Suitability of a dual COI marker for marine zooplankton DNA metabarcoding

Anna Schroeder [a,b], Alberto Pallavicini [b,c,*], Paolo Edomi [b], Marco Pansera [a,c], Elisa Camatti [a]

[a] National Research Council, Institute of Marine Science (CNR ISMAR) Venice, Arsenale Tesa 104, Castello 2737/F, 30122, Venice, Italy
[b] University of Trieste, Department of Life Sciences, Via Licio Giorgieri 5, 34127, Trieste, Italy
[c] Stazione Zoologica Anton Dohrn, Villa Comunale, 80121, Naples, Italy

ARTICLE INFO

Keywords:
Biomonitoring
COI
Marker
Metabarcoding
mlCOIint
Primer
Zooplankton

ABSTRACT

As DNA metabarcoding has become an emerging tool for surveying biodiversity, including its application in legally binding assessments, reliable and efficient barcodes are requested, especially for the highly diverse group of zooplankton. This study focuses on comparing the efficiency of two mitochondrial COI barcodes based on the internal primers mlCOIintF and mlCOIintR utilizing mesozooplankton samples collected in a Mediterranean lagoon. Our results indicate that after a slight adjustment, the mlCOIintR primer performs in combination with jdgLCO1490 (herein) very comparably to the much more widely used primer system mlCOIintF/jgHCO2198+dgHCO2198, in terms of level of taxonomic resolution, species detection and their relative abundance in terms of numbers of reads. As for some groups, like Ctenophora, this barcode is not suitable; a combination of them may be the best option to rely on the Folmer region in its entirety without the risk of losing information for a limited primer match.

1. Introduction

In recent years, the estimation of biodiversity with DNA metabarcoding using high-throughput sequencing (HTS) is becoming a promising tool for surveying biodiversity (Taberlet et al., 2012), especially in the assessment of zooplankton biodiversity in various marine environments (e.g., Bucklin et al., 2019; Harvey et al., 2017; Questel et al., 2021). Until now, out of the more than two thousand research articles published on metabarcoding, 65 were related to zooplankton research (source: Scopus, April 2021). Its success is due to increased sample processing speed, allowing to expand the sampling effort with sustainable costs paired with a broad taxonomic coverage (Brannock et al., 2014; Coissac et al., 2012). As well as due to the constantly decreasing taxonomic expertise necessary to adequately process zooplankton samples.

In general, the mitochondrial cytochrome oxidase subunit I (COI) region (Hebert et al., 2003) is the most frequent DNA barcode region used (Hebert et al., 2003)for accurate and positive species identification, as most taxa show a significantly different intra-versus inter-specific variability, allowing to discriminate between closely related species. In the case of marine zooplankton, as this group consists of animals from almost all phyla, it is challenging to find an appropriate primer covering the immense biodiversity of zooplankton. Different gene regions have been used to describe with metabarcoding the diversity of mixed zooplankton assemblages: the four commonly used gene regions to describe patterns across different systematic levels are the mitochondrial 16S rRNA gene (Clarke et al., 2017; Lindeque et al., 2006), the nuclear genes 28S rRNA (Harvey et al., 2017; Hirai et al, 2013, 2020), and 18S rRNA (Blanco-Bercial, 2020; Chain et al., 2016; Lindeque et al., 2013; Sommer et al., 2017) and the mitochondrial COI gene (Carroll et al., 2019; Schroeder et al., 2020; Stefanni et al., 2018; Zaiko et al., 2015). The latter two are the most used ones: the 18S V9 region is a hypervariable region flanked by highly conserved sections, indicating a broad range of taxonomic groups (Amaral-Zettler et al., 2009; Medlin et al., 1988) but allowing family level identification at best. However, the V9 region has been shown to probably not be optimal as 18S marker for zooplankton biodiversity assessment (Blanco-Bercial, 2020) and Questel et al. (2021), in fact, found the V4 region to have a greater taxonomic resolution. Due to the low genetic diversity and consequently limited taxonomic resolution of 18S rRNA regions, more and more zooplankton metabarcoding studies have also begun to use the COI gene, which shows excellent taxonomic resolution, but with a drawback of

* Corresponding author. University of Trieste, Department of Life Sciences, Via Licio Giorgieri 5, 34127, Trieste, Italy.
E-mail addresses: anna.schroeder@units.it (A. Schroeder), pallavic@units.it (A. Pallavicini), edomi@units.it (P. Edomi), marco.pansera@szn.it (M. Pansera), elisa.camatti@ve.ismar.cnr.it (E. Camatti).

reduced amplification success (Clarke et al., 2017). Indeed, several studies use a multi-marker approach to reduce the bias resulting from differing amplification success between various taxonomic groups. However, Clarke et al. (2017) demonstrated that COI resolved up to threefold more taxa to the species level than 18S rRNA. Also, COI has the benefit of an enormous database for COI sequences (>3.5 million sequences deposited on Genbank), again increasing the suitability of COI as a genetic marker for metabarcoding. Moreover, a new COI sequence database for zooplankton communities called MetaZooGene Barcode Atlas and Database (MZGdb) (https://metazoogene.org/database) has become recently available (Bucklin et al., 2021), which has mined the extensive GenBank and BOLD repositories, removing errors found within these databases, and has created worldwide and geographically specific reference sequence databases for use in zooplankton metabarcoding studies. Therefore, this study focuses on the use of two COI barcodes, based on the internal COI primers proposed by Leray et al. (2013), mlCOIintF and mlCOIintR (5′-GGRGGRTASACSGTTCASCCSGTSCC-3'), which were utilized in several studies. However, as Leray et al. (2013) found "the reverse primer to perform poorly", almost all studies implemented the forward primer in combination with HCO2198 (or its degenerate versions dgHCO2198 and jgHCO2198). Only a few studies, mainly on terrestrial arthropods, started utilizing the proposed reverse mlCOIintR with contrasting success: while Brandon-Mong et al. (2015), which excluded this primer after weak amplification success, and Krehenwinkel et al. (2017) utilized the mlCOIintR primer as proposed by Leray et al. (2013), other authors adapted the mlCOIintR sequence by modifying e.g. all "S" nucleotides to "W" as Günther et al. (2018) (on marine eDNA), to "N" as Wang et al. (2019) or to an "I" as Shokralla et al. (2015), as in fact, the proposed primer sequence is not the proper reverse complement to mlCOIintF.

This study focuses on testing the efficiency of the reverse mlCOIintR primer in combination with the degenerated jdgLCO1490 primer by comparing it with the first half (utilizing the forward mlCOIintF primer) of the COI gene that has been previously evaluated for zooplankton biodiversity assessments by Schroeder et al. (2020) by comparing it with morphological identification.

2. Material and methods

2.1. Data collection

For this study, 44 mesozooplankton samples, a subset of a more extensive survey conducted monthly from April 2018 to March 2019 in the Venice Lagoon (Italy), was used (Fig. 1). This subset corresponds to the samples, where also individuals of *Mnemiopsis leidyi* were collected

