGA	Arachnida	(Barrett & Hebert, 2005)
Barret2 (chelicerateforward1;chelicerateverse2)	TACTCTACTAATCATAAAGACATTGG GGATGGCCAAAAAATCAAAATAAATG	Arachnida	(Barrett & Hebert, 2005)
Bely1 (COI-A+;COI-B-)	CCTGTTCTTGCTGGTGCTATTACNAT TAGTCAGAATATCGCCGAGGTATNCC	Clitellata	(Bely & Wray, 2004)
Blanco (Blancof;Blancor)	GAGCCTGGTCAGGAATAATCG GGTCTCCTCCTCCTCCAACAT	Clausocalanus	(Blanco-Bercial & Álvarez-Marqués, 2007)
Bucklin_Folmer1 (LCO-1490;HCO-2607)	GGTCAACAAATCATAAAGATATTGG ACATAGTGGAAATGTGCTACAACATA	Holozooplankton	(Bucklin *et al.*, 2010a; Folmer *et al.*, 1994)
Bucklin_Folmer2 (LCO-1490;HCO-Co-2358)	GGTCAACAAATCATAAAGATATTGG CCHACDGTAAAYATRTGRTG	Holozooplankton	(Bucklin *et al.*, 2010a; Folmer *et al.*, 1994)
Bucklin_Folmer3 (Cop-COI-1498F;HCO-2198)	AAYCATAAAGAYATYGGDAC TAAACTTCAGCCTGACCAAAAAATCA	Copepoda	(Bucklin *et al.*, 2010b; Folmer *et al.*, 1994)
Bucklin_Folmer4 (LCO-1490;Cop-COI-2105R)	GGTCAACAAATCATAAAGATATTGG CGRTCHGTHARNARYATDGTAATDGC	Copepoda	(Bucklin *et al.*, 2010b; Folmer *et al.*, 1994)
Bucklin_Folmer5 (LCO-1490;Cop-COI-2189R)	GGTCAACAAATCATAAAGATATTGG GGGTGACCAAAAAATCARAA	Copepoda	(Bucklin *et al.*, 2010b; Folmer *et al.*, 1994)
Bucklin_Folmer6 (LCO-1490;Crus-COI-2198R)	GGTCAACAAATCATAAAGATATTGG CCHACDGTAAAYATRTGRTG	Copepoda	(Bucklin *et al.*, 2010b; Folmer *et al.*, 1994)
Bucklin_Folmer7 (LCO-1490;Crus-COI-2428R)	GGTCAACAAATCATAAAGATATTGG TTAATHCCHGTDGGNACVGCAAT	Copepoda	(Bucklin *et al.*, 2010b; Folmer *et al.*, 1994)
Bucklin_Folmer8 (Ost-COI-1535F;HCO-2198)	GGDGCHTGAAGWGCWATGYTAGG TAAACTTCAGCCTGACCAAAAAATCA	Ostracoda	(Bucklin *et al.*, 2010b; Folmer *et al.*, 1994)
Bucklin_Folmer9 (LCO-1490;Eup-COI-2000R)	GGTCAACAAATCATAAAGATATTGG CADACAAAYARWGGDATTCGGTCTAT	Euphausiacea	(Bucklin *et al.*, 2010b; Folmer *et al.*, 1994)
Cohen (Cohen-Fwd;Cohen-Rev)	ATTYTBCCNGGRTTTGG TACCCYCGNCAAAAAC	Brachiopoda	(Cohen *et al.*, 2011)
Costa_Folmer1 (CrustF1;HCO2198)	TTTTCTACAAATCATAAAGACATTGG TAAACTTCAGGGTGACCAAAAAATCA	Crustacea	(Costa *et al.*, 2007; Folmer *et al.*, 1994)
Costa_Folmer2 (CrustF2;HCO2198)	GGTTCTTCTCCACCAACCACAARGAYATHGG TAAACTTCAGGGTGACCAAAAAATCA	Crustacea	(Costa *et al.*, 2007; Folmer *et al.*, 1994)
Dawson1 (LCOjf;HCOcato)	GGTCAACAAATCATAAAGATATTGGAAC CCTCCAGCAGGATCAAAGAAAG	Scyphozoa	(Dawson, 2005)
Dawson2 (LCOjf;HCO-2198)	GGTCAACAAATCATAAAGATATTGGAAC TAAACTTCAGCCTGACCAAAAAATCA	Scyphozoa	(Dawson, 2005)
Figueroa (H2612-COI;L592-RCOI)	AGGCCTAGGAAATGTATAGGGAAA AACCTTAATACATCTTTTTATGATG	Copepoda (primarly calanoida)	(Figueroa, 2011)
Folmer (LCO-1490;HCO-2198)	GGTCAACAAATCATAAAGATATTGG TAAACTTCAGCCTGACCAAAAAATCA	Metazoan invertebrates	(Folmer *et al.*, 1994)

