Status and prospects of marine NIS detection and monitoring through (e)DNA metabarcoding

Sofia Duarte[1,2*], Pedro E. Vieira[1,2], Ana S. Lavrador[1,2], Filipe O. Costa[1,2]

*corresponding author
e-mail: sduarte@bio.uminho.pt
Telephone: 00351 253601522
Fax: 00351 253678980

[1]Centre of Molecular and Environmental Biology (CBMA), Department of Biology, University of Minho, Campus de Gualtar 4710-057 Braga Portugal

[2]Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Campus de Gualtar 4710-057 Braga Portugal

Abstract

In coastal ecosystems, non-indigenous species (NIS) are recognized as a major threat to biodiversity, ecosystem functioning and socio-economic activities. Here we present a systematic review on the use of metabarcoding for NIS surveillance in marine and coastal ecosystems, through the analysis of 42 publications. Metabarcoding has been mainly applied to environmental DNA (eDNA) from water samples, but also to DNA extracted from bulk organismal samples. DNA extraction kits have been widely used and the 18S rRNA and the COI genes the most employed markers, but less than half of the studies targeted more than one marker loci. The Illumina MiSeq platform has been used in more than half of the publications. Current weaknesses include the primer-biased or faulty DNA amplification and the incompleteness of reference libraries, which can restrict the resolution and detection capacity of NIS at the species assignment step. However, occurrence of false negatives, due to faulty amplification or gaps in reference libraries, is particularly concerning in the case of NIS, where species level detection is critical. Until these weaknesses are resolved, NIS metabarcoding is ought to be supported by complementary approaches. Even so, metabarcoding has already proved to be a very sensitive and effective tool to detect small organisms or undifferentiated life stages across a wide taxonomic range; improve the spatial and temporal resolution of NIS surveillance; and be employed in ballast water management. However, before this technique can be routinely applied standardized protocols should be developed. Although specific protocols may be required for targeting particular NIS, for its general monitoring it would be vital to settle on a standard protocol able to generate comparable results among surveillance campaigns and regions of the globe,

seeking the best approach for detecting the broadest range of species, while minimizing the chances of false positives and false negatives detection.

Keywords: Marine and coastal ecosystems; non-indigenous species; biosurveillance; (e)DNA metabarcoding

1. Introduction

Along with global climate change, overexploitation, pollution and habitat destruction, the spread of non-indigenous species (NIS) is among the major threats to marine and coastal ecosystems (Molnar et al. 2008, Rilov and Crooks 2009). Non-indigenous species, which are introduced from outside of their natural (past or present) distributional range, deliberately or unintentionally by humans or other agents, can spread rapidly in the recipient system, become invasive and displace and out-compete native species, triggering severe ecological changes that threaten ecosystem integrity (e.g. biodiversity loss, loss of ecosystem services) (Molnar et al. 2008, Rilov and Crooks 2009, Simberloff et al. 2013). One of the major vectors, responsible for the transfer of marine NIS, is transoceanic shipping via ballast waters and hull fouling, or canals connecting large water bodies (Ruiz et al. 2000, Molnar et al. 2008, Hulme 2009). At smaller scales, intraregional or leisure boating may also enhance the spread of NIS (Fletcher et al. 2017, Pochon et al. 2017, Huhn et al. 2020). The ports and marinas, where these vessels dock, often act as hubs for the spread of NIS (Borrell et al. 2017, Grey et al. 2018, Holman et al. 2019), where artificial marine infrastructures (e.g. pontoons), which may be less attractive for native taxa, provide empty niches for opportunistic NIS where they can settle, establish prosperous populations and spread to close vicinities (Dafforn et al. 2015, Olenin et al. 2016).

While visual surveys based on morphological identification of taxa have been largely contributing to ascertain the current status of NIS in marine and coastal habitats (e.g. Borrell et al. 2018, von Ammon et al. 2018b), this process is expertise-demanding, laborious and time consuming. The lack of available expertise on many different taxonomic groups can also pose a major constraint to bioinvasion assessments (e.g. protists: Pagenkopp Lohan et al. 2016, 2017). Moreover, an accurate identification in aquatic systems may be hindered by low visibility waters, as well as by the presence of various life stages not amenable to morphological identification (e.g. eggs, propagules, planktonic larvae, juveniles), because organisms are not large and distinctive (e.g. microalgae, zooplankton, protists) (Pochon et al. 2013, Zaiko et al. 2015b, Pagenkopp Lohan et al. 2016, 2017) or due to low density (e.g. newly arriving NIS) (Darling and Blum 2007).

The detection of an invasive species soon after its introduction, when the population is still confined to a small area and at a low density, maximises the probability of eradication or effective local management (Simberloff 2001, Anderson 2005, Olenin et al. 2011, Pochon et al. 2013). Once established, NIS are often extremely difficult and expensive to manage (Thresher and Kuris

2004, Olenin et al. 2011). For instance, the EU has spent around 12.5 billion euros annually for controlling and mitigating the damages caused by NIS (Kettunen et al. 2009). In the U.S., the annual costs of bioinvasions have been estimated at over 100 billion dollars (Pimentel et al. 2001). Due to the above-mentioned reasons, the development of novel detection methods capable of overcoming some of these challenges becomes a priority, particularly of approaches allowing an early detection to contain and eradicate NIS before they become established, and before impacting ecosystems and provoking considerable economic costs (Darling and Blum 2007, Darling and Mahon 2011, Xiong et al. 2016, Darling and Frederick 2018). Namely, the combination of DNA barcoding (Hebert et al. 2003) with high-throughput sequencing (HTS) gives DNA metabarcoding the capacity to bolster current biodiversity assessment techniques (Hajibabaei 2012, Cristescu 2014). In (e)DNA metabarcoding, DNA is extracted from bulk collections of organisms or environmental samples (e.g. water, sediments), one short barcode locus is amplified, and resultant amplicon libraries are sequenced in HTS platforms. Thousands to millions of sequences can be generated in the same reaction, which are processed through a bioinformatic pipeline, where the sequences can be clustered into operational taxonomic units (OTUs) and these OTUs or the sequences directly (amplicon sequence variants - ASV) compared to a reference sequence database to get taxonomic identities (Hajibabaei 2012, Taberlet et al. 2012, Cristescu 2014, Creer et al. 2016), ideally at species level.

DNA metabarcoding is rapidly becoming an important approach for direct measurement of biodiversity from environments such as soil, water and air (Baird and Hajibabaei 2012, Taberlet et al. 2012, Deiner et al. 2017) and promises a number of potential benefits over traditional methods. In addition, this approach is easy to implement, which makes DNA metabarcoding one of the tools of choice for the 21st century's next-generation biodiversity monitoring (Leese et al. 2018, Zinger et al. 2019). In particular, the potential for a high sensitivity, greater throughput and accuracy, and propensity to standardisation and automation, as well as faster responses and cost effectiveness (Ji et al. 2013, Pochon et al. 2015, Leese et al. 2016, Zinger et al. 2019), makes DNA metabarcoding a particularly well-suited approach for the early detection of NIS and to improve NIS monitoring in biosurveillance programmes (Zaiko et al. 2015a,b, 2016, von Ammon et al. 2018b, Rey et al. 2019, 2020).

Despite the great potential of DNA metabarcoding to improve NIS monitoring in marine and coastal ecosystems (Pochon et al. 2013, Ardura et al. 2015, Zaiko et al. 2015a,b,c, Borrell et al. 2018, Westfall et al. 2020), its wider and routine implementation is still highly dependent on reliable and cost-effective standardization (Leese et al. 2016, Rey et al. 2020). In the current study we undertake a comprehensive compilation and evaluation of methodologies employed across the full workflow (e.g. sampling strategy, DNA extraction, markers targeted, platform, sequences processing), of DNA metabarcoding applied to eukaryotic NIS detection. This information will allow to better appraise the current potential for application of metabarcoding in bioinvasion

ecology and assess major weaknesses, challenges and hindrances that still need to be addressed for its wider implementation in NIS detection in aquatic ecosystems.

2. Methodology

To reach our objective we collected papers published in the Web of Science database on 28th March 2020. We searched for all possible combinations of terms that can be used to designate non-indigenous species (i.e. non-indigenous species, invasive species, alien species, exotic species, marine pest) and DNA metabarcoding (i.e. metabarcoding, high throughput sequencing and next generation sequencing) plus the term "marine" or "coastal", with the exception of "marine pest", which was used only in combination with the metabarcoding terms. The terms "biosecurity", "biosurveillance" and "biofouling" were also used in combination with the metabarcoding terms plus "marine" or coastal". The term "ballast water" was also used in combination with the metabarcoding terms.

The information retrieved from each selected publication included (but was not limited to): i) the geographic area, ii) the sampled substrates and organisms targeted; iii) the sampling strategy (e.g. nº of sites, nº of sample replicates, nº of technical replicates), iv) the DNA extraction protocols, v) the marker loci targeted and primer pairs used, vi) the sequencing platforms and vii) the bioinformatic pipelines employed to analyse the data.

3. Results and discussion

Our literature search yielded 121 papers published between 2010 and 2020. The terms "invasive species" or "non-indigenous species" in combination with "metabarcoding" or "high throughput sequencing" and "marine", and "ballast water" in combination with "metabarcoding" or "high throughput sequencing", were those that displayed more publications (Fig. S1). After a careful inspection we found 34 papers that indeed employed metabarcoding for NIS surveillance in marine and coastal ecosystems and that were retained to perform our analysis, to which we added 8 additional papers (Chain et al. 2016, Ardura et al. 2017, Deiner et al. 2018, Grey et al. 2018, Koziol et al. 2018, Lacoursière-Roussel et al. 2018, Leduc et al. 2019, Suarez-Menendez et al. 2020), which were not displayed during our search, but that have relevant information (Table S1). The number of studies employing metabarcoding in NIS surveillance have been increasing over the years, at a rate of ca. 6.3 papers/year (Fig. S2). Top countries, where most of the studies have been conducted, include New Zealand, Canada, Spain, USA and Australia (Fig. S3), which are countries that since very early have adopted the use of molecular methods to support management decisions (Darling and Mahon 2011, Xiong et al. 2016). New Zealand is one of the worlds' countries with the highest rates of non-indigenous species introduction in marine waters (Wood et al. 2013). For instance, the cumulative number of species from 2009 to 2015 has risen 10% (http://www.environmentguide.org.nz/issues/marine/marine-biosecurity/invasive-marine-

species/), while in North America, the number of invasive species has more than tripled since the beginning of the century (MEA 2005, Leduc et al. 2019).

3.1. Sampled locations, substrates and organisms

Most of the studies have been conducted in ports or marinas or in the close vicinities (e.g. estuaries), since they are hubs of maritime traffic, where vessels from all continents stop for days or months unintentionally releasing undesired species to the environment (Table S1, Fig. 1A). In these studies, most sampled substrates were sea water and ships ballast water and the majority also applied metabarcoding to environmental DNA (eDNA) (Table S1, Fig. 1B) (e.g. Borrell et al. 2017, Deiner et al. 2018, Grey et al. 2018, Rey et al. 2019, Wood et al. 2019, Suarez-Menendez et al. 2020) or to zooplankton (e.g. Zaiko et al. 2015a,b, Abad et al. 2016, Ghabooli et al. 2016, Darling et al. 2018, Lin et al. 2020).

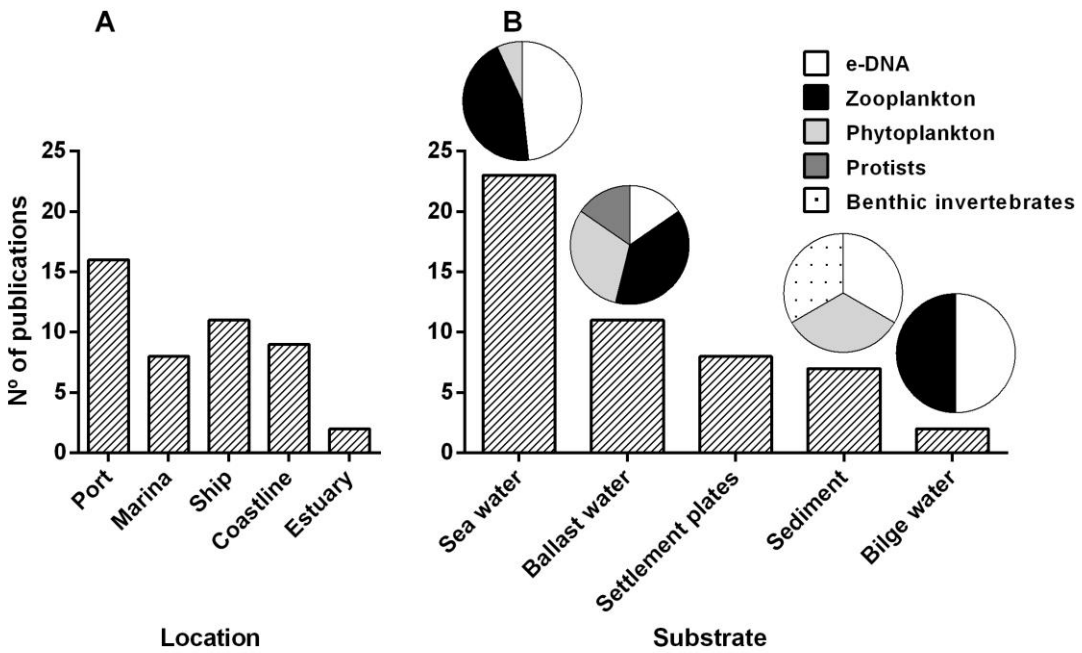

Figure 1. Type of locations (A) and substrates and organisms (B) sampled in publications using DNA metabarcoding in NIS surveillance in marine and coastal ecosystems.

