FROM **MOLECULES** TO **MONITORING**

Integrating genetic tools into freshwater quality assessments

Kevin K. Beentjes

FROM **MOLECULES** TO **MONITORING**

Integrating genetic tools into freshwater quality assessments

Kevin K. Beentjes

Kevin K. Beentjes, 2021. From molecules to monitoring: Integrating genetic tools into freshwater quality assessments.
PhD thesis, University of Leiden, The Netherlands

ISBN: 978-94-6423-157-1

Layout and cover design by Kevin K. Beentjes
Printed by ProefschriftMaken

This research was carried out at Naturalis Biodiversity Center and the Institute of Biology Leiden at Leiden University. It was part of the project DNA Waterscan, funded by the Gieskes-Strijbis Fonds.

FROM MOLECULES TO MONITORING
Integrating genetic tools into freshwater quality assessments

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden
op gezag van rector magnificus prof.dr.ir. H. Bijl
volgens besluit van het college voor promoties
te verdedigen op donderdag 8 april 2021
klokke 10:00 uur

door

Kevin Kay Beentjes
geboren te Zaanstad, Nederland
in 1987

Promotor	Prof. Dr. Menno Schilthuizen *Universiteit Leiden*
Copromotor	Dr. Arjen G.C.L. Speksnijder *Naturalis Biodiversity Center*
Promotiecommissie	Prof. Dr. Gilles P. van Wezel (voorzitter) *Universiteit Leiden*
	Prof. Dr. Koos Biesmeijer (secretaris) *Universiteit Leiden*
	Prof. Dr. Kristy Deiner *ETH Zurich*
	Prof. Dr. Florian Leese *Universität Duisburg-Essen*
	Prof. Dr. Nicole J. de Voogd *Universiteit Leiden*

Voor mijn ouders

CONTENTS

CHAPTER 1

General introduction and literature review

Kevin K. Beentjes

Parts of this text and its illustrations have been adapted to
STOWA Deltafact: DNA-technieken voor waterbeheerders.
www.stowa.nl/deltafacts/waterkwaliteit/diversen/dna-technieken-voor-waterbeheerders

1.1 THE STATE OF FRESHWATER AND ITS INHABITANTS

The continuing anthropogenic decline of the Earth's biodiversity (Barnosky et al. 2011) is one of the most serious threats of the 21st century, with no outlook on significant reduction in the rate of biodiversity loss (Butchart et al. 2010). Human influences have created extinction rates that go beyond those of pre-human periods, current estimates being a thousand-fold the expected background rate, in what has been dubbed "the sixth extinction wave" (Pimm et al. 2014, Dirzo et al. 2014).

Freshwater species appear to be at a greater threat than terrestrial and marine species. Freshwater ecosystems contain a rich diversity of both taxa and habitats, despite the fact that they cover less than one percent of the Earth's surface. Of all water, 2.5% is freshwater, with only 1.2% of freshwater being surface water (Gleick 1993). The 2014 Living Planet Report presented an average decline in size of monitored populations of 76% in freshwater, against 39% for both terrestrial and marine biomes (WWF 2014). The main drivers that threaten freshwater species are habitat loss or degradation, pollution of water, over-exploitation, flow modification and invasive species, the first being the most prevalent by far (Dudgeon et al. 2006, Collen et al. 2014).

Freshwater habitats are, in essence, islands within a sea of dry land or salt water, creating barriers that are unbridgeable for many species living in these ecosystems. This physical isolation makes for limited dispersal opportunities across these islands. The insular nature of freshwater ecosystems has led to the evolution of species with small geographic ranges, and resulted in biotas with high rates of endemism and turnover (Strayer 2006). This fragmentation and relatively high proportion of endemism greatly reduce the ability of freshwater taxa to respond to environmental change, as they limit the ability to freely disperse and re-establish local populations that have been extirpated. This makes those freshwater species that do not have large geographic ranges especially sensitive to human impacts (Strayer & Dudgeon 2010).

Freshwater invertebrates form a phylogenetically diverse group, which are usually not well studied in terms of conservation biology. Hence, they often receive different or less protection than their vertebrate co-occupants of freshwater habitats. Invertebrates live in most freshwater sources, save for the most polluted waters. Densities of all freshwater invertebrates together range between 10^5 and 10^6 individuals per cubic meter (Wetzel 2001), and although the inventories of freshwater invertebrates, even macroinvertebrates, are often incomplete, local faunas may contain hundreds, if not thousands of species (Strayer 2006). The distribution, species richness, and threatened-species richness data for vertebrate taxa show little congruence with those of invertebrate taxa in freshwater (Collen et al. 2014), making

the well-studied fish and amphibians imperfect indicators of macroinvertebrate communities. Conservation statuses for freshwater vertebrates may therefore not be suitable proxies for those of invertebrate taxa in the same habitats (Dudgeon et al. 2006). Data on the geographical distribution ranges and relative extinction risks are limited, but it is expected that the small ranges that many aquatic invertebrates exhibit will be even more dissimilar from the few large-bodied groups that have been studied (amphibians, fish, mammals, reptiles, and crustaceans) (Collen et al. 2014).

Freshwater is not only important for the life that it contains, but also for most other organisms living on our world, including human beings. Our species already uses over half of the accessible global freshwater runoff, with demand steadily increasing (Jackson et al. 2001). Rapid changes in the use of freshwater are causing dramatic changes in patterns of water stress, and we are close to overstepping the limits set in the planetary boundaries for global sustainability (Alcamo et al. 2008, Rockström et al. 2009). Hence, managing water quality is not only important for the aquatic flora and fauna, but also for the ecosystem services that are essential to the well-being and health of mankind (Corvalan et al. 2005).

1.2 THE MONITORING OF BIOLOGICAL QUALITY

Because of its importance, in the past decades a range of monitoring methods have been developed to assess the "health" of freshwater, ranging from abiotic properties to multimetric biological indices. The concept of health within an ecological context has seen much debate in the last decade of the 20th century, with critics insisting that it is not an observable ecological property, and merely a property of the organisms within an ecosystem, or that "preferred ecosystem states" cannot be well-defined (Suter 1993, Scrimgeour & Wicklum 1996). The discussion was complicated by the involvement of societal values, where people argued that health is dependent on human values, and that efforts to protect the health of ecosystems should consider the "human uses and amenities derived from the system" (Rapport 1989, Regier 1993). Supporters of the concept of health in an ecosystem setting have often looked for more objective and scientifically relevant arguments, such as a system's primary productivity, species diversity and connectivity, and resiliency to stress, as well as the interactions between such variables (Costanza 1992). Others have argued that the heavy reliance on ecological theories without any form of validation in the real world might lead to inadequate public environmental policies, and in essence, mislead society (Karr 1999). To illustrate his case, Karr provides examples of situations in which a tropical forest may be classified as more healthy than a spruce-fir forest,

based solely on the fact that it is more diverse and has a higher primary production, or where a community of oligochaete worms at a wastewater treatment outflow may be classified as healthy because of their resiliency to disturbances.

Most parties involved in freshwater quality monitoring and management can agree with the fact that health is an important aspect of systems, especially in those that are of importance to human health, such as freshwater bodies. Directive 2000/60/EC of the European Parliament and of the Council established "a framework for [European] Community action in the field of water policy". In this EU Water Framework Directive (WFD), the European Parliament states that water is "not a commercial product like any other but, rather, a heritage which must be protected, defended and treated as such". The WFD reiterates the declaration of the 1991 Ministerial Seminar on groundwater, that argued for a need for "action to avoid long-term deterioration of freshwater quality and quantity" and called for a program of actions aiming at sustainable management to be implemented by the year 2000 (European Union 2000). The WFD emphasizes the importance of freshwater organisms, as the composition of their communities is now used to determine the condition of water bodies, and therefore defines the need for restoration efforts and investments. The annexes of the WFD provide normative definitions of ecological status classifications, which include quality elements from hydromorphology (e.g. hydrological regime, river continuity), physico-chemistry (e.g. specific synthetic or non-synthetic pollutants) and biology (e.g. phytoplankton, benthic invertebrate fauna, fish fauna). Each quality element can be categorized as high, good, moderate, poor or bad, but all are defined in comparison to totally, or almost totally, undisturbed conditions. The comparison to undisturbed, ideal communities as a reference, however, also makes that this approach provides a general valuation of the biodiversity itself, and not just evaluates the classical response of indicator species (Schmidt-Kloiber & Hering 2015). Almost 300 different assessment methods for biological quality are in use in Europe alone (Birk et al. 2012), many focusing on invertebrate surveys to calculate Ecological Quality Ratios (EQRs). In the Netherlands, the measures set forth in the WFD are implemented in the Kader Richtlijn Water (KRW). The KRW assesses quality on a scale of 0 to 1, subdivided into five quality classes: "bad"(EQR <0.2), "poor" (0.2–0.4), "moderate" (0.4–0.6), "good" (0.6–0.8) and "high" (0.8–1.0) (Evers et al. 2012, van der Molen et al. 2016) (Figure 1.1).