Fukami (MCOIF;MCOIR)	TCTACAAATCATAAAGACATAGG GAGAAATTATACCAAAACCAGG	Scleractinia, Cnidaria	(Fukami *et al.*, 2007)
Geller (jgLCO1490;jgHCO2198)	TNTCNACNAAYCAYAARGAYATTGG TANACYTCNGGRTGNCCRAARAAYCA	Marine invertebrates	(Geller *et al.*, 2013)
Geller_Leray (jgLCO1490;LerayCOIintR)	TNTCNACNAAYCAYAARGAYATTGG GGRGGRTASACSGTTCASCCSGTSCC	Metazoa	(Geller *et al.*, 2013; Leray *et al.*, 2013b)
Goetze1 (RNI;VH)	GTAGTNGTAACWGCTCATGC TAAACTTCAGGGTGACCAAAAAATCA	Eucalanidae	(Goetze & Bradford-Grieve, 2005)
Goetze2 (PLXIVH;PLXIVL)	CCAAACGTTTCTTTCTTCCC TCAGCCAGGGTCTTTAATTGG	Pleuromamma	(Goetze, 2011)
Hebert_Hajibabei (LEP-F1;EPT-long-univR)	ATTCAACCAATCATAAAGATAT AARAAAATYATAAYAAANGCGTGNANNGT	River benthos macroinvertebrates	(Hajibabaei *et al.*, 2011; Hebert *et al.*, 2004b)
Hebert1 (LEP-F1;LEP-R1)	ATTCAACCAATCATAAAGATAT TAAACTTCTGGATGTCCAAAAA	Lepidoptera	(Hebert *et al.*, 2004b)
Hebert2 (LEP-F1;LEP-R2)	ATTCAACCAATCATAAAGATAT CTTATATTATTTATTCGTGGGAAAGC	Lepidoptera	(Hebert *et al.*, 2004b)
Hebert3 (BirdF1;BirdR1)	TTCTCCAACCACAAAGACATTGGCAC ACGTGGGAGATAATTCCAAATCCTG	Birds	(Hebert *et al.*, 2004a)
Hebert4 (BirdF2;BirdR2)	TTCTCCAACCACAAAGACATTGGCAC ACTACATGTGAGATGATTCCGAATCCAG	Birds	(Hebert *et al.*, 2004a)
Hebert5 (BirdF3;BirdR3)	TTCTCCAACCACAAAGACATTGGCAC AGGAGTTTGCTAGTACGATGCC	Birds	(Hebert *et al.*, 2004a)
Hill_Folmer1 (LCO-1703;HCO-2198)	GGTCAACAAATCATAAAGATATTGG TAAACTTCAGCCTGACCAAAAAATCA	Calanoida	(Folmer *et al.*, 1994; Hill *et al.*, 2011)
Hill_Folmer2 (LCO-1719;HCO-2198)	GGATTTGGTAACTGATTAGTGCC TAAACTTCAGCCTGACCAAAAAATCA	Calanoida	(Folmer *et al.*, 1994; Hill *et al.*, 2011)
Hoareau (COIceF;COIceR)	ACTGCCCACGCCCTAGTAATGATATTTTTTATGGTNATGCC TCGTGTGTCTACGTCCATTCCTACTGTRAACATRTG	Echinodermata	(Hoareau & Boissin, 2010)
Jungbluth (L1384;HCO2198)	GGTCATGTAATCATAAAGATATTG TAAACTTCAGGGTGACCAAAAAATCA	Copepod	(Jungbluth & Lenz, 2013)
Kojima1 (COI-7;COI-D)	ACNAAYAARCAYGAYATYGGNAC TCTGGGTGTCCRAARAAYCARAA	Vestimentiferan	(Kojima *et al.*, 1997)
Kojima2 (TW-1;TW-2)	CGAGTYCCTYTNTTYGTNTG ACTACRTARTANGTRTCRTG	Vestimentiferan	(Kojima *et al.*, 1997)
Kojima3 (COI-3;COI-6)	GTNTGRGCNCAYCAYATRTTYACNGT GGRTARTCNSWRTANCGNCGNGGYAT	Vestimentiferan	(Kojima *et al.*, 1997)
Kon (AmphL109;AmphH1325)	ATTCGNGCNGAAYTNTCNCAGCC TCNGAATAYCGNCGWGGTATNCC	Cephalochordata	(Kon *et al.*, 2006)
Leray_Folmer1 (LerayCOIintF;HCO-2198)	GGWACWGGWTGAACWGTWTAYCCYCC TAAACTTCAGCCTGACCAAAAAATCA	Metazoa	(Folmer *et al.*, 1994; Leray *et al.*, 2013b)
Leray_Folmer2 (LCO-1490;LerayCOIintR)	GGTCAACAAATCATAAAGATATTGG GGRGGRTASACSGTTCASCCSGTSCC	Metazoa	(Folmer *et al.*, 1994; Leray *et al.*, 2013b)
Leray_Geller (LerayCOIintF;jgHCO2198)	GGWACWGGWTGAACWGTWTAYCCYCC TANACYTCNGGRTGNCCRAARAAYCA	Metazoa	(Folmer *et al.*, 1994; Leray *et al.*, 2013b)
Leray_Meyer1 (LerayCOIintF;dgHCO2198)	GGWACWGGWTGAACWGTWTAYCCYCC TAAACTTCAGGGTGACCAAARAAYCA	Metazoa	(Folmer *et al.*, 1994; Leray *et al.*, 2013b)
Leray_Meyer2 (dgLCO1490;LerayCOIintR)	GGTCAACAAATCATAAAGAYATYGG GGRGGRTASACSGTTCASCCSGTSCC	Metazoa	(Folmer *et al.*, 1994; Leray *et al.*, 2013b)