Environmental DNA can be extracted directly from environmental samples, such as water and sediments, and can be originated from organisms' body parts, mucus, tissues, cells, secreted faeces or extracellular DNA, leaking to the environment through the normal life or upon death of organisms (Turner et al. 2014, Sassoubre et al. 2016, Holman et al. 2019). Targeting eDNA can be particularly advantageous to provide a rapid snapshot or baseline information of local species composition and to detect early the presence of NIS, without the need for sampling individual organisms (Ficetola et al. 2008, 2018, Hajibabaei et al. 2011, Taberlet et al. 2012, Singer et al.

2019) and has been applied exclusively in 11 biosurveillance studies employing metabarcoding (Table S1). Repeated eDNA surveys may have also the potential to be used to evaluate long-term biodiversity changes and NIS range expansions (Lacoursière-Roussel et al. 2018, Leduc et al. 2019), or to monitor NIS in ballast waters or to test the efficiency of treatment systems (Ardura et al. 2015, Rey et al. 2019, Lin et al. 2020). However, studies sampling both eDNA and bulk organismal samples suggest that eDNA alone cannot substitute for organismal sampling (e.g. Leduc et al. 2019, Huhn et al. 2020, Rey et al. 2020, Westfall et al. 2020). Currently, it would probably work best as an early alert system or an exploratory method to be used as a complement of conventional approaches such as physical *in situ* confirmation (Borrell et al. 2018, Holman et al. 2019, Rey et al. 2020), specimens collections (Leduc et al. 2019, Huhn et al. 2020) or followed by active surveillance with digital drop PCR (ddPCR) or quantitative PCR (qPCR) (Wood et al. 2019).

Zooplankton has been by far the most targeted organisms in marine NIS surveillance using metabarcoding (Fig. 1B), particularly in ships ballast water (Zaiko et al. 2015a,b, Ghabooli et al. 2016, Darling et al. 2018, Lin et al. 2020), boats bilge water (Fletcher et al. 2017), and in sea water in ports and marinas or nearby natural habitats (Pochon et al. 2013, Zaiko et al. 2015c, Abad et al. 2016, Brown et al. 2016, Chain et al. 2016, Stefanni et al. 2018, Couton et al. 2019, Leduc et al. 2019, Pagenkopp Lohan et al. 2019, Rey et al. 2020, Westfall et al. 2020). The monitoring of larvae in plankton samples in ports or in the close vicinities, may provide key information about NIS introduction status and detect early its presence (Couton et al. 2019). However, targeting dispersive life stages of NIS in the water column has been mostly unpracticable in routine monitoring using traditional methods, due to the small size and cryptic nature of planktonic taxa and larval stages, as well as the lack of specific taxonomical expertise, making them well-suited candidates to be identified through metabarcoding. For instance, in the pioneer study of Pochon et al. (2013), a single spiked *Asterias amurensis* larvae was detected in plankton samples through metabarcoding, despite the co-occurrence of a large number of other eukaryotes. Several other planktonic larval NIS stages have been exclusively detected through metabarcoding (Zaiko et al. 2015c, Fletcher et al. 2017, Couton et al. 2019), revealing its enormous potential to detect NIS taxa represented by early-life stages and small juveniles, including short-lived larval stages (e.g. *Bugula neritina*, *Corella eumyota*; Couton et al. 2019), which otherwise might remained unnoticed for long time. Metabarcoding is also becoming an accurate tool in the detection and identification of many phytoplankton species (Zaiko et al. 2015a,b, Wright et al. 2019), that do not fall into the regulated 10-50 µm size class, and in the identification of parasitic protists, particularly toxic dinoflagellates (Pagenkopp Lohan et al. 2016, 2017, Shang et al. 2019), transported in the ballast water of ships.

The identification of hard-substratum dwelling organisms, that become part of biofouling communities in natural and artificial substrates (e.g. settlement plates, Table S1, Fig. 1B), can

also greatly benefit from the use of metabarcoding (Pochon et al. 2015, Zaiko et al. 2016, Koziol et al. 2018, von Ammon et al. 2018a,b, 2019, Wood et al. 2019, Rey et al. 2020). Metabarcoding can be particularly advantageous in detecting early biofouling assemblages from scarce genetic material (Pochon et al. 2015). For instance, many metazoan genera containing NIS or cryptogenic species were detected through metabarcoding in early fouling communities (e.g. *Ciona* sp., Pochon et al. 2015; *Watersipora* sp., *Ciona* sp., *Molgula* sp. and other Bothryllid ascidians; Zaiko et al. 2016), in settlement plates deployed in New Zealand's marinas (Pochon et al. 2015, Zaiko et al. 2016). In addition, results were consistent with the traditional approach at high taxonomic levels, but DNA metabarcoding was able to detect many more taxa at the genera/species level (Zaiko et al. 2016). This is extremely relevant for the early detection of NIS in vessels or equipment from high-risk regions that can be rapidly checked prior to arrival at destination (Pochon et al. 2015), when the species are still at low densities and morphologically unidentifiable.

Compared to water, the collection of sediments to survey eukaryotic organisms has been receiving much less attention, and to our knowledge has been attempted only in 6 studies (Table S1, Fig. 1B, Koziol et al. 2018, Holman et al. 2019, Leduc et al. 2019, Shang et al. 2019, Shaw et al. 2019, Rey et al. 2020). Metabarcoding of eDNA extracted from sediments can provide a rapid screen for a wide range of taxa in ballast tanks and marine port sediments (Holman et al. 2019, Shang et al. 2019, Shaw et al. 2019). In particular, cysts of toxic dinoflagellates and diatoms resting stages are prime candidates for successful transport in sediments from ballast tanks (McCarthy and Crowder 2000, Shang et al. 2019). Its early detection through metabarcoding can help to combat its spread to new locations, where they can cause toxic blooms and have devastating impacts on ecosystems (Shang et al. 2019, Shaw et al. 2019). On the other hand, benthic invertebrates have been only surveyed in two studies using metabarcoding in NIS surveillance (Leduc et al. 2019, Rey et al. 2020). Benthic habitats can be very complex and benthic communities very patchy, which increases the possibility of missing taxa when sampling invertebrates from sediments and, thus, a less used methodology to survey NIS (Leduc et al. 2019). In addition, many marine benthic NIS have a biphasic bentho-pelagic life cycle and targeting their pelagic larval stages in zooplankton its easier and may allow its early detection before establishment and potential spread (Zaiko et al. 2015c, Couton et al. 2019).

3.2. Sampling strategy

We found that most of the studies have sampled multiple ports or marinas (e.g. Brown et al. 2016, Chain et al. 2016, Borrell et al. 2017, Leduct et al. 2019) or estuaries (Borrell et al. 2018), and multiple sites within each port/marina/estuary (Abad et al. 2016, Brown et al. 2016, Leduct et al. 2019, Rey et al. 2020). In order to be comprehensive, biodiversity surveys may need an extensive spatial and temporal coverage and sampling of different substrates (e.g. water, sediment, bulk

samples), so that most biodiversity is captured by the collected samples, in particular rare and/or patchy species (Koziol et al. 2018, Huhn et al. 2020, Rey et al. 2020, Westfall et al. 2020). However, the sampling strategy varied significantly across studies and very few have compared the effects of the sampling effort or of using different substrates in NIS detection (see Zaiko et al. 2016, Koziol et al. 2018, Lacoursière-Roussel et al. 2018, Wood et al. 2019, Rey et al. 2020, Westfall et al. 2020) (Table S1). Conducting trial experiments would be particularly helpful to assess the appropriateness of the sampling strategies employed and how these factors can affect NIS detection at particular sites in coastal and marine ecosystems (Deiner et al. 2018, Koziol et al. 2018, Pagenkopp Lohan et al. 2019, Huhn et al. 2020, Rey et al. 2020). For instance, based on an extensive survey conducted in the port of Bilbao (Spain), Rey et al. (2020) found a strong influence of the sampling site in the recovered biodiversity, suggesting that spatially comprehensive sampling was crucial to recover biodiversity, and thus increase the success of NIS detection. Also, increased temporal coverage has been shown to increase the diversity recovered, as well as the ability to detect NIS (Chain et al. 2016, Lacoursière-Roussel et al. 2018, Couton et al. 2019, Rey et al. 2020, Westfall et al. 2020). Marine communities can shift in community composition across seasons and reproductive cycles (Sutherland and Karlson 1977, Bachelet et al. 2009). However, since increasing the spatial and temporal scales will also highly increase the associated costs, due to extensive sampling and sequencing, in the case of NIS detection specific seasons may be selected. Rey et al. (2020) suggested that sampling in temperate latitudes may be shortened to spring and late summer. Yet, the metabarcoding approach has not been applied to survey a broad variety of NIS across a wide geographic range and in complex natural communities (Table S1, Fig. S3). Thus, further characterizations are needed over multiple years and geographic locations to reach a consensus to design adequate protocols in order to maximize NIS detection.

In our review, we found that DNA metabarcoding has been applied to DNA extracted from bulk organismal samples or to eDNA, namely water (see previous sub-section). However, marine NIS can have different life history traits, such as habitat preferences and life cycles, which will likely influence the detection of taxa in a given substrate. Preliminary studies that have compared different substrates found that the total diversity recovered and distinct subsets of NIS, even within specific taxonomic groups, have been accomplished only when combining different methods (Koziol et al. 2018, Leduc et al. 2019, Huhn et al. 2020, Rey et al. 2020, Westfall et al. 2020). For example, Koziol et al. (2018) found that fish were preferentially detected in water, nematodes in sediments and crustacea in planktonic tows in two Australian coastal ports. Bryozoans and ascidians are incrusting organisms and thus highest detection rates are expected to be found in settlement plates or close to hard substrata. Rey et al. (2020) also found that different sampling methods (i.e. PVC plates deployment, plankton nets, water and sediment sampling) resulted in distinct communities surveyed in the port of Bilbao (Spain). Thus, the substrate selected can have important implications for an effective detection of NIS and

combining different sampling methods may be crucial to maximize NIS detection at a particular site (Koziol et al. 2018, Huhn et al. 2020, Rey et al. 2020, Westfall et al. 2020). Within each site, the use of several replicates per substrate may also be better suited to maximize the chances of NIS detection (Koziol et al. 2018, Holman et al. 2019, Westfall et al. 2020).

3.3. Sample pre-processing before DNA extraction

We found very few studies that have analysed the effects of sample pre-processing (i.e. samples homogenization before DNA extraction, samples refrigeration or freezing before water filtration, type of filters used for water filtration) in marine NIS surveillance studies through DNA metabarcoding (but see Deiner et al. 2018, Pagenkopp Lohan et al. 2019) (Table S1, S2). Pagenkopp Lohan et al. (2019) found that homogenizing zooplankton samples with a mortar and a pestle collected from USA bays, prior to DNA extraction, increased the recovered diversity, particularly of rare or low abundance OTUs. Several other authors have pre-processed their samples prior to DNA extraction, resulting in a broader detection of taxonomic richness (Deiner et al. 2018, Shaw et al. 2019). For example, Deiner et al. (2018) mixed 13 L of sea water collected in the port of Southampton (UK) in a container, that was maintained under continuous stirring before water filtration, to ensure homogeneity among experimental replicates. Also, Shaw et al. (2019) homogenised all sediment samples, within Falcon tubes, collected from ballast water tanks from ships arriving to Australian ports, prior to DNA extraction.

In what respects water filtration, different volumes and types of filters with different composition and pore size have been used in eDNA studies (Table S2, Fig. 2). Since marine water masses are enormous and highly dynamic (exposed to tides, currents), the eDNA of rare species may be highly diluted and a high-water sampling volume may be needed to increase the chances of species detection (Borrell et al. 2017). This can constitute a practical problem in routine surveys, because such a large volume of water cannot be easily refrigerated or frozen until analysis, so should be filtered *in situ* or frozen (Jerde et al. 2011, Goldberg et al. 2016, Hinlo et al. 2017, Majaneva et al. 2018), and most authors have opted for smaller volumes (30 mL to 6 L) (Table S2), choosing a compromise between feasibility, in terms of sample storage and processing time, and taxa representativeness for a given location. One option to increase the throughput volume of capture filters and reduce filtration time may be provided by the use of pre-filters with a larger pore size, followed by a second filtration using a smaller volume through a smaller pore (Ardura et al. 2015, von Ammon et al. 2019, Wood et al. 2019) (Table S2). Another option may be filtering small volumes through independent filters, in order to avoid filter clogging, but pooling results together during sequence analysis (Leduc et al. 2019, Suarez-Menendez et al. 2020).

