# Biases in bulk: DNA metabarcoding of marine communities and the methodology involved

Luna van der Loos[1] and Reindert Nijland[1,2]

[1]Wageningen Universiteit en Research
[2]Wageningen University

June 26, 2020

## Abstract

With the growing anthropogenic pressure on marine ecosystems, the need for efficient monitoring of biodiversity grows stronger. DNA metabarcoding of bulk samples is increasingly implemented in ecosystem assessments and is more cost-efficient and less time-consuming than monitoring based on morphology. However, before raw sequences are obtained from bulk samples, a profound number of methodological choices must be made. Here, we critically review the recent methods used for metabarcoding of marine bulk samples (including benthic, plankton and diet samples) and indicate how potential biases can be introduced throughout sampling, pre-processing, DNA extraction, marker and primer selection, PCR amplification and sequencing. From a total of 64 studies evaluated, our recommendations for best practices include to (a) consider DESS as a fixative instead of ethanol, (b) use the DNeasy PowerSoil kit for any samples containing traces of sediment, (c) not limit the marker selection to COI only, but preferably include multiple markers for higher taxonomic resolution, (d) avoid touchdown PCR profiles, (e) use a fixed annealing temperature for each primer pair when comparing across studies or institutes, (f) use a minimum of 3 PCR replicates and (g) include both negative and positive controls. Although the implementation of DNA metabarcoding still faces several technical complexities, we foresee wide-ranging advances in the near future, including improved bioinformatics for taxonomic assignment, sequencing of longer fragments, and the use of whole-genome information. Despite the bulk of biases involved in metabarcoding of bulk samples, it is clear that DNA metabarcoding provides a valuable tool in ecosystem assessments.

Running title: methodology of marine DNA metabarcoding

Luna M. van der Loos[1,2], Reindert Nijland[1]

[1] Marine Animal Ecology group, Wageningen University, P.O. box 338, 6700 AH Wageningen, The Netherlands

[2] Present address: Department of Biology, Phycology Research Group, Ghent University, Ghent, Belgium

Corresponding Author: Reindert Nijland, email: reindert.nijland@wur.nl

## Abstract

With the growing anthropogenic pressure on marine ecosystems, the need for efficient monitoring of biodiversity grows stronger. DNA metabarcoding of bulk samples is increasingly implemented in ecosystem assessments and is more cost-efficient and less time-consuming than monitoring based on morphology. However, before raw sequences are obtained from bulk samples, a profound number of methodological choices must be made. Here, we critically review the recent methods used for metabarcoding of marine bulk samples (including benthic, plankton and diet samples) and indicate how potential biases can be introduced throughout sampling, pre-processing, DNA extraction, marker and primer selection, PCR amplification and sequencing. From a total of 64 studies evaluated, our recommendations for best practices include to (a) consider DESS as a fixative instead of ethanol, (b) use the DNeasy PowerSoil kit for any samples containing traces of sediment, (c) not limit the marker selection to COI only, but preferably include multiple markers for higher taxonomic resolution, (d) avoid touchdown PCR profiles, (e) use a fixed annealing temperature for each primer pair when comparing across studies or institutes, (f) use a minimum of 3 PCR replicates, and (g) include both negative and positive controls. Although the implementation of DNA metabarcoding still faces several technical complexities, we foresee wide-ranging advances in the near future, including improved bioinformatics for taxonomic assignment, sequencing of longer fragments, and the use of whole-genome information. Despite the bulk of biases involved in metabarcoding of bulk samples, it is clear that DNA metabarcoding provides a valuable tool in ecosystem assessments.

Keywords: biodiversity, DNA metabarcoding, marine monitoring, technical biases

## 1. INTRODUCTION

Marine ecosystem health is inextricably linked to biodiversity (Goodwin et al., 2017; Porter & Hajibabaei, 2018b). Monitoring marine biodiversity in an accurate and cost-effective fashion is essential in order to assess ecosystem quality (Aylagas, Borja, & Rodríguez-Ezpeleta, 2014). This becomes increasingly important in the face of anthropogenic pressure and management of human activities (Leray & Knowlton, 2016). However, traditional biodiversity assessments require identifying each specimen in a sample using morphological characteristics and taxonomic keys. This labour-intensive and time-consuming work needs experienced taxonomists and is therefore an expensive method (Leray & Knowlton, 2016). DNA metabarcoding, which is the simultaneous identification of many taxa within the same sample based on high-throughput sequencing of the pooled genomic DNA, provides a time- and cost-effective alternative (Aylagas, Borja, Muxika, & Rodríguez-Ezpeleta, 2018).

Genomic DNA can be extracted from either bulk samples (e.g. organisms collected with sediment cores or plankton trawls) or from environmental samples (e.g. water or sediment samples) (Creer et al., 2016). The DNA in bulk samples, often referred to as community DNA, is isolated from the tissue of the specimens present in the sample. Environmental DNA (eDNA) on the other hand concerns DNA obtained directly from environmental samples without the presence biological source material (Deiner et al., 2017). Although metabarcoding often refers to both type of samples, they differ widely in the quantity and quality of the DNA and the applications. Bulk samples contain whole specimens and thus whole genomic DNA of relative high quality, whereas environmental samples contain very little relevant DNA which also is mostly degraded (Deiner et al., 2017). As a consequence, metabarcoding of bulk samples (DNA metabarcoding) is mainly used when samples are relatively easy to obtain, whereas metabarcoding of environmental samples (eDNA metabarcoding) is more useful when targeting organisms that are difficult to sample (e.g. fish). Samples concerning dietary analyses (i.e. using faeces and/or stomach content to study the diet of an organism) are in the grey area between bulk and environmental samples, as they do contain biological source material, but the DNA is often degraded and therefore of lower quality (Creer et al., 2016; Pompanon et al., 2012). Here, metabarcoding is especially valuable, as morphological characteristics are often hard to identify in the partially digested organisms in stomach or faeces content.

2

DNA metabarcoding of bulk samples has been used in assessing marine biodiversity for over a decade. The first study that described bacterial diversity (Sogin et al., 2006) was soon followed by studies on benthic invertebrates in sediment (see e.g. Chariton, Court, Hartley, Colloff, & Hardy, 2010; V. G. Fonseca et al., 2010), plankton communities (see e.g. Nanjappa, Audic, Romac, Kooistra, & Zingone, 2014; Zaiko, Samuilo-viene, Ardura, & Garcia-Vazquez, 2015), hard substrate samples (see e.g. Leray & Knowlton, 2015; Pearman, Anlauf, Irigoien, & Carvalho, 2016) and dietary samples (see e.g. Albaina, Aguirre, Abad, Santos, & Estonba, 2016; Yoon et al., 2017). Since then, DNA metabarcoding has been increasingly used worldwide in marine samples, as is demonstrated by projects such as the Genetic tools for Ecosystem health Assessment in the North Sea region (GEANS) and the Global ARMS Program with Autonomous Reef Monitoring Structures.

The increased use of DNA metabarcoding allows for comparisons across studies and within time-series but simultaneously calls for the need of harmonized and standardized protocols and the implementation of DNA-based methods in standard routine monitoring programs. Many methodological steps separate the moment of sampling from obtaining raw sequences and taxonomic assignments from these bulk samples. Starting with the sample method itself, many choices have to be made (e.g. choices of target species group, sampling period, sample type, and (sub)sample size). This continues throughout the process with different pre-processing options (e.g. fixation method, separating the organic from the inorganic compound, size fractioning), DNA extraction methods, DNA marker regions and primers, PCR amplification profiles, and finally sequencing (summarized in Figure 1). A wide array of methods is currently in use, resulting in an overwhelming number of methodological options to choose from (Figure 2). Deciding which method or choice is most suitable during each step is often highly complex, as the complete experimental design and sample types will have to be considered. Technical biases can be introduced with every step in this process and are likely to have a big impact on the obtained results. Harmonization and standardization of methods are therefore essential for DNA metabarcoding to become a standard for biodiversity monitoring. Although standardization of methods will not remove these biases, it will increase the degree to which the results of different studies can be compared. In addition, harmonization of methods is not limited to specifying the standard operating procedures, but also extends to what practices should be avoided. Understanding when and how technical biases can be introduced is a first step towards minimizing the impact of these technical biases and reaching harmonized protocols.

In this review, we focus on the methodology behind DNA metabarcoding of marine bulk samples, including benthic communities of sediment and hard substrate, plankton samples and dietary samples. Although metabarcoding of eDNA from environmental samples has greatly increased in popularity, bulk samples of community DNA yield better results for benthic communities as of now. The methodology of eDNA has received a lot of previous attention [see e.g. Creer et al. (2016) and Deiner et al. (2017)], but the differences in quality and quantity of bulkDNA and eDNA will likely result in different best practice protocols for bulk DNA metabarcoding. A concise review focusing on the technical biases and bottlenecks associated with DNA metabarcoding of bulk samples was hitherto lacking. The aim of this review is threefold: (1) provide an overview of currently employed methods and the potential biases involved with bulk DNA metabarcoding, based on a comprehensive literature survey, (2) offer recommendations for best practice protocols and standard procedures, and (3) critically assess the current bottlenecks and expected future improvements in relation to the implementation of bulk DNA metabarcoding in existing monitoring programs.

## 2. METHODS

In order to synthesise and do a 'stock take' of recent approaches and methods used for DNA metabarcoding of bulk samples, we conducted a systematic literature survey. We searched the database Web of Science in April 2019 using the key words "DNA metabarcoding", "marine", "benthos", "fauna", "biodiversity", "meiofauna", "macrofauna", "sediment", "diet" and "hard substrate" in various combinations of 2-3 keywords, resulting in 518 returned records. Duplicate papers were removed as well as those that were not primary research articles, not published in peer-reviewed journals and not written in English. Following initial filtering, a

total of 133 studies were considered for future reading, with data extracted only from papers that concerned (a) eukaryotic organisms; (b) organisms sampled in a marine environment, and (c) DNA metabarcoding of bulk samples. After excluding all studies that did not meet these criteria, 64 studies published between 2010 and 2019 were analysed. For each study, we extracted information about the sample type (plankton, soft sediment, hard substrate, or stomach content/faeces), the target species groups, sample pre-processing, the preservation method, the DNA extraction method, the selected markers (target DNA region and primers), the PCR thermal profile, the number of technical PCR replicates, and the sequencing platform used (see Supplementary Table 1). The results have been visualised in two types of graphs: a Sankey diagram and bar graphs. The Sankey diagram (Figure 2) provides an overview of the methods used by the 64 studies included in the literature review using flows of which the size is proportional to the relative frequencies. This diagram was created using sankeymatic (http://sankeymatic.com/build/). The bar graphs (Figure 3-6) provide more detailed information showing the number of studies using a specific method in relation to the sample type and target taxa. Note that the number of studies in the bar graphs do not always equal 64, as not all methods applied to every single study (e.g. preservation method of bulk samples did not apply to studies using an artificially assembled mock community). All bar graphs have been created using R 3.6.1 (R Core Team, 2019) with the package ggplot2 (Wickham, 2016).

