Vol. 584: 31–43, 2017
https://doi.org/10.3354/meps12373

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published December 7

18S V9 metabarcoding correctly depicts plankton estuarine community drivers

David Abad[1,*], Aitor Albaina[1,2], Mikel Aguirre[1], Andone Estonba[1]

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

[2]Environmental Studies Center (CEA), Vitoria-Gasteiz City Council, Vitoria-Gasteiz, Spain

ABSTRACT: Metabarcoding is a time-saving and cost-effective approach that promises to overcome issues associated with traditional plankton taxonomy (i.e. lack of specialized personnel, time-consuming methodologies, difficulties in assignment of larval stages, and detection of cryptic species). In this regard, we applied metabarcoding using the18S rDNA V9 region to samples collected throughout 1 yr from the Estuary of Bilbao (Basque Country, Spain) in order to characterize the annual cycle of the eukaryotic plankton community. We found clear patterns of spatial and seasonal environmental variability that drive the distribution and abundance of the plankton assemblage throughout the year, thus confirming results of previous studies using microscopic identification of the planktonic species. Our results also suggest that the low oxygen period during summer in the inner part of the estuary (salinity 30) and the thermal variation from winter to summer are among the main environmental drivers of the plankton community of the Estuary of Bilbao. Finally, we report misidentification of some species (e.g. *Cyclopina gracilis*, *Maristentor dinoferus*), which highlights the need for more comprehensive reference sequence databases in order to overcome this limitation of metabarcoding.

KEY WORDS: Estuary of Bilbao · Bay of Biscay · Spatial patterns · Seasonal variability · Low oxygen · Thermal variation · Misidentification

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Estuaries are ecosystems of high interest for ecological and conservation studies because of their transitional nature, complex dynamics, and species richness (McLusky & Elliott 2004). These environments are characterized by salinity gradients and physicochemical parameters, such as temperature or dissolved oxygen (DO) concentration, that show a clear seasonality (e.g. Morán et al. 2013). Moreover, estuaries are among the most threatened habitats worldwide; their closeness to cities and harbors contributes to habitat alteration and changes in the structure and dynamics of biotic communities (Kennish 2002), mainly due to pollution and/or introduction of non-indigenous species.

In terms of abundance and biomass, estuarine water masses are dominated by planktonic communities, which are essential for the functioning of the ecosystem, playing a crucial role in food webs and biogeochemical cycles (Ward et al. 2012). Although previous studies have provided evidence that salinity is one of the main variables driving variation of these communities (e.g. Kimmerer 2002, Muylaert et al. 2009), precipitation or temperature variation have also been related to shifts in the structure and composition of estuarine plankton assemblages (e.g. Shen et al. 2011). The ability of plankton to rapidly respond to these environmental shifts is precisely the reason why they have been used as indicators of ecosystem change for monitoring purposes (Taylor et al. 2002).

*Corresponding author: dabad87@gmail.com

© Inter-Research 2017 · www.int-res.com

Ecosystem monitoring programs rely on robust information regarding species composition. Until recently, the identification of planktonic organisms has relied on the observation of morphological characteristics by means of microscopy. Apart from the complexity and expertise required for this task, many plankton communities are often comprised of a few dominant species and numerous very rare species, which increases the difficulty of detecting and identifying all taxa (e.g. Cheung et al. 2010). In this context, metabarcoding has emerged during the last few years as a promising approach for the characterization of species composition in a diverse range of aquatic community samples (e.g. Lindeque et al. 2013, Logares et al. 2014, Hirai et al. 2015, Abad et al. 2016, Aguirre et al. 2017). The capability of metabarcoding to generate millions of sequences from a single sample at affordable costs, along with its high sensitivity (capable of detecting DNA traces) and at least comparable taxonomic resolution, provides an alternative to surmount the issues associated with traditional monitoring (Baird & Hajibabaei 2012, Zhan et al. 2013).

The Estuary of Bilbao, situated in Basque Country (south Bay of Biscay), is a ~20 km long channel that crosses a metropolitan area of about 1 million inhabitants and several industrial zones before flowing into the Cantabrian Sea (Uriarte et al. 2014). Land reclamation (especially since the mid-19th century) together with pollution coming from the city of Bilbao and factories in the vicinity reduced the original estuary and modified the ecosystem (Cearreta et al. 2000), altering abiotic processes and seasonal patterns in the planktonic community (Uriarte et al. 2014). All of these factors transformed the Estuary of Bilbao into one of the most polluted estuaries in Europe. However, since 1979, the estuary has been subjected to the Comprehensive Plan for the Sanitation of the Metropolitan Area of Bilbao. Although the pollutant concentrations are still significant, the plan has resulted in a notable improvement of water and sediment quality, and recovery of biodiversity (for a review, see Cajaraville et al. 2016).

Except for short periods of high river discharge, euhaline waters (salinity >30) dominate within the estuary (Villate et al. 2013). The seasonal patterns of this estuary are determined mostly by temperature and precipitation; between November and May, the temperatures are lower and precipitation is higher than during the rest of the year.

In the present study, we used the hyper-variable V9 region of the nuclear 18S rDNA gene (hereafter 18S V9) to characterize the planktonic eukaryotic community associated with the inner (salinity 30) and outer (salinity 35) areas of the Estuary of Bilbao. We conducted temporal monitoring by collecting samples throughout an annual cycle in order to define the key determinants that drive seasonal changes in plankton community structure. Finally, we describe the effect of the different seasonal periods on these communities and compare our findings with those previously reported in microscopy-based surveys (e.g. Villate 1994, Albaina et al. 2009).

MATERIALS AND METHODS

Sampling

Sampling was carried out from September 2013 to September 2014 in areas with salinities of 30 and 35 during neap tides; January and February collections at 35 salinity were not possible due to bad weather conditions. As the distribution and depth of each salinity mass varied from season to season (see Fig. 2 in Intxausti et al. 2012 for further information), measurements were made every 0.5 m depth in order to define the water column profile. Samples were then collected with Niskin bottles and a 200 µm mesh net when the desired salinity mass was reached (sampling depths ranging from 2 to 10 m). Once in the laboratory, 3 plankton size fractions were obtained (0.22–20, 20–200, and >200 µm). While the latter came directly from the plankton net, the Niskin bottle samples were pre-screened with a 200 µm mesh prior to the processing of the 2 lower size fractions (see Abad et al. 2016 for further details). Water samples for chlorophyll *a* (chl *a*) determination (Jeffrey & Mantoura 1997) were also collected with Niskin bottles at each salinity. Furthermore, the values for the different environmental variables and physicochemical parameters (temperature, precipitation, pH, etc.) were measured.

Metabarcoding

DNA was extracted using a modified salt protocol for the 20–200 and >200 µm size fractions, and a commercial kit (MOBIO PowerSoil®) for the 0.22 µm filters. The 18S V9 region (~150 bp) was amplified using the primers 1391f and EukBr from Stoeck et al. (2010). Sequencing data of the samples corresponding to September and October 2013 have already been published (Abad et al. 2016). The rest of

the samples were sequenced at the SGIKER facilities of the University of the Basque Country (UPV/EHU) using Illumina MiSeq 2×150 bp (sequencing information is available at the Sequence Read Archive (SRA); https://www.ncbi.nlm.nih.gov/sra; PRJNA385805).

