

# Considerations for metabarcoding-based port biological baseline surveys aimed at marine nonindigenous species monitoring and risk assessments

Anaïs Rey | Oihane C. Basurko  | Naiara Rodriguez-Ezpeleta 

Marine Research Division, AZTI, Sukarrieta, Spain

## Correspondence

Naiara Rodriguez-Ezpeleta, Marine Research Division, AZTI, Txatxarramendi ugartea z/g, Sukarrieta 48395, Bizkaia, Spain.  
Email: nrodriguez@azti.es

## Funding information

European Commission, Grant/Award Number: 642197

## Abstract

1. Monitoring introduction and spread of nonindigenous species via maritime transport and performing risk assessments require port biological baseline surveys. Yet, the comprehensiveness of these surveys is often compromised by the large number of habitats present in a port, the seasonal variability, and the time-consuming morphological approach used for taxonomic identification. Metabarcoding represents a promising alternative for rapid comprehensive port biological baseline surveys, but its application in this context requires further assessments.
2. We applied metabarcoding (based on barcodes of the cytochrome c oxidase subunit I and of the 18S ribosomal RNA gene) to 192 port samples collected (a) from diverse habitats (water column—including environmental DNA and zooplankton, sediment, and fouling structures), (b) at different sites (from inner to outer estuary), and (iii) during the four seasons of the year.
3. By comparing the biodiversity metrics derived from each sample group, we show that each sampling method resulted in a distinct community profile and that environmental DNA alone cannot substitute for organismal sampling, and that, although sampling at different seasons and locations resulted in higher observed biodiversity, operational results can be obtained by sampling selected locations and seasons.
4. By assessing the taxonomic composition of the samples, we show that metabarcoding data allowed the detection of previously recorded nonindigenous species as well as to reveal presence of new ones, even if in low abundance.
5. *Synthesis and application.* Our comprehensive assessment of metabarcoding for port biological baseline surveys sets the basics for cost-effective, standardized, and comprehensive monitoring of nonindigenous species and for performing risk assessments in ports. This development will contribute to the implementation of the recently entered into force International Convention for the Control and Management of Ships' Ballast Water and Sediments.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

## KEYWORDS

ballast water, environmental DNA, environmental monitoring, invasive species, metabarcoding, nonindigenous species, port biological baseline surveys

## 1 | INTRODUCTION

Globalization has led to an increased maritime transportation, with more and larger ships than ever before transferring species into ports via ballast water and hull fouling (Seebens, Gastner, & Blasius, 2013). Among the thousands species arriving daily (Carlton, 1999), some are nonindigenous species (NIS) and can become invasive, disrupting native populations (Bax, Williamson, Agüero, Gonzalez, & Geeves, 2003). Thus, being important gateways for introduction of NIS, ports need to be monitored to provide information required by legal frameworks aiming at controlling biological invasions (Lehtiniemi et al., 2015).

One of these legal frameworks is the International Convention for the Control and Management of Ships' Ballast Water and Sediments (IMO, 2004), which requires that ships treat their ballast water before its release to port, unless they show that the risk of transferring NIS between the donor and recipient ports is limited (David, Gollasch, & Pavliha, 2013). Such risk assessment requires cataloging biodiversity through port biological baseline surveys (PBBS), which requires collecting samples using various methods (Kraus et al., 2018) and at different spatial and temporal scales (Lehtiniemi et al., 2015) given the diverse range of habitats (such as soft sediment, water column, or exposed and sheltered artificial structures) allowing presence of distinct organismal groups (such as benthic macrofauna, fouling organisms, or planktonic organisms).

Global initiatives have been initiated to standardize sampling for PBBS (Awad, Haag, Anil, & Abdulla, 2014; HELCOM/OSPAR, 2013; Kraus et al., 2018) and are based on morphological taxonomic identification of the species found. Yet, this method lacks accuracy for identification of larvae and eggs, developmental stages at which many NIS are transported (Gittenberger, Rensing, Niemantsverdriet, Schrieken, & Stegenga, 2014), and relies on taxonomists who are often specialized on the local biota, but who have limited knowledge of alien taxa (Pyšek et al., 2013) and/or on specific taxonomic groups coexisting in a port (Bishop & Hutchings, 2011). Moreover, morphological identification is time-consuming (Mandelik, Roll, & Fleischer, 2010), especially when considering the several tens of samples required to characterize a port. These limitations translate into PBBS being completed several years after sample collection (Bott, 2015), reducing the effectiveness of prevention of NIS introduction and spread control strategies.

DNA metabarcoding (Taberlet, Coissac, Pompanon, Brochmann, & Willerslev, 2012), the simultaneous identification of taxa present in a complex environmental sample based on a conserved DNA fragment, is revolutionizing traditional biodiversity monitoring (Creer et al., 2016; Elbrecht, Vamos, Meissner, Aroviita, & Leese, 2017) and could represent the cost-effective and possible to standardize alternative required for PBBS (Lehtiniemi et al., 2015). Metabarcoding

can be applied to DNA extracted from bulk organismal samples (community DNA) or to environmental DNA (eDNA), which is extraorganismal DNA released in the environment in form of cells, feces, skin, saliva, mucus, etc. (Shaw, Weyrich, & Cooper, 2016). The later has received particular attention lately for its potential to cost-effectively and noninvasively survey species richness from many ecosystems (Deiner, Fronhofer, Mächler, Walser, & Altermatt, 2016), and for being able to detect spatially discrete communities (Jeunen, Knapp, Spencer, Lamare, et al., 2019).

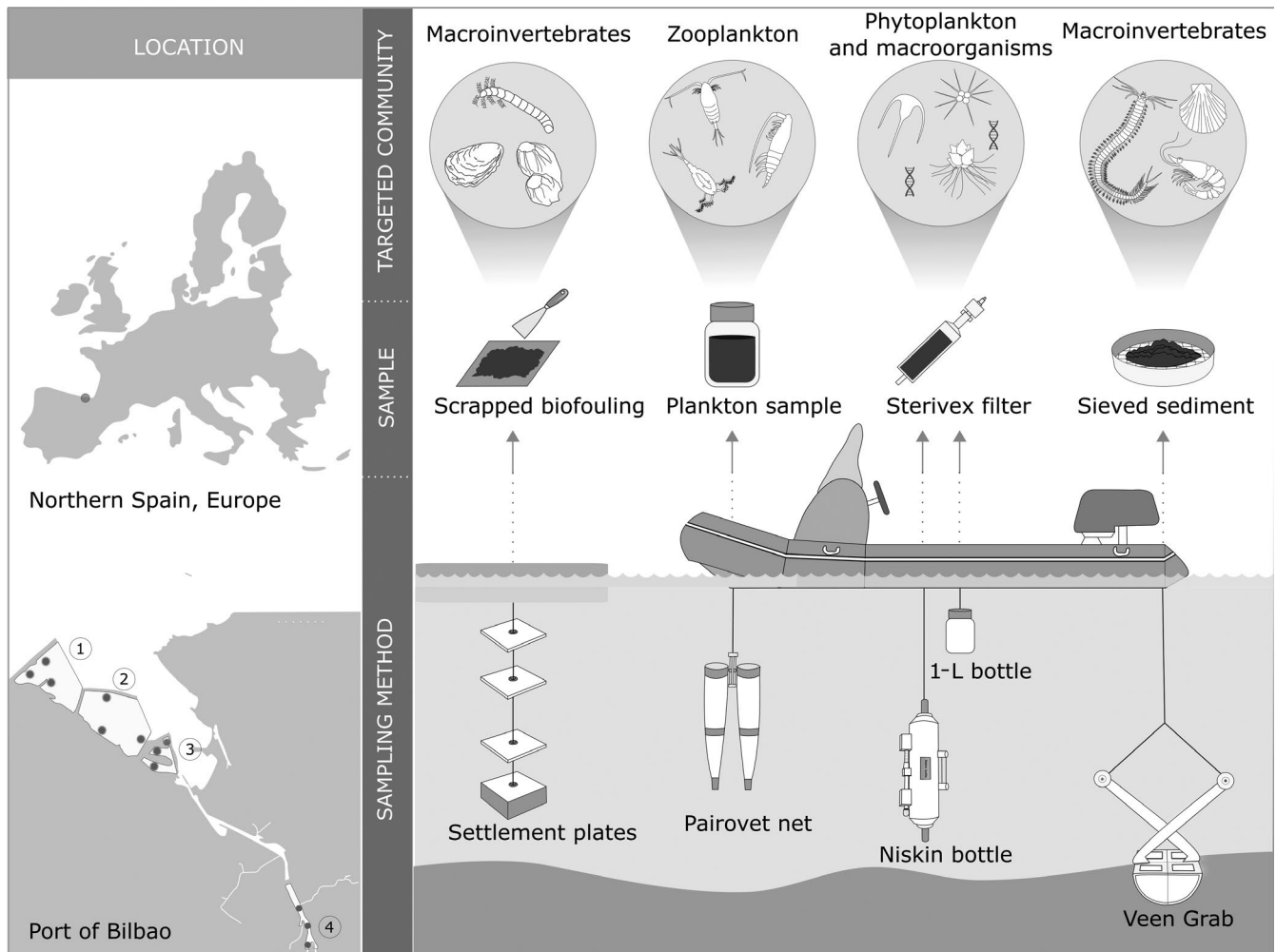
Metabarcoding has been extensively explored and validated as an alternative tool to morphological identification (Aylagas, Borja, Irigoien, & Rodríguez-Ezpeleta, 2016; Aylagas, Borja, Muxika, & Rodríguez-Ezpeleta, 2018) and described as a promising tool for NIS monitoring (Comtet, Sandionigi, Viard, & Casiraghi, 2015). Recent studies have highlighted the potential of metabarcoding applied to eDNA to perform port biodiversity surveys (Borrell, Miralles, Huu, Mohammed-Geba, & Garcia-Vazquez, 2017; Grey et al., 2018; Lacoursière-Roussel et al., 2018). Yet, very few studies have performed port surveys based on additional sampling substrates (Brown, Chain, Zhan, MacIsaac, & Cristescu, 2016; Zaiko et al., 2016) and only one has compared biodiversity assessments obtained by the different substrates (Kozioł et al., 2019). Additionally, few studies have considered the spatial and temporal variability of ports (e.g., eDNA in arctic ports Lacoursière-Roussel et al. (2018) and settlement plates in austral temperate port Zaiko et al. (2016)) but none has fully evaluated the impact of spatial and temporal variability on all the different communities of a port, which is critical to ensure maximum biodiversity recovery in PBBS.