The publications included in this review span nearly a decade (2010-2019). During this time, metabarcoding techniques have significantly improved and methods used ten years ago are likely to be outdated by now. To identify the most current methods used by the metabarcoding community, nine European institutes were asked to fill in a questionnaire related to the processing and molecular methods of marine samples. These institutes currently collaborate in the EU Interreg North Sea region project 'Genetic tools for Ecosystem health Assessment in the North Sea region' (GEANS). The responses (n=9) to the questionnaire provided insights supplementing the results from the literature survey, and identified the current commonly used methods, as well as a few key points for improvements and recommendations. Although all the graphs in this manuscript only present the results of the literature review, we combined the information from both the literature survey and the questionnaire in designing the recommendations given in chapter 4. Details of the questionnaire and a summary of the answers are given in Supplementary Material 2 (Figure S1-S6 and Table S1).

# 3. RESULTS AND DISCUSSION

## 3.1 Sample type and target species

DNA metabarcoding of bulk samples is used to unveil the biodiversity associated with a wide variety of habitats and sample types in the marine environment. The main sample types include soft sediment (27% of the publications included in this study), hard substrate (19%), diet (i.e. faeces or stomach contents; 24% of the studies) and plankton samples (30%). DNA metabarcoding of bulk samples is especially useful for the latter two sample types, as the organisms (and their remnants) in these samples are often hard to identify. Soft sediment is relatively easy to sample using cores or grabs (Aylagas, Mendibil, Borja, & Rodríguez-Ezpeleta, 2016). Sampling hard substrate is more complex due to the three-dimensional environment. Classic methods include scraping quadrats or using settlement plates. Since 2015, Autonomous Reef Monitoring Structures (ARMS) have been deployed for use in DNA metabarcoding (Leray & Knowlton, 2015). These structures, made off 9 stacked PVC plates, are similar to settlement plates, but the three-dimensional character of the structure also attracts motile organisms. This additionally allows for sampling of organisms that hide in cracks and crevices and are otherwise hard to count during visual surveys.

The targeted species group is another important concern in the study design, as this also affects the methods used in processing the samples, and the primer set and marker selection later on. Traditionally, benthic meiofauna and macrofauna are used as bio-indicators in morphological-based studies to assess environmental status of many soft sediment dominated ecosystems (Dauvin, Bellan, & Bellan-Santini, 2010; Moreno et

al., 2008). Hence, the focus of DNA metabarcoding fauna from soft sediment samples is predominantly on invertebrates (Figure 3). Hard substrate samples typically target eukaryotes in general, i.e. they are not restricted to invertebrates, but also include protists such as Rhodophyta. Plankton is mainly targeted in the context of early detection of non-indigenous species. Studies analysing diet composition target a range of different taxa, depending on the feeding habit of the study organism (e.g. herbivorous, piscivorous, planktivorous). DNA metabarcoding of bulk samples aiming for fish species is only applied in diet-based studies of their predatory species (Figure 3), as metabarcoding of eDNA in water samples is generally regarded more suitable for monitoring fish diversity due to the high sampling effort associated with traditional fish samples (Evans et al., 2017).

## 3.2 Sample preservation

The quality of DNA in a sample can degrade quickly, especially when working in high temperatures (Dawson, Raskoff, & Jacobs, 1998). Samples are therefore required to be either immediately preserved after collection in the field, or immediately pre-processed on-site and subsequently fixed. There are several options available for the preservation of samples, including ethanol, freezing or freeze-drying, RNAlater and DESS (salt saturated DMSO buffers containing EDTA). Our literature survey showed that 96-100% ethanol is the most widely used method (50% of analysed studies), especially for plankton and hard substrate (Figure 4). In sediment and diet samples (stomach/faeces content), freezing is also widely used. DESS is only used in 9% of all studies, predominantly with meiofauna and nematode research. The reason DESS is overall less popular than ethanol can partly be attributed to habits. Formalin has a long tradition in preservation of specimens for morphological research but was replaced by ethanol for molecular studies after it became clear that formalin was unsuitable for molecular analyses. In addition, ethanol is very accessible as it is available in every lab. Ethanol however is less adequate in preserving the morphological characters of specimens and is furthermore a restricted chemical in relation to shipping. DMSO on the other hand is not a restricted chemical (Williams, 2007) and samples conserved in DESS are thus easier transported an shipped than samples preserved with ethanol. In search for a more ideal method than formalin and ethanol, Yoder et al. (2006) showed that DESS ensured adequate preservation of both DNA and morphology with higher quality than either ethanol or formalin. Subsequently, DESS has been recommended as primary choice by a number of additional studies (e.g. G. Fonseca & Fehlauer-Ale, 2012 on nematodes; Gaither, Szabó, Crepeau, Bird, & Toonen, 2011 on corals; and Ransome et al., 2017 on hard substrate communuties). These studies showed that DESS preserved samples yielded higher quantity and quality DNA compared to samples preserved in ethanol and RNAlater (Dawson et al., 1998; Ransome et al., 2017), as well as a larger yield of PCR products (Gaither et al., 2011; Kilpatrick, 2002). This may be due to the fact that DMSO enhances the uptake of nuclease inhibiting substances EDTA and NaCl in the cells of the samples, whereas ethanol dehydrates the samples, resulting in denaturation of proteins (Dawson et al., 1998). In addition, DESS is easy to use in subsequent DNA extraction, as it can be simply discarded after centrifugation after which lysis buffer can be added directly, whereas ethanol needs to dry first. The superiority of DESS over ethanol may therefore cause a shift in the predominant preservation method in the future.

## 3.3 Sample pre-processing

DNA extractions are usually not performed on raw bulk samples, as these samples contain large amounts of sediment or other inorganic and organic compounds, the samples are not homogeneous, and the amount of recovered biomass is often too large. The samples are therefore pre-processed before DNA extraction. This generally consists of separating the organisms from the inorganic compound, size fractioning and homogenization.

For soft sediments, it is important to separate the organic compound from the inorganic compound (i.e. the sediment), as general inhibitors present in the sediment can hamper the enzymes active in PCR reactions, possibly leading to false negatives. Sample separation is based on traditional methods used in morphotaxonomy, as separating the fauna from the matrix helps to isolate the species, and can be achieved by sieving

(based on differences in particle size), decanting (based on density) and/or elutriation (based on a different response to a continuous water stream) (Uhlig, Thiel, & Gray, 1973). Isolating fauna from the sediment is generally more effective for hard-bodied fauna (e.g. Copepoda, Isopoda), because soft-bodied fauna (e.g. Platyhelminthes, Polychaeta) can be easily damaged (Rumohr, 1999). Molluscs are an exception in the hard-bodied fauna, especially when using decantation, as they are likely retained with the sediment due to their relatively high weight. This will introduce in important bias leading to the underrepresentation of molluscs in the results. Incorporating extra steps to separate mollusc shells from the remaining sediment should be able to evade this bias. In addition, the successfulness of these methods depends on the sediment type: decantation for example works well for coarse sand but is often only partially successful with muddy sediment (Aylagas, Mendibil, et al., 2016). Fonseca et al. (2010) and Fonseca et al. (2014) therefore also include a Ludox protocol to isolate meiofauna from fine silt. This method is based on the differences between the specific density of the sediment particles and the fauna present in the sediment and involves mixing the sample with a solution that approximates the specific density of the fauna, causing it to float, while the sediment sinks (Burgess, 2001). In our literature survey, most studies used sieving or a combination of sieving and decanting (63%), while 25% used a combination of sieving and a Ludox step, and 17% used both sieving, decanting and a Ludox protocol.

In order to detect whether organisms are retained after sieving and to accelerate the process of separating organisms from the sediment, staining is often used to colour the organic fraction. Fonseca and Fehlauer-Ale (2012) showed that Rose Bengal inhibits amplification during PCR and, as the effects of other stains are not well known, it is generally recommended to avoid staining when samples will be used for molecular analyses (Carugati, Corinaldesi, Dell'Anno, & Danovaro, 2015).

Even after separating the inorganic from the organic fraction, the sample composition is often heterogenic and detecting all species present would demand performing a DNA extraction on the entire sample. The amount of retrieved biomass is generally too large to make this feasible. Instead, the sample can be homogenized, and a well-mixed subsample can subsequently be used for DNA extraction. The homogenization is required to be thoroughly done, as the subsample should be representative for the whole community DNA. Depending on the volume of the total sample, Aylagas et al. (2016b) recommend using a mortar and pestle for the homogenization of small samples and the use of a mixer or blender for large amounts of biomass. Other options include the use of an Ultra-Turrax homogenizer, a smoothie maker, or bead-beating (Brannock & Halanych, 2015; Brannock, Waits, Sharma, & Halanych, 2014). The equipment should be thoroughly sterilized with bleach and soap between handling separate samples to avoid cross-contamination.

A subsample from a well-mixed bulk sample is assumed to contain the DNA of all species present. Larger animals and more abundant species however contribute more tissue to the homogenized biomass, which may cause a bias towards detecting larger or more abundant species, while smaller and rarer animals may remain undetected (Deagle, Clarke, Kitchener, Polanowski, & Davidson, 2018). This is a problem that is more pronounced in high diversity samples, as well as in samples with complex communities (composed of multiple individuals) compared to simple communities composed of a single specimen per species (Hollatz et al., 2017). Elbrecht et al. (2017) showed that DNA metabarcoding of freshwater macroinvertebrate bulk samples sorted by size recovered 30% more taxa than unsorted samples. This emphasizes the importance of size fractioning prior to DNA extraction, especially when specimens differ in size multiple orders of magnitude. Of the studies analysed in our literature survey, the hard substrate samples were almost always separated based on life style (sessile versus motile), and the motile fraction subsequently fractioned based on size (see e.g. Leray and Knowlton 2015; Al-Rshaidat et al. 2016). On the contrary, soft sediment samples mainly targeted a specific size-group (e.g. meiofauna), and therefore did not require additional size fractioning [but see Cowart et al. 2015 on significant differences between 0.5mm, 1mm and 2mm mesh sizes]. It is important to note here that while large species may contribute more to the total DNA, it is still possible for these larger animals to be under-represented in the amount of amplified DNA relative to small organisms due to primer amplification bias (discussed later in more detail).

6

## 3.4 DNA extraction

A wide variety of DNA extraction kits is available on the market, as well as non-commercial methods such as CTAB-procedures and phenol-chloroform procedures. Our results showed that the DNeasy PowerSoil kit (Qiagen) was the most used method for sediment and hard substrate samples (Figure 5). In plankton and diet samples, there was more variation in the extraction methods that were used (amongst others the DNeasy Blood & Tissue Kit and non-commercial methods; Figure 5).