Raw reads were pre-processed (trimming, paired-end merging, and removing chimeras) with Sickle v1.33 (quality threshold = 20; Joshi & Fass 2011), Pear v0.9.5 (minimum overlap of 15 bp and a cut-off p-value of 0.01; Zhang et al. 2014) and UCHIME (using our custom database; Edgar et al. 2011), respectively. The resulting reads were clustered into operational taxonomic units (OTUs) with UCLUST (Edgar 2010) in Qiime v1.9 (Caporaso et al. 2010), using a de novo approach with a 99% identity threshold (Abad et al. 2016). OTU assignment was performed with BLAST (Altschul et al. 1990) with a minimum of 90% identity, against a Silva 119 custom database (with the addition of representative sequences from key local species from the Estuary of Bilbao; Abad et al. 2016). Finally, a core community analysis was performed to detect the OTUs present in at least 90% of the samples collected from each water mass over the annual cycle.

Statistical analysis

A supervised learning analysis (confusion matrix; Table 1) was performed using a random forest classifier (Knights et al. 2011) with OTUs as predictors and size fractions as class labels; this method uses a subset of samples to train a model that identifies unique features within communities to predict putative similarities among size fractions. In addition, to determine the community dissimilarity for each sample, a Bray-Curtis distance network was carried out using the R package Phyloseq v1.14 (McMurdie & Holmes 2013). Alpha diversities were also calculated with Phyloseq. The bar charts representing relative abundances (see Fig. 2) and alpha diversities (see Fig. 5) were created with ggplot2 (Wickham 2009) in R (R Core Team 2017).

Finally, canonical correspondence analysis (CCA) of the OTUs showing >5% relative abundance of reads in a particular sample was carried out using CANOCO v4.5 (ter Braak & Smilauer 2002); square root transformations were used to normalize data among samples.

Table 1. Confusion matrix for the size fractions of the Estuary of Bilbao based on all samples collected during the annual cycle. The classification (Class.) error from machine learning was defined by the proportion of samples that were not clustered into their own size fraction; the higher the value of the classification error, the lower the similarity of that size. Values are ±SD

Size	>200 µm	20–200 µm	0.22–20 µm	Class. error
>200 µm	0.6524 ± 0.1515	0.3348 ± 0.1423	0.0128 ± 0.0135	0.16666667
20–200 µm	0.3387 ± 0.1201	0.5828 ± 0.0964	0.0785 ± 0.0602	0.125
0.22–20 µm	0.0567 ± 0.0398	0.1487 ± 0.0772	0.7946 ± 0.1129	0

RESULTS

OTU assignment

Only 0.74% of the reads were lost after quality filtering, and 0.14% were eliminated due to their putative chimeric nature, resulting in 3 848 144 total reads (64 136 ± 20 729 reads sample^{-1}). After read clustering, 4984 OTUs with assigned taxonomy were obtained. In all, 1859 singletons were discarded, yielding a total of 3125 OTUs for further analysis.

The core community analysis revealed that only 11 and 8 OTUs were present in at least 90% of the samples throughout the annual cycle collected from salinities of 30 and 35, respectively. Six of these OTUs (*Acartia clausi*, *A. tonsa*, *Calanipeda aquaedulcis*, *Cyclopina gracilis*, *Stomatolepas praegustator*, and Appendicularia) were shared between the 2 salinities, whereas only 5 (*Hyperamoeba flagellata,* Chrysophyceae, Paraphysomonas, *Adula californiensis*, *Polydora ciliata*) and 2 OTUs (*Paracalanus parvus* and Maxillopoda) were unique for the salinity 30 and 35 samples, respectively.

Size-fraction similarity

The Bray-Curtis distance network (Fig. 1) showed that the communities from the >200 and 20–200 µm size fractions were more similar to each other than to the 0.22–20 µm fraction. To further support this result, machine learning-based classification was carried out to determine the variability of each size fraction: the model showed that the 0.22–20 µm size fraction had a high similarity and that all samples grouped together. Although the samples from the other 2 size fractions were usually classified together (as shown in Table 1), there were some errors in classification (4 for the >200 µm and 3 for the 20–200 µm fraction, with 0.167 and 0.125 class error respectively, Table 1).

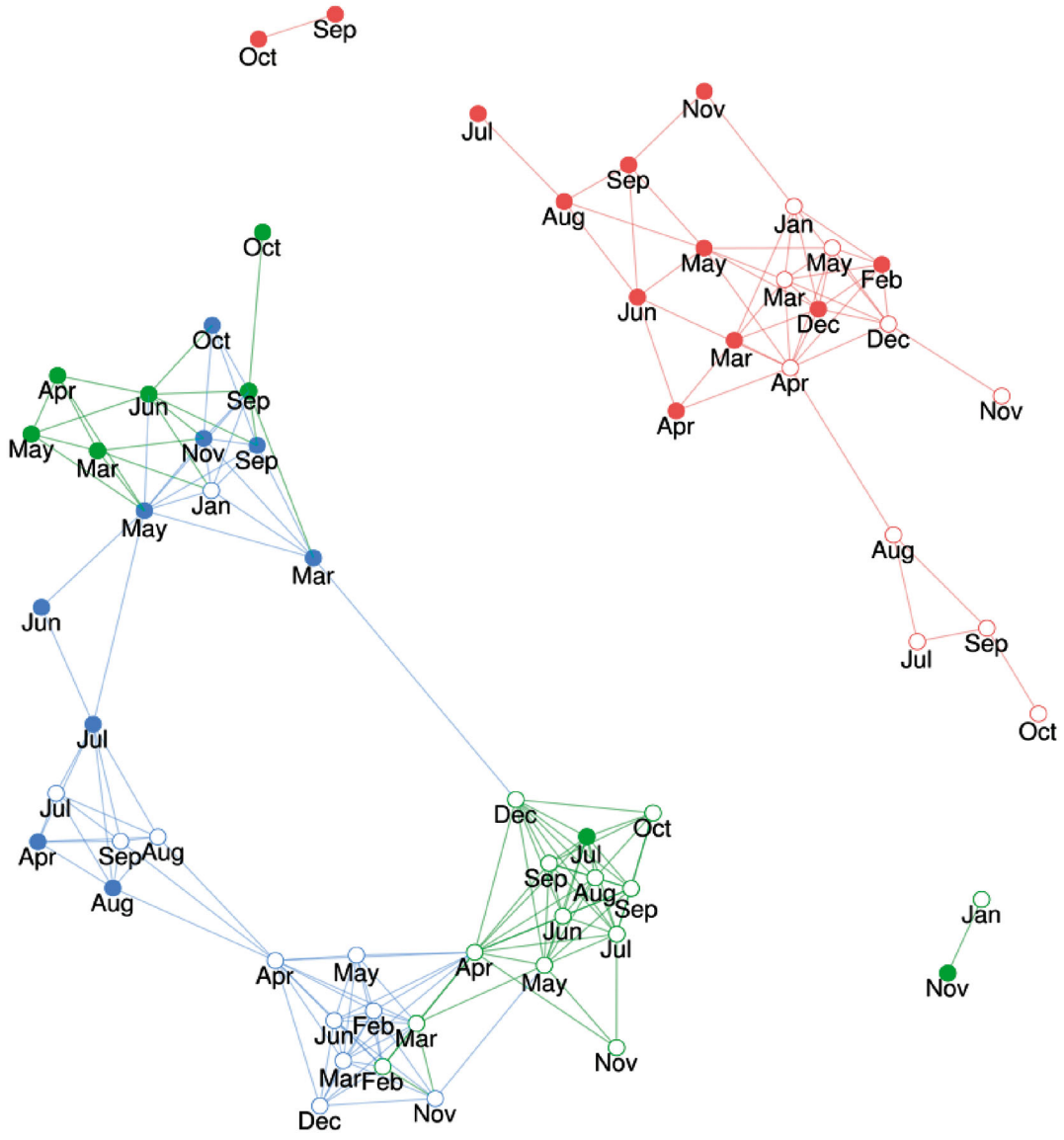

Fig. 1. Bray-Curtis distance network. Each node represents a specific sample (open and solid symbols for salinities 30 and 35, respectively). Different colors are used for the >200 (blue), 20–200 (green), and 0.22–20 µm (red) size fractions

