Mar Biol (2016) 163:149
DOI 10.1007/s00227-016-2920-0

ORIGINAL PAPER

Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy

David Abad[1] · Aitor Albaina[1,2] · Mikel Aguirre[1] · Aitor Laza-Martínez[3] · Ibon Uriarte[4] · Arantza Iriarte[4] · Fernando Villate[5] · Andone Estonba[1]

Received: 10 March 2016 / Accepted: 22 May 2016 / Published online: 14 June 2016

Abstract Metabarcoding is becoming an increasingly valuable alternative approach to biodiversity assessment, due to the combination of extreme sensitivity and potential for the highest taxonomic resolution in a cost- and time-effective methodology. To evaluate the capacity of metabarcoding for estuarine plankton monitoring, a comparison between the results obtained with this approach were compared with those based on traditional taxonomic analysis (microscopy). Database incompleteness, one of the main limitations of metabarcoding, was somewhat overcome by the addition of DNA sequences for local species, which increased the taxonomic assignment success from 23.7 to 50.5 %. When the communities were studied along with environmental variables, similar spatial and temporal trends of taxonomic diversity were observed for metabarcoding and microscopic studies of zooplankton, but not for phytoplankton. This is most likely attributable to the lack of representative sequences for phytoplankton species in current databases. In addition, there was high correspondence in community composition when comparing abundances estimated from metabarcoding and microscopy, suggesting semiquantitative potential for metabarcoding. Furthermore, metabarcoding allowed the detection and identification of two non-indigenous species (NIS) found in the study area at abundances hardly detectable by microscopy. Overall, our results indicate that metabarcoding is a powerful approach with excellent possibilities for use in plankton monitoring, early detection of NIS and plankton biodiversity shifts.

Responsible Editor: T. Reusch.

Reviewed by Undisclosed experts.

Electronic supplementary material The online version of this article (doi:10.1007/s00227-016-2920-0) contains supplementary material, which is available to authorized users.

✉ David Abad
david.abad@ehu.eus

1 Genomic Resources Group, Department of Genetics, Physical Anthropology and Animal Physiology, University of the Basque Country, UPV/EHU, Leioa, Spain

2 Environmental Studies Centre (CEA), Vitoria-Gasteiz City Council, Vitoria-Gasteiz, Spain

3 Phytoplankton Group, Department of Plant Biology and Ecology, University of the Basque Country, UPV/EHU, Leioa, Spain

4 Zooplankton Group, Laboratory of Ecology of Zooplankton, Department of Plant Biology and Ecology, University of the Basque Country, UPV/EHU, Gasteiz, Spain

5 Zooplankton Group, Laboratory of Ecology of Zooplankton, Department of Plant Biology and Ecology, University of the Basque Country, UPV/EHU, Leioa, Spain

Introduction

Plankton communities are essential for aquatic ecosystem functioning, playing a crucial role in food webs and biogeochemical cycles (Ward et al. 2012). Furthermore, due to their rapid response to environmental variation, planktonic organisms have been used as indicators of ecosystem change for monitoring purposes (e.g., Taylor et al. 2002). These features highlight the interest of studying plankton community structure, biodiversity and responses to environmental factors.

On the one hand, phytoplankton biomass has been mainly estimated using chlorophyll *a* (Chl-*a*) as a proxy (e.g., Bricker et al. 2003). It has also been recently included in the Marine Strategy Framework Directive

(MSFDe2008/56/EC) (e.g., Ferreira et al. 2011; Uriarte et al. 2015). However, biodiversity of phytoplankton is very difficult to estimate and monitoring has usually been limited to certain groups (e.g., Amorim Visco et al. 2015; Eiler et al. 2013). On the other hand, and despite its ecological importance, zooplankton biodiversity is not yet included in European marine environmental policies (Borja et al. 2011). This is mostly related to constraints of microscopy-based identification, as the identification of morphological characteristics by light microscopy is complicated and time-consuming and requires wide expertise due to morphological similarities between species and restricted diagnostic features (e.g., Lindeque et al. 2006). Microscopy-based biodiversity assessment is also subjected to an unpredictable, but probably significant, bias due to the presence of cryptic species (e.g., Chen and Hare 2008).

The advent of next-generation sequencing (NGS) technologies has provided an alternative to overcome issues associated with microscopy-based monitoring (Baird and Hajibabaei 2012). These NGS technologies allow the use of metabarcoding, where a short DNA region is sequenced for a whole community sample and the obtained sequences are used to measure biodiversity at an affordable cost (e.g., Lindeque et al. 2013; Hirai et al. 2015). The reported high sensitivity of NGS (Zhan et al. 2013) makes this technique ideal for the detection of rare taxa. Therefore, NGS-based metabarcoding generates large amounts of biodiversity information and is capable of identifying species at any life stage in taxonomically complex assemblages (Comtet et al. 2015), including the precise identification of cryptic species, and those overlooked by traditional methods because they are either too fragile or too small. A particularly useful application of metabarcoding is the early detection of non-indigenous species (NIS), which are a cause of great concern in monitoring vulnerable ecosystems, such as estuaries.

An essential requirement for metabarcoding studies is a deep and curated database of reference DNA sequences for identified specimens. Currently, this necessity represents one of the main drawbacks, since some groups of organisms have none or very few publicly available sequences. In addition, it is also known that the copy number variation (CNV) associated with rDNA genes could affect the abundance estimates when using metabarcoding (Kembel et al. 2012), explaining the lack of correlation between this approach and microscopy in some cases (e.g., Stoeck et al. 2014); other technical biases introduced during DNA extraction (Roh et al. 2006) or the PCR amplification step (Gonzalez et al. 2012) influence these estimates as well.

Recent studies have applied metabarcoding to characterize different groups of organisms in aquatic ecosystems, including bacteria (e.g., Herlemann et al. 2011; Gilbert et al. 2012), protozoa (Bachy et al. 2013; Massana et al. 2015), microalgae (Amorim Visco et al. 2015; Eiler et al. 2013) and zooplankton (e.g., Lindeque et al. 2013; Hirai et al. 2015). These studies have revealed previously hidden taxonomic richness, including rare species and parasites (Lindeque et al. 2013; Logares et al. 2014), and provided much higher biodiversity estimates than traditional methods (Bachy et al. 2013), highlighting the sensitivity and higher taxonomic resolution of NGS-based metabarcoding. Moreover, its application for NIS monitoring has been recently demonstrated (Zaiko et al. 2015c).

In contrast to many previous studies, which focused on a particular size fraction and/or limited number of taxonomic groups, we analyzed the entire eukaryotic plankton community (0.22–20, 20–200 and >200 μm size fractions) of an estuary. We selected the V9 region of the 18S rDNA gene (18S V9) primarily because of its broad amplification range among eukaryotes (de Vargas et al. 2015; Albaina et al. 2016a), but also because the Earth Microbiome Project (EMP; http://www.earthmicrobiome.org) designed a protocol for Illumina platforms that has markedly increased sequencing depth compared to the previously dominant NGS technology (i.e., Roche's 454) in the metabarcoding field (Mahé et al. 2015). Metabarcoding using the 18S V9 has recently allowed the characterization of marine plankton biodiversity within the Tara Oceans (http://www.embl.de/tara-oceans/start/; Massana et al. 2015) and Biomarks initiatives (http://www.biomarks.eu/; de Vargas et al. 2015).

The main objective of this study was to compare the results of plankton community taxonomic composition based on metabarcoding versus microscopy in order to assess the usefulness of metabarcoding for estuarine plankton monitoring. We also used the 18S V9 to analyze the community spatiotemporal structure in relation to environmental parameters. Moreover, we studied the effects of completeness of the reference database on taxon assignment by adding local species, and the sensitivity of metabarcoding for NIS detection.