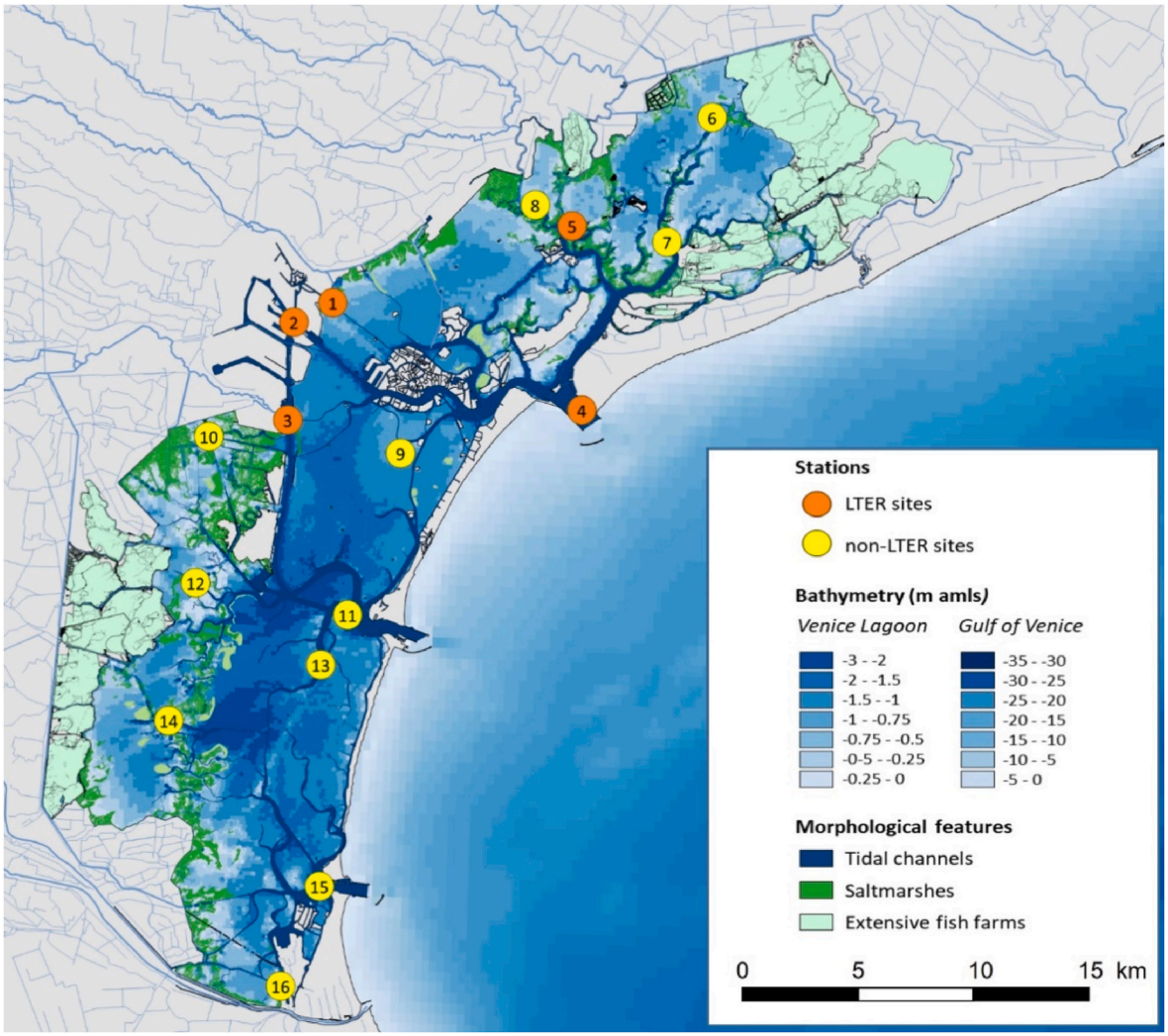

Fig. 1. Overview and bathymetry of the study area (Venice Lagoon, Italy). Orange dots: stations being part of the LTER network (LTER_EU_IT_016, http://www.lter-europe.net), yellow dots: 11 additional non-LTER stations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

for gut content analyses (study in preparation). The sampling was performed using an Apstein net with 0.4 m opening diameter and 200 μm mesh, and the samples were preserved in 96% ethanol for genetic analyses.

The Venice Lagoon is a lagoon with a semi-diurnal microtidal regime characterized by a complex system of microhabitats (Tagliapietra et al., 2009) with salinities ranging from 9 to 36, with high spatial fluctuations due to tides and river discharges as well as remarkable temporal variations in other environmental parameters (e.g., temperatures ranging from 3 to 30 °C). These, among other factors, result in substantial seasonal variability in the zooplankton community and a variety of zooplankton species, ranging from holo-to meroplanktonic organisms, and from brackish to more typically marine species (Schroeder et al., 2020). This peculiarity makes these samples an excellent application for comparing the abilities of the two COI barcodes in assessing the zooplankton community.

2.2. Molecular analysis

As a first step, a representative subsample was taken (about one-third of the total sample, see Table 1) and the ethanol was removed by centrifugation. Afterwards, the samples were rinsed with PBS (1x) and homogenized by bead-beating. The genomic DNA was extracted using the E.Z.N.A.® Mollusc DNA kit (Omega Bio-Tek) following the manufacturer's instructions and increasing the initial volume of reagents (lysis and binding buffer) provided by the kit proportionally to the sample volume. The quality and quantity of the extracted DNA were assessed with a NanoDrop 2000 Spectrophotometer (ThermoScientific).

For each sample, the same DNA extracts were utilized for the amplification of the COI fragments using two sets of primer pairs (positions: 0–319 and 345–568): the amplification was performed A) using a degenerated forward primer jdgLCO1490 (herein) in combination with a reverse internal primer designed from the mlCOIintR proposed by Leray et al. (2013) with a target length of 319 bp (hereon called P1). However, mlCOIintR by Leray et al. (2013) was modified compared to the original to match the forward internal primer mlCOIintF by interchanging the "S" with "W" nucleotides: 5′-GGRGGRTAWACWGTTCAWCCWGTWCC-3′ instead of 5′-GGRGGRTASACSGTTCASCCSGTSCC-3′, and B) using the internal primer mlCOIintF (Leray et al., 2013) together with a combination of degenerated primers dgHCOI2198 and jgHCOI2198 with a target length of 313 bp (hereon called P2) (Table 2).

The amplification of both barcodes was performed in triplicates (to reduce stochastic effects) by a two-step PCR (94 °C for 3 min followed by 5 cycles of denaturation at 94 °C for 10 s, annealing at 46 °C for 20 s, extension at 72 °C for 30 s, and followed by additional 20–30 cycles of denaturation at 94 °C for 10 s, annealing at 54 °C for 20 s and extension at 72 °C for 30 s, with a final extension at 72 °C for 3 min) with the AccuStart II PCR ToughMix (2X) polymerase (Quantabio). After the secondary PCR, where the sample-tags were bint, the library was prepared for HTS by pooling an equimolar amount (quantification performed with a Qubit™ dsDNA HS Assay Kit (ThermoFisher Scientific)) of all previously purified products (Mag-Bind® TotalPure NGS; Omega Bio-Tek). Next, the library was size-selected (range 150–400 bp) and purified using E.Z.N.A.® Size Select-IT kit (Omega Bio-Tek). Emulsion PCR was conducted using the Ion One Touch System (Life Technologies) following the manufacturer's recommendations, and DNA was bound to Ion Sphere particles (Life Technologies) for clonal amplification automatically enriched with the Ion OneTouch ES system (Life Technologies). For sequencing, the library was loaded on a 316™ chip with 650 flows in a PGM (Life Technologies).

Table 2
Primers used in this study.

primer	Sequence (5' - 3')	author	barcode
jdgLCO1490	TCAACAAAYCAYAARGAYATYGG	herein	Forward P1
mlCOIintR-mod	GGRGGRTAWACWGTTCAWCCWGTWCC	herein	Reverse P1
mlCOIintF	GGWACWGGWTGAACWGTWTAYCCYCC	Leray et al. (2013)	Forward P2
dgHCO2198	TAAACTTCAGGGTGACCAAARAAYCA	Meyer (2003)	Reverse P2
jgHCO2198	TAIACYTCIGGRTGICCRAARAAYCA	Geller et al. (2013)	Reverse P2

2.3. Bioinformatics and statistical analyses

The raw COI sequencing reads were first demultiplexed, truncated (tags and primers), sequences <200 bp were excluded, and sequences were trimmed to 320 bp using the CLC Genomics Workbench 20.0 (Qiagen). Afterwards, as suggested by Song et al. (2017), two different sequence corrections were sequentially applied. First *pollux*, a *k*-spectrum-based method that divides reads into *k*-mer lengths and generates a

Table 1
List of the 44 samples.