Another technical aspect that varied across studies is the depth at which water sampling is conducted (Table S2). For instance, most studies sampled surface water (10-30 cm) (Borrell et al. 2017, 2018, Deiner et al. 2018, Grey et al. 2018, Koziol et al. 2018, Holman et al. 2019, Huhn et

al. 2020, Suarez-Menendez et al. 2020), but many others opted to sample different depths in order to increase sample representativeness, since vertical eDNA distribution in the water column may vary as a function of several factors (e.g. life cycle of species, transport, and other complex hydrodynamic processes) (Lacoursière-Roussel 2018, Leduc et al. 2019, von Ammon et al. 2019, Rey et al. 2020, Westfall et al. 2020) and DNA may be less degraded at higher depths.

In addition, the amount of sediment used for eDNA extraction varied considerably (Table S3, 0.25-10 g), depending on the target organisms (Koziol et al. 2018, Holman et al. 2019, Shang et al. 2019, Shaw et al. 2019). In general, studies targeting macro-eukaryotes (Pochon et al. 2013, Holman et al. 2019) used higher amounts of sediment (5-10 g), while lower amounts have been used for targeting micro-eukaryotes (0.25-0.45 g) (Shang et al. 2019, Shaw et al. 2019).

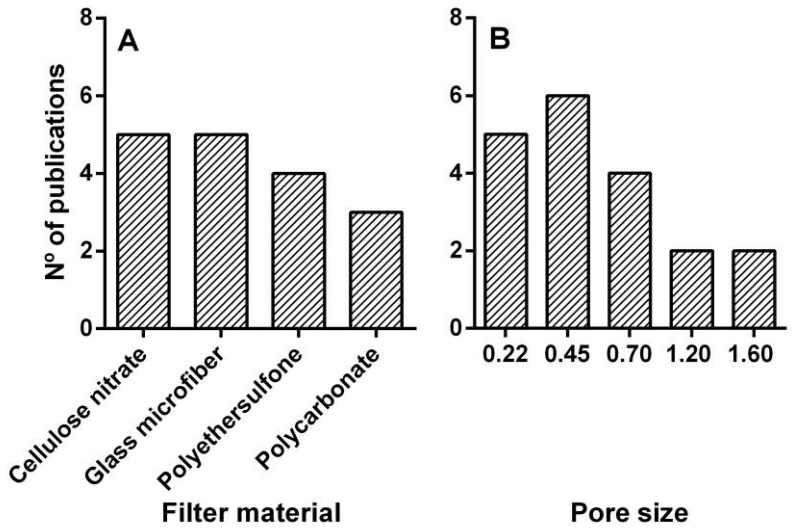

Figure 2. Filter type (A) and pore size (B) used to filter water samples in marine and coastal ecosystems in NIS surveillance publications using eDNA metabarcoding.

Concerning the effects of the filter material and pore size, to our knowledge there was only one study that tested its effects on eDNA from NIS detected in water samples (Deiner et al. 2018). The authors found that the filter material (glass microfiber *versus* cellulose nitrate), but not the pore size (0.2 to 1.2 µm), affected OTUs richness detected from seawater using eDNA metabarcoding in the port of Southampton (UK) (Deiner et al. 2018). Both types of filters and a 0.45 µm pore size, have been widely used in NIS surveillance through eDNA metabarcoding (Fig. 2). Increasing the pore size (e.g. 0.8 µm) may widely improve filtration time without compromising the captured eDNA (Li et al. 2018). But, a greater OTU recovery was found with cellulose nitrate filters in the study of Deiner et al. (2018). Thus, more targeted studies addressing technical details regarding eDNA sampling collection and processing are required, in particular the choice of membrane filters, where both the material and pore size can play important roles in

eDNA capture (Deiner et al. 2018). The use of enclosed filters may reduce handling and exposure to outside stress compared with other filters, contributing to better eDNA results (Spens et al. 2017) and may be better suited to perform on-site filtration (Li et al. 2018). But, to our knowledge, they were employed only in two studies using eDNA metabarcoding in NIS surveillance (Holman et al. 2019, Rey et al. 2020).

3.4. DNA extraction

In what respects DNA extraction, most of the studies have used DNA isolation kits (39 publications, Table S1, Fig. 3). In particular, MoBio/Qiagen kits have been used in 32 out of 42 studies (Table S1, Fig. 3), which have been reported to yield high quality DNA for barcoding or metabarcoding applications (Borrell et al. 2017, Djurhuus et al. 2017). In addition, PowerSoil and PowerMax soil DNA isolation kits are particularly useful for sediment samples and other complex samples due to the copurification of inhibitors during the protocol, that without removal may affect considerably the detectable diversity (Hermans et al. 2018). Commercial DNA isolation kits have the advantage of the use of limited and smaller amounts of chemicals, shortened isolation steps, and more rapid achievement of results. However, the use of kits has usually associated high-costs. Cost-effective extraction methods may increase the ability to process a large number of samples and can also produce good results, but may need extra purification steps (Deiner et al. 2018, Lacoursière-Roussel et al. 2018, Leduc et al. 2019).

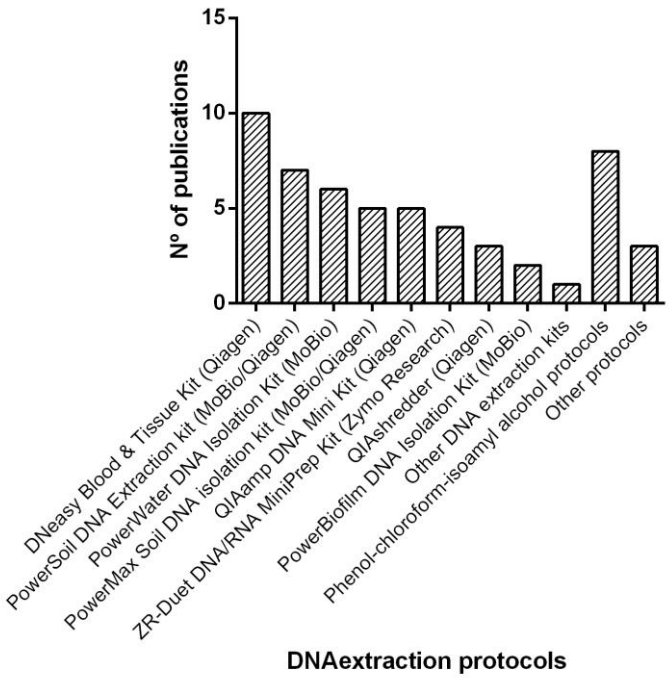

Figure 3. DNA extraction protocols used in publications using DNA metabarcoding in NIS surveillance in marine and coastal ecosystems.

To our best knowledge only one study assessed the efficiency of different DNA extraction methods in NIS detection rates (Deiner et al. 2018). Deiner et al. (2018) found that both filter material (see the previous sub-section) and DNA extraction protocols significantly affected the detection rate of the invasive ascidian *Styela clava*, in the port of Southampton (UK) (Deiner et al. 2018). Many organisms may indeed remain unidentified using DNA metabarcoding due to deficient DNA extraction (e.g. those that are rare, small or calcified). The polychaete worm *Spirorbis* sp. that was reported through traditional methods on 22% of PVC deployed plates in a New Zealand's major shipping port, was not detected by using DNA metabarcoding, although represented by two reference sequences in the Protist Ribosomal Reference database (PR2) (Zaiko et al. 2016). The authors pointed out that one possible reason may be due to inefficient DNA extraction, though faulty amplification cannot also be excluded.

The inclusion of technical replicates during DNA extraction of the same sample (e.g. Brown et al. 2016, Ghabooli et al. 2016, Shaw et al. 2019), in particular if DNA concentrations are low, and also later during PCR for the same DNA extract (e.g. Brown et al. 2016, Chain et al. 2016, Ghabooli et al. 2016, Pagenkopp Lohan et al. 2016, Koziol et al. 2018, Lacoursière-Roussel et al. 2018, Shaw et al. 2019, Rey et al. 2020), should be considered during the experimental workflow to disentangle the effects of technical variance and increase the rate of NIS detection (Borrell et al. 2017, Lanzén et al. 2017). For instance, Pagenkopp Lohan et al. (2019) found a significant effect on the recovered zooplankton diversity, collected at San Diego bays (USA), by increasing the number of DNA extraction replicates. Deiner et al. (2018) used a statistical power analysis *a priori* of their study to determine the number of experimental replicates needed for DNA extraction, in order to increase the confidence in their results. Since the DNA extraction method can strongly influence the biotic composition of the samples (Hermans et al. 2018), different methods might need standardization at some steps of the purification protocols in order to yield comparable results. For example, Koziol et al. (2018) standardized different DNA extraction methods by using a standardized silica-binding purification process to all methods, in order to better compare the recovered diversity from different substrates.

3.5. Use of controls

The inclusion of negative and positive controls, during the sampling steps and that are concurrently analysed together with the biological samples through the complete workflow of DNA metabarcoding (i.e. DNA extraction, PCR and sequencing), can facilitate the detection of contamination or potential artefacts introduced at any step (Zinger et al. 2019). Negative controls (i.e. DNA storage or extraction buffers, sterilized ultra-pure water) have been frequently used in NIS surveillance studies employing (e)DNA metabarcoding, from field sampling until sequence processing, while positive controls (i.e. DNA from mock communities or samples with *a priori* known composition) have not been so frequently used (Table S1). However, positive controls or

prior expectations of the taxa that should occur in the system - which in the case of NIS updated lists would be mandatory - can be used to evaluate the effectiveness of the metabarcoding approach (Zinger et al. 2019). For instance, Zaiko et al. (2016) included DNA samples from 3 NIS (*Sabella spallanzanii*, *Ciona savignyi* and *Perna perna*) as internal quality controls. After purification and quantification of PCR amplicons, these were combined at 3 distinct concentrations and 2 replicates of each were incorporated into the HTS library. The taxonomic assignment of the internal control samples confirmed the robustness of the sequencing and analytical pipeline employed in the studies, with more than 99.9% of the sequences produced being assigned to the target taxa (Zaiko et al. 2016). This procedure was adopted in further studies by others (Fletcher et al. 2017, Pochon et al. 2017). Grey et al. (2018) used mock communities composed by DNA from fish species to assess sequencing efficiency. For instance, all species were sequenced and correctly assigned in the COI data, but only 3 species were sequenced and none correctly assigned in the 18S data (Grey et al. 2018).

PCR inhibition testing controls could also be performed to ensure that extracted DNA is free of PCR inhibitors, by using qPCR (Holman et al. 2019, Wood et al. 2019, Westfall et al. 2020). In addition to the positive controls above mentioned, the introduction of a known amount of DNA from a non-target species, very unlikely to occur in the surveyed environment, or of artificially synthetized DNA, known *a priori* to amplify with the primers in use could also be used.

Inhibition can be overcome through the addition of PCR enhancers (e.g. BSA, Suarez-Menéndez et al. 2020), through dilution of sample DNA (not always advisable when working with low concentrations) (Wood et al. 2019), or through additional inhibitor removal steps after DNA extraction (Deiner et al. 2018, Wright et al. 2019).

The inclusion of these controls through the analytical chain of the metabarcoding approach are extremely important in order to understand the nature of two types of errors that may arise during the use of DNA-based tools: i) the Type-I errors or false positives namely due to contaminations (field, DNA extraction, PCR), “tag/index jumps”, or PCR and sequencing errors, and ii) the Type-II errors or false negatives, due for example to partial sampling, DNA extraction, amplification or sequencing bias (Zinger et al. 2019), as both can trigger action or inaction when not required, causing a potentially large burden on entities responsible for invasive species mitigation and control (Pochon et al. 2017). The concomitant analysis with the morphological approach (e.g. Zaiko et al. 2015b,c, 2016, Abad et al. 2016, Fletcher et al. 2017, Huhn et al. 2020) or at least coarse visual inspections (Borrel et al. 2018, Holman et al. 2019) can be also very important to better detect metabarcoding false negatives and to understand its nature (incompleteness of reference libraries *versus* methodological limitations such as the above-mentioned), but have been considered in less than half of the studies (Table S1).

In addition, finding evidence of DNA from a species, particularly in environmental samples, may not necessarily warrant that it belongs to a living organism (Pochon et al. 2017, von Ammon et

al. 2019), that can in fact become an invader and pose risks to ecosystems. The use of RNA may offer an alternative for identifying living organisms in samples, since it is directly linked with active gene expression of metabolic pathways and degrades rapidly upon cell death (Pawlowski et al. 2014, Pochon et al. 2017, von Ammon et al. 2019, Wood et al. 2020). In a recent study the authors found eRNA persisting in the marine environment longer than expected, which may provide new opportunities for improving biodiversity surveys through reducing false positives (Wood et al. 2020). However, working with RNA requires particular storage conditions of samples, and expensive and time-consuming workflow protocols and much less used (von Ammon et al. 2019, Wood et al. 2019).