Martinez_Folmer (LCOjf;HCO-2198)	GGTCAACAAATCATAAAGATATTGGAAC TAAACTTCAGCCTGACCAAAAAATCA	Medusozoa	(Folmer *et al.*, 1994; Martinez *et al.*, 2010)
Messing (M13F;M13R)	TGTAAAACGACGGCCAGT CAGGAAACAGCTATGAC	Zooplankton, freshwater microcrustacean	(Messing, 1983)
Meusnier (Uni-MinibarF1;Uni-MinibarR1)	TCCACTAATCACAARGATATTGGTAC GAAAATCATAATGAAGGCATGAGC	Universal	(Meusnier *et al.*, 2008)
Meyer (dgLCO1490;dgHCO2198)	GGTCAACAAATCATAAAGAYATYGG TAAACTTCAGGGTGACCAAARAAYCA	Gastropods	(Meyer *et al.*, 2005)
Misof1 (C1-J1751;C1-N2191)	GGATCACCTGATATAGCATTCCC CCCGGTAAAATTAAAATATAAACTTC	Panorpa	(Misof *et al.*, 2000)
Misof2 (C1-J 2165;TL2-N 3014)	GAAGTTTATATTTTAATTTTACCDGG TCCAATGCACTAATCTGCCATATTA	Panorpa	(Misof *et al.*, 2000)
Nelson (COIf;COIr)	TCMACTAATCAYAARGAYATTGGNAC CCDCTTAGWCCTARRAARTGTTGNGG	Vestimentiferan	(Nelson & Fisher, 2000)
Obst (CycF;CycR)	CGRATGGARCTYTCTCAYCC TTAAAATTACGRTCTGTYAAAAG	Cicliophora	(Obst *et al.*, 2005)
Ortman (MedCOIR;HCO-2607)	GGAACTGCTATAATCATAGTTGC ACATAGTGGAAATGTGCTACAACATA	Medusozoa	(Ortman *et al.*, 2010)
Ortman_Folmer1 (LCO-1490;HCO-Med-2414)	GGTCAACAAATCATAAAGATATTGG GGAACTGCTATAATCATAGTTGC	Medusozoa, Ctenophora	(Folmer *et al.*, 1994; Ortman, 2008)
Ortman_Folmer2 (LCO-1490;HCO-2424)	GGTCAACAAATCATAAAGATATTGG TTAATACCGTTAGGAACGGCAATAATTAT	Ctenophora	(Folmer *et al.*, 1994; Ortman, 2008)
Papadopoulos (ChelgCOI-F;ChelgCOI-R)	GGCCAAAACAGGGAGAGATA CGGGACTCAGTATAATTATTCGTCTA	Calanus helgolandicus	(Papadopoulos *et al.*, 2005)
Prosser (ZplankF1_t1;ZplankR1_t1)	TGTAAAACGACGGCCAGTTCTASWAATCATAARGATATTGG CAGGAAACAGCTATGACTTCAGGRTGRCCRAARAATCA	Zooplankton, freshwater microcrustacean	(Prosser *et al.*, 2013)
Radulovici (CrustDF1;CrustDR1)	GGTCWACAAAYCATAAAGAYATTGG TAAACYTCAGGRTGACCRAARAAYCA	Crustacea	(Radulovici *et al.*, 2009)
Schroth (L5;H5)	CTCTTGTAAGGTGAAGCC CATAATTCAACATCGAGG	Aurelia sp	(Schroth *et al.*, 2002)
Schwendinger_Folmer (LCO-1490;HCOoutout)	GGTCAACAAATCATAAAGATATTGG GTAAATATATGRTGDGCTC	Tardigrada, Opiliones, Myriapoda	(Folmer *et al.*, 1994; Schwendinger & Giribet, 2005)
Simon_Folmer (LCO-1490;Nancy)	GGTCAACAAATCATAAAGATATTGG CCCGGTAAAATTAAAATATAAACTTC	Marine invertebrates	(Folmer *et al.*, 1994; Simon *et al.*, 1994)
Stefaniak (Tunf;Tunr)	TCGACTAATCATAAAGATATTAG AACTTGTATTTAAATTACGATC	Tunicata	(Stefaniak *et al.*, 2009)
Ward1 (FishF1;FishR1)	TCAACCAACCACAAAGACATTGGCAC TAGACTTCTGGGTGGCCAAAGAATCA	Fish	(Ward *et al.*, 2005)
Ward2 (FishF2;FishR2)	TCGACTAATCATAAAGATATCGGCAC ACTTCAGGGTGACCGAAGAATCAGAA	Fish	(Ward *et al.*, 2005)
Ward4 (FishF1;FishR2)	TCAACCAACCACAAAGACATTGGCAC ACTTCAGGGTGACCGAAGAATCAGAA	Fish	(Ward *et al.*, 2005)
Ward5 (FishF2;FishR1)	TCGACTAATCATAAAGATATCGGCAC TAGACTTCTGGGTGGCCAAAGAATCA	Fish	(Ward *et al.*, 2005)
Zeale (ZBJ-ArtF1c;ZBJ-ArtR2c)	AGATATTGGAACWTTATATTTTATTTTTGG WACTAATCAATTWCCAAATCCTCC	Arachnida, Insecta	(Zeale *et al.*, 2011)

Table 3. Details of the primer pairs tested for 18S rRNA amplification. For each primer pair, primer names, forward and reverse sequences, taxonomic group for which it was designed and reference are indicated.