ID	Station	month	season	ID	station	month	season
W259.05	5	June	summer	**W261.16**	16	August	summer
W259.14	14	June	summer	**W262.09**	9	September	autumn
W260.09	9	July	summer	**W262.10**	10	September	autumn
W260.10	10	July	summer	**W262.12**	12	September	autumn
W260.11	11	July	summer	**W262.14**	14	September	autumn
W260.12	12	July	summer	**W262.16**	16	September	autumn
W260.14	14	July	summer	**W263.01**	1	October	autumn
W260.16	16	July	summer	**W263.04**	4	October	autumn
W261.01	1	August	summer	**W263.05**	5	October	autumn
W261.02	2	August	summer	**W263.06**	6	October	autumn
W261.04	4	August	summer	**W263.07**	7	October	autumn
W261.05	5	August	summer	**W263.08**	8	October	autumn
W261.06	6	August	summer	**W263.11**	11	October	autumn
W261.07	7	August	summer	**W263.14**	14	October	autumn
W261.08	8	August	summer	**W264.05**	5	November	autumn
W261.09	9	August	summer	**W264.10**	10	November	autumn
W261.10	10	August	summer	**W264.12**	12	November	autumn
W261.11	11	August	summer	**W265.07**	7	December	winter
W261.12	12	August	summer	**W265.11**	11	December	winter
W261.13	13	August	summer	**W265.12**	12	December	winter
W261.14	14	August	summer	**W265.13**	13	December	winter
W261.15	15	August	summer	**W267.09**	9	February	winter

k-mer depth profile to correct the *k*-mer profile (Marinier et al., 2015), was used to detect and remove homopolymers and indel-errors (insertion/deletion), typical for In Torrent PGM (Bragg et al., 2013) and then *fiona*, a suffix array/tree-based method that uses a suffix tree to detect and correct substitution errors (Schulz et al., 2014). The filtering of chimeric feature sequences on previously dereplicated data was performed with q2-vsearch (Rognes et al., 2016) excluding chimeras and "borderline" chimeras, and finally, the sequences were then rereplicated.

To achieve the optimal exploitation of the sequence data and therefore guarantee an assessment that is taxonomically as inclusive as possible and taking into account a series of intrinsic difficulties, e.g., complete absent or geographically not representative reference sequences, a multistep approach based on the recovery of all putative sequences was performed enabling a more reliable taxonomic assignment. A graphical representation of the bioinformatic workflow can be found in Fig. 2.

The taxonomic assignment of the quality filtered COI reads was performed by aligning them against a COI reference database of marine metazoan sequences deposited in GenBank following the query used by Schroeder et al. (2020) (downloaded 17.11.2020; including 660.477 reference sequences) with a naive LCA-assignment algorithm implemented in the MEGAN6 alignment tool (MALT) (Huson et al., 2016). First, sequences were aligned at a similarity threshold of 97%, while sequences not aligned at 97% were again aligned at 94%, and sequences not aligning at 94% were again aligned at 85% to include putative metazoan operational taxonomic unit (OTUs) that could not be recovered above the 94% similarity threshold. Reads not matching any metazoan reference sequence at a threshold of 85% were considered non-metazoan and discarded, while reads with a hit were deemed to be putative metazoan reads. Those OTUs were clustered de-novo at 97% (q2-vsearch (Rognes et al., 2016)) and only those OTUs counting at least 10 reads in the whole dataset were compared against the GenBank database with BlastN+ (Camacho et al., 2009) and only metazoan assignments were kept. Also, only OTUs that had at least a query cover of 70% and an identity of at least 90% were kept, while all others were discarded. Finally, the three datasets were pooled together: above 97% similarity threshold hits were considered as reliable OTUs, while, if a taxon was first recovered between 97% and 94% similarity threshold, it was considered as less confident and so a "cf." was added to its taxonomy, while taxa detected both with 97% and 94% similarity threshold were merged. Also, putative metazoan taxonomy with a BLASTn

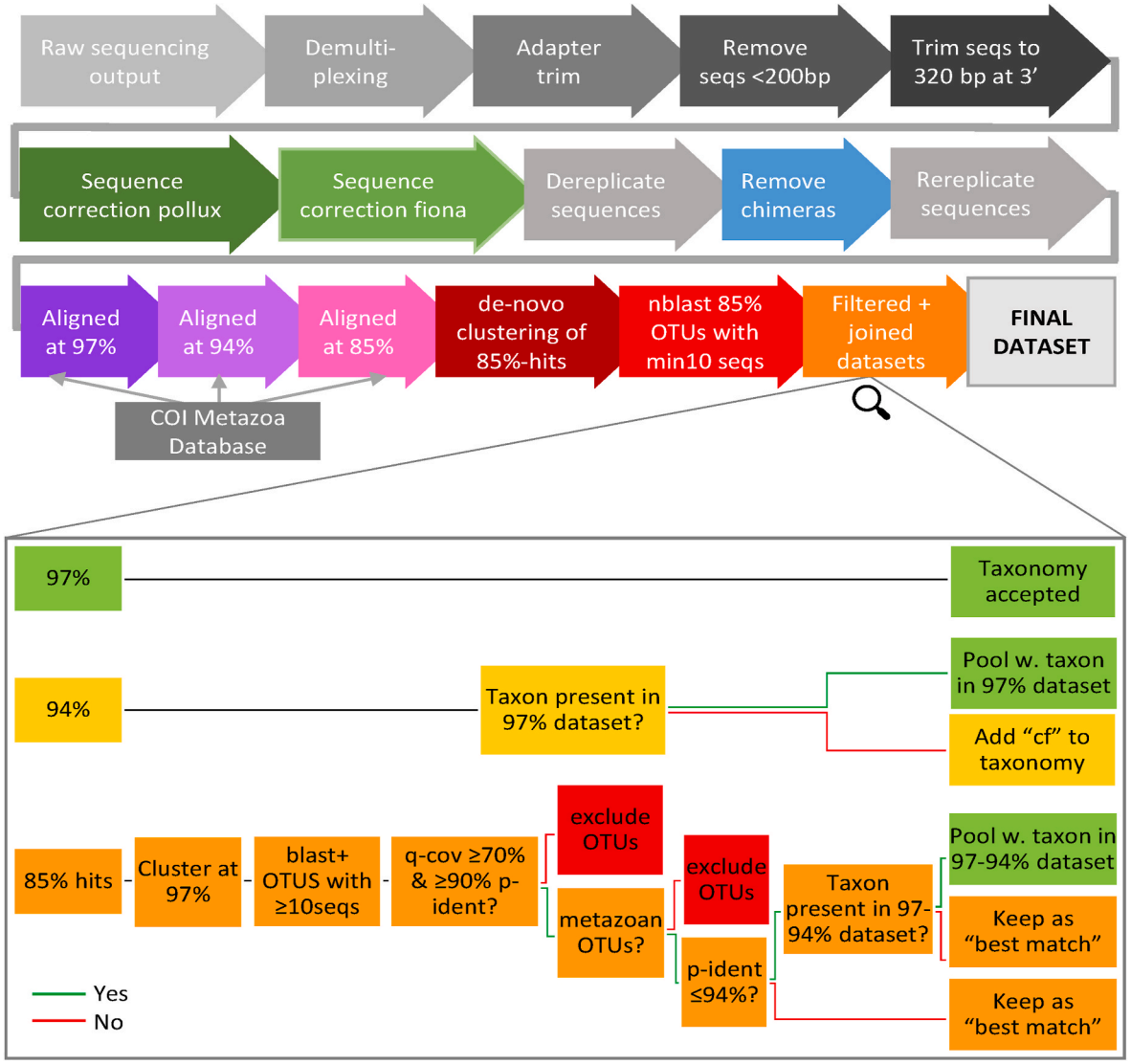

Fig. 2. Bioinformatic workflow. Green boxes: taxa considered as confidential, yellow boxes: considered as less confidential and therefore a "cf." was added to the taxonomy, orange boxes: even less confidential and therefore not considered in some analyses, red boxes: OTUs were excluded as they were not considered putative metazoan sequences. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

p-identity of at least 94% were joined, while OTUs with a BLASTn p-identity of <94% and >90% were considered as "best match" (Fig. 2). As the 2 barcodes are almost of the same length (319 bp for P1 and 313 bp for P2), the bioinformatic pipeline did not differ between the two barcodes.

Based on the final dataset, the number of sequences and OTUs that were detected at each of the steps of the taxonomic assignment and the number of OTUs detected at the different taxonomic levels were calculated and plotted in a bar chart. Species richness per sample was quantified according to the measure of the first Hill number – MOTU/taxa richness (q ¼0) using the R package iNEXT (Chao et al., 2014; Hsieh et al., 2019) (Fig. S1). All following analyses did not consider the "best match" OTUs, as they are not reliable taxonomic assignments but rather a representation of putative metazoans. To investigate the OTUs detected by the two barcodes (P1 and P2), only the OTUs assigned at species and genus level were taken into consideration. Furthermore, we repeat this analysis but aggregated it to family level. For the taxa that are not shared by both barcodes, the mean and standard deviation for each category, P1, both and P2 were calculated to investigate if general low abundances could explain missed detections. Moreover, for those taxa that were detected only with P1, the presence in two other existing P2 datasets, the dataset of the article Schroeder et al. (2020) and the extended dataset comprehending the monthly samples of all 16 stations (unpublished) was checked. The taxonomic tree was visualized and edited using the R package ggtree (Yu et al., 2017). Diversity estimates, alpha-diversity based on Shannon Wiener Index and beta-diversity based on Bray-Curtis similarities, were calculated on square-root transformed data and the correlation between the two barcodes was evaluated based on Pearson's correlations. Pearson's correlations were also calculated of relative abundances of phyla, selected classes and orders between the two barcodes from percentages of square-root transformed data.