Methods

Study area

The estuary of Bilbao is located in the southeast part of the Bay of Biscay (within 43°23′N–43°14′N and 3°07′W–2°55′W). It is a small (~23 km long), narrow (25–270 m), shallow (0.5 to >10 m) and highly stratified channel that crosses urban and industrial settlements and drains into a wide coastal embayment (Uriarte et al. 2014). It was one of the most polluted estuaries in Europe, but since 1979 it has undergone a significant improvement of water/sediment quality and recovering of biodiversity (Borja et al. 2006; Villate et al. 2013). This transition has allowed

the recolonization by a mixture of neritic and estuarine species (Albaina et al. 2009; Uriarte et al. 2015). Among them, there are NIS such as the copepods *Acartia tonsa* (Calanoida, Acartiidae), which was first described in this estuary in 2001 and became dominant the following year (displacing congeneric species; Aravena et al. 2009), and *Pseudodiaptomus marinus* (Calanoida, Pseudodiaptomidae), which was recently collected for the first time in the estuary of Bilbao (Albaina et al. 2016b) and whose effect on the community cannot yet be predicted.

Sampling

Sampling was carried out in summer (June and July) and autumn (September and October) of 2013 from water at salinities 30 and 35 during neap tides. Salinity, temperature (°C), dissolved oxygen (DO, mg/L) and pH at each sampling point were measured with a YSI 556 MPS multi-parameter probe. Water transparency was measured with a Secchi disk. Chl-*a* concentrations (mg/L) were calculated from spectrophotometric measurements on acetone extracts, following the monochromatic method with acidification (Jeffrey and Mantoura 1997). In addition, precipitation (mL/m^2) data were provided by the Hydrometeorology Service of the Regional Council of Bizkaia.

To obtain the 0.22–20 and 20–200 μm size fractions, a Niskin bottle was used to collect 10 L of water at each salinity (sampling depth depended on the water mass location). Samples were filtered through a 20-μm mesh (Millipore Nylon Nets) and, consecutively, approximately 1 L was filtered through a 0.22-μm Durapore membrane (Millipore) using a Kitasato flask and a vacuum pump (Millipore). Meshes and membranes were kept in cryogenic tubes and frozen at −80 °C until further use for metabarcoding. For the microscopy analysis of phytoplankton samples, a 250-mL bottle with 1 mL of Lugol was filled directly with water from the Niskin bottle.

For the >200 μm size fraction, samples were collected at about 3 m depth by ~5-min horizontal tows of a 200-μm plankton net (mouth diameter 0.25 m) equipped with a Hydrobios flow meter. Once in the laboratory, each sample was divided using a plankton splitter. One half was kept in buffered formalin (4 %) for microscopy analysis. One quarter was filtered through a 180-μm mesh (Millipore Nylon Nets) and preserved in ethanol for DNA barcoding analysis of some selected species. Finally, the remaining quarter was also filtered through a 180-μm mesh, kept in a falcon tube and frozen at −80 °C until further use for metabarcoding.

Microscopy

Both Lugol-fixed (non-filtered) and formalin-fixed (>200 μm) plankton samples were identified at the lowest taxonomic resolution possible. The phytoplankton community was characterized from the Lugol-fixed bottle samples through the Utermöhl or sedimentation method (Edler and Elbrächter 2010). Additionally, living subsamples were observed under light microscopy on the day of sampling to determine the presence of species difficult to identify after fixation. Fixed phytoplankton cells from the settled samples were identified and counted under a Nikon Diaphot TMD (Nikon Corporation, Tokyo, Japan) inverted microscope. Heterotrophic dinoflagellates and some non-photosynthetic nanoflagellates such as kathablepharids and choanoflagellates were included in the analyses. For salinities 30 and 35, 10 mL and 50 mL (to obtain sufficient abundances) were settled, respectively. The entire chamber area was examined at 100× magnification, and transects were performed at 200×–400×. A minimum of 300 cells (average of 593 cells per sample) were counted. Species biomass was calculated using formulas from the Baltic Marine Environment Protection Commission (HELCOM; Olenina et al. 2006).

For zooplankton, identification of the >200-μm sample was carried out under an inverted stereo-microscope. A minimum of 100 individuals of the most abundant taxa were counted before finishing subsampling or, if not possible, the whole sample was examined. Absolute and relative abundances were computed for copepods. Biomass was also estimated for copepods based on the average size of individuals (http://copepodes.obs-banyuls.fr), assuming the carbon content was 40 % of the dry weight (Båmstedt 1986) and following the formula from Gaudy and Boucher (1983).

DNA extraction

A modified salt protocol (Aljanabi and Martinez 1997) was used to extract the DNA from the 20 and 200 μm size fractions. Meshes were defrosted and cleaned in a falcon tube with distilled water injected through a wash bottle to remove any possible attached organisms or DNA-containing fragments. The filters were held with each tube lid and centrifuged at 4000 rpm for 30 min to create a pellet. The mesh and supernatant water were then carefully removed. Proteinase K (20 mg/mL) digestion was conducted on the pelleted organisms overnight, and the extraction was continued according to the protocol. Samples from the 0.22-μm mesh were extracted following the instructions of the MOBIO PowerSoil® DNA Isolation Kit.

Extracted DNA was first quantified by spectrophotometry using Nanodrop (ND-1000; Thermo Scientific) and then by fluorometry using Qubit 1.0 (Thermo Scientific) to determine the amount of double-stranded DNA. DNA was stored at −20 °C until further processing.

Table 1 GenBank accession numbers of the 18S V9 sequences generated in this study

Species	GenBank accession numbers		
Acartia tonsa	KP768152	KP768153	KP768154
Acartia clausi	KR919781	KR919782	
Calanipeda aquaedulcis	KP768157	KP768158	
Euterpina acutifrons	KR919779	KR919780	
Pseudodiaptomus marinus	KP768155	KP768156	
Evadne nordmanni	KR919787		
Evadne spinifera	KR919783	KR919784	
Podon spp.	KR919785	KR919786	

DNA barcoding

After checking the microscopy results and previous studies on the area (Albaina et al. 2009; Uriarte et al. 2015), we noted the absence of some key/abundant species (e.g., *Pseudodiaptomus marinus*) of the estuary of Bilbao in publicly available databases. In order not to miss these taxa in our analysis of community composition by metabarcoding, we generated 18S V9 reference sequences (by Sanger sequencing) to include them in the database. Five copepods and three cladocerans species were isolated from the ethanol-preserved sample splits (Table 1; GenBank accession numbers KP768152–KP768156 and KR919779–KR919787). For *Acartia clausi*, *Euterpina acutifrons* and the cladocerans, 10 individuals were pooled in each extraction tube. We used the EMP primers 1391f (5′-GTACACACCGCC CGTC-3′) and EukBr (5′-TGATCCTTCTGCAGGTTCAC CTAC-3′), based on Stoeck et al. (2010), for the amplification of the 18S V9 (~150-bp fragment). Polymerase chain reactions (PCRs) were performed in a 25-µL volume containing 7.5 µL of distilled water, 5.4 µL of buffer (5×), 2.7 µL of $MgCl_2$ (25 mM), 2.7 µL of dNTPs (10 µM), 2.7 µL of each primer (10 µM), 0.3 µL of Taq polymerase (Promega) and 1 µL of template DNA. PCR cycling included initial denaturation at 92 °C for 3 min, followed by 30 cycles of 45-s denaturation at 92 °C, 1-min annealing at 57 °C and 1.5-min extension at 72 °C. A final extension step was performed at 72 °C for 10 min. The purified PCR products were sequenced in both directions on an ABI 31309 capillary electrophoresis analyzer with ABI BigDye Terminator version 3.1 chemistry (Applied Biosystems) at the SGIKER (UPV/EHU).