3.6. Marker loci and primer pairs

The identification at higher taxonomic levels (e.g. family or order), is often enough to broadly estimate biodiversity through DNA metabarcoding (e.g. Creer et al. 2010, Bik et al. 2012). But, in the case of NIS detection, species-level identification is often mandatory, and this is highly dependent on the choice of marker and primers. The so-called “universal” or “near universal” primers, targeting the nuclear SSU or 18S rRNA (18S) and the mitochondrial cytochrome C oxidase subunit I (COI) genes, have been the most widely used in studies of DNA metabarcoding targeting NIS in marine and coastal ecosystems (in 31 and 22 publications, respectively, Table S1, Fig. 4).

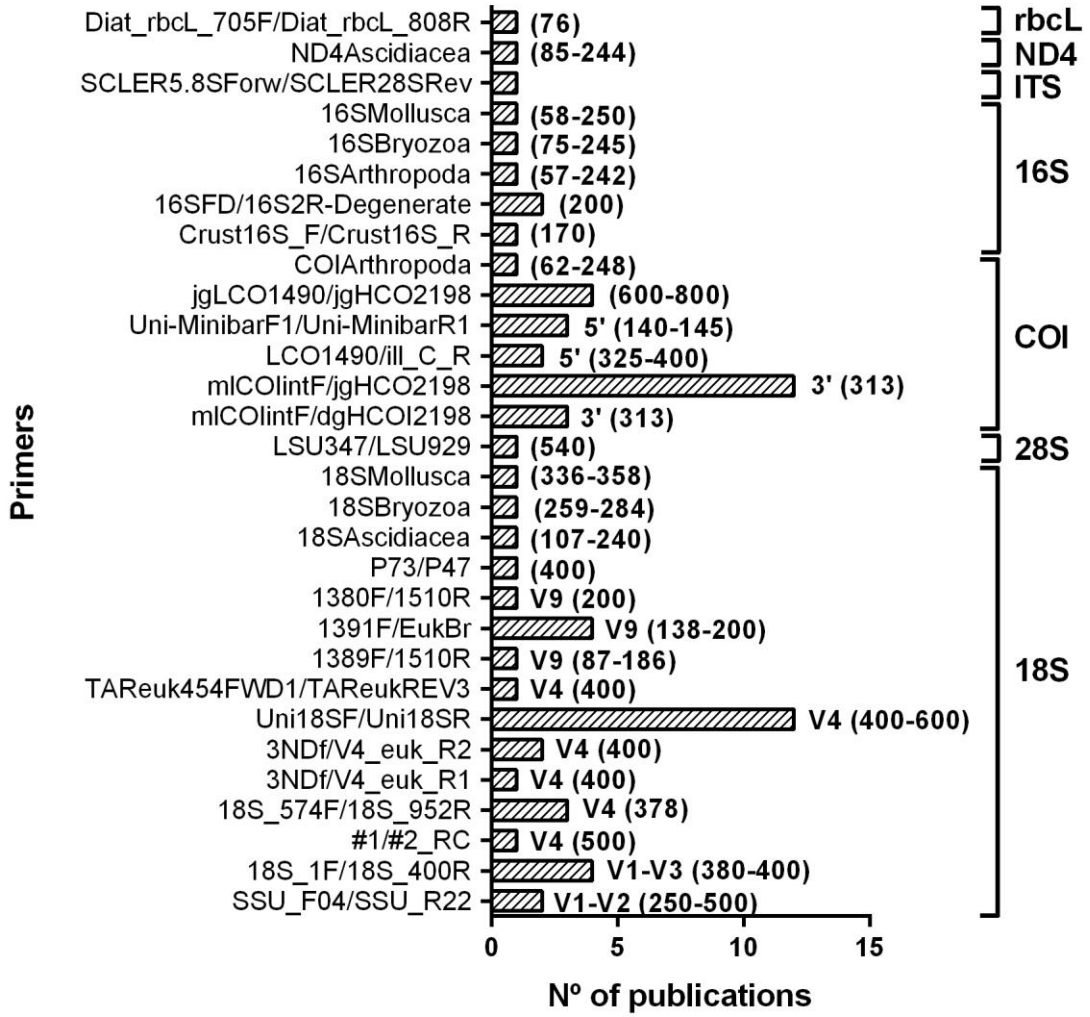

Figure 4. Marker loci and primer pairs used in publications using DNA metabarcoding in NIS surveillance in marine and coastal ecosystems. Within each marker, specific regions used are indicated in the top of each bar, as well as the approximate amplicon length between brackets.

Different regions of 18S have been predominantly used when considering marine microbial eukaryotic diversity (Amaral-Zettler et al. 2009), since it allows assessment of both metazoan and protist communities (Darling et al. 2018), and is the formal DNA barcode adopted for the later (Pawloski et al. 2012). Most of these studies employing 18S targeted the V4 hypervariable region of 18S (Fig. 4) that displays a high amplification success covering major invertebrate groups of Crustacea, Mollusca and Tunicata (Zhan et al. 2013). High amplification success was also found for additional relevant taxonomic groups such as Ciliophora, Alveolata, but also for main groups of algae, i.e. Chlorophyta or Rhodophyta (Briand et al. 2018). Although higher successful amplification rates are usually achieved with 18S due to higher conservation of primer-binding sequences, its species discrimination ability is comparatively lower and insufficient, particularly in some groups (Pochon et al. 2013, 2015, Zhan et al. 2013, Brown et al. 2016, Borrell et al.

2017). For example, no differentiation has been found among invasive Mollusca species in previous studies targeting NIS with 18S (Pochon et al. 2013, Fletcher et al. 2017, Couton et al. 2019) or between species from the genus *Botrylloides* (Couton et al. 2019). The 18S marker also lacked resolution for tunicate species surveyed through eDNA metabarcoding in Indonesian islands (Huhn et al. 2020).

On the other hand, COI has been formally established as the animal barcode (Hebert et al. 2003), enabling species-level identification in large majority of marine metazoan taxa, even using short amplicon segments in metabarcoding approaches (Hebert et al. 2003, Hollatz et al. 2017, Lobo et al. 2013, 2017). The high abundance of mitochondrial DNA copies per cell also potentially increases the success of detecting rare species (Galtier et al. 2009). But particular COI primers have also been reported to overlook some groups (e.g. Copepoda, Cladocera, Nematoda, Creer et al. 2010, Tang et al. 2012, Zhan et al. 2013), amplifying a greater percentage of some species while failing to amplify others, due to the lack of conserved primer binding sites. Primers targeting the 3' end of the COI region have been the most used (Fig. 4), in particular the primer mICOIintF in combination with jgHCO2198 and also, but to a less extent, with dgHCO2198, which have been previously found to perform better across metazoan phylogenetic diversity than the mICOIintF reverse complement with LCO1490 (and its degenerate versions dgLCO1490 and jgLCO1490) (Leray et al. 2013). A primer pair targeting exactly this same region, but using instead Lobo R have been shown to produce similar or higher taxa recovery (e.g. Lobo et al. 2013, 2017, Hollatz et al. 2017, Fais et al. 2020), but it has not been tested yet in NIS-focused studies.

At least 12 studies have employed both markers (Table S1) (Borrell et al. 2017, Grey et al. 2018, Koziol et al. 2018, Stefanni et al. 2018, von Ammon et al. 2018b, Couton et al. 2019, Holman et al. 2019, Leduc et al. 2019, Rey et al. 2019, 2020, Wood et al. 2019, Huhn et al. 2020, Westfall et al. 2020); the 18S for detecting a broader range of taxa and the COI for improved special-level discrimination of metazoan species. In addition, many species have reference sequences for only one or the other of these two markers, and thus both markers can act complimentarily (Borrell et al. 2017, Grey et al. 2018, Couton et al. 2019). A few authors also opted for the use of more than one primer pair targeting the same region to circumvent primer biases and improve the breadth of taxonomic recovery (Pagenkopp Lohan et al. 2016, Lacoursière-Roussel et al. 2018, Leduc et al. 2019). Four studies had employed other markers than COI or 18S for detecting more specific taxonomic groups (rbcL: Zaiko et al. 2015b; 16S: Koziol et al. 2018, Huhn et al. 2020 and Westfall et al. 2020, ITS2; Huhn et al. 2020 and ND4: Westfall et al. 2020) (Table S1). For example, Zaiko et al. (2015b) targeted invasive diatom species by using the rbcL gene in the ballast water of a vessel that crossed the Atlantic Ocean. Without its use, diatoms, green and yellow algae, would remain undetected or highly underestimated (Zaiko et al. 2015b). Koziol et al. (2018) targeted crustacean species by employing specific primers for the 16S mitochondrial gene. This gene remains less explored in bioinvasion studies, but it can provide a good balance

between taxonomic resolution and successful amplification breadth for NIS assessments (Westfall et al. 2020), although the reference library is more incipient than other widely used markers such as COI or 18S.

The inadequate choice of primers when targeting a broad taxonomic range, can lead to the underrepresentation or even failed detections of NIS, resulting in false negatives of potential taxa of concern (Darling et al. 2018, von Ammon et al. 2018b, Wood et al. 2019). For example, most bryozoan taxa detected through morphology, in a New Zealand's marina, had several entries in NCBI for 18S and COI, but were not detected by DNA metabarcoding (von Ammon et al. 2018b). In addition, although highly abundant, low detection rates of the invasive species *Sabella spallanzanii* were found through DNA metabarcoding, although reference sequences were present for the genetic regions targeted (18S and COI) (von Ammon et al. 2019, Wood et al. 2019). These studies suggest that amplification may have possibly failed due to incompatibility with primer binding sites or primer affinity (possibly the metabarcoding primers differed from the primers used to generate the full-length reference sequences). Thus, the characterization of diverse eukaryotic NIS communities will greatly benefit with the application of multiple marker approaches (Borrell et al. 2018, Grey et al. 2018, Steffani et al. 2018, von Ammon et al. 2018b, Westfall et al. 2020), multiple primer pairs targeting the same genetic region (Pagenkopp Lohan et al. 2016, Lacoursière-Roussel et al. 2018) or by designing more NIS-specific primers (Ardura et al. 2017, Westfall et al. 2020), which might limit the false identifications of NIS metabarcoding-derived assignments.

The use of "universal" primers may also fail to distinguish between NIS and closely-related native congeners (Westfall et al. 2020). For instance, using 18S, Shaw et al. (2019) were able to identify 19 OTUs known to contain strains of harmful bloom-causing algae in sediments, but could not distinguish between harmful and nonharmful strains, such as the various genotypes within the *Alexandrium* genus. To circumvent this type of limitation, Westfall et al. (2020) developed a biosurveillance tool tailored for detection of NIS in the Northwest Pacific. These authors designed 8 primer sets targeting multiple gene regions (16S, 18S, COI and ND4) of 4 main groups comprising NIS: Arthropoda, Bryozoa, Mollusca and Ascidiacea. After *in silico* evaluation and refinement, the primers were successfully tested *in situ* with DNA from reference specimens and validated in the field through metabarcoding of water's eDNA and zooplankton. By combining this approach with a comprehensive sampling of different substrates, they were able to detect 12 NIS in Departure Bay, Pacific Canada, and 7 potential new records, across a wide taxonomic range. However, high intraspecific variation due to multiple sympatric lineages of NIS lead to failed amplification with species-specific primers (Cowart et al. 2018, von Ammon et al. 2019). Thus, in the future, the use of metabarcoding in biosurveillance may be greatly improved by optimizing the tool for a particular biogeographic region, using a multi-marker approach, which may accelerate its implementation in regular monitoring of NIS worldwide (Westfall et al. 2020).

3.7. Sequencing platform and sequencing depth

Three main platforms have been used in studies of NIS surveillance in coastal and marine ecosystems through DNA metabarcoding: 454 Roche (11 publications), Ion Torrent (6 publications) and Illumina MiSeq (27 publications) (Table S1), with Illumina MiSeq being the most widely used. In Ion Torrent and 454 Roche, DNA fragments are sequenced-by-synthesis after clonally amplified by emulsion PCR on the surface of microbeads (Shokralla et al. 2012, Zaiko et al. 2015a), while Illumina has also adopted a sequencing-by-synthesis approach, but utilizing fluorescently labelled reversible-terminator nucleotides on clonally amplified DNA templates immobilized on the surface of a flow cell (Shokralla et al. 2012, Zaiko et al. 2015a).

The discrepancy that can be found between sequencing platforms is one of the problems that must be circumvented for a generalized use of DNA metabarcoding data for routine monitoring of biological invasions (Ardura et al. 2015, Zaiko et al. 2015a). For instance, no consistency was found between 454 Roche and Ion Torrent platforms (Zaiko et al. 2015a). Future research should explore more how the use of different platforms may affect the detection of NIS (Deiner et al. 2018, Pagenkopp Lohan et al. 2019, Zinger et al. 2019), especially as DNA metabarcoding moves to a routine surveillance and monitoring tool (Leese et al. 2018). On the other hand, the Illumina's newest high-capacity platform NovaSeq, with a sequencing depth 700 times greater than MiSeq, may be more appropriate for the early detection of NIS and was able to detect many more taxa, even when the sequencing depth matched (Singer et al. 2019). However, the NovaSeq is ca. ten times more expensive than a MiSeq instrument, as well as the sequencing runs cost, and may be out of reach for most laboratories (Singer et al. 2019).