In a comparison of a lab-based protocol and two commercial kits (Qiagen's DNeasy Blood & Tissue and PowerSoil kit) to extract DNA from nematode bulk samples, Dell'Anno et al. (2015) found that commercial kits were preferable, as they consistently provided higher quality DNA. The results were also more reproducible, and the processing time was reduced. Between the two commercial kits, the DNeasy Blood & Tissue kit yielded a higher concentration than the PowerSoil kit. Similar results were found by Jeunen et al. (2019), who extracted eDNA from water samples using seven different methods. The DNeasy Blood & Tissue kit resulted in the highest DNA quantity, whereas the PowerSoil kit and a silica-based protocol yielded the lowest concentrations. CTAB may also perform well in terms of yield, as showed in a comparison study by Schiebelhut et al. (2017). However, the DNA extractions in this comparison were conducted on single ethanol-preserved specimens, instead of community samples, and therefore likely did not contain inhibiting substances. Although the DNeasy PowerSoil kit results in lower amounts of DNA, it is still widely used for sediment samples, likely due to the effectivity of the kit to obtain high purity DNA and to remove the general inhibitors that sediments usually contain. Vasselon et al. (2017) for example extracted DNA from freshwater diatoms with four different commercial kits and a non-kit protocol and quantified not only the yield, but additionally estimated the number of inhibitors during the subsequent PCR. The non-commercial protocol yielded the highest concentration, but also showed significantly higher PCR inhibition. As only a small amount of DNA is needed in the PCR, quality of the extracted DNA could be more important than quantity.

## 3.5 Marker selection

Deciding which DNA marker region to target is one of the key considerations when using DNA metabarcoding (Leray & Knowlton, 2016). A good barcode is taxonomically informative, meaning that the barcode needs to be able to distinguish between species (i.e. that the DNA region should mutate at the right rate). At the same time, conserved primer binding areas, or degenerate primers, are required for the primers to be able to attach to the DNA of all the organisms in the sample (Carugati et al., 2015). Many DNA regions have been suggested and discussions have elaborated on which is the most suitable one [see e.g. Deagle et al. 2014; Andújar et al. 2018]. The mitochondrial cytochrome oxidase subunit I (COI) is considered one of the standard markers, especially for metazoans, and is used in worldwide projects such as the International Barcode of Life Project (https://ibol.org/). Traditionally a 658 base-pair fragment, the Folmer fragment, is targeted. However, the design of a shorter 313 base-pair fragment (Leray et al., 2013), providing paired-end overlap with the Illumina MiSeq sequencing platform, has resulted in a shift toward using this shorter fragment for bulk metabarcoding of marine metazoa. Aylagas et al. (2016a) showed that the Leray fragment (using the mlCOIintF × dgHCO2198 primer pair) outperformed the traditional Folmer fragment (using the dgLCO1490 × dgHCO2198 primer pair), as the former primer pair retrieved a higher number of metazoan taxa benchmarked with morphological methods and in addition returned less unassigned reads. Other mitochondrial markers include the 12S rRNA gene and the 16S rRNA gene. These are more conserved than COI and have not been as widely used, resulting in a less-complete reference database, but may nevertheless be good complementary markers (Devloo-Delva et al., 2019). The nuclear 18S marker is as commonly used as COI. Like the 12S and 16S regions, it is more conserved than COI, making it easier to develop primers for these regions, but they have less discrimination power on species level (Creer et al., 2016).

Choosing a barcode or a primer set is finding the balance between a trade-off: greater taxonomic coverage comes at the cost of taxonomic resolution. The organisms targeted in marine monitoring studies often range over different phyla, and the choice of marker therefore depends on the target taxa. Our literature survey

for example showed that when only benthic meiofauna was targeted, either the 18S V1-V2 region or the 18S V9 region was chosen, whereas for benthic macrofauna different regions of COI were used (Figure 6). For studies focusing on eukaryotes in general the COI 313bp fragment, 18S V4 or other 18S regions dominated. In samples targeting micro-eukaryotes and algae mainly different 18S markers were used and in zooplankton samples a range of different COI and 18S regions. 16S was predominantly used in fish samples, and the 12S marker was solely used in samples targeting fish. Finally, rbcL, tufA and 23S were only used in algal samples, as they are plastid markers.

Overall, only 28% of all studies in our literature survey used a COI fragment, whereas 59% targeted a hypervariable region of the nuclear 18S marker. However, studies comparing the performance of COI and 18S showed that 18S generally underestimates the number of species due to poor species resolution (Leray & Knowlton, 2016). For example, in a zooplankton study, mitochondrial COI and nuclear 18S showed similar coverage across phyla, but COI resolved up to three times more at species level (Clarke, Beard, Swadling, & Deagle, 2017). Similarly, Wangensteen et al. (2018) found higher species resolution and more reliable species assignment with COI, and obtained four times more MOTUs when COI was amplified compared to 18S. However, COI sequences also contained a higher percentage of unassigned MOTUs and were in addition unable to identify microscopic groups that were detected by 18S, such as Apicomplexa, Protalveolata and Foraminifera.

It becomes increasingly clear that a multi-marker approach, using several primer sets, rather than basing conclusions on a single marker, will result in a more reliable estimation of species richness (Alberdi, Aizpurua, Gilbert, & Bohmann, 2018). Yet, 75% of all studies analysed in our literature survey still only used a single marker, 18% used two markers and 7% tested more than two markers. In a study comparing the amplification success of five primer sets of different COI regions, Hollatz et al. (2017) found that three primer sets combined recovered up to 90% of the species in the artificially assembled communities, whereas the success from a single primer set ranged between 47-62%. It should be noted that using multiple markers also increases workload and costs. Whether the gain (i.e. better recovery of species richness) outweighs the costs (i.e. a lower number of samples with the same budget) will depend on the research question and aim of the study. A cost-effective solution is to pool the amplified products of several primer pairs prior to indexing. Zhang et al. (2018) for example used four primers pairs (three COI regions and the 18S V4 region) in a single Illumina run to amplify zooplankton species in a mock community. Using a single marker and single primer pair, a maximum of 77% of the species was recovered. Using only the three COI fragments, 62-83% of the species were detected, compared to 73-75% when only the 18S fragment was used. Species recovery increased to 89-93% if both markers were combined. This shows that for the optimal balance between taxonomic coverage and taxonomic resolution, multiple or combined markers with several primer sets need to be used.

## 3.6 PCR amplification

The two final steps in the process from sample to sequence – amplification by PCR and the sequencing itself – are both important sources of technical bias. During these processes, DNA molecules are randomly sampled causing a stochasticity that has an especially great impact on the detection of rare sequences, and thus on the reproducibility of the dataset (Leray & Knowlton, 2017). In addition, the PCR cycling parameters need to be optimal for efficient amplification, because suboptimal parameters can affect taxonomic resolution and species recovery.

The annealing temperature is one of the most critical parameters of the PCR thermal profile. Finding the right temperature is often difficult. Choosing a temperature that is too low will result in amplification of aspecific products, while setting the temperature too high will reduce the yield and will therefore result in a bias as not all the wanted products are amplified. To overcome this problem, a touchdown profile is sometimes used instead (Korbie & Mattick, 2008). This profile starts with a higher annealing temperature in the initial cycles and decreases the temperature with each cycle. However, several studies found that using a single annealing temperature actually yielded a higher number of zooplankton OTUs (Clarke et al., 2017) and a higher number of matches between morphology- and metabarcoding-based results of benthic

macrofauna (Aylagas, Borja, et al., 2016). Of the two single annealing temperatures compared – 46 and 50°C – the lower temperature performed best for COI (Aylagas, Borja, et al., 2016). Similar results were found for bacterial communities: a bias towards amplification of taxa with less primer mismatches almost exponentially increased with increasing annealing temperature from 47 to 61°C (Sipos et al., 2007). A touchdown profile can likewise create a bias, as the high annealing temperature in the initial cycles likely increases preferential amplification (Clarke et al., 2017)

Our literature survey showed that overall, 80% used a single annealing temperature, 14% a touchdown profile and 6% used two different annealing temperatures. The touchdown profile is predominantly used in combination with the Leray fragment (the 313 bp COI region), as that is the thermal profile described in the original paper (Leray et al., 2013). Single annealing temperatures used for this fragment were either 55-56°C or 45-46°C. For the COI Folmer fragment, annealing temperatures below 50°C were used. To amplify 18S fragments, a single annealing temperature was used in all but one study. 57°C was predominantly used for the 18S V9 and 18S V1-V2 regions, while 50°C was mostly used for the 18S V4 region. The optimal annealing temperature depends amongst others on the G/C content of the primers, but also on the concentration of the primers and the DNA polymerase used (Nichols et al. 2018). This is especially important when using several primer pairs in one PCR. Doi et al. (2019) for example used two primer pairs to respectively detect bony and cartilaginous fish with eDNA. They found that, when considering a range of annealing temperatures, the detection rate of bony fish was higher at temperatures above 57°C, but cartilaginous fish detection was maximised at 58-59°C and decreased when the temperature was either higher or lower. This was likely due to differences in G/C content of the primers. We are unable to cover all the physics of primer design and annealing temperatures in this review, but see e.g. Hollatz et al., (2017), Joanna R. Freeland, (2017) and Lobo, Shokralla, Costa, Hajibabaei, & Costa (2017) for a more detailed discussion.

To circumvent annealing issues with primers or when working with low amounts of DNA, the number of PCR cycles are often increased above 30 cycles. In our literature survey, the majority used 30 cycles or more (10%, 41% and 22% used 30, 35 and 40 cycles, respectively) and only 27% used less than 30 cycles. Although raising the number of PCR cycles does not affect the amount of mutations due to polymerase errors (Vierna et al., 2017), it may increase the formation of chimeras or increase amplification bias and has therefore been discouraged by (Ahn, Kim, Song, & Weon, 2012; Aylagas, Borja, et al., 2016). However, other studies show that a reduction of PCR cycles has little impact on the obtained results (Krehenwinkel et al., 2017; Sipos et al., 2007).

To account for the stochasticity of PCR amplification, it has been strongly recommended to use at least three PCR replicates in bulk metabarcoding (Bourlat, Haenel, Finnman, & Leray, 2016). Yet, other studies argue that stochastic processes during PCR should have limited effects on bulk samples as those contain high amounts of DNA (as opposed to eDNA samples, which contain traces of DNA, and is therefore more subjected to stochastic events) and that increasing sequencing depth is more likely to yield robust results (Smith & Peay, 2014). In a bulk metabarcoding study of fungi in soil for example, increasing PCR replicates up to 16 did not have a significant effect on species richness or turnover (alpha and beta diversity), whereas resequencing the same sample with increased sampling depth improved the reproducibility (Smith & Peay, 2014). A study on marine artificial mock communities likewise shows that stochastic events during sequencing contributes most to differences in presence-absence due to random sampling of rare taxa, while biases during PCR predominantly affect read numbers (Leray & Knowlton, 2017).

Of all the studies analysed during this review, 48% did not utilise PCR replicates, 23% of the studies used two replicates, 24% used three replicates and 5% more than three PCR replicates. As with many of the methodological decision that must be made during the research design, choosing the number of PCR replicates will depend on the research question or goal of the study. Increasing PCR replicates will also increase costs and laboratory time, so when finding the majority of species present in the sample will provide enough information (e.g. when doing ecosystem assessment studies), or when interested in using presence-absence datasets, increasing PCR replicates are not likely to improve the results. If the aim is to detect rare species (such as nonindigenous species in plankton samples), increasing sequencing depth is more important

than using a higher number of PCR replicates. However, when the main goal is to obtain an as complete possible picture of the biodiversity or when conclusions on changes in relative read numbers are important, using at least 3 PCR replicates could improve the obtained results and outweigh the costs. In studies using dietDNA, the target DNA is at least partly degraded and present in smaller amounts. Stochastic processes are therefore more likely to take place and dietDNA studies could likewise benefit from an increased number of PCR replicates. Alberdi et al. (2018) for example studied the diet of bats using faeces samples and found not only substantial difference between PCR replicates, but also showed that the way the PCR replicates are processed in the bioinformatical pipeline significantly alters results. They recommend using a relaxed approach of restrictive PCR replicate filtering to find the optimal balance between additive filtering (leading to inflated diversity) and strict restrictive filtering (resulting in erroneously removed sequences).