Here, we have applied metabarcoding to 192 samples collected from the port of Bilbao (Northern Spain) to evaluate the effect of (a) using various sampling methods to capture biodiversity found in different substrates, (b) sampling at different sites from outer to inner estuary, (c) sampling during different seasons of the year, and (d) using alternative genetic markers. Our results provide relevant information for performing nonindigenous species monitoring and risk assessments in ports in response to the International Convention for the Control and Management of Ships' Ballast Water and Sediments.

## 2 | MATERIALS AND METHODS

### 2.1 | Sampling

Zooplankton, sediment, fouling, and water samples were collected from the port of Bilbao in autumn 2016, winter 2017, spring 2017, and late summer 2017 at four sites (Figure 1 and Appendix S1). Sites 1, 2, and 3 represent the outer parts of the port with busy berths and are characterized by deep waters (10–30 m) and salinity (>30



**FIGURE 1** Sampling protocol representing sampling location (Port of Bilbao), sites (1, 2, 3, and 4), and points per site (dots); and illustrating the sampling methods used and the targeted biological communities

PSU); site 4 represents the inner, less busy part of the port and is characterized by low water depth (6–9 m) and lower salinity (>20 PSU). Zooplankton, sediment, and fouling sampling was based on the protocol designed by HELCOM/OSPAR (2013). Zooplankton samples were collected in vertical tows using Pairovet nets (mesh sizes of 60 and 280  $\mu\text{m}$ ) at three points per site, mixing the collected material to have one sample per mesh size per site. The collected zooplankton was ground with a mortar until no integer organism could be appreciated and was stored in 96% ethanol at  $-20^{\circ}\text{C}$ . Sediment samples (3 samples per site) were collected by sieving material collected from a Van Veen grab (0.07–0.1  $\text{m}^2$ ) with a 1-mm-mesh size sieve. Retained material was processed following Aylagas, Borja, et al. (2016) and Aylagas, Mendibil, et al. (2016) based on decantation and homogenization phases, and retrieved benthic macroinvertebrates were stored in 96% ethanol at  $-20^{\circ}\text{C}$ . Homogenization was done by blending the decanted organic material in a PHILIPS hr2095 700 W 2-L glass jar with 96% ethanol until no fragments of animals and other organic materials could be observed. During decantation and homogenization steps, all laboratory equipment used were cleaned between samples by soaking

in 10% bleach for 15 min and thoroughly rinsing with deionized water. Fouling samples were collected by placing 15 x 15 cm polyvinyl chloride plates at 1, 3, and 7 m depth in a suspended array in each site. Plates were deployed in winter and spring and recovered in spring and late summer, respectively (Table S1). At recovery, plates were placed in individual sterile plastic bags, soaked in 96% ethanol, and thoroughly scrapped both sides with a scalpel to retrieve fouling organisms attached to it. The detached organisms were then homogenized with a blender (Conair™ Waring™ Mini-sample Containers for Blender) in 96% ethanol and stored in 96% ethanol at  $-20^{\circ}\text{C}$ . Cleaning procedure was similar as the one performed for sediment samples. At each sampling site, 1 L surface water sample was collected from each sampling point with a bottle. Samples from the same site and from the same sampling method were combined into a single sample by site. In sites 1 and 3, three additional samples were taken at one meter above the bottom using a Niskin bottle and combined. Each 3 L combined water sample was filtered using a 0.45- $\mu\text{m}$  Sterivex filter unit (Merck Chemicals & Life Science), which was stored at  $-80^{\circ}\text{C}$  until further use. Additionally, 3 L surface water sample from three

sampling sites from the ports of Vigo and A Coruña, located in Galicia, Northwestern Spain, following the same sampling procedure than for Bilbao by combining 1 L water from three sampling point per sampling site, was collected in March 2017 and filtered as described above.

## 2.2 | DNA extraction, library preparation, and sequencing

Total genomic DNA was isolated from zooplankton samples with the DNeasy blood & tissue kit (QIAGEN), benthic macroinvertebrate, and fouling organism samples with, respectively, PowerMax and PowerSoil DNA Isolation Kits (MOBIO) replacing the initial bead-beating step by an overnight incubation at 56°C with proteinase K (0.4 mg/ml). Filter samples were extracted with the DNeasy blood & tissue kit (QIAGEN) following the “SX filters without preservation buffer” developed by Spens et al. (2017). Negative controls (preservation buffer) were included in each batch of DNA extraction and followed the same procedure as all other samples. DNA concentration was measured with the Quant-iT dsDNA HS assay kit (Thermo Scientific) using a Qubit 2.0 Fluorometer (Life Technologies), purity was inferred from 260/280 and 260/230 absorbance ratios with the ND-1000 NanoDrop (Thermo Scientific), and integrity was assessed by electrophoresis in 0.7% agarose. Two primer pairs widely used to assess eukaryotic biodiversity (Amaral-Zettler, McCliment, Ducklow, & Huse, 2009; Hebert, Ratnasingham, & Waard, 2003) were used: mICOLintF/dgHCO2198 (COI primers), targeting a 313-bp fragment of the cytochrome c oxidase subunit I (COI) gene (Leray et al., 2013; Meyer, 2003) and 1389F/1510R (18S primers), targeting a variable length fragment (87–186 bp) of the V9 hypervariable region of the 18S rRNA gene (Amaral-Zettler et al., 2009). DNA extracted from zooplankton, fouling organism, and filter samples were amplified with both primer pairs, whereas DNA from sediment was only amplified with the COI primers. PCR amplifications were performed in two rounds. For the first PCR, 1 µl of genomic DNA (5 ng/µl) was added to a mix consisting in 5 µl of 2X Phusion Master Mix (Thermo Scientific), 0.2 µl of each primer (0.2 µM), and 2.6 µl of Milli-Q water. For the 18S primers, PCR conditions consisted on an initial 3-min denaturation step at 98°C, followed by 25 cycles of 10 s at 98°C, 30 s at 57°C, and 30 s at 72°C and finally 10 min at 72°C. For the COI primers, PCR conditions consisted of an initial 3-min denaturation step at 98°C, followed by 35 cycles of 10 s at 98°C, 30 s at 46°C, and 45 s at 72°C and finally 5 min at 72°C. Negative controls of PCR (no template) were included within each set of PCRs. For each DNA extract, three PCR amplifications were performed and pooled. Once purified using AMPure XP beads (Beckman Coulter), the mixed PCR products were used as template for the generation of dual-indexed amplicons in a second PCR round following the “16S Metagenomic Sequence Library Preparation” protocol (Illumina) using the Nextera XT Index Kit (Illumina). Multiplexed PCR products were purified again using the AMPure XP beads, quantified using Quant-iT dsDNA HS assay kit and a Qubit 2.0 Fluorometer (Life Technologies), normalized to equal concentration, and sequenced

using the 2 x 300 paired-end MiSeq (Illumina). Reads were demultiplexed based on their barcode sequences.

## 2.3 | Raw read preprocessing, clustering, and taxonomic assignment

After quality checking of demultiplexed paired-end reads with FastQC (Andrews, 2010), forward and reverse primers were removed using Cutadapt (Martin, 2011) with the anchored 5' adapter and for “paired-end reads” options and with the linked adapter option for COI and 18S, respectively. Forward and reverse reads were merged using PEAR (Zhang, Kobert, Flouri, & Stamatakis, 2014) with a minimum sequence overlap of 217 bp and a maximum amplicon length of 313 bp and of 80 and 190 bp for COI and 18S barcodes, respectively. Merged reads with average Phred quality score lower than 20 were removed with Trimmomatic (Bolger, Lohse, & Usadel, 2014). Using Mothur (Schloss et al., 2009), sequences without ambiguous bases were aligned to BOLD (<https://www.boldsystems.org>) or SILVA (<https://www.arb-silva.de/documentation/release-132/>) for COI and 18S, respectively, and only those covering the barcode region were kept. Chimeras, detected using de novo mode of UCHIME (Edgar, Haas, Clemente, Quince, & Knight, 2011), were removed, and remaining reads were clustered into OTUs using Swarm 2.2.1 with the step-by-step aggregation clustering algorithm implemented with default settings (Mahé, Rognes, Quince, de Vargas, & Dunthorn, 2014). SWARM algorithm does not rely on a fixed threshold for delimiting OTUs, which is pertinent when performing PBBS where highly diverse biodiversity can be found. “Singleton” OTUs, composed by a single read, were removed. No rarefaction was performed to avoid decreasing sensitivity by choosing an arbitrary minimum library size (McMurdie & Holmes, 2014). The remaining OTUs were taxonomically assigned using the Naïve Bayes Classifier (Wang, Garrity, Tiedje, & Cole, 2007) using BOLD (accessed in May 2018) or PR2 (release 4.10.0) databases for COI and 18S, respectively.