Preparation of the custom reference databases

In order to illustrate the effects of reference database completeness, we studied the assignation rate in two versions of Silva clustered at 99 % identity (http://www.arb-silva.de/; Quast et al. 2013) with and without the inclusion of the herein generated 18S V9 sequences (Table 1). Silva 111 and 119 were the last two available releases at the time the study was performed; the difference in over 1 million reference sequences is explained by their publication dates (July 2012 and 2014, respectively). To further prove the effect of adding sequences corresponding to key/local species, we also included 9 copepod species inhabiting northeast Atlantic (NEA) neritic waters (Laakmann et al. 2013; ESM_1) and representative sequences from *Oithona brevicornis*, *O. nana*, *O. similis* and *Oncaea media* (GenBank accession numbers JF288757, HQ008734, KF153700 and AM114421; ESM_1), for which we could not find other reliable sources.

Metabarcoding, OTU definition and taxonomic assignation

Sequencing of the 18S V9 region was carried out at the Argonne National Laboratory (Lemont, IL, USA) following the EMP protocols and using Illumina MiSeq 2 × 150 bp. Raw reads were trimmed with Sickle version 1.33 (Joshi and Fass 2011), using a quality threshold of 20. For paired-end merging, Pear version 0.9.5 (Zhang et al. 2014) was used with a minimum overlapping of 15 bp and a cutoff p value of 0.01. The barcodes from the sequences discarded in the previous steps were removed by fastq-barcode.pl (Smith, 2012). Chimeras were removed with UCHIME (Edgar et al. 2011), using a reference-based chimeric detection against Silva 119 custom.

Merged reads were processed using Qiime version 1.9 (Caporaso et al. 2010): Sequences were clustered into operational taxonomic units (OTUs) with UCLUST (Edgar 2010), using both *de novo* and closed reference approaches with 97 and 99 % identity thresholds. The *de novo* approach groups sequences based on sequence identity (Navas-Molina et al. 2013), and taxonomy is then assigned to the obtained representative sequences with BLAST (Altschul et al. 1990). The closed reference approach matches sequences to an existing database of reference sequences (Silva, in our case); if a sequence fails to match the database, it is discarded (Navas-Molina et al. 2013). This approach assigns the taxonomy with UCLUST (Edgar 2010).

Statistical analysis

Canonical correspondence analysis (CCA) of the OTUs showing >1 % relative abundance were performed using CANOCO version 4.5 (ter Braak and Smilauer 2002) to investigate the relationship between taxon abundances from metabarcoding versus microscopy (0.22–200 and >200 µm size fractions) in relation to samples and measured environmental variables. The 18S V9 sequences from the 0.22- and

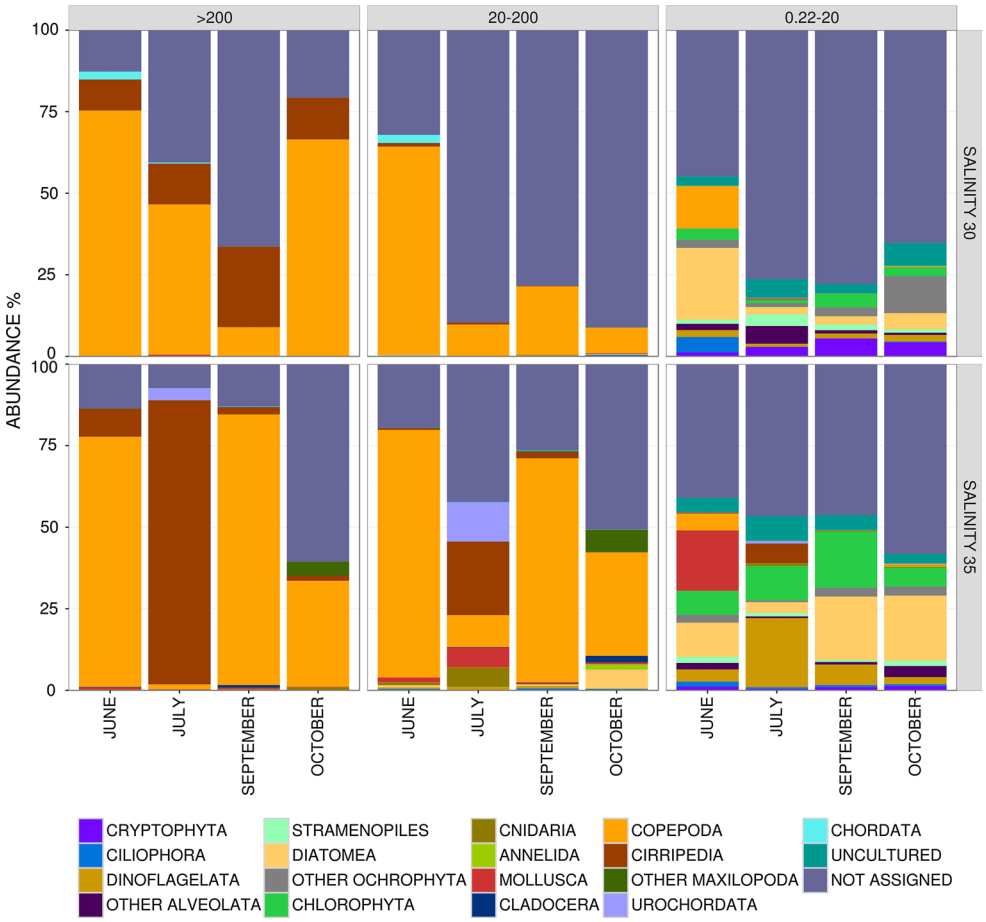

Fig. 1 Proportion of taxonomic groups in each sample based on the metabarcoding approach. A total of 17 taxonomic groups (>1 % abundance) are shown. Samples are arranged by salinity (35 and 30) and plankton size fraction (0.22–20, 20–200 and >200 μm)

20-μm filters were merged (after rarefaction) so that the comparison with phytoplankton microscopy-based results (Utermöhl) can be possible. All the CCAs were constructed using relative abundance data, with square root transformations used to normalize the samples.

Spearman's rank correlation coefficient (ρ) was calculated for the comparison of relative abundances retrieved by microscopy (both counts and biomass measurements) and metabarcoding (18S V9 reads) using cor.test in R (R Core Team 2015); the correlations were limited to taxa uncovered by both methods (ESM_2).

Results

Metabarcoding

Only 0.24 % of the reads were discarded due to poor quality. Of the remaining, 89.89 % were successfully merged. In all, 0.02 % of the reads were eliminated due to their putative chimeric nature. Once the OTU table was constructed, 229 singletons were discarded from further analysis. Finally, the rarefaction curves (ESM_3) showed that the plateau was reached at 3000 reads in most of the samples.

The closed reference approach produced 1174 and 831 OTUs for the 97 and 99 % identity thresholds, respectively. The comparison of the different databases (Table 3) showed that "Silva 119 custom" had the highest proportion of taxonomic assignments with 53.5 %, whereas "Silva 111" only reached 23.7 %. Although the assignment increased to 66.4 % with the 97 % identity threshold, the 99 % (as in Albaina et al. 2016a) was considered to be optimal, given its taxonomic resolution and stringency for the 18S V9 region (below the 1 % machine error rate; Quail et al. 2012).

When using the de novo approach instead, a taxonomic correspondence was obtained for approximately 100 % of the reads, yielding a total of 2139 and 2318 OTUs for 97 and 99 % identity thresholds, respectively. However, due to the low identity values of the taxonomy obtained for the sequences that were unassigned in the closed reference, we focused on the results produced by this latter method.