The sequencing depth employed in DNA metabarcoding, although not specified in most studies, can also greatly compromise the results obtained, which may vary depending on the purpose of the survey (Grey et al. 2018). The sequencing depth needed for less diverse sites or more specific primers is probably lower, but this should be evaluated beforehand by sequencing a few samples. For instance, Singer et al. (2019) found that a typical sequencing depth of 60,000 reads captured about half of the diversity detectable by Illumina MiSeq in water samples from coastal ecosystems in Canada. For the detection of rare or low abundance species, particularly if a NIS did not spread yet into the ecosystem, a higher sequencing depth may be needed to increase the chance of its early detection. Grey et al. (2018) found that a sequencing depth of 150,000 reads was needed to get ca. 80% of estimated richness, including NIS detection, for 18S and COI primers, in most water samples collected from ports from 4 different geographic regions (Australia, Canada, Singapore and USA). Whether or not this has a significant impact on a study will depend on the nature of that study (Singer et al. 2019). But, in studies that have an interest in the early detection of NIS could be significantly impacted by the platform and sequencing depth employed.

3.8. Bioinformatic pipelines

The bioinformatic workflow is a complex process, highly sensitive sequence processing and taxonomic assignment parameterization (Zhan et al. 2014, Flynn et al. 2015, Brown et al. 2015). In our review, we found that different pipelines have been employed in the bioinformatic workflows (Hatzenbuhler et al. 2017, Scott et al. 2018, von Ammon et al. 2018b) (Table S4), which further complicates comparison among studies.

Most pipelines implemented sequence clustering prior to taxonomic assignments, to account for intraspecific variation and artefactual sequences, produced during PCR and sequencing (e.g. Brown et al. 2016, Ghabooli et al. 2016, Pagenkopp Lohan et al. 2016, 2019, Deiner et al. 2018, Holman et al. 2019, Shang et al. 2019). The algorithms used for OTU clustering can lead to both under- and overestimations of biodiversity, due to inherent challenges in properly aligning and clustering of sequences of variable length (e.g. 18S rRNA gene), and the variability that may exist within and among species (Brown et al. 2015, Flynn et al. 2015). A great proportion of the studies employing metabarcoding in NIS surveillance used the UPARSE pipeline for OTU clustering (Table S4) (Brown et al. 2016, Ghabooli et al. 2016, Pagenkopp Lohan et al. 2016, 2019, Deiner et al. 2018, Holman et al. 2019, Shang et al. 2019), which is allegedly among the most reliable methods available for this analysis (Edgar 2013). It is estimated that with UPARSE, OTUs are produced with ≤1% incorrect bases *versus* >3% generated by other methods (e.g. Mothur, QIIME), which tend to overestimate OTU number.

The similarity thresholds used to cluster the sequences can also strongly influence the final output (von Ammon et al. 2018b). Although a 97% threshold has been commonly employed in most biodiversity assessments, and was indeed used in most of the publications analysed in the current review (Table S4) (Zaiko et al. 2016, Pagenkopp Lohan et al. 2017, Pochon et al. 2017), a higher threshold (e.g. 99%) may increase rare taxa detection, which may be better suited to target NIS at early incursion stages (Abad et al. 2016, von Ammon et al. 2018b). Although further additional effort would be required in the analysis, it may enhance the detection of potential important sequences during early incursion stages, increasing the chance of a successful management intervention (von Ammon et al. 2018b). In fact, sequence processing can have a significant effect on species detection through DNA metabarcoding (Holman et al. 2019). For instance, the false-negative eDNA detection of *Bugula neritina* (but detected through morphology) in water and sediments in UK marinas was found to be a result of the parameters chosen during the bioinformatic analyses.

Recent bioinformatic algorithms, which use amplicon sequence variants or unique sequences instead of clustering sequences, are becoming preferred approaches to improve the accuracy and sensitivity of taxonomic assignments compared to methods involving pre-clustering of the sequences (Brown et al. 2016, Chain et al. 2016). Brown et al. (2016) were able to detect 6 NIS identified when reads were BLASTed against their local database, but that remained undetected

when OTUs were used. The reads matching these NIS formed OTUs with closely related species (low interspecific sequence divergence). Thus, the direct taxonomic assignment of each read without pre-clustering, may be more effective in distinguishing between closely related species and more sensitive to detect low abundance taxa, both crucial elements for detecting early stages of species invasions.

Concerning the databases used for taxonomic assignment of OTUs or sequences, to date, all remain incomplete or are often tailored for specific genetic markers or taxonomic groups (e.g. PR2 database for the 18S, BOLD database for COI) (von Ammon et al. 2018b, Rey et al. 2019, 2020). GenBank may contain reference sequences from many different genetic markers (Briski et al. 2016, Ardura 2019) and includes all domains of life, but it is more prone to errors, since it contains a high number of non-curated data entries (López-Escardó et al. 2018). For example, von Ammon et al. (2018b) were able to detect the invasive species *Sabella spallanzanii* in a New Zealand's marina by using 2 approaches: 18S with taxonomic assignment to PR2 and COI with taxonomic assignment to BOLD, but the taxonomic assignment in Genbank delivered matches only at the genus level. Couton et al. (2019) were able to detect 3 NIS (*Crepipatella dilatata*, *Watersipora subtorquata* and *Botrylloides diegensis*), when using a custom-designed database, but that were missing when the taxonomic assignments were conducted on GenBank. Until genetic databases are closer to completion, a combination of approaches is probably the best way to deal with this limitation (von Ammon et al. 2018b).

3.9. Incompleteness of reference sequences databases for NIS

Another issue that can have a tremendous impact on the success of the DNA metabarcoding approach on NIS surveillance in coastal ecosystems is the availability, taxonomic coverage and reliability of reference sequences databases used for taxonomic assignment (Zaiko et al. 2015c, 2016, Steffani et al. 2018, von Ammon et al. 2018b, Rey et al. 2019, 2020). Under-representation of some taxonomic groups in databases may lead to biased results in DNA metabarcoding biodiversity assessments and restrict the resolution and detection capacity of NIS at the taxonomic assignment step (Pochon et al. 2015, Briski et al. 2016, Chain et al. 2016, Ghabooli et al. 2016, Zaiko et al. 2016, Fletcher et al. 2017, Lacoursière-Roussel et al. 2018, von Ammon et al. 2018b, Ardura 2019).

For instance, DNA metabarcoding provided similar spatial and temporal trends for zooplankton analysed through morphology, but not for phytoplankton, surveyed in the estuary of Bilbao (Spain), probably due to the lack of representative sequences in genetic databases for the latter group (Abad et al. 2016). In addition, the absence of 5 zooplankton species, that were detected by morphology, but were missed in the DNA metabarcoding dataset in the Baltic sea (Zaiko et al. 2015c), or the substantial number of OTUs that matched unknown eukaryotes in biofilms scraped from a New Zealand's marina (ca. 25%, Pochon et al. 2015), were mostly attributed to the lack

of reference sequences in BOLD. In fact, nematodes and tunicates (Chain et al. 2016), Rotifera and Porifera (Lacoursière-Roussel et al. 2018) and bryozoans (Leduc et al. 2019), have been pointed out as taxonomic groups greatly missing reference sequences in genetic databases and, concomitantly, these groups may remain undetected when using only the metabarcoding approach.

Rey et al. (2020) found that from the total number of species included in the AquaNIS (http://www.corpi.ku.lt/databases/index.php/aquanis/) database (n=1083), which provides an information system on aquatic non-indigenous and cryptogenic species at a worldwide scale, only ca. 43% and 34% were covered by sequences in BOLD and PR2 databases, respectively. This gap-analysis highlights the need for completing genetic databases with reference NIS sequences. Definitely, expanding genetic databases with reference sequences is mandatory to make the most of the potential of the DNA metabarcoding approach in NIS surveillance in coastal ecosystems. Up to now, at least to our best knowledge, no dedicated reference sequences database exists for invasive species. Ideally, this should cover the full sweep of species in the target ecosystem, with a balanced representation of specimens across each species distribution range (native and recipient locations), to account for possible regional variability in targeted barcode genes. As long as reference databases are far from completion, taxonomic attribution of high throughput sequences may be improved by generating local reference sequences libraries (Abad et al. 2016). Abad et al. (2016) was able to increase the taxonomic assignment success from 23.7 to 50.5%, of plankton species in the estuary of Bilbao (Spain), by generating DNA barcodes for local species. Although completing the gaps in reference libraries is essential to make the most of the potential of the metabarcoding tool, a careful compilation, verification and annotation of available sequences is fundamental to assemble large curated and reliable reference libraries that provide support for rigorous species identifications (Couton et al. 2019, Weigand et al. 2019, Leite et al. 2020). To this end, taxonomic expertise remains crucial for an accurate identification of the specimens used to originate the reference sequences. For instance, sequences from *Acartia tonsa*, which is an invasive species recorded in many ecoregions of the world, form several distinct clades, some of which cluster with *A. hudsonica* (Lacoursière-Roussel et al. 2018). Viard et al. (2019) also found sequences of *Botrylloides diegensis* erroneously assigned to *B. leachii*. *Ciona intestinalis* and *C. robusta* were recently reclassified and, thus, there are many sequences on GenBank of *C. robusta* falsely attributed to *C. intestinalis* (Gissi et al. 2017, Couton et al. 2019). In addition, species complexes have been uncovered for many invasive species (e.g. *Bugula neritina*, *Watersipora* sp., Mackie et al. 2012, Fehlauer-Ale et al. 2014), and thus obtaining reference sequences for local specimens can be extremely important to ensure correct species assignments (Couton et al. 2019). These observations have major implications as introduced species can be misidentified as putative native species. Unfortunately, these database errors can be frequent and, thus, auditing existing records and building/compiling curated reference libraries

(Leite et al. 2020), with good-quality sequences, is just as or even more important than generating new reference sequences for NIS.

4. Final considerations

During our review we found several strengths and opportunities of the metabarcoding approach, highlighting its great potential to become a valuable monitoring tool of biological invasions in marine and coastal ecosystems (Fig. 5). These include:

i) increased resolution, enabling the identification of morphologically similar species and possibly cryptic species, due to a high taxonomic resolution;

ii) increased ability to detect smaller organisms, undifferentiated life stages and rare species, enabling the detection of potential NIS at early incursion stages, which are hardly noticeable through morphology;

iii) a higher throughput than morphology-based approaches, allowing the simultaneous analysis of a larger number of samples and a higher spatial and temporal resolution, which is mandatory for upscaling survey efforts for early detection of populations in biosurveillance programmes;

iv) faster responses, since less time is required to analyse the samples, and thus, more responsive to immediate regulatory and management needs;

v) possibility to use non-destructive surveys by analysing eDNA from environmental samples, which in the future may be conducted "*in situ*";

vi) short learning and training curve, compared with morphology-based identification;

vii) after optimization to better fit the target ecosystem, subsequent employment can be executed at a relatively low cost per sample, particularly when analysing a large number of samples;

viii) more prone to automation since the data generated is more amenable to audit and less vulnerable to variability in taxonomic expertise among studies and;

ix) increased taxonomic breadth allowing whole-ecosystem approaches, that document the composition and distribution of NIS across a wide taxonomic range by surveying different substrates and using different target markers (e.g. zooplankton, protists, harmful algae, sessile and benthic fauna).

DNA metabarcoding in NIS biosurveillance	
Strengths • Increased sensitivity • Species detection in all life stages • High-throughput • Fast response • Requires comparatively low training • Cost-effective for large sample numbers	**Weaknesses** • Lack of standard protocols • Lack of universal primers • Incompleteness of reference libraries • Taxonomic incongruences in reference libraries • False negatives and false positives • Sequencing platforms constant update • Inability to generate absolute abundance data • Inability to differentiate life stages
Opportunities • Higher species differentiation ability • NIS early detection • Increased taxonomic breadth • High spatial and temporal resolution in biosurveillance programs • Use of non-destructive DNA extraction procedures • Possibility of portability and *in situ* analyses • Prone for automation	**Threats** • Comparison among studies hard to acomplish (due to lack of standardization) • Incomplete or erroneous taxonomic assignments • Overselling of benefits and outputs compared to other approaches • High cost of newly developed platforms • Risk of contamination and misleading results if not properly implemented • Difficulties with legal implementation

Figure 5. SWOT analysis of DNA metabarcoding in NIS surveillance in marine and coastal ecosystems.