Primer pair name (forward name; reverse name)	Primer pair sequences in sense 5'-3' (top: forward and bottom: reverse)	Taxa for which it was designed	Reference
Apakupakul1 (18SrDNAA;18SrDNAB)	AACCTGGTTGATCCTGCCAGT TGATCCTTCCGCAGGTTCACCT	Euhirudinea	(Apakupakul *et al.*, 1999)
Apakupakul2 (18SrDNAL;18SrDNAC)	CCAACTACGAGCTTT CGGTAATTCCAGCTC	Euhirudinea	(Apakupakul *et al.*, 1999)
Burnette1 (18F509;18R925D)	CCCCGTAATTGGAATGAGTACA GATCYAAGAATTTCACCTCT	Annelida, Poeobiidae	(Burnette *et al.*, 2005)
Carranza1 (5F;9R)	GCGAAAGCATTTGCCAAGAA GATCCTTCCGCAGGTTCACCTAC	Platyhelminth	(Carranza *et al.*, 1996)
Carranza2 (1F;9R)	TACCTGGTTGATCCTGCCAGTAG GATCCTTCCGCAGGTTCACCTAC	Platyhelminth	
Cohen (F1;R1)	CCGAATTCGTCGACAACTGGTTGATCCTGCCAG CGAGATCTTAATGATGTCA	Brachiopoda, Phoronida	(Cohen *et al.*, 1998)
Figueroa (L10319-CYB;H10648-CYB)	CCTTGGGGKCAGATGTCTTTTTGGG GATAAAATTTTCWGGGTC	Calanoida	(Figueroa, 2011)
Floyd (Nem_18S_F;Nem_18S_R)	CGCGAATRGCTCATTACAACAGC GGGCGGTATCTGATCGCC	Nematodes	(Floyd *et al.*, 2005)
Fonseca (FO4;R22)	GCTTGTCTCAAAGATTAAGCC GCCTGCTGCCTTCCTTGGA	Marine metazoa	(Fonseca *et al.*, 2010)
Guillou1 (StoeckV4f;StoeckV4r)	CCAGCASCYGCGGTAATTCC ACTTTCGTTCTTGATYRA	Universal	(Stoeck *et al.*, 2010)
Guillou2 (StoeckV9Af;StoeckV9Ar)	GTACACACCGCCCGTC TGATCCTTCTGCAGGTTCACCTAC	Universal	(Stoeck *et al.*, 2010)
Guillou3 (StoeckV9Bf;StoeckV9Br)	TTGTACACACCGCCC CCTTCYGCAGGTTCACCTAC	Universal	(Stoeck *et al.*, 2010)
Hamilton (18S11b;18S2A)	GTCAGAGGTTCGAAGGCG GATCCTTCCGCAGGTTCACC	Metazoa	(Hamilton, 2003)
Hillis_Bucklin (18SE;18S-693R)	CTGGTTGATCCTGCCAGT AAACCTCTGGCAAAACTACG	Calanoida	(Bucklin *et al.*, 2003; Hillis & Dixon, 1991)
Hillis_Halanych (18SE;18L)	CTGGTTGATCCTGCCAGT GAATTACCGCGGCTGCTGGCACC	Pisionidae	(Halanych *et al.*, 1995; Hillis & Dixon, 1991)
Hillis_Hamby (18SE;18SL)	CTGGTTGATCCTGCCAGT CACCTACGGAAACCTTGTTACGACTT	Marine invertebrates	(Hamby & Zimmer, 1988; Hillis & Dixon, 1991)
Holland (Mollusca1991f;Mollusca1991r)	GCCAGTAGCATATGCTTGTCTC AGACTTGCCTCCAATGGATCC	Mollusc, Hemichordata, Echinodermata	(Holland *et al.*, 1991)
Kelly (SP4;SP5)	GGTGGTGCCCTTCCGTCAATTCCTTTAAGTTT GTCTGGTGCCAGCAGCCGCGGTAATTC	Demospongiae	(Kelly-Borges & Pomponi, 1994)
Machida1 (1;2RC)	CTGGTGCCAGCAGCCGCGGYAA TCCGTCAATTYCTTTAAGTT	Metazoa	(Machida & Knowlton, 2012)
Machida2 (3;4RC)	GYGGTGCATGGCCGTTSKTRGTT CKRAGGGCATYACWGACCTGTTAT	Metazoa	(Machida & Knowlton, 2012)
Machida3 (3;5RC)	GYGGTGCATGGCCGTTSKTRGTT GTGTGYACAAAGGBCAGGGAC	Metazoa	(Machida & Knowlton, 2012)
Machida4 (1;5RC)	CTGGTGCCAGCAGCCGCGGYAA GTGTGYACAAAGGBCAGGGAC	Metazoa	(Machida & Knowlton, 2012)
Mullin (18S1.2;18Sr2b)	GGCGATCAGATACCGCCCTAGTT TACAAAGGGCAGGGACGTAAT	Nematodes	(Mullin *et al.*, 2003)
Nishida (18S1A;18S2A)	ATCTGGTTGATCCTGCCAGT GATCCTTCCGCAGGTTCACC	Fungi	(Nishida & Sugiyama, 1993)

Palero (18S1f;18Sb2.9)	TACCTGGTTGATCCTGCCAGTAG TATCTGATCGCCTTCGAACCTCT	Achelata	(Palero *et al.*, 2009)
Porazinska (A-NF1;B-18Sr2b)	GCCTCCCTCGCGCCATCAGGGTGGTGCATGGCCGTTCTTAGTT GCCTTGCCAGCCCGCTCAGTACAAAGGGCAGGGACGTAAT	Nematodes	(Porazinska *et al.*, 2009)
Raupach (18A1;1800)	CTGGTTGATCCTGCCAGTCATATGC GATCCTTCCGCAGGTTCACCTACG	Marine invertebrates	(Raupach *et al.*, 2009)
Spears (18s329;18sI)	TAATGATCCTTCCGCAGGTT AACTYAAAGGAATTGACGG	Brachyura, Mesocyclops	(Spears *et al.*, 1992)
Struck (18F35;18R399)	TCTCAAAGATTAAGCCATGCA CCCTCTCCGGAATCGAACCCTGAT	Aeolosomatidae, Parergodrilidae	(Struck *et al.*, 2002)
Thum1 (18S-1F;18S-1R)	AACCTGGTTGATCCTGCCAGT TGGTGCCCTTCCGTCAATTCCT	Diaptomidae	(Thum, 2004)
Thum2 (18S-2F;18S-2R)	CTGGTGCCAGCAGCCGCGG TTGATCCTTCTGCAGGTTCACCTAC	Diaptomidae	(Thum, 2004)