While some common methods will employ the use of physical and chemical properties of the water, such as levels of dissolved oxygen, acidity or turbidity, these parameters only offer a snapshot of the actual conditions, and fail to provide a more integrative measure of the overall condition of a water body. It may therefore be

insufficient to recognize impaired waters (Kenney et al. 2009). Instead, biological indicators—or bio-criteria—are better capable of offering a more integrated assessment of the health of water bodies (Karr 1999). These bio-criteria use measures of biological communities, spanning multiple trophic levels. Policies such as the EU WFD have already adopted the use of different trophic levels, as witnessed by the inclusion of phytoplankton, macrophytes and phytobenthos, macroalgae, benthic invertebrate fauna, and fish fauna as biological quality elements for the different water types defined in the annexes of its establishing directive. Barbour et al. (1999) summarized the advantages of using biological communities for monitoring, or bio-surveys, as reflecting overall ecological integrity and integrating stressors and stresses

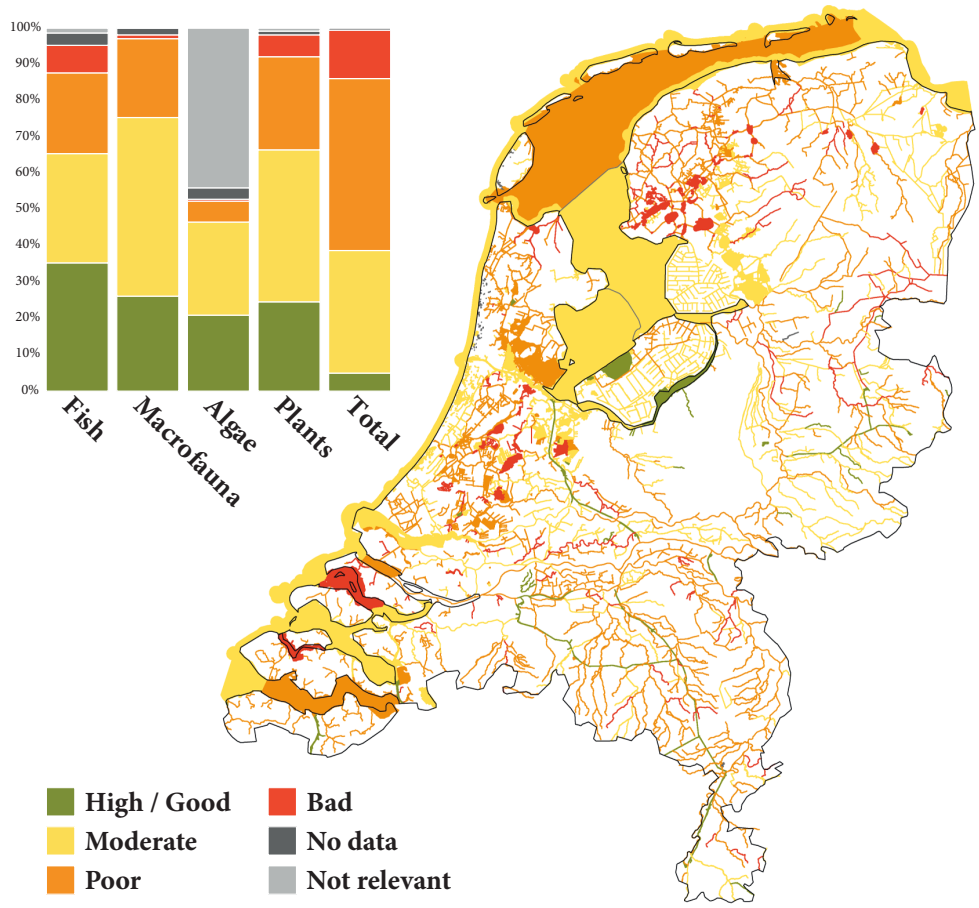


FIGURE 1.1. Ecological quality ratio (EQR) scoring of Dutch water systems in 2015, based on fish, macrofauna, algae, and plants. Map represents the total EQR score of all four elements combined. Data and map obtained from Planbureau voor de Leefomgeving (www.clo.nl/nl142003).

over time to provide a measure of fluctuating environmental conditions. In addition, the routine monitoring of biological communities is argued to be relatively expensive when compared to the assessment of toxic pollutants with chemical tests.

The use of benthic macroinvertebrates for biomonitoring has several advantages. Macroinvertebrate communities are abundant, and reflect localized conditions due to the limited migration patterns of many taxa. This makes them suitable for the assessment of site-specific impacts, such as those measured in upstream-downstream studies. The invertebrate communities are made up of species that represent a broad range of trophic levels, ecological functions and tolerances to stressors. Experienced identifiers can easily recognize most taxa, including the taxa that are most sensitive to changing conditions, allowing for even cursory examinations to yield insights into water quality conditions. In comparison to fish monitoring, the sampling of benthic macroinvertebrates is relatively easy and cheap, with minimal effect of the sampling on the resident biota (Barbour et al. 1999, Kenney et al. 2009).

The advantages already highlight one of the major disadvantages of the use of benthic macroinvertebrates, or any taxonomic group in that respect, as it calls for experienced identifiers. The sheer amount of species in the macroinvertebrate assessments, spread over a huge range of taxonomic groups, requires multiple specialists that divide the workload among the different taxa, or identifiers who are familiar with at least hundreds of species. Such expertise is rare and decreasing: for example, an inquiry among British taxonomists revealed a continuing decline both professional and amateur taxonomists (Hopkins & Freckleton 2002). This taxonomic impediment is furthermore seen in the decreasing number of taxonomic courses offered at universities, and the difficulties encountered by when applying for funding for taxonomic activities by researchers (Drew 2011).

1.3 THE QUALITY OF BIOLOGICAL MONITORING

Variations observed in the macroinvertebrate community—or any biological community—can have several origins. First of all, there are the effects of pollution or other environmental stressors, which are usually the variations that water quality assessments attempt to detect and quantify. Second, there is a natural variation in time, caused by other factors than stress or pollution. Seasonality is the main cause behind this temporal variation. An assessment of macroinvertebrate communities during a one-year period revealed that the best results are obtained by sampling twice a year, in early spring and in late autumn, whereas sampling in summer and winter months is discouraged due to strong seasonal influences and logistical reasons,

respectively (Šporka et al. 2006). Finally, there is variation that occurs during the assessment itself. These have been classified as (1) variations in sampling or sampling methods, (2) sample processing errors, and (3) sample identification errors (Clarke & Hering 2006). The last two steps seem to be the source of most inconsistencies: regular quality control of sample processing is lacking in most laboratories, and only appears to be implemented in the United Kingdom (Haase et al. 2010). Both are labor-intensive, making them susceptible to human errors. Other stages of the water quality assessments at risk of human error include site selection, data entry, and interpretation of the data (Clarke & Hering 2006). Studies have found a significant amount of human errors in the sorting and identification processes, which impacted most of the functional metrics used in water quality assessments (Haase et al. 2006).

During a national survey of streams in the United States, a detailed evaluation of 74 benthic macroinvertebrate samples revealed a taxonomic disagreement of 21% between primary analyst and auditor. This percentage decreased in a second round of evaluations, after primary analyst and auditor communicated and corrective actions were implemented (Stribling et al. 2008). Similar conclusions were drawn in Germany, where an audit on the water quality monitoring program of German streams and rivers was performed. In this audit, 50 out of 414 macroinvertebrate samples were scrutinized on sorting level, identification level, and the combination of both levels. Samples were collected by different commercial laboratories using EU WFD protocols. The human errors were substantial, with 29% of all specimens overlooked during the sorting process by the primary analysts, which led to one in five species being excluded from further analysis. The identification audit revealed that roughly one third of the taxa were different between the primary analyst and the auditor. One of the surprising results was that the error rate was not higher in taxa considered difficult to identify, as compared to those considered easy to recognize. It is postulated that this is caused by the fact that identifiers unconsciously paid less attention to "easy taxa". In the end, about a sixth of all samples was placed into a different ecological assessment compared to the original assessment (Haase et al. 2010).

Several studies identify similar taxonomic groups that are difficult to identify, such as Baetidae, Chironomidae and Hydropsychidae, which are dependent on freshwater during their larval stages. This indicates that the underlying problem is not just a lack of expertise in the audited studies, but that these groups may pose a challenge in general. Inventories of streams and rivers in the United States indicate that there may be several hundreds, or even over a thousand species that live in monitored sections of water, with most of them only identifiable using adult male specimens or relatively late juvenile stages (Jackson et al. 2014). As a considerable amount of

the collected material consists of those stages hardest to identify to species level on morphological grounds, the results of the few published audits do not come as a surprise. This susceptibility to errors in traditional assessments urges the monitoring efforts of freshwater macroinvertebrates to find alternatives for the identification of collected material, or better yet, skip the time-consuming collecting and sample handling completely.

1.4 INTEGRATION OF MOLECULAR TOOLS

Identifying specimens from freshwater samples to species level based on morphology alone remains a challenge. The taxonomic knowledge about various groups of organisms is often rudimentary, and dichotomous keys are of limited use due to the variation in morphology within benthic macroinvertebrate species. In addition, most of these taxa are small, have few accessible morphological characters, often have closely related species, and often need to be reared to adulthood since pre-adult stages are usually not covered in identification keys (Jones 2008). Several studies have looked into the costs of morphological identifications, and the general conclusion is that the difference in cost between order- or family-level on the one hand and genus- or species-level identifications on the other is considerable, whereas the additional costs to increase the resolution from genus- to species-level are relatively modest (Marshall et al. 2006).