## 2.4 | Community analyses

Apart from the complete dataset, we created two subsets: a taxa targeted through PBBS dataset, including only reads classified to the class level and excluding those matching to nontargeted groups for PBBS such as Mammalia, Aves, Insecta, Collembola, Arachnida, and all classes of Fungi; and a NICS dataset, including only reads matching either the 68 nonindigenous and cryptogenic species (NICS) previously detected in the port of Bilbao (Adarraga & Martínez, 2011, 2012; Butrón, Orive, & Madariaga, 2011; Martínez & Adarraga, 2006; Tajadura, Bustamante, & Salinas, 2016; Zorita et al., 2013) or the 1,083 species present in the AquaNIS database (AquaNIS. Editorial Board, 2015). Most analyses were conducted using RStudio (RStudio Team, 2015) with *vegan* (Oksanen et al., 2019), *adespatial* (Dray et al., 2018), and *indicspecies* (De Cáceres & Legendre, 2009) libraries. Indicator species analyses (Dufrene &

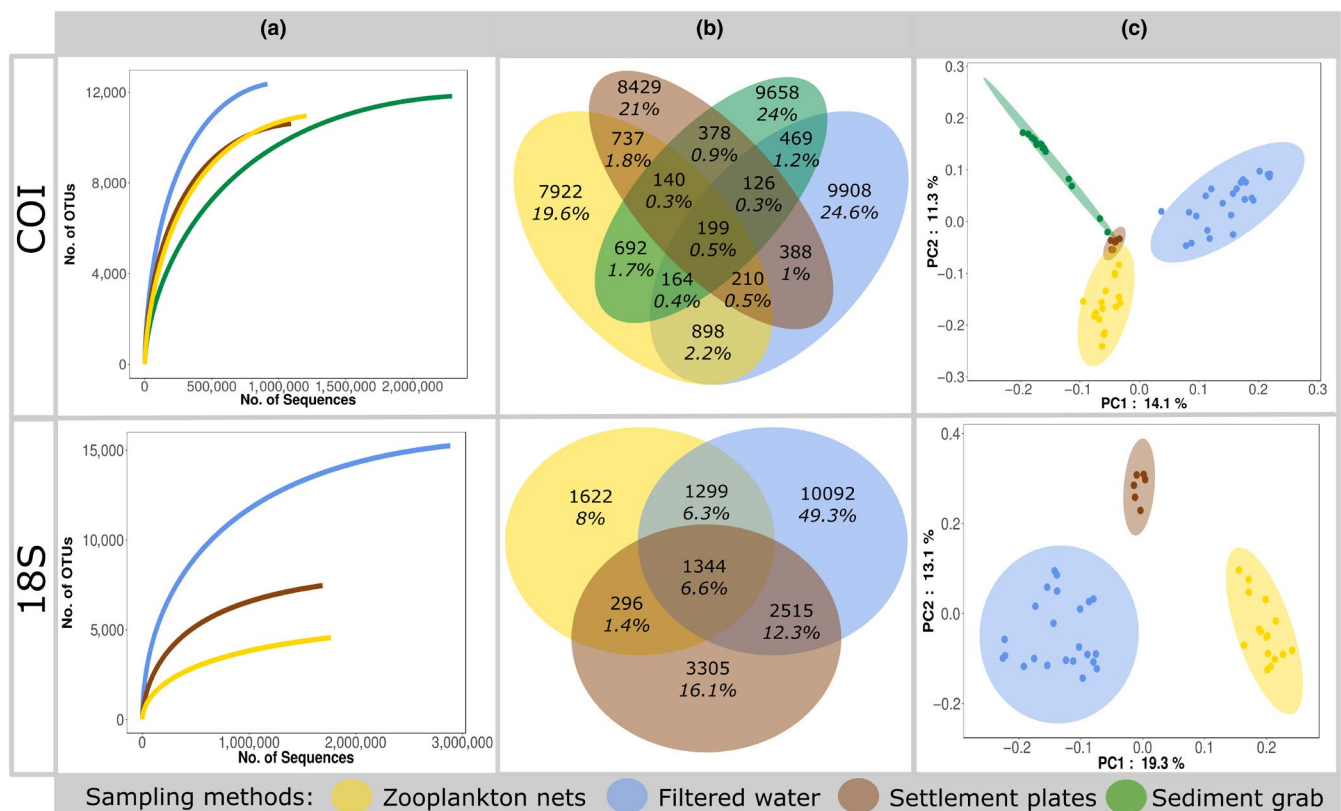
Legendre, 1997) were performed on (a) the taxa targeted through PBBS dataset to identify indicator taxa of each sampling method and (b) the NICS dataset to identify nonindigenous indicator taxa of each sampling site. These analyses were based on the IndVal index calculated as the product of the degree of specificity (measuring the uniqueness to a sampling method or site) and the degree of fidelity (measuring the frequency of occurrence within a sampling method or site) of an OTU to a given sampling condition. Statistical significance of associations was assessed by performing 10,000 permutations. The effects of season and locality on taxa targeted through PBBS communities were tested for significance using a permutational multivariate analysis of variance (PERMANOVA) after checking for multivariate homogeneity of group dispersions (betadisper). PERMANOVA and betadisper were performed on Euclidean distances on Hellinger-transformed OTU abundances (Hellinger distances), which are appropriate for community ordination and clustering (Legendre & Gallagher, 2001) and on Jaccard dissimilarities based on OTU presence/absence. The contribution of replacement (changes in OTU identity) and nestedness (richness differences where one sample is a subset of a richer sample) to beta diversity of taxa targeted through PBBS between seasons and between sites was computed using the relativized nestedness index of Podani and Schmera (2011) based on Jaccard dissimilarity matrix. For each sampling method, replacement and nestedness

were calculated between pairwise comparisons of (a) samples belonging to the same site but sampled at different seasons (season variation of beta diversity), and (b) samples belonging to the same season but sampled at different sites (spatial variation of beta diversity). The mean proportion of nestedness and replacement contribution to beta diversity between sites and between seasons was then calculated. For all OTUs assigned to a NICS, we blasted their representative sequences against BOLD and PR2 databases, respectively, for COI and 18S barcodes.

### 3 | RESULTS

#### 3.1 | Communities retrieved by each sampling method and genetic marker

A total of 5,718,639 and 7,055,675 COI and 18S barcode reads were kept for analysis (Table S1). Negative controls showed little contamination (with only 164, 66 and 213 of COI and 2031 and 163 18S quality-filtered reads). The reads corresponding to the 192 samples collected at the port of Bilbao resulted in 40,318 and 20,473 OTUs for COI and 18S, respectively. For all sampling methods, OTU accumulation curves of OTUs against reads approached saturation, suggesting that adding more sequencing effort would provide limited increase in diversity



**FIGURE 2** Overall description of community detected per barcode. (a) OTU accumulation curves per sampling method. (b) Venn diagrams of the number and percentage of OTUs shared between sampling methods. (c) Principal component analyses of the Hellinger-transformed abundances of OTUs. Sample scores are displayed in scaling 1 with ellipses representing the 95% confidence dispersion of each sampling method

(Figure 2a and Appendix S2). The majority of the OTUs, 89% for COI and 73% for 18S, were unique to one sampling method (Figure 2b), and, for both barcodes, but particularly for 18S, filtered water (aimed at retrieving eDNA from macroorganisms and microbial eukaryotes) resulted in higher OTU richness and unique OTUs than other sampling methods (Figure 2a,b and Appendix S2). Different sampling methods retrieved distinct biological communities as observed with the PCA (Figure 2c) and confirmed with PERMANOVA analyses (COI:  $R^2 = .28$ ,  $p$ -value =  $9.99e-05$ ; 18S:  $R^2 = .30$ ,  $p$ -value =  $9.99e-05$ ).

The percentage of reads assigned to species or higher taxonomic levels differed between sampling methods and barcodes. The majority of the reads were assigned to species level for zooplankton nets (COI: 84% and 18S: 75%), settlement plates (COI: 69% and 18S: 69%), and sediment grabs (COI: 64%); for filtered water, although 51% of reads were assigned to species level with 18S, only 8% could be assigned to a species with COI (Figure 3a). Filtered water analyzed with COI had also the largest proportion of reads that could not be assigned to phylum (87%). This was due to the nonspecific amplification of prokaryotic and nontarget eukaryotic DNAs (Appendix S3), which is known for COI (Collins et al., 2019).

Both barcodes detected a wide range of eukaryotic groups (Figure 3b). With COI, the greatest majority of reads belonged to metazoans, with Arthropoda dominating all sampling methods except sediment grabs, dominated by Annelida. With 18S, a more diverse spectrum of taxa was retrieved, including phytoplankton and macroalgae; as expected, the number of reads assigned to these phyla in zooplankton nets and settlement plates was very low in comparison with filtered water.