Bibliography

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipmanl, D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.*, 215: 403-410

Apakupakul, K., Siddall, M. E. & Burreson, E. M. 1999. Higher level relationships of leeches (annelida: Clitellata: Euhirudinea) based on morphology and gene sequences. *Molecular phylogenetics and evolution*, 12 (3): 350-359

Barrett, R. D. H. & Hebert, P. D. N. 2005. Identifying spiders through DNA barcodes. *Canadian Journal of Zoology*, 83 (3): 481-491

Bely, A. E. & Wray, G. A. 2004. Molecular phylogeny of naidid worms (annelida: Clitellata) based on cytochrome oxidase i. *Molecular phylogenetics and evolution*, 30 (1): 50-63

Blanco-Bercial, L. & Álvarez-Marqués, F. 2007. Rflp procedure to discriminate between *clausocalanus* giesbrecht, 1888 (copepoda, calanoida) species in the central cantabrian sea. *Journal of Experimental Marine Biology and Ecology*, 344 (1): 73-77

Bourlat, S. J., Borja, A., Gilbert, J., Taylor, M. I., Davies, N., Weisberg, S. B., Griffith, J. F., Lettieri, T., Field, D., Benzie, J., Glockner, F. O., Rodriguez-Ezpeleta, N., Faith, D. P., Bean, T. P. & Obst, M. 2013. Genomics in marine monitoring: New opportunities for assessing marine health status. *Marine pollution bulletin*, 74: 19-31

Brusca, R. C. & Brusca, G. J. 2003. *Invertebrates*. 2nd. Sinauer Associates, Inc., United States of America

Bucklin, A., Frost, B. W. & Bradford-Grieve, J. 2003. Molecular systematic and phylogenetic assessment of 34 calanoid copepod species of the calanidae and clausocalanidae. *Marine Biology*, 142: 333-343

Bucklin, A., Hopcroft, R. R., Kosobokova, K. N., Nigro, L. M., Ortman, B. D., Jennings, R. M. & Sweetman, C. J. 2010a. DNA barcoding of arctic ocean holozooplankton for species identification and recognition. *Deep Sea Research Part II: Topical Studies in Oceanography*, 57 (1-2): 40-48

Bucklin, A., Ortman, B. D., Jennings, R. M., Nigro, L. M., Sweetman, C. J., Copley, N. J., Sutton, T. & Wiebe, P. H. 2010b. A "rosetta stone" for metazoan zooplankton: DNA barcode analysis of species diversity of the sargasso sea (northwest atlantic ocean). *Deep Sea Research Part II: Topical Studies in Oceanography*, 57 (24-26): 2234-2247

Bucklin, A., Steinke, D. & Blanco-Bercial, L. 2011. DNA barcoding of marine metazoa. *Annual Review of Marine Science*, 3 (1): 471-508

Burnette, A. B., Struck, T. H. & Halanych, K. M. 2005. Holopelagic *poeobius meseres* ("poeobiidae," annelida) is derived from benthic flabelligerid worms. *Biol. Bull.*, 208: 213-220

Carlotti, F., Giske, J. & Werner, F. 2000. Modeling zooplankton dynamics. 571-668. *In*: Harris, Wiebe, Lenz, Skjoldal&Huntley. *Ices zooplankton methodology manual*. Academic Press, United Kingdom

Carranza, S., Giribet, G., Ribera, C., Baguii, J. & Riutort, M. 1996. Evidence that two types of 18s rdna coexist in the genome of *dugesi(schmidtea) mediterranea* (platyhelminthes, turbellaria, tricladida). *Molecular biology and evolution*, 13 (6): 824-832

Cohen, B. L., Bitner, M. A., Harper, E. M., Lee, D. E., Mutschke, E. & Sellanes, J. 2011. Vicariance and convergence in magellanic and new zealand long-looped brachiopod clades (pan-brachiopoda: Terebratelloidea). *Zoological Journal of the Linnean Society*, 162 (3): 631-645

Cohen, B. L., Gawthrop, A. & Cavalier-Smith, T. 1998. Molecular phylogeny of brachiopods and phoronids based on nuclear-encoded small

subunit ribosomal rna gene sequences. *Phil.Trans. R. Soc. Lond. B*, 353: 2039-2061

Costa, F. O., deWaard, J. R., Boutillier, J., Ratnasingham, S., Dooh, R. T., Hajibabaei, M. & Hebert, P. D. N. 2007. Biological identifications through DNA barcodes: The case of the crustacea. *Canadian Journal of Fisheries and Aquatic Sciences*, 64 (2): 272-295

Creer, S. & Sinniger, F. 2012. Cosmopolitanism of microbial eukaryotes in the global deep seas. *Molecular ecology*, 21 (5): 1033-1035

Dawson, M. N. 2005. Incipient speciation of *catostylus mosaicus* (scyphozoa, rhizostomeae, catostylidae), comparative phylogeography and biogeography in south-east australia. *Journal of Biogeography*, 32 (3): 515-533

Dunthorn, M., Klier, J., Bunge, J. & Stoeck, T. 2012. Comparing the hypervariable v4 and v9 regions of the small subunit rdna for assessment of ciliate environmental diversity. *The Journal of eukaryotic microbiology*, 59 (2): 185-187