Jackson et al. (2014) state that most studies that assign macroinvertebrates to the lowest possible taxonomic level, generally leave around half of the individuals identified at genus level or higher. They argue that the use of molecular methods will enable assessments to take full advantage of all collected specimens, and in turn may even lead to new species-specific insights on ecology and regulations. Species designations and delineations based on DNA barcodes seem to be in good agreement with those based on morphology, ecology or even behavior. DNA barcoding, the technique of using short fragments of molecular data to identify species, has been around for decades. The use of DNA barcodes for species identification grew tremendously after the introduction of the roughly 650 base pair long mitochondrial COI barcode in the early 2000s (Hebert et al. 2003). While there was some initial doubt about the acceptance of DNA barcoding (DeSalle et al. 2005), the continued growth of the Barcode of Life Database (Ratnasingham & Hebert 2007), together with the sheer number of citations of the original publication from 2003 (well over 11,000 at the time of writing), can be seen as proof of the effectiveness and acceptance by the scientific community. The technique has become embedded in the daily work of

many biologists, and many papers about DNA barcoding are still published each year (Figure 1.2). DNA barcoding has enabled the use of improved taxonomic resolutions, reduced costs and a reduction in the human error in identifications (Pauls et al. 2014). Pauls et al. summarized the benefits of the application of molecular tools in freshwater science as (1) the ability to characterize spatial patterns in diversity on a broader range of taxa, with much greater resolution, (2) the ability to assess functional genetic variation and responses to environmental changes, and (3) increased speed and taxonomic resolution in assessing current status of freshwater.

In the case of Sweeney et al. (2011), for example, the use of DNA barcodes allowed for the identification of many more taxa than with morphology alone. When comparing DNA barcode generated taxon data to expert level inventories on genus and species level, they found a 125% and 70% increase, respectively. When comparing to amateur level identifications they even found a 475% (124 taxa) increase. Using barcodes also revealed additional species that were not described in larval keys, as well as coexisting congeners that may well have been missed due to morphological similarity. Increases in species richness in taxonomic inventories were reported by others as well, such as Jackson et al. (2014). They recovered 104 more species based on DNA barcoding, which amounted to a 108% increase in species richness estimations. Results were best for some of those groups which have been described as "difficult",

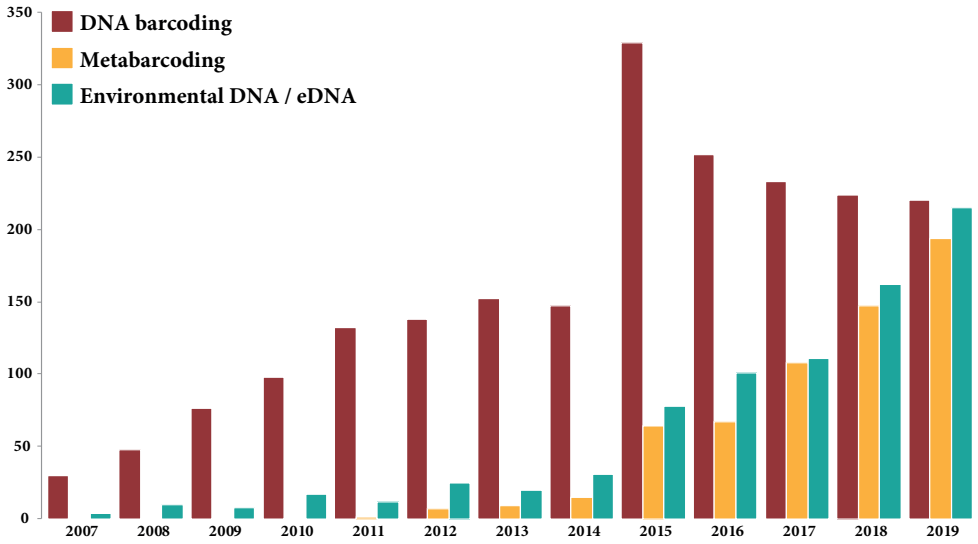


FIGURE 1.2. The number of indexed papers published from 2007 to 2019 on DNA barcoding, metabarcoding, and environmental DNA. Data was retrieved from Web of Science (https://www.webofknowledge.com), based on papers with titles containing "DNA barcoding", "metabarcoding", and "environmental DNA" or "eDNA", respectively.

such as the Chironomidae (194% increase) and Ephemeroptera (77% increase), but also Acari (200% increase). Species identified with DNA were often species known to be uncommon or usually only found in small numbers. Similarly, when creating species lists for alpine lakes, Deiner et al. (2013) found that in about 75% of the cases where young or damaged individuals could not be identified using morphological characters, DNA barcodes allowed for identification up to species or genus level.

Being able to generate species-level taxon lists for freshwater communities also enables the use of species-level ecological characteristics and traits. Even though these may not yet exist for all taxa, as the term "species traits" is often used to refer to genus- or family-level characteristics, their use would greatly improve the ability to reliably identify subtle changes in community structure, and therefore in water quality (Jackson et al. 2014). It has already been shown that even without binomial taxonomic names, DNA barcoding can distinguish between putative species that show differing responses to environmental stressors. DNA barcoding of mayflies in New Zealand stream sites revealed up to twelve different clades or cryptic species, which had contrasting tolerances to common environmental stressors (Macher et al. 2016). Similarly, sequencing of chironomids from a mesocosm experiment showed different response patterns for different biological entities, even though the majority of these operational taxonomic units (OTUs) could not be identified due to lacking references (Beermann et al. 2018). These studies indicate that even with an incomplete reference library or unresolved cryptic species complexes, DNA barcodes provide higher-resolution taxonomic information that can be used for assessments.

1.4.1 High-Throughput Sequencing

Methodological advances in the past decade have led to a situation in which research can focus more on the merger of molecular biology and ecology, and less on the design of studies around technical restrictions. As shown before, the use of DNA barcodes can provide easier and more reliable (at least more standardized) identifications of macroinvertebrates, especially where it concerns pre-adult life stages. To increase the applicability of molecular identification techniques further, DNA barcodes, or any marker for that matter, can be used for simultaneous identification of multiple taxa in complex samples, via DNA metabarcoding (Taberlet et al. 2012a). At the base of this lies what is often called next-generation sequencing (NGS), even though such techniques are nowadays common use, and better refered to as high-throughput sequencing (HTS). Limitations in sequencing platforms at the time, however, made the full COI barcode region unsuitable for use, as its length exceeded the maximum of most platforms. This again spurred some debate as to whether COI was the right

marker for molecular approaches, as it may not contain many suitable conserved regions for broad-spectrum taxonomic coverage (Deagle et al. 2014). Alternatives, such as ribosomal RNA were offered in place of COI, mainly due to the possibility to obtain shorter amplicons. However, the benefit of the COI barcode library, which has much better taxonomic coverage than any other gene for metazoan diversity, has made that COI is still the marker of choice in many studies (Elbrecht et al. 2016, Andújar et al. 2018b), with the exception of fish, where ribosomal markers (such as 12S or 16S) are used, as they allow for better primer design (e.g. Valentini et al. 2015, Fujii et al. 2019).

While COI reference libraries are far from complete, they are sufficiently populated with most of the commonly observed freshwater macroinvertebrates

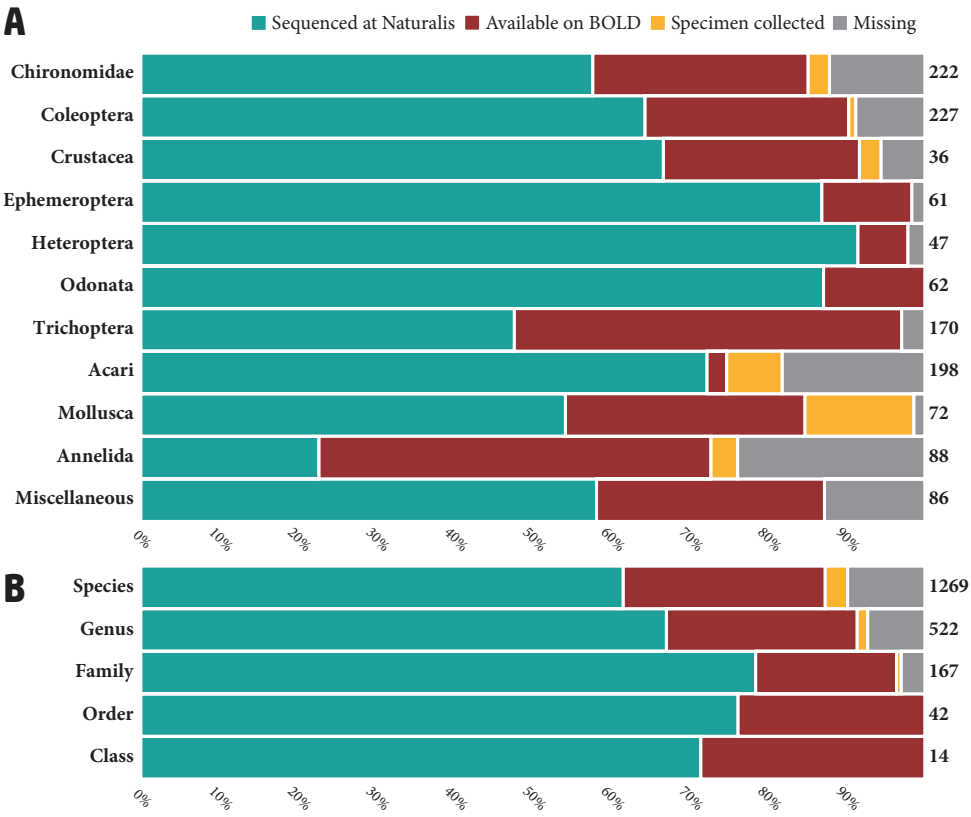


FIGURE 1.3. Percentages of taxa from the Dutch WFD list covered by DNA reference libraries, from the Dutch barcoding campaign at Naturalis Biodiversity Center (as of May 2020), and public data available on the Barcode of Life Database (http://boldsystems.org/). Data is shown for (A) the most important groups of taxa and (B) five taxonomic levels of identification. Indicated to the right of each bar is the number of taxa for each category.

