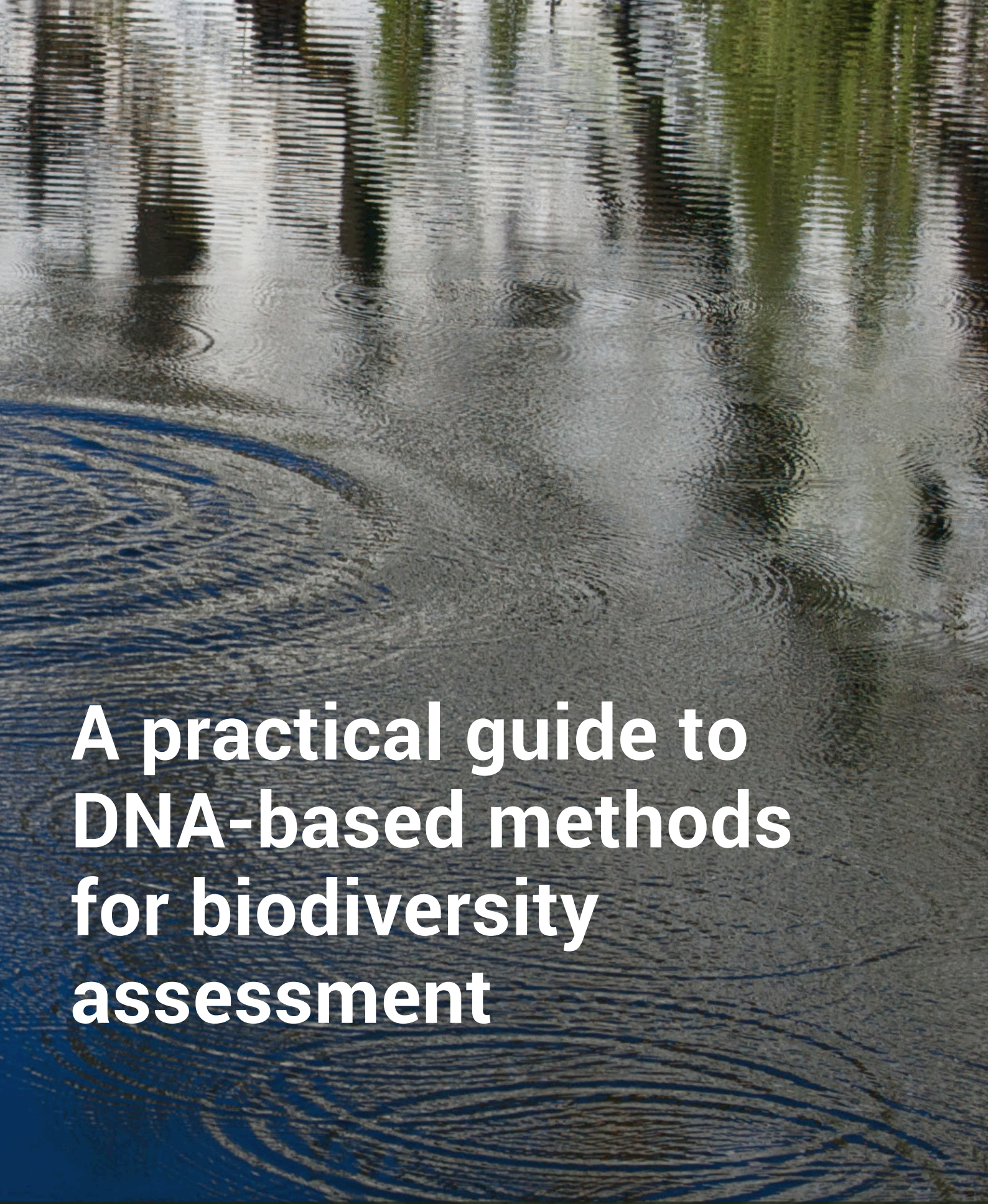

A practical guide to DNA-based methods for biodiversity assessment

Kat Bruce, Rosetta C. Blackman, Sarah J. Bourlat, Micaela Hellström, Judith Bakker, Iliana Bista, Kristine Bohmann, Agnès Bouchez, Rein Brys, Katie Clark, Vasco Elbrecht, Stefano Fazi, Vera G. Fonseca, Bernd Hänfling, Florian Leese, Elvira Mächler, Andrew R. Mahon, Kristian Meissner, Kristel Panksep, Jan Pawlowski, Paul Luis Schmidt Yáñez, Mathew Seymour, Bettina Thalinger, Alice Valentini, Paul Woodcock, Michael Traugott, Valentin Vasselon, Kristy Deiner

PENSOFT®

A practical guide to DNA-based methods for biodiversity assessment

Kat Bruce, Rosetta C. Blackman, Sarah J. Bourlat, Micaela Hellström, Judith Bakker, Iliana Bista, Kristine Bohmann, Agnès Bouchez, Rein Brys, Katie Clark, Vasco Elbrecht, Stefano Fazi, Vera G. Fonseca, Bernd Hänfling, Florian Leese, Elvira Mächler, Andrew R. Mahon, Kristian Meissner, Kristel Panksep, Jan Pawlowski, Paul Luis Schmidt Yáñez, Mathew Seymour, Bettina Thalinger, Alice Valentini, Paul Woodcock, Michael Traugott, Valentin Vasselon, Kristy Deiner

A practical guide to DNA-based methods for biodiversity assessment

Kat Bruce, Rosetta C. Blackman, Sarah J. Bourlat, Micaela Hellström, Judith Bakker, Iliana Bista, Kristine Bohmann, Agnès Bouchez, Rein Brys, Katie Clark, Vasco Elbrecht, Stefano Fazi, Vera G. Fonseca, Bernd Hänfling, Florian Leese, Elvira Mächler, Andrew R. Mahon, Kristian Meissner, Kristel Panksep, Jan Pawlowski, Paul Luis Schmidt Yáñez, Mathew Seymour, Bettina Thalinger, Alice Valentini, Paul Woodcock, Michael Traugott, Valentin Vasselon, Kristy Deiner

Sofia, Bulgaria

2021

A practical guide to DNA-based methods for biodiversity assessment

Kat Bruce, Rosetta C. Blackman, Sarah J. Bourlat, Micaela Hellström, Judith Bakker, Iliana Bista, Kristine Bohmann, Agnès Bouchez, Rein Brys, Katie Clark, Vasco Elbrecht, Stefano Fazi, Vera G. Fonseca, Bernd Hänfling, Florian Leese, Elvira Mächler, Andrew R. Mahon, Kristian Meissner, Kristel Panksep, Jan Pawlowski, Paul Luis Schmidt Yáñez, Mathew Seymour, Bettina Thalinger, Alice Valentini, Paul Woodcock, Michael Traugott, Valentin Vasselon, Kristy Deiner

Cover photo: © P. Bjorn Erikson. Location: Alpine lake in the Sierra Nevada Mountain range, CA, USA.

First published 2021

ISBN 978-619-248-052-3 (paperback)

ISBN 978-619-248-053-0 (e-book)

DOI: 10.3897/ab.e68634

Pensoft Publishers

12, Prof. Georgi Zlatarski Str.

1111 Sofia, Bulgaria

www.pensoft.net

This publication is based upon work from COST Action DNAqua-Net, supported by COST (European Cooperation in Science and Technology).

COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation.

www.cost.eu

EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Funded by the Horizon 2020 Framework Programme of the European Union

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This publication is an output from EU COST Action DNAqua-Net (CA 15219 - Developing new genetic tools for bioassessment of aquatic ecosystems in Europe) and would not have been possible without the opportunities for international collaboration provided by the network, supported by COST (European Cooperation in Science and Technology). Therefore, our highest gratitude is due to Florian Leese and Agnès Bouchez who designed and led DNAqua-Net, and to programme managers Alex Weigand, Sarah Kückmann and Charlotte Frie who coordinated it.

In addition to the authors, hundreds of researchers and practitioners from across Europe and further afield have contributed to the body of knowledge synthesised herein. DNAqua-Net workshops have served as the primary mechanism for consolidating knowledge and were particularly valuable for bringing together research scientists with regulators and end-users, which helped to emphasise the practical considerations in the implementation of DNA-based monitoring programmes. Workshops that ultimately fed into this publication were hosted in Germany (Florian Leese; University of Duisburg-Essen, 2017), Bosnia and Herzegovenia (Belma Kalamujić; University of Sarajevo, 2017), Hungary (Zoltán Csabai; University of Pécs, 2018), Austria (Michael Traugott; University of Innsbruck, 2018), Portugal (Pedro Beja; CIBIO, 2018), Italy (Stefano Fazi; Water Research Institute IRSA-CNR, 2019) and Cyprus (Marlen Vasquez; Cyprus University of Technology, 2019). The workshops highlighted what a collaborative community has emerged among researchers in this field, enabled to a large degree by programmes like DNAqua-Net as well as by a strong collective sense that our research has important real-world applications and is building a foundation for the years to come when we need every tool in the box to promote the protection and recovery of the natural world.

We are particularly grateful to all those non-expert users of environmental DNA who fed back to us their experiences and challenges in engaging with these new methods and provided wider context as to the practical, logistical and financial constraints of routine monitoring (Iwan Jones, Simon Vitecek, Willie Duncan, Kerry Walsh and Martyn Kelly, to name just a few). These insights have helped to guide and shape research priorities, and we hope that this guide will prove a useful resource for these users as they begin to integrate these new technologies into the suite of tools at their disposal.

1. INTRODUCTION

DNA-based methods for species detection and identification have revolutionised our ability to assess biodiversity in terrestrial, freshwater and marine ecosystems. Starting from the seminal study that used eDNA to detect invasive american bullfrogs in France (Ficetola et al. 2008), research conducted over the last decade has demonstrated the power of these approaches for surveying a wide range of species and groups. Early applications included the use of eDNA to monitor Asian Carp in the USA (Jerde et al. 2013). Following heavy scrutiny, the method was eventually adopted, and is still employed today by the United States Geological Survey (USGS). A flurry of research followed, with tests designed for many threatened and invasive species including New zealand mudsnails (Goldberg et al. 2013), american crayfish (Geerts et al. 2018), gammarids (Blackman et al. 2017), and great crested newts (Biggs et al. 2015). The great crested newt eDNA test has been employed for regulatory monitoring in the UK since 2014. During the same time period, there was a proliferation of research studies that used high-throughput sequencing approaches to describe whole communities of organisms from mixed species and environmental samples, using an approach termed DNA metabarcoding (Taberlet et al. 2012c).

As the field developed fast and the approaches were applied to a wide range of research and monitoring objectives, a high level of methodological variation was introduced at all stages of the workflow (Seymour 2019). Thus, while a significant level of consensus on scientific best-practice now exists in many areas, this may not be readily discerned from the now-extensive body of research literature.

As environmental practitioners and policy makers are now increasingly starting to integrate DNA-based methods into routine monitoring applications including protected species licensing[1], statutory monitoring[2] (Hänfling et al. 2016) and environmental impact assessment[3], various national and international efforts have been undertaken to standardise methods and integrate them into monitoring frameworks (Pilliod et al. 2019, Loeza-Quintana et al. 2020, Minamoto et al. 2021, Pawlowski et al. 2020a[4]). In Europe, the EU COST Action DNAqua-Net (Leese et al. 2018) has been working towards incorporating molecular monitoring tools for Biological Quality Elements (BQEs, e.g., fish, macroinvertebrates and phytoplankton-benthos) into the Water Framework Directive (WFD, 2000/60/EC)[5] and the Marine Strategy Framework Directive (MSFD, 2008/56/EC)[6].

Thus, emphasis now shifts from fundamental research to robust and efficient application of DNA-based methods for operational use at large scales. This requires that scientific robustness is balanced with consideration of the practical realities faced by environmental managers. Moreover, there is increased need for strong quality assurance in a setting where non-expert field samplers and commercial laboratories are involved with the generation of data that non-specialist decision-makers then rely on to inform potentially costly action (or non-action). This places increased emphasis on robustness, replicability, traceability and ease-of-use, which may not always be the central focus of studies carried out in the academic research environment.

This document aims to summarise the scientific consensus relating to every step of the field and laboratory workflows involved in the most common types of samples and analyses. We do not go into great detail regarding bioinformatics (computational processing of sequence data) and data analysis since these

1 https://naturalengland-defra.opendata.arcgis.com/datasets/ffba3805a4d9439c95351ef7f26ab33c_0

2 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/575833/A_DNA_based_monitoring_method_for_fish_in_lakes_-_report.pdf

3 https://cieem.net/wp-content/uploads/2020/03/InPractice99_Mar2018.pdf

4 https://www.bafu.admin.ch/bafu/en/home/topics/water/water--publications/publications-water/environmental-dna-applications-in-biomonitoring-and-bioassessment.html

5 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32000L0060

6 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32008L0056

are extensive topics in their own right. We uniquely set the field and lab steps in the context of the practical and logistical constraints faced by environmental managers in terms of cost, logistics, safety, ease-of-use, and quality assurance, highlighting key decisions to be made and the inherent trade-offs associated with the various options. We hope that this will support non-experts, and those new to the field, to navigate the key considerations associated with planning or evaluating monitoring programmes using DNA-based monitoring methods. Additionally, it will aid decision-makers in writing and evaluating tenders and proposals, ensuring that the methods used for a given project are fit-for-purpose and that results are correctly interpreted.

Alongside the many areas of emerging consensus, there remain some areas where further research is still required to balance scientific best-practice with the constraints and priorities of end-users. We hope that by shining a light on the importance of these issues, the research community will be encouraged to address them. More generally, we hope to inspire researchers in this now highly-applied scientific field to consider end-user constraints when designing and implementing research projects. This will help to accelerate uptake by users and maximise the impact of research.

DNA-based bioassessment methods continue to evolve, and there are several emerging technologies that show exciting promise to move beyond even what is possible today. Examples include in-field sequencing using the MinION device from Oxford Nanopore Technologies (Pomerantz et al. 2018, Davidov et al. 2020, Hatfield et al. 2020), PCR-free metagenomic approaches (Bista et al. 2018, Giebner et al. 2020) and CRISPR for rapid detection of species, which is particularly relevant for invasive an non-native species monitoring (Williams et al. 2019, 2020). We recognise the potential of these methods, but do not consider them in detail here, since they are not yet far enough developed for routine application.

1.1 The application of DNA-based methods to biomonitoring

The two main challenges of bioassessment are (1) the detection of species and (2) their correct taxonomic identification. DNA-based monitoring tools can help address both aspects. For small-bodied and species-diverse groups such as benthic macroinvertebrates and diatoms, the monitoring challenge lies not so much in species detection but in the need for rapid, cost-effective and accurate identification of taxa. The best-validated approach for DNA-based biomonitoring of these groups is to follow established sample collection protocols (as outlined in existing standards e.g. ISO 16665:2014; ISO 10870:2012; ISO 10870:2012; ISO 10870:2012; CEN/EN 13946:2014; CEN/EN 14407:2014), substituting morphological identification of taxa with metabarcoding and DNA-based taxonomy (Hering et al. 2018). A Technical Report on sampling benthic diatoms has been adapted for metabarcoding and published by the European Committee for Standardisation[7], and in 2019 a new working group was established within the CEN Technical Committee on Water Analysis (CEN/TC 230) for the development of standards for DNA-based assessment of aquatic biodiversity. This represents a clear marker of the appetite for uptake of these tools.

For fish and invasive non-native species, the challenge for monitoring lies principally in detection rather than identification. Conventional fish survey methods (e.g., electrofishing and netting) are labour-intensive, inefficient for community assessment, and often cause harm or stress for the fish (Snyder 2003), while surveillance for invasive species lacks the sensitivity to detect target species at low population levels (e.g., when first introduced and not yet established) when a rapid management response could minimise the overall cost and impact of the invasion. This creates a strong driver for new sampling methods that increase detection sensitivity for these groups while lessening the physical impact on fish. Therefore, for these target groups we focus on surveys using aquatic eDNA.

---

[7] CEN, 2018. CEN/TR 17245: Water quality –Technical report for the routine sampling of benthic diatoms from rivers and lakes adapted for metabarcoding analysis. CEN/TC 230/WG23, pp. 1–8.

1.2 Document structure

We first consider field sampling and preservation methods for each of four sample types:

- aquatic eDNA
- bulk invertebrates
- benthic periphytic diatoms
- soils/sediments

We then outline key quality control checks to be applied to DNA extracts and a framework for positive and negative controls to be integrated into the workflow.

Next we give a detailed overview of the laboratory steps, decisions and trade-offs associated with the two broad approaches to sample analysis:

- Targeted species detection using qPCR and allied methods.
- Community assessment using metabarcoding

For completeness, we give a brief overview of the major choices and considerations in bioinformatics processing, focusing on those that materially affect the results obtained.

Finally, we summarise the key factors that influence methodological decision-making and outline key practical recommendations for DNA-based biomonitoring (See **Figure 1** for an overview of the document contents).

1.3 Sources and states of DNA

DNA can be captured in various states, and the state in which it is captured influences how it needs to be handled, processed and interpreted. In particular, we make a key distinction between organismal DNA, which is captured in the form of whole organisms, and extra-organismal DNA, which is captured in the absence of the organism they originated from.

While various definitions have been proposed and employed, we define environmental DNA (eDNA) as genetic material that has been isolated from environmental samples such as water, soil or air (Taberlet et al. 2012a, Pawlowski et al. 2020b). A key feature of eDNA is the presence of both organismal DNA from microscopic organisms such as protists and bacteria (organismal DNA) and extra-organismal DNA from larger organisms (Pont et al. 2018, Rodriguez-Ezpeleta et al. 2021).

Spatial and temporal interpretations based on detection of species from extra-organismal DNA are complex because the DNA may have travelled away from the point at which it was released from the organism. Extra-organismal DNA is also typically present at very low concentrations in a sample, which makes it highly vulnerable to contamination. Special precautions need to be taken both in the field and the lab in order to mitigate this risk (Goldberg et al. 2016). The inference challenge and low concentrations have particular implications for the design and validation of methods used for the detection of DNA from environmental samples (i.e., eDNA assay) and for the level of replication required to overcome inherent stochasticity associated with very low target concentrations. We cover these implications in depth below, but briefly outline the main principles here.

In the field, key considerations in working with eDNA include:

- Avoiding situations where the same sampling equipment comes into direct contact with multiple independent samples, unless the equipment is decontaminated in between.
- If sampling from a boat, collecting water from the bows to minimise the risk of contamination from the boat itself. In rivers, start sampling at the most downstream point and travel upstream.
- In lentic water bodies (non-flowing water), avoid entering the sampling area prior to or during sample collection in order to avoid transfer of DNA from footwear or clothing. In lotic (flowing) water bodies where it is necessary to enter the water, standing downstream of the water you collect.
- Wearing disposable gloves to avoid introducing your own DNA to the sample and to reduce the risk of cross-contamination between samples.
- Frequent use of negative field controls, especially where equipment is being decontaminated and reused (see section 6.3 for more details).

In the lab, key aspects of working with eDNA include:

- Dedicated clean-room facilities for working with low quantity DNA.

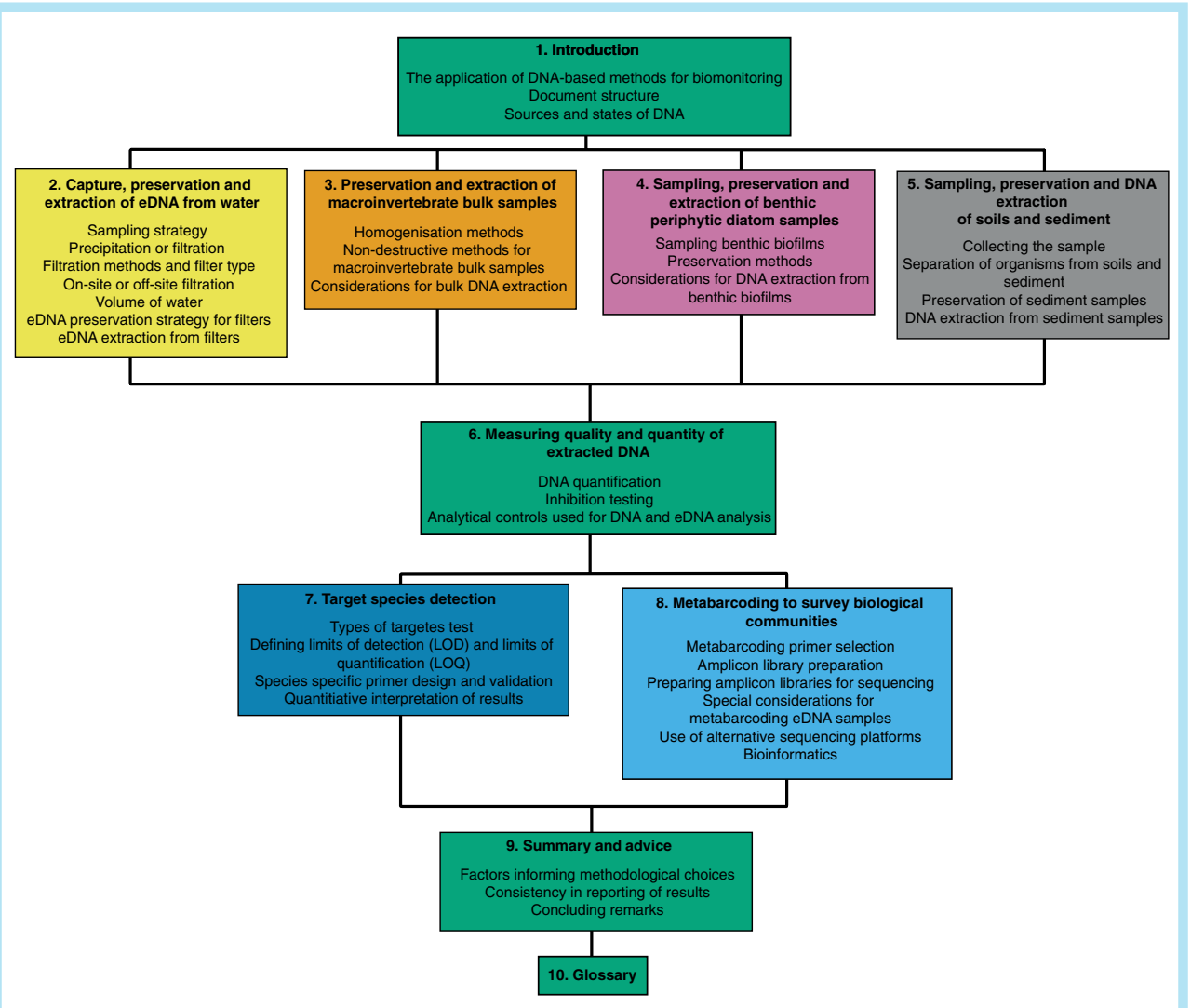

Figure 1. Practical Guide overview.

- Physical separation of the spaces used for different stages of the analysis, and implementation of a unidirectional workflow, with particular attention to separating pre- and post-PCR processes.
- Frequent decontamination of surfaces and equipment.
- Diligent use of gloves, which are regularly changed to reduce the risk of cross-contamination between samples, and face masks to avoid introducing human DNA where this could interfere with the analysis (e.g. if analysis targets a taxonomic group that includes humans such as primers amplifying mammals, vertebrates, and even general eukaryotic primers).
- Frequent use of negative lab controls at each new step of the process (i.e., filtration, extraction, PCR and indexing samples if pooled onto a single sequence run).

2. CAPTURE, PRESERVATION AND EXTRACTION OF eDNA FROM WATER

2.1 Sampling strategy

A variety of different water collection strategies provide robust data with good detection probabilities. These range from sampling continuously across the area of the waterbody for a set period of time, or pooling subsamples from different point locations into single merged sample (Pont et al. 2019), to taking multiple discrete samples spread out across the area of the waterbody (Hänfling et al. 2016). The latter provides more information about the spatial distribution of species and enables analysis of occupancy, but the greater number of samples will increase overall cost.

As with any ecological survey, robust sampling design prior to field collections is essential to ensure the data obtained are fit for the purpose required. A significant advantage of an eDNA approach is that biological replication - crucial for robust statistical analysis - is easily incorporated into survey design.

Environmental DNA sampling design needs to account for (1) the physical and chemical properties of the matrix from which it is isolated, (2) environmental variability, and (3) the ecology of the target species to be surveyed. While many studies have calculated sampling effort for particular species in given environments using occupancy modelling and allied methods (e.g. Erickson et al. 2019), conclusions are often difficult to extrapolate to other environments, species groups and analysis workflows, which can vary in efficiency. Here, we outline some general principles that can be applied to identify circumstances when greater sampling effort is likely required for species detection.

Environmental DNA persistence in space and time is influenced by a multitude of factors. These include season, waterbody size and depth, temperature, stratification, connectivity, substrate, water chemistry, and flow. It is often difficult to tease apart specific effects, especially in natural settings, since combinations of factors will work synergistically or antagonistically to directly or indirectly facilitate degradation of eDNA (Stewart 2019). We do not attempt to provide a complete review of these factors in this guide (see Harrison et al. 2019 and Torti et al. 2015 for reviews), but instead highlight those most likely to impact species detectability.

One of the most important aspects of eDNA is its spatial distribution in the environment, which integrates how far eDNA travels from its "point of release" and how well mixed it is in the water column (Deiner and Altermatt 2014, Deiner et al. 2016, Hänfling et al. 2016, Jerde et al. 2016, Shogren et al. 2017, Macher and Leese 2018, Pont et al. 2018). This is worth considering in some detail for each of the major types of water body (summarised in **Table 1** and **Figure 2**):

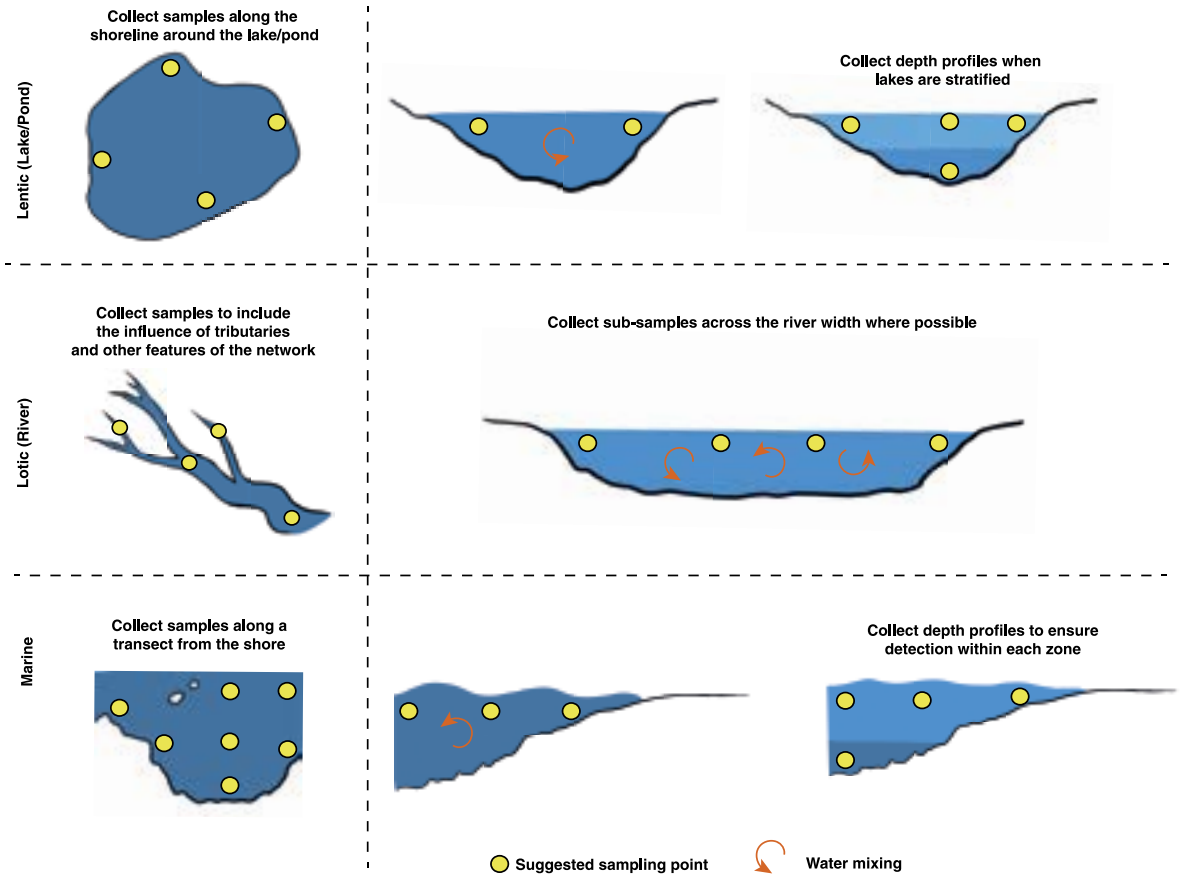

Figure 2. Water eDNA sampling design. Sampling locations should consider the spatial distribution of eDNA within the environment being sampled, i.e., lentic, lotic or marine conditions. This figure indicates possible sample collection points in each of these environments.

Table 1. Summary of considerations for sampling eDNA from water in aquatic systems.

To consider	Applies to all types	Lentic	Lotic	Marine
When to sample?	• When the target species is most likely to be in the waterbody based on what it is known about its ecology and life history • During spring and summer higher bacterial and algal load may interfere with the analyses • To coincide with statutory monitoring	• Consider seasonal thermal stratification classification (7 groups) patterns; In seasonally stratified lakes a more efficient sampling strategy can be deployed when the waters are mixed and samples from depth are not required	• Sample during typical flow levels - avoiding low flows/ drought and flood conditions • Consider seasonal patterns of migratory species	• Nearshore- consider season. Many fish species move to shallow waters for mating and move to deeper waters in the winter, simultaneously some species prefer cold deep water in the summer season • Consider migration patterns, mating & spawning / breeding sites

To consider	Applies to all types	Lentic	Lotic	Marine
Where to sample?	• Avoid entering the water prior to sampling to avoid transfer of DNA on footwear or clothing, transfer of pathogens from other sites and disturbing the substrate • If necessary to enter lotic water, stay downstream of sampling points • Consider sewage pipes, nearshore restaurants and close human dwellings as contaminants from these sources may affect the eDNA results	• Collect samples from around the shoreline/edge of the lake/pond • Also collect from the middle of the lake if there is large variation in depth or if the lake is stratified	• Where possible, collect subsamples across the river width, including flow types such as riffles and pools • Sample at regular intervals along the river network • Consider tributaries, any connected lentic water bodies and changes in elevation	• Sample collection should consider depth profiles, habitat heterogeneity and current/tidal influences • Several depths should be included if trying to capture the full community
Sample number	• The number of samples should reflect the spatial complexity, size of the system and access to the area you wish to represent • Also aim to include subsampling of distinct sub-habitats (such as areas of differing flow or vegetation) to reflect the habitat as a whole • To reduce the number of samples, the extent of pooling of subsamples can be increased. However, this reduces statistical power, detection probability and spatial resolution of the data	• The number of samples will depend on: • Size of the waterbody and accessibility • Water sampling strategy (whether or not merging subsamples and how much water can be passed through a filter) • Topographic variation of shoreline - more samples are required in more complex habitats	• Flowing water should be sampled at regular intervals along the river length to ensure the collection of eDNA prior to degradation or dropping out of the water column • Some replication at each site is always recommended to increase confidence in data	• Depends on the spatial scale of the area you wish to represent • Deeper water will require more samples to cover the different depth zones of the water column • Offshore sampling will need more samples to account for the very high dilution factor
Sample volume	• The sample volume is dependent on a number of factors including turbidity, access to pumps and onsite or lab filtering protocols. Studies show a wide variation in the volumes chosen from 500ml - 50 L, however most studies which filter water process between 500 ml to 5 L per technical sample • Sample volume also depends on sampling strategy; pooled subsamples may require fewer replicates	• Small ponds may require less sampling volume than larger lakes, however filterable volume may be lower in small ponds due to turbidity - aim to maximise the volume filtered	• eDNA distribution in river systems may be very stochastic and dilute compared to lentic samples • Taking regular samples/subsamples is important	• eDNA in marine systems is very dilute therefore you should maximise your sample volume to be representative of the environment • Good results have been obtained with 2-5 L samples but depends on the target taxa (microbial taxa usually require smaller volumes)

To consider	Applies to all types	Lentic	Lotic	Marine
Turbidity	• Turbidity found in water samples causes a number of problems while sampling (i.e., filtering time) and can also cause inhibition	• Freshwater samples can often be turbid. Avoid disturbing the substrate when sampling. Consider using a prefilter or larger pore size, and avoid sampling after rainfall or during algal bloom events		• Usually less of a problem in marine waters although some inshore areas can become turbid due to coastal run-off and wave action disturbing the seafloor

- In ponds and lakes, eDNA is patchily distributed at small-scale intervals (Brys et al. 2020a), meaning that water needs to be sampled (or sub-sampled) from multiple points to maximise detection probability. Seasonal circulation patterns can influence vertical distribution of eDNA in the water column of large lakes. This is largely driven by water temperature, which determines the extent to which the water column is mixed or stratified (e.g. thermal stratification), and seven different thermomix groups have been described (Thomas et al. 1996). Environmental DNA is well mixed during periods of thermal circulation, meaning that surface samples are representative of the whole water column, while during stratification periods the water is layered and eDNA needs to be collected from the different layers for complete sampling. For instance, studies from lakes on both sides of the Atlantic ocean have demonstrated that in monomictic lakes eDNA is stratified during the summer months but homogeneously distributed in the water column in winter (Handley et al. 2019, Littlefair et al. 2020). In practical terms, this means cold-water species occupying the deeper water can be detected in surface water or shoreline samples collected in these lakes during the winter, but only from water collected below the thermocline in summer. Many Scandinavian lakes are dimictic, meaning that they undergo two circulation periods (spring and autumn), and two thermal stratification phases (summer and winter).
- In rivers, eDNA is more evenly mixed in the water column but sampling design needs to consider downstream transportation and tributary dilution of eDNA from its point of release. This means deciding whether to sample upstream or downstream of tributaries or potential sources of environmental contamination, depending on the aim of the monitoring. For instance, sampling immediately downstream of inhabited areas carries a risk of detecting species whose DNA in the water originates from food items via wastewater rather than the presence of the species itself in the local aquatic ecosystem. Modelling the spatial distribution of eDNA in rivers is challenging (Harrison et al. 2019) and it is not yet clear how results from particular studies can be transformed into general models that are applicable across catchments. Thus, some element of spatial uncertainty must usually be accepted when sampling eDNA in rivers. As a guiding principle, the greater the flow rate, the larger the upstream area represented by a sample. This may vary from a few hundred metres in slow-flowing lowland streams and rivers to tens of kilometers in fast-flowing systems (Pont et al. 2018, Seymour et al. 2018). When evaluating the impact of barriers in rivers, eDNA is powerful for indicating lack of upstream movement, but caution should be applied when interpreting results relating to downstream movement. Robust statistical approaches should be integrated into the design of monitoring programmes, particularly where downstream movement of species is of relevance to the aims of the monitoring, and the specific hydrology of the catchment area should

be taken into consideration. Detailed discussion of sampling strategy in rivers is given by (Carraro et al. 2020).