Evans, K. M., Chepurnov, V. A., Sluiman, H. J., Thomas, S. J., Spears, B. M. & Mann, D. G. 2009. Highly differentiated populations of the freshwater diatom *sellaphora capitata* suggest limited dispersal and opportunities for allopatric speciation. *Protist*, 160 (3): 386-396

Evans, N. & Paulay, G. 2012. DNA barcoding methods for invertebrates. 47-77. *In*: *DNA barcodes: Methods and protocols in molecular biology*. 2012/06/12. Methods Mol Biol

Figueroa, D. F. 2011. Phylogenetic analysis of *ridgewayia* (copepoda: Calanoida) from the galapagos and of a new species from the florida keys with a reevaluation of the phylogeny of calanoida. *Journal of Crustacean Biology*, 31 (1): 153-165

Floyd, R. M., Rogers, A. D., Lambshead, P. J. D. & Smith, C. R. 2005. Nematode-specific pcr primers for the 18s small subunit rrna gene. *Molecular Ecology Notes*, 5 (3): 611-612

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 (5): 294-299

Fonseca, V. G., Carvalho, G. R., Sung, W., Johnson, H. F., Power, D. M., Neill, S. P., Packer, M., Blaxter, M. L., Lambshead, P. J., Thomas, W. K. & Creer, S. 2010. Second-generation environmental sequencing unmasks marine metazoan biodiversity. *Nature communications*, 1: 98

Fukami, H., Chen, C. A., Chiou, C.-Y. & Knowlton, N. 2007. Novel group i introns encoding a putative homing endonuclease in the mitochondrial cox1 gene of scleractinian corals. *Journal of molecular evolution*, 64 (5): 591-600

Geller, J., Meyer, C., Parker, M. & Hawk, H. 2013. Redesign of pcr primers for mitochondrial cytochrome c oxidase subunit i for marine invertebrates and application in all-taxa biotic surveys. *Molecular ecology resources*: doi: 10.1111

Goetze, E. 2011. Population differentiation in the open sea: Insights from the pelagic copepod *pleuromamma xiphias*. *Integrative and comparative biology*, 51 (4): 580-597

Goetze, E. & Bradford-Grieve, J. 2005. Genetic and morphological description of *eucalanus spinifer* t. Scott, 1894 (calanoida: Eucalanidae), a circumglobal sister species of the copepod *e. Hyalinus* s.S. (claus, 1866). *Progress in Oceanography*, 65 (1): 55-87

Green, M. R. & Sambrook, J. 2012. *Molecular cloning: A laboratory manual*. 4th edition. Cold spring Harbor Laboratory press, United States of America

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., Del Campo, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W. H., Lara, E., Le Bescot, N., Logares, R., Mahe, F., Massana, R., Montresor, M., Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A.-L., Siano, R., Stoeck, T., Vaulot, D., Zimmermann, P. & Christen, R. 2012. The protist ribosomal reference database (pr2): A catalog of unicellular eukaryote small sub-unit rrna sequences with curated taxonomy. *Nucleic acids research*: 1-8

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A. & Baird, D. J. 2011. Environmental barcoding: A next-generation sequencing approach for biomonitoring applications using river benthos. *PloS one*, 6 (4): e17497

Halanych, K. M., Bacheller, J. D., Aguinaldo, A. M. A., Liva, S. M., Hillis, D. M. & Lake, J. A. 1995. Evidence from 18s ribosomal DNA that the lophophorates are protostome animals. *Science*, 267: 1641-1643

Hamby, K. R. & Zimmer, E. A. 1988. Ribosomal rna sequences for inferring phylogeny within the grass family (poaceae) *Pl. Syst. Evol.*, 160: 29-37

Hamilton, H. C. 2003. *A molecular method for assessing meiofauna diversity in marine sediments*. Master of Science. University of South Florida. United States of America

Hebert, P. D., Ratnasingham, S. & deWaard, J. R. 2003. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings. Biological sciences / The Royal Society*, 270 Suppl 1: S96-99

Hebert, P. D., Stoeckle, M. Y. S., Zemlak, T. S. & Francis, C. M. 2004a. Identification of birds through DNA barcodes. *PLoS Biology*, 2 (10): 1657-1663

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. & Hallwachs, W. 2004b. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *astraptes fulgerator*. *Proceedings of the National Academy of Sciences*, 101 (41): 14812-14817

Hill, R. S., Allen, L. D. & Bucklin, A. 2011. Multiplexed species-specific pcr protocol to discriminate four n. Atlantic *calanus* species, with an mitcoi gene tree for ten *calanus* species. *Marine Biology*, 139: 279-287

Hillis, D. M. & Dixon, M. T. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. *The quarterly review of biology*, 66 (4): 411-453

Hoareau, T. B. & Boissin, E. 2010. Design of phylum-specific hybrid primers for DNA barcoding: Addressing the need for efficient coi amplification in the echinodermata. *Molecular ecology resources*, 10 (6): 960-967

Holland, P. W. H., hacker, A. M. & Williams, N. A. 1991. A molecular analysis of the phylogenetic affinities of *saccoglossus cambrensis* brambell & cole (hemichordata). *Phil. Trans. R. Soc. Land. B*, 332: 185-189

Jungbluth, M. J. & Lenz, P. H. 2013. Copepod diversity in a subtropical bay based on a fragment of the mitochondrial coi gene. *Journal of Plankton Research*, 35 (3): 630-643

Kelly-Borges, M. & Pomponi, S. 1994. Phylogeny and classification of lithistid sponges (*porifera: Demospongiae*): A preliminary assessment using ribosomal DNA sequence comparisons*. *Molecular Marine Biology and Biotechnology*, 3 (2): 87-103