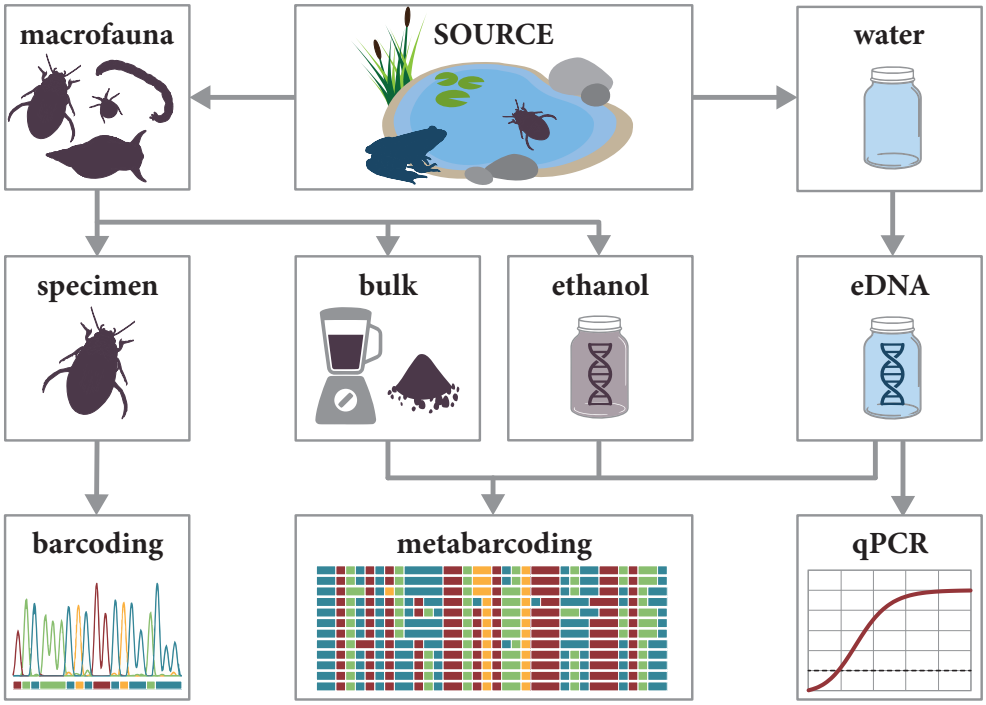


FIGURE 1.4. A schematic overview of the main techniques used in water quality monitoring, based on the type of samples. Individual specimens collected using traditional techniques are essential for the creation of a reliable reference database using single-species DNA barcoding. Bulk, ethanol from bulk, and directly collected environmental DNA can be used for DNA metabarcoding. Species-specific PCR detection is best performed on eDNA samples. Original illustration.

to allow for the use of COI in routine monitoring applications (Curry et al. 2018) (figure 1.3). Shorter fragments of COI can still be used to separate closely-related taxa (Meusnier et al. 2008), and several primer sets have been developed on shorter amplicon lengths and have proven to be successful in capturing relevant groups for biodiversity monitoring, such as marine metazoa (Leray et al. 2013) and freshwater macroinvertebrates (Elbrecht & Leese 2017). DNA metabarcoding techniques have been used for identifications of specimens in bulk samples, simply by homogenizing the samples and performing DNA extractions on the resulting "slurry" (Hajibabaei et al. 2011, Gibson et al. 2015) (figure 1.4). In addition to the bulk samples obtained from biodiversity monitoring programs, the use of DNA metabarcoding has also proven its worth in diet studies, where the original bulk data is impractical to use for morphological identification, due to the degraded state of most tissue found in gut contents or fecal matter (Pompanon et al. 2012, Gibson et al. 2014, Corse et al. 2017).

Non-destructive DNA metabarcoding has also been performed on storage ethanol of bulk samples, leaving the specimens intact and available for further study (figure 1.4). In 2010, it was shown specimens stored in ethanol "leak" DNA into the preservative, when DNA barcodes were obtained from both freshly stored and archival specimens (Shokralla et al. 2010). Metabarcoding studies on preservative ethanol show promise, in some cases obtaining more species than bulk metabarcoding, although there are still differences with traditional morphological assessments, especially where it considers species that represent low proportions of the total biomass of a sample (Hajibabaei et al. 2012, Erdozain et al. 2019).

1.4.2 Environmental DNA

While DNA metabarcoding can potentially replace morphological identifications, and the sampling of preservative ethanol is a non-destructive approach, both still require traditional sampling of specimens. However, in the last decade a new method has become popular in the field of molecular biomonitoring of multicellular organisms: environmental DNA (eDNA) (figure 1.4). Inspired by studies that retrieved ancient DNA from sediment or ice cores (e.g. Willerslev et al. 2003), Ficetola et al. (2008) showed they were able to detect the presence of the invasive American bullfrog (*Lithobates catesbeianus*) in both controlled aquarium setups and natural ponds in France by sampling water and precipitating organic material contained therein. Since then, the use of environmental DNA for the detection of species diversity has increased rapidly (Taberlet et al. 2012b, Thomsen & Willerslev 2015), which is reflected in the number of papers growing steadily each year (Figure 1.2).

Many early papers dealing with eDNA in freshwater and the marine environment focused on single-species detection using specific primer/probe sets. These allowed for the amplification of only target DNA in real-time PCR, resulting in a "yes" or "no" (and quantitative indication) without the need for sequencing any DNA. A fair number of these studies used invasive species as a model organism, as they are relevant for ecosystem management. In these cases, eDNA could provide an "early warning" insight system, in which it would theoretically be possible to detect presence of invasive species in early stages without intensively sampling systems using traditional methods. The majority of the papers employing eDNA for such detections focused on amphibians (e.g. Dejean et al. 2012, Smart et al. 2015) or fish (e.g. Jerde et al. 2011, Takahara et al. 2013), organisms known to shed relatively large amounts of DNA into the water column, in comparison to many hard-bodied macroinvertebrates, although there have been several studies that showed eDNA is also usable to detect invasive crayfish (Tréguier et al. 2014, Agersnap et al. 2017). In

a similar fashion, species-specific PCR assays have also been used for the detection of rare, endangered or policy-relevant species. This effectively allows monitoring agencies to cover more terrain by simply sampling water at any location of interest and foregoing the invasive and labor-intensive traditional assessment methods, and has been coined "conservation in a cup of water" (Lodge et al. 2012). As with the invasive species, most of the target organisms in these eDNA studies are vertebrates, ranging from amphibians (Goldberg et al. 2011, Spear et al. 2015) to fish (Sigsgaard et al. 2015) and cetaceans (Stewart et al. 2017), although several inquiries have been made into invertebrates as well (Thomsen et al. 2012b, Mächler et al. 2014), and even aquatic plants (Matsuhashi et al. 2016). The use of eDNA for the detection of single species has been shown to be less labor-intensive and more cost-effective than traditional monitoring, but require well-planned sampling strategies adapted to the target organisms (Smart et al. 2016, Evans et al. 2017a, Lugg et al. 2018).

Important with species-specific assays is the specificity and sensitivity of the primer/probe sets used for detection, especially when there are closely-related species that might provide false positive signal (Wilcox et al. 2013). Thomsen et al. (2012) also showed, however, that high-throughput sequencing was also possible for species detection of fish and amphibians, which makes that primers may not necessarily have to be species-specific. On the contrary, HTS allows for the simultaneous sequencing of multiple organisms, so in theory it is best used in combination with primers that are not species-specific. Environmental DNA metabarcoding has become more and more mainstream, and forms the final stage in the transition from traditional monitoring to molecular biomonitoring: (1) replace taxon identification by DNA barcoding, (2) replace specimen handling by DNA metabarcoding, and (3) replace traditional sampling by environmental DNA. Several primer sets have been developed for the monitoring of fish (and other vertebrates), both freshwater and marine (Thomsen et al. 2012a, Miya et al. 2015, Valentini et al. 2015, Andruszkiewicz et al. 2017), which often show congruence with traditional inventories (e.g. Hänfling et al. 2016). In addition to fish, group-specific assays have been developed for several other relevant organism groups for biomonitoring, such as indicator chironomids (Bista et al. 2017), invasive molluscs (Klymus et al. 2017), and mosquitoes that act as disease vectors (Schneider et al. 2016, Krol et al. 2019). For good comparison with traditional monitoring, such group-specific primers would be optimal, as general primers often amplify a wide range of non-target taxa.

1.5 NEXT-GENERATION BIOMONITORING

The developments in sequencing and alternative DNA sources also brought with them an increased resolution of information to be obtained from natural systems. While many European systems that assess macroinvertebrates rely on species-level identifications for the resulting EQR (at least on paper), there are still others in which only higher-level identifications are used, as well as many taxonomic groups for which species-level identifications are not always possible. However, some challenges still need to be overcome to allow for full incorporation into standard monitoring practice, depending on the taxonomic groups assessed (Hering et al. 2018). Whereas the use of DNA barcoding may not be cheaper than morphological analysis, costs will be driven down by adopting metabarcoding as method for taxon identification, especially when compared to species-level identifications using morphology (Jones 2008, Stein et al. 2014, Aylagas et al. 2018). Newly emerging techniques for high-throughput sequencing will make sequencing even cheaper in the near future. It also opens up possibilities for groups that are now largely ignored due to their difficulty with identifications, such as planktonic taxa, to be included in routine monitoring or impact studies. While the inclusion of any such groups needs a completely new method of assessment, the tools are already largely available to start working towards this "biomonitoring 2.0" (Baird & Hajibabaei 2012, Pawlowski et al. 2018). It has already been shown in several studies that genetic assessments of biodiversity can yield significantly different results compared to morphological assessments, although both genetic and traditional surveys can complement each other and present a more complete view of the ecosystem (Shaw et al. 2016, Kelly et al. 2017).