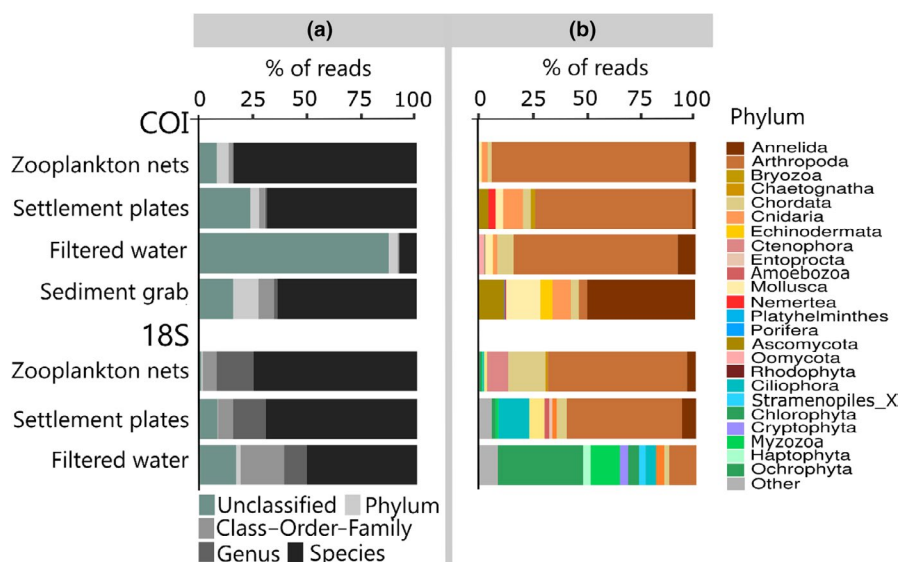
### 3.2 | Distribution of taxa targeted through port biological baseline survey

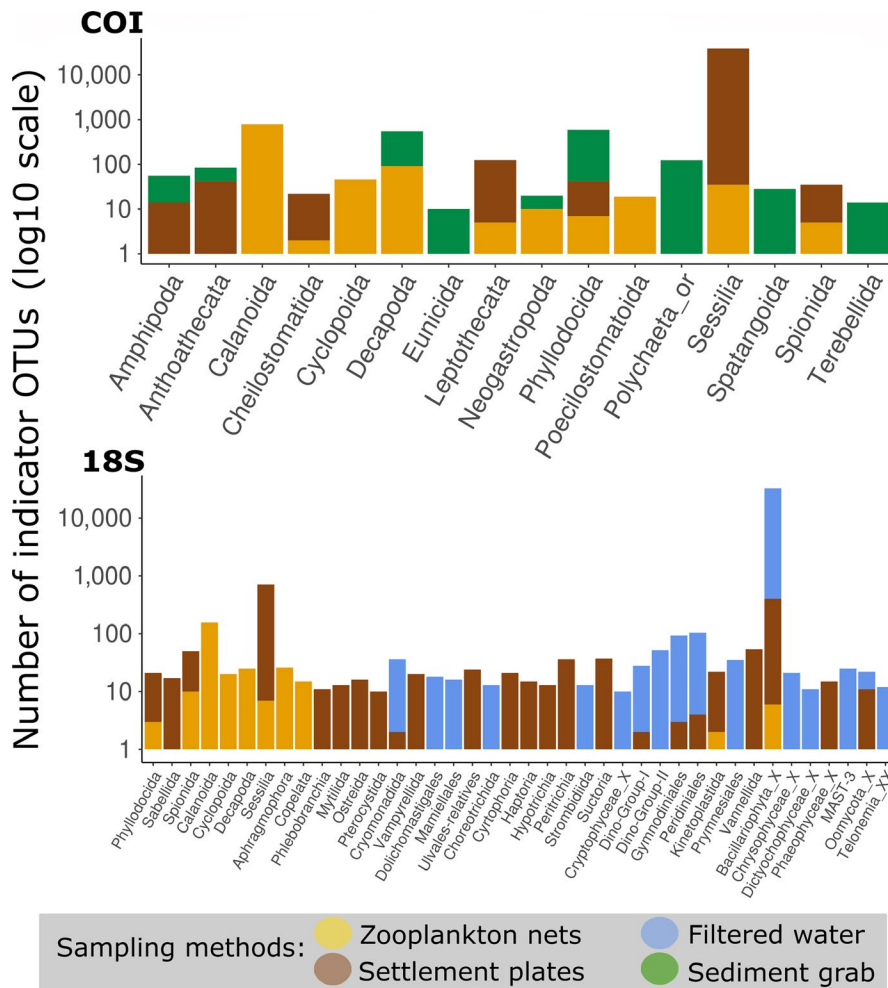
A total of 16,828 and 9,091 OTUs were kept as taxa targeted through PBBS for COI and 18S, respectively. From those,

indicator analysis identified, respectively, 2,600 (15%) and 1,700 (19%) OTUs significantly associated with one of the sampling methods. Settlement plates were associated with 1,268 (COI) and 808 (18S) OTUs, zooplankton nets with 1,001 (COI) and 315 (18S) OTUs, and sediment grabs with 238 (COI) OTUs. Only 12 OTUs (9 of which were metazoans) were indicators of filtered water with COI, while 580 (3 of which were metazoans) were associated with this sampling method with 18S. We observed expected strong associations of some taxa to one particular sampling method: fouling bivalves (Ostreoida, Mytilida) with settlement plates, copepods (Calanoida, Cyclopoida) with zooplankton nets, and sea urchins (Spatangoida) with sediment grabs (Figure 4). Yet, less obvious associations were also observed such as barnacles (Sessilia) and two polychaeta orders (Spionida, Phyllodocida) with zooplankton nets or dinoflagellates (Peridinales and Gymnodinales) with settlement plates, illustrating the advantages of a complementary sampling approach to recover the diversity of these taxonomic groups.

The need for complementarity of sampling methods was further confirmed by PCA on PBBS target taxa (Figure 5). For both barcodes, expected taxonomic differences were found between the sampling methods: Sediment grabs were characterized by the polychaeta *Maldane glebifex* (Figure 5b), zooplankton nets were distinguished by several copepods (Figure 5b,d), and settlement plates, by the encrusting *Semibalanus balanoides* (Figure 5d). Filtered water communities were characterized differently according to each barcode. For COI, filtered water samples were not differentiated as a distinct group and were in general close to those retrieved with zooplankton samples collected at the same season (Figure 5a). In contrast, for 18S, filtered water was different due to the presence of the phytoplankton species *Phaeocystis globosa* (Figure 5d); yet, when targeting only metazoan taxa, the patterns observed with COI were similarly to the ones retrieved with 18S (Appendix S4). For filtered samples, the proportion of metazoan OTUs detected was largely different compared to the other methods (Appendix S5). For instance, Decapoda orders had a low diversity with filtered water (COI: 0.7%; 18S: 7%), whereas more diversity could be

**FIGURE 3** Taxonomic assignment per barcode and per sampling method. (a) Percentage of reads assigned to each taxonomic level. (b) Relative abundance of reads classified to at least Phylum level for the 10 most abundant phyla





**FIGURE 4** Distribution of indicator OTUs associated with each sampling method at the order level (in log<sub>10</sub> scale) for COI and 18S. Only orders with at least 10 indicator OTUs are shown

detected for Calanoida (COI: 8%; 18S: 35%), Sabellida (COI: 15%; 18S: 28%), and Leptothecata (COI: 13%; 18S: 58%). In general, filtered water recovered the smallest metazoan diversity.

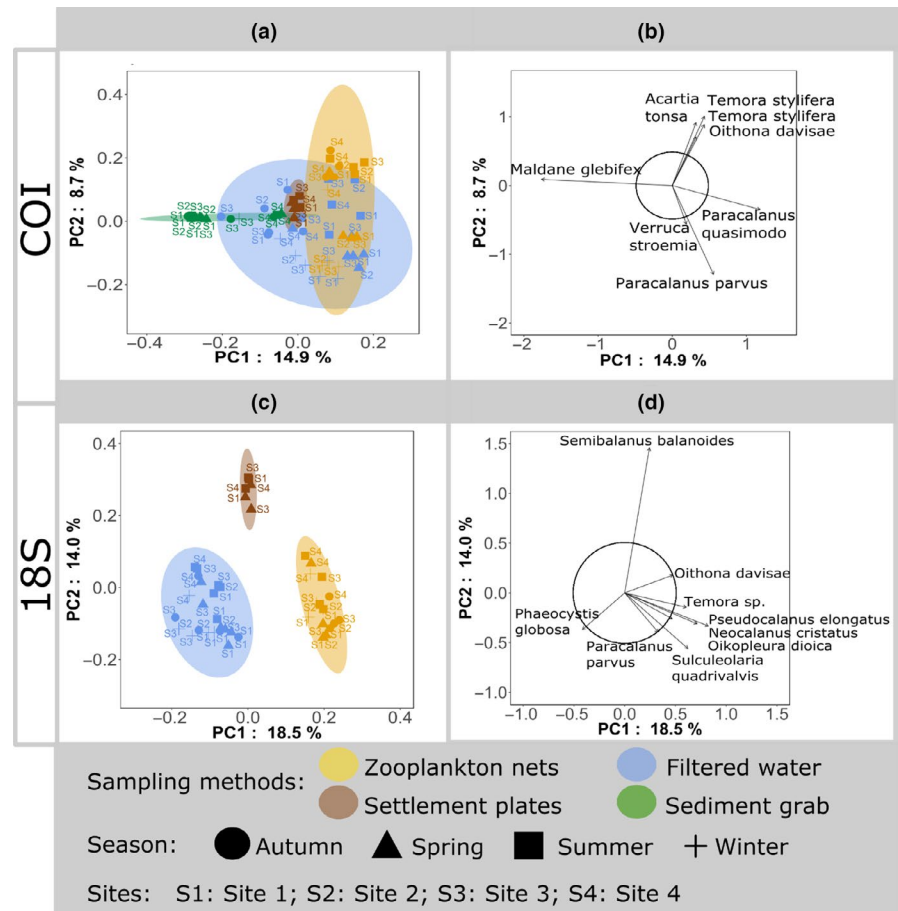
### 3.3 | Influence of sampling seasonality and locality on detected biodiversity

Seasonal variations contributed significantly to differences in community composition of zooplankton nets and filtered water but not of sediment grab (Table S2). Total richness and unique richness varied between seasons for all sampling methods (Figure 6a,b). Spring and late summer were generally richer than autumn and winter, excepted for filtered water communities represented with 18S. For sediment grabs and zooplankton nets with 18S, a strong proportion of taxa found in autumn and winter were a subset of taxa retrieved in late summer and spring. Indeed, the nestedness component of sediment grab assemblages represented between 60% and 70% of total compositional variations among these pairs of seasons (Figure 6d). For zooplankton net assemblages, 63% and 72% of nestedness were observed between late summer and spring and between late summer and autumn, respectively, suggesting that late summer recovered a majority of spring and autumn diversity (Figure 6d). Thus,

the combination of spring and late summer for sediment grabs retrieved 85% of the total OTU richness over the four seasons, while for zooplankton nets, in late summer alone 71% was retrieved (Figure 6c,e). For filtered water and settlement plates, seasonal community changes were driven by both OTU replacement and nestedness, with similar relative contributions (Figure 6d). The seasonal influence observed in intraport samples was also observed between ports (Appendix S6). The differences in communities between Bilbao, A Coruña, and Vigo (ports belonging to the same ecoregion but separated by over 500 km) during the same season were generally smaller than those between seasons in the same port, indicating that communities were driven by seasonality rather than location (Appendix S6 and Table S3). This pattern was more pronounced with 18S than with COI.