- In oceans, less is known about how hydrological systems affect eDNA transportation and distribution and a lot more research is required to fully understand how to optimise marine sampling strategies and spatially interpret results. However, several studies have shown that communities obtained from eDNA metabarcoding are surprisingly representative of the immediate local habitat where the sample was collected (e.g. Port et al. 2016, Yamamoto et al. 2017, Jeunen et al. 2019, Djurhuus et al. 2020). Like in lakes, vertical stratification in the water column (e.g. due to thermoclines and haloclines) restricts mixing of DNA, meaning that water samples should be collected from each different depth zone to fully characterise marine communities at a particular point location (Jeunen et al. 2020).

Chemical, physical and biotic factors influence the persistence of eDNA in the environment by affecting the rate at which it is degraded. Faster degradation reduces the time window for species detection, which carries advantages and disadvantages. On the one hand, it gives less opportunity for the DNA to travel long distances from the point of release, giving somewhat greater precision in temporal and spatial inference. On the other hand, sampling may need to occur more frequently or with greater spatial sampling effort to fully characterise communities.

Factors that affect the rate of eDNA degradation include:

- Low pH is generally associated with a faster rate of degradation (Strickler et al. 2015, Seymour et al. 2018). A mechanistic assessment of pH-related degradation of eDNA has yet to be conducted, but some speculation can be derived from classical molecular biology. Degradation of DNA is particularly likely when positively charged enzymes, indicative of acidic conditions (i.e., low pH), are present, which is why preservation buffers for DNA extracts are typically alkaline (e.g., Tris, EDTA buffer at pH 9). Moreover, extracted DNA will degrade if left in water due to acid hydrolysis, particularly below pH 7.5 (Torti et al. 2015).
- eDNA degrades faster in warmer water because of increased microbial activity, which is a strong driver of eDNA degradation (Zulkefli et al. 2019). However, in practice the faster degradation of eDNA in warmer waters will usually be offset by increased rate of eDNA production, linked to greater activity under these conditions (discussed below). Oxygenation levels may also play a role in degradation rate, and this is also linked to water temperature, with oxygen saturation decreasing as water temperature rises (Bozinovic and Pörtner 2015). DNA structure is highly stable in dry anoxic conditions, but decays rapidly via hydrolysis in oxygenated environments (Torti et al. 2015).
- High nutrient loading is often associated with increased microbial activity, which is expected to negatively impact eDNA detection through increased consumption or absorption of the freely available genetic material (Barnes and Turner 2015). This means that eDNA may degrade faster in environments with high nutrient inputs including agricultural run-offs, though this has yet to be tested.

It is also important to note that different subcellular components degrade at different rates, so detectability may vary according to the gene region targeted for analysis. Most commonly-used mitochondrial gene fragments (e.g. COI, 12S and 16S) will persist in the environment longer than ribosomal DNA fragments (e.g. 18S) due to the more resilient structure of the mitochondria once cells start to degrade. However, the higher abundance of ribosomal genes may offer a better alternative for localized monitoring under certain conditions (Jo et al. 2019b, Moushomi et al. 2019).

Factors related to the biology and ecology of the target organism(s) can also affect detection probability. These include:

- Habitat preferences and patterns of seasonal activity and movement, which affect the probability a particular species will be present or detectable in a given water body at a given time of year.
- Aspects of life-history mean that some lifestages may be more detectable than others. For instance, some species or groups of organisms shed very little DNA as adults (e.g. crayfish; Rusch et al. 2020) but are readily detected during breeding season when adults are more active, eggs and sperm provide an abundant source of eDNA, and juveniles are growing and moulting.
- Biological and physiological traits affect eDNA detectability in several ways:
  - Some animal forms shed more DNA than others. For instance, animals with exoskeletons seem to shed less DNA (Allan et al. 2020), as do those with dry, scaly skin such as reptiles. Conversely, those that produce lots of mucous or shed skin / scales, are more readily detectable.
  - Faeces represent a major source of eDNA, meaning that toilet habits are a key predictor of detectability. Some semi-aquatic species that use latrines on land tend to be underrepresented in eDNA surveys (e.g. otters), while terrestrial animals that commonly defecate in or above the water may be overrepresented (e.g. tapirs). This also means that frequency of feeding and defecation (i.e. energy use) are linked to detectability of aquatic species (Klymus et al. 2015). It therefore follows that ectothermic (cold-blooded) organisms are expected to shed less DNA than endothermic (warm-blooded) ones.

Behavioural factors interact with life-history and physiology to affect the amount of DNA released by a given species at any one time. Animals have been documented to release more eDNA when they are stressed, when they are active, and when they are warm (Jo et al. 2019a, Thalinger et al. 2021a).

- Thus, if animals become less active during cold weather, they may be expected to release less eDNA and therefore to become less detectable during these times, although note that this may be offset to some extent by slower degradation of eDNA at lower water temperatures, as mentioned above.

Taking these factors into account allows estimation of how detection probability is likely to vary temporally and spatially for a given species, and more intensive sampling regimes may be required when conditions or timing are not optimal for detection, or when the species' behavioural or physiological traits mean that it is likely to be underrepresented in aquatic eDNA. **Table 2** summarises factors that could be expected to reduce detection probability and recommends how to adjust sampling strategies to account for these. Note that some of these factors have not yet been studied in great depth and they will not affect all target species to the same extent in all environments. However, where several of these factors apply, it should be considered that greater sampling effort may be required in order to achieve a high detection probability.

2.2 Precipitation or filtration

Two principal methods have been used for capture of eDNA from water.

- Ethanol precipitation involves adding water to an ethanol and salt (usually sodium acetate) solution. At low temperature, the salt changes the charge of dissolved DNA so it becomes hydrophobic and comes out of solution (precipitates). Precipitated DNA is then forced into a pellet at the bottom

Table 2. Factors that could be expected to reduce detection probability for aquatic eDNA.

	Factor	Reason for possible lower detection	Mechanism	Counteracting factors	Recommendations
Habitat properties	**Cold water**	Reduced eDNA production & reduced mixing in lakes	Reduced activity of some animals Stratification of water column	Greater persistence (accumulation) of eDNA through reduced microbial activity. Some groups have higher detection probability in cold water	• Increase sampling effort if targeting species that are likely to be less active in cold water • Collect samples from different depths if thermocline likely to be present
	Warm water	Faster degradation of eDNA	Increased microbial activity	Increased activity of some animals in warmer conditions leads to greater production of eDNA	• Increase sampling effort in warm water especially if target is expected to be rare or transient, and if it is not expected to be more active in warmer conditions
	Large water volume	Reduced eDNA concentration	Dilution		• Increase sampling effort in line with water body size
	Low pH	Faster eDNA degradation	Positively charged enzymes		• Consider increasing sampling effort in acidic environments
	High nutrient inputs	Faster eDNA degradation	Microbial activity		• Consider increasing sampling effort in water bodies with high nutrient input
Target species properties	**Exoskeleton**	Reduced eDNA production	Physical barrier to eDNA release	Moulting, release of gametes	• Greater sampling effort needed for arthropods • Sample during and after breeding season when juveniles are growing
	Ectothermic	Reduced eDNA production	Lower metabolism & less shedding	Production of mucous or shedding of skin/ scales	• Increase sampling effort, especially for reptiles
	Low activity	Reduced eDNA production	Lower metabolism & less shedding	Greater accumulation of eDNA	• Increase sampling effort when activity is expected to be low
	Terrestrial latrine	Reduced eDNA input to water	Major source of eDNA lacking	May be more detectable after rainfall	• Increase sampling effort • Try to sample after rainfall
	Not fully aquatic	Reduced eDNA input to water	Inconsistent release of eDNA	May be seasonal	• Increase sampling effort and align sampling with species' expected use of aquatic habitats

of the tube through centrifugation and the ethanol solution discarded. DNA extraction is applied to the pellet.

- Filtration involves passing water through a fine porous membrane. Cellular and subcellular material is captured on the membrane and preserved ready for DNA extraction.

Although early studies (e.g. Ficetola et al. 2008) used ethanol precipitation as the primary capture method for eDNA, there is now broad consensus that filtration is a more effective approach for detection of aquatic species in most environments (Tsuji et al. 2019; see Text Box 1 for more details), and we recommend that this is the first method to consider for aquatic eDNA monitoring programmes until different or new knowledge is gained about the benefits of other methods. Henceforth in this guide, we focus on the various decisions that one must make when using filtration-based capture methods.

Note that in certain environments where cellular breakdown happens very fast (e.g. environments of extreme heat or acidity), precipitation may be a more effective method because a greater proportion of the available DNA will be extracellular.

Why use filtration instead of ethanol precipitation?

Here, we give a detailed explanation of the reasons for our recommendation to use filtration for eDNA capture because it runs contrary to one of the few cases in which eDNA is currently applied within a regulated monitoring context - detection of great crested newts (*Triturus cristatus*) in the UK - for which ethanol precipitation is stipulated in the standard protocol.

Sensitivity: The likely greatest limitation of the precipitation approach is the volume of water that can be processed because the corresponding volume of ethanol required for precipitation quickly becomes prohibitive in terms of both cost and logistics. For instance, the precipitation-based protocol widely employed in the UK for capturing eDNA of the Great Crested Newt (*Triturus cristatus*) (Biggs et al. 2015) samples a total of just 90 ml of water from each pond up to 1 ha in size. This small volume limits detection rates compared with a filtration-based approach (Muha et al. 2019). Moreover, multiple studies (e.g. Deiner et al. 2015, Spens et al. 2017) have shown that filtration outperforms precipitation even when applied to equal volumes of water.

Contamination risk: Precipitation-based methods commonly use multiple separate collection tubes per sample to maximise the volume of water tested, but the DNA must then be combined into a single tube during extraction, which typically involves vortexing to dislodge the DNA from the tube surface and then pouring all the DNA pellets into a single tube. This is an imprecise process compared to other processes used in molecular biology, and poses a high risk of cross-contamination. DNA extraction protocols from filters are typically simpler, quicker and more contained, which both lowers the labour cost per sample and reduces the risk of sample cross-contamination.

Logistics, safety & disposal: The ethanol precipitation approach requires relatively large volumes of ethanol, which is subject to extremely heavy taxation in some countries unless it can be procured under a duty-free licence. It is also a flammable liquid and therefore classed as dangerous goods for transportation purposes (molecular grade ethanol: UN1173, class 3, packing group II) which means that specialist couriers are required, specific packing requirements apply (especially for air transport under IATA regulations), and shipment costs can become high. From a safety perspective, ethanol fumes from a single kit pose a negligible risk in the field during sampling, but pose a much greater risk in the lab, especially when large numbers of samples are processed. This kind of work requires adequately-ventilated laboratory spaces and should ideally take places in fume hoods with extractors. Where testing is taking place at large scales, huge volumes of ethanol waste are generated, which is highly flammable and must be properly stored and disposed of through specialist waste companies at significant cost. All these issues can be avoided by the use of a filtration-based eDNA capture approach.

2.3 Filtration method and filter type

A wide variety of equipment is used for filtration-based capture of eDNA from water, including different filter membrane materials, pore sizes, and filtration mechanisms, different transportation, storage and preservation methods, and different DNA extraction protocols. Due to the number of variables in a given workflow, it is difficult to robustly assess the importance of the choices made at any particular step. Moreover, there are almost certainly interactions between different elements of the process; for

instance, certain membrane materials or filter designs may be best suited to use with particular pore sizes or DNA extraction methods, while others work optimally with a different combination of choices (Deiner et al. 2018).

Our overarching message is that the first step should be to identify the main constraints for the project (e.g. time, remoteness, budget, availability of equipment), the characteristics of the study system, and the key aspects of the methodology that they affect. Next, optimisation should be carried out to determine the most effective combination of choices in the rest of the workflow, given the identified constraints. Reassuringly, many studies now show that with appropriate workflow optimisation, eDNA analyses are highly robust to different choices made (e.g. Li et al. 2019, Di Muri et al. 2020).

We identify three distinct categories of filters used for eDNA capture (See **Figure 3**):

1. **Open filters** are membranes that are exposed to the air during filtering. These are commonly used with vacuum pumps when filtering in a laboratory setting, and also include cup-type filters such as

Figure 3. Water eDNA filter types. 1: Open filters are exposed to the air during filtration either in the field or lab (a and b). **2: Housed filters** are a membrane placed in a solid unit during filtration (c). Filters from Open and Housed units need to be removed from the filtration unit and stored in a petri dish or eppendorf tube until extraction (c and d). **3: Enclosed filters** are systems in which the membrane is enclosed within the outer housing (e and f). Extraction is carried out directly from the enclosed filtration unit.

the Analytical Test Filter Funnel from Nalgene, which is commonly used with peristaltic pumps in the field. These filters carry the greatest risk of contamination because not only are they exposed during filtering, but they must usually be handled during the preservation process and usually again during DNA extraction. This handling must be carried out with the utmost care to avoid cross-contamination, and the use of negative filter controls is highly recommended (Goldberg et al. 2016).

2. **Housed filters** describe systems in which a filter membrane is protected within a solid housing during the filtration process. Their advantage is these filters can easily be opened and the filter removed for preservation and later processing. When opened in the field, care must be taken not to contaminate the samples. Preservative solutions can be added directly to the filter membrane inside the housing and transported to the laboratory, cleaned on the outside and opened in the laboratory in clean room conditions in order to avoid contamination. As with the open filters, transferring the filter carries a risk of contamination and must be carried out with great care and appropriate negative controls to check for contamination. However, this type of housing allows for the greatest flexibility of filter types and extraction methods, including those that involve bead beating.
3. **Enclosed filters** are systems in which the membrane is enclosed within an outer housing from which it is not removed, and no handling of the filter membrane is required during the process. Preservative solutions can be added directly to the filter membrane inside the housing. Examples include Millipore's Sterivex units and various varieties of disc filters. These systems are the most robust to field contamination and are likely to be the favoured option for large-scale surveys where samples are being collected by contractors or volunteers who are not molecular biologists. However, limited combinations of pore sizes and membrane materials are available, and the units themselves tend to be more expensive than open or housed filters. Enclosed filters are incompatible with mechanical lysis (the physical breakdown of cell membranes), which is needed for DNA extraction from diatoms captured in the water sample. Therefore, open or housed filters should be used if you aim to assess pelagic diatom from the eDNA samples.

The combination of membrane material and pore size is important for determining how much volume can be filtered through a single unit. A membrane must be hydrophilic for water to easily pass through, and note that some membrane materials (e.g. PVDF) can be purchased in both hydrophilic and hydrophobic versions.

Commonly used materials include Cellulose Nitrate (CN), Polyethersulphane (PES), Polyvinylidene Difluoride (PVDF), glass fiber (GF) and Polycarbonate Track Etched (PCTE). Some membrane material may introduce constraints in relation to other parameters; for instance, GF membranes cannot be produced with consistent pore sizes due to the matrix-like nature of the material, so the stated pore sizes are nominal only and are rarely below 0.7 μm. PCTE membranes are difficult to incorporate into enclosed filters because the material is less strong than others and cannot withstand the same amount of pressure during filtration without additional support. Cellulose nitrate filter membranes have been found to disintegrate when stored for long periods in ethanol.

Pore size selected for eDNA capture varies substantially but is most commonly below 1 μm. The volume of water that can be passed through a filter membrane generally increases with pore size due to a reduced rate of clogging, but the trade-off is that the smallest particles containing eDNA may pass through the filter. Experiments on fish eDNA (Turner et al. 2014, Wilcox et al. 2015) have shown that most eDNA-containing particles are in the range of 1-10 μm, so eDNA loss is expected to be minimal for pore sizes up to 1 μm and even a little larger (Deiner et al. 2018). Several studies targeting macrofauna have shown that filtering more water with a larger pore size is more efficient than filtering less water with a smaller pore size (Hosler 2017, Sepulveda et al. 2019), and even very large pore sizes 64 μm (Schabacker et al. 2020) have been used with success to capture fish eDNA. However, smaller pore sizes (<1 μm) should

be used to target species such as macroinvertebrates, which produce smaller quantities of eDNA than fish (Moushomi et al. 2019).

The most commonly used pore sizes in published studies are 0.22 µm and 0.45 µm. While there are often good reasons for choosing these smaller pore sizes, this is also partly a function of the limited range of commercially available filters with larger pore sizes. For instance, the widely-used Sterivex filters are only available in 0.22 µm or 0.45 µm. Note that if you aim to capture total microbial diversity as well as eDNA from larger organisms, a smaller pore size is advised (generally around 0.2 µm; Lee et al. 2010).

In particularly turbid waters, or where a small pore size is required, pre-filtering through a membrane with a larger pore-size can be an effective way of maximising sample volume by removing larger particles of sediment and plant material before passing the water through the main filter for capturing eDNA. This carries a risk of losing some eDNA particles during pre-filtering so it is recommended that DNA is extracted from the pre-filter and processed alongside that from the main filter, at least until it can be determined that results from the main filter are not negatively impacted by the prefiltration.

2.4 On-site or off-site filtration

Many published studies describe collecting water in sealed containers and transporting it to a clean laboratory for filtration using vacuum pumps (Jerde et al. 2011, 2013, Hänfling et al. 2016). This approach is attractive for speed and simplicity in the field. However, since eDNA degrades quickly, the water must be kept refrigerated during transportation and either filtered on the same day as collection or frozen for storage. This can be impractical and expensive, especially for large-scale sampling campaigns.

The alternative is to perform filtration on-site, either manually with syringes or hand pumps, or with the aid of a powered pump (vacuum or peristaltic) (**Figure 4**). Manual filtration represents an inexpensive and universally applicable solution, but it can be hard physical work and time consuming depending on the targeted volume of water per sample, the particle load in the water and number of sites to be completed. Filtration with powered pumps makes it easier to process larger volumes of water, but may be unfeasible in situations where sampling is carried out by multiple field teams in parallel, or for use in remote areas where it is not possible to carry in the equipment and power supply needed. Sampling equipment is an area of rapid innovation and more portable and fully integrated eDNA sampling systems are now starting to become available (e.g. the ANDe eDNA sampling backpack; Thomas et al. 2018).

Note that as filtration pressure increases (whether using pump or syringe filtration), there is some evidence of reduced DNA retention on the filter membrane, as more molecules are forced through. However, this effect seems to be offset by the benefits of processing higher water volumes (Thomas et al. 2018).

2.5 Volume of water

The volume of water filtered in published studies ranges from as little as 15 ml to over 100 l, but the most common volumes are between 500 ml and 5 l. There is little consensus on the minimum viable filtration volume, which will depend to some extent on other factors, such as:

1. The sampling strategy employed, including number of sampling replicates and spatial representativeness of each sample.
2. Practical constraints, such as turbidity of the water, which causes filters to clog.
3. The size of the waterbody, which affects the concentration of eDNA in the water, and therefore detection probability (McClenaghan et al. 2020).

4. The aim of the monitoring and the management cost of failure to detect low-abundance species, which may be high in contexts such as biosecurity but less serious for establishing community metrics.
5. The efficiency of the downstream processes for DNA extraction and analysis.

Figure 4. Water eDNA collection and filtration methods. Water collection should cause minimum disruption to the substrate of the water body, either by collecting the sample from the bank (a and b) without entering the water or by collecting the sample upstream of the sampler (c and d). Pairing filtration or sampling in the field with an enclosed filter is recommended to minimise potential contamination. Filtration in the field can be carried out with a peristaltic pump (e) or by using a disposable syringe (f, g and h).

While for any given sampling system the volume of water filtered tends to correlate positively with the amount of DNA recovered and detection probability of rare species (McClenaghan et al. 2020), this is by no means a linear relationship, and community composition can be well recovered even when relatively small volumes of water are filtered (Mächler et al. 2016, Muha et al. 2019). It is likely that the sampling strategy - including the spatial representativeness of the filtered water - is of greater importance than volume *per se*, particularly in lentic systems (standing waters such as lakes or ponds), where DNA can be patchily distributed (Li et al. 2019). Note that in marine environments, where eDNA concentration is often extremely low, it is generally recommended to process larger volumes of water.

In general, as detection probabilities decrease due to environmental, physical or biological factors, sampling effort should be increased. This can be achieved through both increasing the number of samples collected and increasing the volume of water filtered per sample. Increasing the number of samples is often likely to be more effective in increasing detection probability, and has the added benefit of enabling assessment of frequency or occupancy of species across the replicates. It is also usually easier to achieve given that volume may be restricted by filter clogging, although there are some high-volume sampling systems that enable filtration of much larger volumes of water in a single sample (Cilleros et al. 2019, Schabacker et al. 2020).

2.6 eDNA preservation strategy for filters

eDNA on filters is preserved for transportation and storage by either freezing, drying, or adding liquid preservative to the filter.

- **Freezing** is effective but requires immediate access to cooling equipment and the ability to keep the samples cooled or frozen during transportation to the laboratory. Freezing may have a positive effect on eDNA recovery compared with samples that are extracted immediately after filtering, possibly due to a lysis effect resulting from cell bursting (Mauvisseau et al. 2021), but multiple freeze-thaw cycles should be avoided. This is likely to be most applicable to samples that have been filtered in the laboratory.
- **Drying** the filter membrane requires either silica gel, a desiccator, or paper that absorbs water. This is a less commonly used preservation approach, but is attractive in its ease and simplicity as it allows storage at room temperature for several weeks or even months (Thomas et al. 2019). It is challenging to completely dry the membrane within an enclosed, disc-shaped filter capsule, so alternative approaches are advisable if working with this type of filter.
- **Preservation liquids** can be broadly assigned to two different categories: pure preservatives such as ethanol and RNAlater, or lysis agents, including Longmire's buffer (Longmire et al. 1997, Wegleitner et al. 2015) and Sarkosyl buffer (Civade et al. 2016), which release DNA into solution at the same time as preventing degradation. It is critical to know whether or not the buffer you use is lysing the cells or organelle membranes, since this affects how you approach the early stages of the DNA extraction process in the laboratory. It is also important to consider if you will analyse the microbial portion of diversity collected in the sample. If microbes are targeted then it will be important to arrest microbial growth at the time of preservation, and not all preservative solutions will achieve this. The properties of four commonly-used preservative solutions are summarised in **Table 3** below. An advantage of using preservative solutions is that positive control DNA can be incorporated into the solution and used to check DNA preservation and extraction.

Table 3. Properties of four commonly-used preservative solutions, with recipes where applicable.

Preservative solution	Recipe	Lyses cells?	Kills microbes?	Preserves RNA?	Practical considerations
Ethanol	NA	No	Yes	No	Flammable liquid subject to dangerous goods transport regulations (UN1170, Class 3, packing group II) Can inhibit downstream reactions if samples are not completely dried prior to DNA extraction
RNAlater	25 mM Sodium Citrate, 10 mM EDTA, 70 g ammonium sulfate/100 ml solution, pH 5.2	No	Yes	Yes	DNA extraction can be challenging. Requires specific optimisation
Longmire's buffer	0.1 M Tris-HCL at pH 8.0, 0.1 M EDTA, 0.1 M NaCl, 0.5% w/v SDS	Yes	aided with addition of (hazardous) sodium azide	No	Precipitates at low temperatures (< 10 °C), but will return to solution if warmed
Sarkosyl buffer	100 mM Tris, 100 mM EDTA, 10 mM NaCl, 1% sodium N-lauroylsarcosinate	Yes	Yes	No	Does not precipitate at low temperatures, making it an attractive alternative to Longmire's
LifeGuard™ Soil Preservation Solution	Commercially purchased from Qiagen	No	Yes	Yes	Mostly used for small-volume sediment samples (< 1g). Cost is likely to be prohibitive for larger-volume samples

Exogenous internal positive control (IPC) DNA can be added to the sample to check that DNA has been adequately preserved. This DNA, which can be purchased commercially along with the qPCR primers and probes needed for amplification, can be added, in a well defined concentration to the filter capsule shortly after filtering. If a liquid preservative is used then the IPC can be pre-mixed into the preservative and efficiently added this way. Note that IPC added to the water prior to filtering may not be captured effectively in the filter since it is not in the same state as eDNA (predominantly cellular, subcellular or particle bound).

After DNA extraction, IPC concentration can be quantified using qPCR or ddPCR to check that it is recovered at the expected concentration. Testing should be carried out using the specific sampling and DNA extraction method to be employed, to ensure IPC recovery, and to determine the concentration to be added to the sample and the expected results in the absence of DNA degradation and inhibition.

We recommend that commercial IPCs are used where possible, and this should always be the case for analyses carried out in a regulatory or management context. If custom controls are to be used in a research setting, the DNA used as IPC should be completely absent in the study system, and should not interfere with the DNA of target organisms and/or downstream applications. If the IPC is designed to be

analysed alongside the target group in a metabarcoding analysis, it should be ensured that there are no primer mismatches, as this could reduce efficiency of recovery in some circumstances.

2.7 eDNA extraction from filters

eDNA extraction protocols from filters can be based on a number of different commercial DNA extraction kits (e.g. Qiagen DNeasy PowerWater or PowerSoil kits, Macherey Nagel NucleoMag Water kits), custom column-based methods (Sellers et al. 2018), or liquid phase methods (Renshaw et al. 2015, Deiner et al. 2018). For routine monitoring applications in regulatory or industry contexts, the use of reagents from commercial DNA extraction kits is generally advised, since they are expected to be standardised and certified as DNA-free. Protocols such as liquid phase extractions with phenol-chloroform-isoamyl are effective and produce greater DNA yields (Deiner et al. 2015), but they carry significant health and safety concerns that most commercial laboratories will seek to avoid.

Note that this is not to say that commercial kits are without health and safety concerns and must be handled appropriately and need proper disposal of waste. For example, the most widely used kits contain guanidine thiocyanate or guanidine hydrochloride which are reactive with sodium hypochlorite (i.e., bleach) to produce chloramines, chlorine and hydrogen cyanide gases. Due to the common use of bleach as a decontamination agent in most laboratories, this presents a major health risk if laboratory personnel are not properly trained to avoid contact between commercial extraction kit liquids and bleach (e.g. they mistakenly wipe the bench top with bleach first instead of with a mild detergent after an extraction procedure).

Initial steps in extracting DNA from filters must be optimised according to the type of filter and preservation strategy used and the biological targets for analysis.

First, the target group(s) must be considered in the selection of a lysis method. Chemical lysis is sufficient for extraction of animal DNA, but mechanical lysis is required for disrupting cell walls of some unicellular groups such as diatoms. Since mechanical lysis cannot easily be applied to enclosed filters, open or housed filters are recommended if you plan to target such groups in your samples, and a DNA extraction that includes bead beating (a type of mechanical lysis using beads to break down cell walls) or equivalent is required.

Second, the preservative solution and filter type will influence the lysis procedure. It is vital to know whether the storage solution has lysed the cells.

- Pure preservatives such ethanol and RNAlater can be discarded, leaving the material containing DNA on the filter membrane ready for subsequent lysis (Spens et al. 2017).
- Ethanol should be completely evaporated otherwise this can cause inhibition of later stages in the extraction. To maximise DNA recovery from samples contained in ethanol it is possible to carry out a precipitation step on the ethanol in addition to extraction from the filter membrane, although this will increase costs.
- Lysis buffers used for preservation (e.g. Longmire's or Sarkosyl buffer) cause much of the DNA to be in solution by the time extraction begins in the lab. In this case, discarding the storage solution will result in catastrophic loss of DNA. Instead, the filter should be incubated in the storage solution, which is then used as lysate for the next stages of the DNA extraction. Note that some DNA extraction kits require particular salt concentrations in the lysis buffer, so it is important to check compatibility between the buffer solution and kit chosen.

An additional consideration is that not all extraction methods will release DNA bound to particles such as clay. If it is suspected that much of the DNA in a sample is particle bound (e.g. in highly turbid waters),

then using kits optimized for soil extraction (e.g. Qiagen PowerSoil) or lysis buffers containing trisodium phosphate are needed to release adsorbed DNA (Ogram et al. 1987, Sellers et al. 2018). If not, then adsorbed DNA will not be extracted.

Organic compounds co-extracted from the sample along with the DNA may inhibit downstream PCR reactions. This particularly affects samples from turbid waters and small water bodies containing lots of rotting leaves, which introduce tannins and other dissolved organic compounds to the water. Other sources of organic material that may cause inhibition include faeces from livestock (e.g. cattle; Wilson 1997, Rapp 2010). Extraction kits designed for use on water, soil or faecal samples often incorporate inhibitor removal as part of the standard process. Others - including those designed for use of pure biological samples of tissue or blood - do not, and a separate step may be required to remove inhibitors post-extraction using kits such as the Zymo OneStep PCR inhibitor removal kit or Qiagen PowerClean kit, or with custom protocols (Abbaszadegan et al. 1993). Inhibition can also sometimes be overcome by dilution of the DNA extract. This reduces the concentration of inhibitors but also reduces the concentration of target DNA, which can affect detection probability. A common approach is to run a dilution series to determine the minimum level of dilution that can counteract the inhibition. Note that the use of clean-up kits also results in some DNA loss, so these measures should only be applied to eDNA samples that are affected by inhibition.

Extraction efficiency and the presence of inhibitors can be assessed by including internal positive control DNA in the lysis buffer and checking via qPCR that it is recovered in the expected quantity after DNA extraction. This is discussed in more detail in Section 6.2 below.