### 3.7 Sequencing

Several high-throughput sequencing platforms are available, of which IonTorrent, Roche 454 (no longer available) and Illumina are the most well-known. The Illumina technology currently dominates the market (>65% of the studies analyses in this review), especially after the shutdown of the Roche 454 platform in 2013. Several studies found that sequence platform (Roche 454, Illumina MiSeq, Ion Torrent PGM and Ion Torrent S5) did not affect species recovery (Braukmann et al., 2019; Eckert et al., 2018), although sequences obtained by Illumina had a higher quality (Braukmann et al., 2019). On the contrary, a recent study by Singer et al. (2019) on eDNA metabarcoding of seawater, found a much higher diversity in samples sequenced with the Illumina NovaSeq than with the MiSeq. This was partly due to sequencing depth, as the former has a 700 times greater sequencing depth than the MiSeq. The same held true though even when the NovaSeq and MiSeq were compared at similar sequencing depths. Possibly explanations for this include differences in hardware, signal processing software and flow cells (Singer et al., 2019). The authors conclude that a great portion of the biodiversity in samples could be missed while using the MiSeq at lower sequencing depths (Singer et al., 2019). Increasing the detected diversity by elevating the sequencing depth is especially important in samples with lower traces of DNA (e.g. dietDNA samples) or samples with high diversity. In addition, sequencing depth has been proposed to mitigate PCR amplification and primer bias (Alberdi et al., 2018). However, the dissimilarity between PCR replicates likewise increases with sequencing depth, indicating a higher number of sequence artefacts that can erroneously inflate the detected diversity (Alberdi et al., 2018). Thus, simply raising the sequencing depth to high levels comes at the cost of accurately characterizing the diversity and requires adjusting the bioinformatic processing to the correct level of sequencing depth.

Although widely used, the above-mentioned platforms have a limited read length, and associated to that, a limited taxonomic resolving power. The PacBio platform allows for obtaining longer reads, such as the 658bp COI Folmer fragment and beyond (Hebert et al., 2018). Very recently, nanopore sequencing using the Oxford Nanopore sequencing technology has also become an interesting alternative sequencing platform. In contrast to the widely used Illumina sequencing platforms, Nanopore sequencing is able to sequence DNA molecules of any length, and as such, it can provide a better resolution as longer fragments can be analysed (Johnson et al., 2019). This would allow for the use of a complete 18S marker, or an amplicon combining both 12S and 16S as a single marker. Obviously, for this to be efficient, primer pairs need to be developed that can cover such longer fragments. Compared to Illumina sequencing, the PacBio and Oxford Nanopore platforms both have a lower raw read accuracy. However, this lower read accuracy can be overcome by the longer fragment length, or by using bioinformatics approached that generate a consensus sequence from multiple raw reads (Nijland et al, in prep).

### 3.8 Current practices at European North Sea institutes

The results of the literature review showed that a wide variety of methods have been used and are currently in used (Figure 2). As the publication included in this review span nearly a decade (2010-2019), it is not surprising that some methods are outdated by now. An obvious example is the 454 Roche sequencing platform, which was used in 25% of the studies but is no longer available. In order to discriminate between

outdated and current approaches, nine European institutes were asked to fill in a questionnaire related to the processing and molecular methods of marine samples. These results are summarised in the following section. Detailed answers to the questionnaire are given in Supplementary Materials 2.

Similar to the results based on the literature, a wide variety of sample pre-processing methods were used. Of these methods, sieving – sometimes combined with Ludox protocol – were most used for soft sediment and plankton, and size fractioning was used for all sample types. Despite older and recent publications on the success of DESS, ethanol was still the most popular preservation method. However, several institutes were at the moment of the questionnaire performing pilot experiments comparing ethanol and DESS.

The majority of the institutes used the DNeasy Powersoil kit for DNA extraction, both for hard substrate and sediment samples. The DNeasyBlood & Tissue Kit (Qiagen), the E.Z.N.A. Mollusk DNA Kit (OmegaBio-tek) and Macherey-Nagel kits were also used for plankton.

A wide variety of DNA markers was targeted, including 12S, 16S (V3-V4 region), different lengths of COI and 18S, and larger fragments or complete mitogenomes. Overall, the COI marker was predominantly used, due to the higher taxonomic resolution and the availability of a more extended reference database. Other considerations included the availability of primers in the literature or ease of designing primers, the use of markers in previous datasets for which work was continued, and how well the specific target taxa are amplified and resolved by each marker. Only one institute combined the use of 18S and COI in order to include both phylogenetic backbone and high resolution.

PCR amplification profiles varied extensively, including the used PCR polymerase mix. Examples of used polymerases include the Promega GoTaq DNA Polymerase, KAPA HiFi HotStart ReadyMix, Phusion HiFi Polymerase and the Thermofisher PhireTissue Direct PCR MasterMix. In addition to the previously discussed technical biases introduced during PCR, the choice of polymerase may also influence the obtained results. A study on the diet of shrimps for example found that the Bioline MyTaq Red Mix yielded better results, as the Invitrogen Platinum Taq reaction was more susceptible to PCR inhibitors (van der Reis, Laroche, Jeffs, & Lavery, 2018). A study on eDNA from marine sediment further showed that polymerase choice affects both occurrence and relative abundance estimates, likely due to a preference of the polymerase for sequences with a specific GC content. Some polymerase kits are also incompatible with certain nucleosides, such as the KAPA HiFi kit with primers involving inosine (Zhang et al., 2018). Biases due to polymerase choice have received relatively little attention in the literature and are not always reported in the methods section.

Finally, five out of nine institutes used PCR replicates (mainly 3 replicates, but experimenting up to 12 replicates) and eight out of nine institutes performed sequencing on the Illumina platform. One institute fully used the Oxford Nanopore Technology, with two other institutes experimenting with this platform as well (especially considering primer free approaches).

The most important points for improvement were the use of negative controls and positive controls. Whereas the majority of the institutes used negative controls with each PCR run, this was not standardly implemented amongst all institutes. Negative controls are an essential quality control check and can be easily implemented as they do not take away sequencing depth if actually negative (i.e. there is no loss for the biological samples). Positive controls consisting of mock samples, synthesised DNA or older samples were only implemented by a minority. For positive controls, 1-2 samples will likely be sufficient. They too provide an important quality check (i.e. checking and correcting for cross-contamination between samples during amplification and library preparation, as well as detecting tag switching) (Barbato, Kovacs, Coleman, Broadhurst, & de Bruyn, 2019). Positive samples should be chosen carefully and are ideally from a different ecosystem to avoid cross-contamination between samples (Elbrecht & Steinke, 2019).

11

# 4. CONCLUSIONS AND RECOMMENDATIONS

Our results demonstrate a wide variety of methodological protocols that can be used in DNA metabarcoding of bulk samples (Figure 2). Technical biases can be introduced with almost every step and can likely not be avoided in its entirety. This calls for an urgent need to incorporate harmonized protocols. To minimize the impact of the technical biases however, and in order to obtain results comparable within and between studies, standard procedures should be followed. This is essential when implementing DNA metabarcoding in standard monitoring. We would furthermore like to stress that standardization starts with reporting laboratory protocols and bioinformatical methods clearly and accurately. In addition, making demultiplexed raw data, OTU tables and taxonomy read counts available would greatly improve the reproducibility of each study.

Although finetuning is needed for each specific sample or study design, we offer the following generic recommendations for best practice protocols and standard procedures for DNA metabarcoding of marine bulk samples:

- The sample methods naturally largely depend on the study design and account for biological variation more than technical variation. However, it is advised to specify the target organismal group before sampling, as this will influence further choices made.
- If staining is necessary for additional morphological analyses, it is recommended to split the samples before staining and use an unstained sample for DNA analyses.

- It is important to homogenize the samples thoroughly before subsampling and DNA extraction in order to obtain a subsample representing the whole community present in the larger sample. Homogenization is preferably done with a blender or smoothie-maker for larger samples and a mortar and pestle or ultra-turrax for small samples.
- DESS is the preferred preservation solution, as this results in the highest quantity and quality of DNA, is easy to use in the field even at remote locations and does not pose any restrictions on shipping.
- The Qiagen DNeasy PowerSoil kit is recommended, as previous studies showed that this kit is highly effective to obtain high purity DNA and to remove the general inhibitors. This is especially important for samples containing traces of sediments, but to minimize differences between protocols and therefore minimize biases, this extraction kit is also recommended for other types of bulk samples.
- In metabarcoding studies, at least COI (313bp fragment) should be used, and if needed 12S, 16S, 18S and/or longer fragments in addition. When the research aim is to discover as much of the biodiversity in a system as possible, using multiple markers (or deeper sequencing as an alternative) is strongly advised.
- The PCR thermal profile is one of the most bias-sensitive steps throughout the process from sample to sequence and results can only be compared across samples or studies if PCR protocols are strictly similar.
- The PCR thermal profile should preferably have a fixed annealing temperature, as touchdown protocols should be avoided.
- The number of cycles in the PCR thermal profile should preferably be fewer than 35 cycles.
- The advisable number of PCR replicates depends on the aim or research question of the study and whether the costs will outweigh the benefits. For every sample of extracted DNA, it is advisable to use at least 3 PCR replicates when:

1. The aim is to find an as complete possible picture of the biodiversity
2. The analyses will take into account changes in relative read counts
3. Working with degraded or low input target DNA (i.e. dietDNA obtained from stomach/gut content or faeces)

Also note that the way PCR replicates are processed further down the bioinformatical pipeline will have a marked impact on the results.

Using zero PCR replicates will be outweigh the benefits when:

1. Finding the majority of species present provides enough information (e.g. ecosystem assessment studies)
2. The analyses will take into account presence-absence data frames

- Include several negative controls and at least 1-2 positive controls. Including a positive control (e.g. a mock sample or synthesized DNA) can help to test for optimal bioinformatic parameters further down the pipeline, to define amplification efficiency, calibrate for PCR amplification bias and to detect tag switching. Positive samples should be chosen carefully and are ideally from a different ecosystem to avoid cross-contamination between samples.
- When increasing sequencing depth in order to detect a higher number of taxa, the number of artefactual sequences also increase. This has to be accounted for, e.g. by setting copy number thresholds, using relative copy number thresholds or by rarefying the dataset to the lowest number of sequences that a sample contained.