Taxonomic composition

The OTUs identified by metabarcoding were assigned to 29 categories (Fig. 2; see Table S1 in the Supplement at www.int-res.com/articles/suppl/m584p031_supp.xls for detailed relative abundances). Maxillopoda was the most frequently observed group in the 20–200 and >200 µm size fractions for samples from both salinities: more concretely, copepods represented 51.7 and 57.1% of the OTUs, respectively. Chrysophyceae (11.7%) and Cirripedia (15.6%) were the second most abundant groups in the 20–200 and >200 µm size fractions, respectively.

The >200 and 20–200 µm communities

C. gracilis was the most abundant copepod species in salinity 30 (16.1% of the total relative abundance combined for both size fractions), followed closely by *C. aquaedulcis* (15.4%); the third most abundant species was *A. tonsa* (11.6%). For this salinity, there was a clear succession between *A. tonsa* and *C. gracilis* from the end of summer to the beginning of autumn, followed by the dominance of *C. aquaedulcis* during late winter and spring (Fig. 3); there was also a peak of *A. clausi* during winter (3.5%).

A. clausi dominated the samples from salinity 35, accounting for 18.3% of the total combined abun-

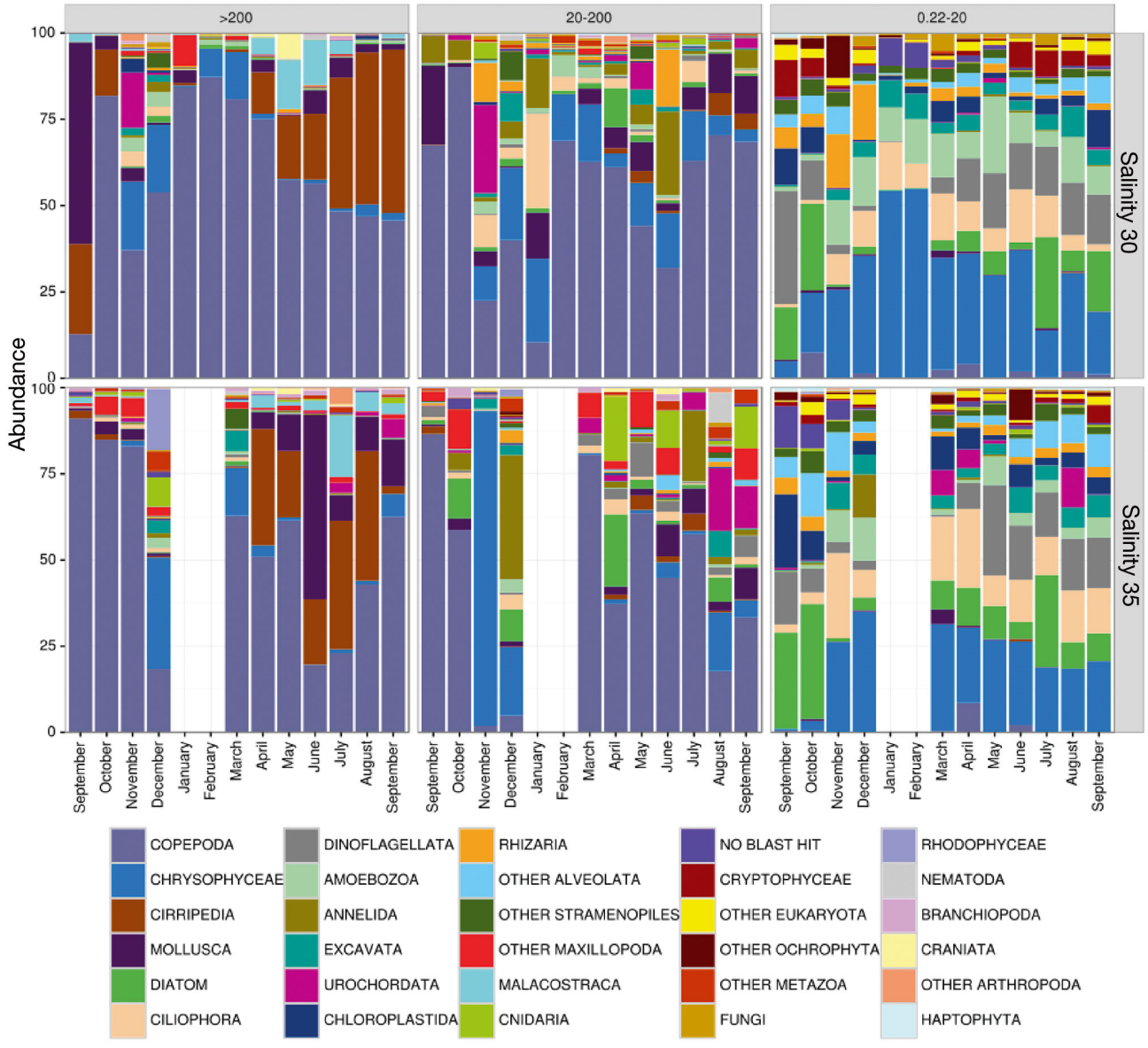

Fig. 2. Relative abundances (percentage of reads) of taxonomic groups by sample from September 2013 to September 2014. A total of 29 taxonomic groups are shown. Samples are arranged by size fraction (>200, 20–200, and 0.22–20 µm) and salinity (30 and 35). A category with sequences that had no database match is labeled 'no blast hit.' Samples from January and February in salinity 35 are missing due to bad weather conditions

dance, while *Paracalanus parvus* was the secondary dominant species at 8.8%. There was also replacement of both species by *A. tonsa* (3.6%) and *C. gracilis* (1.7%) during spring and summer–autumn, respectively, but the pattern was not as clear as in the sample from salinity 30 (Fig. 3).

Although copepods dominated these assemblages, a barnacle bloom was observed in the >200 µm size fraction in both salinities during April/May (Fig. 2) and was dominated mainly by the species *S. praegustator* (13.1%).