3. Results

3.1. Taxonomic assignment

The number of raw sequences of the 44 samples was about 3.2×10^6 reads (348 bp ± 34) for P1 and about 2.3×10^6 reads (384 bp ± 35) for P2, while after trimming, sequence correction and chimera filtering, 2,267,551 reads (98.9%) and 795,027 OTUs for P1 and 2,061,251 (99.5%) and 943,983 OTUs for P2 were retained and used as input for the taxonomic assignments (Table 3).

The taxonomic assignment resulted in comparable proportions of sequences and OTUs at the different levels of assignments (Fig. 3A and B). With P1, 1,486,969 sequences (65.6%) that could be assigned to metazoans, out of which, 1,464,182 sequences (98.5%) and 177 confidential OTUs (detected at 97% threshold), 641 sequences (0.04%) and 36 OTUs assigned at 94% similarity threshold only ("cf." taxa), and 22,146 sequences (1.49%) and 20 OTUs as putative metazoans ("best match"). While with P2, 1,432,433 sequences (69.5%) could be assigned to metazoans. Regarding the different levels of assignments, 1,429,240 sequences (99.8%) and 246 OTUs were considered as confidential, 238 sequences (0.02%) and 10 OTUs were assigned at 94% similarity threshold only, and 2655 sequences (0.19%) and 28 OTUs were considered as putative metazoans ("best match") (Fig. 3A and B, Table 3). Per sample, the mean number of assigned sequences was 35×10^3 ($\pm 15 \times 10^3$) for P1 and 33×10^3 ($\pm 11 \times 10^3$) for P2. The unassigned reads, 34.4% and 30.5% for P1 and P2, respectively, belong most probably to the SAR supergroup, non (marine) metazoans, or low-quality reads.

3.2. Taxonomic richness

In terms of the number of detected taxa, the taxonomic richness showed similar results for both barcodes, with slightly higher values for P2. P1 resulted in 213 taxa (233 taxa including "best matches"), while P2 resulted in 256 taxa (284 taxa including "best matches"). Also, the taxonomic resolution shows comparable results, with the main proportion composed by species level (both 83%) and genus level (11% and 13%, respectively) assignments (Fig. 3C).

However, in terms of shared taxa, considering only species and genus level assignments and excluding "best matches", the two barcodes share 161 taxa, while 38 taxa were only detected by P1 and 85 taxa only by P2. However, the 161 shared taxa represent the vast majority, 97.3%, of the sequences, while the non-shared taxa are composed by 0.02% and 2.7% for P1 and P2, respectively. To further investigate if taxa that are not shared by both barcodes have in general low abundances, the mean for each category (P1-only, shared and P2-only taxa) were calculated (excluding absence in samples), resulting in considerably higher mean abundances for the shared taxa in comparison to the non-shared taxa: 18-fold higher than P1-only and 9-fold higher than P2-only taxa. Similarly, 112 families were identified by both barcodes, while 48 only by P2 and 18 only by P1 (Fig. 3D).

As graphically evident in Fig. 4, the taxonomic richness is comparable between the two methods as almost all phyla present a similar number of branches. The P1 barcode, however, in contrast to the P2, was also able to detect the phylum Gastrotricha, while P2 in contrast to P1 did detect the phyla Rotifera, Platyhelminthes and Ctenophora. At species level, examples of species detected with P1 only are the shrimp *Palaemon macrodactylus*, the decapod *Pilumnus spinifer*, the fish *Sprattus sprattus* and *Serranus hepatus*, the anemone *Anemonia viridis* and the order of sponges *Haplosclerida*. On the other side, taxa detected with P2 only were the gobies *Knipowitschia panizzae*, *Ninnigobius canestrinii* and *Pomatoschistus marmoratus* as well as the sea needle *Belone belone*, and the bivalves Teredinidae, *Lithophaga lithophaga* and *Pinna nobilis*. Further details regarding the species detection with P1 and P2 can be found in the supplemental table (Table S1). Within copepods, *Paracartia latisetosa, Ditrichocorycaeus anglicus*, *Oithona plumifera* and *O. davisae* and *Mesocyclops pehpeiensis* were not detected with P1 and *Candacia* (*C.* cf. *bipinnata* - "cf." assignment) was not detected using the P2 system in this dataset even though previously spotted in Schroeder et al. (2020) (Table S1).

Table 3
Stepwise change in number of sequences and OTUs through bioinformatic analyses and taxonomic assignment.

process	*step*	P1			P2		
		sequences	% of sequences	OTUs	sequences	% of sequences	OTUs
Sequence preparation	**Raw reads**	3,204,993	–	–	2,330,003	–	–
	After trim and seqs corr.	2,293,472	71.6%	815,369	2,071,489	88.9%	952,422
	After chimera filtering	2,267,551	98.9%	795,027	2,061,251	99.5%	943,983
Taxonomic assignment	**Confidential taxa**	1,464,182	98.5%	177	1,429,240	99.8%	246
	cf. taxa	641	0.04%	36	238	0.02%	10
	"best match"	22,146	1.49%	20	2655	0.19%	28
	assigned	1,486,969	65.6%	233	1,432,433	69.5%	284
	unassigned	780,582	34.4%	–	628,818	30.5%	–

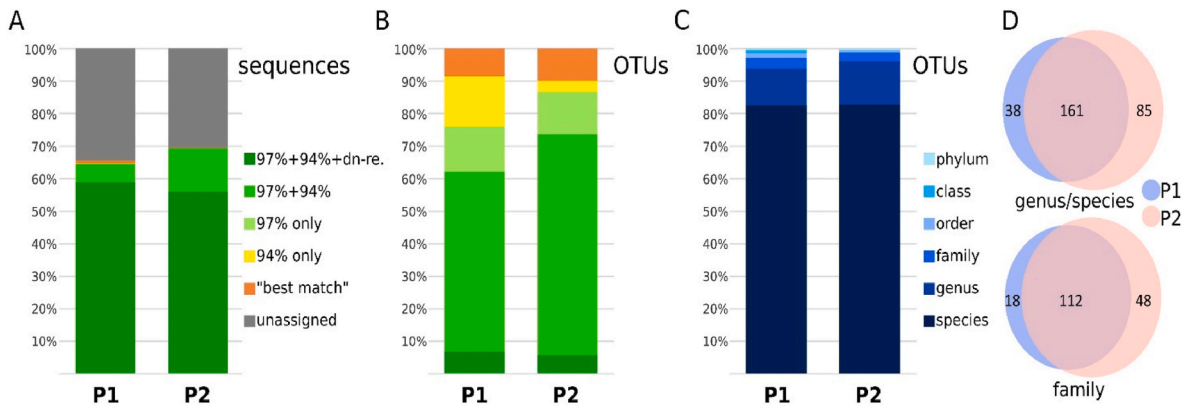

Fig. 3. A Proportion of sequences and **B** of OTUs for each step in the taxonomic assignments for each barcode, P1 and P2. **C** Proportion of OTUs for each taxonomic level of assignment. **D** Proportion of OTUs assigned at species or genus level (above) and, below, those aggregated at family level, detected with P1 only, with both barcodes, or with P2 only.