The resulting 831 OTUs (produced by the closed reference at 99 % similarity) were classified into 17 taxonomic groups (those representing less than 1 % abundance were not plotted), uncultured taxa and unassigned reads (Fig. 1). Interestingly, the percentage of reads for which the taxonomy was assigned was higher for salinity 35 (64.8 %)

Table 2 Percentage of sequences that were assigned to taxonomy using four different databases

	Silva 111			Silva 111 custom			Silva 119			Silva 119 custom		
	0.20–20	20–200	>200	0.20–20	20–200	>200	0.20–20	20–200	>200	0.20–20	20–200	>200
June 30	28.21	5.25	14.46	40.96	67.99	87.34	55.60	5.63	14.67	55.69	68.12	87.34
June 35	50.71	17.38	24.26	55.62	80.59	86.81	55.26	22.96	48.81	60.09	80.52	86.49
July 30	42.38	1.16	13.69	42.42	10.79	59.68	23.95	0.98	14.85	23.99	10.36	59.47
July 35	46.03	35.28	88.17	46.05	43.39	89.68	53.61	51.20	91.24	53.62	57.81	92.64
Sept 30	22.53	0.75	24.97	22.57	21.67	33.7	22.78	6.55	29.91	22.80	21.68	33.71
Sept 35	38.21	21.30	10.58	38.23	72.84	86.58	54.06	24.55	12.81	54.08	73.71	87.13
Octo 30	30.36	2.31	13.35	30.63	10.16	79.31	35.11	2.44	76.93	35.14	8.85	79.31
Octo 35	25.05	6.63	6.54	25.48	39.69	35.48	42.18	16.38	19.58	42.59	49.41	39.62
Mean	35.44	11.26	24.5	37.75	43.39	69.82	42.82	16.34	38.60	43.50	46.31	70.71
Total	23.73			50.32			32.58			53.51		

Similarity threshold was set at 99 %. Total assignment percentage for each database is shown along with those for each specific size fraction (0.22–20, 20–200 and >200 μm), salinity (30 and 35) and sampling month (June–October)

than for 30 (42.2 %). In general, the percentage of unassigned reads was lower as the size fraction increased (Table 2): 56.5, 53.7 and 29.3 % for the 0.22–20, 20–200 and >200 μm, respectively.

Maxillopoda (mainly copepods and barnacles) predominated at the 20–200 and >200 μm size fractions, while a more diverse assemblage characterized the 0.22–20 μm size fraction (Fig. 1). Copepods represented 2.3, 36 and 48.6 % of the OTUs, while phytoplankton groups (e.g., Bacillariophyceae, Dinophyceae, Cryptophyceae) were 26.1, 1.5 and <0.1 % of each size fraction (0.22–20, 20–200 and >200 μm, respectively).

Metabarcoding versus microscopy

The microscopic analysis identified 180 taxa for the Utermöhl method and 100 for the >200-μm zooplankton net. When the resulting taxa identified by metabarcoding and microscopy were compared, 44 of them were found in common. However, if the comparison was performed for the taxa with a >1 % abundance in at least one of the samples (Table 3), only eleven species (three from phytoplankton and eight from zooplankton) were detected as abundant by both methodologies. Twelve taxa (six from phytoplankton and another six from zooplankton) were detected as abundant by microscopy, but not as abundant in metabarcoding. Finally, two species (the diatom *Skeletonema menzellii* and the copepod *Centropages hamatus*) were detected as abundant by metabarcoding, but not as abundant in microscopy.

While microscopy was unable to identify below genus level in *Thalassiosira*, metabarcoding was able to distinguish congeneric species (e.g., *T. allenii*, *T. delicatula*); the same occurred with the genus *Skeletonema* (*S. pseudocostatum*). Conversely, the microscopy-based analysis reported several species that were not identified using metabarcoding (e.g., *Apedinella radians, Teleaulax gracilis, Teleaulax minuta, Oithona davisae*).

The same spatial (salinity) and temporal (date, seasonal variation) trends were described for those species detected by both methodologies in the >200 μm size fraction (Fig. 2a, b): While the higher dissolved oxygen (DO) and water transparency (Secchi disk) values were associated with salinity, the highest precipitation could be associated with date (summer-to-autumn transition). The concentration of chlorophyll a (Chl-*a*) decreased with the advance of the season. However, in the 0.22–200 μm size fraction neither approach identified a temporal pattern (Fig. 2c, d); a spatial pattern was discriminated only by microscopy (Fig. 2d).

When comparing the relative abundance of all taxa within each particular sample obtained by both approaches (Table 4), significant correlations were reported in most cases. No difference was found between comparisons against microscopy-based counts or biomass (Table 4).

Non-indigenous species (NIS)

We compared the performance of metabarcoding and microscopy to detect two NIS: *A. tonsa* and *P. marinus* in the >200 μm size fraction. While similar relative abundances were found for *A. tonsa* in salinity 30 by both approaches (Fig. 3a), it was only detected by metabarcoding in salinity 35 sample (Fig. 3b). Regarding *P. marinus*, metabarcoding was capable of detecting the species in all the samples, while its presence was detected by microscopy only in two of them (Fig. 3c, d). Finally, negative controls and extraction blanks showed no sequences corresponding to these two organisms, giving further support to the herein reported data.

Table 3 List of most abundant taxa from metabarcoding and microscopy

Metabarcoding	Microscopy	Acronym	Metabarcoding	Microscopy	Acronym
	Centric diatoms	CENT		*Scenedesmus* spp.	SCEN*
	Chaetoceros curvisetus	CCUR		*Pyramimonas* spp.	PYRA
	Chaetoceros debilis	CDEB	*Ostreococcus tauri*		OTAU
	Chaetoceros socialis	CSOC	*Ulva intestinalis*		UINT
	Chaetoceros spp.	CHAE	Uncultured phytoplankton		UPHY
	Conticribra weissflogii	CWEI		Choanoflagellates	CHOA
Cyclotella choctawhatcheeana		CCHO		*Leucocryptos* spp.	LEUC
	Cyclotella meneghiniana	CMEN*	*Strombidium basimorphum*		SBAS
Leptocylindrus hargravesii		LHAR	*Chelophyes appendiculata*		CAPP
	L. danicus/hargravesii	LDAN		Unidentified Siphonophora	USIP
	Leptocylindrus aporus	LAPO*		*Evadne nordmanii*	ENOR*
	Leptocylindrus convexus	LCON*		*Evadne spinifera*	ESPI*
	Melosira varians	MVAR	*Podon* spp.		PODO
	Navicula radiosa	NRAD		*Podon intermedius*	PINT
	Pennate diatoms	PENN	*Balanus balanus*		BALA
	Proboscia alata	PALA	*Peltogaster paguri*		PPAG
	Pseudo-nitzschia multistriata	PMUL		Cirripedia nauplius larvae	CNAL
	Pseudo-nitzschia spp.	PSEU		Cirripedia cypris larvae	CCYL
	Skeletonema sp.	SKEL	*Acartia clausi*	*Acartia clausi*	ACLA*
Skeletonema menzellii		SMEN*	*Acartia tonsa*	*Acartia tonsa*	ATON*
Skeletonema pseudocostatum		SPSE		*Acartia* sp. (copepodite)	ASCO
Tenuicylindrus belgicus	*Tenuicylindrus belgicus*	TBEL*	*Calanipeda aquaedulcis*	*Calanipeda aquaedulcis*	CAQU*
	Thalassiosira sp.	THAL	*Calanus helgolandicus*		CHEL
Thalassiosira allenii		TALL	*Centropages hamatus*		CENT*
Thalassiosira delicatula		TDEL		*Cyclops* sp.	CYCL*
Thalassiosira guillardii		TGUI	*Euterpina acutifrons*	*Euterpina acutifrons*	EACU*
Thalassiosira lundiana		TLUN		*Oithona davisae*	OBRE
Heterosigma akashiwo		HAKA	*Oithona nana*	*Oithona nana*	ONAN*
	Apedinella radians	ASPI		*Oithona similis*	OSIM*
	Hemiselmis sp.	HEMI*		*Oncaea media*	OMED*
	Plagioselmis sp.	PLAG	*Paracalanus parvus*	*Paracalanus parvus*	PPAR*
Teleaulax acuta	*Teleaulax acuta*	TACU*		P-calanus (copepodite)	PCAL
Teleaulax amphioxeia	*Teleaulax amphioxeia*	TAMP*	*Pseudocalanus elongatus*		PELO
	Teleaulax gracilis	TGRA	*Pseudodiaptomus marinus*	*Pseudodiaptomus marinus*	PMAR*
	Teleaulax minuta	TMIN	*Temora longicornis*	*Temora longicornis*	TLON*
	Teleaulax spp.	TELE		Copepod nauplius	CNAU
	Chrysochromulina spp.	CHRY*		Unidentified brachiura larvae	UBRL
	Prymnesiales	PRYM	*Crassostrea gigas*		CGIG
	Gymnodiniales	GYMN	*Mytilus edulis*		MEDU
Gymnodinium aureolum		GAUR	*Littorina littorea*		LITT
	Gyrodinium flagellare	GFLA		Gastropod veliger larvae	GVEL
	Gyrodinium sp.	GYRO		Bivalve veliger larvae	BVEL
	Heterocapsa rotundata	HROT	Uncultured zooplankton		UZOO
	Heterocapsa sp.	HETE		*Oikopleura* sp.	OIKO*
	Katodinium spp.	KATO	*Sabellaria alveolata*		SALV
	Pfiesteria-like	PFIE		*Sagitta* sp.	SAGI
Micromonas pusilla		MPUS	*Scyliorhinus torazame*		STOR