The 0.22–20 µm community

A more diverse assemblage characterized the 0.22–20 µm size fraction, as shown in the taxonomic composition (Fig. 2); furthermore, the network analysis (Fig. 1) also showed that the communities from both salinities were quite similar. Chrysophytes (23.3% of the total relative abundance for the combined salinities) were the most abundant group of the whole community. Phytoplankton components such as diatoms and cryptophytes accounted for 9.7 and 2.3% of the

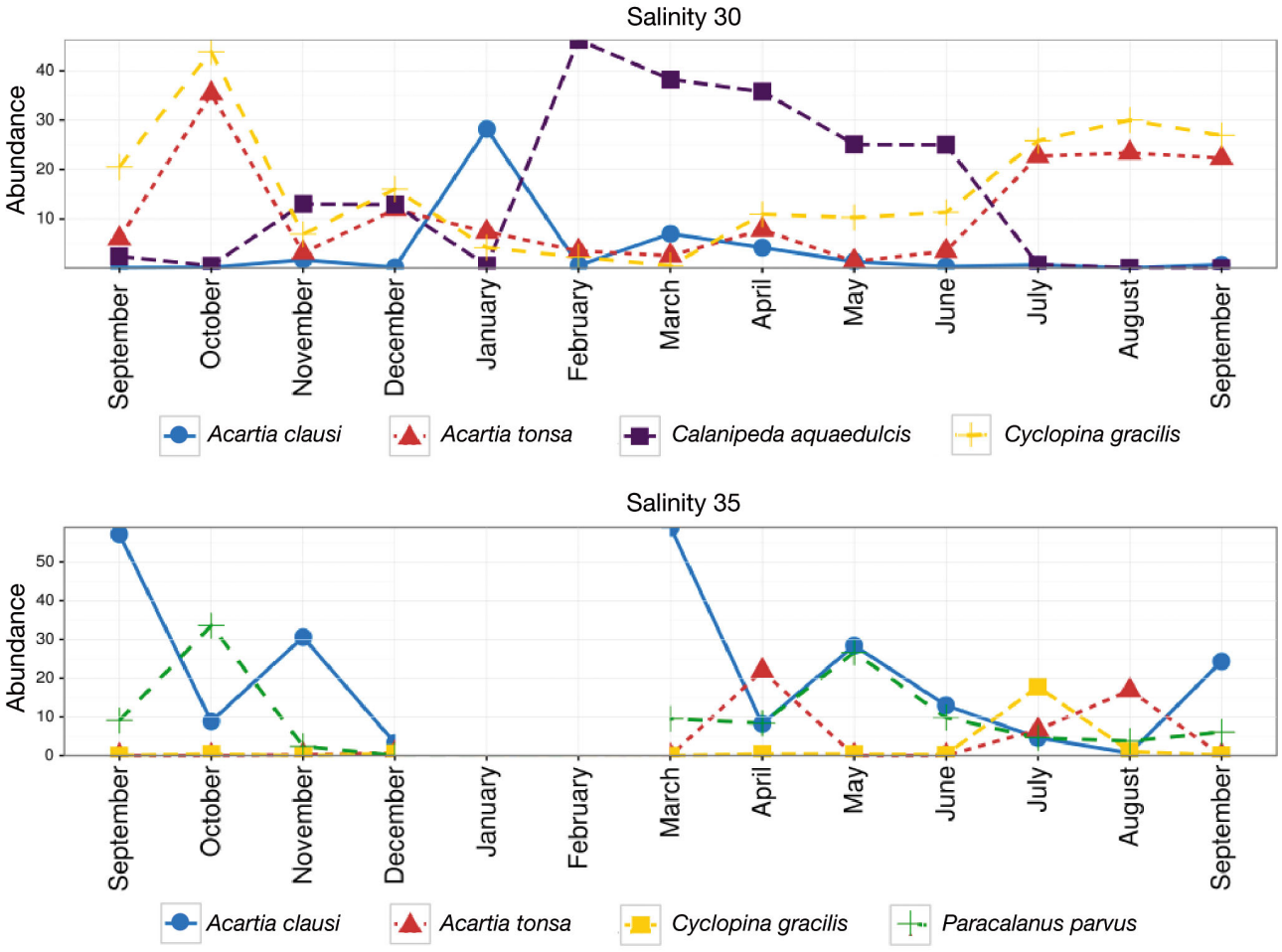

Fig. 3. Relative abundances (percentage of reads) of copepods during the annual cycle from September 2013 to September 2014. The 6 most abundant copepods of the community are shown by sample, arranged by salinity. Each species has a distinct color and symbol shape. The data included here correspond to the 20–200 and >200 µm size fractions

total community, respectively. Dinoflagellates (10%) and ciliates (9.4%) were also abundant.

The dominant diatoms were the species *Papiliocellulus elegans* (2%) and the genus *Skeletonema* (1.8%). Among the dinoflagellates, the genera *Gyrodinium* (5.4%) and *Protoperidinium* (1.7%) were the most abundant. The species *Maristentor dinoferus* was the main ciliate, accounting for 6.4% of the community. Finally, the heterotrophic genus *Paraphysomonas* (23.4%) was the most abundant in these samples and the dominant among the chrysophyceans. In this size fraction, the chrysophytes became the dominant group in both salinities during the whole year, with occasional exceptions of diatoms (Fig. 4). The similarity between communities was also reflected in the annual cycle (Fig. 4); the only exception was the ciliate group in salinity 35 during winter and some occasional peaks from other groups (e.g. dinoflagellates or diatoms).

The Shannon index for each sample is represented in Fig. 5. Overall, there are higher values and more homogeneous diversities throughout the year in the 0.22–20 µm size fraction. In contrast, the 20–200 and the >200 µm samples showed greater fluctuation in the samples from salinity 35 than in those from salinity 30.

Environmental drivers

A total of 64 taxa, which contributed to a minimum of 5% relative abundance in at least 1 sample, made up 80.3% of the total community throughout the annual cycle. The selected taxa for the CCA (Fig. 6) consisted primarily of copepods, diatoms, and mollusks (for a listing of groups, see Table S2 in the Supplement).

According to the results of the forward selection procedure in CCA, all selected environmental