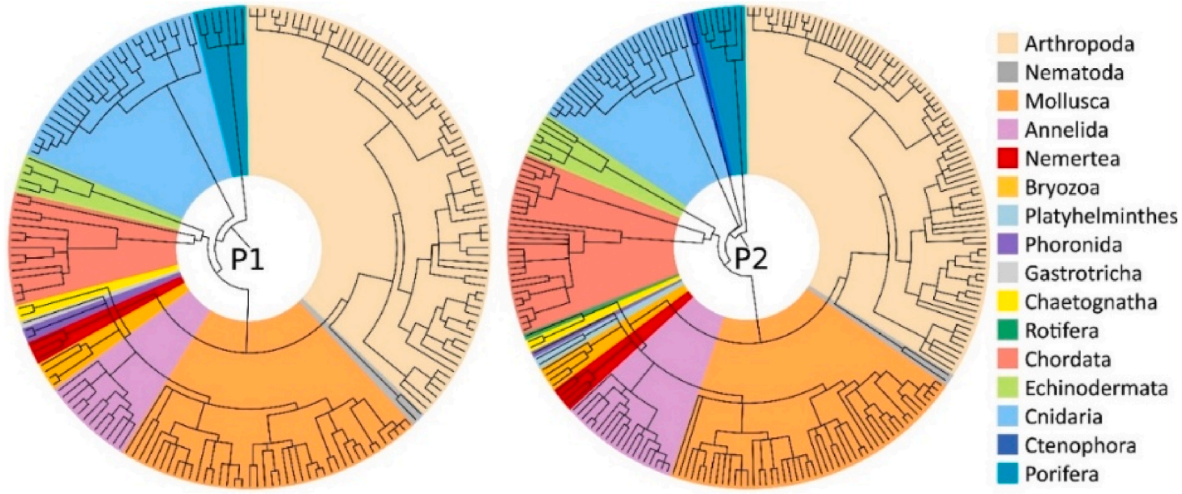

Fig. 4. Taxonomic tree representing the taxonomic richness revealed with the barcodes P1 and P2, respectively.

3.3. Relative composition

In terms of relative abundances, the two barcodes show similar patterns at class and phylum level (Fig. 5). As expected, containing the very abundant group of copepods, Arthropoda is the dominant phylum (89.4% ± 16.6 and 87.9% ± 11.8 with P1 and P2, respectively). Also, some seasonal patterns can be noticed with both barcodes, especially in the winter samples in December and January, where the zooplankton community slightly differs from the rest of the dataset as well as the higher fish abundance in same 2 samples (July station 16 and August station 2), as well as the single appearance of e.g., Hydrozoa and Ophiuroidea in the same samples. Moreover, for almost all phyla, and selected classes and orders, the relative abundance in the two barcodes is significantly correlated (Table 4). The only groups where the two barcodes are not significantly correlated are the classes Ascidiacea and Bivalvia. The most evident difference, though, is the noticeable absence of Ctenophora with the P1 barcode. Within the very abundant group of copepods, the only genus detected with both barcodes, but not showing a significant correlation in its relative abundance between the two barcodes was the genus *Pseudodiaptomus* (*P. marinus*) (Fig. 5). The relative abundance of copepod at genus level confirms that the two barcodes are highly correlated ($R = 0.95$, $p < 2.2e\text{-}16$) (Fig. 6B).

The alpha-diversity estimates based on Shannon Wiener Index of the P1 samples (2.6 ± 0.45) were not significantly different to the P2 samples (2.9 ± 0.38) (KW: $chi^2 = 43$, $df = 43$, $p = 0.47$), and significantly correlated (Pearson's correlation: $R = 0.8$, $p = 1.3e\text{-}9$) (Fig. 6A) as well as the taxa richness per sample ($R = 0.6$, $p = 5.3e\text{-}5$). The beta-diversity based on Bray-Curtis similarities was also significantly correlated (Mantel test based on Pearson's correlation: $R = 0.9$, $p = 0.001$).

4. Discussion

The efficiency and suitability of the barcode proposed by Leray et al. (2013), the forward internal primer mlCOIintF in combination with dgHCO2198 and jgHCO2198 (herein called P2), in assessing the zooplankton biodiversity was already approved by comparing it with morphological identifications in Schroeder et al. (2020) and can therefore be considered a reliable barcode for this type of studies. This study aims to test the efficiency of the second barcode proposed by Leray et al. (2013), the reverse internal primer mlCOIintR in combination jdgLCO1490, slightly modified and herein called P1, by comparing it with the already mentioned P2 barcode. In contrast to the findings of Leray et al. (2013), which found the reverse mlCOIintR primer to "poorly" perform whether it was used with LCO1490, dgLCO1490 or jgLCO1490, this study could show that once modified, the P1 barcode seems to be a helpful and reliable barcode for zooplankton metabarcoding studies.

On the one hand, the two barcodes showed slight differences: i) While the proportions of taxonomically assigned vs. non-assigned sequences are comparable, P1 has slightly fewer assignments which can be

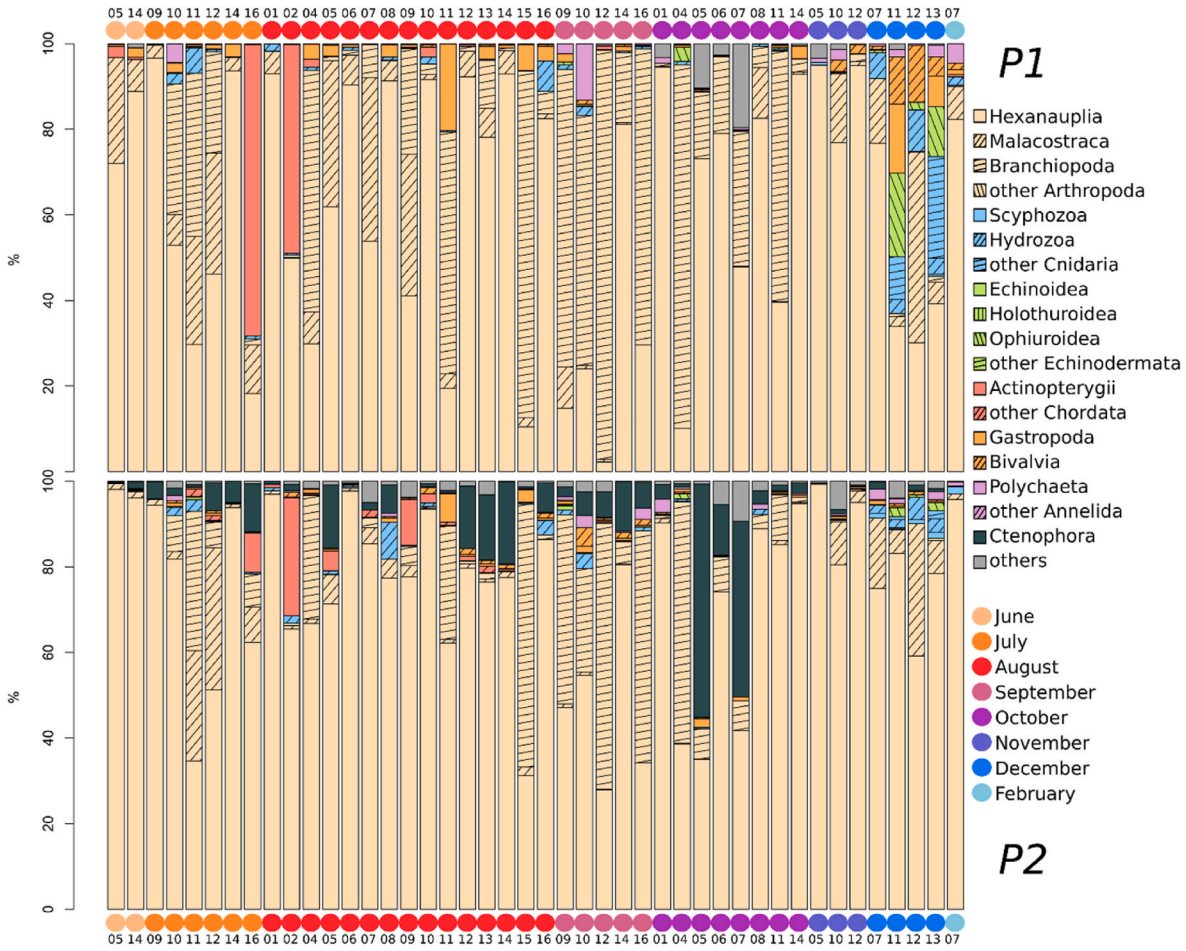

Fig. 5. Relative composition of the most abundant phyla (colours) and their classes (pattern) for each sample for P1 and P2. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