Genetic assessments have already been performed and compared to traditional monitoring in several studies. For marine monitoring, a genetics-based version of the AZTI's Marine Biotic Index (gAMBI) has been proposed and compared to the traditional AMBI index, in which there was a moderately strong correlation between the two (Aylagas et al. 2016). The best performing DNA-based method, using the 313 bp fragment by Leray et al. (2013), resulted in a 62.4% match between the morphological and molecular taxa list, and 76.3% when only looking at species level. A similar study by the same authors only found about half the taxa in metabarcoding compared to traditional identifications (54.4%), but metabarcoding and traditional data led to similar assessments, comparable to the earlier study. Additionally, the metabarcoding was calculated to be around 55% less costly and 72% less time consuming (Aylagas et al. 2018). A study comparing morphological and metabarcoding-based stream assessment in Finnish monitoring sites, on the other hand, found twice as many

taxa with DNA as with morphology, although that conclusion is somewhat inflated due to the fact that some species-rich groups like chironomids were only identified up to family level in the morphological analysis. The EQR and other assessment metrics were significantly correlated (Elbrecht et al. 2017a), indicating that DNA metabarcoding-based approaches certainly have potential to replace traditional monitoring with the necessary recalibration of metrics involved.

Micro-organisms, such as planktonic taxa or bacteria, are an ideal target group for the use of molecular tools. The sampling methods for the collection of samples are relatively easy, compared to the traditional collection methods employed for macroinvertebrates, and the identification requires microscopic inspection by specialists. These groups are often also more diverse than most macroinvertebrate groups, meaning they might provide a better insight into the ecosystem due to the increased resolution they provide in comparison to relatively species-poor vertebrate and macro-invertebrate groups used in traditional surveys. Diatom indices inferred from metabarcoding data have already been shown to be significantly correlated to morphological assessments, demonstrating the feasibility of applying metabarcoding in such surveys (Visco et al. 2015). The relatively high diversity can, however, also be a limiting factor, since it is likely that taxon groups such as plankton may hide a large proportion of undescribed or understudied taxa, which could cause issues inferring quality assessments. There is still ample opportunity to incorporate these unknown taxa into biomonitoring, using so-called "taxonomy-free" methods. A paper on diatoms showed that three-quarters of the examined sites could be assessed correctly using molecular operational taxonomic units (MOTUs), rather than classic taxonomic assignments (Apothéloz-Perret-Gentil et al. 2017). Unassigned MOTUs were also used to infer assessments based on benthic foraminifera via machine learning techniques, again leading to a similar ecological status as the traditional monitoring using macroinvertebrates (Cordier et al. 2017). The same machine learning techniques were later used to show that different markers (both prokaryote and eukaryote) could accurately predict environmental impact of marine aquaculture, and all outperformed the assessment based on traditional methods (Cordier et al. 2018).

Such impact assessments are a prime target for the use of (environmental) DNA metabarcoding. Not only have these molecular tools the potential to provide much higher resolution views on an ecosystem, impact assessment studies are not necessarily bound to traditional survey methods like the ecological quality assessments dictated by the EU WFD. Several studies have already been performed using metabarcoding for impact assessments in the marine realm, dealing with impacts from fish farms (Pochon et al. 2015, Stoeck et al. 2018, Cordier et al. 2018) and offshore oil and gas

drilling (Laroche et al. 2018). Metabarcoding has also been successfully applied in various impact studies in freshwater in recent years, ranging from metabarcoding of bulk macroinvertebrates samples to investigate pesticide spills (Andújar et al. 2018a) to the collection of planktonic organisms to study the effects of urbanization and wastewater (Chonova et al. 2019, Hanashiro et al. 2019). Environmental DNA obtained from water samples contained few metazoans, making comparison to traditional methods difficult, but nonetheless the MOTUs could still be used to reveal impaired sites (Bagley et al. 2019), demonstrating again that eDNA is a powerful tool in water quality assessment and management.

While DNA techniques have proven useful in various analyses performed for water management, from the detection of invasive species, to the assessment of stressor impacts, there is still some way to go before molecular tools can be integrated into water quality monitoring across the board. Many assessments, especially those performed under the WFD, are still performed using traditional methods, simply because that is, from a legal perspective, the golden standard. The current implementation of the WFD runs until 2027, meaning that there is still time before DNA methods can be legally entered into the standard practices. This leaves room for much needed standardization in some approaches, which would arguably make it easier for policy makers and stakeholders to accept DNA techniques as actual "standards". Some of the discussions about any such standardization are currently taking place in DNAqua-Net, a COST Action network aimed to "nucleate a group of researchers across disciplines with the task to identify gold-standard genomic tools and novel eco- genomic indices for routine application in biodiversity assessments of European fresh- and marine water bodies" (Leese et al. 2016).

There are as many different protocols for DNA metabarcoding as there are labs, and probably even more. The idea that one protocol will arise as "the golden standard" and consequently be used by all players in the field of water quality management sounds utopian. Different nations will likely have their own interpretations of any WFD protocol, as is already the case with the traditional monitoring today. Nonetheless, it is reasonable to have some standards that allow for cross-comparison of results, and to implement some basic guidelines which improve reliability and reproducibility of results. Several papers have already been published that try to further this agenda, by suggesting recommendations for sampling and analysis, and minimum recommended reporting guidelines for eDNA studies (Goldberg et al. 2016, Nicholson et al. 2020), or highlighting the need to take the necessary controls in each step of the eDNA metabarcoding process (Zinger et al. 2019). In addition to the standardization of techniques for successful integration into quality assessments,

it is imperative that end-users become familiar with new technologies and their terminology, and are also shown the shortcomings of current methodology (Bush et al. 2019). The terminology that molecular tools bring with them may be unfamiliar for practitioners and policy makers, especially regarding NGS technology, which can hinder the successful implementation of these tools into water quality assessments. For instance, reporting "presence/absence" when using eDNA may be incorrect, and terms like "detected/not detected" would be better suited (Roussel et al. 2015), since false positives and negatives are still commonplace in eDNA analyses (Ficetola et al. 2016).

1.6 THE ECOLOGY OF EDNA

Besides terminology and the need for better reporting, there are several key challenges in the use of environmental DNA and DNA metabarcoding. For eDNA, some of the main challenges deal with the ecology of DNA within the environment, in particular the origin, state, transport, and fate (Figure 1.5), which have significant impact on the sampling design and DNA extraction methods. Where it concerns DNA metabarcoding, one of the key discussion topics surrounds the inference of abundance from molecular data; another deals with the use of MOTU data and genetic variation between taxonomic units, and both have considerable impact on the subsequent (ecological) analyses and assessments. Since the first papers emerged that demonstrated the potential of eDNA for the detection of aquatic organisms, there has been an increasing number of publications delving into the behavior of environmental DNA (see also Figure 1.2). Central questions in the "ecology of eDNA" (figure 1.5) deal with the state and fate of eDNA. Factors that play a role in this ecology of eDNA, such as transportation and degradation of eDNA, as well as spatial and temporal distribution patterns are key to the interpretation of eDNA results.

There are many potential sources of environmental DNA. Living organisms release DNA into the environment via shedding and extraction. The slimy coating that fish and amphibians use as a form of self-defense seems to be a decent source of eDNA, considering the many papers that use eDNA for the detection of these organism groups (e.g. Ficetola et al. 2008, Jerde et al. 2011). Other organisms that are known for shedding relatively high quantities of eDNA are bivalves, which not only continuously filter water, but also have periods of spawning that introduce large quantities of genetic material into the water column (Sansom & Sassoubre 2017). Aquatic insects that have larval stages in the water column introduce eDNA by molting and pupating, such as the larvae and pupae of Chironomidae (Bista et al. 2017). Feces are also a source of eDNA, from both prey and predator species, as are any dead

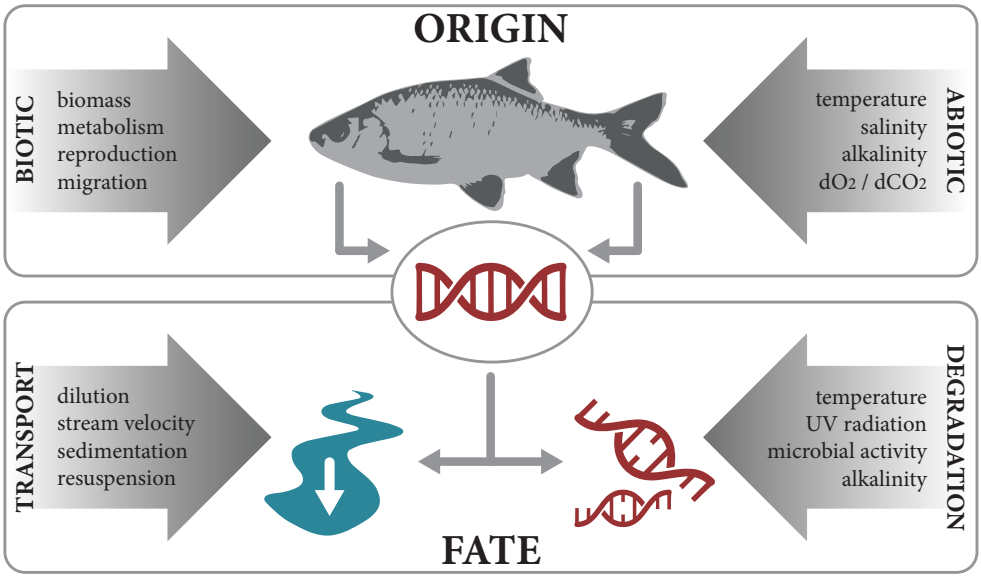


FIGURE 1.5. The main components in the cycle of eDNA in the water column: the origin of eDNA and the fate of eDNA through transportation and degradation, as well as some of the most important biotic and abiotic factors that influence the ecology of eDNA. Original illustration, based on Barnes & Turner (2015) and Stewart (2019).