However, some technical issues, that we consider the main weaknesses and threats, still exist and must be well acknowledged and managed in order to correct and/or reduce the errors to facilitate a greater adoption of the technique in biosurveillance programs:

i) the workflows of the DNA metabarcoding analytical chain (i.e. sampling strategy, DNA extraction, target regions and primer sets, high-throughput sequences processing) vary considerably, which make comparisons across studies hard to accomplish, thus, standardized and harmonized approaches must be developed;

ii) the lack of the so-called “universal” primers may compromise species-level detections and in biosurveillance species-level identifications would be preferable;

iii) the incompleteness of reference libraries and the existence of taxonomic incongruences, can restrict the resolution and detection capacity of NIS at the taxonomic assignment step, leading to false negatives and/or erroneous identifications;

iv) the presence of Type I (false positives) and Type II (false negatives) errors, may mislead data interpretations, as both can trigger action or inaction when not required, and, thus, proper controls (negative controls and positive controls) must be designed and used through all the metabarcoding workflow;

v) the update of sequencing platforms and the high costs of runs associated may be out of reach for most laboratories and;

vi) the inability to generate absolute abundance data due for example (but no exclusively) to primer bias and/or preferential amplification of particular species, due to high primer compatibility, during PCR reactions.

In addition, a lot of sequence data is generated through metabarcoding to find in the end a relatively small number of sequences matching potential invaders. Difficulties in determining which species are more likely to establish to prioritize their early eradication or prevent future introductions, may arise. Usually there is a delay between the introduction of a new technology and its implementation in environmental surveys and, because of this, many would argue that it is still premature to replace the knowledge and expertise of experienced taxonomists by DNA metabarcoding in NIS surveillance, delaying its legal implementation.

Until these weaknesses are resolved, NIS metabarcoding is ought to be supported by complementary approaches, where positive detections can trigger in-depth morphological analysis (Borrell et al. 2017) or the application of more targeted molecular methods (e.g. qPCR, ddPCR) (von Ammon et al. 2018b, Wood et al. 2019). Further refinements to the current methodology and improvements in reference sequence databases will facilitate the uptake of this technique for routine NIS surveillance in marine and coastal ecosystems. Also, continued research should be conducted to understand and reduce the nature of false positives and false negatives. In addition, as methods improve and become standardized, and the cost of sequencing decreases, researchers and natural resource managers will have access to much greater resolution data at a fraction of the cost and time of current traditional monitoring surveys. This will definitely facilitate a greater adoption of DNA metabarcoding for NIS surveillance in coastal ecosystems.

Acknowledgements

This work was supported by national funds through the Portuguese Foundation for Science and Technology (FCT, I.P.) in the scope of the project PTDC/BIA-BMA/29754/2017.

References

Abad D, Albaina A, Aguirre M, Laza-Martínez A, Uriarte I, Uriarte A, Villate F, Estonba A (2016) Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. Marine Biology 163: 149. https://doi.org/10.1007/s00227-016-2920-0

Amaral-Zettler LA, Mccliment EA, Ducklow HW, Huse SM (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS ONE 4: e6372. https://doi.org/10.1371/journal.pone.0006372

Anderson L (2005) California's reaction to *Caulerpa taxifolia*: a model for invasive species rapid response. Biological Invasions 7: 1003-1016. https://doi.org/10.1007/s10530-004-3123-z

Ardura A (2019) Species-specific markers for early detection of marine invertebrate invaders through eDNA methods: Gaps and priorities in GenBank as database example. Journal for Nature Conservation 47: 51–57. https://doi.org/10.1016/j.jnc.2018.11.005

Ardura A, Zaiko A, Martinez JL, Samuiloviene A, Borrell Y, Garcia-Vasquez E (2015) Environmental DNA evidence of transfer of North Sea molluscs across tropical waters through ballast water. Journal of Molluscan Studies 81: 495–501. https://doi.org/10.1093/mollus/eyv022

Ardura A, Zaiko A, Borrell YJ, Samuiloviene A, Garcia-Vazquez E (2017) Novel tools for early detection of a global aquatic invasive, the zebra mussel *Dreissena polymorpha*. Aquatic Conservation: Marine and Freshwater Ecosystems 27: 165-176. https://doi.org/10.1002/aqc.2655

Bachelet G, Blanchet H, Cottet M, Dang C, de Montaudouin X, Queirós AM, Gouillieux B, Lavesque N (2009) A round-the-world tour almost completed: first records of the invasive mussel *Musculista senhousia* in the northeast Atlantic (southern Bay of Biscay). Marine Biodiversity Records 2. https://doi.org/10.1017/s1755267209001080

Baird DJ, Hajibabaei M (2012) Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. Molecular Ecology 21: 2039-2044. https://doi.org/10.1111/j.1365-294x.2012.05519.x

Bik HM, Sung W, De Ley P, Baldwin JG, Sharma J, Rocha-Olivares A, Thomas WK (2012) Metagenetic community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and shallow water sediments. Molecular Ecology 21: 1048-59. https://doi.org/10.1111/j.1365-294X.2011.05297.x.

Borrell YJ, Miralles L, Do Huu H, Mohammed-Geba K, Garcia-Vasquez E (2017) DNA in a bottle—Rapid metabarcoding survey for early alerts of invasive species in ports. PLoS ONE 12: e0183347. https://doi.org/10.1371/journal.pone.0183347

Borrell YJ, Miralles L, Mártinez-Marqués A, Semeraro A, Arias A, Carleos CE, García-Vázques E (2018) Metabarcoding and post-sampling strategies to discover non-indigenous species: a case study in the estuaries of the central south Bay of Biscay. Journal for Nature Conservation 42: 67-74. https://doi.org/10.1016/j.jnc.2017.07.002

Briand J-F, Pochon X, Wood SA, Bressy C, Garnier C, Réhel K, Urvois F, Culioli G, Zaiko A (2018) Metabarcoding and metabolomics offer complementarity in deciphering marine eukaryotic biofouling community shifts. Biofouling 34: 657-672. https://doi.org/10.1080/08927014.2018.1480757

Briski E, Ghabooli S, Bailey SA, MacIsaac HJ (2016) Are genetic databases sufficiently populated to detect non-indigenous species? Biological Invasions 18: 1911-1922. https://doi.org/10.1007/s10530-016-1134-1

Brown EA, Chain FJJ, Crease TJ, MacIsaac HJ, Cristescu ME (2015) Divergence thresholds and divergent biodiversity estimates: can metabarcoding reliably describe zooplankton communities? Ecology and Evolution 5: 2234-2251. https://doi.org/10.1002/ece3.1485

Brown EA, Chain FJJ, Zhan A, MacIsaac HJ, Cristescu ME (2016) Early detection of aquatic invaders using metabarcoding reveals a high number of non-indigenous species in Canadian ports (2016) Diversity and Distributions 22: 1045-1059. https://doi.org/10.1111/ddi.12465

Chain FJJ, Brown EA, MacIsaac HJ, Cristescu ME (2016) Metabarcoding reveals strong spatial structure and temporal turnover of zooplankton communities among marine and freshwater ports. Diversity and Distributions 22: 493-504. https://doi.org/10.1111/ddi.12427

Couton M, Comtet T, Le Cam S, Corre E, Viard F (2019) Metabarcoding on planktonic larval stages: an efficient approach for detecting and investigating life cycle dynamics of benthic aliens. Management of Biological Invasions 10: 657-689. https://doi.org/10.3391/mbi.2019.10.4.06

Cowart DA, Renshaw MA, Gantz CA, Umek J, Chandra S, Egan SP, Lodge DM, Larson ER (2018) Development and field validation of an environmental DNA (eDNA) assay for invasive clams of the genus *Corbicula*. Management of Biological Invasions 9: 27-37. https://doi.org/10.3391/mbi.2018.9.1.03

Creer C, Deiner K, Frey S, Porazinska D, Taberlet P, Thomas WK, Potter C, Bik HM (2016) The ecologist's field guide to sequence-based identification of biodiversity. Methods in Ecology and Evolution 7: 1008-1018. https://doi.org/10.1111/2041-210X.12574

Creer S, Fonseca VG, Porazinska DL, Giblin-Davis RM, Sung W, Power DM, Packer M, Carvalho GR, Blaxter ML, Lambshead PJ, Thomas WK (2010) Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises. Molecular Ecology 19: 4-20. https://doi.org/10.1111/j.1365-294X.2009.04473.x

Cristescu ME (2014) From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. Trends in Ecology and Evolution 29: 566-571. https://doi.org/10.1016/j.tree.2014.08.001

Dafforn KA, Glasby TM, Airoldi L, Rivero NK, Mayer-Pinto M, Johnston EL (2015) Marine urbanization: an ecological framework for designing multifunctional artificial structures. Frontiers in Ecology and the Environment 13: 82-90. https://doi.org/10.1890/140050

Darling JA, Blum MJ (2007) DNA-based methods for monitoring invasive species: a review and prospectus. Biological Invasions 9: 751-765. https://doi.org/10.1007/s10530-006-9079-4

Darling JA, Frederick RM (2018) Nucleic acids-based tools for ballast water surveillance, monitoring, and research. Journal of Sea Research 133: 43-52. https://doi.org/10.1016/j.seares.2017.02.005

Darling JA, Mahon AR (2011) From molecules to management: Adopting DNA-based methods for monitoring biological invasions in aquatic environments. Environmental Research 111: 978-988. https://doi.org/10.1016/j.envres.2011.02.001

Darling JA, Martinson J, Gong Y, Okum S, Pilgrim E, Pagenkopp Lohan KM, Carney KJ, Ruiz GM (2018) Ballast water exchange and invasion risk posed by intracoastal vessel traffic: an evaluation using high throughput sequencing. Environmental Science and Technology 52: 9926-9936. https://doi.org/10.1021/acs.est.8b02108

Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, Creer S, Bista I, Lodge DM, de Vere N, Pfrender ME, Bernatchez L (2017) Environmental DNA metabarcoding: transforming how we survey animal and plant communities. Molecular Ecology 26: 5872-5895. https://doi.org/10.1111/mec.14350

Deiner K, Lopez J, Bourne S, Holman LE, Seymour M, Grey EK, Lacoursière-Roussel A, Li Y, Renshaw MA, Pfrender ME, Rius M, Bernatchez L, Lodge DM (2018) Optimising the detection of marine taxonomic richness using environmental DNA metabarcoding: the effects of filter material, pore size and extraction method. Metabarcoding and Metagenomics 2: 1-15. https://doi.org/10.3897/mbmg.2.28963

Djurhuus A, Port J, Closek CJ, Yamahara KM, Romero-Maraccini O, Walz KR, Goldsmith DB, Michisaki R, Breitbart M, Boehm AB, Chavez FP (2017) Evaluation of fltration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. Frontiers in Marine Science 4: 314. https://doi.org/10.3389/fmars.2017.00314

Edgar RC (2013) UPARSE: Highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10: 996-998. https://doi.org/10.1038/nmeth.2604

Fais M, Duarte S, Vieira PE, Sousa R, Hajibabaei M, Canchaya CA, Costa FO (2020) Small-scale spatial variation of meiofaunal communities in Lima estuary (NW Portugal) assessed through metabarcoding. Estuarine, Coastal and Shelf Science. 238: 106683. https://doi.org/10.1016/j.ecss.2020.106683

Fehlauer-Ale KH, Mackie JA, Lim-Fong GE, Ale E, Pie MR, Waeschenbach A (2014) Cryptic species in the cosmopolitan Bugula neritina complex (Bryozoa, Cheilostomata). Zoologica Scripta 43: 193-205. https://doi.org/10.1111/zsc.12042

Ficetola GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental DNA from water samples. Biology Letters 4: 423-425. https://doi.org/10.1098/rsbl.2008.0118

Ficetola GF, Poulenard J, Sabatier P, Messager E, Gielly L, Leloup A, Etienne D, Bakke J, Malet E, Fanget B, Støren E, Reyss J-L, Taberlet P, Arnaud F (2018) DNA from lake sediments reveals long-term ecosystem changes after a biological invasion. Science Advances 4: eaar4292. https://doi.org/10.1126/sciadv.aar4292

Fletcher LM, Zaiko A, Atalah J, Richter I, Dufour CM, Pochon X, Wood SA, Hopkins GA (2017) Bilge water as a vector for the spread of marine pests: a morphological, metabarcoding and experimental assessment. Biological Invasions 19: 2851-2867. https://doi.org/10.1007/s10530-017-1489-y

Flynn JM, Brown EA, Chain FJJ, MacIsaac HJ, Cristescu ME (2015) Toward accurate molecular identification of species in complex environmental samples: Testing the performance of sequence filtering and clustering methods. Ecology and Evolution 5: 2252-2266. https://doi.org/10.1002/ece3.1497

Galtier N, Nabholz B, Glémin S, Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Molecular Ecology 18: 4541-4550. https://doi.org/10.1111/j.1365-294X.2009.04380.x.