# 5. CHALLENGES AND PERSPECTIVES

## 5.1 Current limitations

Despite the fact that the use of DNA metabarcoding has been clearly demonstrated in terms of costs and timewise efficiency (Aylagas et al., 2018), its implementation is limited in current monitoring programmes. This is caused by the following:

### Matching to morphological based methods

The first main bottleneck is the long tradition of morphological-based methods in existing monitoring programmes and ecosystem status assessments. Apart from technical biases that may limit the implementation of DNA metabarcoding in current monitoring programmes, or future improvements that may improve implementation speed, a link with historical monitoring is important to keep in mind. Many current monitoring efforts are part of time-series and long-term databases that are fully based on traditional morphological data. Were these monitoring campaigns to suddenly switch to metabarcoding, this would almost certainly cause a shift in the observed species composition. Incorporating DNA metabarcoding is therefore only considered if the results obtained with metabarcoding can be matched to the results obtained with traditional morphotaxonomy.

Benchmarking metabarcoding with morphological assessments has been applied in several marine studies, mainly targeting soft sediment samples. A notable result is that metabarcoding in the majority of cases outperformed morphotaxonomy concerning the number of taxa retrieved (see e.g. Cowart et al. 2015; Lobo et al. 2017; Vause et al. 2019). DNA metabarcoding in seagrass sediment samples for example yielded hundreds to thousands of MOTUs (molecular operational taxonomic units) compared to only 323 species morphologically identified (Cowart et al., 2015). Likewise, Vause et al. (2019) found 62 orders in sediment samples from Antarctica with metabarcoding compared to 37 orders based on morphology. Similar results were found for plankton communities (Abad et al., 2016; Deagle et al., 2018) and hard substrate samples (Pearman et al., 2016).

Whilst often finding more taxa, metabarcoding does not necessarily cover the full taxonomic composition found with traditional morphology. In a study on zooplankton for example, 34 species were detected with metabarcoding and 21 with morphology, but the two methods only had 12 species in common (Deagle et al., 2018). This means that using only morphology, 22 species would have been missed, but using metabarcoding as sole method, still 9 species go undetected. Likewise, Lobo et al. (2017) found that metabarcoding soft-bottom macrofauna communities yielded 21-26 species that could not be detected with morphology, but using morphology in addition resulted in an extra 3-6 species.

13

Differences between results obtained by morphology or metabarcoding may partly be attributed to erroneous taxonomic assignment (with either method) or the presence of parasites, cryptic species (i.e. species that look morphologically similar but are genetically different) and other non-targeted species that are hard to detect by eye. Metabarcoding for example does not only identify larger species belonging to Cnidaria, Porifera and Rhodophyta, but also detects small eukaryotic groups such as Syndiniophyceae, Mamiellophyceae and Bacillariophyceae that would have gone undetected when using visual inspection only (Pearman et al., 2016). In addition, metabarcoding has a high resolution for problematic taxonomic groups, as well as specimens at immature stages or specimens that are physically damaged (Abad et al., 2016; Deagle et al., 2018; Lobo et al., 2017), and often has a higher specificity to detect non-indigenous species (Sundberg, Obst, Bourlat, Bergkvist, & Magnusson, 2018). This shows that metabarcoding allows for inclusion of taxonomic groups in biodiversity analyses that were formerly excluded when using classic monitoring methods.

However, morphology-based methods may currently still yield better results for the traditionally monitored taxonomic groups. Aylagas et al. (2018) for example detected higher alpha diversity with metabarcoding when taking into account all taxa (including those belonging to Fungi, Metazoa, Rhodophyta and Heterokontophyta), but when focusing on metazoans only, morphology yielded nearly twice as many benthic macrofauna taxa than metabarcoding. Similarly, Dell'Anno et al. (2015) conducted a study targeting deep-sea nematodes, but found that only 32% of the OCTUs (operational clustered taxonomic units) were assigned to Nematoda and a similar percentage of OCTUs belonged to non-targeted Fungi (29%). In this case, morphological analyses actually resulted in a higher number of nematode species. This is likely due to an incomplete or inaccurate reference sequence database, which is especially problematic for inconspicuous taxa our relatively understudied species, such as deep-sea nematodes (Dell'Anno et al., 2015; Faria et al., 2018).

### Incomplete reference databases

The existence of a less than perfect reference database (both in completeness and correctness) is a second major bottleneck. Two of the most commonly used databases are the Barcode of Life Data Systems (BOLD) (Ratnasingham & Hebert, 2007) and the NCBI GenBank (Benson et al., 2013). BOLD currently contains ~8 million sequences (accessed on 27/02/2020), is well-curated and subjected to strict quality checks, and contains a lot of metadata including photos. A large proportion of the data (~40%) is private however and not freely usable, which greatly restricts the accessibility. GenBank is considerable larger as it includes sequences of all genetic markers (not all of those usable as barcoding genes however), and currently contains >216 million sequences (accessed on 27/02/2020) but is also thought to be less reliable and more prone to contain erroneous sequences and annotations (Bidartondo, 2008; Harris, 2003). Despite these concerns, recent analyses showed that metazoan identifications in GenBank are accurate with an error rate likely below 1% at genus level and can therefore be used reliably (Leray, Knowlton, Ho, Nguyen, & Machida, 2019). The accuracy of both GenBank and BOLD do differ per taxonomic group, as GenBank for example outperformed BOLD for the identification of insect taxa, but both databases function equally well for plants and macro-fungi (Meiklejohn, Damaso, & Robertson, 2019).

The curation of reference databases is an ongoing task and debates continue over which genes should be included or should receive priority, how to assure accuracy of species identification, what metadata should be required, and which species should be prioritized. A recent gap-analysis for example showed that while freshwater vascular plants and freshwater fish have a high coverage in BOLD and GenBank, only 22% of the marine invertebrates on the European Register of Marine Species list were covered (Weigand et al., 2019). In addition, it showed that up to 50% of some marine taxonomic groups (e.g. Annelida and Nemertea) were only available in private datasets in BOLD. This raises questions on how to ensure open access data whilst simultaneously protect data ownership.

Even without any technical biases, the availability of a well-curated reference database is crucial and solving this limitation may increase the number of species that can be detected by metabarcoding. All sequences in BOLD are automatically submitted to GenBank and BOLD regularly mines sequences from Genbank, but not all records are fully synced. Improving cross-referencing has been suggested to increase the usability of

both databases (Porter & Hajibabaei, 2018a). Another solution is to combine different reference databases when assigning taxonomic identification to raw sequences (Macher, Macher, & Leese, 2017), but this is only possible when the data is freely available. We also emphasize that taxonomists are essential to obtain a well-curated reference database and to enable DNA metabarcoding to become a standard.

**Quantitative correlations**

The final main bottleneck is the lack of correlation between the proportion of reads and the original proportions of species in the sample. Establishing whether this quantitative correlation can be found has been the focus of many recent studies and in-depth meta-analyses [see e.g. (Deagle et al., 2019; Lamb et al., 2019)]. While some studies found a good correlation between the proportion of reads and species abundance (Aylagas et al., 2018), sometimes after transformation of the data (Rossel, Khodami, & Martínez Arbizu, 2019), the majority of studies did not (Hollatz et al., 2017; Leray & Knowlton, 2017) and there is still little consensus. A weak correlation was found between read abundance and biomass, but not with number of individuals (Lamb et al., 2019). This may explain why studies with low differences in size/biomass between sampled specimens do show a correlation between species abundance and read abundance (Rossel et al., 2019). Likely even more important than biomass, are the biases introduced during PCR and primer bias (Nichols et al., 2018). In the future these might be circumvented using PCR-free methods (see section in Future improvements). Other possible solutions include the use of mock communities to infer information on the quantitative signal in the concerning study (Lamb et al., 2019) and defining correction factors (Thomas, Deagle, Eveson, Harsch, & Trites, 2016).

Being able to quantitate species abundance is often a requirement in biodiversity monitoring and the inability to use metabarcoding for this purpose could increase the hesitation to incorporate DNA metabarcoding in current monitoring programmes. Although being able to obtain relative abundance from sequencing data is often regarded as the holy grail of DNA metabarcoding, it is not always necessary. The detection of non-indigenous species for example simply needs confirmation of the presence or absence of these species. Likewise, certain ecosystem indices function well without inclusion of abundance data (Aylagas et al., 2014). Currently, DNA metabarcoding still has practical limitations. Some of these are inherent to the techniques used and are unlikely to fully disappear. In contrast, we expect limitations such as the lack of reference databases and the matching to current monitoring practices to be tackled in the coming years. Below we provide an overview of the improvements that can be expected over the next years.

## 5.2 Future improvements

DNA metabarcoding techniques are still improving and hold exciting prospects in the (near) future. The main big improvements in DNA bases taxa detection and identification that we foresee are: 1) the bioinformatics involved in processing raw sequences and obtaining taxonomic assignment; 2) Improvement of reference databases and markers used; 3) removing the PCR step; 4) using whole-genome information.

**Bioinformatics for taxonomic assignment.**

This review focused on the biases that can be introduced in the methodological process of obtaining raw sequences from biological samples. After obtaining the raw sequences however, the bioinformatical approaches employed to obtain taxonomic assignment will also influence the recovered community. These bioinformatical aspects have been reviewed previously in more detail [see e.g. (Alberdi et al., 2018; Coissac, Riaz, & Puillandre, 2012; Liu, Clarke, Baker, Jordan, & Burridge, 2019)] and we therefore confine ourselves here to a brief summary. Several bioinformatical pipelines and software packages are available to process raw sequences, including MOTHUR (Schloss et al., 2009), QIIME (Caporaso et al., 2010), USEARCH (Edgar, 2010), the open-source alternative VSEARCH (Rognes, Flouri, Nichols, Quince, & Mahé, 2016), and the R-package DADA2 (Callahan et al., 2016). Post-processing of raw sequences includes amongst others read assembly, setting a minimum sequence copy number threshold to remove artefactual sequences (Leray & Knowlton, 2017), detecting and removing chimeric sequences that may have formed during PCR (Cordier

et al., 2018), sequence clustering (Pauvert et al., 2019), and choosing an approach or algorithm to assign taxonomy (Coissac et al., 2012). When the selected thresholds are too conservative, rare species are erroneously deleted, whereas overly liberal parameters will result in an inflated diversity. Alberdi et al. (2018) in a study on bat droppings showed that the highest proportion of biases are introduced through setting the copy number threshold and sequence clustering threshold, with the number of species detected differing up to a fourfold depending on the chosen threshold (e.g. 300 species were detected with copy number threshold = 1, while only 75 species were detected with copy number threshold = 100), whereas the effect of chimera removal was very limited. Although clustering sequencing into Operational Taxonomic Units (OTUs) using a fixed dissimilarity threshold was one of the most common approaches, the use of exact Amplicon Sequence Variants (ASVs) is gradually increasing. ASVs can distinguish sequence variants with single-nucleotide differences and thus resolve genetic variation on a finer-scale (Callahan, McMurdie, & Holmes, 2017; Pauvert et al., 2019). With the continuous development of new methods and bioinformatical tools, metabarcoding is increasing in resolution and specificity.

As the bioinformatical choices can highly influence the recovered biodiversity, the lack of a standardized bioinformatical pipeline thus increases the existing difficulty in comparing between studies. Setting firm thresholds in a standardized protocol however will not benefit the results and approaches will need to be customised to each dataset, taking into account amongst others the marker choice, study design and target taxa (Alberdi et al., 2018). Including a mock sample (positive sample) can not only aid to define amplification efficiency and calibrate for PCR amplification bias but can also help to test for optimal bioinformatic parameters further down the pipeline.