Summary: Capture, preservation and extraction of eDNA from water

Key takeaways:

- There are multiple methodological variations and trade-offs, and choice of method will often be driven by the goals, logistics, and practicalities of the monitoring programme, rather than by the difference in quality of data produced. Research has shown many methods to be effective at detection of species.
- As for any sampling method, eDNA sampling requires good ecological survey design to ensure the data will meet the needs of the project. However, it also needs to account for how eDNA moves within the environment once it has been released from the organisms, and how this affects inference of species presence.
- DNA yield is affected by multiple interconnected factors, including filter material, pore size, preservation strategy, and DNA extraction method. To ensure the best quality of results, any given workflow needs to be optimised for the specific choices made, and the optimisation process should be repeated if any aspect of the capture and extraction protocol is changed.
- Many options exist for storing and preserving eDNA after filtration, but storage in a preservation buffer has been shown to be practical and robust. Choice of preservation buffer affects the first steps of the DNA extraction process, so these aspects must always be considered together when designing a new workflow. In particular, it is essential to know whether the preservation solution also functions as a lysis buffer in order to avoid catastrophic loss of DNA from mistakenly discarding the solution.
- DNA extraction must consider the environmental context from which the DNA was sampled due to co-extraction of inhibitors, which can reduce the efficiency of DNA amplification.

Priorities for future research:

- Modelling the distribution and dynamics of eDNA in different environments to inform optimal sampling design, especially in large water bodies such as oceans.
- Improved understanding of the state of eDNA in different environments and how this affects the performance of different eDNA capture and extraction methods.
- Increased understanding in the variance among paired capture, preservation and extraction methods in their performance across different ecosystem types.

3. PRESERVATION AND EXTRACTION OF MACROINVERTEBRATE BULK SAMPLES

Well-established methods already exist for capturing aquatic macroinvertebrates and terrestrial arthropods by passive trapping (e.g. Malaise traps for flying insects, pitfall traps for ground invertebrates) or active sampling (e.g. kick sampling for benthic macroinvertebrates in streams). By-and-large the use of molecular methods for species identification does not demand development of new sampling approaches for these groups - although it should be noted that methods that yield clean samples with minimal detritus are highly preferred - so we do not go into detail about sampling methods here.

3.1 Preservation methods

The first point at which DNA- and morphology-based assessment processes for bulk samples differ is in the preservation of bulk samples for storage and transportation. However, there is currently no clear consensus on the ideal preservation and storage strategy. Fixation in formaldehyde, which is commonly used to preserve bulk samples for morphological processing, is incompatible with DNA-based analyses and should be avoided at all cost.

Research studies commonly use ethanol for sample preservation. Ethanol works as a preservative by replacing water in biological tissues, but water drawn out from the tissues serves to dilute the ethanol, reducing its effectiveness for long-term storage. The dilution effect decreases as the ethanol to sample volume ratio increases so it is important to add ethanol in at least double the sample volume and it is standard practice to replace the ethanol for long-term storage of samples. However, while ethanol is certainly an effective preservative, it also poses considerable difficulties for use in routine biomonitoring. Pure, undenatured ethanol can be expensive to purchase in countries where alcohol duties are applied, requires special storage conditions, and is difficult to transport because it is a flammable liquid that is subject to dangerous goods regulations. This also raises health and safety concerns for many organisations that carry out fieldwork. Moreover, the large volumes of ethanol required for preserving bulk macroinvertebrate samples are expensive to dispose of correctly, and the sample needs to be completely dried prior to DNA extraction because ethanol residues interfere with the extraction chemistry.

The cost and accessibility barrier can be overcome through the use of denatured alcohols such as industrial denatured alcohol (IDA), industrial methylated spirits (IMS), and isopropanol. However, some denatured alcohols (e.g. IMS) seem to be unreliable as preservatives for DNA because of their capacity to gradually degrade dsDNA (Carter et al. 1997). They may also act as inhibitors in downstream steps. Moreover, denatured alcohols vary substantially in their chemical composition and to date there has been a lack of systematic testing that would enable general principles to be outlined regarding the suitability

of use for DNA preservation. Like ethanol, they are also highly flammable and subject to regulations for transport, storage and disposal.

Non-flammable preservation solutions include:

- Propylene glycol (Robinson et al. 2021, Weigand et al. 2021), which can be purchased in large volumes and at low cost as an antifreeze. Note that it can be viscous in some forms, which complicates laboratory processing.
- RNAlater, which is prohibitively expensive to purchase commercially at the volumes needed for preserving bulk invertebrate samples but potentially represents a good option if it can be made up in the laboratory (see **Table 3** for details). It is non-hazardous, making it more logistically viable than ethanol in many cases. Testing is still required to demonstrate use of RNAlater for preservation of this sample type.

Note that for collection of terrestrial invertebrates using passive trapping methods, the fluid serves as a collection and killing agent as well as a preservative. In some types of traps (e.g. Malaise traps), ethanol currently remains the preferred solution for this reason, and any alternative should be tested to check that it does not reduce trapping efficiency. If ethanol is used for collection, it can subsequently be filtered off for dry transportation of samples if necessary, but will still need to be disposed of correctly.

An alternative approach is to store the sample in a lysis buffer, which effectively maintains DNA integrity during storage. These solutions are usually non-hazardous and easy to transport. There are commercially available buffers as well as those that can be made up in the laboratory, such as Longmire's Solution (Longmire et al. 1997, Wegleitner et al. 2015). However, note that soft-bodied organisms will disintegrate if stored in lysis buffers, making it impossible to morphologically confirm species identity or archive voucher specimens.

Non-liquid-based preservation strategies include immediate freezing or crushing of organisms to isolate mitochondria (Macher et al. 2018a). These strategies require in-field equipment that may not always be readily available or easily transported to site, so they do not currently represent realistic options for standard protocols. A more practical option may be to dry the samples at 40-50°C as soon as possible after collection, as this will stabilise the DNA for short- or medium-term storage and transportation at the same time as reducing the weight and volume of the sample. This requires further testing to fully understand its effectiveness in preserving DNA, and how it affects downstream processes. For instance, when a large amount of debris is collected along with the organisms it may be harder to separate this from a dried sample.

3.2 Homogenisation methods

DNA extraction from bulk samples typically starts with homogenisation, which can be achieved either using bead beating or blending (**Figure 5**). Organisms must be ground extremely finely to facilitate unbiased extraction. Specimens can be homogenised either dry or in a preservation solution e.g. (Pereira-da-Conceicoa et al. 2021), but this will depend on your choice of preservation strategy and homogenisation equipment. This can also happen in ethanol, but fire risk should be considered if grinding samples in ethanol as significant heat can be created, particularly if samples contain stones and other hard objects, and ethanol has a very low flash-point of 37°C.

A practical challenge to the large-scale operationalisation of this process is that bulk samples are often of considerable volume - much larger than can be accommodated in most tissue homogenisers or bead mills, and especially those that can process multiple samples in parallel. Therefore, large samples

Figure 5. DNA homogenisation methods. DNA sample homogenisation protocols vary depending on equipment available, here we present a number of lab set-ups which show differing equipment for the homogenisation of tissue samples (dry and wet) , IKA Tube-Mill 100 (a) with blending blades (b), dry malaise trap sample before and after homogenisation with the IKA Tube-Mill 100 (c and d), Qiagen TissueLyser II with plate attachment (bead mill also available) (e), and stainless steel wet-grinding blender (f and g).

must either be homogenised one-by-one in a large-volume blender (which requires decontamination between samples) or split into smaller subsamples for this stage, which introduces additional consumables cost and reduces the capacity for parallel processing of large numbers of samples.

Moreover, it is usually necessary to clean the sample prior to homogenisation, removing organic and inorganic detritus (and organisms with thick calcium carbonate shells) to reduce the volume of the sample prior to homogenisation and improve homogenisation efficiency (although see See Pereira-da-Conceicoa et al. 2021 and Buchner et al. 2021 who successfully blended samples without cleaning). This also removes potential sources of inhibition that could affect PCR efficiency downstream. Like with size-sorting, this manual sample processing step reduces overall time and cost efficiency of the DNA-based approach, while also risking the loss of small organisms if due care is not taken to retain them.

Lastly, variation in body size presents a major challenge when working with bulk samples. This is particularly relevant to benthic macroinvertebrates, which vary in biomass by orders of magnitude. If DNA is extracted from bulk samples without size sorting, this leads to the DNA of very small organisms

being overwhelmed by that of larger ones, resulting in detection bias weighted towards large-bodied taxa. Size sorting the organisms prior to extraction helps to fully recover the diversity of a sample, but it also requires a significant time investment as well as multiple DNA extractions per sample, which adds substantially to the cost, so there is a trade-off between taxonomic completeness and processing time/cost. Different sorting schemes have been proposed, ranging from very strict (separate into multiple size classes, extract DNA from each separately, and then pool the DNA extracts proportionally) to very coarse (remove large individuals, leaving only a leg or other body part in the bulk sample) (Elbrecht et al. 2017a, 2021) or even sorting according to taxonomic group (Moriniere et al. 2016, Beentjes et al. 2019). An alternative to size sorting is to greatly increase sequencing depth during metabarcoding to improve recovery of small organisms. These options carry significant costs, which may in some cases negate the economic advantage of employing a molecular approach over a morphological one.

3.3 Non-destructive methods for macroinvertebrate bulk samples

As discussed above, various aspects of the sample-handling and homogenisation process limit the throughput of the DNA-based approach for bulk samples (specifically the time taken in drying, cleaning and size-sorting the samples, and capacity issues introduced by the need to divide up large samples). Furthermore, a potential barrier to the uptake of DNA-based methods for macroinvertebrate monitoring under the Water Framework Directive is the requirement to retain voucher specimens (a preserved specimen that serves as a verifiable and permanent record), which is incompatible with sample homogenisation.

An alternative to homogenisation is to retrieve DNA non-destructively by extracting it from the liquid solution in which the sample has been stored. Although frequently used for DNA extraction from individual museum specimens, this has only recently been widely applied to bulk samples in metabarcoding studies (e.g., Hajibabaei et al. 2012, Linard et al. 2016, Carew et al. 2018, Erdozain et al. 2019, Ji et al. 2019, Martins et al. 2019, Zizka et al. 2019b, Zenker et al. 2020).

Many tests of this approach (e.g. Hajibabaei et al. 2012, Martins et al. 2019, Zenker et al. 2020) have focused on retrieving DNA from samples stored in ethanol, with varied levels of success. More recent studies have worked with samples stored in a lysis buffer, with extremely promising results (Ji et al. 2019, Nielsen et al. 2019). However, different types of organisms have been observed to contribute DNA into the preservation solution at different rates, with soft bodied organisms shedding DNA more readily than those with an exoskeleton or operculum (Carew et al. 2018, Zizka et al. 2019b). This, together with differential amplification efficiencies, will mean that quantitative information will almost certainly be lost during the process and data should be analysed based on presence-absence only. The concentration of DNA extracted from a storage solution will also be lower than that extracted directly from homogenised tissues and it will be necessary to account for amplification stochasticity and increased vulnerability to contamination.

Since this method is in the early stages of development, a number of fundamental questions remain regarding the relative effectiveness of different buffer solutions and how best to maximise release of DNA from the organisms into the solution without completely destroying the specimens (if they are to be retained as voucher specimens or for subsequent morphological identification). For instance, it may be possible to greatly increase extraction efficiency through agitation of the sample or the addition of proteinases, but this will destroy soft-bodied organisms. It is still unclear how long a sample should be left in the preservative solution and what volume of solution should be used for DNA extraction (Martins et al. 2019, Zizka et al. 2019b). This approach therefore requires further research attention, but has the potential to increase efficiency and facilitate high-throughput monitoring of macroinvertebrate communities.

3.4 Considerations for bulk DNA extraction

A wide variety of different commercial kits or liquid phase extraction methods are used for DNA extraction from homogenised macroinvertebrate samples. Thus, there are no clear recommendations as to which kit or method is preferable. However, due to health and safety concerns of working with toxic chemicals such as Phenol, commercial kit-based extraction approaches are generally preferred in commercial or routine monitoring laboratories. For DNA extraction from lysis buffer solutions, kits that enable a large volume of lysate to be processed may be preferred (e.g. QiAamp DNA Blood Maxi Kit, which uses 3-10 ml of lysate; Fonseca et al. 2011).

Early stages of the DNA extraction process are influenced by choice of preservation strategy, since some solutions, including ethanol, can interact with the chemistry during initial extraction stages, causing reduced extraction efficiency. Thus, samples preserved in ethanol must be fully dried prior to lysis, and this can pose a cross-contamination risk as electrostatic charges can cause small dried fragments to jump considerable distances. Using a lysis buffer (without SDS or other surfactants) or wet grinding in ethanol (Buchner et al. 2021) for sample preservation avoids this issue and can streamline the extraction process.

The volume of material used for extraction is also an important consideration. If the sample has been homogenised, using too much of the sample can cause PCR efficiency to be reduced through inhibition. Assuming that homogenisation has been sufficiently complete, representative communities can be described from as little as 0.3 g of the total homogenate (Hajibabaei et al. 2019). If DNA is extracted from a storage solution, there is a decision to be made regarding the volume of solution that is processed in the extraction. Using a larger volume may increase detection of rare species, but will significantly slow down the extraction process when large numbers of samples are being analysed. Further testing is required to optimise this.

Summary: Sampling & DNA extraction for Macroinvertebrate bulk samples

Key takeaways:

- Capture of insects and aquatic macroinvertebrates broadly follows established methods, but the samples must be stored in a solution that effectively preserves DNA. Sample volume often represents a practical challenge, both in terms of the volume of preservative required (which introduces cost and logistical constraints) and in terms of the machinery needed to homogenise the samples.
- The most important trade-off is choice of preservation solution and its compatibility with downstream homogenization and extraction methods. If ethanol is used, this must be removed by drying to not inhibit downstream methods.
- When optimizing for time, the choice to blend the whole sample versus size sorting and debris removal has important consideration for detection of rare or small-bodied species. This can be partially mitigated by greater sequencing depth. Both size sorting and increasing sequence depth are associated with increased costs.
- Extraction of DNA direct from storage solution represents a rapidly-developing area of research with significant potential for accelerating and scaling up the processing of bulk invertebrate samples.

Research priorities:

- Further systematic testing of non-ethanol preservative liquids for storage of bulk invertebrate samples.

- Systematic testing of different options for extracting DNA from storage solutions, including the use of different buffer solutions, the ratio of sample to solution volume, varying levels of sample agitation, optimal incubation time, and volume of solution to be used as input for DNA extraction.

4. SAMPLING, PRESERVATION AND EXTRACTION OF BENTHIC PERIPHYTIC DIATOMS

Freshwater periphytic samples (a.k.a. benthic biofilm) are characterized by the presence of a wide diversity of organisms representing all the tree of life domains (Battin et al. 2016). This includes both microorganisms (bacteria, archaea, micro-eukaryotes) and macro-organisms (macroinvertebrates). This organismal DNA is also supplemented by extra-organismal DNA that settles out from the water column and becomes trapped in the biofilm. Thus, biofilm samples can convey a lot of information about a wide diversity of organisms, although note that the spatial and temporal representativeness of information derived from extra-organismal DNA in aquatic biofilms is not clearly understood (e.g. Shogren et al. 2018, Harrison et al. 2019).

Benthic diatoms are one of the major components of biofilm diversity in aquatic ecosystems and are commonly used for morphology-based ecological assessment (e.g. Battin et al. 2016, Morin et al. 2016). The majority of diatom DNA extracted from biofilm comes from living cells. This enables the characterization of community structure, and the detection and relative quantification of taxa, which are the key parameters used to compute diatom water quality indices. It also means that a high degree of spatial and temporal precision can be inferred during interpretation of the results.

4.1 Sampling benthic biofilms

DNA-based assessment of benthic diatom communities uses the same sampling methodology as morphology-based assessments applied for monitoring under the WFD, for which standard protocols have been set (NF EN 13946 - April 2014). These documents describe how to collect biofilm samples (number and type of substrate collected, habitat, biofilm surface collected), while a Technical Report from the European Committee for Standardisation (CEN/TR 17245, 2018) provides recommendations to ensure that biofilm samples are collected and stored in such a way as to be compatible with molecular analysis (**Figure 6**).

This sampling strategy described in the standards was statistically optimized to obtain the best ecological assessment of evaluated rivers or lakes for minimal sampling effort. In the scope of routine monitoring and WFD requirements, this method is easy to apply and the same sample can be used for both DNA- and morphology-based assessment.

Diatom community DNA is dominant in the samples, so the vulnerability to contamination is low, which means that precautions can be lighter than for aquatic eDNA sampling: gloves are not mandatory, and collection of substrates generally requires entering the water (walking against the water flow), see the technical report (CEN/TR 17245, 2018) for more details. However, if biofilm samples are used to characterise prokaryotic communities or for analyses of extra-organismal DNA (e.g. from fish), considerations described in section 1.3 should be carefully applied to limit contamination.

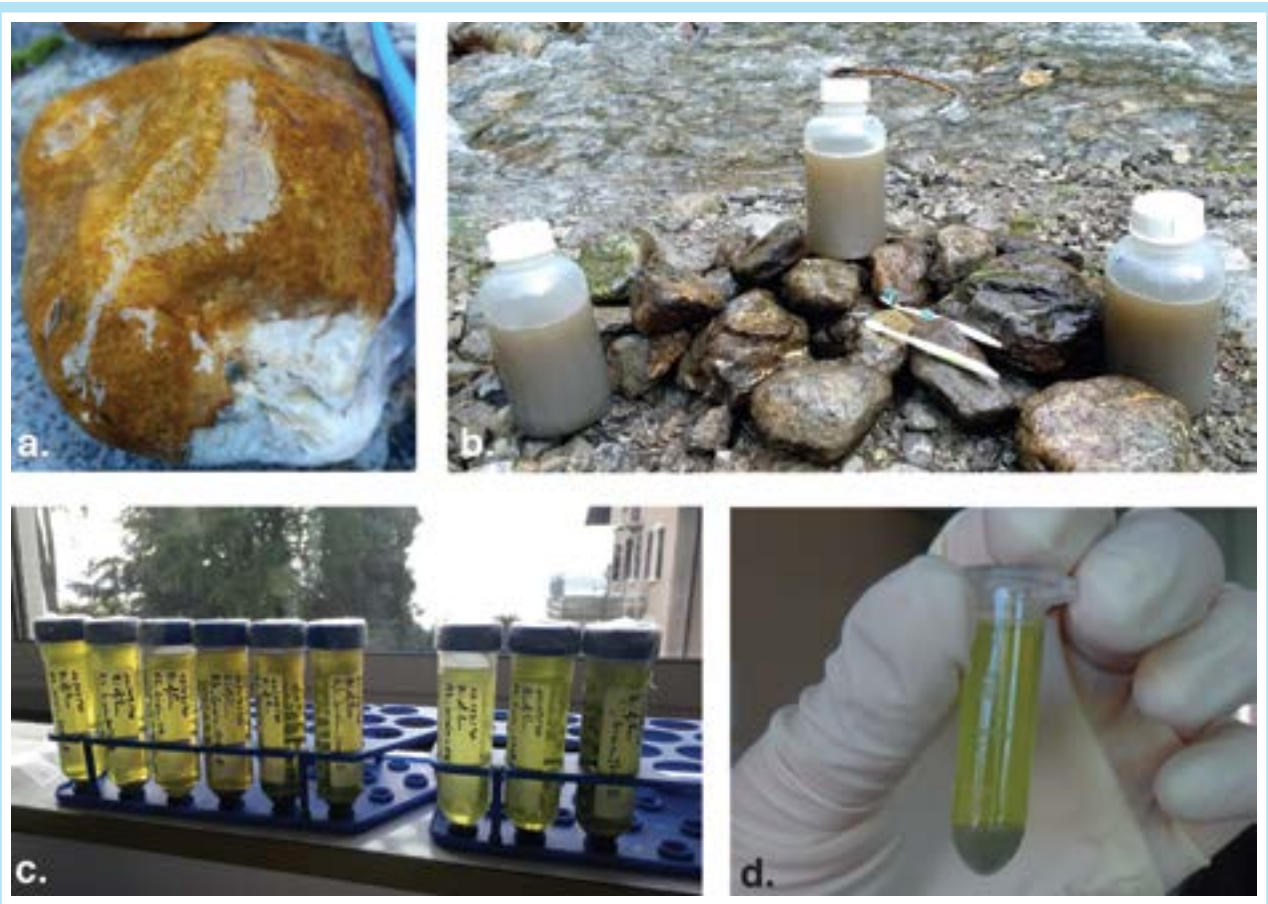

Figure 6. Biofilm sampling. Biofilm samples are collected from submerged stones (a) in a waterbody using toothbrush (b). Diatom sub-samples are then collected from larger samples (c.) and then pelleted using centrifugation prior for DNA extraction (d).

4.2 Preservation methods

Historically, biofilm samples were preserved using formaldehyde or Lugol solutions. These are compatible with morphological processing of samples, since they preserve the silica walls surrounding the diatom cells (frustules), which are used for identification with light microscopy. Formaldehyde solution is known to be incompatible with DNA preservation, and the use of Lugol solution for such application is still under debate (Mäki et al. 2017).

The CEN technical report (CEN/TR 17245, 2018) recommends the use of pure and undenatured ethanol solution with a final concentration >70% to preserve biofilm samples collected in lakes and rivers. Several studies have applied this protocol successfully for benthic diatom ecological assessment in lakes and rivers using DNA metabarcoding (Vasselon et al. 2017a, Rivera et al. 2018, Bailet et al. 2019, Mortágua et al. 2019, Apothéloz-Perret-Gentil et al. 2020, Pérez-Burillo et al. 2020, Rivera et al. 2020). This choice is mainly driven by the fact that ethanol preservation remains easy to apply and enables the use of the same biofilm sample for both morphology- and DNA-based approaches. However, further experiments are required to evaluate the stability of ethanol preserved samples over time and the effect of different storage conditions on DNA preservation.

Alternative preservation approaches that have been used for benthic diatom samples include RNA-later (Hamilton et al. 2015, Kelly et al. 2018) and direct freezing (Apothéloz-Perret-Gentil et al. 2017, Mora et al. 2019). An ongoing experiment, conducted as part of the DNAqua-Net COST Action, is

evaluating the performance of these 2 approaches in comparison to ethanol over a period of one year. Results will help to update guidelines and recommendations.

4.3 Considerations for DNA extraction from benthic biofilms

Growth of aquatic biofilms can be affected by a wide variety of environmental factors including the nature of the substratum (Mora-Gómez et al. 2016). This means that biofilms collected with the same protocol may vary in physicochemical properties and biological composition, resulting in different structure and biomass development. Thus, biofilm samples are usually homogenized using a vortex and subsampled in smaller aliquots using 1 - 2 ml of preserved biofilm. This provides enough biological material to characterise diatom community structure without reaching maximum starting sample amount for most DNA extraction protocols. For biofilm samples with low biomass, higher volumes could be used to increase starting material amount.

The subsampled biofilm aliquot is pelleted using centrifugation (a separation process that relies on the action of centrifugal force to separate particles in a solid–liquid mixture) and the supernatant (i.e. the preservative solution) is discarded.

Like other environmental samples, biofilm samples are characterized by the presence of organic matter, humic acids and polyphenols that are known to inhibit molecular methods like PCR, so this needs to be accounted for during or after the DNA extraction process. In addition, diatom cells are protected by a frustule made of silica that is hard to break, and this can reduce the efficiency of DNA extraction if the cell lysis step is insufficient (Hamm et al. 2003).

A wide variety of DNA extraction methods from benthic biofilms are found in the literature (e.g. Apothéloz-Perret-Gentil et al. 2017, Vasselon et al. 2017a, Mora et al. 2019, Rivera et al. 2020, Kelly et al. 2018). Studies have compared the efficiency of different elements of these extraction protocols (Nguyen et al. 2011, Vasselon et al. 2017b), and shown that although extraction method does affect the relative quantification of diatom taxa obtained after metabarcoding, results were similar in terms of ecological quality scores based on diatom indexes.

Summary: Capture and DNA extraction for benthic periphytic diatom samples

Key takeaways:

- Diatom metabarcoding is applied to organismal samples (collections of diatoms) that are collected following the same methods as used for conventional morpho-taxonomic assessment using light microscopy.
- Diatoms have strong cell walls that may require mechanical lysis to disrupt for effective DNA extraction.
- Periphytic samples are characterised by the presence of organic and inorganic compounds that might be co-extracted with DNA (e.g. humic acids, carbonates). These can react with chemicals used in DNA extraction kits or act as inhibitors for downstream molecular processes like PCR amplification. For complex periphytic samples, particular attention should be paid to the choice of DNA extraction and DNA purification methods.
- Optimal sample volume for DNA extraction is 1-2 ml of preserved sample. Larger volumes can be used for samples with less developed biofilms and small amounts of biological material.

Research priorities:

- Further systematic testing of non-ethanol preservation methods, including lugol, lysis buffers, RNAlater and freezing.
- Evaluate the stability of preserved samples over the time in order to propose guidelines for storage conditions (temperature, duration, light).
- Understanding the performance of mechanical vs chemical lysis of samples.
- Temporal variation of benthic diatom communities is not monitored, sampling strategies should be adjusted according to season and flow velocity.

5. SAMPLING, PRESERVATION AND EXTRACTION OF SOILS AND SEDIMENTS

Soil and aquatic sediment samples typically contain a high diversity of living organisms (organismal DNA) as well as DNA from larger organisms that has been shed into the sediment (extra-organismal DNA) and DNA from dead or dormant organisms. In aquatic systems, surface sediments also contain DNA from pelagic organisms or their cells that have settled from the water column. This makes soil and sediment samples a rich source of data across the entire spectrum of biodiversity (Weigand and Macher 2018) but can also complicate the interpretation of results.

Sediment metabarcoding that targets Metazoa will tend to be dominated by organismal DNA, which is present in much higher concentrations than environmental DNA (extra-organismal). This means that the datasets will predominantly comprise meiofaunal taxa rather than macrofauna. This is advantageous from the point of view of maximising the statistical power to show community change in response to impact or land use change, since meiofauna are often more diverse and more abundant than macrofauna, but it makes it difficult to directly compare metabarcoding results with those obtained from conventional surveys of benthic macrofauna or ground insects.

Molecular analysis of sediment samples is complicated by the fact that complex organic molecules and inorganic particles are able to bind, adsorb, and stabilize free DNA in sediments (Romanowski et al. 1991). Consequently, the residence time of extracellular DNA in soils and sediments can be much longer than in water (Corinaldesi et al. 2005), meaning that detections may not always indicate contemporary presence of the species. This does not necessarily pose a problem for projects focusing on community ecology, since a background signal of historic DNA does not obscure temporal trends and environmental responses in community composition (Brandt et al. 2020). However, it presents a greater risk in contexts such as biosecurity, when detection of even trace amounts of a particular species' DNA may trigger expensive management responses.

5.1 Collecting the sample

5.1.1 Sample volume

The volume of soil or sediment collected per sample will usually be decided according to the portion of biodiversity that is targeted and the spatial scale at which it operates. For microorganisms such as bacteria and single-celled eukaryotes, it is common to collect only very small-volume samples, which can be as small as 0.25 g and usually not more than 1 g. The advantage of such small-volume samples

is that they can be easily and cheaply preserved and are compatible with high-throughput (i.e. automated) DNA extraction systems. If larger-bodied organisms belonging to the meiofauna or macrofauna are targeted, larger volumes (> 10 g) are recommended in order to achieve a representative sample (Dopheide et al. 2019, Brandt et al. 2020), and these are somewhat more challenging and expensive to handle. The maximum volume of sediment that can be extracted using commercial kits is 10 g (Qiagen PowerMax Soil kit).

Since DNA does not mix well in sediments, it is usually necessary to collect subsamples from across the area that the sample aims to represent. Sediment subsamples are most commonly collected using either a spoon/spatula (ideal for targeting a shallow layer of surface sediment) or a small coring device such as can be fashioned from a syringe. Syringe corers are ideal for targeting a deeper sediment profile, and the plunger on the syringe can be used to create suction enabling cores of loose or wet sediment to be collected. Subsamples can either be treated separately to maximise statistical power or merged and re-sampled to give a single smaller-volume sample that is representative of a wider area of sediment. The latter approach is more cost effective.

Very large volumes of sediment have been used in some studies (1 l or more) to target extracellular or extra-organismal DNA, which is present in very low concentrations in both terrestrial soils and marine sediments (e.g. Taberlet et al. 2012b, Guardiola et al. 2015, Leempoel et al. 2020).

5.1.2 Sampling depth

Samples are usually collected from the surface of the soil or sediment. If a large primary corer (e.g., grab sampler or boxcorer) is used to recover sediment from deep water, subsamples should be taken away from the edges of the corer, targeting the minimally disturbed parts of the sediment that have not come into direct contact with the equipment. There is no clear consensus as to the ideal vertical depth of the sample, and this may vary between target groups and ecosystems. For instance:

- Environmental responses of microbial communities to aquatic pollution are likely to be concentrated in the surface layer of the sediment, so this layer is commonly targeted for microbial assessments in aquatic sediments.
- Larger organisms (meiofauna and macrofauna) are more mobile and operate at larger spatial scales in both the horizontal and vertical planes. Therefore, is it common (though by no means universal) to sample to depths of around 5 cm for marine benthic meiofauna.
- In aquatic sediments, there is an argument for discarding the very surface layer which may contain DNA of more transient species or those that have settled from the water column and are not truly representative of the benthic community. There is also an argument for avoiding sampling from the anoxic sediment layer that sits beneath the oxygenated layer at the surface, since this is a much more inhospitable environment. The anoxic layer can be recognised by its darker colouration and foul smell, and can feasibly be discarded from a transparent corer where the sediment profile has been well retained.
- In terrestrial soils, the vertical distribution of biodiversity is often more complex, partly because of the structure provided by plant roots. Therefore, it is more common to sample at greater depths (e.g. Arribas et al. 2016, Treonis et al. 2018), and more research is needed to determine the best way to standardise a sampling strategy for terrestrial soils that balances ease of collection with the structural complexities of this environment.

More research is needed to establish the optimal depth of core samples for targeting different portions of sediment biodiversity, and this may vary depending on the environmental, physical and chemical characteristics of the sediment.

5.2 Separation of organisms from soils and sediments

One way to deal with large volumes of soil or sediment samples is to separate the organisms (macrofauna and/or meiofauna) from the soil or sediment itself, which is usually achieved through a series of flotation, decanting and sieving steps (Creer et al. 2010, Fonseca et al. 2011, Creer et al. 2016, Haenel et al. 2017). The organisms can then be processed as bulk samples, as described above. If this approach is taken, it is important to consider the risk of cross-contamination from the equipment used to clean and sort the sample. It has also been shown that the method used for separating the organisms from the sediment significantly affects community composition of the sample (Haenel et al. 2017), so whatever method is chosen should be maintained throughout the monitoring programme.

Separating the organisms from the soil or sediment allows a larger volume of sediment to be processed, but this needs to be balanced against the consideration that it is a labour-intensive process, and this may limit the scope of monitoring programmes in terms of the number of samples that can be handled.

Thus, extracting DNA directly from the soil or sediment itself is preferable in many ways: the process is more readily standardised and scalable, and requires less handling of the sample, which reduces contamination risk. DNA can be directly extracted from soil or sediment in volumes of up to 10 g using commercial DNA extraction kits (e.g. DNeasy PowerMax Soil Kit). Further work is needed to establish the optimal combination of sample size and replication to account for spatial heterogeneity in sediment communities, although various studies have examined this in certain environments (Nascimento et al. 2018).

5.3 Preservation of sediment samples

Samples must be preserved for transportation to the laboratory and storage prior to DNA extraction. Rapid and effective preservation of soil/sediment samples is particularly important if either eDNA, RNA, or microorganisms are targeted. Common preservation strategies include freezing at -80°C, and the use of preservative solutions such as ethanol or Qiagen's LifeGuard Soil Preservation Solution.

Freezing is widely accepted to be an effective way of preserving samples (Domaizon et al. 2013, Agasild et al. 2018) but relies on having immediate access to the necessary equipment and the logistical capabilities to keep the samples frozen during transport to the lab. This may make it an impractical choice for large-scale monitoring programmes in which a range of different parties will be collecting samples, and not all may have ready access to freezers. Note that a recent paper suggests that microbial communities are preserved with less extreme cooling (Delavaux et al. 2020).