Ghabooli S, Zhan A, Paolucci E, Hernandez MR, Briski E, Cristescu ME, MacIsaac HJ (2016) Population attenuation in zooplankton communities during transoceanic transfer in ballast water. Ecology and Evolution 6: 6170-6177. https://doi.org/10.1002/ece3.2349

Gissi C, Hastings KEM, Gasparini F, Stach T, Pennati R, Manni L (2017) An unprecedented taxonomic revision of a model organism: the paradigmatic case of *Ciona robusta* and *Ciona intestinalis*. Zoologica Scripta 46: 521-522. https://doi.org/10.1111/zsc.12233

Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, Spear SF, McKee A, Oyler-McCance SJ, Cornman RS, Laramie MB, Mahon AR, Lance RF, Pilliod DS, Strickler KM, Waits LP, Fremier AK, Takahara T, Herder JE, Taberlet P (2016) Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods in Ecology and Evolution 7: 1299-1307. https://doi.org/10.1111/2041-210x.12595

Grey EK, Bernatchez L, Cassey P, Deiner K, Deveney M, Howland KL, Lacoursière-Roussel A, Leong SCY, Li Y, Olds B, Pfrender ME, Prowse TAA, Renshaw MA, Lodge DM (2018) Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. Scientific Reports 8: 8843. https://doi.org/10.1038/s41598-018-27048-2

Hajibabaei M (2012) The golden age of DNA metasystematics. Trends in Genetics 28: 535-537. https://doi.org/10.1016/j.tig.2012.08.001

Hajibabaei M, Shokralla S, Zhou X, Singer GAC, Baird DJ (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS ONE 6: e17497. https://doi.org/10.1371/journal.pone.0017497

Hatzenbuhler C, Kelly JR, Martinson J, Okum S, Pilgrim E (2017) Sensitivity and accuracy of high-throughput metabarcoding methods for early detection of invasive fish species. Scientific Reports 7: 46393. https://doi.org/10.1038/srep46393

Hebert PDN, Ratnasingham S, DeWaard JR (2003) Barcoding animal life, cytochrome c oxidase subunit I divergences among closely related species. Proceedings of the Royal Society of

London, series B Biological Sciences 270: S596-S599. https://doi.org/10.1098/rsbl.2003.0025

Hermans SM, Buckley H L, Lear G (2018) Optimal extraction methods for the simultaneous analysis of DNA from diverse organisms and sample types. Molecular Ecology Resources. 18: 557-569. https://doi.org/10.1111/1755-0998.12762

Hinlo R, Gleeson D, Lintermans M, Furlan E (2017) Methods to maximise recovery of environmental DNA from water samples. PLoS ONE 12: e0179251. https://doi.org/10.1371/journal.pone.0179251

Hollatz C, Leite BR, Lobo J, Froufe H, Egas C, Costa FO (2017) Priming of a DNA metabarcoding approach for species identification and inventory in marine macrobenthic communities. Genome 60: 260-271. https://doi.org/10.1139/gen-2015-0220

Holman LE, de Bruyn M, Creer S, Carvalho G, Robidart J, Rius M (2019) Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. Scientific Reports 9: 11559. https://doi.org/10.1038/s41598-019-47899-7

Hulme PE (2009) Trade, transport and trouble: Managing invasive species pathways in an era of globalization. Journal of Applied Ecology 46: 10-18. https://doi.org/10.1111/j.1365-2664.2008.01600.x

Huhn M, Madduppa HH, Khair M, Sabrian A, Irawati Y, Anggraini NP, Wilkinson SP, Simpson T, Iwasaki K, Setiamarga DHE, Dias JP (2020) Keeping up with introduced marine species at a remote biodiversity hotspot: awareness, training and collaboration across different sectors is key. Biological Invasions 22: 749-771. https://doi.org/10.1007/s10530-019-02126-2

Jerde CL, Mahon AR, Chadderton WL, Lodge DM. “Sight-unseen” detection of rare aquatic species using environmental DNA (2011) Conservation Letters 4: 150-157. https://doi.org/10.1111/j.1755-263X.2010.00158.x

Ji Y, Ashton L, Pedley SM, Edwards DP, Tang Y, Nakamura A, Kitching R, Dolman PM, Woodcock P, Edwards FA, Larsen TH, Hsu WW, Benedick S, Hamer KC, Wilcove DS, Bruce C, Wang X, Levi T, Lott M, Emerson BC, Yu DW (2013) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. Ecology Letters 16: 1245. https://doi.org/10.1111/ele.12162

Kettunen M, Genovesi P, Gollasch S, Pagados S, Starfinger U, Brink P, Shine C (2009) Technical support to EU strategy on invasive species (IAS) – Assessment of the impacts of IAS in Europe and the EU (final module report for the European Commission). Brussels (Belgium): Institute for European Environmental Policy (IEEP), 123 pp.

Koziol A, Stat M, Simpson T, Jarman S, DiBattista JD, Harvey ES, Marnane M, McDonald J, Bunce M (2018) Environmental DNA metabarcoding studies are critically affected by substrate selection. Molecular Ecology Resources 19: 366-376. https://doi.org/10.1111/1755-0998.12971

Lacoursière-Roussel A, Howland K, Normandeau E, Grey EK, Archambault P, Deiner K, Lodge DM, Hernandez C, Leduc N, 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

Lanzén A, Lekang K, Jonassen I, Thompson EM, Troedsson C (2017) DNA extraction replicates improve diversity and compositional dissimilarity in metabarcoding of eukaryotes in marine sediments. PLoS ONE 12: e0179443. https://doi.org/10.1371/journal.pone.0179443

Leduc N, Lacoursière-Roussel A, Howland KL, Archambault P, Sevellec M, Normandeau E, Dispas A, Winkler G, McKindsey CW, Simard N, Bernatchez L (2019) Comparing eDNA metabarcoding and species collection for documenting Arctic metazoan biodiversity. Environmental DNA 4: 342-358. https://doi.org/10.1002/edn3.35

Leese F, Altermatt F, Bouchez A, Ekrem T, Hering D, Meissner K, Mergen P, Pawlowski J, Piggott J, Rimet F, Steinke D, Taberlet P, Weigand A, Abarenkov K, Beja P, Bervoets L, Björnsdóttir S, Boets P, Boggero A, Bones A, Borja Á, Bruce K, Bursić V, Carlsson J, Čiampor F, Čiamporová-Zatovičová Z, Coissac E, Costa F, Costache M, Creer S, Csabai Z, Deiner K, DelValls Á, Drakare S, Duarte S, Eleršek T, Fazi S, Fišer C, Flot J, Fonseca V, Fontaneto D, Grabowski M, Graf W, Guðbrandsson J, Hellström M, Hershkovitz Y, Hollingsworth P, Japoshvili B, Jones J, Kahlert M, Kalamujic Stroil B, Kasapidis P, Kelly M, Kelly-Quinn M, Keskin E, Kõljalg U, Ljubešić Z, Maček I, Mächler E, Mahon A, Marečková M, Mejdandzic M, Mircheva G, Montagna M, Moritz C, Mulk V, Naumoski A, Navodaru I, Padisák J, Pálsson S, Panksep K, Penev L, Petrusek A, Pfannkuchen M, Primmer C, Rinkevich B, Rotter A, Schmidt-Kloiber A, Segurado P, Speksnijder A, Stoev P, Strand M, Šulčius S, Sundberg P, Traugott M, Tsigenopoulos C, Turon X, Valentini A, van der Hoorn B, Várbíró G, Vasquez Hadjilyra M, Viguri J, Vitonytė I, Vogler A, Vrålstad T, Wägele W, Wenne R, Winding A, Woodward G, Zegura B, Zimmermann J (2016) DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. Research Ideas and Outcomes 2: e11321. https://doi.org/10.3897/rio.2.e11321

Leese F, Bouchez A, Abarenkov K, Altermatt F, Borja Á, Bruce K, Ekrem T, Čiampor F, Čiamporová-Zaťovičová Z, Costa FO, Duarte S, Elbrecht V, Fontaneto D, Franc A, Geiger MF, Hering D, Kahlert M, Kalamujić Stroil B, Kelly M, Keskin E, Liska I, Mergen P, Meissner K, Pawlowski J, Penev L, Reyjol Y, Rotter A, Steinke D, van der Wal B, Vitecek S, Zimmermann J, Weigand AM (2018) Chapter Two – Why we need sustainable networks bridging countries, disciplines, cultures and generations for aquatic biomonitoring 2.0: a perspective derived from the DNAqua-Net COST action, in: Bohan DA, Dumbrell AJ, Woodward G, Jackson M (Eds.), Advances in Ecological Research, Next Generation Biomonitoring: Part 1. Academic Press, pp. 63–99. https://doi.org/10.1016/bs.aecr.2018.01.001

Leite B, Vieira PE, Teixeira MAL, Lobo-Arteaga J, Hollatz C, Borges LMS, Duarte S, Troncoso JS, Costa FO (2020) Gap-analysis and annotated reference library for supporting macroinvertebrate metabarcoding in Atlantic Iberia. Regional Studies in Marine Science 36: 101307. https://doi.org/10.1016/j.rsma.2020.101307

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (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

Li J, Handley L-JL, Read DS, Hänfling B (2018) The effect of filtration method on the efficiency of environmental DNA capture and quantification via metabarcoding. Molecular Ecology Resources 18: 1102-1114. https://doi.org/10.1111/1755-0998.12899

Lin Y, Zhan A, Hernandez MR, Paolucci E, MacIsaac HJ, Briski E (2020) Can chlorination of ballast water reduce biological invasions? Journal of Applied Ecology 57: 331-343. https://doi.org/10.1111/1365-2664.13528

Lobo J, Costa PM, Teixeira MAL, Ferreira MS, Costa MH, Costa FO (2013) Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. BMC Ecology 13: 34. https://doi.org/10.1186/1472-6785-13-34

Lobo J, Shokralla S, Costa MH, Hajibabaei M, Costa FO (2017) DNA metabarcoding for highthroughput monitoring of estuarine macrobenthic communities. Scientific Reports 7: 15618. https://doi.org/10.1038/s41598-017-15823-6

López-Escardó D, Paps J, de Vargas C, Massana R, Ruiz-Trillo I, del Campo J (2018) Metabarcoding analysis on European coastal samples reveals new molecular metazoan diversity. Scientific Reports 8: 9106. https://doi.org/10.1038/s41598-018-27509-8

Mackie JA, Darling JA, Geller JB (2012) Ecology of cryptic invasions: latitudinal segregation among *Watersipora* (Bryozoa) species. Scientific Reports 2: 871. https://doi.org/10.1038/srep00871

Majaneva M, Diserud OH, Eagle SHC, Boström E, Hajibabaei M, Ekrem T (2018) Environmental DNA filtration techniques affect recovered biodiversity. Scientific Reports 8: 4682. https://doi.org/10.1038/s41598-018-23052-8

McCarthy HP, Crowder LB (2000) An overlooked scale of global transport: Phytoplankton species richness in ships' ballast water. Biological Invasions 2: 321-322. https://doi.org/10.1023/A:1011418432256

MEA (Millennium Ecosystem Assessment) (2005) Ecosystems and Human Wellbeing: Biodiversity Synthesis. World Resources Institute, Washington, DC. https://www.millenniumassessment.org/documents/document.356.aspx.pdf

Molnar JL, Gamboa RL, Revenga C, Spalding MD (2008) Assessing the global threat of invasive species to marine biodiversity. Frontiers in Ecology and the Environment 6: 485-492. https://doi.org/10.1890/070064

Olenin S, Elliott M, Bysveen I, Culverhouse PF, Daunys D, Dubelaar GBJ, Gollasch S, Goulletquer P, Jelmert A, Kantor Y, BringsvorMezeth K, Minchin D, Occhipinti-Ambrogi A, Olenina I, Vandekerkhove J (2011) Recommendations on methods for the detection and control of biological pollution in marine coastal waters. Marine Pollution Bulletin 62: 2598-2604. https://doi.org/10.1016/j.marpolbul.2011.08.011

Olenin S, Narščius A, Gollasch S, Lehtiniemi M, Marchini A, Minchin D, Srėbalienė G (2016) New arrivals: an indicator for non-indigenous species introductions at different geographical scales. Frontiers in Marine Science 3: 208. https://doi.org/10.3389/fmars.2016.00208

Pagenkopp Lohan KM, Fleischer RC, Carney KJ, Holzer KK, Ruiz GM (2016) Amplicon-based pyrosequencing reveals high diversity of protistan parasites in ships' ballast water: implications for biogeography and infectious diseases. Microbial Ecology 71: 530-542. https://doi.org/10.1007/s00248-015-0684-6

Pagenkopp Lohan KM, Fleischer RC, Carney KJ, Holzer KK, Ruiz GM (2017) Molecular characterisation of protistan species and communities in ships' ballast water across three U.S. coasts. Diversity and Distributions 23: 680-691. https://doi.org/10.1111/ddi.12550

Pagenkopp Lohan KM, Campbell TL, Guo J, Wheelock M, DiMaria RA, Geller JB (2019) Intact vs. homogenized subsampling: testing impacts of pre-extraction processing of multi-species samples on invasive species detection. Management of Biological Invasions 10: 324-341. https://doi.org/10.3391/mbi.2019.10.2.08

Pawlowski J, Audic S, Adl S, Bass D, Belbahri L, Berney C, Bowser SS, Cepicka I, Decelle J, Dunthorn M, Fiore-Donno AM, Gile GH, Holzmann M, Jahn R, Jirků M, Keeling PJ, Kostka M, Kudryavtsev A, Lara E, Lukeš J, Mann DG, Mitchell EAD, Nitsche F, Romeralo M, Saunders GW, Simpson AGB, Smirnov AV, Spouge JL, Stern RF, Stoeck T, Zimmermann J, Schindel D, de Vargas C (2012) CBOL Protist Working Group: Barcoding Eukaryotic Richness beyond the Animal, Plant, and Fungal Kingdoms. PLoS Biology 10: e1001419. https://doi.org/10.1371/journal.pbio.1001419