Locality within the Bilbao port appeared to impact on benthic assemblages (Table S2), since sites outside the estuary (sites 1–3) were different from site 4 inside the estuary; mainly, the polychaete *Maldane glebifex* was less abundant in site 4 (Figure 5b). Regarding zooplankton nets, site 4 was the main driver of difference in communities as, when not considered, no significant differences between sites were observed (Table S2). Each site harbored a similar proportion of the total OTU richness found by each sampling method (Figure 7c). This proportion did not exceed 60%. Indeed,

**FIGURE 5** Principal component analyses of Hellinger-transformed abundances of OTUs included in the *PBBS targeted taxa dataset* for COI (a, b) and 18S (c, d). (a & c) Samples scores in scaling 1 with ellipses representing the 95% confidence dispersion of each sampling method. (b & d) OTU scores in scaling 1 with the circle of equilibrium contribution. Only OTUs whose projected length in these two principal component axes exceeding the value of equilibrium contribution are represented



OTU replacement contributed more to community variation among sites than nestedness, especially with sediment grab and settlement plates (Figure 7d). OTU replacement was more important when comparing sites from outside the estuary (sites 1, 2, and 5) with site 4 inside the estuary, while it contributed less to community variation among sites 1, 2, and 3. This is congruent with site 4 having generally more unique OTUs than the other sites (Figure 7b). An exception was observed for filtered water with COI, where site 4 had the lowest OTU richness and unique OTU richness, and where nestedness contributed more to the variation in community composition in comparison with sites 1 and 2 (Figure 7a,b,d).

### 3.4 | Detection of nonindigenous and cryptogenic species

Our port baseline biological survey detected 79 putative NICS, among which 29 of the 68 previously recorded in the port Bilbao were found (Tables S4–S6). Most of the other species (43) were previously detected NICS inside the port's Large Marine Ecoregion (LME) but, for 7 of them (*Ammothea hilgendorfi*, *Bugulina fulva*, *Grandidierella japonica*, *Melita nitida*, *Neodexiospira brasiliensis*, *Pseudochattonella verruculosa*, and *Tubificoides pseudogaster*), it was, to our knowledge, the first report in the LME. The indicator species analysis performed on reads corresponding to nonindigenous species revealed that site 4 was

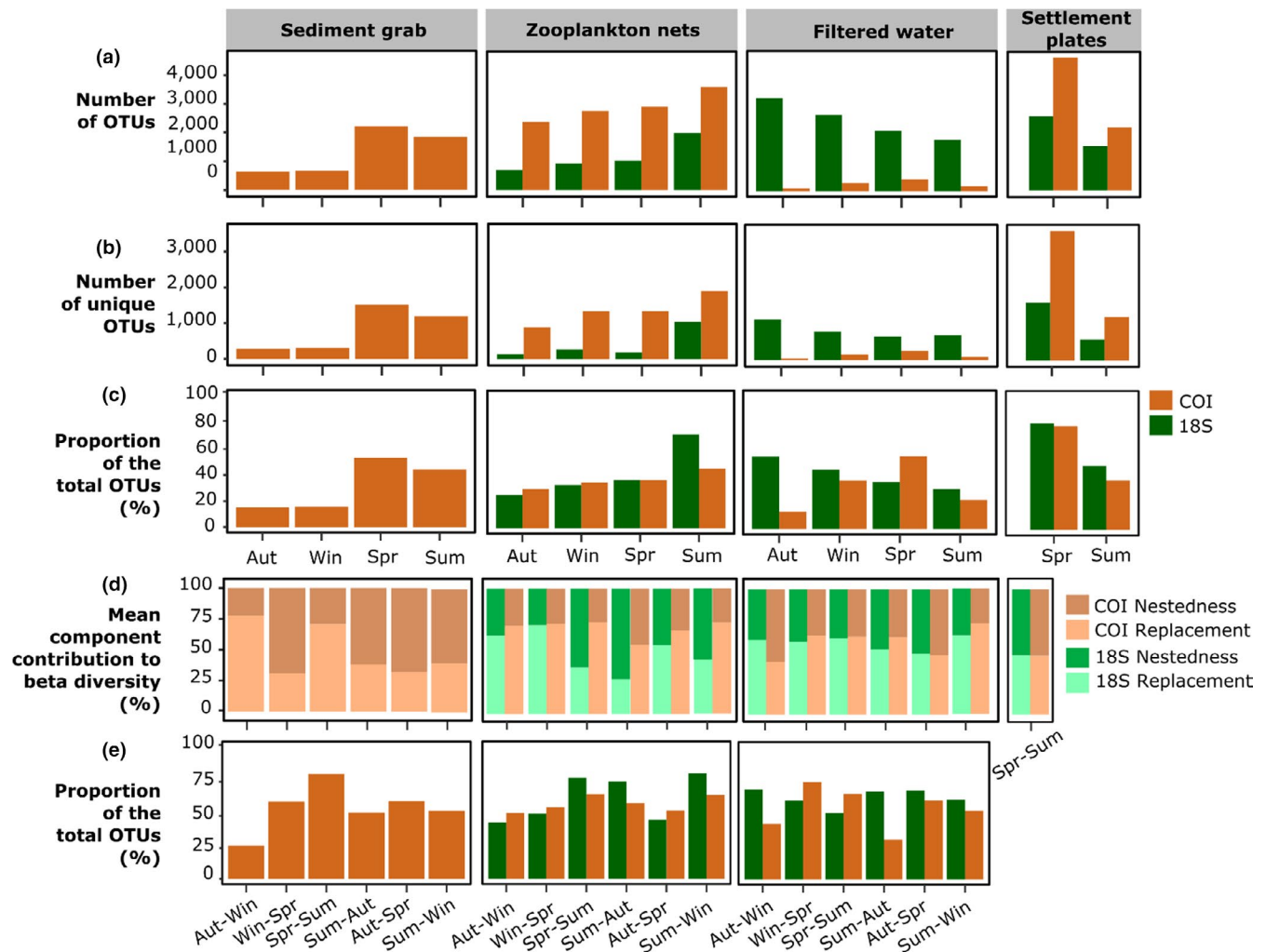
associated with the highest number of nonindigenous and cryptogenic species, which were *Oithona davisae*, *Acartia tonsa*, and *Allita succinea*, found by both barcodes, *Ficopomatus enigmaticus*, found with 18S and *Amphibalanus eburneus*, *Grandidierella japonica*, *Polydora cornuta*, *Austrominius modestus*, *Monocorophium acherusicum*, *Xenostrobus securis*, found with COI. Regarding the three other sites, for COI, site 1 was associated with *Clytia hemisphaerica* and to *Balanus trigonus*, site 2, to *Clytia hemisphaerica* and *Mytilus edulis*, and site 3, to *Balanus trigonus*. For 18S, only site 1 was associated with *Gymnodinium aureolum*.