### Improvement of reference databases and markers used

The second big improvement in DNA metabarcoding will be the vast expansion of the reference databases. Not only will more and more species be included, we expect that complete mitochondrial genomes and full length nuclear ribosomal sequences (18S-5.8S-28S) will be added instead of only short single markers. The rapid decrease of sequencing costs and the maturation of long read sequencing technologies such as nanopore sequencing will enable this transition. Importantly, not only the reference databases will obtain more completeness and better coverage of all species, also the metabarcodes themselves could become longer, providing higher resolution. Again, present and future developments in long read sequencing technology will enable this. The obvious advantages can already be seen in the microbial field, where the switch from short regions of the prokaryotic 16S ribosomal marker to full length 16S enables the metabarcoding resolution to move from genus level to species and even strain level (Johnson et al., 2019).

### Removing the PCR step

Some technical biases are more important than other biases and amplification by PCR caused by amongst other primer biases is one of the largest sources. Direct sequencing of mitochondrial DNA and/or ribosomal genes without amplification is the obvious solution. This could be achieved through target enrichment strategies enriching the DNA pool before sequencing. The recent development of adaptive sampling approaches is highly interesting in this context. Adaptive sampling or read-until approaches allow for a short fragment of the DNA being read by the sequencer, and the DNA molecule is only sequenced completely if it is classified as being of interest for the particular experiment (Kovaka, Fan, Ni, Timp, & Schatz, 2020; Payne et al., 2020). This technique allows for implementing an enrichment step during the sequencing process. For biodiversity assessment, enrichment could be aimed at only the DNA fragments containing data required for taxonomic assignment. If these techniques mature, it will become possible to skip the amplification step and sequence the isolated DNA directly, removing a large source of bias from the DNA metabarcoding analysis. The relevance of sequencing longer fragments and full mitogenome or ribosomal gene clusters will be more pronounced for bulk sample DNA compared to dietDNA and especially eDNA, because in the latter samples most DNA fragments will have been degraded.

**Whole genome data**

Currently, sequencing technology and data processing are becoming cost effective enough to omit specific amplification or enrichment, and instead just sequence the DNA from the sample directly. Again, for this to result in an effective taxonomic assignment, the reference database will be the major bottleneck. In this context the development of shotgun metagenomics approaches based on genome skims are worth mentioning. Here, the genome of target organisms is sequenced with a very low coverage depth (~0.5x). Subsequently, long reads obtained from the samples are compared to this incomplete reference database, allowing for many matches to the partially sequenced genome (Peel et al., 2019). This version of shotgun metagenomics approach allows for the use of incomplete whole genome sequencing at much lower costs, while already taking advantage of the information present in the full genome. Eventually, the databases will mature, and complete genomes will be added for more and more species, enabling shotgun metagenomics to take over the metabarcoding strategies. Bulk metabarcoding is the perfect application for this, and, in contrast to for instance eDNA, in bulk samples there is relatively little DNA present in the samples which is not relevant for the biodiversity assessment of the sample.

## 5.3 Metabarcoding as complementary or stand-alone method?

The key question related to metabarcoding is often whether we can use metabarcoding as a stand-alone method to monitor biodiversity. To answer this question, it is natural to compare the results to those we would have obtained using traditional methods (see section on Current limitations). Although we can certainly learn a lot about the effectiveness of both metabarcoding and morphotaxonomy by making these comparisons, using morphology to "benchmark" whether metabarcoding provides a true estimate of alpha diversity may provide an incorrect contrast.

First of all, it is easy to forget that morphology – like metabarcoding – has biases as well. Classic monitoring of biodiversity for example often focuses on well-defined groups (i.e. meiofauna or macrofauna in sediment samples, which requires sampling methods that in itself are biased), while excluding species that are harder to detect. These organisms however may have an enormous impact on the ecosystem (e.g. fungi, parasites, phytoplankton and bacteria), or indeed may be key to functioning of these ecosystems and would be well worth it to include in monitoring programmes. In addition, by using well-framed but limited groups as a measure of biodiversity, classic monitoring methods actually provide an underestimation of true alpha diversity as we currently accept it (V. G. Fonseca et al., 2010; Lobo et al., 2017) and is as such not appropriate to benchmark another method. Thus, metabarcoding leads us to think about biodiversity and species richness in other terms, and on a whole different scale again.

At the same time, a one-to-one comparison of morphology and metabarcoding is problematic, as the concept of "species" can often not be literally translated to unique sequences. Several concepts of molecular species are used across metabarcoding studies, including OTUs (operational taxonomic units), MOTUs (molecular operational taxonomic units), OCTUs (operational clustered taxonomic units), and ASVs (Amplicon Sequence Variants) and one taxonomic unit does not always correspond directly to a single morphological species (Blaxter et al., 2005; Brown, Chain, Crease, Macisaac, & Cristescu, 2015). By using metabarcoding, we may have to rethink the concept "species" again. However, both morphological species and taxonomic units can be used to calculate ecosystem assessment indices, such as AMBI (AZTI Marine Biotic Index). The latter is based on the pollution tolerances of species present in soft-bottom communities (Borja, Franco, & Pérez, 2000) and has been shown to perform equally well, despite differences in alpha richness estimates, when derived from either genetic or morphological data (Aylagas, Borja, et al., 2016; Aylagas et al., 2018). Therefore, the choice whether to use metabarcoding as a stand-alone method or as a complementary method depends on the goal of the study. Either way, it is clear that metabarcoding provides a valuable tool to monitor marine life that has been increasingly used over the past years and it is likely that its role will only increase in the near future.

17

# Acknowledgements

# References

Abad, D., Albaina, A., Aguirre, M., Laza-Martínez, A., Uriarte, I., Iriarte, A., ... Estonba, A. (2016). Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. *Marine Biology* , *163* (7), 1–13. doi: 10.1007/s00227-016-2920-0

Ahn, J. H., Kim, B. Y., Song, J., & Weon, H. Y. (2012). Effects of PCR cycle number and DNA polymerase type on the 16S rRNA gene pyrosequencing analysis of bacterial communities. *Journal of Microbiology* ,*50* (6), 1071–1074. doi: 10.1007/s12275-012-2642-z

Al-Rshaidat, M. M. D., Snider, A., Rosebraugh, S., Devine, A. M., Devine, T. D., Plaisance, L., ... Leray, M. (2016). Deep COI sequencing of standardized benthic samples unveils overlooked diversity of Jordanian coral reefs in the northern Red Sea. *Genome* ,*59* (9), 724–737. doi: 10.1139/gen-2015-0208

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

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution* , *9* (1), 134–147. doi: 10.1111/2041-210X.12849

Andújar, C., Arribas, P., Yu, D. W., Vogler, A. P., & Emerson, B. C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology* , *27* (20), 3968–3975. doi: 10.1111/mec.14844

Aylagas, E., Borja, Á., Irigoien, X., & Rodríguez-Ezpeleta, N. (2016). Benchmarking DNA metabarcoding for biodiversity-based monitoring and assessment. *Frontiers in Marine Science* , *3* (JUN), 1–12. doi: 10.3389/fmars.2016.00096

Aylagas, E., Borja, Á., Muxika, I., & Rodríguez-Ezpeleta, N. (2018). Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. *Ecological Indicators* ,*95* (July), 194–202. doi: 10.1016/j.ecolind.2018.07.044

Aylagas, E., Borja, Á., & Rodríguez-Ezpeleta, N. (2014). Environmental status assessment using DNA metabarcoding: Towards a genetics based marine biotic index (gAMBI). *PLoS ONE* , *9* (3). doi: 10.1371/journal.pone.0090529

Aylagas, E., Mendibil, I., Borja, Á., & Rodríguez-Ezpeleta, N. (2016). Marine sediment sample pre-processing for macroinvertebrates metabarcoding: Mechanical enrichment and homogenization. *Frontiers in Marine Science* , *3* (OCT), 1–8. doi: 10.3389/fmars.2016.00203

Barbato, M., Kovacs, T., Coleman, M. A., Broadhurst, M. K., & de Bruyn, M. (2019). Metabarcoding for stomach-content analyses of Pygmy devil ray (Mobula kuhlii cf. eregoodootenkee): Comparing tissue and ethanol preservative-derived DNA. *Ecology and Evolution* , *9* (5), 2678–2687. doi: 10.1002/ece3.4934

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2013). GenBank. *Nucleic Acids Research* , *41* (D1), 36–42. doi: 10.1093/nar/gks1195

Bidartondo, M. I. (2008). Preserving Accuracy in GenBank.*Science* , *319* , 1616.

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., & Abebe, E. (2005). Defining operational taxonomic units using DNA barcode data. *Philosophical Transactions of the Royal Society B: Biological Sciences* , *360* (1462), 1935–1943. doi: 10.1098/rstb.2005.1725

Borja, A., Franco, J., & Pérez, V. (2000). A marine Biotic Index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. *Marine Pollution Bulletin* ,*40* (12), 1100–1114. doi: 10.1016/S0025-326X(00)00061-8

Bourlat, S., Haenel, Q., Finnman, J., & Leray, M. (2016). Preparation of Amplicon Libraries for Metabarcoding of Marine Eukaryotes Using Illumina MiSeq: The Dual-PCR Method. In *Marine Genomics: Methods and Protocols* (pp. 197–208). New York: Springer.

Brannock, P. M., & Halanych, K. M. (2015). Meiofaunal community analysis by high-throughput sequencing: Comparison of extraction, quality filtering, and clustering methods. *Marine Genomics* ,*23* (2015), 67–75. doi: 10.1016/j.margen.2015.05.007

Brannock, P. M., Waits, D. S., Sharma, J., & Halanych, K. M. (2014). High-throughput sequencing characterizes intertidal meiofaunal communities in Northern Gulf of Mexico (Dauphin Island and Mobile Bay, Alabama). *Biological Bulletin* , *227* (2), 161–174. doi: 10.1086/BBLv227n2p161

Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Elbrecht, V., Steinke, D., Ratnasingham, S., . . . Hebert, P. D. N. (2019). Metabarcoding a diverse arthropod mock community. *Molecular Ecology Resources* , *19* (3), 711–727. doi: 10.1111/1755-0998.13008

Brown, E. A., Chain, F. J. J., Crease, T. J., Macisaac, H. J., & Cristescu, M. E. (2015). Divergence thresholds and divergent biodiversity estimates: Can metabarcoding reliably describe zooplankton communities? *Ecology and Evolution* , *5* (11), 2234–2251. doi: 10.1002/ece3.1485

Burgess, R. (2001). An improved protocol for separating meiofauna from sediments using colloidal silica sols. *Marine Ecology Progress Series* , *214* (Iler 1979), 161–165. doi: 10.3354/meps214161

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal* , *11* (12), 2639–2643. doi: 10.1038/ismej.2017.119

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* , *13* (7), 581–583. doi: 10.1038/nmeth.3869

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., . . . Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* , *7* (5), 335–336. doi: 10.1038/nmeth.f.303.