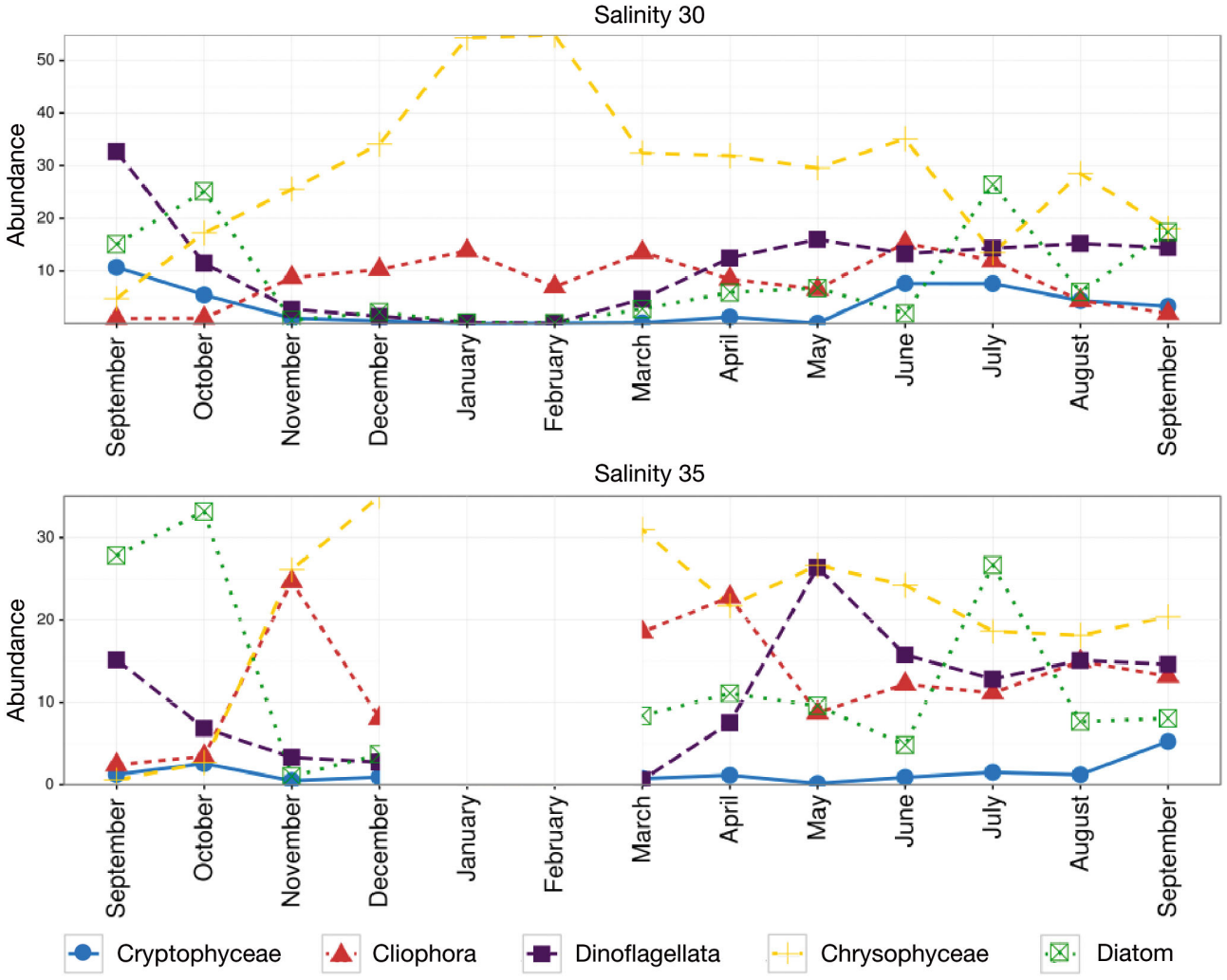

Fig. 4. Annual cycle of the 0.22–20 µm size fraction community from September 2013 to September 2014. Relative abundances (percentage of reads) of the 5 most abundant groups from this size fraction are shown. Samples are arranged by salinity. Groups are indicated by symbols of different colors and shapes

variables (pH, DO, temperature, salinity, chl *a*, and precipitation) were significantly correlated with the most abundant OTUs of the plankton community (Fig. 6). Specifically, axis 1 explained 33.9% of the species–environment relation; this axis was strongly determined by DO, pH and salinity. Axis 2 explained 20.3% of the variation in the species data, which was determined by temperature, precipitation, and chl *a*. The CCA analysis also showed a clear spatial (salinity) and temporal (seasonal) separation for the most abundant taxa throughout the year: the samples from summer and autumn were grouped together for each salinity, as were those from winter and spring. Higher DO and pH values were associated with salinity; the highest chl *a* peak was linked to temperature. As expected, precipitation varies in opposition with temperature.

According to the CCA of community composition and environmental variables (Fig. 6), the copepods *A. tonsa* and *C. gracilis*, barnacle *S. praegustator,* dinoflagellate genus *Gyrodinium* and diatom genus *Skeletonema* were inversely correlated with salinity and precipitation, but positively correlated with temperature; the copepod *C. aquaedulcis* and the chrysophyte genus *Paraphysomonas* had negative associations with temperature and salinity. The ciliate *M. dinoferus* and copepod *A. clausi* were positively correlated with salinity and precipitation, but negatively correlated with temperature. Finally, the copepod *P. parvus* was related positively with salinity.

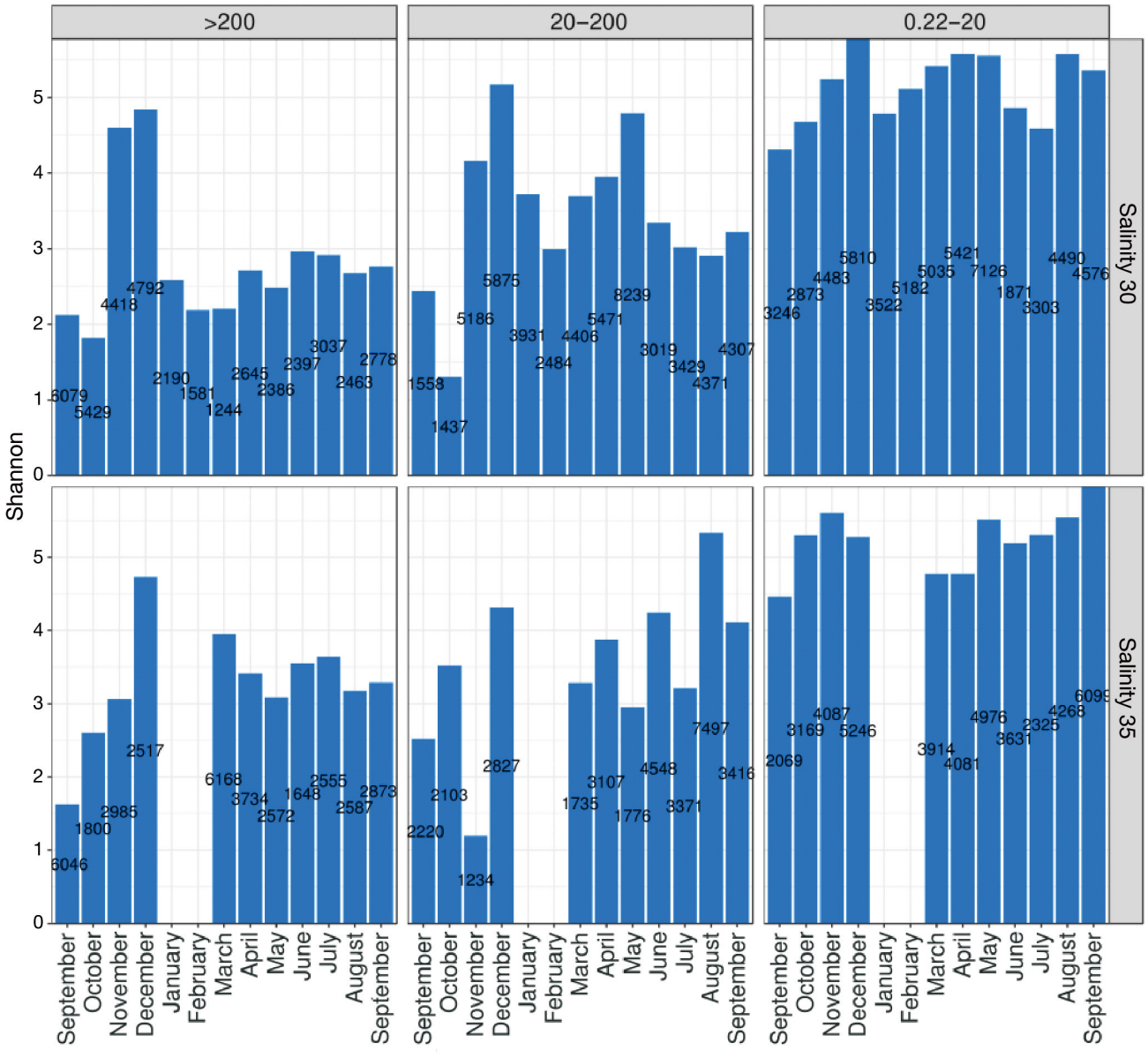

Fig. 5. Alpha diversities over the annual cycle from September 2013 to September 2014. Values shown are the Shannon diversity index for each sample, based on relative abundances of operational taxonomic units (OTUs). Samples are organized by size fraction (0.22–20, 20–200, and >200 µm), salinity (30, 35), and month sampled. Number of OTUs per sample is included within each column

DISCUSSION

The >200 and 20–200 µm communities

Although the Estuary of Bilbao used to be one of the most polluted in Europe, its water/sediment quality has improved significantly and biodiversity has recovered well since 1979 (e.g. Villate et al. 2013). This transition from a polluted to a rehabilitated area has allowed the recolonization of the water column by a mixture of neritic and brackish-water species, including non-indigenous species (Albaina et al. 2009, 2016a, Aravena et al. 2009, Uriarte et al. 2016).