## 4 | DISCUSSION

### 4.1 | Metabarcoding-based port baseline surveys require a combination of sampling methods and should not rely solely on eDNA

Most of the past attempts on using metabarcoding for port monitoring have relied on a single sampling method (Borrell et al., 2017; Brown et al., 2016; Grey et al., 2018; Lacoursière-Roussel et al., 2018; Zaiko et al., 2016) and only one evaluated the importance of using different sampling methods (Kozioł et al., 2019). In agreement with the latter, our metabarcoding analysis shows that each sampling method recovered a distinct subset of the port community, and that, despite some taxa being expectedly associated with a given sampling





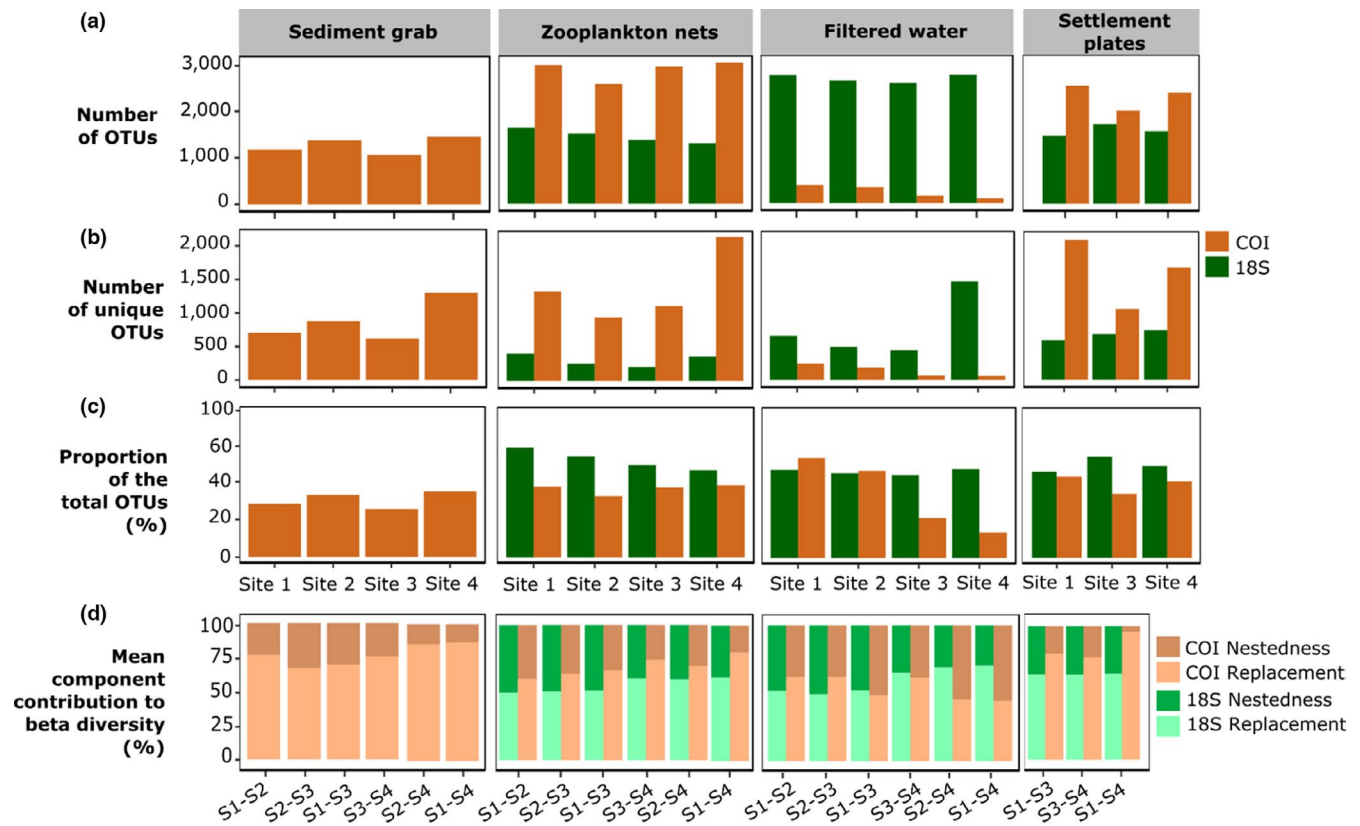
**FIGURE 6** Seasonal variation of alpha and beta diversity for each sampling method with COI and 18S. (a) Total OTU richness recovered at each season. (b) OTU richness unique to each season. (c) Proportion of the total OTU richness detected with one season. (d) Decomposition of between-season beta diversity into replacement and nestedness components. (e) Proportion of the total OTU richness detected with two seasons

method, the total diversity, even within specific taxonomic groups, was only recovered by combining different methods. Interestingly, despite the documented potential of eDNA metabarcoding to capture a large fraction of the macroorganismal diversity with limited effort (Bista et al., 2017; O'Donnell et al., 2017), including detection of nonindigenous taxa (Klymus, Marshall, & Stepien, 2017) and port surveys (Borrell et al., 2017; Grey et al., 2018; Lacoursière-Roussel et al., 2018), our analyses show that, compared to bulk sample metabarcoding, eDNA metabarcoding recovered only a subset of the metazoan diversity and did not provide additional information on targeted groups. This confirms previous findings in coral reef sites (Djurhuus et al., 2018) and freshwater streams (Macher et al., 2018) and recently in ports (Kozioł et al., 2019), suggesting that despite requiring less sampling effort, COI- and 18S-based metabarcoding of eDNA obtained from filtered water should not be used to replace the need of obtaining bulk samples (using zooplankton nets, grabs, and settlement plates) for PBBS taxa detection. Increasing evidences showed that the use of universal and degenerated COI primers is

leading to nonspecific amplification of prokaryotic and nontarget eukaryotic DNA (Collins et al., 2019). Yet, in this context, at least two alternatives are possible to improve eDNA-based biodiversity assessments: (a) increasing sequencing depth (Grey et al., 2018) or (b) using group-specific markers (Jeunen, Knapp, Spencer, Taylor, et al., 2019). Yet, none of them ensures full biodiversity recovery and both significantly increase costs. Thus, the decision between increasing sampling depth and/or using multiple group-specific primers for eDNA and adding multiple substrate sampling needs to be carefully considered and will be an important area of future research to optimize metabarcoding-based port monitoring.

#### 4.2 | Metabarcoding-based port baseline surveys should include spatiotemporal sampling

While it is expected that increased temporal coverage will retrieve more taxa, the HELCOM/OSPAR protocol, a widely applied



**FIGURE 7** Spatial variation of alpha and beta diversity for each sampling method with COI and 18S. (a) Total OTU richness recovered at each site. (b) OTU richness unique to each site. (c) Proportion of the total OTU richness detected with one site. (d) Decomposition of between-site beta diversity into replacement and nestedness components

protocol developed for PBBS ([http://jointbwmexemptions.org/ballast\\_water\\_RA/apex/f?p=104:13](http://jointbwmexemptions.org/ballast_water_RA/apex/f?p=104:13)), limits sampling to late summer for sediment and fouling, and to spring and late summer for plankton in order to reduce costs. Yet, so far, no studies using metabarcoding have been performed to support this decision. Here, we show that for sediment and fouling, spring sampling provides higher diversity than late summer. Although sampling in late summer could be appropriate for morphological taxonomy because of the more abundance of adult individuals, for metabarcoding, sampling in spring is preferable because during this season, (a) sizes of organisms are less variable and thus metabarcoding is less likely to under detect small organisms (Elbrecht, Peinert, & Leese, 2017) and (b) species diversity is at its maxima due to being a high recruitment period with abundant organisms at early life stages (Bijleveld et al., 2018). For zooplankton diversity, our results show that the HELCOM/OSPAR sampling in spring and late summer produces the highest diversity. However, the data obtained from filtered water were inconclusive. Previous studies have already shown that seasonal variations are important considerations for eDNA studies (Lacoursière-Roussel et al., 2018), but further characterizations are needed over multiple years to design an adequate protocol for maximizing biodiversity recovery with eDNA.

Concerning spatial sampling, for all sampling methods, OTU replacement generally contributed more to community variation among sites than nestedness, suggesting that spatially comprehensive

sampling is crucial to recover the port's biodiversity. Interestingly, site 4 was not only the most different from all four sites but was also the one that recorded the largest number of NIS. In coherence with our findings, it has been observed that brackish environments favor NIS settlement (Zorita et al., 2013) because these species usually support wider range of salinity (Cardeccia et al., 2018). Thus, samples for port monitoring should include those with a wide range of abiotic conditions and covering priority sampling, such as highly active ship berths, potential reservoirs of newly arrived NIS (Hewitt & Martin, 2001).

### 4.3 | Metabarcoding provides valuable information on nonindigenous and cryptogenic species

Our metabarcoding port biological baseline survey detected NICS previously recorded in the port of Bilbao and NICS known to be present in the port's marine ecoregion. Importantly, it also unveiled presence of seven NICS for which no records exist so far in this ecoregion, highlighting the potential of metabarcoding for early NICS detection. Nonetheless, not all NICS species previously recorded in Bilbao were found by our analyses. This might be due to these species not being present in the port at the time of our survey, to biases of the metabarcoding process such as differential DNA extraction and primer nonspecificity (Xiong, Li, & Zhan, 2016), or to

database incompleteness. Primer bias and/or database incompleteness could also explain why 35% and 42% of NICS were uniquely detected with COI and 18S, respectively. Yet, from the total species included in the AquaNIS database ( $n = 1,083$ ), only 460 and 369 are included in the BOLD and PR2 databases, respectively. This stresses the need of increasing reference databases and of using multiple universal primers for species detection (Grey et al., 2018). Importantly, our study confirmed that metabarcoding can detect species occurring at low abundance (Pochon, Bott, Smith, & Wood, 2013) as we found the nonindigenous amphipod *Melita nitida* with only 18 reads, while it took intensive surveys in 2013, 2014, and 2016 in three distinct sampling regions of the Bay of Biscay to record 76 individuals (Gouillieux, Lavesque, Blanchet, & Bachelet, 2016).

## 5 | CONCLUSIONS

The Ballast Water Management Convention, entered into force in September 2017, aims at preventing the spread of nonindigenous species from one region to another (IMO, 2004). Yet, the implementation of this convention still requires technological developments for assessing compliance and granting exemptions, for which genetic methods have been suggested promising (Rey, Basurko, & Rodríguez-Ezpeleta, 2018). Based on the comparative analysis of 192 samples assessing the use of alternative sampling methods, of sampling at different seasons and at different port locations, we have demonstrated the suitability of metabarcoding for port biodiversity surveys and NIS monitoring and settled the guidelines for future studies. We show that (a) combining two pairs of universal primers provides a more holistic view of the port biodiversity, (b) a combination of sampling methods is necessary to recover the different taxonomic groups, (c) environmental DNA cannot replace traditional sampling; (d) sampling should take place in spring and late summer preferably, and (e) spatial coverage should cover the port's salinity gradient. Considering the cost-effectiveness of metabarcoding with respect to morphological identification, these guidelines and considerations are particularly relevant for performing the risk assessment required for granting exemptions within the International Convention for the Control and Management of Ships' Ballast Water and Sediments.