Preservation using liquid preservative solutions may represent a more practical option, but choice of preservative solution is critical.

- It is important to consider whether or not the microbial component of biodiversity is being targeted (or could be in the future), since not all preservative solutions will arrest the growth of microbial communities. If these continue to grow and develop after sampling, the community will cease to reflect that of the environment in which it was sampled.
- Analysis of RNA requires specific preservation solutions, such as LifeGuard™, which is very expensive for use with samples of the volume recommended for meiofaunal assessment.
- Ethanol is a good preservative, but for use with sediment samples (which have a high water content), it should ideally be changed once, and the final concentration should be > 80% for effective DNA preservation. As mentioned elsewhere, ethanol also presents logistical and safety challenges

when applied to routine or industrial use and interferes with the chemistry of the DNA extraction process if not completely removed from the sample.
- Salt-based lysis solutions such as DESS (Yoder et al. 2006, Tatangelo et al. 2014) are effective for preservation of metazoan DNA, but require the addition of sodium azide or similar chemicals to arrest growth of some microbial groups, and this brings health and safety concerns both in the field and in the laboratory. Further systematic testing of the effectiveness of different solutions for preserving communities of various groups of organisms in sediments is required.
- RNAlater is not recommended as a preservative for soil and sediment samples because storage in this medium has been shown to lead to changes in composition of microbial communities and loss of diversity (Nilsson et al. 2019, Delavaux et al. 2020).

5.4 DNA extraction from sediment samples

Initial steps in DNA extraction will depend on the volume of soil or sediment collected, the target group for analysis, and the preservation method used. If a preservation liquid such as ethanol or LifeGuard™ has been used for sample preservation then this must first be removed from the sample, usually by centrifugation and discarding the supernatant. Subsequent wash steps may be needed to ensure that all traces of preservative are removed since they may interfere with the chemistry of the extraction kit. This is especially important with ethanol. If a salt-based buffer such as DESS has been used to preserve DNA, then the sample can sometimes be introduced more directly into an extraction process (Fonseca et al. 2011, van der Loos and Nijland 2020).

The maximum volume of soil or sediment that can be directly extracted in commercial kits is currently 10 g. Where a larger volume has been collected, thorough mixing prior to subsampling for extraction will help to ensure that the extracted DNA is representative of the sample as a whole. It may be worth carrying out multiple extractions per sample in this case, at least until it is clear to what extent a single extraction is representative of the whole sample. Soils and sediments typically contain PCR inhibitors, and heavily polluted sediments are often associated with particularly high levels of inhibition. Therefore, most extraction protocols need to include an inhibitor removal step. This is incorporated into commercial kit protocols designed for soils and sediments, but a specific clean-up step will almost certainly need to be incorporated where custom extraction protocols are used (e.g. Sellers et al. 2018).

DNA from very large volume sediment samples are typically extracted using a phosphate buffer approach as outlined by (Taberlet et al. 2012b), although note that this specifically targets extracellular DNA.

Summary: Capture & DNA extraction for soils & sediments

Key takeaways:
- Soil and sediment samples contain a mixture of organismal and extra-organismal DNA, which can be of varying ages. It is not currently possible to separate the different sources and states of DNA.
- The surface of aquatic sediments represents a complex interface between the pelagic and benthic environments, containing DNA that originates from both.
- Optimal sample volume depends on the target group, with larger volumes of soil/sediment usually required for capturing a representative community of multi-cellular organisms (e.g. meiofauna) than for microbial communities.
- Preservation strategy needs to take into account sample volume, whether the analysis will target microbial communities, availability of freezers and the feasibility of keeping samples frozen from

the point of collection to arrival at the lab, and whether or not it is necessary to preserve RNA as well as DNA.
- Choice of preservation methods also affects the methods used at the beginning of the DNA extraction process. Ethanol will need to be removed from the sample prior to beginning extraction.
- Some organisms in the soil have hard cell walls or cuticles, which may require thorough mechanical lysis for efficient extraction of DNA.

Research priorities:
- Improved understanding of optimal sampling depth in different types of soils and sediments under different conditions and when targeting different components of biodiversity.
- Understanding seasonal effects on soil biodiversity to guide sampling strategies in multi-year monitoring programmes.
- Improved understanding of the effect of different sampling and subsampling strategies in yielding representative biological communities.
- Systematic testing of different cost-effective preservation methods for larger-volume (> 2 g) sediment samples. These should aim to achieve consistent performance without the need for flammable liquids or precise temperature control.
- Efficient methods for DNA extraction from large-volume sediment samples.

6. MEASURING QUALITY AND QUANTITY OF EXTRACTED DNA

Before starting the main analysis, it is common and recommended to carry out some preliminary tests to characterise the DNA that has been extracted. This includes DNA quantification and testing for inhibition.

6.1 DNA quantification

The most commonly-used platforms for DNA quantification are the Qubit fluorometer (Invitrogen, Carlsbad CA) and the Nanodrop spectrophotometer (Thermo Scientific, Waltham MA).

- Nanodrop concentration usually gives an inflated estimate of DNA concentration because it measures even single nucleotides, so it is not recommended for this purpose. However, the 260/280 ratio given by the Nanodrop can be useful for indicating RNA and protein contamination (values between 1.8 and 2. denote pure DNA; below 1.8 indicates phenol, salt, protein or polysaccharide contamination; Olson and Morrow 2012).
- The Qubit uses an intercalating dye that binds directly to DNA and thus measures the quantity of double or single stranded DNA precisely even at very low concentrations, although it provides no estimate of DNA purity. Users should be aware that readings from the Qubit are affected by temperature, so care needs to be taken to ensure consistency in this regard. Where both instruments are widely available, it is recommended to use Qubit for quantifying DNA concentrations, while the Nanodrop or other spectrophotometer readings can be useful for assessing DNA purity.

When targeting extra-organismal DNA extracted from environmental samples (water and sediment), it is important to bear in mind that a significant portion of the DNA extracted can belong to non-target organisms (Capo et al. 2020). Thus, even when a relatively high total DNA concentration is obtained, the concentration of target DNA can be extremely low. Where it is necessary to quantify target DNA in preparation for metabarcoding, this can be achieved by running a qPCR assay or equivalent with the chosen metabarcoding primers, although the accuracy of this analysis will depend on primer specificity to the target group. Quality control analyses such as running agarose or digital gels (e.g. FragmentAnalyzer, Agilent, Santa Clara CA) can also be useful to assess potential DNA degradation and the fragment size distribution of the DNA.

6.2 Inhibition testing

A common cause of false negative results is PCR inhibition (Jane et al. 2015, McKee et al. 2015), which is caused by chemicals and compounds in the sample that interact with the PCR and either reduce its efficiency or cause it to fail completely, even when target DNA is present. Inhibitors are usually present in soil, sediment and faecal samples, and often in water samples. In soils and water, complex humic substances are the main known inhibitor of PCR (Braid et al. 2003), although much research remains to be done to complete our understanding of this. They tend to be more prevalent in eutrophic waters than in oligotrophic ones and are also associated with high sediment loads, shallow waters that contain a lot of organic material such as faeces from livestock, heavily polluted waters, and samples containing a high concentration of calcium carbonate (e.g. DNA extracted from samples that contain bivalves and gastropods; Schrader et al. 2012). Thus, it is important to test extracted DNA for inhibition in order to avoid false negative results.

This can be achieved through the use of an exogenous internal positive control (IPC) added at known concentration to the extracted DNA (Hoorfar et al. 2004; Furlan et al. 2016; Brys et al. 2020b). These non-target standardised IPC DNAs can be purchased commercially along with the primers and probe required for the qPCR/ddPCR reaction, and they can be multiplexed with the tests for the target species (i.e. tested for in the same reaction to save time and cost). If amplification of the IPC fails or is delayed (i.e. Ct value is higher than expected), the sample should be assumed to be inhibited (Hartman et al. 2005). The difference between expected and observed Ct value can give some indication of the strength of the inhibition.

Inhibition can often be overcome by using additional purification steps (using commercial kits such as the Zymo OneStep PCR Inhibitor Removal Kits or Qiagen PowerClean kits), use of chemical enhancers such as bovine serum albumin (BSA) and dimethyl sulphoxide (DMSO) in the PCR reaction, or by dilution of the DNA. The inhibition test should then be repeated to check that the inhibition has been overcome.

While dilution of the DNA reduces the chance of one type of false negative result (inhibition) it can also increase the chance of false negatives occurring due to stochasticity by reducing the concentration of target DNA. This can be especially consequential when working with eDNA, where target DNA is already likely to be at very low concentrations. Dilution should therefore be compensated by increasing the number of PCR replicates performed or the volume of extracted DNA added to each reaction.

6.3 Analytical controls used for DNA and eDNA analyses

Data generated from eDNA analyses could be used as evidence for species presence (or when rigorously understood, absence) in contexts including permits for new development or construction of infrastructure, the designation of protected areas, protected species licensing, assessing illegal species introductions, and decisions to remove connectivity barriers. Many of these applications sit within a regulated framework of

requirements, with legal consequences for environmental mismanagement, and this means that data derived from DNA-based analyses may have to be defended in court. Moreover, management responses to detection of certain species can be expensive and disruptive, so confidence in data quality is paramount.

Thus, the use of DNA-based monitoring methods to determine species presence or absence at a given location for environmental management purposes requires rigid quality assurance protocols, similar to those applied in other high-consequence sectors that use molecular methods, such as forensics (Board 2000, Brandhagen et al. 2020) and medicine (Precone et al. 2018). This typically goes beyond the main principles of sound field and laboratory conduct (Griffiths et al. 2015, Goldberg et al. 2016), and involves the use of analytical controls at every step of the workflow, from field planning to reporting and interpretation frameworks. The controls are particularly important as the analyses move out of the research sphere and into commercial and operational use where the field, laboratory, data analysis, and interpretation may be undertaken by different actors. In this scenario, the decision-maker must have access to evidence that each stage of the process has performed as expected, that there is a system in place to prove chain of custody such that a judgement can be made as to the accuracy and authenticity of the results, and any errors can be traced back to a specific point in the workflow. We do not discuss chain of custody issues, but a rigorous quality assurance document can lay out this process. For example, the US Fish and Wildlife Service has established such a quality assurance document for the monitoring of the economically important invasive asian carps (Woldt et al. 2020). Here we present a summary of the controls necessary for such cases in the application of DNA-based monitoring (**Table 4**).

Table 4. A summary of the positive and negative controls that can be used at each stage of the process.

Stage in workflow	Positive control	Negative control	Referred to in this document	Notes
Site controls	Sample from site with known **presence** of target species	Sample from site with known **absence** of target species		Not always needed, but important when using new or partially-validated assays. Site negative is more relevant for single-species tests than for metabarcoding
Filter controls (aquatic eDNA only)		Distilled or bottled water processed under the same filtering conditions as the field samples to check that contamination is not being introduced in the field	2.1.12	Highly recommended for eDNA samples, especially when using open filters
Laboratory controls: DNA extraction	Add IPC (internal positive control) at known concentration to the lysis buffer (or other extraction solution) to control for successful DNA extraction. Amplify the IPC with qPCR or ddPCR to check the expected concentration is obtained (after controlling for inhibition).	DNA extraction protocol applied in the absence of the biological sample. Must be carried out alongside sample extractions using the same equipment and materials. Include IPC for comparison with samples. For each set of samples extracted together, add at least one negative extraction control	3.3	Negative controls mandatory Positive controls highly recommended

Stage in workflow	Positive control	Negative control	Referred to in this document	Notes
Laboratory controls: Inhibition testing	Use IPC added to extracted DNA and amplify with qPCR / ddPCR to check efficiency. If IPC amplification fails or is delayed, purification steps should be undertaken and the test repeated. Negative results from inhibited samples should be reported as inconclusive		3.4.2 3.4.3	Highly recommended for all sample types; mandatory for environmental samples (water and sediment)
Laboratory controls: amplification with cPCR ddPCR, qPCR	Use DNA of the target species as a positive control for qPCR or ddPCR analysis. This should be run on every plate of samples. Standard curves are needed for DNA quantification in qPCR but not in ddPCR	Multiple template negative controls should be included in each plate of samples, using nuclease free water as a template in the PCR reaction. All other equipment and materials should be identical to those used for the samples	3.4.2 3.4.3	Mandatory
Laboratory controls: Amplification for Metabarcoding	Amplify DNA from a mock community with known species composition. The mock community species should not be expected to occur in the samples on the same sequencing run. This should be included at least once in every PCR plate and included in all downstream processes	Multiple template negative controls should be included in each plate of samples, using nuclease free water as template in the PCR reaction. All other equipment and materials should be identical to those used for the samples	3.5.3	Recommended
Laboratory controls: Sequencing	Sequence the mock community used as the positive PCR control	Sequence all negative controls included above	3.5.3	Negative controls Mandatory Positive controls highly recommended

We define analytical positive and negative controls as follows:

A **negative site control** refers to a sample collected from a field site where the target taxon is known to be absent. It is not always required but is a core part of assay validation (Thalinger et al. 2021b) and can be useful when using a new or semi-validated targeted detection assay (e.g. qPCR or ddPCR tests that do not involve sequencing). If positive results are obtained for a negative site control, this indicates an increased likelihood of false positive results being obtained (i.e., species being detected when they are absent; a Type 1 error). Note that DNA from sources such as sewage, restaurants and bird faeces can in rare cases be detected. This reflects true presence of DNA from contaminant sources in the environment but does not indicate occurence of living individuals at the

site. Knowledge about the location of wastewater outlets, human dwellings, presence of bird flocks and restaurants close to the sampling site is useful in order to omit these data and to verify the true presence of a species, and this should be particularly taken into account when selecting negative site control locations. Similarly, a **positive site control** refers to a sample collected from a location where the target is known to be present and expected to be detected. Regardless of the analysis type (targeted detection or metabarcoding), positive site controls are useful for confirming performance of the chosen workflow in detecting the presence of the target species.

A negative filtration control is a sample where DNA-free water is filtered alongside the eDNA samples to check that DNA is not transferred between samples. It is especially important when any piece of equipment comes into direct contact with consecutive samples, using decontamination procedures to avoid cross-contamination.

Negative laboratory controls consist of DNA-free samples processed alongside the test samples at each stage of the process to check for (cross-)contamination. All negative controls should be processed to the end of the workflow, and new negatives should be added at each stage so that any contamination detected can be traced back to a specific point in the analysis.

Positive laboratory controls consist of a known concentration of pre-prepared DNA of one or more species that are expected to be amplified efficiently using the selected PCR protocol. In some cases they can also consist of purpose-designed synthetic DNA. If the positive control does not amplify as expected, it implies a high risk of false negative results (i.e., species not being detected when their DNA is present in the sample; a Type II error). Positive laboratory controls can represent a potential source of contamination for test samples, so they should be designed and handled with care. Where mock communities are used as positive DNA controls in metabarcoding workflows, they should contain species that are not expected in the test samples (e.g. from different geographic regions). This mitigates the risk that the positive control samples themselves represent a contamination risk, and additionally allows them to also be useful for detecting cross-contamination, functioning as additional negative controls.

Summary: Measuring the quality and quantity of extracted DNA

Key takeaways:

- Quantification of DNA post extraction can be achieved in multiple ways, but it is important to remember that the total quantity of DNA does not always correlate with the quantity of target DNA.
- Inhibition is a major cause of false negative results. Inhibition can usually be overcome through the use of clean-up kits or dilution so internal positive controls should be routinely used to check for inhibition before amplifying DNA.
- Negative and positive controls should be integrated throughout the workflow to provide assurance in the reliability of results. Certain use cases may require a greater number and variety of controls than others. For instance more controls would be needed where there is a legal or regulatory relevance and results may be scrutinised in a court of law, than would be required for a research study.

Priorities for future research:

- Standardisation of a quality control framework to govern the type and quantity of controls that should be used in different use cases.

- Analysis of different types of internal positive controls to identify those that best reflect the state of environmental DNA and can estimate losses.

7. TARGET SPECIES DETECTION

Targeted species detection methods refer to assays that are used to screen DNA samples for the presence of one or a few pre-defined species. While in principle bulk sample extracted DNA could be screened for indicator taxa, this is rarely practiced, and these samples are usually analysed with metabarcoding. Therefore, we focus here on single species detection from aquatic eDNA.

7.1 Analysis methods for target species detection

While various technology platforms can be used, the basic principle is that species presence is inferred based on successful amplification using a taxon-specific primer set. The same primer set can be adapted for use with different technologies, the most commonly-used being:

- Conventional PCR (cPCR; also referred to as endpoint or diagnostic PCR),
- Quantitative PCR (qPCR)
- Digital droplet PCR (ddPCR).

Note that other methods exist and are being developed, but have yet to be fully integrated into mainstream eDNA applications. Examples include loop-mediated isothermal amplification (LAMP), nucleic acid sequence based amplification (NASBA), self-sustained sequence replication (3SR), rolling circle amplification (RCA), and CRISPR-based assays (Williams et al. 2017, 2019).

The benefits of these single-taxon screening methods lies primarily in the simplicity of the laboratory and data analyses, and potentially in their sensitivity compared to metabarcoding approaches for some taxonomic groups (Simmons et al. 2016, Blackman et al. 2020). Tests can be carried out fast - even to the point of being able to conduct them in the field - which is useful where management decisions need to be taken quickly. Principal disadvantages include the need for extensive and costly validation studies to ensure specificity of the assay prior to operational use (Thalinger et al. 2021b), and the limited information obtained, which relates only to the target species. Thus, cost mounts rapidly as the number of target species increases.

7.1.1 Conventional PCR (cPCR) assays

cPCR does not incorporate any fluorescent dyes during the amplification process. The generated amplicon is visualized on agarose gels or with capillary electrophoresis machines. Using agarose gels, interpretation of the results is restricted to presence/absence inference, while analysis with capillary electrophoresis allows a fluorescence threshold to be set for positive detection, and the signal strength of the target band can be used as a semi-quantitative measure of target DNA.

While some eDNA studies have used this method, it most commonly serves as a cost-effective tool for preliminary tests during the assay validation process. Specificity to the target relies exclusively on primer design, so non-target amplification is a substantial risk with this approach if primer design has not been sufficiently optimised and validated (see below). Positive amplification of the target species can easily be confirmed by sanger sequencing of the PCR product, which should be carried out for a subsample of all successful amplifications.

7.1.2 Quantitative real-time PCR (qPCR)

qPCR is currently the most widely used technique for the detection of single taxa from environmental samples (Langlois et al. 2020). Unlike with cPCR and ddPCR, amplification and visualization occur simultaneously during qPCR. Quantification of target DNA can be achieved using standard curves generated with defined quantities of DNA, so long as target DNA is present at levels above the limit of quantification (LOQ) (see below). There are two common forms of qPCR assay, which both rely on the use of fluorescent dyes (See also: Real-time PCR handbook. 2012 Life Technologies Corporation):

1. Sybr Green technology intercalates the fluorescent dye between double-stranded DNA strands as they anneal. Like with cPCR approaches, assay specificity relies primarily on primer design since no probe is used. However, the melt curve profile (which charts the change in fluorescence as temperature rises) can be used as an additional source of information, since this is determined by the GC-content, length and other attributes of the amplicon sequence. Indeed, multiple species can be amplified with the same primer set and separated based on melt curve profiles given sufficient difference in amplicon sequences. Although easier to develop and cheaper to run, these assays still rely on visual recognition of melting curve profiles typical to the target species, and ambiguous results are common.
2. qPCR assays using hydrolysis probes (often referred to as TaqMan™ probes) represent the most commonly-used tests for individual target species. Specificity is achieved both through primer design and via a target-specific probe that anneals to the single stranded DNA during PCR, in a region between the forward and reverse primer. The fluorescent dye is released only when the probe anneals to the matching target sequence, meaning that fluorescence is only detected when both the primers and probe match the target sequence. This improves specificity and reduces the chance of false-positive results that arise from non-target amplifications. However, it increases the complexity of the optimisation and validation process since the primers and probe must be designed and optimised together to maximise efficiency of the assay. Moreover, it is important to be aware that there are two different types of hydrolysis probes - minor-groove binding (MGB) and non-MGB probes - and these have different properties:
    1. Non-MGB probe sequences are longer than the primer sequences because they require a higher annealing temperature than the primers. A single base pair mismatch can usually be tolerated, meaning that fluorescence still occurs even if the target sequence varies slightly from that for which the probe was designed.
    2. MGB probes do not require a higher annealing temperature than the primers. They are shorter than the primer sequences, and a single-base mismatch has a much greater impact on annealing success, allowing higher stringency to be achieved. They are also more expensive than Non-MGB probes and must be licenced for commercial use. Note that an MGB probe sequence manufactured with non-MGB technology will lead to qPCR failure because it will be too short to function as intended under the cycling conditions required for the primers.

7.1.3 Droplet digital PCR (ddPCR) assays

Similar to cPCR, ddPCR combines thermo cycling amplification (PCR) with subsequent evaluation of the generated amplicons in a second, separate step. ddPCR is a highly sensitive method that is capable of detecting single copies of target DNA. Furthermore, ddPCR is an endpoint reaction, and allows absolute quantification of target DNA copy number without the need for standard curves (Hindson et al. 2011).

ddPCR randomly partitions 20 µl of PCR master mix (including up to 10 µl DNA extract) into ~20,000 individual droplets. The droplet matrix is then subjected to thermo cycling. Like qPCR, ddPCR is based on either i) probe hydrolysis, or ii) intercalating dyes (Sybr Green/Evagreen) technology. After amplification,

droplets that contain target DNA are detected by fluorescence, allowing absolute quantification through poisson statistical analysis of the ratio of positive to negative droplets (Baker 2012).

ddPCR instruments are more expensive and less widely accessible than qPCR instruments, but the approach is reported to be less prone to the effects of PCR inhibition. As a result, ddPCR can often be run with higher relative extract volumes without any sign of inhibition, which may at least partly explain the reported higher sensitivity of ddPCR compared to qPCR (Brys et al. 2020b). In addition, ddPCR allows exact quantification at very low concentrations (Hindson et al. 2011, Doi et al. 2015). Hence, ddPCR is particularly useful for detecting very low abundant target DNA and has shown great potential for detecting species that pose significant detection challenges, such as sharks (Lafferty et al. 2018, Schweiss et al. 2020).

7.2 Defining limits of detection (LOD) and quantification (LOQ)

LOD and LOQ should be determined after an assay has been optimized and must be reported in order for an assay's results to be interpreted correctly (Klymus et al. 2020). The method for determining LOD and LOQ varies across platforms.

For qPCR, estimating target DNA concentration requires normalization against standard curves. These should consist of at least 5 different concentrations of DNA that contains primer and probe binding sites identical to the target species. Because high-concentration positive control DNA poses a contamination risk (especially when working with low template samples such as eDNA), the standard DNA should ideally be created using purified amplicons (e.g. Currier et al. 2018), or ideally double-stranded synthetic amplicons (e.g. gBlocks™; Langlois et al. 2020). For calculation of LOD and LOQ, the number of technical replicates per sample should be as high as possible (5-12) in order to properly assess the mean and associated variance per sample, with 3 replicates being the absolute minimum required to calculate a standard deviation. LOD and LOQ can be defined in various ways (e.g. Agersnap et al. 2017, Hunter et al. 2017), but a recent effort to standardize these definitions ultimately defined LOD as being the lowest concentration at which 95% of the technical replicates of the standard amplify, while LOQ was defined as being the lowest concentration for which the coefficient of variation (CV) value is < 35% for the used standard DNA (Klymus et al. 2020); also adopted by (Thalinger et al. 2021b). The CV represents the ratio of the standard deviation to the mean, and allows the amount of variation to be compared between replicates of low versus high target concentration samples. Based on these definitions, it is still possible (indeed common) to detect target DNA at concentrations below the theoretical LOD, especially when multiple technical replicates are run for each environmental sample. Detections below the LOD may be attributed a lower confidence level in interpretation of results. Since definitions of LOD and LOQ are still being established insofar as they relate to eDNA, it is important to report not only the actual values of these two parameters for each assay, but also the way in which they have been defined and calculated.

- For cPCR with capillary electrophoresis, the LOD is selected to reliably separate target bands from background fluorescence and is commonly around 0.08 Relative Fluorescence Units (RFU) (Thalinger et al. 2019). A positive PCR result can typically be generated from 10-30 target DNA double strands present at the beginning of PCR (Deprez et al. 2016, Hunter et al. 2017, Thalinger et al. 2021c).
- For ddPCR, the lowest concentration that can theoretically be detected and measured is one target molecule per reaction (Deprez et al. 2016, Hunter et al. 2017, Thalinger et al. 2021c), and each reaction can include up to 10 µl of DNA template because of the reduced impact of inhibitors compared with in qPCR. The same practical definition of LOD as used in qPCR can be applied to ddPCR (95% positive detections of standard DNA) meaning that here too it is still possible to obtain valid detec-

tions below the LOD. Note that when calculating the mean target concentration of samples, it is important to include non-amplifications as zero-estimates.

7.3 Species specific primer design and validation

Good design of species-specific primers and probes is critical for any targeted species approach because interpretation of the results often depends solely on whether or not amplification occurs. Non-target amplification can therefore lead to species presence being wrongly inferred, with potentially costly consequences for environmental management. The challenge with the use of target-based approaches for routine monitoring at large geographic scales lies in designing and validating species-specific primers that are reliable across diverse ecological systems. This can be achieved, but it involves significant investment of time and effort, along with resources to fully test and validate assay performance in the intended environment.

As a guiding principle, the design and implementation of PCR primers should follow the established guidelines for the Minimum Information for the Publication of Quantitative Real-Time PCR Experiments, which have been adapted for use in eDNA applications (Bustin et al. 2011, Goldberg et al. 2016, Langlois et al. 2020). However, in practice the degree to which published eDNA research follows these guidelines is highly variable, since research study objectives often do not require the high level of robustness and quality assurance needed for routine biomonitoring. When assays are implemented outside the geographic region for which they were originally developed, the validation status of the primer pair should be considered with great care and additional tests for specificity and sensitivity are likely to be required (Thalinger et al. 2021b) recently carried out an extensive meta-analysis of published assays used for targeted eDNA detection, and proposed a set of validation levels along with clear guidelines for how to both develop new assays and implement previously developed assays, taking into account their limitations. This extensive set of guidelines can be used as the basis on which to standardize reporting guidelines for targeted eDNA assays, facilitating their use for the entire scientific community and environmental managers (https://edna-validation.com).

7.4 Quantitative interpretation of results

The concentration of target DNA in the DNA extract (as measured by qPCR and ddPCR) or fluorescence strength of the target band (in cPCR) has been linked to species abundances under controlled experimental settings, but this is extremely complex to extend to natural systems. Use of IPCs for internal normalisation of quantitative estimates, together with models based on allometric scaling of species' body sizes (Yates et al. 2020) can help to calibrate estimates, but caution should still be applied given the multitude of factors that affect the amount of DNA shed by a given number of individuals and captured using a given sampling and capture strategy (see section 2.1 above).

Summary: Target species detection (active surveillance)

Key takeaways:

- The key advantages of targeted assays are the straightforward laboratory and data analysis, enabling fast turn-around of results, and high sensitivity to detect targets at very low concentrations.
- The key disadvantages are that extensive validation of assays is necessary prior to application because amplification of non-target organisms cannot be distinguished from target amplifica-

tion in the results output, so assay specificity needs to be rigorously tested in field settings and primer design is of critical importance. In addition, costs scale with the number of targets so the approach becomes prohibitively expensive for the surveillance of multiple taxa.
- The most commonly used technologies are conventional PCR (cPCR), quantitative PCR (qPCR), and digital droplet PCR (ddPCR). cPCR is cost-effective and useful for research or in the early stages of assay validation but is not recommended for use in routine surveillance applications. qPCR using hydrolysis probes is the most commonly-used method for targeted detection of taxa, but users should be aware of the differences between MGB and non-MGB probes. Like qPCR, ddPCR can be used with probes to maximise specificity, but is less affected by inhibition and can provide absolute quantification of target DNA. ddPCR equipment is expensive to purchase but this is a good solution for detection of very low abundance target DNA.
- Limits of detection (LOD) and limits of quantification (LOQ) can be defined in various ways. These are useful measurements for comparing relative sensitivity of different assays, but are not straightforward to extend to the probability of detection in a given field setting.

Priorities for future research:
- Standardize reporting to ensure robust application and advancement of previously developed assays
- Establish cost-effective high-throughput approaches and portable devices without forfeiting sensitivity (e.g. high-thoughput qPCR, Wilcox et al. 2020). In particular, there is a need for methods to multiplex tests so that multiple targets can be screened for at once while remaining cost-effective.
- Further research into the relationship between species abundance and eDNA concentrations in natural environments.

8. METABARCODING TO SURVEY BIOLOGICAL COMMUNITIES

Metabarcoding allows the simultaneous taxonomic identification of organism assemblages from a biological sample using high throughput sequencing of a standardised gene fragment (Yu et al. 2012). Originally developed for the assessment of microbial community diversity (Pace et al. 1986), metabarcoding was first applied to seawater samples to characterise natural bacterial communities (Sogin et al. 2006). It has since been applied to a wide range of macroorganism communities and is already beginning to revolutionise biomonitoring (Deiner et al. 2017b, Seymour et al. 2020), because it enables the generation of high-quality biodiversity data for all taxonomic groups occurring in an ecosystem in a consistent and standardised manner (Baird and Hajibabaei 2012, Gibson et al. 2015). Recent studies comparing the statistical power of metabarcoding with traditional approaches in terms of taxonomic resolution, sample similarity, taxon misidentification, and taxon abundance, show that metabarcoding improves the quality and utility of ecological data and allows new insights into the assembly and structure of communities (Bush et al. 2019, McElroy et al. 2020). Metabarcoding data are already being used for the calculation of various biotic indices encompassing species richness and taxonomic composition (Aylagas et al. 2014, 2018, Elbrecht et al. 2017b), and there is further potential for the development of new molecular metrics for routine aquatic biomonitoring (reviewed in Pawlowski et al. 2018, applied in Seymour et al. 2020; see also Text Box 3).

Metabarcoding involves three principle laboratory steps:

1. Designing and selecting primers
2. Amplicon library preparation
3. High-throughput sequencing

Here we focus on the laboratory steps for the metabarcoding of taxa routinely used as ecological indicators. Most biomonitoring efforts will involve large numbers of samples and recent studies have already started addressing the practical and technical challenges of scaling up metabarcoding workflows for freshwater monitoring (Elbrecht and Steinke 2018, Hering et al. 2018, Leese et al. 2018, Porter and Hajibabaei 2018).

8.1 Metabarcoding primer selection

The goal of any metabarcoding analysis is to record and identify all species within a particular taxonomic group from a sample. To accomplish high taxonomic resolution (i.e. species level assignments) and maximise detection probability for rare species, designing or selecting appropriate primers is crucial.

Primers can be designed to target narrow taxonomic groups through to very broad ones. At the narrowest end of the spectrum, primers can also be designed to target a single species, using fast-evolving DNA regions that allow identification of intraspecific (population level) genetic diversity (e.g. Tsuji et al. 2020). At the other extreme, primers can be designed to encompass entire kingdoms or domains (e.g. Eukaryotes or Bacteria). Between these extremes, commonly used metabarcoding assays target intermediate-level groups, such as animals (Metazoa), vertebrates, fish, diatoms, etc. Some of those commonly used in aquatic biomonitoring are listed in **Table 5**.

To target broad taxonomic groups, degenerate primers are often used, which incorporate some level of flexibility in the priming sequence via so-called 'mixed' or 'wobble' bases. This allows for some sequence variation in the primer binding region of the target organisms so a more diverse group can be targeted. Degenerate primers will increase the number of amplified taxa, but can lead to increased amplification of non-target organisms when applied to environmental samples (e.g. Deiner et al. 2016, Macher et al. 2018b, Wangensteen et al. 2018).

The most challenging aspect of primer design often relates to intermediate taxonomic groups such as Metazoa or paraphyletic groups such as 'macroinvertebrates' when they are targeted in environmental samples. In this case, primers have to be sufficiently broad to capture a taxonomically diverse group of target organisms, while not being so permissive as to amplify non-target groups such as bacteria and algae, which may dominate in terms of quantity of DNA isolated from environmental samples. High levels of non-target amplification in environmental samples are often mitigated by increasing sequencing depth. The increased read depth allows for sequences from non-target groups to be discarded during bioinformatic filtering, leaving a sufficient amount of sequence data derived from the target group. While this can work, it is an inefficient solution increasing overall costs and careful primer design is preferred where possible, with the aim of increasing target specificity and minimising non-target amplification.