Our study shows that in this estuary there is clear dominance of the *Acartia* complex copepod species among the mesozooplankton, as demonstrated in previous morphological studies (e.g. Villate 1994, Uriarte & Villate 2004, Albaina et al. 2009, Aravena et al. 2009, Uriarte et al. 2016). In our case, the 18S V9 region allowed us to decipher the current status of this complex, in which *A. tonsa* dominates during most of the year in samples collected from salinity 30,

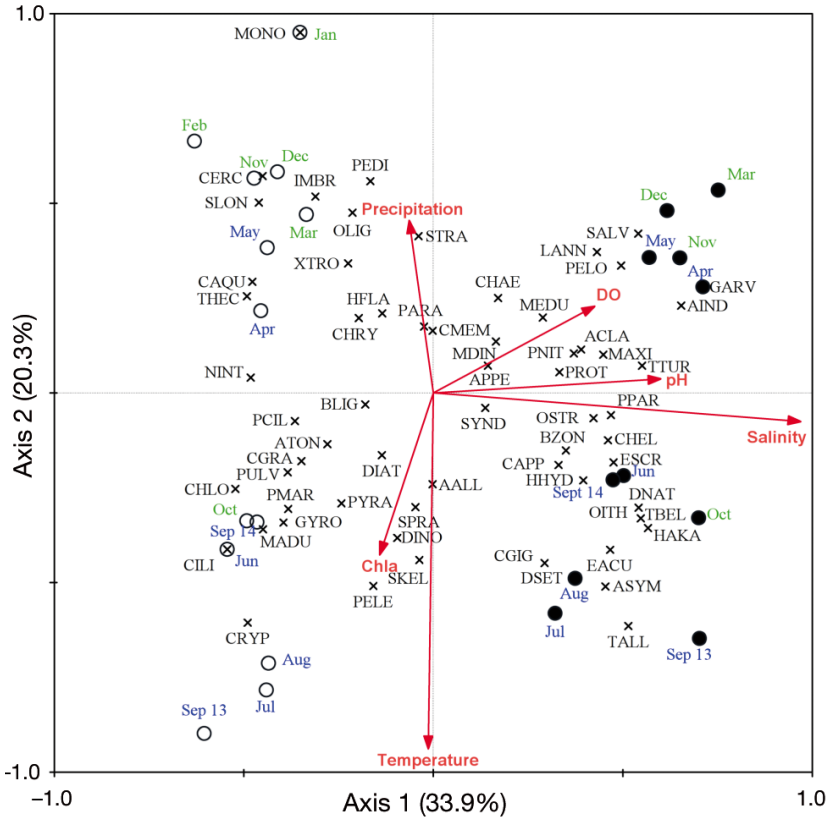

Fig. 6. Multivariate analysis of the most abundant operational taxonomic units (OTUs) and environmental variables. Only OTUs representing >5% of the abundance in any size fraction in at least 1 sample were included in the analysis. OTUs are identified as cross-marks (see acronyms in Table S2 in the Supplement at www.int-res.com/articles/suppl/m584p031_supp.xls); sampling stations are depicted by circles (open and solid for salinities 30 and 35, respectively) with the particular sampling month depicted in blue (summer and spring stations) or green (winter and autumn stations). Environmental variable gradients are represented by red arrows (DO: dissolved oxygen)

while *A. clausi* is the most abundant in salinity 35. This spatial separation of *A. clausi* and *A. tonsa* in higher and lower salinity waters, as well as the seasonal segregation, agrees with previous studies of the area (Fig. 3; Aravena et al. 2009) and has also been observed in other estuaries (e.g. Gaudy et al. 2000, Azeiteiro et al. 2005). The 18S V9 genetic marker has been shown to provide sufficient taxonomic resolution for *Acartia* species (Abad et al. 2016); the observed discriminatory power probably results from the isolation of this brackish-water genus (Chen & Hare 2008).

Furthermore, the copepod genera *Paracalanus, Clausocalanus, Pseudocalanus*, and *Ctenocalanus* usually represent a large percentage of the abundance in planktonic communities in temperate waters and are commonly grouped together due to identification difficulties (e.g. Albaina & Irigoien 2004, Gonçalves et al. 2012). Interestingly, the metabarcoding approach does not have the bias regarding the developmental stage or the cryptic species issue that affects this category, so potentially it would be able to estimate abundance more accurately. Nonetheless, there are some cases in which the reliability of the 18S V9 for copepods is compromised: a 100% identity was detected between 2 *Centropages* species, as well as among 8 copepod species corresponding to 2 sister families, Aetideidae and Euchaetidae (Albaina et al. 2016b).

Such inaccuracies of OTU assignment by metabarcoding could explain the finding of *Cyclopina gracilis*, for which there was no previous record in the area (Villate et al. 1997, Albaina et al. 2009, Uriarte et al. 2016). Given its abundance, this is most likely an error in taxonomic assignment, due to the absence of a comprehensive SILVA reference database. A search of the GenBank repository revealed that the sequence belonging to this OTU is most likely *Oithona davisae* (accession number: KJ814022) for which there is a recent citation in the Estuary of Bilbao (Uriarte et al. 2016). The genus *Oithona* is among the main constituents of the >100 µm copepod assemblage in this system (Intxausti et al. 2012) and hence mostly falls into the microzooplankton, a fraction that has been less studied to date. The difficulty of identifying early stages of this genus implies that these organisms would be classified as *Oithona* spp. rather than a particular species. As expected, this OTU is more abundant in the 20–200 µm size fraction (Table S2) and presents similar patterns in seasonal abundance, with peaks during summer/autumn, as described in previous studies of the area for *Oithona* spp. (Intxausti et al. 2012, Uriarte et al. 2016).

Difficulties in identifying developmental stages and cryptic species are more evident within the microplankton fraction (20–200 µm): for example, in a zooplankton study of the Estuary of Bilbao carried out by Intxausti et al. (2012), the identified organisms were grouped in broad taxonomic categories since some of the larval and immature forms (nauplii and copepodites) that dominated this lower size fraction

could not be assigned to species without time-consuming examination. Metabarcoding does not have this limitation, and is thus capable of assigning early stages to a certain taxonomic classification, as long as there is a reference sequence for the organism in the database. This is particularly useful for detecting non-indigenous species at very low abundances (e.g. Abad et al. 2016).

Our finding of another abundant copepod species, namely *Calanipeda aquaedulcis*, also agreed with previous studies (Albaina et al. 2009, Aravena et al. 2009, Uriarte et al. 2016). This species has contributed significantly to the increase in the total number of copepods in the Estuary of Bilbao during the last few years (Uriarte et al. 2016). *C. aquaedulcis* is known to attain peak abundances from March (Uriarte et al. 2016) and, as shown in this study, until June. Apart from this, the seasonal succession of the inner estuary zooplankton assemblage in the present study corresponded to a low oxygen period that is commonly reported during part of the summer following stratification (Intxausti et al. 2012). This condition, along with an increase in temperature may have favored the settlement of 2 species with a higher tolerance to some degree of hypoxia: *A. tonsa* and *O. davisae* (Roman et al. 1993, Itoh et al. 2011). During winter and spring, the dominance shifts to *C. aquaedulcis*, which is considered to be eurythermal but usually prefers cooler temperatures (Frisch et al. 2006).