organisms (Merkes et al. 2014). Shedding rates of living organisms vary throughout the seasons, and are influenced by different environmental factors, such as increases in water temperature (Jo et al. 2019). A study using bighead and silver carp also showed that feeding patterns of fish had an influence on the amount of eDNA shed into the water, probably due to increased excretion of sloughed gut cells with higher food intakes (Klymus et al. 2015). Other biotic factors, such as metabolism, age, stress, reproductive state, and migration, as well as a wide variety of abiotic factors like temperature, salinity, alkalinity, and levels of dissolved O_2 and CO_2 have been linked to variations in eDNA release and shedding from its source organism (Stewart 2019) (Figure 1.5). Environmental DNA sourced from dead animals can in some cases provide difficulties, for example in case of monitoring of invasive species from ballast water. Ballast water from transport ships is often treated to kill any remaining organisms, which are potential invaders. Environmental DNA might not be sufficient to separate signals from living invasive species and dead species dumped with ballast water, and some limited work has been done to circumvent these difficulties by using the more instable RNA as environmental signal (Pochon et al. 2017, Cristescu 2019).

The source of eDNA also influences its state. Whilst eDNA is often thought of as

DNA strands that exist freely in the water column, a large part of the environmental DNA is in fact still bound in cellular remains and adsorbed on particulate matter. Extracellular eDNA is relatively unstable, and exposed to the elements that break down DNA. Conflicting reports are found in the literature regarding the particle size distributions of eDNA. Fractioned filtration of water samples with filters with decreasing pore sizes revealed that fish eDNA was most abundant in fractions that would not pass through 1.0 μm pores (Turner et al. 2014a, Wilcox et al. 2015). These findings are consistent with hypotheses that a large proportion of eDNA is still bound in cellular remains or to particles, and also explains the heterogeneity observed in water samples. Turner et al. (2014a), however, also showed that total eDNA yields were highest using 0.2 μm pore size, as did a paper looking at particle size distribution in water fleas (Moushomi et al. 2019), which would suggest that eDNA is predominantly subcellular. Many studies, however, do not distinguish between intra- and extracellular DNA when assessing the processes within the ecology of eDNA, such as the factors playing a role in the persistence (Barnes et al. 2014). Often such studies use model organisms that are removed from an aquarium setup at a certain point in time to evaluate the persistence of DNA under various conditions, but use eDNA extraction methods that do not allow for distinctions between intra- and extracellular DNA (e.g. Dejean et al. 2011). This is not necessarily an issue, although it does highlight the continuing discussion around the use of the term "environmental DNA". Many papers use "eDNA" in a way that covers both intra- and extracellular DNA, but sometimes bulk-collected specimens are also treated as "environmental DNA". Fortunately, most studies around the ecology of eDNA use vertebrate or macroinvertebrate model organisms, in combination with collection methods that are not aimed at collecting bulk specimens. Generally, persistence of eDNA in the water column is relatively low, with signals becoming undetectable within days or weeks after removal of the source organisms (Dejean et al. 2011). Environments that have lower temperatures, higher pH, and are more protected from UV radiation have been found to allow a longer persistence time than water bodies with higher temperatures, lower pH, and more exposure to UV radiation (Strickler et al. 2015, Goldberg et al. 2018), all processes that either directly or indirectly (via increased microbial activity) influence the degradation of eDNA.

Transport of eDNA is another factor that plays an important role in the analysis of eDNA results, and something that should be considered during the sampling strategy design. Transport of eDNA is most obvious in lotic systems, in which eDNA has been shown to be transported from a point source with the flow of the water. DNA of two species living in a lake in Switzerland was observed up to almost 10

kilometers downstream in an outflowing river (Deiner & Altermatt 2014). Similar studies with different organisms found different detection ranges, from several kilometers for fish in a river in France (Civade et al. 2016), down to only five meters for amphibians in a stream in the United States (Pilliod et al. 2014). Several factors explain these differences, such as those involved in the persistence, but also density of organisms at the source, as well as stream velocity and turbulence, and their effect on sedimentation rates of particulate matter with adsorbed DNA. The influence of such factors were also postulated by Jane et al. (2015) from an experiment with caged trout in two headwater streams. In these experiments, they found that low stream velocities resulted in high eDNA concentrations near the cage, with concentrations quickly dropping further downstream, whereas high stream velocities resulted in low eDNA concentrations both near the cage and downstream. A study by Pont et al. (2018) combined observations from literature and their own field data into a model that showed that eDNA in lotic waters behaves much like fine particulate organic matter, and reported detection distances of up to a hundred kilometers. In a follow-up study in the catchment system in Switzerland, Deiner et al. (2016) concluded that eDNA was better able to provide information of a catchment area than kicknet sampling, with eDNA samples provided higher richness estimates in samples further downstream (and thus representing larger catchment area), as well as lower community dissimilarities compared to kicknet sampling.

The number of publications on the transport of eDNA in lentic systems is limited, although various studies have looked at the spatial distribution patterns of eDNA in freshwater bodies. The general conclusions from these studies is that the distribution of eDNA in lentic systems is quite heterogeneous, with very local signals of eDNA representing local presence of aquatic organisms. Optimal sampling requires spatial replicates to improve detection of organisms (Thomsen et al. 2012b, Harper et al. 2019b) or increase the number of taxa detected with NGS (Evans et al. 2017b, Grey et al. 2018, Lawson Handley et al. 2019). This suggests that the transport of eDNA in lakes and non-flowing ditches based on diffusion is limited. Similar findings of spatial dissimilarities have been reported for marine sampling, which is not surprising seeing how study sites there are larger than most freshwater systems (Guardiola et al. 2016, O'Donnell et al. 2017, Stat et al. 2019).

In addition to spatial distribution and transportation of eDNA, seasonal differences in eDNA have been studied in more detail the last few years. Some work has been done to highlight the need for spatial and temporal replicate sampling macroinvertebrate communities, although seasonal differences are mainly attributed to phenological patterns that have already been observed in morphological monitoring (Šporka et al.

2006), especially for those groups that are aquatic only in the larval stages (Bista et al. 2017). Most papers that study seasonal differences in aquatic eDNA, however, focus on fish (Stoeckle et al. 2017, Sigsgaard et al. 2017) or amphibians (Rees et al. 2017, Buxton et al. 2018). Environmental DNA seems to be more abundant in warmer seasons of the year, requiring fewer spatial replicates for successful detection of organisms (De Souza et al. 2016). Differences between winter and summer have been attributed to different factors, most of all the increased activity of many organisms, including reproductive activity (Figure 1.5). Other factors that are involved in seasonal differences of eDNA detectability are abiotic factors that influence persistence of eDNA, or physical processes like stratification and admixture in lakes (Lawson Handley et al. 2019).

In addition to the influence of the ecology of eDNA on DNA-based biodiversity monitoring, there are numerous practical and technical considerations, from the handling of samples in the lab to the interpretation of data during the analyses, which will be reflected upon in this thesis. The increasing number of studies published in the field of eDNA, metabarcoding (Figure 1.2), and their applications in biomonitoring, however, is a sign that the research field is making progress.

1.7 OUTLINE OF THIS THESIS

With this thesis I aim to shed some light on a few important considerations when dealing with molecular data, but also show the potential of these techniques. Since the DNA barcode reference libraries for Dutch aquatic macrofauna are relatively complete (Figure 1.3), using Dutch freshwaters as a study focus was an obvious choice. The Netherlands also has a long history of standardized WFD monitoring, which allowed us to make use of historic data and samples in the research presented in this thesis.

Chapter 2

One of the major obstacles in using data generated by DNA metabarcoding is the unreliable abundance data obtained from next-generation sequencing reads. Abundance is an important factor for WFD quality assessments under the current benchmarks. Developing methods that provide accurate abundance information from metabarcoding data is difficult, if not impossible, and problems are probably even greater when dealing with eDNA. So, if DNA metabarcoding or eDNA metabarcoding is ever to play a role in WFD monitoring and the calculations of EQRs, it would be necessary to make a transition to an EQR scoring system independent of abundance

data. In Chapter 2, titled "The influence of macroinvertebrate abundance on the assessment of freshwater quality in the Netherlands", I used historical monitoring data from several water authorities in the Netherlands (which were based on abundance data), and transformed these into presence/absence data. By directly comparing EQR scores calculated on both abundance and presence/absence data, I concluded that for macroinvertebrates, the importance of abundance data was only limited, and, perhaps surprisingly, removing abundances had little impact on the resulting EQR scores.