There is an inherent trade-off in that as the target group broadens, a wider cross-section of biodiversity is obtained, but this comes at the expense of the completeness of the data in each of the different groups (Macher et al. 2018b, Hajibabaei et al. 2019, Gleason et al. 2020). For example, a eukaryote assay applied to marine sediment samples will profile biodiversity across metazoans, algae including diatoms, protists (e.g. foraminiferans), and marine fungi, yielding an integrated biodiversity signal that is likely to reflect overall ecosystem status (Grey et al. 2018). Nonetheless, a metazoan-specific primer set would more comprehensively characterise the animal diversity contained in the samples. Similarly, a vertebrate assay applied to aquatic eDNA samples is attractive for its ability to reveal both fish and mammals in a single assay, but ap-

Table 5. Barcode markers commonly used for metabarcoding in various organism groups. For a comprehensive list see (Deiner et al. 2017b).

Primer name	Target group	Gene region	Amplicon length	Citation	Forward primer sequence	Reverse primer sequence	Strengths	Weaknesses
Batra	Amphibians	12S	c.a.60 bp	(Valentini et al. 2016, Taberlet 2018)	ACACCGCCCGTCACCCT	GTAYACTTACCATGTTAC-GACTT	High taxonomic coverage	Do not discriminate Peophylax species, need a human blocking primer
515F–806R	Bacteria	16S V4	c.a. 390 bp	(Caporaso et al. 2011)	GTGYCAGCMG-CCGCGGTAA	GGACTACNVGG-GTWTCTAAT	Very widely used and standardised	
[Diat_rbcL_708F_1+Diat_rbcL_708F_2+Diat_rbcL_708F_3] + [R3_1+R3_2]	Diatoms	rbcL	312 bp	(Vasselon et al. 2017a)	AGGTGAAGTAAAAGGTT-CWTACTTAAA AGGTGAAACTAAAGGTT-CWTACTTAAA AGGTGAAGTTAAAGGTT-CWTAYTTAAA	CCTTCTAATTTACCWAC-WACTG CCTTCTAATTTAC-CWACAACAG	Diatom specific rbcL primers used in biomonitoring - high specificity and high coverage for diatoms - fully adapted to the use of Diat. barcode reference library - enable the evaluation of relative abundance of taxa (Vasselon et al 2018)	High specificity and coverage are obtained by pooling 3 forward and 2 reverse primers in equimolar proportions
DIV4for - DIV4rev	Diatoms	18S V4	280-300 bp	(Visco et al. 2015) modified from (Zimmermann et al. 2011)	GCGGTAATTCCAG-CTCCAATAG	CTCTGACAATGGAATAC-GAATA	Diatom specific 18S primers used in biomonitoring	
1389F - 1510R	Eukaryotes	18S V9	c.a. 150 bp	(Amaral-Zettler et al. 2009)	TTGTACACACCGCCC	CCTTCYGCAGGTTCACCT-AC	Universal eukaryotic primers	
SSU_F04+SSU_R22	Eukaryotes (marine meiofauna)	18S	380bp	(Blaxter et al. 1998)	GCTTGTCTCAAAGAT-TAAGCC	CCTGCTGCCTTCCTTRGA	Good taxonomic coverage for marine meiofauna (V1-V2 region)	

Primer name	Target group	Gene region	Amplicon length	Citation	Forward primer sequence	Reverse primer sequence	Strengths	Weaknesses
SSU_F04 - SSU_R22mod	Eukaryotes (marine meiofauna)	18S	c.a. 400 bp	(Sinniger et al. 2016) modified from (Blaxter et al. 1998)	GCTTGWCTCAAAGATTAAGCC	CCTGCTGCCTTCCTTRGA	Eukaryotic 18S primers, commonly used for metabarcoding marine meiofauna	
TAReuk454FWD1 - TAReukREV3	Eukaryotes (marine plankton)	18S V4	c.a. 400 bp	(Stoeck et al. 2010)	CCAGCASCYGCGGTAATTCC	ACTTTCGTTCTTGATYRA	Universal eukaryotic primers commonly used for marine plankton (Tara Oceans) but not only.	
Teleo or Tele01	Fish	12S	c.a. 80 bp	(Valentini et al. 2016, Taberlet 2018)	ACACCGCCCGTCACTCT	CTTCCGGTACACTTACCATG	High taxonomic coverage, including Actinopterygii and Elasmobranchs	Poor species resolution for same family, need a human blocking primer
MiFish-U	Fish	12S	c.a. 170 bp	(Miya et al. 2015)	GTCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	Good specificity, coverage, and resolution	
s14F1 - s15.3	Foraminifers	18S 37F	120-140 bp	(Pawlowski and Lecroq 2010, Pawlowski et al. 2014)	AAGGGCACCACAAGAACGC	CCACCTATCACAYAATCATG	Foraminifer specific 18S primers commonly used in metabarcoding	
FwhF2 + FwhR2n	Macroinvertebrates	COI	205 bp	(Vamos et al. 2017)	GGDACWGGWTGAACWGTWTAYCCHCC	GTRATWGCHCCDGCTARWACWGG	Short length of amplicon allows reliable amplification also of degraded material, very good taxonomic representation; short length allows to sequence on HiSeq and othe 2x150 bp platforms - cheap!	Not good for eDNA, too many bacteria are amplified

Primer name	Target group	Gene region	Amplicon length	Citation	Forward primer sequence	Reverse primer sequence	Strengths	Weaknesses
fwhF2 + EPTDr2n	Macroinvertebrates, freshwater (especially insects)	COI	142 bp (191 bp)	(Vamos et al. 2017, Leese et al. 2021)	GGDACWGGWTGAACWGTWTAYCCHCC	CAAACAAATARDGGTATTCGDTY	Reduced non target amplification with eDNA	Increased primer bias compared to bulk metabarcoding primers, misses some taxa (e.g. some crustaceans, few trichopterans) due to the increased specificity. Can be more challenging to establish in the lab.
fwhF2 + FoI-degen-rev	Macroinvertebrates, freshwater (especially insects)	COI	313 bp (365 bp)	(Yu et al. 2012, Vamos et al. 2017)	GGDACWGGWTGAACWGTWTAYCCHCC	TANACYTCNGGRTGNCCRAARAAYCA	Good amplicon length for Illumina sequencing	Not good for eDNA, too many bacteria are amplified. Forward primer affected by primer slippage
BF3 + BR2	Macroinvertebrates, freshwater (especially insects)	COI	418 bp (458 bp)	(Elbrecht and Leese 2017, Elbrecht et al. 2019)	CCHGAYATRGCHTTYCCHCG	TCDGGRTGNCCRAARAAYCA	Good taxonomic resolution	Not good for eDNA, too many bacteria are amplified. Amplicon length not ideal for all illumina sequencers
Unio	Molluscs (Unionida)	16S	c.a.130 bp	(Prié et al. 2020)	GCTGTTATCCCCGGGGTAR	AAGACGAAAAGACCCCGC	High taxonomic coverage	Some species cannot be discriminated
Vene	Molluscs (Venerida)	16S	c.a.130 bp	(Prié et al. 2020)	CSCTGTTATCCCYRCGGTA	TTDTAAAAGACGAGAAGACCC	High taxonomic coverage	Some species cannot be discriminated
12S V5	Vertebrates	12S	c.a. 106 bp	(Riaz et al. 2011)	ACTGGGATTAGATACCCC	TAGAACAGGCTCCTCTAG	Very high taxonomic coverage	Lower taxonomic resolution, particularly among bird species, non-target amplification of human

plying two separate assays that target fish and mammals, respectively, will usually recover more diversity in each group. Narrower assays are particularly useful for detecting rare species and those present at low population densities (e.g. Boussarie et al. 2018). A multi-marker approach is usually required for obtaining comprehensive species data across a wide range of taxonomic groups (Hajibabaei et al. 2019, Ficetola et al. 2020, Martins et al. 2020). The number and combination of primer sets selected will depend on the aims of the monitoring and the available budget, since cost scales with the number of primer sets to be used.

The potential for primer bias (also referred to as amplification bias) is an important consideration in primer design and selection. Primers that target a narrow taxonomic group can often be designed to have an exact or near-exact match to all target taxa, which ensures approximately equal amplification efficiency across taxa. This means that sequence read counts will usually correlate well with the relative concentrations of eDNA captured for each species in the sample. Conversely, those targeting a very broad group will vary in amplification efficiency across taxa, which reduces the potential to make (semi-)quantitative inferences from sequence read count data (Bista et al. 2018). Inclusion of positive controls and spike-ins can help with understanding stochasticity and potentially even quantification of metabarcoded data (Ji et al. 2019).

8.1.2 Amplicon length

Amplicon length is a key factor affecting primer performance for metabarcoding, and here too there is an intrinsic trade-off to be negotiated. Shorter amplicons are usually more sensitive for the amplification of degraded DNA (Geller et al. 2013, Leray et al. 2013), and are commonly prefered for eDNA metabarcoding (although note that much longer amplicons can be amplified from eDNA, including mitochondrial genomes; Bista et al. 2017, Deiner et al. 2017a). However, the increased sensitivity comes at the cost of taxonomic resolution, and short amplicons may not be able to separate some closely-related species. This will represent a significant limitation in some monitoring contexts if species level identification is required.

Thus, the balance tips in favour of using longer amplicons for metabarcoding of bulk invertebrates and sediment samples, especially when DNA is primarily derived from organismal DNA and is not expected to be degraded. In fact, for sediment biomonitoring the use of a longer amplicon can also be helpful in preferentially targeting DNA derived from living organisms, as opposed to accumulated eDNA, which tends to persist in short fragments (Jo et al. 2017). Here the upper limit on amplicon length is generally imposed by the sequencing technology. The vast majority of sequencing for metabarcoding is currently carried out on Illumina platforms, due to performance and cost advantages compared with other technologies, but the longest sequences that can be generated using this technology are 2 x 300 base pairs (Illumina MiSeq V3 kit). By the time that sufficient overlap is incorporated to allow reliable merging of the paired ends, and the primers, tags, adapters and any heterogeneity spacers are accounted for at both ends of the amplicon, the length of the amplicon itself is often restricted to little more than 450 bp.

8.2 Amplicon library preparation

Before ordering metabarcoding primers, it is first important to decide which metabarcoding labelling strategy will be employed. This is important because the primers need to be ordered with the corresponding 5' nucleotide additions required for building sequencing libraries. A thorough review of metabarcoding labelling strategies can be found in Bohmann et al. 2021.

8.2.1 Types of library preparation

The library preparation step combines three key processes:

1. PCR amplification of the chosen barcode region

2. Unique labelling of the DNA in each sample. This allows samples to be pooled for sequencing in such a way that the sequences can then be separated back into the samples they came from. Labelling is achieved through adding a unique sequence of nucleotides to each sample, and these are referred to as 'tags' or 'indexes'.
3. The addition of sequencing adaptors to each DNA amplicon. These adaptors bind the DNA onto the flow cell for sequencing.

There are three main strategies with which metabarcoding amplicon libraries can be constructed prior to sequencing: the '**one-step PCR' approach**, the '**two-step PCR' approach** and the '**ligation-based approach**' (Taberlet et al. 2012a, Bourlat et al. 2016, Leray et al. 2016, also reviewed in Bohmann et al. 2021) (See **Figure 7**). In these three approaches, sample labelling is carried out as 5' nucleotide tags and/or as library indexes. In the one-step and two-step PCR approaches, library preparation is carried out during PCR amplification with primers carrying sequencing adapters and indexes, while in the tagged PCR (ligation-based) approach, library preparation is achieved using ligation applied to pools of tagged amplicons.

8.2.1.1 One-step PCR

In the **one-step PCR approach**, sequence adapters and (typically) nucleotide tags are incorporated directly to the synthesis of the forward and reverse primers so that amplification and library preparation is achieved in a single PCR step, see **Figure 7a** (Elbrecht and Leese 2017, Vamos et al. 2017). This provides a highly streamlined workflow and reduces the risk of cross-contamination between amplicons, with the caveat of increasing materials cost for projects with large sample sizes and multiple primer sets. Although usually performed using conventional thermocycling machines, this can also be carried out on a qPCR platform (e.g. Bessey et al. 2020), providing simultaneous quantification of the product and eliminating yet another step that would otherwise be carried out separately.

Figure 7. Metabarcoding amplicon construct. There are three approaches for the identification of samples sequenced in a library, these are: one-step PCR (a), two-step PCR (b, c, and d) and a ligation-based approach (e).

The trade-off with this approach is that the primer sequences become very long when up to 60 nucleotides (sequence adapters and indexes) are added to the primers. This makes them costly to buy, while the long overhangs can decrease efficiency, thereby reducing the detection probability of rare species and the consistency between replicates (Zizka et al. 2019a).

The one-step approach is sometimes referred to as the 'fusion primer approach', but we avoid this terminology because all metabarcoding approaches involve primers that are to some extent a fusion of multiple components.

8.2.1.2 two-step PCR

In the **two-step PCR** approach, shorter and potentially more efficient primers are used for the initial amplification of the DNA extracts, see **Figure 7b**, c and d (Bourlat et al. 2016). However, note that the additions to the primers can still be quite long - typically around 30 nucleotides. In the two-step PCR approach, the primer in the first round of PCR includes at minimum the target-specific primer sequence and an adaptor 'tail' for the second-round PCR primers to bind onto. These most minimal metabarcoding primers are attractive for maximising efficiency (by minimising length) and cost-effectiveness (the same primer is used for all samples). Nonetheless, these advantages need to be weighed against risk of cross-contamination between amplicons, which can easily occur during two-step library preparation and is untraceable if these minimal primers are used. A more robust modification of this approach is to incorporate a short, sample-specific 5' nucleotide tag sequence in first round PCR primers, such that each sample is amplified with a different set of primers incorporating a unique sequence of bases between the primer sequence and the adaptor tail. This has higher up-front costs since a separate primer set needs to be purchased for each unique index used for each sample.

The second stage of the two-step PCR process consists of a short second-round PCR (typically 8-10 cycles; Bourlat et al. 2016) performed with primers that bind to the tail of the first-round primers and incorporate Illumina library indexes and adaptors. If primers with sample-specific tags have been used during the first-round PCR then multiple samples can be combined within a single library index in the second round.

8.2.1.3 Ligation-based approach

The third of the three main strategies for amplicon labelling in metabarcoding studies is '**adapter ligation**' or '**ligation-based approach**', see **Figure 7e** (Leray et al. 2016). Note that this is sometimes referred to as the 'tagged primer approach', but we avoid this terminology since most library preparation approaches incorporate some element of tagging within the primer construction.

In the ligation-based approach, DNA extracts are PCR amplified with primers carrying short 5' nucleotide tags, typically just 6-10 nucleotides in length and unique to each sample. The additions to primers are thus the shortest among the three approaches, which in theory should cause the least reduction in PCR efficiency. Following PCR, the tagged PCR products are pooled, and sequencing adapters are added using a ligase enzyme that covalently links the amplified DNA fragments to the adaptors[8]. Library indexes can be included with the adaptors or added separately via a PCR step (reviewed in Bohmann et al. 2021). Detailed protocols for the robust design and implementation of ligation based library preparation include the Tagsteady protocol (Carøe and Bohmann 2020), and commercial library preparation kits include the Illumina TruSeq DNA PCR free library kit.

In all three approaches, heterogeneity spacers can be added to the primer sequence to improve sequencing performance by increasing the diversity at each base position, see **Figure 7c** and d (De Barba et

[8] https://eu.idtdna.com/pages/technology/next-generation-sequencing/library-preparation/ligation-based-library-prep

al. 2014, Elbrecht and Steinke 2018). These usually take the form of a series of degenerate bases that vary in length between samples.

8.2.2 PCR replication

Regardless of which library preparation is used, it is important to consider the number of PCR replicates to be performed per sample. Species represented in low DNA copy numbers can easily be missed in any given reaction, which introduces an element of stochasticity (Murray et al. 2015). To increase the detection probability of rare species, it is therefore usual to carry out multiple replicate PCR reactions for each sample, which effectively increases the volume of the DNA extract that is analysed. For bulk tissue and sediment samples, it is typical to carry out three PCR replicates per sample, although this does vary somewhat. In general, the lower the concentration of target DNA in the sample, the more PCR replicates will be needed to detect all species present in a sample (Doi et al. 2019). Thus, it is often recommended to carry out more PCR replicates for eDNA analyses than for bulk tissue samples, with a minimum of eight replicates suggested in the former case (Ficetola et al. 2015, Doi et al. 2019). Increasing the number of PCR replicates brings additional costs, and fewer replicates may be used in cases where there is already a high level of biological replication (i.e. increased sampling intensity) to compensate for the reduced detection probability at the individual sample level.

PCR replicates can either be labelled with the same nucleotide tags and/or indexes, meaning that the sequences derived from them will be pooled, or they can be individually tagged/indexed and treated as separate samples for downstream analysis. Sequencing each replicate independently allows for detection confidence to be bioinformatically assessed as a proportion of replicates in which a species occurs (Zepeda-Mendoza et al. 2016), which helps to balance error removal with diversity detection (Alberdi et al. 2018). However, it is an expensive approach when high numbers of replicates are performed for each sample (e.g. for aquatic eDNA analyses), since all of the pre-sequencing purification, quality checking and normalisation steps need to be carried out separately for each replicate. Pooling PCR replicates represents a cost-efficient way to maximise detection probability of rare species.

8.2.3 Choice of polymerase

Another often overlooked consideration is the choice of polymerase, which can introduce significant amplification bias based on varying GC content preferences (Nichols et al. 2018). Moreover, some proofreading polymerases are incompatible with the degenerate base Inosine, leading to reduced amplification efficiency or complete failure when used with primers that contain this base (eg. jgHCO2198 from (Geller et al. 2013), emphasising that methodological choices are not independent of other elements of a workflow. Different polymerases may also introduce different rates of copy errors, which occur when bases are incorrectly copied by the enzymes during PCR, leading to artifactual OTUs and overestimation of diversity (Nichols et al. 2018). This may be controlled for by methodological approaches such as the use of unique molecular identifiers (UMIs; Fields et al. 2020), but this approach is not yet widely used and it remains unclear how well it will work on low-template samples such as those derived from aquatic eDNA. In the meantime, it is recommended to apply an experimental approach to optimising choice of polymerase for particular primer sets to maximise efficiency and minimise error and biases.

8.3 Preparing amplicon libraries for sequencing

The most common reasons for a sequencing run to fail or produce low quality data is overclustering, which reduces the efficiency of base-calling and is usually linked to either insufficiently purified ampl-

icon libraries, too high concentration of libraries due to mis-quantification of libraries, or insufficient nucleotide diversity[9] (Illumina, 2016).

8.3.1 Purification

Due to the preferential sequencing of short reads on high-throughput sequencing (HTS) platforms, it is vital to remove primer dimers and other short fragments of DNA from the amplicon libraries prior to sequencing, otherwise few or no target sequences will be obtained. Amplicon libraries are usually purified using magnetic beads (e.g. AMPure XP, or Solid Phase Reversible Immobilization (SPRI)) or gel extraction, although the latter is labour-intensive and presents a higher contamination risk, especially where primers have been used without individual tags in first-round PCR. Gel extraction may be useful in small projects where a very close non-target band is present, but in general the presence of such a non-target band is a sign that the protocol requires further optimisation before being adopted for routine use. For large scale projects, post-PCR magnetic bead clean-up can be conveniently carried out in 96 well plate format (Elbrecht and Steinke 2018). Libraries should then be validated by verifying average fragment size and making sure there are no additional unwanted fragments present. This should be carried out using a platform such as a Bioanalyzer or Fragment Analyzer, rather than relying on agarose gel visualisation, which lacks the required sensitivity.

8.3.2 Quantification

A high degree of precision is required for accurate normalisation of libraries to ensure an even distribution of sequencing and prevent overclustering. The most accurate quantification of libraries is achieved using qPCR, although Qubit, TapeStation and Bioanalyzer devices are often used. Nanodrop and other spectrofluorometers are not recommended for this step (Illumina, 2016). Note that quantification of libraries carrying adapters such as those resulting from library preparation with the TruSeq DNA PCR free library kit can only be quantified using qPCR.

Following quantification, libraries should be adjusted to equal concentration and pooled for sequencing. Negative controls should be indexed and included in the final pool, but since these are expected to contain no detectable DNA it will usually not be possible to add them at equal concentration. There is not yet a standard approach for pooling negatives, but one suggestion is to calculate the median volume in which the test sample libraries are added (to achieve the equimolar pool) and add the negative controls in this volume.

8.3.3 Amplicon diversity

If all samples on the sequencing run have been amplified with the same primers and heterogeneity spacers have not been incorporated, there is a high risk that base-calling quality will be compromised by a lack of variation in nucleotide identity at each base position[10]. To mitigate this, it is common to add PhiX control v3 library as a spike-in accounting for approximately 5-10% of the final pooled library. The final concentration of the pooled libraries should be checked and adjusted to match that specified for the flow cell to be used.

Sequencing depth achieved for each sample will be a factor of (1) the choice of flow cell and sequencing platform and (2) the number of samples included on each run. Increasing sequencing depth will aid the recovery of rare taxa, especially in high diversity systems, but increases the per-sample cost. Most projects typically aim for a sequencing depth of between 50,000 and 200,000 reads per sample, but there

[9] https://emea.illumina.com/content/dam/illumina-marketing/documents/products/other/miseq-overclustering-primer-770-2014-038.pdf

[10] https://support.illumina.com/bulletins/2016/07/what-is-nucleotide-diversity-and-why-is-it-important.html

are occasions when either shallower or deeper sequencing would be appropriate, depending on the sampling design, number of biological, PCR and sequencing replicates, and the importance of detecting rare species.

The Illumina MiSeq platform remains the most commonly used for metabarcoding and is well suited for most monitoring or research applications. For very large projects (i.e. hundreds or thousands of samples) or when much deeper sequencing is required, cost efficiencies may be gained through the use of higher-throughput platforms including the Illumina HiSeq, NextSeq, and NovaSeq. A comparison of costs per sample, sequence throughputs, and error rate among high-throughput sequencing platforms can be found in (Piper et al. 2019). Many commercial labs offer high throughput sequencing services, so sequencing can be outsourced, often along with some stages of library preparation, which can deliver better quality data through the use of highly-trained, specialist staff.

8.4 Special considerations for metabarcoding aquatic eDNA samples

Although metabarcoding can be applied to DNA extracted from any type of sample, the sample type still influences decisions to be made at several points in the workflow (Deiner et al. 2017b). In particular, metabarcoding of eDNA samples needs to take into account the very low concentration of target DNA in the sample. This makes PCR highly stochastic, meaning that rare targets will often fail to amplify in a given reaction. A much higher number of PCR replicates is required to recover the full community present in the sample than is the case when working with high concentration DNA obtained from bulk samples or even from sediment samples (Doi et al. 2019). Increasing the volume of DNA extract in the reaction can also help to reduce stochasticity, although this simultaneously increases the concentration of inhibitors, which could reduce PCR efficiency, so that samples should be screened for inhibition first, using the same volume of DNA as will be used in the metabarcoding PCR.

8.5 Use of alternative sequencing platforms

While the above text focuses on the use of Illumina platforms for metabarcoding, Oxford Nanopore Technology devices (particularly the MinION) have also been used in several recent metabarcoding studies to detect target species of toxic microalgae (Hatfield et al. 2020), bivalve invasive species (Egeter et al. 2020) and even sharks (Truelove et al. 2019) and macroinvertebrate communities (Baloğlu et al. 2021). This technology is highly portable, does not require an especially stable bench (i.e. can be used at sea on a ship in motion) with a fast turn-around sample processing time. Furthermore it generates longer reads and combined with field-friendly DNA extraction and PCR methods (portable labs e.g. BentoLab www.bento.bio) and is thus poised to provide powerful point-of-care detection capabilities. Although, the use of Nanopore Technology remains some way off in terms of routine operationalisation because the principal disadvantage remains with the relatively high sequencing error rate, and it is slightly more expensive than MiSeq metabarcoding to obtain the same sequencing depth.

The PacBio Sequel platforms also offer long read sequencing capabilities and can be used for metabarcoding. Although the reported error-rate is higher than that in Illumina sequencing, this can be corrected through the use of consensus sequences (Reuter et al. 2015), and there are significant advantages to be gained in taxonomic resolution from the use of longer amplicons. Moreover, for barcode regions that exhibit length variation (e.g. the ITS genes, which are commonly used for fungal and plant metabarcoding), there appears to be a far less length bias in PacBio sequencing than is typical of the Illumina platforms, which are biased towards the shorter amplicons within a sample (Castaño et al. 2020). The principal dis-

advantage of PacBio sequencing is cost, which is over 20 times greater than the equivalent sequencing depth on an Illumina MiSeq platform.

8.6 Bioinformatics

We do not explicitly cover bioinformatics in this document, but we do emphasise the importance of using a well-designed bioinformatics pipeline (a chain of command-line tools and custom scripts) that has been optimised for the specific marker, target group and use case. Even from the same raw sequence data, choice of bioinformatics parameters can make the difference between results that are fit for purpose and those that are not.

For a given metabarcoding project, it will be important to ensure that all samples are processed with the exact same bioinformatics pipeline, and it is particularly important to consider the need to link together datasets generated from different sequencing runs. This is especially relevant for taxonomic groups and markers with incomplete reference databases, meaning taxa cannot be linked based on species names and may influence the choice between use of OTUs (operational taxonomic units) and ESVs (exact sequence variants; Callahan et al. 2016). The OTU approach overcomes PCR and sequencing errors by clustering together highly similar sequences, with the most dominant sequence from each cluster used for taxonomic assignment. In contrast, ESVs keep each unique sequence separate but filter out likely PCR and sequencing errors based on built-in error models. While overall ecological patterns derived from metabarcoding data tend to be fairly robust to choice of approach (Glassman and Martiny 2018), ESVs are more reproducible, and therefore more cross-comparable where linking relies on sequence identity rather than species names (Callahan et al. 2017). Other authors have provided detailed reviews of these terminologies (Glassman and Martiny 2018, Porter and Hajibabaei 2020).

Choice of taxonomic assignment method and taxon acceptance thresholds (i.e., the number or proportion of sequence reads required for an OTU/ESV to be retained in the final dataset) can make a material difference to results obtained. Optimal parameter choices will depend on the characteristics of the marker used, the completeness of the reference database, and the purpose for which the data is to be used. For instance, if the aim is to assess overall ecological patterns then more aggressive filtering may be chosen to reduce noise and there is a relatively low cost to inaccurate taxonomic identification. However, if the aim is to detect invasive or endangered species, even very weak detections should be retained and species need to be identified with a high degree of accuracy.

There is considerable interest in the extent to which metabarcoding data can be used to determine intraspecific genetic diversity by differentiating between haplotypes. While the methods mentioned above go some way to minimise the impact of PCR and sequencing errors on metabarcoding data, precise choice of parameters can have a significant influence on effectiveness, and there is usually a tradeoff between minimising these errors and retaining true diversity, especially for low-abundance taxa. Therefore, extreme caution should be applied in interpreting sequence variants that match to the same species as evidence of intraspecific variation, particularly when using a metabarcoding marker designed for species-level identification, such as those listed in **Table 5**. However, studies using alternative approaches (e.g. long range PCR (Deiner et al. 2017a), hybrid capture (Jensen et al. 2020)) have shown that genuine haplotype data can be obtained from water samples. To obtain intraspecific gene data from metabarcoding requires the use of markers in faster-evolving gene regions than typically used for metabarcoding which are informative at the population level (e.g. Tsuji et al. 2020).

Text Box 3: Ecological indices inferred from metabarcoding

Ecological indices, which are typically referred to as diversity indices, are quantitative measures that serve as statistical descriptors of biodiversity. In general, ecological indices either use presence/absence data to calculate richness, the number of unique biological units (e.g. species, functional group, etc.), or use proportional differences in abundances among biological groups to calculate diversity. There are many different abundance based ecological indices, due to the different levels of emphasis placed on rare species, but the central calculation framework is rooted in what are referred to as Hill numbers (Chao et al. 2014). For the purposes of regulated monitoring (e.g. under the Water Framework Directive), ecological indices are drastically different from typical ecological indices as they are custom designed to a set of organisms. In short, biomonitoring ecological indices are designed by assigning environmental sensitivity scores to specific individual biological categories, which are often a mixed assortment of species, genus and family level groups. The sensitivity scores themselves are based on known or observed associations/trends between specific groups and environmental parameters of interest, usually pollution intolerance (e.g. Kelly 1998, Mandaville and Soil & Water Conservation Society of Metro Halifax 2002).

Traditionally, ecological indices have been calculated from data that catalogue the occurrence of captured individuals, such that the counts, abundances or frequencies of each unique biological unit (species) are directly linked to the individuals observed. In contrast, metabarcoding derived index scores, are not directly linked to abundance of each unique biological unit, but sometimes take into account sequence read counts. Read counts do not always closely correlate with abundance (Bista et al. 2018), however comparing proportional changes between samples (e.g. sites) that utilize similar protocols can be informative in assessing changes in species and community abundances (e.g. Hänfling et al. 2016). Often datasets are normalized to their common minimum or medium sequencing depth in order to make inferences on taxon relative abundance changes and allow direct ecological comparisons (de Cárcer et al. 2011, Wangensteen and Turon 2017). It may also be possible to link phenotypic information (observable characteristics or traits of an organism) to read abundances, to infer biological abundances across phylogenetically related groups (e.g. Yates et al. 2020), however this may not be widely applicable across traditional biomonitoring groups. Because of the abundance dislink between traditional and metabarcode derived ecological data, the calculation of existing ecological indices from metabarcoding data needs to be considered carefully.

Inferrering community ecological indices from metabarcoding data can be relatively straightforward to adapt, at least from an analysis standpoint. Taxon richness can be easily calculated from metabarcoding data (Seymour et al. 2021), while abundance based indices, including Shannon and Simpson indices, can be derived by calculating Hill numbers from relative read abundances using an index based approach to existing statistical methods (Chao et al. 2014, Alberdi and Gilbert 2019). Biomonitoring indices are inherently more specific in their methodology, meaning each index needs to be assessed individually for whether a simple presence absence conversion of the metabarcoding derived community is sufficient to calculate the index or if a relative read conversion can be used to account for abundance based metrics (Seymour et al. 2020). A key caveat with interpreting biomonitoring indices is that the biomonitoring scores are often based on abundance frequencies linked to traditional sampling methods, which are known to differ from metabarcoding, particularly eDNA (Valentini et al. 2016, Tapolczai et al. 2019, Seymour et al. 2021). The inherent difference in biomonitoring scores between traditional and metabarcode based data may therefore correlate, but can differ when interpretations are made regarding the environmental status of a given site if not scaled properly (Tapolczai et al. 2019, Seymour et al. 2020). Alternatively, current efforts in moving

away from taxon specific bioscores to realtime biodiversity and environmental analyses via interactive network analyses are looking to alleviate the difficulties in assigning scores to an increasing list of environmentally sensitive organisms, many previously unnoticed using traditional methods (Cordier et al. 2017, Seymour et al. 2020).

Summary: Metabarcoding to survey biological communities

Key takeaways:

- Design or selection of primers is critical for maximising the completeness of data for the target group. There is a trade-off between the taxonomic breadth of the group targeted by the primer set and the completeness of the assessment obtained, particularly in terms of the probability of detecting rare species. Shorter amplicons being more sensitive but less able to separate closely related species.
- To enable sequences to be bioinformatically sorted back to the samples they came from, nucleotide tags and/or indexes are added to the amplicons along with sequencing adaptors that bind the amplicons to the flow cell for sequencing. There are three main approaches to carry out labelling and library preparation in metabarcoding studies. They all have trade-offs in terms of amplification efficiency, cost, time and contamination risk.
- Low template concentration increases the stochasticity of PCR, so a greater number of PCR replicates is advised when working with aquatic eDNA samples.
- Polymerase needs to be selected depending on the primers used. Different polymerases have different GC content preferences and some are incompatible with certain degenerate bases in the primer sequences.
- The quality of sequencing data can be reduced by factors such as lack of nucleotide diversity, inaccurate quantification and normalisation of libraries, incomplete removal of primer dimers and other impurities.
- Many choices are made in the bioinformatic processing of sequencing data, and these need to be kept consistent across all samples analysed within the same project or dataset to ensure comparability of results. Key choices such as whether or not to cluster sequences into OTUs, thresholds for detection, and taxonomic assignment methods will depend on factors such as target group, management context or research question, and the completeness of reference databases.