The 0.22–20 µm community

Previous studies of the picoplankton in the Estuary of Bilbao have focused mainly on taxonomic or phylogenetic analysis of specific groups (Seoane et al. 2005, Laza-Martinez et al. 2007, Orive et al. 2010, Hevia-Orube et al. 2016), since the time and cost constraints of morphological identification have prevented studies entailing analysis of samples with sufficient volume required to detect the whole community's spatial and temporal cycles. Metabarcoding using the 18S V9 region, although subject to the aforementioned taxonomic resolution limitations (but see the genus *Acartia* case), allowed us to analyze the entire community assemblage through a year, and thereby to reveal previously unreported patterns of variation.

Our results showed that the chrysophytes are the most abundant group throughout the year: the heterotrophic *Paraphysomonas* was the dominant genus, not unusual for partially eutrophic estuaries (Bazin et al. 2014). Colorless chrysomonads, such as *Paraphysomonas*, are the major phagotrophs in freshwater and soil food webs, but they are also widespread in marine environments (Scoble & Cavalier-Smith 2014), and have been found in the Bay of Biscay (Artolozaga et al. 2000) and the Estuary of Bilbao (Cajaraville et al. 2016).

Furthermore, the naked dinoflagellate genus *Gyrodinium* is among the least-known groups of marine protists (Kubiszyn & Wiktor 2016). In contrast, *Protoperidinium* is a large and ubiquitous genus of marine heterotrophic dinoflagellates, whose species typically follow diatom blooms and generally exhibit coastal distributions (Taylor 1990). Both genera were previously described in other studies of the area, but were not followed during a complete year, as in the present study (e.g. Seoane et al. 2005).

Among the diatoms, the tiny *Papiliocellulus elegans* is a marine organism commonly found in coastal environments. Its small size requires electron microscopy for its identification, so this species has typically been overlooked and the extent of its habitat is not yet well known, although it has been regarded as possibly planktonic (Round et al. 1990). It could be present in the Estuary of Bilbao but it has not been previously reported. On the other hand, the genus *Skeletonema* occurs in coastal waters throughout the world, where it can be extremely common (Round et al. 1990) and is usually found in this estuary (Seoane et al. 2005, Laza-Martinez et al. 2007, Hevia-Orube et al. 2016).

Finally, the benthic species *Maristentor dinoferus* was the dominant ciliate, but, as in the case of *C. gracilis*, this is most likely an error of taxonomic assignment, because this organism was recently discovered on coral reefs (Lobban et al. 2002). A GenBank search resulted in matches of the sequences belonging to this OTU to uncultured phytoplankton, so it is entirely possible that this could be another species. Completion of a reference database is needed to solve problems associated with taxonomic identification by metabarcoding (e.g. Abad et al. 2016, Albaina et al. 2016b).

In this size fraction, there is little variation between the community compositions of both salinities, suggesting that the low oxygen period does not have the same influence as in the zooplankton. *Paraphysomonas*, the most abundant group throughout the year, are important feeders on bacteria (but not exclusively restricted to them) and peaked during the coldest months in the Estuary of Bilbao (November/December to March), when the lack of nutrients and sunlight prevents the proliferation of autotrophic phytoplankton and turbulence can increase grazing rates of protozoa on bacteria (Rose & Caron 2007). Diatoms, on

the other hand, showed peaks in abundance during the summer (July to October), when the temperature was higher and precipitation resulted in nutrient input from the tributaries. Non-photosynthetic species of dinoflagellates feed on diatoms or other protists (Jeong et al. 2010), which would explain why they begin to be more abundant during spring (Cajaraville et al. 2016). Finally, ciliates seem to have a peak during winter in samples collected from salinity 35 (Fig. 3), but we cannot be certain of this, due to the lack of data from January and February.

Conclusions

The metabarcoding analysis of the plankton communities present in the Estuary of Bilbao revealed that their distribution and abundance throughout the year were due to spatial and seasonal environmental variability, confirming results of previous studies using traditional techniques. The low oxygen period during summer in salinity 30 and the thermal variation from winter to summer are among the main environmental drivers of zooplankton, while temperature and precipitation are the main drivers for phytoplankton. Furthermore, we also reported misidentification of some species (e.g. *Cyclopina gracilis*, *Maristentor dinoferus*), which highlights the need for completing reference sequence databases to overcome this limitation. In light of these results, we think that metabarcoding can be useful for plankton monitoring, but that the findings obtained should be interpreted carefully until further improvement of the approach.

Acknowledgements. We thank Ann Bucklin (University of Connecticut) for her comments on the manuscript. We also thank the SGIker (UPV/EHU) for the technical and human support provided, and the Hydro-meteorology Service of the Regional Council of Bizkaia for the precipitation data. The work of A.A. and D.A. 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). 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.

LITERATURE CITED

Abad D, Albaina A, Aguirre M, Laza-Martínez A and others (2016) Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. Mar Biol 163:149, doi:10.1007/s00227-016-2920-0

Aguirre M, Abad D, Albaina A, Cralle L, Goñi-Urriza MS, Estonba A, Zarraonaindia I (2017) Unraveling the environmental and anthropogenic drivers of bacterial community changes in the Estuary of Bilbao and its tributaries. PLOS ONE 12:e0178755

Albaina A, Irigoien X (2004) Relationships between frontal structures and zooplankton communities along a cross-shelf transect in the Bay of Biscay (1995 to 2003). Mar Ecol Prog Ser 284:65–75

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, Uriarte I, Aguirre M, Abad D, Iriarte A, Villate F, Estonba A (2016a) Insights on the origin of invasive copepods colonizing Basque estuaries; a DNA barcoding approach. Mar Biodivers Rec 9:51

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

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

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. Estuar Coast Shelf Sci 83: 621–628

Artolozaga I, Ayo B, Latatu A, Azúa I, Unanue M, Iriberri J (2000) Spatial distribution of protists in the presence of macroaggregates in a marine system. FEMS Microbiol Ecol 33:191–196

Azeiteiro UM, Marques SC, Vieira LMR, Pastorinho MRD, Ré PAB, Pereira MJ, Morgado FMR (2005) Dynamics of the *Acartia* genus (Calanoida: Copepoda) in a temperate shallow estuary (the Mondego estuary) on the western coast of Portugal. Acta Adriat 46:7–20

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

Bazin P, Jouenne F, Friedl T, Deton-Cabanillas AF, Le Roy B, Véron B (2014) Phytoplankton diversity and community composition along the estuarine gradient of a temperate macrotidal ecosystem: combined morphological and molecular approaches. PLOS ONE 9:e94110

Cajaraville MP, Orive E, Villate F, Laza-Martínez A and others (2016) Health status of the Bilbao estuary: a review of data from a multidisciplinary approach. Estuar Coast Shelf Sci 179:124–134

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K and others (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336

Cearreta A, Irabien MJ, Leorri E, Yusta I, Croudace IW, Cundy AB (2000) Recent anthropogenic impacts on the Bilbao Estuary, northern Spain: geochemical and microfaunal evidence. Estuar Coast Shelf Sci 50:571–592

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

Cheung MK, Au CH, Chu KH, Kwan HS, Wong CK (2010) Composition and genetic diversity of picoeukaryotes in subtropical coastal waters as revealed by 454 pyrosequencing. ISME J 4:1053–1059