a result of the missing detected of the relatively abundant ctenophore *Mnemiopsis leidyi*. In fact, some phyla, like Ctenophora, Platyhelminthes and Rotifera were not detected by P1. Investigating on the primer binding sites of the P1 primers, no binding site for jdgLCO1490 was found for Ctenophora (tested on M. leidyi (Acc.N. NC_016117) and Boreo cucumis (Acc.N. NC_045305)), as well as for Platyhelminthes (tested on Benedenia humboldti (Acc.N. CM028216), Paragonimus westermani (Acc.N. CM017921), Schistosoma bovis (Acc.N. CM014335), Hymenolepis microstoma (Acc.N. LR215992)), while in the phylum Rotifera, checked exemplarily on Trichocera bimacula (Acc.N. JN861750) the primer jdgLCO1490 could theoretically anneal. In fact, the mitochondrial (mt) genome of *M. leidyi* strongly diverges from the "typical" mt genome as it is of minimal size (10–11 kb) (Pett et al., 2011) and reveals an extremely high evolutionary rate and for Platyhelminthes COI was shown to have a poor primer performance for the Folmer region (Vanhove et al., 2013). These findings highlight that the suitability of a barcode, also the considered "universal" COI, depends on the target species e.g., for cnidarians and ctenophores (Bucklin et al., 2011; Lindsay et al., 2015) and for groups such as Appendicularia, which in fact were neither detected with P1 nor with P2, and pelagic tunicates (Goodall-Copestake, 2017) COI has a proven difficulty in amplifying. This suggests opting for a multi-marker approach using both barcodes, P1/P2 or in combination with other genes, like 18S or 12S, instead of using the P1 as stand along marker. In fact, the benefit of using multiple markers has been shown in several studies (e.g. Stefanni et al., 2018 and Carroll et al., 2019 (both on COI and 18S–V9); Laroche et al., 2020 (COI and 18S–V4); Questel et al., 2021 (COI and 18S–V4+V9); Clarke et al., 2017 (COI, 18S–V4 and 16S); Lobo et al., 2017 (on different COI markers) or Miya et al., 2020 (COI and 12S)). ii) Both in terms of numbers of sequences and OTUs for each level of assignment ("confidential" taxa, "cf." taxa and "best matches")the two barcodes performed similarly, but with a slightly higher proportion of 94%-only assignment ("cf." assignments) for the P1 barcode. This finding could probably indicate a higher variability in P1 region compared to P2. iii) The total number of OTUs was slightly different, and some could be detected only with one or the other barcode. For example, considering only species and genus level assignments, 38 taxa were detected only with P1, 161 were found with both barcodes and 85 were detected only with P2. However, most cases are explicable by very low abundances and may result from stochastic effects during the PCR. Indeed, 16 out of the 38 taxa detected only with P1 were detected in the large dataset comprehending the monthly sampling of all 16 stations (unpublished data) or in Schroeder et al. (2020)). As on these datasets only the P2 barcode has been applied, we cannot make assumptions on taxa detected with P2 only. The largest differences in the mean relative abundance, were found in Branchiopoda (Cladocera), Decapoda, Actinopterygii, Anthozoa, Hydrozoa, Ophiuroidea, Bivalvia, Gastropoda and Porifera with in higher relative abundance in P1 compared to P2, and in Calanoida, Sessilia and Chaetognatha with lower relative abundance in P1 compared to P2.

On the other hand, the two barcodes performed very comparably: i) the proportion of the different levels of taxonomic assignment are very similar, which depends both on the taxonomic level of the reference sequence as well as on the potential assignment of the same OTUs to different taxa, consequently assigned to the next higher taxonomic level due to the LCA algorithm. ii) The two taxonomic trees showed similar

Table 4
Mean relative presence of all phyla and selected classes and orders and its standard deviation for each barcode and Pearson's correlation of the relative abundance between the two barcodes (based on square-rooted data). NA: group present in 1 single sample only (same sample for both barcodes). ***: highly significant ($p < 0.001$).

taxon	level	P1		P2		Pearson's correlation (DF = 42)			
		mean [%]	SD	mean [%]	SD	*r*	*t*	*p*	*Sign.*
Annelida	phylum	0.87	2.19	0.47	0.79	0.619	5.10	7.58e-06	***
Arthropoda	phylum	89.38	16.62	87.94	11.84	0.355	2.46	0.018	
Branchiopoda	class	19.85	28.13	10.65	18.30	0.968	24.91	<2.2e-16	***
Hexanauplia	class	60.83	29.99	72.68	21.06	0.799	8.62	7.8e-11	***
Calanoida	order	56.00	29.13	65.82	22.27	0.686	6.11	2.7e-07	***
Cyclopoida	order	0.47	1.44	0.03	0.05	0.635	5.32	3.7e-06	***
Harpacticoida	order	0.89	2.87	1.29	3.84	0.904	13.67	<2.2e-16	***
Poecilostomatoida	order	0.02	0.05	0.28	0.79	0.823	9.38	7.4e-12	***
Sessilia	order	3.46	5.68	5.25	8.85	0.939	17.71	<2.2e-16	***
Malacostraca	class	8.69	11.58	4.62	8.38	0.750	7.34	4.8e-09	***
Amphipoda	order	0.07	0.15	0.05	0.11	0.863	11.06	5.1e-14	***
Decapoda	order	8.61	11.55	4.54	8.41	0.765	7.71	1.4e-09	***
Isopoda	order	–	–	0.0001	0.001	–	–	–	–
Mysida	order	0.01	0.05	0.02	0.16	1	NA	<2.2e-16	***
Pycnogonida	class	–	–	0.0004	0.003	–	–	–	–
Bryozoa	phylum	0.04	0.14	0.02	0.06	0.953	20.40	<2.2e-16	***
Chaetognatha	phylum	0.13	0.27	1.42	1.85	0.764	7.67	1.6e-09	***
Chordata	phylum	2.90	12.45	1.51	4.59	0.768	7.77	1.2e-09	***
Actinopterygii	class	2.88	12.45	1.32	4.60	0.780	8.09	4.2e-10	***
Ascidiacea	class	0.02	0.05	0.20	0.44	0.054	0.35	0.729	
Cnidaria	phylum	2.05	4.87	1.16	1.87	0.823	9.39	7.04e-12	***
Anthozoa	class	0.78	3.84	0.06	0.29	0.985	37.26	<2.2e-16	***
Hydrozoa	class	1.24	2.15	0.98	1.65	0.776	7.97	6.2e-10	***
Scyphozoa	class	0.03	0.09	0.12	0.35	0.846	10.27	5.0e-13	***
Ctenophora	phylum	–	–	6.14	10.53	–	–	–	–
Echinodermata	phylum	0.88	3.40	0.19	0.48	0.923	15.50	<2.2e-16	***
Echinoidea	class	0.0002	0.0009	0.04	0.17	0.903	13.59	<2.2e-16	***
Ophiuroidea	class	0.88	3.40	0.15	0.45	0.986	38.99	<2.2e-16	***
Gastrotricha	phylum	0.0003	0.002	–	–	–	–	–	–
Mollusca	phylum	2.85	5.40	1.06	1.36	0.524	3.99	0.0003	***
Bivalvia	class	0.91	2.63	0.47	0.74	0.276	1.86	0.070	
Cephalopoda	class	0.0003	0.002	0.0004	0.003	0.813	9.03	2.1e-11	***
Gastropoda	class	1.94	3.97	0.58	1.09	0.746	7.30	5.5e-09	***
Scaphopoda	class	0.0003	0.002	0.0004	0.003	1	NA	<2.2e-16	***
Nematoda	phylum	0.01	0.03	0.002	0.01	0.788	8.30	2.2e-10	***
Nemertea	phylum	0.02	0.04	0.01	0.01	0.561	4.40	7.4e-05	***
Phoronida	phylum	0.002	0.01	0.002	0.01	0.546	4.23	0.00012	***
Platyhelminthes	phylum	–	–	0.01	0.02	–	–	–	–
Porifera	phylum	0.87	3.14	0.06	0.18	0.971	26.09	<2.2e-16	***
Rotifera	phylum	–	–	0.01	0.04	–	–	–	–