Chapter 3

The next step was to start comparing molecular tools for the identification of macroinvertebrates with traditional analyses of WFD samples. Current practices for WFD monitoring rely on the cumbersome and time-consuming visual identification of all the specimens collected at the monitoring sites. In addition to their time-consuming nature, morphological assessments are reliant on the expertise of individual assessors, and thus prone to error (see also paragraph 1.3). While DNA barcodes may not be able to distinguish all of the >1000 Dutch aquatic macroinvertebrate species, especially with the gaps in the reference libraries as they are, DNA-based methods have the benefit of not relying on individual taxonomic expertise.

In Chapter 3, titled "Increased performance of DNA metabarcoding of macroinvertebrates by taxonomic sorting", I tried to limit the influence of preferential amplification by using taxonomically sorted samples. Historical WFD samples were obtained, and kept the specimens separated according to the taxonomic categories used during the morphological analysis; Annelida, Crustacea, Heteroptera/Coleoptera, Mollusca, Trichoptera/Odonata/Ephemeroptera and "rest" (predominantly Chironomidae and other Diptera). Sorting specimens into these six basic groups before DNA extraction and amplification improved taxon recovery by 46.5%. When comparing the species lists obtained with DNA metabarcoding to those from the morphological assessment, there were considerable differences, although the numbers of taxa detected were similar. With an average overlap of 56.8% between the two, the impact on the EQR calculations was severe. However, for a few taxonomic groups the use of DNA barcodes resulted in much more detailed information, especially in those groups that are considered difficult to identify based on morphology.

Chapter 4

While using a blender and DNA metabarcoding WFD samples in bulk seems to be successful, when certain steps are taken to overcome the worst of the preferential

amplification, such methods would still require the specimens to be collected. While not as time-consuming as the identification process, the collecting of specimens still imposes considerable effort on the part of the monitoring agency. Sampling of eDNA would alleviate this. However, the approach is still under heavy scrutiny, especially where it concerns the sampling design. In Chapter 4, titled "The effects of spatial and temporal replicate sampling on eDNA metabarcoding", I investigated the effects of replication at three different levels during collecting and processing of eDNA samples: spatial replicate sampling, temporal replicate sampling, and PCR replicates. Two undisturbed natural lakes in the dunes of Wassenaar were sampled over 20 consecutive weeks, and eDNA was analyzed using the same general COI primer set used in the analysis of bulk samples. While our initial observation was that these general primers were not optimal for the detection of macroinvertebrates in eDNA samples due to the amplification of many non-target taxa, the replicate patterns were clear. Temporal differences over intervals larger than two weeks were larger than differences in spatial replicates, suggesting that turnover effects might be more important for the dynamics of eDNA than its heterogeneity within a study site. PCR replicates also showed dissimilarities, although not as profound as the replicate sampling.

Chapter 5

The effects of temporal replication were also witnessed in the study presented in Chapter 5 of this thesis, titled "Environmental DNA metabarcoding reveals comparable responses to agricultural stressors on different trophic levels of a freshwater community". In this study, I put eDNA to the test as a tool for monitoring impact of agricultural stressors on aquatic ecosystems. Because previous research had already shown that general macroinvertebrate primers for COI showed a lot of "bycatch", I opted to have an in-depth look into the ecosystem on three different trophic levels. Bacteria represented the decomposers, phytoplankton acted as a representative for the primary producers, and Chironomidae represented the macrofaunal community. Chironomidae are a key indicator group for water quality, and relatively well-represented in the custom reference databases created over the duration of the PhD project. All three groups are often understudied in water quality assessments, mostly due to the difficulties with morphological identifications.

To get a good grip on the individual and combined effects of two major agricultural stressors, the neonicotinoid insecticide thiacloprid, and fertilizer influx, this study was performed in the "Living Lab" facility of the University of Leiden. This unique setup strikes a balance between controlled laboratory mesocosms, and a natural field

situation. Replication of single and combined treatments, and undisturbed control situations, allowed us to disentangle the stressor effects, as well as differences caused by turnover through time. All three groups under assessment showed significant impact from both agricultural stressors, where even bacteria and phytoplankton communities were influenced by insecticides at concentrations regularly observed in surface waters in the Netherlands. Concurrently with the eDNA research, the Living Lab samples were also analyzed using traditional morphology, allowing us to directly compare results. The impact patterns seen with traditional assessments were comparable to those observed with eDNA, but at a much lower resolution (i.e., fewer taxa), and more replicates were needed to come to the same conclusions. The use of three understudied groups also allowed to us to uncover potential new bioindicators for freshwater stressors, although except for the Chironomidae, these usually were unidentified MOTUs. Large parts of the freshwater biodiversity, especially microorganisms, are still underrepresented in DNA reference databases.

Chapter 6

In the final chapter of this thesis, Chapter 6, I combine the insights into the applicability of DNA-based methods in freshwater monitoring obtained within my PhD project with results and conclusions from research in this field from the past decade. I discuss the potential and pitfalls of the use of environmental DNA for different assessments into the quality of freshwater, especially where it concerns technical considerations. Additionally, I give directions for future research to increase the understanding of eDNA, and how this knowledge should be integrated into methods that are suitable for direct application in the field.

CHAPTER 2

The influence of macroinvertebrate abundance on the assessment of freshwater quality in the Netherlands

Kevin K. Beentjes[1,2], Arjen G. C. L. Speksnijder[1], Menno Schilthuizen[1,2], Bartholomeus E. M. Schaub[3], Berry B. van der Hoorn[1]

Metabarcoding and Metagenomics 2:e26744 (2018)
https://doi.org/10.3897/mbmg.2.26744

[1] Naturalis Biodiversity Center, Leiden, The Netherlands
[2] Institute of Biology Leiden, Leiden University, Leiden, The Netherlands
[3] Hoogheemraadschap van Rijnland, Leiden, The Netherlands.

ABSTRACT

The use of molecular tools for the detection and identification of invertebrate species enables the development of more easily standardizable inventories of biological elements for water quality assessments, as it circumvents human-based bias and errors in species identifications. Current Ecological Quality Ratio (EQR) assessments methods, however, often rely on abundance data. Translating metabarcoding sequence data into biomass or specimen abundances has proven difficult, as PCR amplification bias due to primer mismatching often provides skewed proportions of read abundances. While some potential solutions have been proposed in previous research, we instead looked at the necessity of abundance data in EQR assessments. In this study, we used historical monitoring data from natural (lakes, rivers and streams) and artificial (ditches and canals) water bodies to assess the impact of species abundances on the EQR scores for macroinvertebrates in the Water Framework Directive (WFD) monitoring program of The Netherlands. By removing all the abundance data from the taxon observations, we simulated presence/absence-based monitoring, for which EQRs were calculated according to traditional methods. Our results showed a strong correlation between abundance-based and presence/absence-based EQRs. EQR scores were generally higher without abundances (75.8% of all samples), which resulted in 9.1% of samples being assigned to a higher quality class. The majority of the samples (89.7%) were assigned to the same quality class in both cases. These results are valuable for the incorporation of presence/absence metabarcoding data into water quality assessment methodology, potentially eliminating the need to translate metabarcoding data into biomass or absolute specimen counts for EQR assessments.

2.1 INTRODUCTION

Quality monitoring of freshwater ecosystems is prescribed under the European Union Water Framework Directive of 2000 (EU WFD; Directive 2000/60/EC) and focuses on monitoring of biological quality elements (BQEs) (European Union 2000). In Europe, benthic invertebrates are one of the most prevalently monitored BQE (Birk et al. 2012). Invertebrate communities are made up of species that represent a broad range of trophic levels, ecological functions and tolerances to stressors (Kenney et al. 2009). Traditional monitoring of freshwater macroinvertebrates, however, is labor-intensive and heavily dependent on expert knowledge of the assessors, making it slow, expensive and prone to human-induced bias and errors at all stages of collecting, sorting and identifying (Clarke & Hering 2006, Haase et al. 2010).

The incorporation of DNA barcodes (Hebert et al. 2003) into the identification process seems to have alleviated some of the human-induced issues. The use of these barcodes for identification of species has become more and more accepted, especially given the decline in traditional taxonomists (Hopkins & Freckleton 2002) and the ability of DNA barcodes to provide identifications of non-adult specimens and distinguish between cryptic clades (Sweeney et al. 2011, Jackson et al. 2014, Macher et al. 2016). Recent developments in DNA metabarcoding show high potential to assess biodiversity across many biomes (Taberlet et al. 2012a, Carew et al. 2013, Leray et al. 2013, Gibson et al. 2014, Stein et al. 2014, Pauls et al. 2014).

Now that the actual identification of species in bulk samples with high throughput sequencing (HTS) has shown its efficacy, the focus seems to shift towards solving some of the other issues concerning these novel strategies, especially the relationship between input biomass or specimen counts and output HTS sequence abundances (Amend et al. 2010, Deagle et al. 2013, Kelly et al. 2014, Elbrecht & Leese 2015, Piñol et al. 2015, Gibson et al. 2015, Hering et al. 2018, Aylagas et al. 2018)

The discussion, regarding the use of HTS read counts as an approximation of biomass or specimen abundances, is important for the biological components of the WFD as well. Abundance of (indicator) species or species groups is used in many European assessment metrics (albeit regularly as abundance classes) and is often part of multi-metric approaches (Birk et al. 2012, Hering et al. 2018, Pawlowski et al. 2018). While information on species abundances and evenness are generally considered important ecosystem properties, the often relatively simple WFD scoring systems may abide with a presence/absence-based methodology. Most traditional morphological monitoring relies on specimen count data, rather than biomass abundances, so even in situations where read abundances can be translated into

relative biomass, comparisons are difficult, considering also that most invertebrate taxa differ in biomass depending on their life stage. If presence/absence data can be as useful for WFD scoring as abundance data, it would allow for easier and faster incorporation of molecular techniques, especially now that efforts have been made to infer biotic indices from DNA data (Aylagas et al. 2014, Elbrecht et al. 2017a, Pawlowski et al. 2018).