Priorities for future research:

- Design of new primers to detect different taxonomic groups while minimising non-target amplification and maximising taxonomic resolution at species level.
- Design of primers to characterise intraspecific variation within particular species or species groups.
- Further understanding of how factors such as polymerase choice and number of PCR cycles affects metabarcoding results.
- Further understanding of how different library preparation methods affect the sensitivity of metabarcoding analyses to detect low-abundance taxa.
- Use of alternative, more portable sequencing platforms for metabarcoding (e.g. MinION)

9. SUMMARY AND ADVICE

This practical guide set out to summarize the current state of the art for the field and laboratory workflow that is used for DNA based methods of species detection and monitoring. While the process is complex and at each step there are many decisions to be made, we hope to promote greater understanding of the inherent considerations, trade-offs, and uncertainties so that good choices can be made. This section distills the main advice and steps for moving forward and is discerned from our collective knowledge presented in the guide. We conclude here with the main factors that influence methodological choices, suggestions for how to report results and share raw data in a consistent way. **We end with an open invitation to the community to help us keep this guide as up to date as possible.** The relevance of these methods for applied biomonitoring has already seen a huge increase in both monetary and time investment in solving the outstanding challenges, so we expect that many of these will be addressed within the coming months and years.

9.1 Factors influencing methodological choices

Regardless of sample type and analysis method, DNA-based monitoring of biodiversity is a complex process made up of many interdependent steps, each of which requires optimisation and incorporates choices and trade-offs. It is important to emphasize that given sufficient optimisation, highly reliable and replicable results can be obtained, and this is often robust to different choices being made at certain steps.

Choices made at each step will be influenced by multiple factors, including:

- **Data requirements for the intended use-case**. This is the single most important determinant of methodological choices, since the first priority is to ensure that the data are fit for purpose. Different applications carry different management costs of failure to detect rare species, and this will influence decisions from sampling design and intensity to technical replication in DNA extraction and PCR, stringency of the quality control framework, sequencing depth and bioinformatic filtering stringency.
- **Practicality and logistics**. This will primarily affect sample collection and preservation, with knock-on implications for DNA extraction. Important factors include the accessibility of the sampling locations and number of individual field teams sampling in parallel, the availability of cold-storage facilities, and the ability to safely handle and transport flammable liquids.
- **Cost**. Budget is a major consideration in almost all monitoring programmes. It principally affects the number of samples that can be collected and analysed, and places the emphasis on generating the maximum amount of information from a given number of samples. At first glance, there are many ways in which the costs of a DNA-based monitoring programme can be reduced, but many of these are ultimately false economies and all have tradeoffs that need to be evaluated in the context of the specific programme.

In applied DNA-based biomonitoring, it is vital to provide consistent documentation of methodological choices made and the results of quality control tests conducted throughout the workflow. This helps to provide confidence in the results obtained or to flag results that may be less reliable, and enables assessment of the comparability of results obtained using different workflows.

The exact data to report from any given workflow covered in this guide will vary depending on sample and analysis type. Minimum reporting standards will also vary depending on the goal of the work, which

Table 6. Options for reducing the cost of DNA-based monitoring programmes.

	Cost saving option	Sample / analysis type	Considerations	Impact on budget
Field	Filter a larger volume of water per filter	Aquatic eDNA	May not be a viable option in systems with high suspended solids. Multiple samples still needed to account for uneven distribution of eDNA, especially in lentic habitats	Small in most cases, although may have a larger impact in certain environments (e.g. marine). Large volume samplers are likely to be more expensive so cost saving is only likely to be significant at large scales
Field	Pool subsamples so that each sample is representative of a larger area	All	Lose spatial resolution & decrease probability of detecting rare species. Reduces options to analyse frequency / occupancy or conduct power analysis	Medium to high. A common way to maximise the number of different environments that can be sampled for a given budget
Lab	Extract samples individually but then pool aliquots of DNA extracts for initial analysis. Store remaining DNA extracts individually.	All	Gives you the option to go back to analyse the individual samples for additional insights but start at a coarser level	Medium to high. Still incurs the costs of sample collection and DNA extraction for each sample, but initial analysis cost could be substantially reduced
Lab	Reduce number of PCR replicates	All, but especially aquatic eDNA, for which more PCR replicates are recommended	Increases stochasticity and reduces detection probability	Small. This is usually a false economy since the reduced detection probability usually has to be compensated for by collecting and analysing more replicate samples, which is more expensive overall
Lab	Use a higher-throughput sequencing platform such as Illumina Hi Seq or NovaSeq	Metabarcoding	Requires a large number of indexes for sample multiplexing	Small or negative unless very large numbers of samples give economies of scale. Can slow down the analysis as need to wait for enough samples to be accumulated
Lab	Reduce sequencing depth	Metabarcoding	Reduces detection probability of rare species	Usually small. Sequencing typically only represents a small proportion of the overall cost per sample, and this can be a false economy similar to reducing the number of PCR replicates, requiring more biological sample replicates
Lab	Reducing quality control testing	All	Lowers confidence in output and reduces the opportunity to identify specific steps that could have compromised the quality of results	Small to medium. Not recommended to eliminate the QC steps marked as mandatory in **Table 4**, as this will be a false economy if work needs to be repeated

may range from exploratory surveys to publication in scientific journals and increasingly the provision of evidence meant to stand up in a court of law.

For the laboratory processes covered in this guide, our starting point should be to draw from standards that have already been established in other industries that use the same types of analyses. These include MIMARKS (minimum information about a marker gene sequence) and MIxS (minimum information about any "x" sequence) specifications (Yilmaz et al. 2011) as well as the MIQE guidelines (dMIQE Group and Huggett 2020) for qPCR. The FAIRsharing Collections represents a valuable repository of standards and databases, and includes a standard on the minimum information for biological and biomedical investigations (https://fairsharing.org/collection/MIBBI).

Several efforts have been made to adapt these broader molecular standards for eDNA applications, including the establishment of minimum information criteria for eDNA analysis from water samples (Goldberg et al. 2016), determination of LOD and LOQ for eDNA qPCR assays (Klymus et al. 2020) and the development of a framework for validating single species assays, which covers the entire workflow from field to data interpretation (Thalinger et al. 2021b). More broadly, the ten simple rules for reproducible computational research (Sandve et al. 2013) have been advocated as best practice standards in DNA based species monitoring (Deiner et al. 2017b).

For field sampling steps, however, there is little in the way of standardised reporting requirements beyond best practice for reporting in scientific publications. Goldberg et al. (2016) provided a list of field variables that should be reported as standard for detection of species' DNA from aquatic samples, but these require adaptation for other environments and sample types and will need to be tailored for routine monitoring contexts as opposed to scientific research.

In Text Box 4 we provide an example of key information that should be recorded for all aquatic eDNA samples at the field collection stage and during analysis with a metabarcoding pipeline.

Text Box 4: Example of minimum reporting requirements for eDNA metabarcoded water samples

Note that in commercial settings, some methodological details may be commercially sensitive. These do not necessarily need to be included in reports but providers should ensure that all details are documented and securely stored internally in case they should be required for validation or verification purposes.

In the field:

1. Unique name for sample
2. Type of waterbody (e.g. pond/lake/river/estuary/ocean)
3. Coordinates, coordinate system, date and time
4. Weather conditions at the time of sampling
5. Details of the person who collected the sample
6. Is this sample a Filter Negative Control?
7. Sampling strategy (single point sample/merged subsamples/ntegrated sample).
8. Water depth at which sample collected
9. Details of filter (membrane material/pore size/enclosed, housed or open)
10. Time elapsed between sampling & filtration, and storage conditions during this time
11. Volume of water filtered
12. Preservation method

13. Any physical data associated with the site (e.g. temperature, pH, water depth etc)
14. Other observations (e.g. turbidity, presence of leaf-litter, livestock or obvious chemical pollutants)

In the lab:

15. Total DNA concentration for samples (including negative controls)
16. Results of inhibition test and any other IPCs (e.g. testing extraction efficiency)
17. Barcode gene used, and length of target amplicon
18. Primers, indexes and library preparation protocol
19. Number of PCR replicates performed
20. PCR success and details of how this was determined
21. Performance of negative controls in PCR
22. Details of positive controls used and their performance in PCR
23. Details of any spike-in added to achieve heterogeneity
24. Final concentration of libraries loaded onto flow cell

Bioinformatics

25. Quality statistics related to the sequencing run
26. The number of samples sequenced together on the run
27. For each sample, total number of sequence reads obtained, the number retained after each stage of quality filtering, and the number assigned to the target group.
28. Use of OTUs or ASVs and the method and parameters selected
29. OTU / ASV acceptance thresholds applied
30. Reference database used
31. Taxonomic assignment method and parameters selected
32. Details of data excluded from final results (e.g. livestock species or non-target taxa)
33. Raw FASTq files securely stored and bioinformatic pipeline script fully documented such that it can be assessed by a third party such as a court of law.

9.3 Concluding remarks

This practical guide to DNA-based biomonitoring was a collaboration of many people across the spectrum of basic research to applied contexts and represents the state of advice and knowledge at the time of writing. We acknowledge that at the time of your reading this resource, some of the hurdles and challenges we highlight may have been overcome and there may be new evidence for ways forward that may even contradict the guidance provided here. We intend to update this practical guide periodically, but recommend that readers and users should keep in mind that rapid advancements will be made and it is always a good idea to discuss the current trends and read the newest studies. Please consider the authors of this guide as a resource and do not hesitate to reach out and discuss with us!

Thus, we view the practical advice collected here as to the best of our collective knowledge at time of publication. The chosen publication style as an electronic book format will allow us to update the guide to new versions with ease. We welcome further input from the broader scientific and applied communities of practice in this area to help us keep this document as up to date as possible and contribute to future editions so that it remains relevant and useful.

10. GLOSSARY

Amplicon: A section of DNA that has been amplified through a reaction such as **PCR**. It can also be termed **PCR product**.

Amplification: The process of creating copies of a particular region of DNA (the **amplicon**), usually through a PCR reaction using **primers** and enzymes such as polymerases. **Non-target amplification** refers to the unintended amplification of DNA from taxa that the primers were not designed to amplify (e.g. amplification of bacteria by primers designed to target metazoans).

Buffer: Liquid solutions used to maintain a stable pH, as they can neutralize small quantities of additional acid of base. For examples of buffers commonly used for preservation of DNA see **Table 3**.

Barcoding: Taxonomic identification of a species based on the DNA sequencing of a short gene region that shows variation at the species level. This is known as a **barcode** region (also referred to as a **marker** region). Sequences obtained are compared against a **reference database** (e.g. BOLD www.boldsystems.org) to assign taxonomy.

Bioinformatics: Computational processing of sequence data. A core element of DNA metabarcoding pipelines, in which **high-throughput sequencing** data are quality filtered, summarised and compared against reference databases for taxonomic assignment, yielding a taxon-by-sample table that can be subjected to ecological analysis. A **bioinformatics pipeline** describes a script linking together a chain of software programmes that perform the various steps of data handling.

Bulk sample: A mixed community sample of organisms or their tissues such as would be collected in a net or trap or extracted from an **environmental sample** (e.g. by sieving soil or sediment samples).

CEN standard: CEN (Comité Européen de Normalisation) provides a platform for the development of European Standards and Technical Reports.

Clustering: A step in the bioinformatics process in which highly similar sequence reads are grouped together to form a cluster of highly similar reads that putatively originate from the same species. Sequence variants within a cluster include both real intraspecific genetic variation and sequences that contain errors introduced during PCR and sequencing. See also the definition for **OTU**.

ddPCR: Droplet Digital PCR. A platform that can be used for targeted detection of species (alternative to **qPCR**). A highly sensitive fluorescent probe based approach in which each reaction takes place in 20,000 individual droplets enabling absolute quantification of target DNA copy number without the need for standard curves. See section **7.1.3.**

eDNA / Environmental DNA. DNA isolated from an **environmental sample** such as water or sediment. May include both **Organismal DNA** derived from whole organisms in the sample and **extra-organismal DNA** which is captured separately from the organism it originated from. Extra-organismal DNA may be in the form of cells, organelles, or free-floating DNA, originating from sources such as shed skin, scales, blood, mucus, faeces, urine, saliva and gametes.

DNA extraction: Isolation of DNA from a sample, using chemical methods. The DNA extraction usually incorporates steps to remove impurities from the DNA.

Environmental sample: A sample of an environmental medium, such as seawater, freshwater, soil or air.

ESV /ASV: Exact Sequence Variant or **Amplicon Sequence Variant** (broadly synonymous) are generated in the **bioinformatics pipeline** and represent individual high quality sequences in metabarcoding datasets. Can be used for taxon delimitation as an alternative to clustering into OTUs, with sequences that contain errors filtered out using denoising algorithms.

Filter: Membrane filter for the capture of eDNA constructed out of a wide range of synthetic materials, with specific pore sizes. See section **2.3** for further details.

High-throughput sequencing (HTS): DNA sequencing technology that produces millions of DNA sequence reads in parallel. Enables thousands of different organisms from a mixture of species to be sequenced at once, to obtain community data from a single analysis (i.e. metabarcoding). Various different platforms exist, but the most commonly used is Illumina's MiSeq. Also known as Next-Generation Sequencing (NGS) or parallel sequencing. In contrast, the classic **Sanger sequencing** method produces one sequence at a time and is not suitable for mixed-species samples.

Indexing: Also known as sample **multiplexing**. Allows multiple samples to be pooled on one **high-throughput sequencing** run, by adding a short sequence of nucleotide base pairs to each sample during library preparation. This sequence is different for each sample on the run and enables sequences to be assigned back to the sample they came from after sequencing (known as demultiplexing; see **Figure 7**).

Inhibition: Certain chemical compounds can reduce the efficiency of PCR amplification, or in some cases cause it to fail completely. This can lead to false negative results (i.e. non-detection when a species' DNA is in fact present in the sample). **Inhibitors** may be present in the original sample (eg. in the form of tannins or humic acids) or may be added during sample processing or DNA extraction (eg. SDS, ethanol). Internal positive **controls** can be used to check for the presence of inhibition (see also **section 6.2** on inhibition testing), and inhibitors can usually be removed through purification kits or dilution of the DNA.

ISO standard: The International Organization for Standardization is an international standard-setting body that promotes worldwide proprietary, industrial, and commercial standards.

Library: A molecular biology protocol through which DNA is prepared for sequencing on a **high throughput sequencing** platform. In the case of Illumina **metabarcoding**, this includes PCR amplification of the target DNA region, labelling of samples with unique nucleotide tags so that they can be multiplexed (pooled together for sequencing and bioinformatically separated after sequencing), and the addition of sequencing adaptors so that the DNA can bind to the Illumina flow cell. **Metabarcoding library preparation** can be done following various different approaches (see section **8.2.1** on amplicon library preparation).

Metabarcoding: Taxonomic identification of multiple species simultaneously from a complex (multi-species) sample, using high-throughput **amplicon** sequencing of a standardized gene fragment (e.g. COI). See section **8**.

Metagenomics: The study of genomes recovered from a mixed community of organisms or from environmental samples. Metagenomics usually refers to the study of microbial communities but has also been applied to invertebrate faunal collections.

Mock community: A species community of known composition, usually assembled for use as a positive control. See also section **6.3** on analytical controls used for DNA and eDNA analyses.

Negative control: A negative control is used to check for potential contamination. A **negative site control** refers to a sample collected from a field site where the target taxon is known to be absent. A **negative filtration control** is a sample where DNA-free water is filtered alongside the eDNA samples to check that DNA is not transferred between samples. **Negative laboratory controls** consist of DNA-free samples processed alongside the test samples at each stage of the process to check for (cross-)contamination. In the context of **DNA extraction**, a negative control should not contain a DNA template and in the context of **PCR**, a negative control should not give **amplicons**. See also section **6.3** on analytical controls used for DNA and eDNA analyses.

OTU: Operational Taxonomic Units (OTUs) are proxies for species obtained using **clustering** algorithms to **bioinformatically** process sequencing data obtained from **metabarcoding**. Reads are **clustered** using a sequence similarity threshold (e.g. most often 97%). OTUs are not easily comparable across studies as they depend on the dataset in which they were created.

PCR: Stands for **polymerase chain reaction**. PCR is a method that uses thermal cycling (cyclical variations in temperature) in the presence of a polymerase enzyme to rapidly create millions of copies of a predefined DNA fragment. **Primers** are designed to bind to the DNA of the target group at either end of the chosen DNA fragment and the polymerase creates a copy of the DNA sequence between them. PCR (also termed DNA **amplification**) is a prerequisite for most forms of DNA sequencing and can be used as a diagnostic tool in itself to detect the presence of particular species when species-specific **primers** are used (e.g using **qPCR** or **ddPCR**). During thermal cycling, a series of repeated temperature changes are performed, which variously cause (1) DNA denaturation in which double-stranded DNA separates into single strands, (2) primer annealing where the primers bind onto single-stranded DNA, and (3) elongation where the polymerase synthesises DNA starting from the forward and reverse primers. This series of temperatures is repeated a predetermined number of times (termed **PCR cycles**), with the amount of target DNA doubling with each cycle leading to exponential amplification.

Positive control: A positive control is a sample that is expected to produce a known positive result, and is analysed alongside test samples to check that the analytical process is working as it should (e.g. a **mock community** can be used as a positive control during metabarcoding). **Positive laboratory controls** consist of a known concentration of pre-prepared DNA of one or more species that are expected to be amplified efficiently using the selected PCR protocol. See also section **6.3** on analytical controls used for DNA and eDNA analyses.

Primer or **oligonucleotide:** Short, single-stranded nucleic acid molecule (typically 20 bp or longer) consisting of a sequence of DNA bases that are designed to match the target DNA at a particular point in the genome. PCR usually requires a pair of primers (or **primer set**), one matching the target DNA at either end of the barcode region to be amplified. **Primer mismatches** occur when the primer sequence does not exactly match the target sequence, and this can reduce PCR efficiency or cause false negative results. **Degenerate primers** consist of a mixture of primer sequences that incorporate some variation at certain

base positions so that the primers can bind to more variable target DNA (e.g. a broader taxonomic group) with minimal mismatches.

Probe: Usually refers to hydrolysis probes used in **qPCR** and **ddPCR**. Probes are DNA **oligonucleotides** designed to bind to the target DNA in a location between the PCR **primers**. The probe contains a fluorescent label, which is suppressed until PCR occurs, when it is released. The fluorescence emitted is detected by the instrument and used as a measure of target DNA amplification. Probe-based qPCR and ddPCR require both primers and the probe to bind to the target DNA in order for amplification to be detected, and this increases the **specificity** of assays.

qPCR: Quantitative polymerase chain reaction is a **PCR** reaction that quantifies DNA by means of a fluorescent dye that is measured by a fluorometer in real time throughout the amplification process. Information about relative and absolute amounts of DNA present can be inferred with the use of appropriate standard curves. See section **7.1.2.**

Reference database: A library of DNA sequences derived from specimens of known identity. Sequence data obtained from test samples (e.g. via **metabarcoding**) can be matched against a reference database to assign taxonomic names to the sequences. The Barcode of Life Database (BOLD) is specifically developed for DNA **barcoding** and is highly curated but contains a limited selection of barcode genes. The NCBI database (also known as Genbank) is far more extensive but is not curated and contains a high level of error that must be accounted for in taxonomic assignment pipelines. Custom reference databases can also be made for particular projects to ensure that important species can be confidently identified.

Replicates: Repeat or duplicate samples / analyses used to test repeatability and measure variation, and to improve detection probability by overcoming stochasticity. **Sample replicates** (sometimes called biological replicates) refer to samples collected at the same time and location. **Technical replicates** are repetitions of the same analysis on the same sample - this can include **extraction replicates** where the sample is subdivided and multiple separate DNA extractions carried out, and **PCR replicates**, where the same PCR reaction is applied to multiple subsamples of a single DNA extract.

Sanger sequencing: A method of DNA sequencing developed by Frederick Sanger in 1977, also termed the Sanger 'chain termination' method. The method is based on the incorporation of radioactively or fluorescently labelled chain terminating nucleotides. It produces a single DNA sequence for each reaction, unlike high-throughput sequencing which can produce millions. In the field of DNA-based bioassessment, it is most commonly used for DNA barcoding to identify single specimens and the creation of reference barcodes from specimens of known identity.

Sequence read: a sequence of nucleotide bases (A,G,T,C) representing a DNA fragment. An Illumina MiSeq run generates around 30 million reads, each originating from an individual DNA fragment that was bound onto the surface of a flow cell. Many copies of the same DNA fragment can originate from the same species (even from the same organism), meaning that metabarcoding datasets typically contain many identical sequence reads. The number of sequence reads obtained for a given species in a sample is known as the **read count**. Although read counts often correlate with the relative quantity of species' DNA in a sample, the quantitative interpretations must be made with caution due to technical and biological biases.

Sequencing: The process of determining the nucleotide sequence of a given DNA fragment, which enables species identification. See also definitions for **Sanger sequencing** and **high-throughput sequencing.**

Sensitivity: In diagnostics, sensitivity or true positive rate is a measure of the proportion of positives that are correctly identified. Essentially this refers to the ability of an assay to detect target DNA when it is present at very low concentrations.

Specificity: In diagnostics, specificity or true negative rate measures the proportion of negatives that are correctly identified. In **primer** design, specificity refers to the extent to which the primers (and probes where relevant) bind only to the target DNA without any non-target amplification. Specificity can be affected by the length and GC content of the primers, and by the annealing temperature used in PCR. However, the most important factor is careful primer/probe design to achieve exact complementarity to the target and multiple **primer mismatches** to related taxa that may co-occur.

Validation: A comprehensive set of experiments that evaluate the performance of an assay, including its **sensitivity, specificity,** accuracy, detection limit, range and limits of quantitation. In terms of eDNA, this also extends to field testing to check that expected results are returned under known conditions in the field.

WFD: EU water framework directive. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy.

11. REFERENCES

Abbaszadegan M, Huber MS, Gerba CP, Pepper IL (1993) Detection of enteroviruses in groundwater with the polymerase chain reaction. Applied and Environmental Microbiology 59: 1318–1324. https://doi.org/10.1128/AEM.59.5.1318-1324.1993

Agasild H, Kisand A, Ainelo E, Feldmann T, Timm H, Karus K, Kisand V, Jones RI, Nõges T (2018) Chironomid incorporation of methane-derived carbon in plankton- and macrophyte-dominated habitats in a large shallow lake. Freshwater Biology 63: 1433–1445. https://doi.org/10.1111/fwb.13170

Agersnap S, Larsen WB, Knudsen SW, Strand D, Thomsen PF, Hesselsøe M, Mortensen PB, Vrålstad T, Møller PR (2017) Monitoring of noble, signal and narrow-clawed crayfish using environmental DNA from freshwater samples. PloS One 12: e0179261. https://doi.org/10.1371/journal.pone.0179261

Alberdi A, Gilbert MTP (2019) A guide to the application of Hill numbers to DNA-based diversity analyses. Molecular ecology resources 19: 804–817. https://doi.org/10.1111/1755-0998.13014

Alberdi A, Aizpurua O, Gilbert MTP, Bohmann K (2018) Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods in Ecology and Evolution 9: 134–147. https://doi.org/10.1111/2041-210x.12849

Allan EA, Zhang WG, Lavery A, Govindarajan A (2020) Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. Environmental DNA. https://doi.org/10.1002/edn3.141

Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (2009) A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes. PloS One 4. https://doi.org/10.1371/journal.pone.0006372

Apothéloz-Perret-Gentil L, Cordonier A, Straub F, Iseli J, Esling P, Pawlowski J (2017) Taxonomy-free molecular diatom index for high-throughput eDNA biomonitoring. Molecular Ecology Resources 17: 1231–1242. https://doi.org/10.1111/1755-0998.12668

Apothéloz-Perret-Gentil L, Bouchez A, Cordier T, Cordonier A, Guéguen J, Rimet F, Vasselon V, Pawlowski J (2020) Monitoring the ecological status of rivers with diatom eDNA metabarcoding: a comparison of

taxonomic markers and analytical approaches for the inference of a molecular diatom index. Molecular Ecology, 30: 2959–2968. https://doi.org/10.1111/mec.15646

Arribas P, Andújar C, Hopkins K, Shepherd M, Vogler AP (2016) Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. Methods in Ecology and Evolution 7: 1071–1081. https://doi.org/10.1111/2041-210x.12557

Aylagas E, Borja A, Rodriguez-Ezpeleta N (2014) Environmental Status Assessment Using DNA Metabarcoding: Towards a Genetics Based Marine Biotic Index (gAMBI). PloS One 9: e90529. https://doi.org/10.1371/journal.pone.0090529

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: 194–202. https://doi.org/10.1016/j.ecolind.2018.07.044

Bailet B, Bouchez A, Franc A, Frigerio J-M, Keck F, Karjalainen S-M, Rimet F, Schneider S, Kahlert M (2019) Molecular versus morphological data for benthic diatoms biomonitoring in Northern Europe freshwater and consequences for ecological status. Metabarcoding and Metagenomics 3. https://doi.org/10.3897/mbmg.3.34002

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

Baker M (2012) Digital PCR hits its stride. Nature Methods 9: 541–544. https://doi.org/10.1038/nmeth.2027

Baloğlu B, Chen Z, Elbrecht V, Braukmann T, MacDonald S, Steinke D (2021) A workflow for accurate metabarcoding using nanopore MinION sequencing. https://doi.org/10.1111/2041-210X.13561

Barnes MA, Turner CR (2015) The ecology of environmental DNA and implications for conservation genetics. Conservation Genetics 17: 1–17. https://doi.org/10.1007/s10592-015-0775-4

Battin TJ, Besemer K, Bengtsson MM, Romani AM, Packmann AI (2016) The ecology and biogeochemistry of stream biofilms. Nature reviews. Microbiology 14: 251–263. https://doi.org/10.1038/nrmicro.2016.15

Beentjes KK, Speksnijder AGCL, Schilthuizen M, Hoogeveen M, Pastoor R, van der Hoorn BB (2019) Increased performance of DNA metabarcoding of macroinvertebrates by taxonomic sorting. PloS One 14: e0226527. https://doi.org/10.1371/journal.pone.0226527

Bessey C, Jarman SN, Berry O, Olsen YS, Bunce M, Simpson T, Power M, McLaughlin J, Edgar GJ, Keesing J (2020) Maximizing fish detection with eDNA metabarcoding. Environmental DNA 2: 493–504. https://doi.org/10.1002/edn3.74

Biggs J, Ewald N, Valentini A, Gaboriaud C, Dejean T, Griffiths RA, Foster J, Wilkinson JW, Arnell A, Brotherton P, Williams P, Dunn F (2015) Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (*Triturus cristatus*). Biological Conservation 183: 19–28. https://doi.org/10.1016/j.biocon.2014.11.029

Bista I, Carvalho GR, Walsh K, Seymour M, Hajibabaei M, Lallias D, Christmas M, Creer S (2017) Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. Nature Communications 8: 14087. https://doi.org/10.1038/ncomms14087

Bista I, Carvalho GR, Tang M, Walsh K, Zhou X, Hajibabaei M, Shokralla S, Seymour M, Bradley D, Liu S, Christmas M, Creer S (2018) Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. Molecular Ecology Resources. https://doi.org/10.1111/1755-0998.12888

Blackman R, Constable D, Hahn C, Sheard A, Durkota J, Hänfling B, Handley LL (2017) Detection of a new non-native freshwater species by DNA metabarcoding of environmental samples – first record of Gammarus fossarum in the UK. Aquatic Invasions 12: 177–189. https://doi.org/10.3391/ai.2017.12.2.06

Blackman RC, Ling KKS, Harper LR, Shum P, Hänfling B, Lawson-Handley L (2020) Targeted and passive environmental DNA approaches outperform established methods for detection of quagga mussels, in flowing water. Ecology and Evolution 10: 13248–13259. https://doi.org/10.1002/ece3.6921

Blaxter ML, De Ley P, Garey JR, Liu LX, Scheldeman P, Vierstraete A, Vanfleteren JR, Mackey LY, Dorris M, Frisse LM, Vida JT, Thomas WK (1998) A molecular evolutionary framework for the phylum Nematoda. Nature 392: 71–75. https://doi.org/10.1038/32160

Board DA (2000) Quality assurance standards for forensic DNA testing laboratories. Forensic Sci Commun, 2(3).

Bohmann, K, Elbrecht, V, Carøe, C, Bista, I, Leese, F, Bunce, M, Yu, DW, Seymour, M, Dumbrell, AJ and Creer, S (2021), Strategies for sample labelling and library preparation in DNA metabarcoding studies. Molecular Ecology Resources. Accepted Author Manuscript. https://doi.org/10.1111/1755-0998.13512

Bourlat SJ, Haenel Q, Finnman J, Leray M (2016) Preparation of Amplicon Libraries for Metabarcoding of Marine Eukaryotes Using Illumina MiSeq: The Dual-PCR Method. Methods in Molecular Biology 1452: 197–207. https://doi.org/10.1007/978-1-4939-3774-5_13

Boussarie G, Bakker J, Wangensteen OS, Mariani S, Bonnin L, Juhel J-B, Kiszka JJ, Kulbicki M, Manel S, Robbins WD, Vigliola L, Mouillot D (2018) Environmental DNA illuminates the dark diversity of sharks. Science Advances 4: eaap9661. https://doi.org/10.1126/sciadv.aap9661

Bozinovic F, Pörtner H-O (2015) Physiological ecology meets climate change. Ecology and Evolution 5: 1025–1030. https://doi.org/10.1002/ece3.1403

Braid MD, Daniels LM, Kitts CL (2003) Removal of PCR inhibitors from soil DNA by chemical flocculation. Journal of microbiological methods 52: 389–393. https://doi.org/10.1016/s0167-7012(02)00210-5

Brandhagen MD, Just RS, Irwin JA (2020) Validation of NGS for mitochondrial DNA casework at the FBI Laboratory. Forensic science international. Genetics 44: 102151. https://doi.org/10.1016/j.fsigen.2019.102151

Brandt MI, Trouche B, Henry N, Liautard-Haag C, Maignien L, de Vargas C, Wincker P, Poulain J, Zeppilli D, Arnaud-Haond S (2020) An Assessment of Environmental Metabarcoding Protocols Aiming at Favoring Contemporary Biodiversity in Inventories of Deep-Sea Communities. Frontiers in Marine Science 7. https://doi.org/10.3389/fmars.2020.00234

Brys R, Haegeman A, Halfmaerten D, Neyrinck S, Staelens A, Auwerx J, Ruttink T (2020a) Monitoring of spatiotemporal occupancy patterns of fish and amphibian species in a lentic aquatic system using environmental DNA. Molecular Ecology. https://doi.org/10.1111/mec.15742

Brys R, Halfmaerten D, Neyrinck S, Mauvisseau Q, Auwerx J, Sweet M, Mergeay J (2020b) Reliable eDNA detection and quantification of the European weather loach (*Misgurnus fossilis*). Journal of Fish Biology. https://doi.org/10.1111/jfb.14315

Buchner D, Haase P, Leese F (2021) Wet grinding of invertebrate bulk samples – a scalable and cost-efficient protocol for metabarcoding and metagenomics. Metabarcoding and Metagenomics, 5: e67533. https://doi.org/10.3897/mbmg.5.67533

Bush A, Compson ZG, Monk WA, Porter TM, Steeves R, Emilson E, Gagne N, Hajibabaei M, Roy M, Baird DJ (2019) Studying Ecosystems With DNA Metabarcoding: Lessons From Biomonitoring of Aquatic Macroinvertebrates. Frontiers in Ecology and Evolution 7. https://doi.org/10.3389/fevo.2019.00434

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2011) Primer sequence disclosure: a clarification of the MIQE guidelines. Clinical Chemistry 57: 919–921. https://doi.org/10.1373/clinchem.2011.162958

Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. The ISME journal 11: 2639–2643. https://doi.org/10.1038/ismej.2017.119

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13: 581–583. https://doi.org/10.1038/nmeth.3869

Capo E, Spong G, Königsson H, Byström P (2020) Effects of filtration methods and water volume on the quantification of brown trout (*Salmo trutta*) and Arctic char (*Salvelinus alpinus*) eDNA concentrations via droplet digital PCR. Environmental DNA 2: 152–160. https://doi.org/10.1002/edn3.52

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences of the United States of America 108 Suppl 1: 4516–4522. https://doi.org/10.1073/pnas.1000080107

de Cárcer DA, Denman SE, McSweeney C, Morrison M (2011) Evaluation of Subsampling-Based Normalization Strategies for Tagged High-Throughput Sequencing Data Sets from Gut Microbiomes. Applied and Environmental Microbiology 77: 8795–8798. https://doi.org/10.1128/AEM.05491-11

Carew ME, Coleman RA, Hoffmann AA (2018) Can non-destructive DNA extraction of bulk invertebrate samples be used for metabarcoding? PeerJ 6: e4980. https://doi.org/10.7717/peerj.4980

Carøe C, Bohmann K (2020) Tagsteady: A metabarcoding library preparation protocol to avoid false assignment of sequences to samples. Molecular Ecology Resources 20: 1620–1631. https://doi.org/10.1111/1755-0998.13227

Carraro L, Stauffer JB, Altermatt F (2020) How to design optimal eDNA sampling strategies for biomonitoring in river networks. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/edn3.137.