Only taxa with >1 % abundance in at least one of the samples are shown. The acronyms listed here are used in the multivariate analysis. An asterisk marks those taxa identified by both methodologies, although not all of them are represented in the table (abundance lower than 1 %)

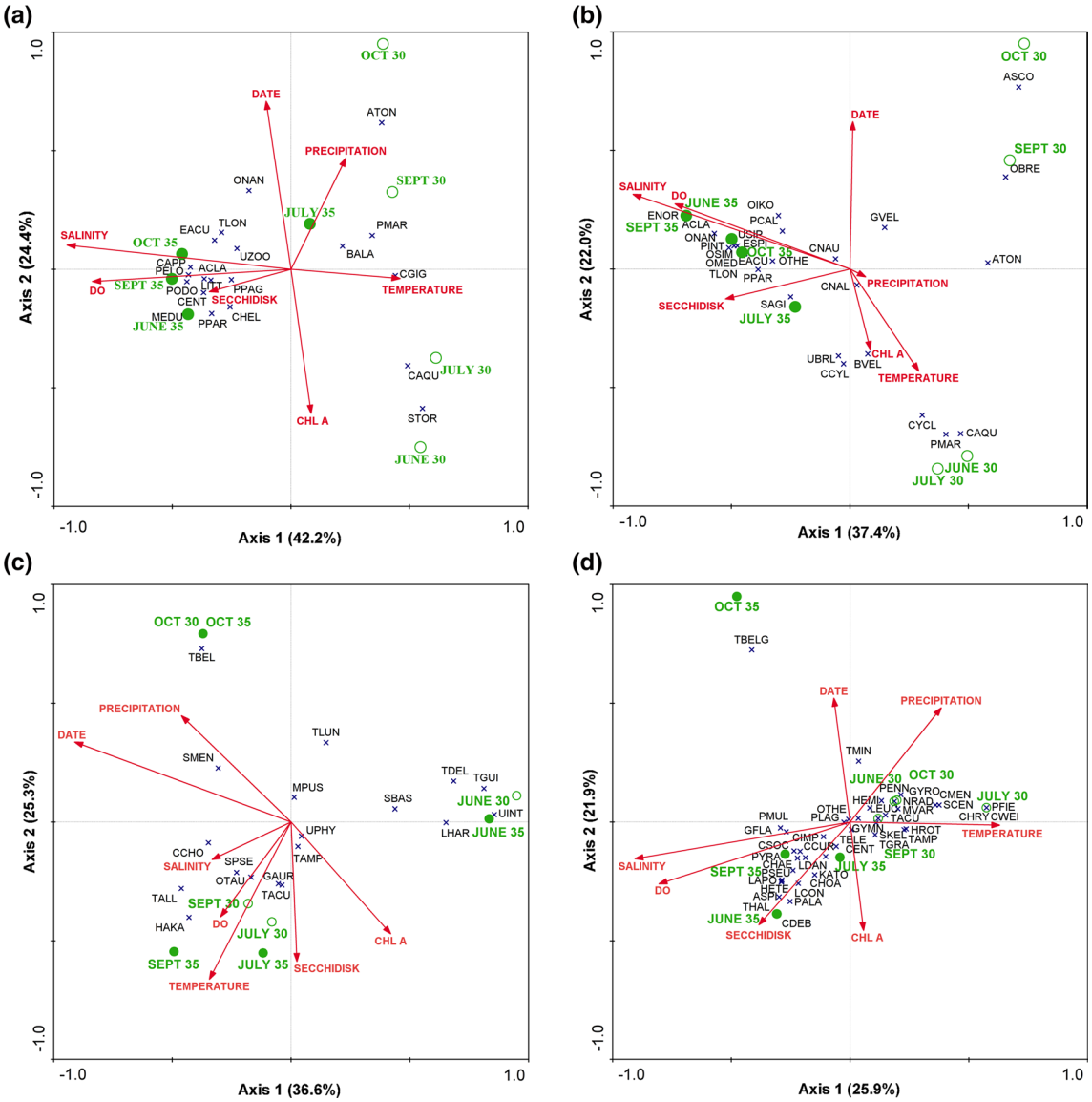

Fig. 2 Metabarcoding and microscopy CCA results. Only taxa with an abundance of 1 % or higher in at least one sample were taken into account. Cross marks identify taxa (see acronyms in Table 3). Sampling months are represented in green (salinity 35 with filled circles). Environmental variables are shown as red arrows. **a** >200 μm metabarcoding, **b** >200 μm microscopy, **c** 0.22–200 μm metabarcoding and **d** 0.22–200 μm microscopy. Date is in Julian days

Discussion

The Marine Strategy Framework Directive establishes a framework for marine environmental policy of the European exclusive economic zone (Ferreira et al. 2011), including a series of indicators that should be monitored to achieve the descriptors. Metabarcoding is especially valuable for some of these indicators, such as NIS management or biodiversity assessment (Bourlat et al. 2013). The performance of metabarcoding in monitoring plankton species, including two NIS, was critically assessed in the present study by comparing results with those of classical taxonomic analysis (microscopy). Our results indicate that metabarcoding is a promising alternative to traditional methods for early detection of NIS and plankton biodiversity shifts. Metabarcoding can be a useful tool for implementation in environmental policies, including the timely design of appropriate adaptation/mitigation measures. However, until more complete reference DNA sequence databases are publicly available, microscopic analysis should be performed in parallel at least for representative samples.

Table 4 Correlations between metabarcoding and microscopy-based analysis of community compositions

Fraction	Salinity (*n*)	Month	ρ (counts)	ρ (biomass)
>200	30 (6)	June	0.77*	0.89**
	30 (6)	July	0.95***	0.88*
	30 (6)	Sept	0.65	0.65
	30 (6)	Oct	0.51	0.51
	35 (12)	June	0.63**	0.63**
	35 (12)	July	−0.27	−0.08
	35 (12)	Sept	0.51*	0.58**
	35 (12)	Oct	0.52*	0.49*
0.22–200	30 (18)	June	0.48**	0.45*
	30 (18)	July	0.44*	0.48**
	30 (18)	Sept	0.67***	0.69***
	30 (18)	Oct	0.75***	0.77***
	35 (25)	June	0.72***	0.73***
	35 (25)	July	0.55***	0.59***
	35 (25)	Sept	0.58***	0.74***
	35 (25)	Oct	0.40**	0.44**

Spearman's rank correlation coefficient (ρ) and p values are shown; $p < 0.01$ (***), $p < 0.05$ (**) and $p < 0.1$ (*). Relative abundances from metabarcoding were compared against both microscopy-based relative abundances and biomass. Each sample was defined by size fraction (0.22–200 and >200 µm), salinity (30 and 35) and sampling month. The number of taxa (*n*) included in the correlations is specified after each salinity

NIS detection

Although the high sensitivity of metabarcoding has been described elsewhere (e.g., Zhan et al. 2013; Pochon et al. 2013), its application for monitoring biological invasions has only recently been demonstrated (Zaiko et al. 2015c). In the present work, we confirm the suitability of metabarcoding for early detection of NIS at extremely low relative abundances (Fig. 3). The reasons behind this are: (1) the ability to analyze bigger sample volumes compared to microscopy-based methods, for which screening the whole sample requires great time and effort, and (2) the capacity to detect the presence of individuals at early life stages, such as eggs or nauplius larvae, whose identification is complicated with traditional methods (Comtet et al. 2015).