Carugati, L., Corinaldesi, C., Dell'Anno, A., & Danovaro, R. (2015). Metagenetic tools for the census of marine meiofaunal biodiversity: An overview. *Marine Genomics* , *24* , 11–20. doi: 10.1016/j.margen.2015.04.010

Chariton, A. A., Court, L. N., Hartley, D. M., Colloff, M. J., & Hardy, C. M. (2010). Ecological assessment of estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. *Frontiers in Ecology and the Environment* , *8* (5), 233–238. doi: 10.1890/090115

Clarke, L. J., Beard, J. M., Swadling, K. M., & Deagle, B. E. (2017). Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. *Ecology and Evolution* , *7* (3), 873–883. doi: 10.1002/ece3.2667

Coissac, E., Riaz, T., & Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular Ecology* , *21* (8), 1834–1847. doi: 10.1111/j.1365-294X.2012.05550.x

Cordier, T., Forster, D., Dufresne, Y., Martins, C. I. M., Stoeck, T., & Pawlowski, J. (2018). Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring.*Molecular Ecology Resources* , *18* (6), 1381–1391. doi: 10.1111/1755-0998.12926

Cowart, D. A., Pinheiro, M., Mouchel, O., Maguer, M., Grall, J., Miné, J., & Arnaud-Haond, S. (2015). Metabarcoding is powerful yet still blind: A comparative analysis of morphological and molecular surveys of seagrass communities. *PLoS ONE* , *10* (2), 1–26. doi: 10.1371/journal.pone.0117562

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., . . . Bik, H. M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution* , *7* (9), 1008–1018. doi: 10.1111/2041-210X.12574

Dauvin, J. C., Bellan, G., & Bellan-Santini, D. (2010). Benthic indicators: From subjectivity to objectivity - Where is the line?*Marine Pollution Bulletin* , *60* (7), 947–953. doi: 10.1016/j.marpolbul.2010.03.028

Dawson, M. N., Raskoff, K. A., & Jacobs, D. K. (1998). Field preservation of marine invertebrate tissue for DNA analyses.*Molecular Marine Biology and Biotechnology* , *7* (2), 145–152.

Deagle, B. E., Clarke, L. J., Kitchener, J. A., Polanowski, A. M., & Davidson, A. T. (2018). Genetic monitoring of open ocean biodiversity: An evaluation of DNA metabarcoding for processing continuous plankton recorder samples. *Molecular Ecology Resources* , *18* (3), 391–406. doi: 10.1111/1755-0998.12740

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters* , *10* (9). doi: 10.1098/rsbl.2014.0562

Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., . . . Eveson, J. P. (2019). Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology* , *28* (2), 391–406. doi: 10.1111/mec.14734

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., . . . Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities.*Molecular Ecology* , *26* (21), 5872–5895. doi: 10.1111/mec.14350

Dell'Anno, A., Carugati, L., Corinaldesi, C., Riccioni, G., & Danovaro, R. (2015). Unveiling the biodiversity of deep-sea nematodes through metabarcoding: Are we ready to bypass the classical taxonomy? *PLoS ONE* , *10* (12), 1–18. doi: 10.1371/journal.pone.0144928

Devloo-Delva, F., Huerlimann, R., Chua, G., Matley, J. K., Heupel, M. R., Simpfendorfer, C. A., & Maes, G. E. (2019). How does marker choice affect your diet analysis: Comparing genetic markers and digestion levels for diet metabarcoding of tropical-reef piscivores. *Marine and Freshwater Research* , *70* (1), 8–18. doi: 10.1071/MF17209

Doi, H., Fukaya, K., Oka, S. ichiro, Sato, K., Kondoh, M., & Miya, M. (2019). Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site occupancy model. *Scientific Reports* ,*9* (1), 1–8. doi: 10.1038/s41598-019-40233-1

Eckert, I. M. K., Littlefair, J. E., Zhang, G. K., Chain, F. J. J., Crease, T. J., & Cristescu, M. E. (2018). Bioinformatics for Biomonitoring: Species Detection and Diversity Estimates Across Next-Generation Sequencing Platforms. *Advances in Ecological Research* , *59* (January 2018), 1–32. doi: 10.1016/bs.aecr.2018.06.002

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* , *26* (19), 2460–2461. doi: 10.1093/bioinformatics/btq461

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding.*Ecology and Evolution* , *7* (17), 6918–6926. doi: 10.1002/ece3.3192

Elbrecht, V., & Steinke, D. (2019). Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. *Freshwater Biology* ,*64* (2), 380–387. doi: 10.1111/fwb.13220

Evans, N. T., Li, Y., Renshaw, M. A., Olds, B. P., Deiner, K., Turner, C. R., . . . Pfrender, M. E. (2017). Fish community assessment with eDNA metabarcoding: Effects of sampling design and bioinformatic filtering. *Canadian Journal of Fisheries and Aquatic Sciences* ,*74* (9), 1362–1374. doi: 10.1139/cjfas-2016-0306

Faria, L. C. de, Di Domenico, M., Andrade, S. C. S., Santos, M. C. dos, Fonseca, G., Zanol, J., & Amaral, A. C. Z. (2018). The use of metabarcoding for meiofauna ecological patterns assessment. *Marine Environmental Research* , *140* (January), 160–168. doi: 10.1016/j.marenvres.2018.06.013

Fonseca, G., & Fehlauer-Ale, K. H. (2012). Three in one: Fixing marine nematodes for ecological, molecular, and morphological studies.*Limnology and Oceanography: Methods* , *10* (JULY), 516–523. doi: 10.4319/lom.2012.10.516

Fonseca, V. G., Carvalho, G. R., Nichols, B., Quince, C., Johnson, H. F., Neill, S. P., . . . Creer, S. (2014). Metagenetic analysis of patterns of distribution and diversity of marine meiobenthic eukaryotes.*Global Ecology and Biogeography* , *23* (11), 1293–1302. doi: 10.1111/geb.12223

Fonseca, V. G., Carvalho, G. R., Sung, W., Johnson, H. F., Power, D. M., Neill, S. P., . . . Creer, S. (2010). Second-generation environmental sequencing unmasks marine metazoan biodiversity. *Nature Communications* , *1* (7), 1–8. doi: 10.1038/ncomms1095

Gaither, M. R., Szabó, Z., Crepeau, M. W., Bird, C. E., & Toonen, R. J. (2011). Preservation of corals in salt-saturated DMSO buffer is superior to ethanol for PCR experiments. *Coral Reefs* , *30* (2), 329–333. doi: 10.1007/s00338-010-0687-1

Goodwin, K. D., Thompson, L. R., Duarte, B., Kahlke, T., Thompson, A. R., Marques, J. C., & Caçador, I. (2017). DNA sequencing as a tool to monitor marine ecological status. *Frontiers in Marine Science* ,*4* (MAY), 1–14. doi: 10.3389/fmars.2017.00107

Harris, D. J. (2003). Can you bank on GenBank? *Trends in Ecology and Evolution* , *18* (7), 317–319. doi: 10.1016/S0169-5347(03)00128-9

Hebert, P. D. N., Braukmann, T. W. A., Prosser, S. W. J., Ratnasingham, S., deWaard, J. R., Ivanova, N. V., . . . Zakharov, E. V. (2018). A Sequel to Sanger: Amplicon sequencing that scales. *BMC Genomics* ,*19* (1), 1–14. doi: 10.1186/s12864-018-4611-3

Hollatz, C., Leite, B. R., Lobo, J., Froufe, H., Egas, C., & Costa, F. O. (2017). Priming of a DNA metabarcoding approach for species identification and inventory in marine macrobenthic communities.*Genome* , *60* (3), 260–271. doi: 10.1139/gen-2015-0220

Jeunen, G. J., Knapp, M., Spencer, H. G., Taylor, H. R., Lamare, M. D., Stat, M., . . . Gemmell, N. J. (2019). Species-level biodiversity assessment using marine environmental DNA metabarcoding requires protocol optimization and standardization. *Ecology and Evolution* , Vol. 9, pp. 1323–1335. doi: 10.1002/ece3.4843

Joanna R. Freeland. (2017). The importance of molecular markers and primer design when characterizing biodiversity from environmental DNA (eDNA). *Genome* , *60* (4), 358–374. doi: https://doi.org/10.1139/gen-2016-0100

Johnson, J. S., Spakowicz, D. J., Hong, B. Y., Petersen, L. M., Demkowicz, P., Chen, L., . . . Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications* , *10* (1), 1–11. doi: 10.1038/s41467-019-13036-1

Kilpatrick, C. W. (2002). Noncryogenic preservation of mammalian tissues for DNA extraction: An assessment of storage methods. *Biochemical Genetics* , *40* (1–2), 53–62. doi: 10.1023/A:1014541222816

Korbie, D. J., & Mattick, J. S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols* , *3* (9), 1452–1456. doi: 10.1038/nprot.2008.133

Kovaka, S., Fan, Y., Ni, B., Timp, W., & Schatz, M. C. (2020). Targeted nanopore sequencing by real-time mapping of raw electrical signal with UNCALLED. *BioRxiv* . doi: 10.1101/2020.02.03.931923

Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding.*Scientific Reports* , *7* (1), 1–12. doi: 10.1038/s41598-017-17333-x

Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., & Taylor, M. I. (2019). How quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology* , *28* (2), 420–430. doi: 10.1111/mec.14920

Leray, M., & Knowlton, N. (2015). DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity.*Proceedings of the National Academy of Sciences of the United States of America* , *112* (7), 2076–2081. doi: 10.1073/pnas.1424997112

Leray, M., & Knowlton, N. (2016). Censusing marine eukaryotic diversity in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* , *371* (1702). doi: 10.1098/rstb.2015.0331

Leray, M., & Knowlton, N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding.*PeerJ* , *2017* (3), 1–27. doi: 10.7717/peerj.3006

Leray, M., Knowlton, N., Ho, S. L., Nguyen, B. N., & Machida, R. J. (2019). GenBank is a reliable resource for 21st century biodiversity research. *Proceedings of the National Academy of Sciences of the United States of America* , *116* (45), 22651–22656. doi: 10.1073/pnas.1911714116

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., . . . Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology* , *10* (1), 1–14. doi: 10.1186/1742-9994-10-34

Liu, M., Clarke, L. J., Baker, S. C., Jordan, G. J., & Burridge, C. P. (2019). A practical guide to DNA metabarcoding for entomological ecologists. *Ecological Entomology* . doi: 10.1111/een.12831

Lobo, J., Shokralla, S., Costa, M. H., Hajibabaei, M., & Costa, F. O. (2017). DNA metabarcoding for high-throughput monitoring of estuarine macrobenthic communities. *Scientific Reports* , *7* (1), 1–13. doi: 10.1038/s41598-017-15823-6

Macher, J.-N., Macher, T.-H., & Leese, F. (2017). Combining NCBI and BOLD databases for OTU assignment in metabarcoding and metagenomic datasets: The BOLD_NCBI _Merger. *Metabarcoding and Metagenomics* , *1* , e22262. doi: 10.3897/mbmg.1.22262

Meiklejohn, K. A., Damaso, N., & Robertson, J. M. (2019). Assessment of BOLD and GenBank – Their accuracy and reliability for the identification of biological materials. *PLoS ONE* , *14* (6), 1–14. doi: 10.1371/journal.pone.0217084