Carter DJ, Vogler AP, Vane-Wright RI (1997) Notes on basic collecting techniques for morphological studies of Lepidoptera. Metamorphosis 8: 99–106.

Castaño C, Berlin A, Brandström Durling M, Ihrmark K, Lindahl BD, Stenlid J, Clemmensen KE, Olson Å (2020) Optimized metabarcoding with Pacific biosciences enables semi-quantitative analysis of fungal communities. The New Phytologist 228. https://doi.org/10.1111/nph.16731

Chao A, Chiu C-H, Jost L (2014) Unifying Species Diversity, Phylogenetic Diversity, Functional Diversity, and Related Similarity and Differentiation Measures Through Hill Numbers. Annual Review of Ecology, Evolution, and Systematics 45: 297–324. https://doi.org/10.1146/annurev-ecolsys-120213-091540

Cilleros K, Valentini A, Allard L, Dejean T, Etienne R, Grenouillet G, Iribar A, Taberlet P, Vigouroux R, Brosse S (2019) Unlocking biodiversity and conservation studies in high-diversity environments using environmental DNA (eDNA): A test with Guianese freshwater fishes. Molecular Ecology Resources 19: 27–46. https://doi.org/10.1111/1755-0998.12900

Civade R, Dejean T, Valentini A, Roset N, Raymond J-C, Bonin A, Taberlet P, Pont D (2016) Spatial Representativeness of Environmental DNA Metabarcoding Signal for Fish Biodiversity Assessment in a Natural Freshwater System. PloS One 11: e0157366. https://doi.org/10.1371/journal.pone.0157366

Cordier T, Esling P, Lejzerowicz F, Visco J, Ouadahi A, Martins C, Cedhagen T, Pawlowski J (2017) Predicting the Ecological Quality Status of Marine Environments from eDNA Metabarcoding Data Using Supervised Machine Learning. Environmental Science & Technology 51: 9118–9126. https://doi.org/10.1021/acs.est.7b01518

Corinaldesi C, Danovaro R, Dell'Anno A (2005) Simultaneous recovery of extracellular and intracellular DNA suitable for molecular studies from marine sediments. Applied and Environmental Microbiology 71: 46–50. https://doi.org/10.1128/AEM.71.1.46-50.2005

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

Creer S, Fonseca VG, Porazinska DL, Giblin-Davis RM, Sung W, Power DM, Packer M, Carvalho GR, Blaxter ML, Lambshead PJD, Thomas WK (2010) Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises. Molecular Ecology 19, Suppl 1: 4–20. https://doi.org/10.1111/j.1365-294X.2009.04473.x

Currier CA, Morris TJ, Wilson CC, Freeland JR (2018) Validation of environmental DNA (eDNA) as a detection tool for at-risk freshwater pearly mussel species (Bivalvia: Unionidae). Aquatic Conserv: Mar Freshw Ecosyst. 28: 545– 558. https://doi.org/10.1002/aqc.2869

Davidov K, Iankelevich-Kounio E, Yakovenko I, Koucherov Y, Rubin-Blum M, Oren M (2020) Identification of plastic-associated species in the Mediterranean Sea using DNA metabarcoding with Nanopore MinION. Scientific Reports 10: 17533. https://doi.org/10.1038/s41598-020-74180-z

De Barba M, Miquel C, Boyer F, Mercier C, Rioux D, Coissac E, Taberlet P (2014) DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. Molecular Ecology Resources 14: 306–323. https://doi.org/10.1111/1755-0998.12188

Deiner K, Altermatt F (2014) Transport Distance of Invertebrate Environmental DNA in a Natural River. PloS One 9: e88786. https://doi.org/10.1371/journal.pone.0088786

Deiner K, Walser J-C, Mächler E, Altermatt F (2015) Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. Biological Conservation 183: 53–63. https://doi.org/10.1016/j.biocon.2014.11.018

Deiner K, Fronhofer EA, Mächler E, Walser J-C, Altermatt F (2016) Environmental DNA reveals that rivers are conveyer belts of biodiversity information. Nature Communications 7: 12544. https://doi.org/10.1038/ncomms12544

Deiner K, Renshaw MA, Li Y, Olds BP, Lodge DM, Pfrender ME (2017a) Long-range PCR allows sequencing of mitochondrial genomes from environmental DNA. Methods in Ecology and Evolution 8: 1888–1898. https://doi.org/10.1111/2041-210x.12836

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

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

Delavaux CS, Bever JD, Karppinen EM, Bainard LD (2020) Keeping it cool: Soil sample cold pack storage and DNA shipment up to 1 month does not impact metabarcoding results. Ecology and Evolution 10: 4652–4664. https://doi.org/10.1002/ece3.6219

Deprez L, Corbisier P, Kortekaas A-M, Mazoua S, Beaz Hidalgo R, Trapmann S, Emons H (2016) Validation of a digital PCR method for quantification of DNA copy number concentrations by using a certified reference material. Biomolecular Detection And Quantification 9: 29–39. https://doi.org/10.1016/j.bdq.2016.08.002

Di Muri C, Handley LL, Bean CW, Li J, Peirson G, Sellers GS, Walsh K, Watson HV, Winfield IJ, Hänfling B (2020) Read counts from environmental DNA (eDNA) metabarcoding reflect fish abundance and biomass in drained ponds. Metabarcoding and Metagenomics 4: e56959. https://doi.org/10.3897/mbmg.4.56959

Djurhuus A, Closek CJ, Kelly RP, Pitz KJ, Michisaki RP, Starks HA, Walz KR, Andruszkiewicz EA, Olesin E, Hubbard K, Montes E, Otis D, Muller-Karger FE, Chavez FP, Boehm AB, Breitbart M (2020) Environmental DNA reveals seasonal shifts and potential interactions in a marine community. Nature Communications 11: 1–9. https://doi.org/10.1038/s41467-019-14105-1

dMIQE Group, Huggett JF (2020) The Digital MIQE Guidelines Update: Minimum Information for Publication of Quantitative Digital PCR Experiments for 2020. Clinical Chemistry 66: 1012–1029. https://doi.org/10.1093/clinchem/hvaa125

Doi H, Uchii K, Takahara T, Matsuhashi S, Yamanaka H, Minamoto T (2015) Use of droplet digital PCR for estimation of fish abundance and biomass in environmental DNA surveys. PloS One 10: e0122763. https://doi.org/10.1371/journal.pone.0122763

Doi H, Fukaya K, Oka S-I, 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: 3581. https://doi.org/10.1038/s41598-019-40233-1

Domaizon I, Savichtcheva O, Debroas D, Arnaud F, Villar C, Pignol C, Alric B, Perga ME (2013) DNA from lake sediments reveals the long-term dynamics and diversity of *Synechococcus* assemblages. Biogeosciences 10: 3817–3838. https://doi.org/10.5194/bg-10-3817-2013

Dopheide A, Xie D, Buckley TR, Drummond AJ, Newcomb RD (2019) Impacts of DNA extraction and PCR on DNA metabarcoding estimates of soil biodiversity. Methods in Ecology and Evolution 10: 120–133. https://doi.org/10.1111/2041-210x.13086

Egeter B, Veríssimo J, Lopes-Lima M, Chaves C, Pinto J, Riccardi N, Beja P, Fonseca NA (2020) Cold Spring Harbor Laboratory Speeding up the detection of invasive aquatic species using environmental DNA and nanopore sequencing. bioRxiv 2020.06.09.142521pp. https://doi.org/10.1101/2020.06.09.142521

Elbrecht V, Leese F (2017) Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. Frontiers of Environmental Science & Engineering in China 5. https://doi.org/10.3389/fenvs.2017.00011

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

Elbrecht V, Vamos EE, Meissner K, Aroviita J, Leese F (2017b) Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. Methods in Ecology and Evolution 8: 1265–1275. https://doi.org/10.1111/2041-210x.12789

Elbrecht V, Steinke D (2018) Scaling up DNAmetabarcoding for freshwater macrozoobenthos monitoring. Freshwater Biology. https://doi.org/10.1111/fwb.13220

Elbrecht V, Bourlat SJ, Hörren T, Lindner A, Mordente A, Noll NW, Schäffler L, Sorg M, Zizka VMA (2021) Pooling size sorted Malaise trap fractions to maximize taxon recovery with metabarcoding. PeerJ 9:e12177 https://doi.org/10.7717/peerj.12177

Elbrecht V, Braukmann TWA, Ivanova NV, Prosser SWJ, Hajibabaei M, Wright M, Zakharov EV, Hebert PDN, Steinke D (2019) Validation of COI metabarcoding primers for terrestrial arthropods. PeerJ 7: e7745. https://doi.org/10.7717/peerj.7745

Erdozain M, Thompson DG, Porter TM, Kidd KA, Kreutzweiser DP, Sibley PK, Swystun T, Chartrand D, Hajibabaei M (2019) Metabarcoding of storage ethanol vs. conventional morphometric identification in relation to the use of stream macroinvertebrates as ecological indicators in forest management. Ecological Indicators 101: 173–184. https://doi.org/10.1016/j.ecolind.2019.01.014

Erickson RA, Merkes CM, Mize EL (2019) Sampling Designs for Landscape-level eDNA Monitoring Programs. Integrated Environmental Assessment and Management 15: 760–771. https://doi.org/10.1002/ieam.4155

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

Ficetola GF, Boyer F, Valentini A, Bonin A, Meyer A, Dejean T, Gaboriaud C, Usseglio-Polatera P, Taberlet P (2020) Comparison of markers for the monitoring of freshwater benthic biodiversity through DNA metabarcoding. Molecular Ecology. https://doi.org/10.1111/mec.15632

Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, De Barba M, Gielly L, Lopes CM, Boyer F, Pompanon F, Rayé G, Taberlet P (2015) Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Molecular Ecology Resources 15: 543–556. https://doi.org/10.1111/1755-0998.12338

Fields B, Moeskjaer S, Friman V-P, Andersen SU, Young JPW (2020) MAUI-seq: Metabarcoding using amplicons with unique molecular identifiers to improve error correction. Molecular Ecology Resources. https://doi.org/10.1111/1755-0998.13294

Fonseca VG, Packer M, Carvalho G, Power D (2011) Isolation of marine meiofauna from sandy sediments From decanting to DNA extraction. Protocol. Available from: http://sapientia.ualg.pt/handle/10400.1/5414.

Furlan EM, Gleeson D, Hardy CM, Duncan RP (2016) A framework for estimating the sensitivity of eDNA surveys. Molecular Ecology Resources 16: 641-654. https://doi.org/10.1111/1755-0998.12483

Geerts AN, Boets P, Van den Heede S, Goethals P, Van der heyden C (2018) A search for standardized protocols to detect alien invasive crayfish based on environmental DNA (eDNA): A lab and field evaluation. Ecological Indicators 84: 564–572. https://doi.org/10.1016/j.ecolind.2017.08.068

Geller J, Meyer C, Parker M, Hawk H (2013) Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Molecular Ecology Resources 13: 851–861. https://doi.org/10.1111/1755-0998.12138

Gibson JF, Shokralla S, Curry C, Baird DJ, Monk WA, King I, Hajibabaei M (2015) Large-Scale Biomonitoring of Remote and Threatened Ecosystems via High-Throughput Sequencing. PloS One 10: e0138432. https://doi.org/10.1371/journal.pone.0138432

Giebner H, Langen K, Bourlat SJ, Kukowka S, Mayer C, Astrin JJ, Misof B, Fonseca VG (2020) Comparing diversity levels in environmental samples: DNA sequence capture and metabarcoding approaches using 18S and COI genes. Molecular Ecology Resources 20: 1333–1345. https://doi.org/10.1111/1755-0998.13201

Glassman SI, Martiny JBH (2018) Broadscale Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units. mSphere 3. https://doi.org/10.1128/mSphere.00148-18

Gleason JE, Elbrecht V, Braukmann TWA, Hanner RH, Cottenie K (2020) Assessment of stream macroinvertebrate communities with eDNA is not congruent with tissue-based metabarcoding. Molecular Ecology. https://doi.org/10.1111/mec.15597

Goldberg CS, Sepulveda A, Ray A, Baumgardt J, Waits LP (2013) Environmental DNA as a new method for early detection of New Zealand mudsnails (*Potamopyrgus antipodarum*). Freshwater Science 32: 792–800. https://doi.org/10.1899/13-046.1

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

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

Griffiths AJF, Wessler SR, Carroll SB, Doebley J (2015) An Introduction to Genetic Analysis. W. H. Freeman, 868 pp. Available from: https://books.google.com/books/about/An_Introduction_to_Genetic_Analysis.html?hl=&id=_AUnrgEACAAJ.

Guardiola M, Uriz MJ, Taberlet P, Coissac E, Wangensteen OS, Turon X (2015) Deep-Sea, Deep-Sequencing: Metabarcoding Extracellular DNA from Sediments of Marine Canyons. PloS One 10: e0139633. https://doi.org/10.1371/journal.pone.0139633

Haenel Q, Holovachov O, Jondelius U, Sundberg P, Bourlat SJ (2017) NGS-based biodiversity and community structure analysis of meiofaunal eukaryotes in shell sand from Hållö island, Smögen, and soft mud from Gullmarn Fjord, Sweden. Biodiversity Data Journal: e12731. https://doi.org/10.3897/BDJ.5.e12731

Hajibabaei M, Spall JL, Shokralla S, van Konynenburg S (2012) Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. BMC Ecology 12: 28. https://doi.org/10.1186/1472-6785-12-28

Hajibabaei M, Porter TM, Wright M, Rudar J (2019) COI metabarcoding primer choice affects richness and recovery of indicator taxa in freshwater systems. PloS One 14: e0220953. https://doi.org/10.1371/journal.pone.0220953

Hamilton PB, Lefebvre KE, Bull RD (2015) Single cell PCR amplification of diatoms using fresh and preserved samples. Frontiers in Microbiology 6: 1084. https://doi.org/10.3389/fmicb.2015.01084

Hamm CE, Merkel R, Springer O, Jurkojc P, Maier C, Prechtel K, Smetacek V (2003) Architecture and material properties of diatom shells provide effective mechanical protection. Nature 421: 841–843. https://doi.org/10.1038/nature01416

Handley LL, Read DS, Winfield IJ, Kimbell H, Johnson H, Li J, Hahn C, Blackman R, Wilcox R, Donnelly R, Szitenberg A, Hänfling B (2019) Temporal and spatial variation in distribution of fish environmental DNA in England's largest lake. Environmental DNA 1: 26–39. https://doi.org/10.1002/edn3.5

Hänfling B, Lawson Handley L, Read DS, Hahn C, Li J, Nichols P, Blackman RC, Oliver A, Winfield IJ (2016) Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. Molecular Ecology 25: 3101–3119. https://doi.org/10.1111/mec.13660

Harrison JB, Sunday JM, Rogers SM (2019) Predicting the fate of eDNA in the environment and implications for studying biodiversity. Proceedings. Biological Sciences / The Royal Society 286: 20191409. https://doi.org/10.1098/rspb.2019.1409

Hartman LJ, Coyne SR, Norwood DA (2005) Development of a novel internal positive control for Taqman® based assays. Molecular and Cellular Probes 19: 51–59. https://doi.org/10.1016/j.mcp.2004.07.006

Hatfield RG, Batista FM, Bean TP, Fonseca VG, Santos A, Turner AD, Lewis A, Dean KJ, Martinez-Urtaza J (2020) The Application of Nanopore Sequencing Technology to the Study of Dinoflagellates: A Proof of Concept Study for Rapid Sequence-Based Discrimination of Potentially Harmful Algae. Frontiers in Microbiology 11. https://doi.org/10.3389/fmicb.2020.00844

Hering D, Borja A, Jones JI, Pont D, Boets P, Bouchez A, Bruce K, Drakare S, Hänfling B, Kahlert M, Leese F, Meissner K, Mergen P, Reyjol Y, Segurado P, Vogler A, Kelly M (2018) Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. Water Research 138: 192–205. https://doi.org/10.1016/j.watres.2018.03.003

Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, Erndt NG, Cauley TH, Koehler RT, So AP, Dube S, Rose KA, Montesclaros L, Wang S, Stumbo DP, Hodges SP, Romine S, Milanovich FP, White HE, Regan JF, Karlin-Neumann GA, Hindson CM, Saxonov S, Colston BW (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Analytical Chemistry 83: 8604–8610. https://doi.org/10.1021/ac202028g

Hoorfar J, Malorny B, Abdulmawjood A, Cook N, Wagner M, Fach P (2004) Practical considerations in design of internal amplification controls for diagnostic PCR assays. Journal of Clinical Microbiology 42(5): 1863-1868.DOI: 10.1128/jcm.42.5.1863-1868.2004

Hosler D (2017) Where is the body? Dreissenid mussels, raw water testing, and the real value of environmental DNA. Management of Biological Invasions: International Journal of Applied Research on Biological Invasions 8: 335–341. https://doi.org/10.3391/mbi.2017.8.3.07

Hunter ME, Dorazio RM, Butterfield JSS, Meigs-Friend G, Nico LG, Ferrante JA (2017) Detection limits of quantitative and digital PCR assays and their influence in presence-absence surveys of environmental DNA. Molecular Ecology Resources 17: 221–229. https://doi.org/10.1111/1755-0998.12619

Illumina (2016) Optimising cluster density on Illumina Sequencing Systems https://emea.illumina.com/content/dam/illumina-marketing/documents/products/other/miseq-overclustering-primer-770-2014-038.pdf

Jane SF, Wilcox TM, McKelvey KS, Young MK, Schwartz MK, Lowe WH, Letcher BH, Whiteley AR (2015) Distance, flow and PCR inhibition: eDNA dynamics in two headwater streams. Molecular Ecology Resources 15: 216–227. https://doi.org/10.1111/1755-0998.12285

Jensen MR, Sigsgaard EE, Liu S, Manica A, Bach SS, Hansen MM, Møller PR, Thomsen PF (2020) Genome-scale target capture of mitochondrial and nuclear environmental DNA from water samples. Molecular Ecology Resources. https://doi.org/10.1111/1755-0998.13293

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

Jerde CL, Olds BP, Shogren AJ, Andruszkiewicz EA, Mahon AR, Bolster D, Tank JL (2016) Influence of Stream Bottom Substrate on Retention and Transport of Vertebrate Environmental DNA. Environmental Science & Technology 50: 8770–8779. https://doi.org/10.1021/acs.est.6b01761

Jerde CL, Lindsay Chadderton W, Mahon AR, Renshaw MA, Corush J, Budny ML, Mysorekar S, Lodge DM (2013) Detection of Asian carp DNA as part of a Great Lakes basin-wide surveillance program. Canadian Journal of Fisheries and Aquatic Sciences 70: 522–526. https://doi.org/10.1139/cjfas-2012-0478

Jeunen G, Lamare MD, Knapp M, Spencer HG, Taylor HR, Stat M, Bunce M, Gemmell NJ (2020) Water stratification in the marine biome restricts vertical environmental DNA (eDNA) signal dispersal. Environmental DNA 2: 99–111. https://doi.org/10.1002/edn3.49

Jeunen G-J, Knapp M, Spencer HG, Lamare MD, Taylor HR, Stat M, Bunce M, Gemmell NJ (2019) Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats connected by water movement. Molecular Ecology Resources 19: 426–438. https://doi.org/10.1111/1755-0998.12982

Ji Y, Huotari T, Roslin T, Schmidt NM, Wang J, Yu DW, Ovaskainen O (2019) SPIKEPIPE: A metagenomic pipeline for the accurate quantification of eukaryotic species occurrences and intraspecific abundance change using DNA barcodes or mitogenomes. Molecular Ecology Resources. https://doi.org/10.1111/1755-0998.13057

Jo T, Murakami H, Yamamoto S, Masuda R, Minamoto T (2019a) Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution. Ecology and Volution 9: 1135–1146. https://doi.org/10.1002/ece3.4802

Jo T, Arimoto M, Murakami H, Masuda R, Minamoto T (2019b) Particle size distribution of environmental DNA from the nuclei of marine fish. Environmental Science & Technology 53: 9947–9956. https://doi.org/10.1021/acs.est.9b02833

Jo T, Murakami H, Masuda R, Sakata MK, Yamamoto S, Minamoto T (2017) Rapid degradation of longer DNA fragments enables the improved estimation of distribution and biomass using environmental DNA. Molecular Ecology Resources 17: e25–e33. https://doi.org/10.1111/1755-0998.12685

Kelly MG (1998) Use of the trophic diatom index to monitor eutrophication in rivers. Water Research 32: 236–242. https://doi.org/10.1016/s0043-1354(97)00157-7

Kelly MG, Boonham N, Juggins S, Kille P, Mann DG, Pass D, Sapp M, Sato S, Glover R (2018) A DNA based diatom metabarcoding approach for Water Framework Directive classification of rivers (Issue March). ISBN: 978-1-84911-406-6

Klymus KE, Richter CA, Chapman DC, Paukert C (2015) Quantification of eDNA shedding rates from invasive bighead carp *Hypophthalmichthys nobilis* and silver carp *Hypophthalmichthys molitrix*. Biological Conservation 183: 77–84. https://doi.org/10.1016/j.biocon.2014.11.020

Klymus KE, Merkes CM, Allison MJ, Goldberg CS, Helbing CC, Hunter ME, Jackson CA, Lance RF, Mangan AM, Monroe EM, Piaggio AJ, Stokdyk JP, Wilson CC, Richter CA (2020) Reporting the limits of detection and quantification for environmental DNA assays. Environmental DNA 2: 271–282. https://doi.org/10.1002/edn3.29

Lafferty KD, Benesh KC, Mahon AR, Jerde CL, Lowe CG (2018) Detecting Southern California's White Sharks With Environmental DNA. Frontiers in Marine Science 5. https://doi.org/10.3389/fmars.2018.00355

Langlois VS, Allison MJ, Bergman LC, To TA, Helbing CC (2020) The need for robust qPCR-based eDNA detection assays in environmental monitoring and species inventories. Environmental DNA. https://doi.org/10.1002/edn3.164

Lee A, McVey J, Faustino P, Lute S, Sweeney N, Pawar V, Khan M, Brorson K, Hussong D (2010) Use of Hydrogenophaga pseudoflava penetration to quantitatively assess the impact of filtration parameters for 0.2-micrometer-pore-size filters. Applied and Environmental Microbiology 76: 695–700. https://doi.org/10.1128/AEM.01825-09

Leempoel K, Hebert T, Hadly EA (2020) A comparison of eDNA to camera trapping for assessment of terrestrial mammal diversity. Proceedings. Biological sciences / The Royal Society 287: 20192353. https://doi.org/10.1098/rspb.2019.2353

Leese F, Sander M, Buchner D, Elbrecht V, Haase P, Zizka VMA (2021) Improved freshwater macroinvertebrate detection from environmental DNA through minimized nontarget amplification. Environmental DNA 3: 261–276. https://doi.org/10.1002/edn3.177

Leese F, Bouchez A, Abarenkov K, Altermatt F, Borja Á, Bruce K, Ekrem T, Čiampor F, Čiamporová-Zaťovičová Z, Costa FO, Duarte S, Elbrecht V, Fontaneto D, Franc A, Geiger MF, Hering D, Kahlert M, Stroil BK, Kelly M, Keskin E, Liska I, Mergen P, Meissner K, Pawlowski J, Penev L, Reyjol Y, Rotter A, Steinke D, van der Wal B, Vitecek S, Zimmermann J, Weigand AM (2018) Why We Need Sustainable Networks Bridging Countries, Disciplines, Cultures and Generations for Aquatic Biomonitoring 2.0: A Perspective Derived From the DNAqua-Net COST Action. Next Generation Biomonitoring: Part 1: 63–99. https://doi.org/10.1016/bs.aecr.2018.01.001

Leray M, Haenel Q, Bourlat SJ (2016) Preparation of Amplicon Libraries for Metabarcoding of Marine Eukaryotes Using Illumina MiSeq: The Adapter Ligation Method. Methods in Molecular Biology 1452: 209–218. https://doi.org/10.1007/978-1-4939-3774-5_14

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Frontiers in Zoology 10: 34. https://doi.org/10.1186/1742-9994-10-34

Li J, Lawson Handley LJ, Harper LR, Brys R, Watson HV, Di Muri C, Zhang X, Hänfling B (2019) Limited dispersion and quick degradation of environmental DNA in fish ponds inferred by metabarcoding. Environmental DNA 1: 238–250. https://doi.org/10.1002/edn3.24

Life Technologies Corporation (2012) Real-time PCR handbook. https://www.gene-quantification.de/real-time-pcr-handbook-life-technologies-update-flr.pdf

Linard B, Arribas P, Andújar C, Crampton-Platt A, Vogler AP (2016) Lessons from genome skimming of arthropod-preserving ethanol. Molecular Ecology Resources 16: 1365–1377. https://doi.org/10.1111/1755-0998.12539

Littlefair JE, Hrenchuk LE, Blanchfield PJ, Rennie MD, Cristescu ME (2020) Thermal stratification and fish thermal preference explain vertical eDNA distributions in lakes. Molecular Ecology. https://doi.org/10.1111/mec.15623

Loeza-Quintana T, Abbott CL, Heath DD, Bernatchez L, Hanner RH (2020) Pathway to Increase Standards and Competency of eDNA Surveys (PISCeS)—Advancing collaboration and standardization efforts in the field of eDNA. Environmental DNA 2: 255–260. https://doi.org/10.1002/edn3.112

Longmire JL, Baker RJ, Maltbie M, Texas Tech University. (1997) Use of "Lysis Buffer" in DNA isolation and its implication for museum collections. https://doi.org/10.5962/bhl.title.143318

van der Loos LM, Nijland R (2020) Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. Molecular Ecology. https://doi.org/10.1111/mec.15592

Macher J-N, Leese F (2018) eDNA metabarcoding of rivers: Is all eDNA everywhere, all the time? Proceedings of the 5th European Congress of Conservation Biology. https://doi.org/10.17011/conference/eccb2018/107411

Macher J-N, Zizka VMA, Weigand AM, Leese F (2018a) A simple centrifugation protocol for metagenomic studies increases mitochondrial DNA yield by two orders of magnitude. Methods in Ecology and Evolution 9: 1070–1074. https://doi.org/10.1111/2041-210x.12937

Macher J-N, Vivancos A, Piggott JJ, Centeno FC, Matthaei CD, Leese F (2018b) Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate cytochrome c oxidase I primers. Molecular Ecology Resources 18: 1456–1468. https://doi.org/10.1111/1755-0998.12940

Mächler E, Deiner K, Spahn F, Altermatt F (2016) Fishing in the Water: Effect of Sampled Water Volume on Environmental DNA-Based Detection of Macroinvertebrates. Environmental Science & Technology 50: 305–312. https://doi.org/10.1021/acs.est.5b04188

Mäki A, Salmi P, Mikkonen A, Kremp A, Tiirola M (2017) Sample Preservation, DNA or RNA Extraction and Data Analysis for High-Throughput Phytoplankton Community Sequencing. Frontiers in Microbiology 8. https://doi.org/10.3389/fmicb.2017.01848

Mandaville SM, Soil & Water Conservation Society of Metro Halifax (2002) Benthic Macroinvertebrates in Freshwaters: Taxa Tolerance Values, Metrics, and Protocols. 48 pp. Available from: https://books.google.com/books/about/Benthic_Macroinvertebrates_in_Freshwater.html?hl=&id=Gdc0OAAACAAJ.

Martins FMS, Galhardo M, Filipe AF, Teixeira A, Pinheiro P, Paupério J, Alves PC, Beja P (2019) Have the cake and eat it: Optimising non-destructive DNA metabarcoding of macroinvertebrate samples for freshwater biomonitoring. Molecular Ecology Resources. https://doi.org/10.1111/1755-0998.13012

Martins FMS, Porto M, Feio MJ, Egeter B, Bonin A, Serra SRQ, Taberlet P, Beja P (2020) Modelling technical and biological biases in macroinvertebrate community assessment from bulk preservative using multiple metabarcoding markers. Molecular Ecology. https://doi.org/10.1111/mec.15620

Mauvisseau Q, Halfmaerten D, Neyrinck S, Burian A, Brys R (2021) Effects of preservation strategies on environmental DNA detection and quantification using ddPCR. Environmental DNA. https://doi.org/10.1002/edn3.188

McClenaghan B, Fahner N, Cote D, Chawarski J, McCarthy A, Rajabi H, Singer G, Hajibabaei M (2020) Harnessing the power of eDNA metabarcoding for the detection of deep-sea fishes. https://doi.org/10.1371/journal.pone.0236540

McElroy ME, Dressler TL, Titcomb GC, Wilson EA, Deiner K, Dudley TL, Eliason EJ, Evans NT, Gaines SD, Lafferty KD, Lamberti GA, Li Y, Lodge DM, Love MS, Mahon AR, Pfrender ME, Renshaw MA, Selkoe KA, Jerde CL (2020) Calibrating Environmental DNA Metabarcoding to Conventional Surveys for Measuring Fish Species Richness. Frontiers in Ecology and Evolution 8. https://doi.org/10.3389/fevo.2020.00276

McKee AM, Spear SF, Pierson TW (2015) The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. Biological Conservation 183: 70–76. https://doi.org/10.1016/j.biocon.2014.11.031

Minamoto T, Miya M, Sado T, Seino S, Doi H, Kondoh M, Nakamura K, Takahara T, Yamamoto S, Yamanaka H, Araki H, Iwasaki W, Kasai A, Masuda R, Uchii K (2021) An illustrated manual for environmental

DNA research: Water sampling guidelines and experimental protocols. Environmental DNA 3: 8–13. https://doi.org/10.1002/edn3.121

Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, Minamoto T, Yamamoto S, Yamanaka H, Araki H, Kondoh M, Iwasaki W (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. Royal Society Open Science 2: 150088. https://doi.org/10.1098/rsos.150088

Mora D, Abarca N, Proft S, Grau JH, Enke N, Carmona J, Skibbe O, Jahn R, Zimmermann J (2019) Morphology and metabarcoding: a test with stream diatoms from Mexico highlights the complementarity of identification methods. Freshwater Science 38: 448–464. https://doi.org/10.1086/704827

Mora-Gómez J, Freixa A, Perujo N, Barral-Fraga L (2016) Limits of the Biofilm Concept and Types of Aquatic Biofilms. Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment: 3–28. https://doi.org/10.21775/9781910190173.01

Morin S, Gómez N, Tornés E, Licursi M, Rosebery J (2016) Benthic Diatom Monitoring and Assessment of Freshwater Environments: Standard Methods and Future Challenges. Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment: 111–124. https://doi.org/10.21775/9781910190173.06

Moriniere J, Cancian de Araujo B, Lam AW, Hausmann A, Balke M, Schmidt S, Hendrich L, Doczkal D, Fartmann B, Arvidsson S, Haszprunar G (2016) Species Identification in Malaise Trap Samples by DNA Barcoding Based on NGS Technologies and a Scoring Matrix. PloS One 11: e0155497. https://doi.org/10.1371/journal.pone.0155497

Mortágua A, Vasselon V, Oliveira R, Elias C, Chardon C, Bouchez A, Rimet F, Feio MJ, Almeida SFP (2019) Applicability of DNA metabarcoding approach in the bioassessment of Portuguese rivers using diatoms. Ecological Indicators 106: 105470. https://doi.org/10.1016/j.ecolind.2019.105470

Moushomi R, Wilgar G, Carvalho G, Creer S, Seymour M (2019) Environmental DNA size sorting and degradation experiment indicates the state of Daphnia magna mitochondrial and nuclear eDNA is subcellular. Scientific Reports 9: 12500. https://doi.org/10.1038/s41598-019-48984-7

Muha TP, Robinson CV, de Leaniz CG, Consuegra S (2019) An optimised eDNA protocol for detecting fish in lentic and lotic freshwaters using a small water volume. PloS One 14: e0219218. https://doi.org/10.1371/journal.pone.0219218

Murray DC, Coghlan ML, Bunce M (2015) From benchtop to desktop: important considerations when designing amplicon sequencing workflows. PloS One 10: e0124671. https://doi.org/10.1371/journal.pone.0124671

Nascimento FJA, Lallias D, Bik HM, Creer S (2018) Sample size effects on the assessment of eukaryotic diversity and community structure in aquatic sediments using high-throughput sequencing. Scientific Reports 8: 11737. https://doi.org/10.1038/s41598-018-30179-1

NF EN 13946 (2014) Water quality - Guidance standard for the routine sampling and pretreatment of benthic diatoms from rivers. European Standard EN 13946: 1–17.