In this sense, plankton monitoring programs are not usually designed to provide an early warning alert of NIS. However, the sensitivity of metabarcoding, combined with the relatively low time and cost associated with this technique (Kelly et al. 2014), results in a promising alternative approach for the rapid detection of plankton biodiversity shifts, opening the possibility for its implementation in environmental policies. As an example, Zaiko et al. (2015a, b) recently suggested the value of combining metabarcoding with current taxonomic analysis for the surveillance and management of ballast water, the main vector of most marine NIS introductions.

Quantitative nature of metabarcoding

Discrepancies between metabarcoding and microscopy-based relative abundances or biomass have been reported for plankton assemblages (Hirai et al. 2015; Massana et al. 2015; Stoeck et al. 2014; Sun et al. 2015), but studies evaluating the quantitative nature of this technique are still scarce. When comparing the relative abundances of all taxa within each sample, we showed that metabarcoding and microscopy data were correlated in most cases (Table 4). The lack of correlation could be explained by technical biases introduced during the DNA extraction (Roh et al. 2006), for which the method's performance can vary with organism type or even development stage, or PCR amplification step (Gonzalez et al. 2012), with a differential amplification that favors abundant taxa. But the copy number variation (CNV) associated with rDNA has been suggested as one of the main factors affecting the quantitative value of metabarcoding (Kembel et al. 2012): Incorporating CNV to the analysis can help to improve abundance estimates. There are also reported correlations between CNV and genome size in eukaryotes (Prokopowich et al. 2003), and between CNV and cell length and biovolume in unicellular organisms (Zhu et al. 2005; Godhe et al. 2008), suggesting a potential way of addressing this drawback in eukaryotes. In the meantime, metabarcoding targeting multi-copy genes will remain as a semiquantitative approach (Amend et al. 2010; Albaina et al. 2016a).

Metabarcoding for community ecology

Metabarcoding analysis of the plankton community replicated the temporal and spatial trends of the Bilbao estuary observed in the morphological (microscopic) analyses better for zooplankton than for phytoplankton. As expected, the main trends driving the community in the estuary of Bilbao are date (seasonal variation) and salinity (Uriarte and Villate 2004). This somewhat reduced performance in the lowest size fractions (also shown in Fig. 2) is probably related to the deficit of representative sequences for these organisms in current databases.

Metabarcoding was able to overcome the lack of resolution of microscopy for picoplankton (0.2–2 µm): Among the most abundant OTUs, the smallest size fraction was represented by the mamiellophyceans *Micromonas pusilla* and *Ostreococcus tauri* (Table 3 and ESM_4), which are known as important components of the picoplankton in temperate waters (Romari and Vaulot 2004). Regarding the taxa that were only identified in the microscopy-based analysis (Table 3), there could be two possible explanations

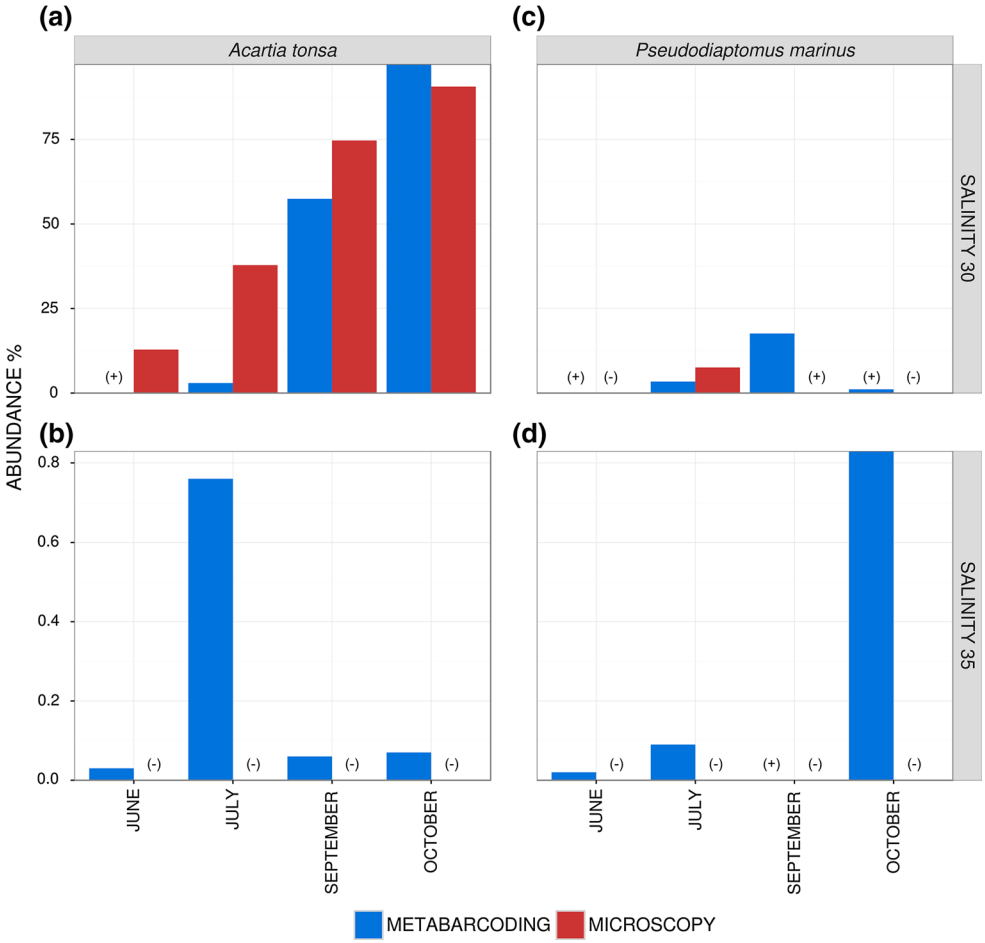

Fig. 3 Comparison of metabarcoding and microscopy when assessing two NIS. *Acartia tonsa* (**a**, **b**) and *Pseudodiaptomus marinus* (**c**, **d**) relative abundances in the >200 µm size fraction are divided by salinity (30 and 35). "+" stands for low detection percentages. "−" is shows when the species was not detected. Note that the *y*-axis scale is different for each salinity

for their absence in the metabarcoding analysis: The taxa had no representative sequence in the database (e.g., *Teleaulax gracilis*, *Oithona davisae*) or the taxa were present, but not assigned (e.g., *Apedinella radians*, *Teleaulax minuta*). In relation to the latter, this could happen if the V9 region is missing or incomplete in the available representative sequence (*Teleaulax minuta*) or due to a possible local variability or misidentification (*Apedinella radians*). In this sense, the addition of local species with no previous representation in the database (as demonstrated in this study; Table 2) significantly increases the assignment success for locally collected field samples and is recommended when designing metabarcoding studies (e.g., Cowart et al. 2015).

Table 3 also shows that among the most abundant taxa identified by microscopy, there were plankton developmental stages, such as copepodites or larvae. In the metabarcoding analysis, those organisms would be assigned to a certain taxonomy, regardless of the developmental stage. For example, the gastropod veliger larvae observed at the microscope could correspond to the benthic *Littorina littorea* or the Cirripedia nauplius/cypris larvae to *Balanus balanus* or *Peltogaster paguri*. This shows the taxonomic potential of metabarcoding versus microscopy and at the same time its limitation to be employed in studies where developmental stages need to be assessed.