Kojima, S., Segawa, R., Hashimoto, J. & Ohta, S. 1997. Molecular phylogeny of vestimentiferans collected around japan, revealed by the nucleotide sequences of mitochondrial DNA. *Marine Biology*, 127: 507-513

Kon, T., Nohara, M., Nishida, M., Sterrer, W. & Nishikawa, T. 2006. Hidden ancient diversification in the circumtropical lancelet asymmetron lucayanum complex. *Marine Biology*, 149 (4): 875-883

Lenz, J. 2000. Introduction. 1-32. *In*: Harris, Wiebe, Lenz, Skjoldal&Huntley. *Ices zooplankton methodology manual*. Academic Press, United Kingdom

Leray, M., Boehm, J. T., Mills, S. C. & Meyer, C. P. 2011. Moorea biocode barcode library as a tool for understanding predator–prey interactions: Insights into the diet of common predatory coral reef fishes. *Coral Reefs*, 31 (2): 383-388

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T. & Machida, R. J. 2013a. A new versatile primer set targeting a short fragment of the mitochondrial coi region for metabarcoding metazoan diversity: Applicationfor characterizing coral reef fish gut contents. *Frontiers in zoology*, 10 (34): 1-14

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T. & Machida, R. J. 2013b. Targeting a short fragment of the mitochondrial coi region for metabarcoding metazoan diversity – application for characterizing coral reef fish gut contents.

Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J. & Atkinson, A. 2013. Next generation sequencing reveals the hidden diversity of zooplankton assemblages. *PloS one*, 8 (11): e81327

Machida, R. J., Hashiguchi, Y., Nishida, M. & Nishida, S. 2009. Zooplankton diversity analysis through single-gene sequencing of a community sample. *BMC genomics*, 10: 438

Machida, R. J. & Knowlton, N. 2012. Pcr primers for metazoan nuclear 18s and 28s ribosomal DNA sequences. *PloS one*, 7 (9): e46180

Manaffar, R., Maleki, R., Zare, S., Agh, N., Soltanian, S., Sehatnia, B., Sorgeloo, P., Bossier, P. & Van Stappen, G. 2010. A new method for rapid DNA extraction from *artemia* (branchipoda, crustacea). *International Journal of Biological and Life Sciences*, 6 (2): 123-127

Martinez, D. E., Iniguez, A. R., Percell, K. M., Willner, J. B., Signorovitch, J. & Campbell, R. D. 2010. Phylogeny and biogeography of *hydra* (cnidaria: Hydridae) using mitochondrial and nuclear DNA sequences. *Molecular phylogenetics and evolution*, 57 (1): 403-410

Messing, J. 1983. New m13 vectors for cloning. *Methods in enzymology*, 101: 20-78

Meusnier, I., Singer, G. A., Landry, J.-F., Hickey, D. A., Hebert, P. D. & Hajibabaei, M. 2008. A universal DNA mini-barcode for biodiversity analysis. *BMC genomics*, 9: 214

Meyer, C. 2003. Molecular systematics of cowries (gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biological Journal of the Linnean Society*, 79: 401-459

Meyer, C. P., Geller, J. B. & Paulay, G. 2005. Fine scale endemism on coral reefs: Archipelagic differentiation in turbinid gastropods. *Evolution*, 59 (1): 113-125

Meyer, C. P. & Paulay, G. 2005. DNA barcoding: Error rates based on comprehensive sampling. *PLoS Biology*, 3 (12): 2229-2238

Misof, B., Erpenbeck, D. & Sauer, K. P. 2000. Mitochondrial gene fragments suggest paraphyly of the genus *panorpa* (mecoptera, panorpidae). *Molecular phylogenetics and evolution*, 17 (1): 76-84

Mullin, P. G., Harris, T. S. & Powers, T. O. 2003. Systematic status of *campydora* cobb, 1920 (nematoda: Campydorina). *Nematodology*, 5 (5): 699-711

Nelson, K. & Fisher, C. R. 2000. Absence cospeciation in the deep-sea vestimentiferan tube worms and their bacterial endosymbionts. *Symbiosis*, 28: 1-15

Nishida, H. & Sugiyama, J. 1993. Phylogenetic relationships among *taphrina, saitoella*, and other higher fungi. *Molecular biology and evolution*, 10 (2): 431-436

Niu, B., Fu, L., Sun, S. & Li, W. 2010. Artificial and natural duplicates in pyrosequencing reads of metagenomic data. *BMC bioinformatics*, 11: 187

Nowrousian, M. 2010. Next-generation sequencing techniques for eukaryotic microorganisms: Sequencing-based solutions to biological problems. *Eukaryotic cell*, 9 (9): 1300-1310

Obst, M., Funch, P. & Giribet, G. 2005. Hidden diversity and host specificity in cyclophorans: A phylogeographic analysis along the north atlantic and mediterranean sea. *Molecular ecology*, 14 (14): 4427-4440

Ortman, B. D. 2008. *DNA barcoding the medusozoa and ctenophora*. UMI Dissertation Publishing, United States of America

Ortman, B. D., Bucklin, A., Pagès, F. & Youngbluth, M. 2010. DNA barcoding the medusozoa using mtcoi. *Deep Sea Research Part II: Topical Studies in Oceanography*, 57 (24-26): 2148-2156

Palero, F., Crandall, K. A., Abello, P., Macpherson, E. & Pascual, M. 2009. Phylogenetic relationships between spiny, slipper and coral lobsters (crustacea, decapoda, achelata). *Molecular phylogenetics and evolution*, 50 (1): 152-162