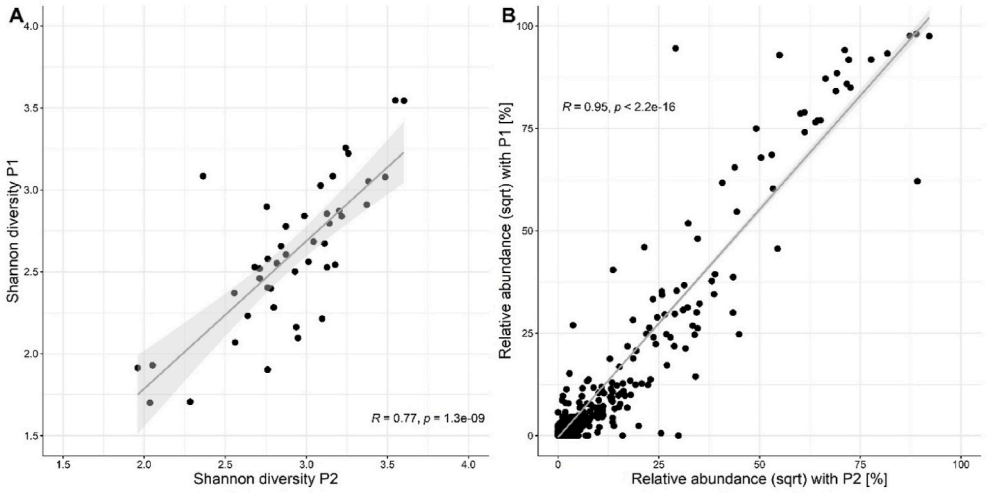

Fig. 6. A Alpha-diversity (Shannon Wiener Index H′ based on square-rooted data) and **B** Relative abundance (based on square-rooted data) of copepods merged at genus level for P1 vs. P2 and its correlation (Pearson) and confidence interval (0.95) (grey shading).

detail and proportions for each phylum, and iii) in terms of relative abundances, P1 and P2 showed similar patterns and the abundance was significantly correlated for almost all investigated groups. Apart from the phyla that were entirely missed by one barcode, only the classes Ascidiacea and Bivalvia were not significantly correlated, suggesting that for these groups P2 may be the better barcode option.

5. Conclusion

The utility of the herein investigated P1 barcode, the modified reverse internal primer mlCOIintR in combination with the forward jdgLCO1490 primer (herein), when studying zooplankton biodiversity, and probably valid for metazoans in general, can be various: first, the P1 barcode could be used in combination with P2: many hopes for better taxonomic resolution lie in the sequencing of large DNA fragments with Pacbio and ONT platforms. Indeed, a large DNA marker sequenced in its entirety dramatically enhances the ability of a correct taxonomic assignment, but our data show that one could risk losing OTUs for a limited primer match in some metazoan groups. Using P1 in combination with P2 could have the benefit to not rely on one barcode only, especially as they are singularly relatively short (319 and 313 bp) and having, therefore, the variability of the whole Folmer region without the problem of a rather large barcode (658 bp), while keeping the advantage of being based on the widely used COI gene. Also, you will overcome taxa specific primer selectivity as the missing primer binding site, e.g., in ctenophores, for the forward jdgLCO1490 primer, as you will have at least the second half of the COI, the P2, impeding to lose information on those taxa.

And third, the missing amplification success for some taxa could be advantageously in particular cases. For example, although missing the detection of potential cannibalism, excluding host sequences when studying the gut content of e.g., the invasive ctenophore *M. leidyi* (unpublished data) can be beneficial to avoid unnecessarily sequence the host's DNA.

Author contributions

A.S. contributed substantially to the study's conceptualization, data acquisition, molecular and bioinformatic analysis, and to the original draft preparation; A.P. contributed to the study's conceptualization and molecular and bioinformatic analysis; P.E. contributed to the molecular analysis; M.P. contributed to the study's conceptualization and data acquisition; and E.C. contributed to the study's conceptualization and data acquisition. All authors have reviewed and approved the final submitted manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study is based on a monitoring program within LTER-ITALY (Italian Long-Term Ecological Research Network) that we wish to acknowledge. A.S. was supported by the joint PhD Program "Environmental Life Sciences" of the University of Trieste and the University of Udine co-financed by the Institute of Marine Sciences - National Research Council ISMAR-CNR. The authors wish to thank F. Bernardi Aubry, G. Zennaro and L. Dametto for the technical support during sampling, D. Stanković for its bioinformatic support and F. Gionechetti and the whole team of the Laboratory of Applied and Comparative Genomics at the University of Trieste.

References

Amaral-Zettler, L.A., McCliment, E.A., Ducklow, H.W., Huse, S.M., 2009. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA Genes. PloS One 4, 1–9. https://doi.org/10.1371/journal.pone.0006372.

Blanco-Bercial, L., 2020. Metabarcoding analyses and seasonality of the zooplankton community at BATS. Front. Mar. Sci. 7, 1–16. https://doi.org/10.3389/fmars.2020.00173.

Bragg, L.M., Stone, G., Butler, M.K., Hugenholtz, P., Tyson, G.W., 2013. Shining a light on dark sequencing: characterising errors in Ion torrent PGM data. PLoS Comput. Biol. 9 https://doi.org/10.1371/journal.pcbi.1003031.

Brandon-Mong, G.J., Gan, H.M., Sing, K.W., Lee, P.S., Lim, P.E., Wilson, J.J., 2015. DNA metabarcoding of insects and allies: an evaluation of primers and pipelines. Bull. Entomol. Res. 105, 717–727. https://doi.org/10.1017/S0007485315000681.

Brannock, P.M., Waits, D.S., Sharma, J., Halanych, K.M., 2014. High-throughput sequencing characterizes intertidal meiofaunal communities in northern gulf of Mexico (dauphin island and mobile bay, Alabama). Biol. Bull. 227, 161–174. https://doi.org/10.1086/BBLv227n2p161.

Bucklin, A., Peijnenburg, K.T.C.A., Kosobokova, K.N., O'Brien, T.D., Blanco-Bercial, L., Cornils, A., Falkenhaug, T., Hopcroft, R.R., Hosia, A., Laakmann, S., Li, C., Martell, L., Questel, J.M., Wall-Palmer, D., Wang, M., Wiebe, P.H., Weydmann-Zwolicka, A., 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.

Bucklin, A., Steinke, D., Blanco-Bercial, L., 2011. DNA barcoding of marine metazoa. Ann. Rev. Mar. Sci. 3, 471–508. https://doi.org/10.1146/annurev-marine-120308-080950.

Bucklin, A., Yeh, H.D., Questel, J.M., Richardson, D.E., Reese, B., Copley, N.J., Wiebe, P. H., 2019. Time-series metabarcoding analysis of zooplankton diversity of the NW Atlantic continental shelf. ICES J. Mar. Sci. 76, 1162–1176. https://doi.org/10.1093/icesjms/fsz021.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinf. 10, 1–9. https://doi.org/10.1186/1471-2105-10-421.

Carroll, E.L., Gallego, R., Sewell, M.A., Zeldis, J., Ranjard, L., Ross, H.A., Tooman, L.K., O'Rorke, R., Newcomb, R.D., Constantine, R., 2019. Multi-locus DNA metabarcoding of zooplankton communities and scat reveal trophic interactions of a generalist predator. Sci. Rep. 9, 1–14. https://doi.org/10.1038/s41598-018-36478-x.

Chain, F.J.J., Brown, E.A., Macisaac, H.J., Cristescu, M.E., 2016. Metabarcoding reveals strong spatial structure and temporal turnover of zooplankton communities among marine and freshwater ports. Divers. Distrib. 22, 493–504. https://doi.org/10.1111/ddi.12427.

Chao, A., Colwell, R.K., Gotelli, N.J., Hsieh, T.C., Sander, E.L., Ma, K.H., Colwell, R.K., Ellison, A.M., 2014. Rarefaction and extrapolation with Hill numbers : a framework for sampling and estimation in species diversity studies. Ecol. Monogr. 84, 45–67. https://doi.org/10.1890/13-0133.1.

Clarke, L.J., Beard, J.M., Swadling, K.M., Deagle, B.E., 2017. Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. Ecol. Evol. 7, 873–883. https://doi.org/10.1002/ece3.2667.

Coissac, E., Riaz, T., Puillandre, N., 2012. Bioinformatic challenges for DNA metabarcoding of plants and animals. Mol. Ecol. 21, 1834–1847. https://doi.org/10.1111/j.1365-294X.2012.05550.x.

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. Mol. Ecol. Resour. 13, 851–861. https://doi.org/10.1111/1755-0998.12138.

Goodall-Copestake, W.P., 2017. One tunic but more than one barcode: evolutionary insights from dynamic mitochondrial DNA in Salpa thompsoni (Tunicata: salpida). Biol. J. Linn. Soc. 120, 637–648. https://doi.org/10.1111/bij.12915.