Finally, a thorough revision of the OTUs uncovered surprising assignments such as the case of *Scyliorhinus torazame* (cloudy cat shark; Table 3 and ESM_4), relatively abundant in several samples (always higher in salinity 30 than in 35), which may be due to the incompleteness of the reference DNA sequence database and/or suboptimal taxonomic resolution of 18S V9 for this organism. In the same way, *Chelophyes appendiculata* (Table 3 and ESM_4) might most likely be *Muggiaea* spp., which are the only siphonophore species reported in the estuaries and coastal waters of the studied area (Villate et al. 2004). This calls for caution when reporting previously undetected species (such as NIS) using metabarcoding, because they may correspond to incorrectly assigned local species; further analysis should be performed to confirm the results.

Conclusion

Metabarcoding identified spatial and temporal trends similar to those resulting from morphological (microscopic)

taxonomic analysis for zooplankton, but not for phytoplankton, probably due to the lack of representative sequences for the latter group of organisms in current databases. The addition of representative sequences from local species resulted in an improvement in taxonomic assignment success, highlighting the need for completing reference sequence databases in order to overcome these limitations. There was a high correspondence between this approach and microscopy-based abundances, suggesting the capacity of metabarcoding for semiquantitative analysis of some taxonomic groups. Regarding the taxonomic resolution issue, while 18S V9 metabarcoding gives a broader range of taxa, its species-level resolution is not complete: A possible solution would include combining the results of 18S V9 with the 18S V1–V2 (avoiding therefore the introduction of a distinct copy number variation bias) or with a high-resolution marker such as COI, for better discrimination between species. Furthermore, the superior sensitivity of this approach allowed the identification of NIS at abundances barely detectable by microscopy. In conclusion, we think that metabarcoding is a rapid and cost-effective assessment tool that can be useful for the timely detection of NIS, which may allow the prevention or mitigation of their effects, and plankton biodiversity shifts.

Acknowledgments The authors thank Ann Bucklin (University of Connecticut) for her comments on the manuscript. The authors also thank the SGIker (UPV/EHU) for the technical and human support provided and the Hydrometeorology Service of the Regional Council of Bizkaia for the precipitation data.

Funding AA and DA work was supported by a contract with the Euskampus Foundation (Euskampus Fundazioa) and a Basque Government doctoral fellowship (UPV/EHU "ZabaldUz" program), respectively. This study was funded by a Basque Government Grant "Grupo de Investigación Consolidado del Sistema Universitario Vasco" to support activities of the Genomic Resources Research Group (IT-558-10), the Phytoplankton (IT-699-13) and Zooplankton (IT-778-13) Ecology Groups. Sampling was also partly financed by the Euskampus Foundation. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with animals performed by any of the authors.

Archiving of data Metabarcoding data (quality-filtered, chimera-free merged reads) are available at Qiita repository (https://qiita.ucsd.edu/; ID 10518).

References

Albaina A, Villate F, Uriarte I (2009) Zooplankton communities in two contrasting Basque estuaries (1999–2001): reporting changes associated with ecosystem health. J Plankton Res 31:739–752

Albaina A, Aguirre M, Abad D, Santos M, Estonba A (2016a) 18S rRNA V9 metabarcoding for diet characterization: a critical evaluation with two sympatric zooplanktivorous fish species. Ecol Evol 6:1809–1824. doi:10.1002/ece3.1986

Albaina A, Uriarte I, Aguirre M, Abad D, Iriarte A, Villate F, Estonba A. (2016b) Insights on the origin of invasive copepods colonizing Basque estuaries; a DNA barcoding approach. Mar Biodivers Rec (**in press**)

Aljanabi SM, Martinez I (1997) Universal and rapid salt extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Res 25:4692–4693

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410

Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: does read abundance count? Mol Ecol 19:5555–5565

Amorim Visco J, Apothoz-Perret-Gentil L, Cordonier A, Esling P, Pillet L, Pawlowski J (2015) Environmental monitoring: inferring diatom index from next-generation sequencing data. Environ Sci Technol 49:7597–7605

Aravena G, Villate F, Uriarte I, Iriarte A, Ibáñez B (2009) Response of Acartia populations to environmental variability and effects of invasive congenerics in the estuary of Bilbao, Bay of Biscay. Est Coast Shelf Sci 83:621–628

Bachy C, Dolan JR, López-García P, Deschamps P, Moreira D (2013) Accuracy of protist diversity assessments: morphology compared with cloning and direct pyrosequencing of 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. ISME J 7(2):244–255

Baird DJ, Hajibabaei M (2012) Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. Mol Ecol 21:2039–2044

Båmstedt U (1986) Chemical composition and energy content. In: Corner EDS, O'Hara SCM (eds) The biological chemistry of marine copepods. Clarendon, Oxford, pp 1–58

Borja A, Muxika I, Franco J (2006) Long-term recovery of soft-bottom benthos following urban and industrial sewage treatment in the Nervión estuary (southern Bay of Biscay). Mar Ecol Prog Ser 313:43–55

Borja A, Elliott M, Carstensen J, Heiskanen AS, van de Bund W (2011) Marine management—towards an integrated implementation of the European Marine Strategy Framework and the Water Framework Directives. Mar Pollut Bull 60:2175–2186

Bourlat SJ, Borja A, Gilbert J, Taylor MI, Davies N, Weisberg SB, Griffith JF, Lettierih T et al (2013) Genomics in marine monitoring: new opportunities for assessing marine health status. Mar Pollut Bull 74(1):19–31

Bricker SB, Ferreira JG, Simas T (2003) An integrated methodology for assessment of estuarine trophic status. Ecol Modell 169:39–60

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7(5):335–336. doi:10.1038/nmeth.f.303

Chen G, Hare MP (2008) Cryptic ecological diversification of a planktonic estuarine copepod, *Acartia tonsa*. Mol Ecol 17:1451–1468

Comtet T, Sandionigi A, Viard F, Casiraghi M (2015) DNA (meta) barcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. Biol Invasions 17:905–922

Cowart DA, Pinheiro M, Mouchel O, Maguer M, Grall J, Miné J, Arnaud-Haond S (2015) Metabarcoding is powerful yet still blind: a comparative analysis of morphological and molecular surveys of seagrass communities. PLoS One 10:e0117562

de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, Lara E, Berney C et al (2015) Eukaryotic plankton diversity