## ACKNOWLEDGMENTS

This manuscript is a result of the Aquainvad-ED (AQUatic INVADers: Early Detection, Control and Management—www.aquainvad-ed.com) project funded by the European Union (Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no 642197). We thank the Port Authority of Bilbao for authorizing sampling in the port and Manuel González for making the necessary arrangements for it. We also thank Iñaki Mendibil and Elisabete Bilbao for laboratory work, Ekaitz Erauskin, Iker Urtizberea, and Lander Larrañaga Juarez for sampling, Anders Lanzen for providing comments on the manuscript, and Carlota Bermejo, student of the postgraduate course in Scientific Illustration of the UPV/EHU, for producing

Figure 1. This manuscript is contribution X [to be completed upon acceptance] from the Marine Research Division of AZTI.

## CONFLICT OF INTEREST

None declared.

## AUTHORS' CONTRIBUTIONS

AR, OCB, and NR-E conceived the study and designed methodology; AR and OCB collected the samples; AR performed laboratory analyses; AR and NR-E analyzed the data; and AR and NR-E wrote the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

## DATA AVAILABILITY STATEMENT

Raw demultiplexed MiSeq reads are available at NCBI SRA Bioproject PRJNA515494. Scripts used for OTU table generation and taxonomic assignment are available at <https://github.com/rodriguez-ezpeleta/metabarcoding-pbbs>.

## ORCID

Oihane C. Basurko  <https://orcid.org/0000-0002-4157-4655>

Naiara Rodríguez-Ezpeleta  <https://orcid.org/0000-0001-6735-6755>

## REFERENCES

- Adarraga, I., & Martínez, J. (2011). First record of invasive *Theora lubrica* (Mollusca: Bivalvia: Semelidae) in the Atlantic Ocean. *Marine Biodiversity Records*, 4, E100 <https://doi.org/10.1017/S1755267211001011>
- Adarraga, I., & Martínez, J. (2012). First record of the invasive brackish water mytilid *Limnoperna securis* (Lamarck, 1819) in the Bay of Biscay. *Aquatic Invasions*, 7(2), 171–180. <https://doi.org/10.3391/ai.2012.7.2.003>
- Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., & Huse, S. M. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS ONE*, 4(7), e6372. <https://doi.org/10.1371/journal.pone.0006372>
- Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- AquaNIS. Editorial Board (2015). *Information system on aquatic non-indigenous and cryptogenic species (World Wide Web Electronic Publication)*. Version, 2.36+. Retrieved from [www.corpi.ku.lt/databases/aquanis](http://www.corpi.ku.lt/databases/aquanis)
- Awad, A., Haag, F., Anil, A. C., & Abdulla, A. (2014). *Guidance on Port Biological Baseline Surveys (PBBS)*. GEF-UNDP-IMO GloBallast Partnerships, IOI, NIO, IUCN.
- Aylagas, E., Borja, Á., Irigoien, X., & Rodríguez-Ezpeleta, N. (2016). Benchmarking DNA metabarcoding for biodiversity-based monitoring and assessment. *Frontiers in Marine Science*, 3, 96 <https://doi.org/10.3389/fmars.2016.00096>
- Aylagas, E., Borja, Á., Muxika, I., & Rodríguez-Ezpeleta, N. (2018). Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. *Ecological Indicators*, 95(1), 194–202. <https://doi.org/10.1016/j.ecolind.2018.07.044>
- Aylagas, E., Mendibil, I., Borja, Á., & Rodríguez-Ezpeleta, N. (2016). Marine sediment sample pre-processing for macroinvertebrates metabarcoding: Mechanical enrichment and homogenization. *Frontiers in Marine Science*, 3, 203. <https://doi.org/10.3389/fmars.2016.00203>

- Bax, N., Williamson, A., Agüero, M., Gonzalez, E., & Geeves, W. (2003). Marine invasive alien species: A threat to global biodiversity. *Marine Policy*, 27(4), 313–323. [https://doi.org/10.1016/S0308-597X\(03\)00041-1](https://doi.org/10.1016/S0308-597X(03)00041-1)
- Bijleveld, A. I., Compton, T. J., Klunder, L., Holthuijsen, S., ten Horn, J., Koolhaas, A., ... van der Veer, H. W. (2018). Presence-absence of marine macrozoobenthos does not generally predict abundance and biomass. *Scientific Reports*, 8(1), 3039. <https://doi.org/10.1038/s41598-018-21285-1>
- Bishop, M. J., & Hutchings, P. A. (2011). How useful are port surveys focused on target pest identification for exotic species management? *Marine Pollution Bulletin*, 62(1), 36–42. <https://doi.org/10.1016/j.marpolbul.2010.09.014>
- Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., ... Creer, S. (2017). Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature Communications*, 8, 14087. <https://doi.org/10.1038/ncomms14087>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Borrell, Y. J., Miralles, L., Huu, H. D., Mohammed-Geba, K., & Garcia-Vazquez, E. (2017). DNA in a bottle—Rapid metabarcoding survey for early alerts of invasive species in ports. *PLoS ONE*, 12(9), e0183347. <https://doi.org/10.1371/journal.pone.0183347>
- Bott, N. J. (2015). Detection of invasive species. In S.-K. Kim (Ed.), *Springer handbook of marine biotechnology* (pp. 441–453). Berlin, Heidelberg, Germany: Springer.
- Brown, E. A., Chain, F. J. J., Zhan, A., Maclsaac, H. J., & Cristescu, M. E. (2016). Early detection of aquatic invaders using metabarcoding reveals a high number of non-indigenous species in Canadian ports. *Diversity and Distributions*, 22(10), 1045–1059. <https://doi.org/10.1111/ddi.12465>
- Butrón, A., Orive, E., & Madariaga, I. (2011). Potential risk of harmful algae transport by ballast waters: The case of Bilbao Harbour. *Marine Pollution Bulletin*, 62(4), 747–757. <https://doi.org/10.1016/j.marpolbul.2011.01.008>
- Cardeccia, A., Marchini, A., Occhipinti-Ambrogi, A., Galil, B., Gollasch, S., Minchin, D., ... Ojaveer, H. (2018). Assessing biological invasions in European Seas: Biological traits of the most widespread non-indigenous species. *Estuarine, Coastal and Shelf Science*, 201, 17–28. <https://doi.org/10.1016/j.ecss.2016.02.014>
- Carlton, J. T. (1999). The scale and ecological consequences of biological invasions in the World's oceans. In *Presented at the population and community biology series* (vol. 24, pp. 195–212). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Collins, R. A., Bakker, J., Wangenstein, O. S., Soto, A. Z., Corrigan, L., Sims, D. W., ... Mariani, S. (2019). Non-specific amplification compromises environmental DNA metabarcoding with COI. *Methods in Ecology and Evolution*, 10(11), 1985–2001. <https://doi.org/10.1111/2041-210X.13276>
- 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. *Biological Invasions*, 17(3), 905–922. <https://doi.org/10.1007/s10530-015-0854-y>
- Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., ... Bik, H. M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution*, 7(9), 1008–1018. <https://doi.org/10.1111/2041-210X.12574>
- David, M., Gollasch, S., & Pavliha, M. (2013). Global ballast water management and the “same location” concept: A clear term or a clear issue? *Ecological Applications*, 23(2), 331–338. <https://doi.org/10.1890/12-0992.1>
- De Cáceres, M. D., & Legendre, P. (2009). Associations between species and groups of sites: Indices and statistical inference. *Ecology*, 90(12), 3566–3574. <https://doi.org/10.1890/08-1823.1>
- Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J.-C., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyor belts of biodiversity information. *Nature Communications*, 7, 12544. <https://doi.org/10.1038/ncomms12544>
- Djurhuus, A., Pitz, K., Sawaya, N. A., Rojas-Márquez, J., Michaud, B., Montes, E., ... Breitbart, M. (2018). Evaluation of marine zooplankton community structure through environmental DNA metabarcoding: Metabarcoding zooplankton from eDNA. *Limnology and Oceanography: Methods*, 16(4), 209–221. <https://doi.org/10.1002/lom3.10237>
- Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., & Wagner, H. H. (2018). *adespatial: Multivariate multiscale spatial analysis*. Retrieved from <https://CRAN.R-project.org/package=adespatial>
- Dufrene, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, 67(3), 345. <https://doi.org/10.2307/2963459>
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27(16), 2194–2200. <https://doi.org/10.1093/bioinformatics/btr381>
- Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, 7(17), 6918–6926. <https://doi.org/10.1002/ece3.3192>
- Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, 8(10), 1265–1275. <https://doi.org/10.1111/2041-210X.12789>
- Gittenberger, A., Rensing, M., Niemantsverdriet, P., Schrieken, N., & Stegenga, H. (2014). *Port of Rotterdam survey and monitoring non-native species conform HELCOM/OSPAR protocol*.
- Gouillieux, B., Lavesque, N., Blanchet, H., & Bachelet, G. (2016). First record of the non-indigenous *Melita nitida* Smith, 1873 (Crustacea: Amphipoda: Melitidae) in the Bay of Biscay (NE Atlantic). *BioInvasions Records*, 5(2), 85–92. <https://doi.org/10.3391/bir.2016.5.2.05>
- Grey, E. K., Bernatchez, L., Cassey, P., Deiner, K., Deveney, M., Howland, K. L., ... Lodge, D. M. (2018). Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. *Scientific Reports*, 8(1), 8843. <https://doi.org/10.1038/s41598-018-27048-2>
- Hebert, P. D., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(Suppl 1), S96–S99.
- HELCOM/OSPAR (2013). *Joint HELCOM/OSPAR Guidelines for the Contracting Parties of OSPAR and HELCOM on the granting of exemptions under International Convention for the Control and Management of Ships' Ballast Water and Sediments, Regulation A-4 (OSPAR Agreement, 2013-09)*. 45 pp.
- Hewitt, C. L., & Martin, R. B. (2001). *Revised protocols for baseline port surveys for introduced marine species—design considerations, sampling protocols and taxonomic sufficiency*. Centre for Research on Introduced Marine Pests Technical Report No. 22. Hobart, TAS: CSIRO Marine Research.
- IMO (2004). *International Convention for the Control and Management of Ships' Ballast Water and Sediments, 2004*. London, UK: International Maritime Organization.
- Jeunen, G.-J., Knapp, M., Spencer, H. G., Lamare, M. D., Taylor, H. R., Stat, M., ... Gemmill, N. J. (2019). Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats

- connected by water movement. *Molecular Ecology Resources*, 19(2), 426–438. <https://doi.org/10.1111/1755-0998.12982>
- Jeunen, G.-J., Knapp, M., Spencer, H. G., Taylor, H. R., Lamare, M. D., Stat, M., ... Gemmill, N. J. (2019). Species-level biodiversity assessment using marine environmental DNA metabarcoding requires protocol optimization and standardization. *Ecology and Evolution*, 9(3), 1323–1335. <https://doi.org/10.1002/ece3.4843>
- Klymus, K. E., Marshall, N. T., & Stepien, C. A. (2017). Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PLoS ONE*, 12(5), e0177643. <https://doi.org/10.1371/journal.pone.0177643>
- Kozioł, A., Stat, M., Simpson, T., Jarman, S., DiBattista, J. D., Harvey, E. S., ... Bunce, M. (2019). Environmental DNA metabarcoding studies are critically affected by substrate selection. *Molecular Ecology Resources*, 19(2), 366–376. <https://doi.org/10.1111/1755-0998.12971>
- Kraus, R., Ninčević-Gladan, Ž., Auriemma, R., Bastianini, M., Bolognini, L., Cabrini, M., ... Žuljević, A. (2018). Strategy of port baseline surveys (PBS) in the Adriatic Sea. *Marine Pollution Bulletin*, 147, 47–58. <https://doi.org/10.1016/j.marpolbul.2018.08.067>
- Lacoursière-Roussel, A., Howland, K., Normandeau, E., Grey, E. K., Archambault, P., Deiner, K., ... Bernatchez, L. (2018). eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *Ecology and Evolution*, 8, 7763–7777. <https://doi.org/10.1002/ece3.4213>
- Legendre, P., & Gallagher, E. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129(2), 271–280. <https://doi.org/10.1007/s004420100716>
- Lehtiniemi, M., Ojaveer, H., David, M., Galil, B., Gollasch, S., McKenzie, C., ... Pederson, J. (2015). Dose of truth—Monitoring marine non-indigenous species to serve legislative requirements. *Marine Policy*, 54, 26–35. <https://doi.org/10.1016/j.marpol.2014.12.015>
- Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., ... Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, 10, 34. <https://doi.org/10.1186/1742-9994-10-34>
- Macher, J.-N., Vivancos, A., Piggott, J. J., Centeno, F. C., Matthaei, C. D., & Leese, F. (2018). Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate cytochrome c oxidase I primers. *Molecular Ecology Resources*, 18, 1456–1468. <https://doi.org/10.1111/1755-0998.12940>
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2014). Swarm: Robust and fast clustering method for amplicon-based studies. *PeerJ*, 2, e593. <https://doi.org/10.7717/peerj.593>
- Mandelik, Y., Roll, U., & Fleischer, A. (2010). Cost-efficiency of biodiversity indicators for Mediterranean ecosystems and the effects of socio-economic factors. *Journal of Applied Ecology*, 47(6), 1179–1188. <https://doi.org/10.1111/j.1365-2664.2010.01864.x>
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, 17(1), 10–12. <https://doi.org/10.14806/ej.17.1.200>
- Martínez, J., & Adarraga, I. (2006). *Programa De Vigilancia Y Control De La Introducción De Especies Invasoras En Los Ecosistemas Litorales De La Costa Vasca*. 1. Costa de Gipuzkoa, 2.
- McMurdie, P. J., & Holmes, S. (2014). Waste not, want not: Why rarefying microbiome data is inadmissible. *PLoS Computational Biology*, 10(4), e1003531. <https://doi.org/10.1371/journal.pcbi.1003531>
- Meyer, C. P. (2003). Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biological Journal of the Linnean Society*, 79(3), 401–459. <https://doi.org/10.1046/j.1095-8312.2003.00197.x>
- O'Donnell, J. L., Kelly, R. P., Shelton, A. O., Samhoury, J. F., Lowell, N. C., & Williams, G. D. (2017). Spatial distribution of environmental DNA in a nearshore marine habitat. *PeerJ*, 5, e3044. <https://doi.org/10.7717/peerj.3044>
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., ... Wagner, H. (2019). *vegan: Community Ecology Package*. R package version 2.5-5. <https://CRAN.R-project.org/package=vegan>
- Pochon, X., Bott, N. J., Smith, K. F., & Wood, S. A. (2013). Evaluating detection limits of next-generation sequencing for the surveillance and monitoring of international marine pests. *PLoS ONE*, 8(9), e73935. <https://doi.org/10.1371/journal.pone.0073935>
- Podani, J., & Schmera, D. (2011). A new conceptual and methodological framework for exploring and explaining pattern in presence – absence data. *Oikos*, 120(11), 1625–1638. <https://doi.org/10.1111/j.1600-0706.2011.19451.x>
- Pysek, P., Hulme, P. E., Meyerson, L. A., Smith, G. F., Boatwright, J. S., Crouch, N. R., ... Wilson, J. R. U. (2013). Hitting the right target: Taxonomic challenges for, and of, plant invasions. *AoB PLANTS*, 5, plt042–plt042. <https://doi.org/10.1093/aobpla/plt042>
- Rey, A., Basurko, O. C., & Rodríguez-Ezpeleta, N. (2018). The challenges and promises of genetic approaches for ballast water management. *Journal of Sea Research*, 133, 134–145. <https://doi.org/10.1016/j.seares.2017.06.001>
- RStudio Team (2015). *RStudio: Integrated development for R*. Boston, MA: RStudio Inc. Retrieved from <http://www.Rstudio.com>
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., ... Weber, C. F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, 75(23), 7537–7541. <https://doi.org/10.1128/AEM.01541-09>
- Seebens, H., Gastner, M. T., & Blasius, B. (2013). The risk of marine bio-invasion caused by global shipping. *Ecology Letters*, 16(6), 782–790. <https://doi.org/10.1111/ele.12111>
- Shaw, J. L. A., Weyrich, L., & Cooper, A. (2016). Using environmental (e)DNA sequencing for aquatic biodiversity surveys: A beginner's guide. *Marine and Freshwater Research*, 68(1), 20. <https://doi.org/10.1071/MF15361>
- Spens, J., Evans, A. R., Halfmaerten, D., Knudsen, S. W., Sengupta, M. E., Mak, S. S. T., ... Hellström, M. (2017). Comparison of capture and storage methods for aqueous microbial eDNA using an optimized extraction protocol: Advantage of enclosed filter. *Methods in Ecology and Evolution*, 8(5), 635–645. <https://doi.org/10.1111/2041-210X.12683>
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, 21(8), 2045–2050. <https://doi.org/10.1111/j.1365-294X.2012.05470.x>
- Tajadura, J., Bustamante, M., & Salinas, J. I. S. (2016). *Temporal and spatial variability in the distribution of non-indigenous zoobenthic species in the intertidal hard substrates in the 'Abra of Bilbao' Poster ISOBAY conference 2016*. Retrieved from <https://www.ehu.es/documents/5019180/0/Book-of-abstract-ISOBAY15.pdf>. ISOBAY Conference
- Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, 73(16), 5261–5267. <https://doi.org/10.1128/AEM.00062-07>
- Xiong, W., Li, H., & Zhan, A. (2016). Early detection of invasive species in marine ecosystems using high-throughput sequencing: Technical challenges and possible solutions. *Marine Biology*, 163(6), 1–12. <https://doi.org/10.1007/s00227-016-2911-1>
- Zaiko, A., Schimanski, K., Pochon, X., Hopkins, G. A., Goldstien, S., Flierl, O., & Wood, S. A. (2016). Metabarcoding improves detection of eukaryotes from early biofouling communities: Implications for pest monitoring and pathway management. *Biofouling*, 32(6), 671–684. <https://doi.org/10.1080/08927014.2016.1186165>
- Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics*, 30(5), 614–620. <https://doi.org/10.1093/bioinformatics/btt593>

Zorita, I., Solaun, O., Borja, A., Franco, J., Muxika, I., & Pascual, M. (2013). Spatial distribution and temporal trends of soft-bottom marine benthic alien species collected during the period 1989–2008 in the Nervión estuary (southeastern Bay of Biscay). *Journal of Sea Research*, 83, 104–110. <https://doi.org/10.1016/j.seares.2013.04.009>

#### SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Rey A, Basurko OC, Rodríguez-Ezpeleta N. Considerations for metabarcoding-based port biological baseline surveys aimed at marine nonindigenous species monitoring and risk assessments. *Ecol Evol*. 2020;10:2452–2465. <https://doi.org/10.1002/ece3.6071>