Nguyen TNM, Berzano M, Gualerzi CO, Spurio R (2011) Development of molecular tools for the detection of freshwater diatoms. Journal of Microbiological Methods 84: 33–40. https://doi.org/10.1016/j.mimet.2010.10.008

Nichols RV, Vollmers C, Newsom LA, Wang Y, Heintzman PD, Leighton M, Green RE, Shapiro B (2018) Minimizing polymerase biases in metabarcoding. Molecular Ecology Resources. https://doi.org/10.1111/1755-0998.12895

Nielsen M, Gilbert MTP, Pape T, Bohmann K (2019) A simplified DNA extraction protocol for unsorted bulk arthropod samples that maintains exoskeletal integrity. Environmental DNA 1: 144–154. https://doi.org/10.1002/edn3.16

Nilsson RH, Anslan S, Bahram M, Wurzbacher C, Baldrian P, Tedersoo L (2019) Mycobiome diversity: high-throughput sequencing and identification of fungi. Nature reviews. Microbiology 17: 95–109. https://doi.org/10.1038/s41579-018-0116-y

Ogram A, Sayler GS, Barkay T (1987) The extraction and purification of microbial DNA from sediments. Journal of Microbiological Methods 7: 57–66. https://doi.org/10.1016/0167-7012(87)90025-x

Olson ND, Morrow JB (2012) DNA extract characterization process for microbial detection methods development and validation. BMC Res Notes 5: 668. https://doi.org/10.1186/1756-0500-5-668

Pace NR, Stahl DA, Lane DJ, Olsen GJ (1986) The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences. Advances in Microbial Ecology: 1–55. https://doi.org/10.1007/978-1-4757-0611-6_1

Pawlowski J, Lecroq B (2010) Short rDNA barcodes for species identification in foraminifera. The Journal of Eukaryotic Microbiology 57: 197–205. https://doi.org/10.1111/j.1550-7408.2009.00468.x

Pawlowski J, Lejzerowicz F, Esling P (2014) Next-generation environmental diversity surveys of foraminifera: preparing the future. The Biological Bulletin 227: 93–106. https://doi.org/10.1086/BBLv227n2p93

Pawlowski J, Apothéloz-Perret-Gentil L, Mächler E & Altermatt F (2020a) Environmental DNA applications in biomonitoring and bio- assessment of aquatic ecosystems. Guidelines. Federal Office for the Environment, Bern. Environmental Studies. no. 2010: 71 pp. https://www.bafu.admin.ch/bafu/en/home/topics/water/water--publications/publications-water/environmental-dna-applications-in-biomonitoring-and-bioassessment.html

Pawlowski J, Apothéloz-Perret-Gentil L, Altermatt F (2020b) Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. Molecular Ecology 29: 4258–4264. https://doi.org/10.1111/mec.15643

Pawlowski J, Kelly-Quinn M, Altermatt F, Apothéloz-Perret-Gentil L, Beja P, Boggero A, Borja A, Bouchez A, Cordier T, Domaizon I, Feio MJ, Filipe AF, Fornaroli R, Graf W, Herder J, van der Hoorn B, Iwan Jones J, Sagova-Mareckova M, Moritz C, Barquín J, Piggott JJ, Pinna M, Rimet F, Rinkevich B, Sousa-Santos C, Specchia V, Trobajo R, Vasselon V, Vitecek S, Zimmerman J, Weigand A, Leese F, Kahlert M (2018) The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. Science of the Total Environment 637–638: 1295–1310. https://doi.org/10.1016/j.scitotenv.2018.05.002

Pereira-da-Conceicoa L, Elbrecht V, Hall A, Briscoe A, Barber-James H, Price B (2021) Metabarcoding unsorted kick-samples facilitates macroinvertebrate-based biomonitoring with increased taxonomic resolution, while outperforming environmental DNA. Environmental DNA 3: 353–371. https://doi.org/10.1002/edn3.116

Pérez-Burillo J, Trobajo R, Vasselon V, Rimet F, Bouchez A, Mann DG (2020) Evaluation and sensitivity analysis of diatom DNA metabarcoding for WFD bioassessment of Mediterranean rivers. The Science of the Total Environment 727: 138445. https://doi.org/10.1016/j.scitotenv.2020.138445

Pilliod DS, Laramie MB, MacCoy D, Maclean S (2019) Integration of eDNA-Based Biological Monitoring within the U.S. Geological Survey's National Streamgage Network. JAWRA Journal of the American Water Resources Association 55: 1505–1518. https://doi.org/10.1111/1752-1688.12800

Piper AM, Batovska J, Cogan NOI, Weiss J, Cunningham JP, Rodoni BC, Blacket MJ (2019) Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. GigaScience 8. https://doi.org/10.1093/gigascience/giz092

Pomerantz A, Peñafiel N, Arteaga A, Bustamante L, Pichardo F, Coloma LA, Barrio-Amorós CL, Salazar-Valenzuela D, Prost S (2018) Real-time DNA barcoding in a rainforest using nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building. GigaScience 7. https://doi.org/10.1093/gigascience/giy033

Pont D, Valentini A, Rocle M, Maire A, Delaigue O, Jean P, Dejean T (2019) The future of fish-based ecological assessment of European rivers: from traditional EU Water Framework Directive compliant methods to eDNA metabarcoding-based approaches. Journal of Fish Biology. https://doi.org/10.1111/jfb.14176

Pont D, Rocle M, Valentini A, Civade R, Jean P, Maire A, Roset N, Schabuss M, Zornig H, Dejean T (2018) Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. Scientific Reports 8: 10361. https://doi.org/10.1038/s41598-018-28424-8

Port JA, O'Donnell JL, Romero-Maraccini OC, Leary PR, Litvin SY, Nickols KJ, Yamahara KM, Kelly RP (2016) Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. Molecular Ecology 25: 527–541. https://doi.org/10.1111/mec.13481

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

Porter TM, Hajibabaei M (2020) Putting COI Metabarcoding in Context: The Utility of Exact Sequence Variants (ESVs) in Biodiversity Analysis. Frontiers in Ecology and Evolution 8. https://doi.org/10.3389/fevo.2020.00248

Postaire BD, Bakker J, Gardiner J, Wiley TR, Chapman DD (2020) Environmental DNA detection tracks established seasonal occurrence of blacktip sharks (*Carcharhinus limbatus*) in a semi-enclosed subtropical bay. Scientific Reports 10: 11847. https://doi.org/10.1038/s41598-020-68843-0

Precone V, Dundar M, Beccari T, Turanli ET, Cecchin S, Marceddu G, Manara E, Bertelli M (2018) Quality assurance of genetic laboratories and the EBTNA practice certification, a simple standardization assurance system for a laboratory network. The EuroBiotech Journal 2: 215–222. https://doi.org/10.2478/ebtj-2018-0052

Prié V, Valentini A, Lopes-Lima M, Froufe E, Rocle M, Poulet N, Taberlet P, Dejean T (2020) Environmental DNA metabarcoding for freshwater bivalves biodiversity assessment: methods and results for the Western Palearctic (European sub-region). Hydrobiologia. https://doi.org/10.1007/s10750-020-04260-8

Rapp D (2010) DNA extraction from bovine faeces: current status and future trends. Journal of Applied Microbiology 108: 1485–1493. https://doi.org/10.1111/j.1365-2672.2009.04606.x

Renshaw MA, Olds BP, Jerde CL, McVeigh MM, Lodge DM (2015) The room temperature preservation of filtered environmental DNA samples and assimilation into a phenol-chloroform-isoamyl alcohol DNA extraction. Molecular Ecology Resources 15: 168–176. https://doi.org/10.1111/1755-0998.12281

Reuter JA, Spacek DV, Snyder MP (2015) High-throughput sequencing technologies. Molecular Cell 58: 586–597. https://doi.org/10.1016/j.molcel.2015.05.004

Riaz T, Shehzad W, Viari A, Pompanon F, Taberlet P, Coissac E (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. Nucleic Acids Research 39: e145. https://doi.org/10.1093/nar/gkr732

Rivera SF, Vasselon V, Bouchez A, Rimet F (2020) Diatom metabarcoding applied to large scale monitoring networks: Optimization of bioinformatics strategies using Mothur software. Ecological Indicators 109: 105775. https://doi.org/10.1016/j.ecolind.2019.105775

Rivera SF, Vasselon V, Jacquet S, Bouchez A, Ariztegui D, Rimet F (2018) Metabarcoding of lake benthic diatoms: from structure assemblages to ecological assessment. Hydrobiologia 807: 37–51. https://doi.org/10.1007/s10750-017-3381-2

Robinson CV, Porter TM, Wright MTG, Hajibabaei M (2021) Propylene glycol-based antifreeze is an effective preservative for DNA metabarcoding of benthic arthropods. Freshwater Science 40: 77–87. https://doi.org/10.1086/712232

Rodriguez-Ezpeleta N, Morissette O, Bean C, Manu S, Banerjee P, Lacoursiere A, Ben K, Alter E, Roger F, Holman L, Stewart K, Monaghan M, Mauvisseau Q, Mirimin L, Wangensteen OS, Antognazza C, Helyar S, de Boer H, Monchamp M-E, Nijland R, Abbott C, Doi H, Barnes M, Leray M, Hablützel P, Deiner K (2021) Trade-offs between reducing complex terminology and producing accurate interpretations from environmental DNA: Comment on "Environmental DNA: What's behind the term?" by Pawlowski et al. (2020). Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/mec.15942.

Romanowski G, Lorenz MG, Wackernagel W (1991) Adsorption of plasmid DNA to mineral surfaces and protection against DNase I. Applied and Environmental Microbiology 57: 1057–1061. https://doi.org/10.1128/AEM.57.4.1057-1061.1991

Rusch JC, Mojžišová M, Strand DA, Svobodová J, Vrålstad T, Petrusek A (2020) Simultaneous detection of native and invasive crayfish and Aphanomyces astaci from environmental DNA samples in a wide range of habitats in Central Europe. NeoBiota 58: 1–32. https://doi.org/10.3897/neobiota.58.49358

Sandve GK, Nekrutenko A, Taylor J, Hovig E (2013) Ten simple rules for reproducible computational research. PLoS Computational Biology 9: e1003285. https://doi.org/10.1371/journal.pcbi.1003285

Schabacker JC, Amish SJ, Ellis BK, Gardner B, Miller DL, Rutledge EA, Sepulveda AJ, Luikart G (2020) Increased eDNA detection sensitivity using a novel high-volume water sampling method. Environmental DNA. https://doi.org/10.1002/edn3.63

Schrader C, Schielke A, Ellerbroek L, Johne R (2012) PCR inhibitors - occurrence, properties and removal. Journal of Applied Microbiology 113: 1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x

Schweiss KE, Lehman RN, Marcus Drymon J, Phillips NM (2020) Development of highly sensitive environmental DNA methods for the detection of Bull Sharks, Carcharhinus leucas (Müller and Henle, 1839), using Droplet Digital™ PCR. Environmental DNA 2: 3–12. https://doi.org/10.1002/edn3.39

Sellers GS, Di Muri C, Gómez A, Hänfling B (2018) Mu-DNA: a modular universal DNA extraction method adaptable for a wide range of sample types. Metabarcoding and Metagenomics 2: e24556. https://doi.org/10.3897/mbmg.2.24556

Sepulveda AJ, Schabacker J, Smith S, Al-Chokhachy R, Luikart G, Amish SJ (2019) Improved detection of rare, endangered and invasive trout in using a new large-volume sampling method for eDNA capture. Environmental DNA 1: 227–237. https://doi.org/10.1002/edn3.23

Seymour M, Durance I, Cosby BJ, Ransom-Jones E, Deiner K, Ormerod SJ, Colbourne JK, Wilgar G, Carvalho GR, de Bruyn M, Edwards F, Emmett BA, Bik HM, Creer S (2018) Acidity promotes degradation of multi-species environmental DNA in lotic mesocosms. Communications Biology 1: 4. https://doi.org/10.1038/s42003-017-0005-3

Seymour M (2019) Rapid progression and future of environmental DNA research. Communications Biology 2: 80. https://doi.org/10.1038/s42003-019-0330-9

Seymour M, Edwards FK, Cosby BJ, Kelly MG, de Bruyn M, Carvalho GR, Creer S (2020) Executing multitaxa eDNA ecological assessment via traditional metrics and interactive networks. The Science of the Total Environment 729: 138801. https://doi.org/10.1016/j.scitotenv.2020.138801

Seymour M, Edwards F, Cosby J, Scarlet P, Bista I, Brailsford F, Glanville H, de Bruyn M, Carvalho G, & Creer, S (2021) Environmental DNA provides greater insight to biodiversity and ecosystem function compared to traditional approaches, via spatio-temporal nestedness and turnover partitioning. Communications Biology, 4(1), p. 512. doi: 10.1038/s42003-021-02031-2

Shogren AJ, Tank JL, Andruszkiewicz E, Olds B, Mahon AR, Jerde CL, Bolster D (2017) Controls on eDNA movement in streams: Transport, Retention, and Resuspension. Scientific Reports 7: 5065. https://doi.org/10.1038/s41598-017-05223-1

Shogren AJ, Tank JL, Egan SP, August O, Rosi EJ, Hanrahan BR, Renshaw MA, Gantz CA, Bolster D (2018) Water Flow and Biofilm Cover Influence Environmental DNA Detection in Recirculating Streams. Environmental Science & Technology 52: 8530–8537. https://doi.org/10.1021/acs.est.8b01822

Simmons M, Tucker A, Lindsay Chadderton W, Jerde CL, Mahon AR (2016) Active and passive environmental DNA surveillance of aquatic invasive species. Canadian Journal of Fisheries and Aquatic Sciences 73: 76–83. https://doi.org/10.1139/cjfas-2015-0262

Singer GAC, Shekarriz S, McCarthy A, Fahner N, Hajibabaei M (2020) The utility of a metagenomics approachfor marine biomonitoring. bioRxiv. https://doi.org/10.1101/2020.03.16.993667

Sinniger F, Pawlowski J, Harii S, Gooday AJ, Yamamoto H, Chevaldonné P, Cedhagen T, Carvalho G, Creer S (2016) Worldwide Analysis of Sedimentary DNA Reveals Major Gaps in Taxonomic Knowledge of Deep-Sea Benthos. Frontiers in Marine Science 3. https://doi.org/10.3389/fmars.2016.00092

Snyder DE (2003) Invited overview: conclusions from a review of electrofishing and its harmful effects on fish. Reviews in Fish Biology and Fisheries 13: 445–453. https://doi.org/10.1007/s11160-004-1095-9

Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR, Arrieta JM, Herndl GJ (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". Proceedings of the National Academy of Sciences of the United States of America 103: 12115–12120. https://doi.org/10.1073/pnas.0605127103

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

Stewart KA (2019) Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. Biodiversity and Conservation 28: 983–1001. https://doi.org/10.1007/s10531-019-01709-8

Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner H-W, Richards TA (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Molecular Ecology 19, Suppl 1: 21–31. https://doi.org/10.1111/j.1365-294X.2009.04480.x

Strickler KM, Fremier AK, Goldberg CS (2015) Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. Biological Conservation 183: 85–92. https://doi.org/10.1016/j.biocon.2014.11.038

Taberlet P (2018) Environmental dna for biodiversity research and monitoring: methods for biodiversity research. 1st edition. Oxford University Press, New York, NY, pages cm pp.

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

Taberlet P, Prud'homme SM, Campione E, Roy J, Miquel C, Shehzad W, Gielly L, Rioux D, Choler P, Clement JC, Melodelima C, Pompanon F, Coissac E (2012b) Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. Molecular Ecology 21: 1816–1820. https://doi.org/10.1111/J.1365-294x.2011.05317.X

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012c). Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology 21: 2045–2050. https://doi.org/10.1111/j.1365-294x.2012.05470.x

Tapolczai K, Keck F, Bouchez A, Rimet F, Kahlert M, Vasselon V (2019) Diatom DNA Metabarcoding forBiomonitoring: Strategies to Avoid Major Taxonomical and Bioinformatical Biases Limiting MolecularIndices Capacities. Frontiers in Ecology and Evolution 7:409. https://doi.org/10.3389/fevo.2019.00409

Tatangelo V, Franzetti A, Gandolfi I, Bestetti G, Ambrosini R (2014) Effect of preservation method on the assessment of bacterial community structure in soil and water samples. FEMS Microbiology Letters 356: 32–38. https://doi.org/10.1111/1574-6968.12475

Thalinger B, Pütz Y, Traugott M (2021c) Endpoint PCR coupled with capillary electrophoresis (celPCR) provides sensitive and quantitative measures of environmental DNA in singleplex and multiplex reactions. PloS One 16(7): e0254356. https://doi.org/10.1371/journal.pone.0254356

Thalinger B, Wolf E, Traugott M, Wanzenböck J (2019) Monitoring spawning migrations of potamodromous fish species via eDNA. Scientific Reports 9: 15388. https://doi.org/10.1038/s41598-019-51398-0

Thalinger B, Rieder A, Teuffenbach A, Pütz Y, Schwerte T, Wanzenböck J, Traugott M (2021a) The effect of activity, energy use, and species identity on environmental DNA shedding of freshwater fish. Frontiers in Ecology and Evolution 9:623718. https://doi.org/10.3389/fevo.2021.623718

Thalinger B, Deiner K, Harper LR, Rees HC, Blackman RC, Sint D, Traugott M, Goldberg CS, Bruce K (2021b) A validation scale to determine the readiness of environmental DNA assays for routine species monitoring. Environmental DNA 3(4): 823-836. https://doi.org/10.1002/edn3.189

Thomas AC, Nguyen PL, Howard J, Goldberg CS (2019) A self-preserving, partially biodegradable eDNA filter. Methods in Ecology and Evolution 10: 1136–1141. https://doi.org/10.1111/2041-210x.13212

Thomas AC, Howard J, Nguyen PL, Seimon TA, Goldberg CS (2018) eDNA Sampler: A fully integrated environmental DNA sampling system. Methods in Ecology and Evolution 9: 1379–1385. https://doi.org/10.1111/2041-210x.12994

Thomas R, Meybeck M, and Beim A (1996) Chapter 7. Lakes. In: Water Quality Assessments - A Guide to Use of Biota, Sediments and Water in Environmental Monitoring - Second Edition. Edited by Deborah Chapman, 1992, 1996 UNESCO/WHO/UNEP. ISBN 0 419 21590 5 (HB) 0 419 21600 6 (PB)

Torti A, Lever MA, Jørgensen BB (2015) Origin, dynamics, and implications of extracellular DNA pools in marine sediments. Marine Genomics 24 Pt 3: 185–196. https://doi.org/10.1016/j.margen.2015.08.007

Treonis AM, Unangst SK, Kepler RM, Buyer JS, Cavigelli MA, Mirsky SB, Maul JE (2018) Characterization of soil nematode communities in three cropping systems through morphological and DNA metabarcoding approaches. Scientific Reports 8: 2004. https://doi.org/10.1038/s41598-018-20366-5

Truelove NK, Andruszkiewicz EA, Block BA (2019) A rapid environmental DNA method for detecting white sharks in the open ocean. Methods in Ecology and Evolution / British Ecological Society 10: 1128–1135. https://doi.org/10.1111/2041-210x.13201

Tsuji S, Takahara T, Doi H, Shibata N, Yamanaka H (2019) The detection of aquatic macroorganisms using environmental DNA analysis—A review of methods for collection, extraction, and detection. Environmental DNA 1: 99–108. https://doi.org/10.1002/edn3.21

Tsuji S, Miya M, Ushio M, Sato H, Minamoto T, Yamanaka H (2020) Evaluating intraspecific genetic diversity using environmental DNA and denoising approach: A case study using tank water. Environmental DNA 2: 42–52. https://doi.org/10.1002/edn3.44

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

Valentini A, Taberlet P, Miaud C, Civade R, Herder J, Thomsen PF, Bellemain E, Besnard A, Coissac E, Boyer F, Gaboriaud C, Jean P, Poulet N, Roset N, Copp GH, Geniez P, Pont D, Argillier C, Baudoin J-M, Peroux T, Crivelli AJ, Olivier A, Acqueberge M, Le Brun M, Møller PR, Willerslev E, Dejean T (2016) Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Molecular Ecology 25: 929–942. https://doi.org/10.1111/mec.13428

Vamos E, Elbrecht V, Leese F (2017) Short COI markers for freshwater macroinvertebrate metabarcoding. Metabarcoding and Metagenomics 1: e14625. https://doi.org/10.3897/mbmg.1.14625

Vasselon V, Rimet F, Tapolczai K, Bouchez A (2017a) Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). Ecological Indicators 82: 1–12. https://doi.org/10.1016/j.ecolind.2017.06.024

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

Visco JA, Apothéloz-Perret-Gentil L, Cordonier A, Esling P, Pillet L, Pawlowski J (2015) Environmental Monitoring: Inferring the Diatom Index from Next-Generation Sequencing Data. Environmental Science & Technology 49: 7597–7605. https://doi.org/10.1021/es506158m

Wangensteen OS, Turon X (2017) Metabarcoding techniques for assessing biodiversity of marine animal forests. Marine animal forests. The Ecology of Benthic Biodiversity Hotspots 1: 445–503. Available

from: http://www.academia.edu/download/50865315/Metabarcoding_Techniques_for_Assessing_Biodiversity_of_Marine_Animal_Forests.pdf.

Wangensteen OS, Palacín C, Guardiola M, Turon X (2018) DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. PeerJ 6: e4705. https://doi.org/10.7717/peerj.4705

Wegleitner BJ, Jerde CL, Tucker A, Lindsay Chadderton W, Mahon AR (2015) Long duration, room temperature preservation of filtered eDNA samples. Conservation Genetics Resources 7: 789–791. https://doi.org/10.1007/s12686-015-0483-x

Weigand AM, Macher J-N (2018) A DNA metabarcoding protocol for hyporheic freshwater meiofauna: Evaluating highly degenerate COI primers and replication strategy. Metabarcoding and Metagenomics 2. https://doi.org/10.3897/mbmg.2.26869

Weigand AM, Desquiotz N, Weigand H, Szucsich N (2021) Application of propylene glycol in DNA-based studies of invertebrates. Metabarcoding and Metagenomics 5. https://doi.org/10.3897/mbmg.5.57278

Wilcox TM, McKelvey KS, Young MK, Lowe WH, Schwartz MK (2015) Environmental DNA particle size distribution from Brook Trout (*Salvelinus fontinalis*). Conservation Genetics Resources 7: 639–641. https://doi.org/10.1007/s12686-015-0465-z

Wilcox TM, McKelvey KS, Young MK, Engkjer C, Lance RF, Lahr A, Eby LA, Schwartz MK (2020) Parallel, targeted analysis of environmental samples via high-throughput quantitative PCR. Environmental DNA 2: 544–553. https://doi.org/10.1002/edn3.80

Williams M, Hernandez C, O'Sullivan AM, April J, Regan F, Bernatchez L, Parle-McDermott A (2020) Comparing CRISPR-Cas and qPCR eDNA assays for the detection of Atlantic salmon (*Salmo salar* L.). Environmental DNA. https://doi.org/10.1002/edn3.174

Williams M-A, O'Grady J, Ball B, Carlsson J, de Eyto E, McGinnity P, Jennings E, Regan F, Parle-McDermott A (2019) The application of CRISPR-Cas for single species identification from environmental DNA. Molecular Ecology Resources 19: 1106–1114. https://doi.org/10.1111/1755-0998.13045

Williams MR, Stedtfeld RD, Engle C, Salach P, Fakher U, Stedtfeld T, Dreelin E, Stevenson RJ, Latimore J, Hashsham SA (2017) Isothermal amplification of environmental DNA (eDNA) for direct field-based monitoring and laboratory confirmation of Dreissena sp. PloS One 12: e0186462. https://doi.org/10.1371/journal.pone.0186462

Wilson IG (1997) Inhibition and facilitation of nucleic acid amplification. Applied and Environmental Microbiology 63: 3741–3751. https://doi.org/10.1128/aem.63.10.3741-3751.1997

Woldt A, McGovern A, Lewis T, Tuttle-Lau M (2020) Quality assurance project plan eDNA monitoring of Bighead and Silver Carps. Prepared for: U.S. Fish and Wildlife Service USFWS Great Lakes Region 3 Bloomington, MN 2020. https://www.fws.gov/midwest/fisheries/eDNA/documents/QAPP.pdf

Yamamoto S, Masuda R, Sato Y, Sado T, Araki H, Kondoh M, Minamoto T, Miya M (2017) Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. Scientific Reports 7. https://doi.org/10.1038/srep40368

Yates MC, Glaser DM, Post JR, Cristescu ME, Fraser DJ, Derry AM (2020) The relationship between eDNA particle concentration and organism abundance in nature is strengthened by allometric scaling. Molecular Ecology. https://doi.org/10.1111/mec.15543

Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G, Vaughan R, Hunter C, Park J, Morrison N, Rocca-Serra P, Sterk P, Arumugam M, Bailey M, Baumgartner L, Birren BW, Blaser MJ, Bonazzi V, Booth T, Bork P, Bushman FD, Buttigieg PL, Chain PSG, Charlson E, Costello EK, Huot-Creasy H, Dawyndt P, DeSantis T, Fierer N, Fuhrman JA, Gallery RE, Gevers D, Gibbs RA, San Gil I, Gonzalez A, Gordon JI, Guralnick R, Hankeln W, Highlander S, Hugenholtz P, Jansson J, Kau AL, Kelley ST, Kennedy J, Knights D, Koren O, Kuczynski J, Kyrpides N, Larsen R, Lauber CL, Legg T, Ley RE, Lozupone CA, Ludwig W, Lyons D, Maguire E, Methé BA, Meyer

F, Muegge B, Nakielny S, Nelson KE, Nemergut D, Neufeld JD, Newbold LK, Oliver AE, Pace NR, Palanisamy G, Peplies J, Petrosino J, Proctor L, Pruesse E, Quast C, Raes J, Ratnasingham S, Ravel J, Relman DA, Assunta-Sansone S, Schloss PD, Schriml L, Sinha R, Smith MI, Sodergren E, Spo A, Stombaugh J, Tiedje JM, Ward DV, Weinstock GM, Wendel D, White O, Whiteley A, Wilke A, Wortman JR, Yatsunenko T, Glöckner FO (2011) Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. Nature Biotechnology 29: 415–420. https://doi.org/10.1038/nbt.1823

Yoder M, King IW, De Ley IT, Mann J, Mundo-Ocampo M, Poiras L, De Ley P, Blaxter M (2006) DESS: a versatile solution for preserving morphology and extractable DNA of nematodes. Nematology 8: 367–376. https://doi.org/10.1163/156854106778493448

Yu DW, Ji YQ, Emerson BC, Wang XY, Ye CX, Yang CY, Ding ZL (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods in Ecology and Evolution / British Ecological Society 3: 613–623. https://doi.org/10.1111/J.2041-210x.2012.00198.X

Zenker MM, Specht A, Fonseca VG (2020) Assessing insect biodiversity with automatic light traps in Brazil: Pearls and pitfalls of metabarcoding samples in preservative ethanol. Ecology and Evolution 10: 2352–2366. https://doi.org/10.1002/ece3.6042

Zepeda-Mendoza ML, Bohmann K, Carmona Baez A, Gilbert MTP (2016) DAMe: a toolkit for the initial processing of datasets with PCR replicates of double-tagged amplicons for DNA metabarcoding analyses. BMC Research Notes 9: 255. https://doi.org/10.1186/s13104-016-2064-9

Zimmermann J, Jahn R, Gemeinholzer B (2011) Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. Organisms Diversity & Evolution 11: 173–192. https://doi.org/10.1007/s13127-011-0050-6

Zizka VMA, Elbrecht V, Macher JN, Leese F (2019a) Assessing the influence of sample tagging and library preparation on DNA metabarcoding. Molecular Ecology Resources. https://doi.org/10.1111/1755-0998.13018

Zizka VMA, Leese F, Peinert B, Geiger MF (2019b) DNA metabarcoding from sample fixative as a quick and voucher-preserving biodiversity assessment method 1. Genome / National Research Council Canada = Genome / Conseil national de recherches Canada 62: 122–136. https://doi.org/10.1139/gen-2018-0048

Zulkefli NS, Kim K-H, Hwang S-J (2019) Effects of Microbial Activity and Environmental Parameters on the Degradation of Extracellular Environmental DNA from a Eutrophic Lake. International journal of Environmental Research and Public Health 16. https://doi.org/10.3390/ijerph16183339

12. FIGURE INFORMATION AND PHOTO CREDITS

Figure 1: Practical Guide overview.
R.C. Blackman, Illustrator version 25.2.3.

Figure 2: Water eDNA sample design.
R.C. Blackman, Illustrator version 25.2.3.

Figure 3: Water eDNA filter types.
a: R.C. Blackman, eawag
b: K. Panksep, Estonian University of Life Sciences
c: T. Macher, University of Duisburg-Essen
d: R.C. Blackman, Eawag
e: K. Bruce, NatureMetrics
f: K. Panksep, Estonian University of Life Sciences

Figure 4: Water eDNA collection and filtration methods.
a: M, Hellstrom, MIX Research
b: M. Seymour, The University of Hong Kong
c: A. Sheard
d: Smith-Root Inc
e: T. Macher, University of Duisburg-Essen
f. K. Panksep, Estonian University of Life Sciences
g: A. Valentini, SPYGEN
h: R. Schuetz, University Duisburg-Essen

Figure 5: DNA homogenisation methods.
a: A. Lindner, Zoological Research Museum Alexander Koenig
b: V. Ebrecht, ETH
c: A. Lindner, Zoological Research Museum Alexander Koenig
d: A. Lindner, Zoological Research Museum Alexander Koenig
e: R. Donnelly, University of Hull
f: F. Leese, University of Duisburg-Essen
g: F. Leese, University of Duisburg-Essen

Figure 6: Biofilm sampling.
a: V. Vasselon, Scimabio Interface
b: A. Bouchez, INRAE
c: V. Vasselon, Scimabio Interface
d: S. Lacroix, SYNAQUA project

Figure 7: Metabarcoding amplicon construct.
R.C. Blackman, Illustrator version 25.2.3.

This book represents a synthesis of knowledge and best practice in the field of DNA-based biomonitoring at the time of writing. It has been written with end-users of molecular tools in mind, as well as those who are new to the field in research settings and are looking to gain an overall grounding in the subject area. For each of the main types of sample (water, soil / sediment, bulk invertebrates and diatoms), and for each stage of the field and laboratory processes, we outline key considerations, decisions that need to be made, factors that might influence those decisions, and trade-offs inherent in the choices made. We hope that this will help users, practitioners, and those commissioning DNA-based monitoring programmes to navigate this large field and critically evaluate the strengths and weaknesses of different analysis workflows based on context, project aims and available resources.

9 786192 480523

Cover photo: © P. Bjorn Erikson

cost
EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Funded by the Horizon 2020 Framework Programme of the European Union

DNAqua-Net
EU COST ACTION

PENSOFT