in the sunlit ocean. Science 348(6237):1261605. doi:10.1126/science.1261605
Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461
Edgar RC, Haas BJ, Clememte JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200
Edler L, Elbrächter M. (2010) The Utermöhl method for quantitative phytoplankton analysis. In: Microscopic and molecular methods for quantitative phytoplankton analysis. IOC Manuals and Guides
Eiler A, Drakare S, Bertilsson S, Pernthaler J, Peura S, Rofner C, Simek K, Yang Y, Znachor P, Lindström ES (2013) Unveiling distribution patterns of freshwater phytoplankton by a next generation sequencing based approach. PLoS One 8:e53516
Ferreira JG, Andersen JH, Borja A, Bricker SB, Camp J, Cardoso da Silva M, Garcés E, Heiskaneng AS et al (2011) Overview of eutrophication indicators to assess environmental status within the European Marine Strategy Framework Directive. Estuar Coast Shelf Sci 93:117–131
Gaudy R, Boucher J (1983) Relation between respiration, excretion (ammonia and inorganic phosphorus) and activity of amylase and trypsin in different species of pelagic copepods from an Indian Ocean equatorial area. Mar Biol 75:37–45
Gilbert JA, Steele JA, Caporaso JG, Steinbrük L, Reeder J, Temperton B, Huse S, McHardy AC et al (2012) Defining seasonal marine microbial community dynamics. ISME J 6:298–308
Godhe A, Asplund ME, Härnström K, Saravanan V, Tyagi A, Karunasagar I (2008) Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. Appl Environ Microbiol 74:7174–7182
Gonzalez JM, Portillo MC, Belda-Ferre P, Mira A (2012) Amplification by PCR artificially reduces the proportion of the rare biosphere in microbial communities. PLoS One 7(1):e29973. doi:10.1371/journal.pone.0029973
Herlemann DP, Labrenz M, Jurgens K, Bertilsson S, Waniek JJ, Andersson AF (2011) Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. ISME J 5:1571–1579
Hirai J, Kuriyama M, Ichikawa T, Hidaka K, Tsuda A (2015) A metagenetic approach for revealing community structure of marine planktonic copepods. Mol Ecol Res 15:68–80
Jeffrey SW, Mantoura RFC (1997) Development of pigment methods for oceanography: SCOR-supported working groups and objectives. In: Jeffrey SW et al (eds) Phytoplankton pigments in oceanography: guidelines to modern methods. Monographs on oceanographic methodology, vol 10. pp 19–36
Joshi NA, Fass JN (2011) Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. https://github.com/najoshi/sickle
Kelly RP, Port JA, Yamahara KM, Martone RG, Lowell N, Thomsen PF, Mach ME, Bennett M et al (2014) Environmental monitoring. Harnessing DNA to improve environmental management. Science 344:1455–1456
Kembel SW, Wu M, Eisen JA, Green JL (2012) Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. PLoS Comput Biol 8(10):e1002743. doi:10.1371/journal.pcbi.1002743
Laakmann S, Gerdts G, Erler R, Knebelsberger T, Martínez Arbizu P, Raupach MJ (2013) Comparison of molecular species identification for North Sea calanoid copepods (Crustacea) using proteome fingerprints and DNA sequences. Mol Ecol Res 13:862–876
Lindeque PK, Hay SJ, Heath MR, Ingvarsdottir A, Rasmussen J, Smerdon GR, Waniek JJ (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
Lindeque PK, Parry HE, Harmer RA, Somerfield PJ, Atkinson A (2013) Next generation sequencing reveals the hidden diversity of zooplankton assemblages. PLoS One 8(11):e81327
Logares R, Audic S, Bass D, Bittner L, Boutte C, Christen R, Claverie JM, Decelle J et al (2014) Patterns of rare and abundant marine microbial eukaryotes. Curr Biol 24:813–821
Mahé F, Mayor J, Bunge J, Chi J, Siemensmeyer T, Stoeck T, Wahl B, Paprotka T et al (2015) Comparing high-throughput platforms for sequencing the V4 region of SSU-rDNA in Environmental Microbial Eukaryotic Diversity surveys. J Eukaryot Microbiol 62:338–345
Massana R, Gober A, Audic S, Bass D, Bittner L, Boutte C, Chambouvet A, Christen R et al (2015) Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. Environ Microbiol 17:4035–4049. doi:10.1111/1462-2920.12955
Navas-Molina JA, Peralta-Sánchez JM, González A, McMurdie PJ, Vázquez-Baeza Y, Xu Z, Ursell LK, Lauber C, Zhou H, Song SJ, Huntley J, Ackermann GL, Berg-Lyons D, Holmes S, Caporaso JG, Knight R (2013) Advancing our understanding of the human microbiome using QIIME. Methods Enzymol 531:371–444
Olenina I, Hajdu S, Edler L, Andersson A, Wasmund N, Busch S, Göbel J, Gromisz S et al (2006) Biovolumes and size-classes of phytoplankton in the Baltic Sea. In: HELCOM Baltic sea environment proceedings no.106, Helsinki, Finland, p 144
Pochon X, Bott NJ, Smith KF, Wood SA (2013) Evaluating detection limits of next-generation sequencing for the surveillance and monitoring of international marine pests. PLoS One 8:e73935
Prokopowich CD, Gregory TR, Crease TJ (2003) The correlation between rDNA copy number and genome size in eukaryotes. Genome 46:48–50
Quail M, Smith ME, Coupland P et al (2012) A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and Illumina MiSeq sequencers. BMC Genom 13:341
Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:590–596
R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
Roh C, Villatte F, Kim B-G, Schmid RD (2006) Comparative study of methods for extraction and purification of environmental DNA from soil and sludge samples. Appl Biochem Biotechnol 134:97–112
Romari K, Vaulot D (2004) Composition and temporal variability of picoeukaryote communities at a coastal site of the English channel from 18SrDNA sequences. Limnol Oceanogr 49:784–798
Smith D (2012) fastq-barcode.pl. [Software]. https://gist.github.com/dansmith01/3920169
Stoeck T, Bass D, Nebel M, Christen R, Jones MD, Breiner HW, Richards TA (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Mol Ecol 19(Suppl 1):21–31
Stoeck T, Breiner HW, Filker S, Ostermaier V, Kammerlander B, Sonntag B (2014) A morphogenetic survey on ciliate plankton from a mountain lake pinpoints the necessity of lineage-specific barcode markers in microbial ecology. Environ Microbiol 1:430–444
Sun C, Zhao Y, Li H, Dong Y, MacIsaac HJ, Zhan A (2015) Unreliable quantitation of species abundance based on high-throughput sequencing data of zooplankton communities. Aquat Biol 24:9–15

Taylor AH, Allen JI, Clark PA (2002) Extraction of a weak climatic signal by an ecosystem. Nature 416:629–632

ter Braak CJF, Smilauer P (2002) CANOCO Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (Version 4.5). Microcomputer Power, Ithaca

Uriarte I, Villate F (2004) Differences in the abundance and distribution of copepods in two estuaries of the Basque coast (Bay of Biscay) in relation to pollution. J Plankton Res 27(9):863–974

Uriarte I, Villate F, Iriarte A, Duque J, Ameztoy I (2014) Seasonal and axial variations of net water circulation and turnover in the estuary of Bilbao. Estuar Coast Shelf Sci 150:312–324

Uriarte I, Villate F, Iriarte A (2015) Zooplankton recolonization of the inner estuary of Bilbao: influence of pollution abatement, climate and non-indigenous species. J Plankton Res. doi:10.1093/plankt/fbv060

Villate F, Uriarte I, Irigoien X, Beaugrand G, Cotano U (2004) Zooplankton communities. In: Borja A, Collins M (eds) Oceanography and marine environment of the Basque country. Elsevier Oceanography Series, vol 70. pp 395–423

Villate F, Iriarte A, Uriarte I, Intxausti L, de la Sota A (2013) Dissolved oxygen in the rehabilitation phase of an estuary: influence of sewage pollution abatement and hydro-climatic factors. Mar Pollut Bull 70(1–2):234–246

Ward BA, Dutkiewicz SA, Jahn O, Follows MJ (2012) A size-structured food-web model of the global ocean. Limnol Oceanogr 57:1877–1891

Zaiko A, Martinez JL, Ardura A, Clusa L, Borrell YJ, Samuiloviene A, Roca A, Garcia-Vazquez E (2015a) Detecting nuisance species using NGST: methodology shortcomings and possible application in ballast water monitoring. Mar Environ Res 112:64–72. doi:10.1016/j.marenvres.2015.07.002

Zaiko A, Martinez JL, Schmidt-Petersend J, Ribicic D, Samuiloviene A, Garcia-Vazquez E (2015b) Metabarcoding approach for the ballast water surveillance—an advantageous solution or an awkward challenge? Mar Pollut Bull 92(1–2):25–34

Zaiko A, Samuiloviene A, Ardura A, Garcia-Vazquez E (2015c) Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. Mar Pollut Bull 100(1):53–59. doi:10.1016/j.marpolbul.2015.09.030

Zhan A, Hulak M, Sylvester F, Huang X, Adebayo AA, Abbott CL, Adamowicz SJ, Heath DD et al (2013) High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. Methods Ecol Evol 4:558–565

Zhang J, Kobert K, Flouri T, Stamatakis A (2014) PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics 30:614–620

Zhu F, Massana R, Not F, Marie D, Vaulot D (2005) Mapping of picoeukaryotes in marine ecosystems with a quantitative PCR of the 18S rRNA gene. FEMS Microbiol Ecol 52(1):79–92