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

Frisch D, Rodríguez-Pérez H, Green AJ (2006) Invasion of artificial ponds in Doñana Natural Park, southwest Spain, by an exotic estuarine copepod. Aquat Conserv 16:483–492

Gaudy R, Cervetto G, Pagano M (2000) Comparison of the metabolism of *Acartia clausi* and *A. tonsa*: influence of temperature and salinity. J Exp Mar Biol Ecol 247:51–65

Gonçalves AM, Pardal MA, Marques SC, Mendes S, Fernández-Gómez MJ, Galindo-Villardón MP, Azeiteiro UM (2012) Responses of Copepoda life-history stages to climatic variability in a Southern-European temperate estuary. Zool Stud 51:321–335

Hevia-Orube J, Orive E, David H, Díez A, Laza-Martínez A, Miguel I, Seoane S (2016) Molecular and morphological analyses of solitary forms of brackish Thalassiosiroid diatoms (Coscinodiscophyceae), with emphasis on their phenotypic plasticity. Eur J Phycol 51:11–30

Hirai J, Kuriyama M, Ichikawa T, Hidaka K, Tsuda A (2015) A metagenetic approach for revealing community structure of marine planktonic copepods. Mol Ecol Resour 15: 68–80

Intxausti L, Villate F, Uriarte I, Iriarte A, Ameztoy I (2012) Size-related response of zooplankton to hydroclimatic variability and water-quality in an organically polluted estuary of the Basque coast (Bay of Biscay). J Mar Syst 94:87–96

Itoh H, Tachibana A, Nomura H, Tanaka Y, Furota T, Ishimaru T (2011) Vertical distribution of planktonic copepods in Tokyo Bay in summer. Plankton Benthos Res 6: 129–134

Jeffrey SW, Mantoura RFC (1997) Development of pigment methods for oceanography: SCOR-supported working groups and objectives. In: Jeffrey SW, Mantoura RFC, Wright SW (eds) Phytoplankton pigments in oceanography: guidelines to modern methods. Monogr Oceanogr Methodol. UNESCO Publishing, Paris, p 19–36

Jeong HJ, Yoo YD, Kim JS, Seong KA, Kang NS, Kim TH (2010) Growth, feeding and ecological roles of the mixotrophic and heterotrophic dinoflagellates in marine planktonic food webs. Ocean Sci J 45:65–91

Joshi NA, Fass JN (2011) Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (v1.33). https://github.com/najoshi/sickle

Kennish MJ (2002) Environmental threats and environmental future of estuaries. Environ Conserv 29:78–107

Kimmerer WJ (2002) Effects of freshwater flow on abundance of estuarine organisms: physical effects or trophic linkages? Mar Ecol Prog Ser 243:39–55

Knights D, Costello EK, Knight R (2011) Supervised classification of human microbiota. FEMS Microbiol Rev 35: 343–359

Kubiszyn AM, Wiktor JM (2015) The *Gymnodinium* and *Gyrodinium* (Dinoflagellata: Gymnodiniaceae) of the West Spitsbergen waters (1999–2010): biodiversity and morphological description of unidentified species. Polar Biol 39:1739–1747

Laza-Martinez A, Seoane S, Zapata M, Orive E (2007) Phytoplankton pigment patterns in a temperate estuary: from unialgal cultures to natural assemblages. J Plankton Res 29:913–929

Lindeque PK, Parry HE, Harmer RA, Somerfield PJ, Atkinson A (2013) Next generation sequencing reveals the hidden diversity of zooplankton assemblages. PLOS ONE 8:e81327

Lobban CM, Schefter M, Simpson A, Pochon X, Pawlowski J, Foissner W (2002) *Maristentor dinoferus* n. gen., n. sp., a giant heterotrich ciliate (Spirotrichea: Heterotrichida) with zooxanthellae, from coral reefs on Guam, Mariana Islands. Mar Biol 140:411–423

Logares R, Audic S, Bass D, Bittner L and others (2014) Patterns of rare and abundant marine microbial eukaryotes. Curr Biol 24:813–821

McLusky DS, Elliott M (2004) The estuarine ecosystem: ecology, threats and management. Oxford University Press, Oxford

McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLOS ONE 8:e61217

Morán XAG, Ducklow HW, Erickson M (2013) Carbon fluxes through estuarine bacteria reflect coupling with phytoplankton. Mar Ecol Prog Ser 489:75–85

Muylaert K, Sabbe K, Vyverman W (2009) Changes in phytoplankton diversity and community composition along the salinity gradient of the Schelde estuary (Belgium/The Netherlands). Estuar Coast Shelf Sci 82: 335–340

Orive E, Laza-Martinez A, Seoane S, Alonso A, Andrade R, Miguel I (2010) Diversity of *Pseudo-nitzschia* in the southeastern Bay of Biscay. Diatom Res 25:125–145

R Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/

Roman MR, Gauzens AL, Rhinehart WK, White JR (1993) Effects of low oxygen waters on Chesapeake Bay zooplankton. Limnol Oceanogr 38:1603–1614

Rose JM, Caron DA (2007) Does low temperature constrain the growth rates of heterotrophic protists? Evidence and implications for algal blooms in cold waters. Limnol Oceanogr 52:886–895

Round FE, Crawford RM, Mann DG (1990) The diatoms — biology & morphology of the genera. Cambridge University Press, Cambridge

Scoble JM, Cavalier-Smith T (2014) Scale evolution in Paraphysomonadida (Chrysophyceae): sequence phylogeny and revised taxonomy of *Paraphysomonas*, new genus *Clathromonas*, and 25 new species. Eur J Protistol 50: 551–592

Seoane S, Laza A, Urrutxurtu I, Orive E (2005) Phytoplankton assemblages and their dominant pigments in the Nervion River Estuary. Hydrobiologia 549:1–13

Shen PP, Li G, Huang LM, Zhang JL, Tan YH (2011) Spatiotemporal variability of phytoplankton assemblages in the Pearl River estuary, with special reference to the influence of turbidity and temperature. Cont Shelf Res 31: 1672–1681

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

Taylor FJR (1990) Phylum Dinoflagellata. In: Margulis L, Corliss JO, Melkonian M, Chapman DJ (eds) Handbook of Protoctista. Jones and Bartlett Publishers, Boston, MA

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 (v4.5). Microcomputer Power, Ithaca, NY

Uriarte I, Villate F (2004) Effects of pollution on zooplankton abundance and distribution in two estuaries of the Basque coast (Bay of Biscay). Mar Pollut Bull 49:220–228

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 (2016) Zooplankton recolonization of the inner estuary of Bilbao: influence of pollution abatement, climate and non-indigenous species. J Plankton Res 38:718–731

Villate F (1994) Temporal variability of the spatial distribution of the zooplankton community in a coastal embayment of the Basque country in relation to physical phenomena. Hydrobiologia 288:79–95

Villate F, Moral M, Valencia V (1997) Mesozooplankton community indicates climate changes in a shelf area of the inner Bay of Biscay throughout 1988 to 1990. J Plankton Res 19:1617–1636

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 hydroclimatic factors. Mar Pollut Bull 70:234–246

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

Wickham H (2009) ggplot2. Springer, New York, NY

Zhan A, Hulak M, Sylvester F, Huang X and others (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

Editorial responsibility: Antonio Bode, A Coruña, Spain

Submitted: December 27, 2016; Accepted: October 17, 2017
Proofs received from author(s): November 28, 2017