Günther, B., Knebelsberger, T., Neumann, H., Laakmann, S., Martínez Arbizu, P., 2018. Metabarcoding of marine environmental DNA based on mitochondrial and nuclear genes. Sci. Rep. 8, 1–13. https://doi.org/10.1038/s41598-018-32917-x.

Harvey, J.B.J., Johnson, S.B., Fisher, J.L., Peterson, W.T., Vrijenhoek, R.C., 2017. Comparison of morphological and next generation DNA sequencing methods for assessing zooplankton assemblages. J. Exp. Mar. Biol. Ecol. 487, 113–126. https://doi.org/10.1016/j.jembe.2016.12.002.

Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. B Biol. Sci. 270, 313–321. https://doi.org/10.1098/rspb.2002.2218.

Hirai, J., Shimode, S., Tsuda, A., 2013. Evaluation of ITS2-28S as a molecular marker for identification of calanoid copepods in the subtropical western North Pacific. J. Plankton Res. 35, 644–656. https://doi.org/10.1093/plankt/fbt016.

Hirai, J., Tachibana, A., Tsuda, A., 2020. Large-scale metabarcoding analysis of epipelagic and mesopelagic copepods in the Pacific. PloS One 15, 1–24. https://doi.org/10.1371/journal.pone.0233189.

Hsieh, T.C., Ma, K.H., Chao, A., 2019. iNEXT: iNterpolation and EXTrapolation for species diversity. R package version 2.0.19.

Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.J., Tappu, R., 2016. MEGAN community edition - interactive exploration and analysis of

large-scale microbiome sequencing data. PLoS Comput. Biol. 12, 1–12. https://doi.org/10.1371/journal.pcbi.1004957.
Krehenwinkel, H., Kennedy, S., Pekár, S., Gillespie, R.G., 2017. A cost-efficient and simple protocol to enrich prey DNA from extractions of predatory arthropods for large-scale gut content analysis by Illumina sequencing. Methods Ecol. Evol. 8, 126–134. https://doi.org/10.1111/2041-210X.12647.
Laroche, O., Kersten, O., Smith, C.R., Goetze, E., 2020. Environmental DNA surveys detect distinct metazoan communities across abyssal plains and seamounts in the western Clarion Clipperton Zone. Mol. Ecol. 29, 4588–4604. https://doi.org/10.1111/mec.15484.
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: application for characterizing coral reef fish gut contents. Front. Zool. 10, 1–14. https://doi.org/10.1186/1742-9994-10-34.
Lindeque, P.K., Hay, S.J., Heath, M.R., Ingvarsdottir, A., Rasmussen, J., Smerdon, G.R., Waniek, J.J., 2006. Integrating conventional microscopy and molecular analysis to analyse the abundance and distribution of four Calanus congeners in the North Atlantic. J. Plankton Res. 28, 221–238. https://doi.org/10.1093/plankt/fbi115.
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, 1–14. https://doi.org/10.1371/journal.pone.0081327.
Lindsay, D., Grossmann, M.M., Nishikawa, J., Bentlage, B., Collins, A., Lindsay, D., Matilda Grossmann, M., 2015. DNA barcoding of pelagic cnidarians: current status and future prospects. Bull. Plankton Soc. Jpn. 62, 39–43. https://doi.org/10.24763/bpsj.62.1_39.
Lobo, J., Shokralla, S., Costa, M.H., Hajibabaei, M., Costa, F.O., 2017. DNA metabarcoding for high-throughput monitoring of estuarine macrobenthic communities. Sci. Rep. 7, 1–13. https://doi.org/10.1038/s41598-017-15823-6.
Marinier, E., Brown, D.G., McConkey, B.J., 2015. Pollux: platform independent error correction of single and mixed genomes. BMC Bioinf. 16, 1–12. https://doi.org/10.1186/s12859-014-0435-6.
Medlin, L., Elwood, H.J., Stickel, S., Sogin, M.L., 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene 71, 491–499. https://doi.org/10.1016/0378-1119(88)90066-2.
Meyer, C.P., 2003. Molecular systematics of cowries (Gastropoda: cypraeidae) and diversification patterns in the tropics. Biol. J. Linn. Soc. 79, 401–459. https://doi.org/10.1046/j.1095-8312.2003.00197.x.
Miya, M., Gotoh, R.O., Sado, T., 2020. MiFish Metabarcoding: a High-Throughput Approach for Simultaneous Detection of Multiple Fish Species from Environmental DNA and Other Samples, Fisheries Science. Springer Japan. https://doi.org/10.1007/s12562-020-01461-x.
Pett, W., Ryan, J.F., Pang, K., Mullikin, J.C., Martindale, M.Q., Baxevanis, A.D., Lavrov, D.V., 2011. Leidyi : insights from mtDNA and the nuclear genome. Fossils 22, 130–142. https://doi.org/10.3109/19401736.2011.624611.Extreme.
Questel, J.M., Hopcroft, R.R., DeHart, H.M., Smoot, C.A., Kosobokova, K.N., Bucklin, A., 2021. Metabarcoding of zooplankton diversity within the Chukchi Borderland, Arctic Ocean: improved resolution from multi-gene markers and region-specific DNA databases. Mar. Biodivers. 51 https://doi.org/10.1007/s12526-020-01136-x.
Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ. https://doi.org/10.7717/peerj.2584, 2016, 1–22.
Schroeder, A., Stanković, D., Pallavicini, A., Gionechetti, F., Pansera, M., Camatti, E., 2020. DNA metabarcoding and morphological analysis - assessment of zooplankton biodiversity in transitional waters. Mar. Environ. Res. 160 https://doi.org/10.1016/j.marenvres.2020.104946.
Schulz, M.H., Weese, D., Holtgrewe, M., Dimitrova, V., Niu, S., Reinert, K., Richard, H., 2014. Fiona: a parallel and automatic strategy for read error correction. Bioinformatics 30, 356–363. https://doi.org/10.1093/bioinformatics/btu440.
Shokralla, S., Porter, T.M., Gibson, J.F., Dobosz, R., Janzen, D.H., Hallwachs, W., Golding, G.B., Hajibabaei, M., 2015. Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. Sci. Rep. 5 https://doi.org/10.1038/srep09687.
Sommer, S.A., Van Woudenberg, L., Lenz, P.H., Cepeda, G., Goetze, E., 2017. Vertical gradients in species richness and community composition across the twilight zone in the North Pacific Subtropical Gyre. Mol. Ecol. 26, 6136–6156. https://doi.org/10.1111/mec.14286.
Song, L., Huang, W., Kang, J., Huang, Y., Ren, H., Ding, K., 2017. Comparison of error correction algorithms for Ion Torrent PGM data: application to hepatitis B virus. Sci. Rep. 7, 1–11. https://doi.org/10.1038/s41598-017-08139-y.
Stefanni, S., Stanković, D., Borme, D., de Olazabal, A., Juretić, T., Pallavicini, A., Tirelli, V., 2018. Multi-marker metabarcoding approach to study mesozooplankton at basin scale. Sci. Rep. 8, 1–13. https://doi.org/10.1038/s41598-018-30157-7.
Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., Willerslev, E., 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. Mol. Ecol. 21, 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x.
Tagliapietra, D., Sigovini, M., Ghirardini, A.V., 2009. A review of terms and definitions to categorise estuaries, lagoons and associated environments. Mar. Freshw. Res. 60, 497–509. https://doi.org/10.1071/MF08088.
Vanhove, M.P.M., Tessens, B., Schoelinck, C., Jondelius, U., Littlewood, D.T.J., Artois, T., Huyse, T., 2013. Problematic barcoding in flatworms: a case-study on monogeneans and rhabdocoels (Platyhelminthes). ZooKeys 365, 355–379. https://doi.org/10.3897/zookeys.365.5776.
Wang, X., Hua, F., Wang, L., Wilcove, D.S., Yu, D.W., 2019. The biodiversity benefit of native forests and mixed-species plantations over monoculture plantations. Divers. Distrib. 25, 1721–1735. https://doi.org/10.1111/ddi.12972.
Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.Y., 2017. Ggtree: an R Package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8, 28–36. https://doi.org/10.1111/2041-210X.12628.
Zaiko, A., Samuiloviene, A., Ardura, A., Garcia-Vazquez, E., 2015. Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. Mar. Pollut. Bull. 100, 53–59. https://doi.org/10.1016/j.marpolbul.2015.09.030.