Papadopoulos, L. N., Peijnenburg, K. T. C. A. & Luttikhuizen, P. C. 2005. Phylogeography of the calanoid copepods *calanus helgolandicus* and *c. Euxinus* suggests pleistocene divergences between atlantic, mediterranean, and black sea populations. *Marine Biology*, 147 (6): 1353-1365

Porazinska, D. L., Giblin-Davis, R. M., Faller, L., Farmerie, W., Kanzaki, N., Morris, K., Powers, T. O., Tucker, A. E., Sung, W. & Thomas, W. K. 2009. Evaluating high-throughput sequencing as a method for metagenomic analysis of nematode diversity. *Molecular ecology resources*, 9 (6): 1439-1450

Postel, L., Fock, H. & Hagen, W. 2000. Biomass and abundence. 83-192. *In*: Harris, Wiebe, Lenz, Skjoldal&Huntley. *Ices zooplankton methodology manual*. Academic Press, United Kingdom

Promega. 2010. Wizard genomic DNA purification kit protocol. 1-21

Prosser, S., Martinez-Arce, A. & Elias-Gutierrez, M. 2013. A new set

of primers for coi amplification from freshwater microcrustaceans. *Molecular ecology resources*:

QIAgen. 2010. Qiaamp DNA mini and blood mini handbook. 72

Radulovici, A. E., Sainte-Marie, B. & Dufresne, F. 2009. DNA barcoding of marine crustaceans from the estuary and gulf of st lawrence: A regional-scale approach. *Molecular ecology resources*, 9 Suppl s1: 181-187

Raupach, M. J., Mayer, C., Malyutina, M. & Wagele, J.-W. 2009. Multiple origins of deep-sea asellota (crustacea: Isopoda) from shallow waters revealed by molecular data. *Proceedings. Biological sciences / The Royal Society*, 276 (1658): 799-808

Salgado, A., Vieiralves, T., Lamarão, F. R. M., Assumpção, L. L. M., Gomes, D., Jascone, L., Valadão, A. L., Albano, R. M. & Lôbo-Hajdu, G. 2007. Field preservation and optimization of a DNA extraction method for porifera. *Porifera Research: Biodiversity, Innovation and Sustainability*: 555-560

Sanger, F., Nicklen, S. & Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.*, 74 (12): 5463-5467

Schroth, W., Jarms, G., Streit, B. & Schierwater, B. 2002. Speciation and phylogeography in the cosmopolitan marine moon jelly, *aurelia* sp. *BMC Evolutionary Biology*, 2 (1): 1-10

Schwendinger, P. J. & Giribet, G. 2005. The systematics of the south-east asian genus *fangensis* rambla (opiliones: Cyphophthalmi: Stylocellidae). *Invertebrate Systematics*, 19: 297-323

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. & Flook, P. 1994. Evolution, weigthing and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reactions primers. *Ann. Entomol. Soc. Am.*, 87 (6): 651-701

Spears, T., Abele, L. G. & Kim, W. 1992. The monophyly of brachyuran crabs a phylogenetic study based on 18s rrna. *Systematic biology*, 41 (4): 446-461

Stefaniak, L., Lambert, G., Gittenberger, A., Zhang, H., Lin, S. & Whitlatch, R. B. 2009. Genetic conspecificity of the worldwide populations of *didemnum vexillum* kott, 2002. *Aquatic Invasions*, 4 (1): 29-44

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D., Breiner, H.-W. & Richards, T. A. 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular ecology*, 19 Suppl 1: 21-31

Struck, T. H., Hessling, R. & Purschke, G. 2002. The phylogenetic position of the aeolosomatidae and parergodrilidae, two enigmatic oligochaete-like taxa of the 'polychaeta', based on molecular data from 18srdna sequences. *J. Zool. Syst. Evol. Research*, 40: 155-163

Thermo-Scientific. 2010a. *Nanodrop spectophotometers: 260/280 and 260/230 ratios*. Thermo Scientific. United States of America

Thermo-Scientific. 2010b. *Nanodrop spectrophotometers: Nucleic acid purity ratios*. Thermo Scientific

United States of America

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F. N., Ferreira de Siquiera, M., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A. S., Midgley, G. F., Miles, L., Ortega-Huerta, A. A., Peterson, T. A., Phillips, O. L. & Williams, S. E. 2004. Extinction risk from climate change. *Nature*, 427: 145-148

Thum, R. A. 2004. Using 18s rdna to resolve diaptomid copepod (copepoda: Calanoida: Diaptomidae) phylogeny: An example with the north american genera. *Hydrobiologia*, 519: 135-141

Walters, W. A., Caporaso, J. G., Lauber, C. L., Berg-Lyons, D., Fierer, N. & Knight, R. 2011. Primerprospector: De novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics*, 27 (8): 1159-1161

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. & Hebert, P. D. 2005. DNA barcoding australia's fish species. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 360 (1462): 1847-1857

Wilson, I. G. 1997. Inhibition and facilitation of nucleic acid amplification. *Applied and environmental microbiology*, 63 (10): 3741-3751

Zeale, M. R. K., Butlin, R. K., Barker, G. L. A., Lees, D. C. & Jones, G. 2011. Taxon-specific pcr for DNA barcoding arthropod prey in bat faeces. *Molecular ecology resources*, 11 (2): 236-244

www.azti.es

azti tecnalia

Txatxarramendi ugartea z/g
48395 Sukarrieta (Bizkaia)
Tel.: +34 94 657 40 00
Fax: +34 94 657 25 55

Herrera Kaia, Portualdea z/g
20110 Pasaia (Gipuzkoa)

Parque Tecnológico de Bizkaia
Astondo bidea. Edificio 609.
48160 Derio (Bizkaia)