Moreno, M., Vezzulli, L., Marin, V., Laconi, P., Albertelli, G., & Fabiano, M. (2008). The use of meiofauna diversity as an indicator of pollution in harbours. *ICES Journal of Marine Science* ,*65* (8), 1428–1435. doi: 10.1093/icesjms/fsn116

Nanjappa, D., Audic, S., Romac, S., Kooistra, W. H. C. F., & Zingone, A. (2014). Assessment of species diversity and distribution of an ancient diatom lineage using a DNA metabarcoding approach. *PLoS ONE* , *9* (8). doi: 10.1371/journal.pone.0103810

Nichols, R. V., Vollmers, C., Newsom, L. A., Wang, Y., Heintzman, P. D., Leighton, M., . . . Shapiro, B. (2018). Minimizing polymerase biases in metabarcoding. *Molecular Ecology Resources* , *18* (5), 927–939. doi: 10.1111/1755-0998.12895

Pauvert, C., Buée, M., Laval, V., Edel-Hermann, V., Fauchery, L., Gautier, A., . . . Vacher, C. (2019). Bioinformatics matters: The accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. *Fungal Ecology* , *41* , 23–33. doi: 10.1016/j.funeco.2019.03.005

Payne, A., Holmes, N., Clarke, T., Munro, R., Debebe, B., & Loose, M. W. (2020). Nanopore adaptive sequencing for mixed samples, whole exome capture and targeted panels. *BioRxiv* , 2020.02.03.926956. doi: 10.1101/2020.02.03.926956

Pearman, J. K., Anlauf, H., Irigoien, X., & Carvalho, S. (2016). Please mind the gap - Visual census and cryptic biodiversity assessment at central Red Sea coral reefs. *Marine Environmental Research* ,*118* , 20–30. doi: 10.1016/j.marenvres.2016.04.011

Peel, N., Dicks, L. V., Clark, M. D., Heavens, D., Percival-Alwyn, L., Cooper, C., . . . Yu, D. W. (2019). Semi-quantitative characterisation of mixed pollen samples using MinION sequencing and Reverse Metagenomics (RevMet). *Methods in Ecology and Evolution* ,*10* (10), 1690–1701. doi: 10.1111/2041-210X.13265

Pompanon, F., Deagle, B. E., Symondson, W. O. C., Brown, D. S., Jarman, S. N., & Taberlet, P. (2012). Who is eating what: Diet assessment using next generation sequencing. *Molecular Ecology* , *21* (8), 1931–1950. doi: 10.1111/j.1365-294X.2011.05403.x

Porter, T. M., & Hajibabaei, M. (2018a). Over 2.5 million COI sequences in GenBank and growing. *PLoS ONE* , *13* (9), 1–16. doi: 10.1371/journal.pone.0200177

Porter, T. M., & Hajibabaei, M. (2018b). Scaling up : A guide to high-throughput genomic approaches for biodiversity analysis.*Molecular Ecology* , *27* , 313–338. doi: 10.1111/mec.14478

R Core Team. (2019). *R: A language and environment for statistical computing. R Foundation for Statistical Computing* . Retrieved from https://www.r-project.org/

Ransome, E., Geller, J. B., Timmers, M., Leray, M., Mahardini, A., Sembiring, A., . . . Meyer, C. P. (2017). The importance of standardization for biodiversity comparisons: A case study using autonomous reef monitoring structures (ARMS) and metabarcoding to measure cryptic diversity on Mo'orea coral reefs, French Polynesia.*PLoS ONE* , *12* (4), 1–19. doi: 10.1371/journal.pone.0175066

Ratnasingham, S., & Hebert, P. D. N. (2007). The Barcode of Life Data System. *Molecular Ecology Notes* , *7* (April 2016), 355–364. doi: 10.1111/j.1471-8286.2006.01678.x

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahe, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ* ,*2016* (10), 1–22. doi: 10.7717/peerj.2584

Rossel, S., Khodami, S., & Martinez Arbizu, P. (2019). Comparison of Rapid Biodiversity Assessment of Meiobenthos Using MALDI-TOF MS and Metabarcoding. *Frontiers in Marine Science* , *6* (November), 1–13. doi: 10.3389/fmars.2019.00659

Rumohr, H. (1999). ICES Techniques in Marine Environmental Sciences. Soft bottom macrofauna: Collection, treatment, and quality assurance of samples. *International Counsil for the Exploration of the Sea* , (27), 1–20.

Schiebelhut, L. M., Abboud, S. S., Gomez Daglio, L. E., Swift, H. F., & Dawson, M. N. (2017). A comparison of DNA extraction methods for high-throughput DNA analyses. *Molecular Ecology Resources* ,*17* (4), 721–729. doi: 10.1111/1755-0998.12620

23

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., ... Weber, C. F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* , *75* (23), 7537–7541. doi: 10.1128/AEM.01541-09

Singer, G. A. C., Fahner, N. A., Barnes, J. G., McCarthy, A., & Hajibabaei, M. (2019). Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. *Scientific Reports* , *9* (1), 1–12. doi: 10.1038/s41598-019-42455-9

Sipos, R., Szekely, A. J., Palatinszky, M., Revesz, S., Marialigeti, K., & Nikolausz, M. (2007). Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. *FEMS Microbiology Ecology* , *60* (2), 341–350. doi: 10.1111/j.1574-6941.2007.00283.x

Smith, D. P., & Peay, K. G. (2014). Sequence depth, not PCR replication, improves ecological inference from next generation DNA sequencing. *PLoS ONE* , *9* (2). doi: 10.1371/journal.pone.0090234

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., ... Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "'rare biosphere."' *Pnas* ,*103* (32), 12115–12120. doi: 10.4278/ajhp.28.4.tahp

Sundberg, P., Obst, M., Bourlat, S., Bergkvist, J., & Magnusson, M. (2018). *Utvardering av ny overvakning av frammande arter. Metodjamforelse mellan traditionell och DNA-baserad identifiering* . Goteborg.

Thomas, A. C., Deagle, B. E., Eveson, J. P., Harsch, C. H., & Trites, A. W. (2016). Quantitative DNA metabarcoding: Improved estimates of species proportional biomass using correction factors derived from control material. *Molecular Ecology Resources* , *16* (3), 714–726. doi: 10.1111/1755-0998.12490

Uhlig, G., Thiel, H., & Gray, J. S. (1973). The quantitative separation of meiofauna. *Helgolander Wissenschaftliche Meeresuntersuchungen* ,*25* (1), 173–195. doi: 10.1007/BF01609968

van der Reis, A. L., Laroche, O., Jeffs, A. G., & Lavery, S. D. (2018). Preliminary analysis of New Zealand scampi (Metanephrops challengeri) diet using metabarcoding. *PeerJ* , *2018* (9), 1–25. doi: 10.7717/peerj.5641

Vasselon, V., Domaizon, I., Rimet, F., Kahlert, M., & Bouchez, A. (2017). Application of high-throughput sequencing (HTS) metabarcoding to diatom biomonitoring: Do DNA extraction methods matter? *Freshwater Science* , *36* (1), 162–177. doi: 10.1086/690649

Vause, B. J., Morley, S. A., Fonseca, V. G., Jazdzewska, A., Ashton, G. V., Barnes, D. K. A., ... Peck, L. S. (2019). Spatial and temporal dynamics of Antarctic shallow soft-bottom benthic communities: Ecological drivers under climate change. *BMC Ecology* ,*19* (1), 1–14. doi: 10.1186/s12898-019-0244-x

Vierna, J., Dona, J., Vizcaino, A., Serrano, D., Jovani, R., & Chain, F. (2017). PCR cycles above routine numbers do not compromise high-throughput DNA barcoding results. *Genome* , *60* (10), 868–873. doi: 10.1139/gen-2017-0081

Wangensteen, O. S., Palacin, C., Guardiola, M., & Turon, X. (2018). DNA metabarcoding of littoral hardbottom communities: High diversity and database gaps revealed by two molecular markers. *PeerJ* ,*2018* (5), 1–30. doi: 10.7717/peerj.4705

Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., ... 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. doi: 10.1016/j.scitotenv.2019.04.247

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* . Retrieved from https://ggplot2.tidyverse.org

Williams, S. T. (2007). Safe and legal shipment of tissue samples: Does it affect DNA quality? *Journal of Molluscan Studies* ,*73* (4), 416–418. doi: 10.1093/mollus/eym039

24

Yoder, M., Tandingan De Ley, I., King, I. W., Mundo-Ocampo, M., Mann, J., Blaxter, M., . . . De Ley, P. (2006). DESS: A versatile solution for preserving morphology and extractable DNA of nematodes. *Nematology* , *8* (3), 367–376. doi: 10.1163/156854106778493448

Yoon, T. H., Kang, H. E., Lee, S. R., Lee, J. B., Baeck, G. W., Park, H., & Kim, H. W. (2017). Metabarcoding analysis of the stomach contents of the Antarctic Toothfish (Dissostichus mawsoni) collected in the Antarctic Ocean. *PeerJ* , *2017* (11), 1–22. doi: 10.7717/peerj.3977

Zaiko, A., Samuiloviene, A., Ardura, A., & Garcia-Vazquez, E. (2015). Metabarcoding approach for non-indigenous species surveillance in marine coastal waters. *Marine Pollution Bulletin* , *100* (1), 53–59. doi: 10.1016/j.marpolbul.2015.09.030

Zhang, G. K., Chain, F. J. J., Abbott, C. L., & Cristescu, M. E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. *Evolutionary Applications* , *11* (10), 1901–1914. doi: 10.1111/eva.12694

# Figures

Figure 1. Schematic overview of the process from bulk sample to raw sequences and taxonomical assignments. Technical biases can potentially be introduced during sampling, pre-processing, DNA extraction, marker selection, PCR amplification and sequencing. (Graphical design by Aline Joustra).

Figure 2. Relative frequencies of methods used during seven methodological stages in DNA metabarcoding of bulk samples from 64 studies published between 2010 and 2019, including sample type, target species group, preservation, DNA extraction, marker, PCR profile and sequencing platform. The thickness of the lines corresponds to the number of papers using this method. 'NA' represents studies where the corresponding stage/method was not reported or was not needed. The diagram was created with sankeymatic (http://sankeymatic.com/build/).

Figure 3. Taxa targeted in bulk DNA metabarcoding studies with different sample types.

Figure 4. Preservation method used in bulk metabarcoding studies with different sample types.

Figure 5. DNA extraction methods used in bulk metabarcoding studies with different sample types.

Figure 6. Different available markers used in bulk metabarcoding studies when targeting different marine groups.

# Potential introductions of technical biases

## SAMPLE TYPE

## PRE-PROCESSING

Isolate  Homogenise

### PRESERVATION

C2H5OH  DMSO

## MARKER SELECTION

Single  Multiple

Longer

## DNA EXTRACTION

Subsampling

## PCR AMPLIFICATION

Denature  Anneal

Temperature  Specificity  Yield

## SEQUENCING

Amplicon  Raw Sequence  ID

GAAGCGCGTTTA

GTTTACACTGA

GGTCGTGAAG

GAAGCGTTTAC