Pawlowski J, Esling P, Lejzerowicz F, Cedhagen T, Wilding TA (2014) Environmental monitoring through protist NGS metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. Molecular Ecology Resources 14: 1129-1140. https://doi.org/10.1111/1755-0998.12261

Petri B, Chaganti SR, Chan P-S, Heath D (2019) Phytoplankton growth characterization in short term MPN culture assays using 18S metabarcoding and qRT-PCR. Water Research 164: 114941. https://doi.org/10.1016/j.watres.2019.114941

Pimentel D, McNair S, Janecka J, Wightman J, Simmonds C, O'Connell C, Wong E, Russel L, Zern J, Aquino A, Tsomondo T (2001) Economic and environmental threats of alien plant, animal, and microbe invasions. Agriculture, Ecosystems and Environment 84: 1-20. https://doi.org/10.1016/S0167-8809(00)00178-X

Pochon X, Bott NJ, Smith KF, Wood SA (2013) Evaluating detection limits of next-generation sequencing for the surveillance and monitoring of international marine pests. PLoS ONE 8: e73935. https://doi.org/10.1371/journal.pone.0073935

Pochon X, Zaiko A, Hopkins GA, Banks JC, Wood SA (2015) Early detection of eukaryotic communities from marine biofilm using high-throughput sequencing: an assessment of different sampling devices. Biofouling 31: 241-251. https://doi.org/10.1080/08927014.2015.1028923

Pochon X, Zaiko A, Fletcher LM, Laroche O, Wood SA (2017) Wanted dead or alive? Using metabarcoding of environmental DNA and RNA to distinguish living assemblages for biosecurity applications. PLoS ONE 12: e0187636. https://doi.org/10.1371/journal.pone.0187636

Rey A, Carney KJ, Quinones LE, Pagenkopp Lohan KM, Ruiz GM, Basurko OC, Rodríguez-Ezpeleta N (2019) Environmental DNA metabarcoding: a promising tool for ballast water monitoring. Environmental Science and Technology 53: 11849-11859. https://doi.org/10.1021/acs.est.9b01855

Rey A, Basurko OC, Rodriguez-Ezpeleta N (2020) Considerations for metabarcoding-based port biological baseline surveys aimed at marine nonindigenous species monitoring and risk assessments. Ecology and Evolution 10: 2452-2465. https://doi.org/10.1002/ece3.6071

Rilov G, Crooks J (2009) Biological Invasions in marine ecosystems: ecological, management and geographic perspectives. Springer, Heidelberg, pp. 109-116.

Ruiz GM, Fofonoff PW, Carlton JT, Wonham MJ, Hines AH (2000) Invasion of coastal marine communities in North America: apparent patterns, processes, and biases. Annual Review of Ecology and Systematics 31: 481-531. https://doi.org/10.1146/annurev.ecolsys.31.1.481

Sassoubre LM, Yamahara KM, Gardner LD, Block BA, Boehm AB (2016) Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. Environmental Science and Technology 50: 10456-10464. https://doi.org/10.1021/acs.est.6b03114

Scott R, Zhan A, Brown EA, Chain FJJ, Cristescu M, Gras R, MacIsaac HJ (2018) Optimization and performance testing of a sequence processing pipeline applied to detection of nonindigenous species. Evolutionary Applications 11: 891-905. https://doi.org/10.1111/eva.12604

Shang L, Hu Z, Deng Y, Liu Y, Zhai X, Chai Z, Liu X, Zhan Z, Dobbs FC, Tang YZ (2019) Metagenomic sequencing identifies highly diverse assemblages of dinoflagellate cysts in

sediments from ships' ballast tanks. Microorganisms 7: 250. https://doi.org/10.3390/microorganisms7080250

Shaw JLA, Weyrich LS, Hallegraeff G, Cooper A (2019) Retrospective eDNA assessment of potentially harmful algae in historical ship ballast tank and marine port sediments. Molecular Ecology 28: 2476-2485. https://doi.org/10.1111/mec.15055

Shokralla S, Spall JL, Gibson JF, Hajibabaei M (2012) Next-generation sequencing technologies for environmental DNA research. Molecular Ecology 21: 1794-1805. https://doi.org/10.1111/j.1365-294X.2012.05538.x

Simberloff D (2001) Eradication of island invasives: practical actions and results achieved. Trends in Ecology and Evolution 16: 273-274. https://doi.org/10.1016/S0169-5347(01)02154-1

Simberloff D, Martin J-L, Genovesi P, Maris V, Wardle D, Aronson J, Courchamp F, Galil B, Garcia-Berthou E, Pascal M, Pyšek P, Sousa R, Tabacchi E, Vilà M (2013) Impacts of biological invasions: what's what and the way forward. Trends in Ecology and Evolution 28: 58-66. https://doi.org/10.1016/j.tree.2012.07.013

Singer GAC, Fahner NA, Barnes JG, McCarthy A, Hajibabaei M (2019) Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study eDNA metabarcoding seawater. Scientific Reports 9: 5991. https://doi.org/10.1038/s41598-019-42455-9

Spens J, Evans AR, Halfmaerten D, Knudsen SW, Sengupta ME, Mak SST, Sigsgaard EE, Hellstrom M (2017) Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. Methods in Ecology and Evolution 8: 635-645. https://doi.org/10.1111/2041-210X.12683

Stefanni S, Stanković, Borne D, de Olazabal A, Juretić T, Pallavicini A, Tirelli V (2018) Multi-marker metabarcoding approach to study mesozooplankton at basin scale. Scientific Reports 8: 12085. https://doi.org/10.1038/s41598-018-30157-7

Sutherland JP, Karlson RH (1977) Development and stability of the fouling community at Beaufort, North Carolina. Ecological Monographs 47: 425-446. https://doi.org/10.2307/1942176

Suarez-Menendez M, Planes S, Garcia-Vazquez E, Ardura A (2020) Early alert of biological risk in a coastal lagoon through eDNA metabarcoding. Frontiers in Ecology and Evolution 8: 9. https://doi.org/10.3389/fevo.2020.00009

Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012) Environmental DNA. Molecular Ecology 21: 1789-1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x

Tang CQ, Leasi F, Obertegger U, Kieneke A, Barraclough TG, Fontaneto D (2012) The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity

surveys of the meiofauna. Proceedings of the National Academy of Sciences of the United States of America. 109: 16208-16212. https://doi.org/10.1073/pnas.1209160109

Thresher R, Kuris A (2004) Options for managing invasive marine species. Biological Invasions. 6: 295-300. https://doi.org/10.1023/B:BINV.0000034598.28718.2e

Turner CR, Barnes MA, Xu CCY, Jones SE, Jerde CL, Lodge DM (2014) Particle size distribution and optimal capture of aqueous macrobial eDNA. Methods in Ecology and Evolution 5: 676-684. https://doi.org/10.1111/2041-210x.12206

Viard F, Roby C, Turon X, Bouchemousse S, Bishop J (2019) Cryptic diversity and database errors challenge non-indigenous species surveys: an illustration with Botrylloides spp. in the English Channel and Mediterranean Sea. Frontiers in Marine Science 6: 615. https://doi.org/10.3389/fmars.2019.00615

von Ammon U, Wood SA, Laroche O, Zaiko A, Tait L, Lavery S, Inglis GJ, Pochon X (2018a) The impact of artificial surfaces on marine bacterial and eukaryotic biofouling assemblages: A high-throughput sequencing analysis. Marine Environmental Research 133: 57-66. https://doi.org/10.1016/j.marenvres.2017.12.003

von Ammon U, Wood SA, Laroche O, Zaiko A, Tait L, Lavery S, Inglis GJ, Pochon X (2018b) Combining morpho-taxonomy and metabarcoding enhances the detection of non-indigenous marine pests in biofouling communities. Scientific Reports 8: 16290. https://doi.org/10.1038/s41598-018-34541-1

von Ammon U, Wood SA, Laroche O, Zaiko A, Tait L, Lavery S, Inglis GJ, Pochon X (2019) Linking environmental DNA and RNA for improved detection of the marine invasive fan worm *Sabella spallanzanii*. Frontiers in Marine Science 6: 621. https://doi.org/10.3389/fmars.2019.00621

Weigand H, Beermann AJ, Čiampor F, Costa FO, Csabai Z, Duarte S, Geiger MF, Grabowski M, Rimet F, Rulik B, Strand M, Szucsich N, Weigand AM, Willassen E, Wyler SA, Bouchez A, Borja A, Čiamporová-Zaťovičová Z, Ferreira F, Dijkstra KD, Eisendle U, Freyhof J, Gadawski P, Graf W, Haegerbaeumer A, van der Hoorn BB, Japoshvili B, Keresztes L, Keskin E, Leese F, Macher J, Mamos T, Paz G, Pešić V, Pfannkuchen DM, Pfannkuchen MA, Price BW, Rinkevich B, Teixeira MAL, Várbíró G, Ekrem T (2019) DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. Science of the Total Environment 678: 499-524. https://doi.org/10.1016/j.scitotenv.2019.04.247

Westfall KM, Therriault TW, Abbot CL (2020) A new approach to molecular biosurveillance of invasive species using DNA metabarcoding. Global Change Biology 26: 1012-1022 https://doi.org/10.1111/gcb.14886

Wright DA, Mitchelmore CL, Place A, Williams E, Orano-Dawson C (2019) Genomic and Microscopic Analysis of Ballast Water in the Great Lakes Region. Applied Sciences 9: 2441. https://doi.org/10.3390/app9122441

Wood SA, Biessy L, Latchford JL, Zaiko A, von Ammon U, Audrezet F, Cristescu ME, Pochon X (2020) Release and degradation of environmental DNA and RNA in a marine system. Science of the Total Environment 704: 135314. https://doi.org/10.1016/j.scitotenv.2019.135314

Wood SA, Pochon X, Laroche O, von Ammon U, Adamson J, Zaiko A (2019) A comparison of droplet digital polymerase chain reaction (PCR), quantitative PCR and metabarcoding for species-specific detection in environmental DNA. Molecular Ecology Resources 19: 1407-1419. https://doi.org/10.1111/1755-0998.13055

Wood SA, Smith KF, Banks JC, Tremblay LA, Rhodes L, Mountfort D, Cary SC, Pochon X (2013) Molecular genetic tools for environmental monitoring of New Zealand's aquatic habitats, past, present and the future. New Zealand Journal of Marine and Freshwater Research 47: 90-119. http://dx.doi.org/10.1080/00288330.2012.745885

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: 139. https://doi.org/10.1007/s00227-016-2911-1

Zaiko A, Martinez JL, Ardura A, Clusa L, Borrell YJ, Samuiloviene A, Roca A, Garcia-Vasquez E (2015a) Detecting nuisance species using NGST: Methodology shortcomings and possible application in ballast water monitoring. Marine Environmental Research 112: 64-72. https://doi.org/10.1016/j.marenvres.2015.07.002

Zaiko A, Martinez JL, Schmidt-Petersen J, Ribicic D, Samuiloviene A, Garcia-Vasquez E (2015b) Metabarcoding approach for the ballast water surveillance — an advantageous solution or an awkward challenge? Marine Pollution Bulletin 92: 25-34. https://doi.org/10.1016/j.marpolbul.2015.01.008

Zaiko A, Samuiloviene A, Ardura A, Garcia-Vazquez E (2015c) Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. Marine Pollution Bulletin 100: 53-59. https://doi.org/10.1016/j.marpolbul.2015.09.030

Zaiko A, Schimanski K, Pochon X, Hopkins GA, Goldstien S, Floerl O, Wood SA (2016) Metabarcoding improves detection of eukaryotes from early biofouling communities: implications for pest monitoring and pathway management. Bioufouling 32: 671-684. https://doi.org/10.1080/08927014.2016.1186165

Zhan A, Hulák M, Sylvester F, Huang X, Adebayo AA, Abbott CL, Adamowicz SJ, Heath DD, Cristescu ME, MacIsaac HJ (2013) High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. Methods in Ecology and Evolution 4: 558-565. https://doi.org/10.1111/2041-210X.12037

Zhan A, Xiong W, Song H, MacIsaac H J (2014) Influence of artifact removal on rare species recovery in natural complex communities using high-throughput sequencing. PLoS ONE 9: e96928. https://doi.org/10.1371/journal.pone.0096928

Zinger L, Bonin A, Alsos IG, Bálint M, Bik H, Boyer F, Chariton AA, Creer S, Coissac E, Deagle BE, De Barba M, Dickie IA, Dumbrell AJ, Ficetola GF, Fierer N, Fumagalli L, Gilbert MTP, Jarman S, Jumpponen A, Kauserud H, Orlando L, Pansu J, Pawlowski J, Tedersoo L, Thomsen PF, Willerslev E, Taberlet P (2019) DNA metabarcoding - Need for robust experimental designs to draw sound ecological conclusions. Molecular Ecology 28: 1857-1862. https://doi.org/10.1111/mec.15060