In this study, therefore, we assessed the influence of species abundances on the Ecological Quality Ratio (EQR) scores for macroinvertebrates in the WFD monitoring program of The Netherlands. The Dutch system uses abundance data (in the form of abundance classes) for macroinvertebrates, where each species is scored as either a positive indicator, a negative indicator, a characteristic species or none of the aforementioned, depending on the type of water body (Evers et al. 2012, Van der Molen et al. 2016). A simple formula is used to calculate the ratio between normalized values for the indicators and expected reference values for the water type, which is expressed as a value between 0 and 1. Using historical records from traditional monitoring, we evaluated whether abundance data and presence/absence-based data produce comparable EQR scores.

2.2 MATERIALS AND METHODS

EQR scores for macrofauna were calculated on historical monitoring data from four Dutch water authorities (Hoogheemraadschap van Rijnland, Waterschap Aa en Maas, Waterschap Brabantse Delta and Waterschap Rivierenland), using morphological macroinvertebrate records from 2009 to 2017. These records are based on traditional macrofauna monitoring using kick-net sampling and morphological identification. The dataset included 877 monitoring locations spanning 23 different water types according to the Dutch classification system. Most locations were monitored more than once (some even annually), creating a total of 1780 macrofauna samples. An overview of the samples is provided in Table 2.1.

EQR macrofauna scores were calculated for all samples. The scoring system is based on the presence and/or abundance of positive indicator (DP), negative indicator (DN) and characteristic (KM) taxa. Most taxa are identified to species level in the Dutch macrofauna metrics, although for some "harder to identify" groups, species aggregates or higher-level taxonomic assignments are used (Evers et al. 2012, Van der Molen et al. 2016). In the most recent version of the Dutch WFD benchmarks, the absolute abundances of the dominant negative and the characteristics species used in the calculation are transformed into abundance classes (van der Hammen 1992). The

TABLE 2.1. Overview of samples. Distribution of samples used in this study, per water authority (includes survey time span), divided into the three categories defined by the EQR calculation: artificial ditches and canals, natural lentic (lakes) waters and natural lotic (rivers and streams) waters. No monitoring sites are present in rivers and streams for Hoogheemraadschap van Rijnland.

	Natural waters		Artificial waters	Total
	Lakes (type M12-M32)	**Rivers / streams** (type R04-R18)	**Ditches / canals** (type M01-M10)	
Hoogheemraadschap Rijnland (2009–2014)	198	n/a	173	371
Waterschap Aa en Maas (2011–2017)	9	221	150	380
Waterschap Brabantse Delta (2011–2016)	139	230	62	431
Waterschap Rivierenland (2011–2017)	8	56	534	598
Total	354	507	919	1780

EQR scores are calculated according to three different methods, based on the water type. Natural water bodies are divided into lentic and lotic. For lentic water bodies, such as lakes, the EQR is calculated according to the formula:

$$\mathbf{EQR_L} = (200*((KM\%)/KMmax)+(100-DN\%)+KMDP\%)/400$$

Where KM% is the percentage of characteristic taxa, KMmax is a constant value representing the expected fraction of characteristic taxa depending on the specific water type, DN% is the percentage of negative indicator individuals and KMDP% is the percentage of characteristic and positive indicator individuals (Van der Molen et al. 2016). Lotic water bodies, such as streams and rivers, are calculated slightly differently, with more emphasis on the negative indicators:

$$\mathbf{EQR_R} = (200*((KM\%)/KMmax)+(2*(100-DN\%))+KMDP\%)/400$$

For artificial water bodies, such as ditches and canals, the calculation is performed according to the following formula:

$$\mathbf{EQR_A} = ((2*PT/PTmax)+(1-(DN\%/DNmax)))/3$$

Where PT is the absolute number of positive indicator taxa, PTmax is the absolute number expected positive indicator taxa, DN% is the percentage of negative indicator individuals and DNmax is a constant value depending on the specific water type (Evers et al. 2012). The resulting score of all formulae is a value between 0 and 1, which is subdivided into five quality classes: "bad" (EQR <0.2), "poor" (0.2–0.4), "moderate" (0.4–0.6), "good" (0.6–0.8) and "high" (0.8–1.0). These scores also reflect how observed conditions compare to reference status (and thus target status) for the assessed water type, where the highest status shows no difference and the lowest status shows substantial differences (Birk et al. 2012). For artificial water bodies, there are only four quality classes, with "good" representing scores between 0.6 and 1.0, as artificial waters have no natural reference status for comparison.

For each of the 1780 samples, EQR scores were calculated using both original data with abundance classes and a manipulated dataset, converted to a presence/absence monitoring scheme by setting all specimen counts to 1. Any duplicate taxa in a given sample (e.g. where both adult and juvenile specimens were recorded separately) were removed to avoid aggregation into abundance classes other than 1 (abundance class 1 indicates a single specimen was found). QBWat software version 5.33 (Pot 2015) was used to compare and score the original and manipulated monitoring lists with predefined positive and negative indicator species lists, as well as the characteristic taxa list and the EQR based on the relevant formula for each water type was calculated. These predefined species lists (positive indicators, negative indicators and characteristic taxa) have been created specifically for EQR assessments and are based on species characteristics described in literature and expert judgements (Van der Molen et al. 2016). EQR scores with and without abundances were compared to determine the effect on the score, as well as the effect on the classification into the five quality classes. Dunn's test was used to investigate the difference between water types and between quality classes.

2.3 RESULTS

The investigated macrofauna samples had an average of 72.1 ± 0.8 (mean ± SEM) species (minimum 1, maximum 217) recorded, with an average of 1221.5 ± 25.8 specimens (minimum 1, maximum 11767). Mean EQRs, calculated with presence/absence-based data, were highly correlated to original EQRs based on abundance class data, for natural lentic sites (Pearson correlation $r = 0.993$, $p < 0.001$) (Figure 2.1A), natural lotic sites (Pearson correlation $r = 0.982$, $p < 0.001$) (Figure 2.1B) and artificial sites (Pearson correlation $r = 0.983$, $p < 0.001$) (Figure 2.1C). Neither of the EQRs,

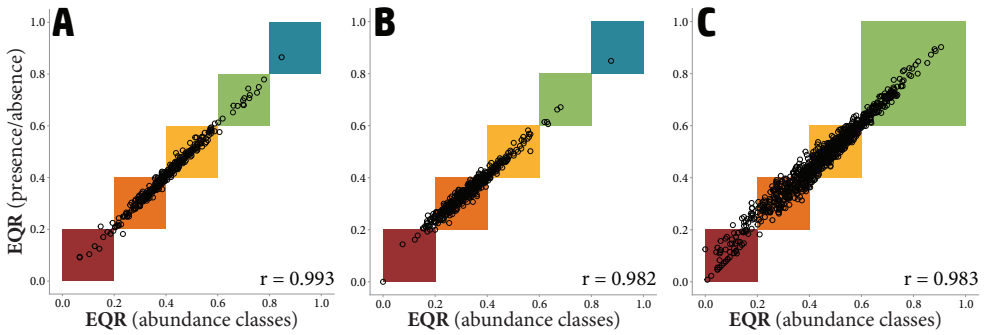


FIGURE 2.1. EQR (presence/absence) versus EQR (abundance classes). Comparison of macroinvertebrate EQR scores in standard assessment using abundance classes and EQR scores in simulated scenarios with presence/absence data for (A) natural lentic waters (lakes, n=354), (B) natural lotic waters (streams and rivers, n=507) and (C) artificial waters (ditches and canals, n=919). Coloured boxes indicate EQR quality classes: "bad" (red), "poor" (orange), "moderate" (yellow), "good" (green) and "high" (blue). For artificial water bodies, there are only four quality classes, with "good" representing scores between 0.6 and 1.0. For all comparisons, the EQR scores of abundance class data and presence/absence data was significantly correlated (Pearson correlation, $p < 0.001$). Pearson correlation values are provided in the panels.

nor the difference between the two scores (ΔEQR) followed a normal distribution. Mean EQR without abundance data was 0.424 ± 0.003, which was significantly higher than the mean EQR calculated with abundance classes (0.404 ± 0.003) (Wilcoxon signed-rank test, $p < 0.001$). The majority of EQRs were higher without abundances (1349 samples, 75.8%), 359 samples scored lower (20.2%) and only 72 out of 1780 samples (4.0%) scored exactly the same (based on scores with three decimal digits). The removal of abundance classes had significantly less impact on the scoring for natural lentic systems (mean ΔEQR 0.006 ± 0.001) than it had on both natural lotic systems (0.021 ± 0.001) and artificial water bodies (0.024 ± 0.001) (Dunn's test, $p < 0.001$). There was no significant difference between the lotic and artificial systems (Figure 2.2A). Removal of abundance data had a stronger effect on samples from the lowest quality class ("bad"), where the mean ΔEQR was significantly higher than all other quality classes (Dunn's test, $p < 0.001$). Mean ΔEQR of the "poor" class, in turn, was significantly higher than those of "moderate" and "good" (Dunn's test, $p < 0.001$), while there was no significant difference in the impact on "moderate" and "good". The "high" class was excluded from this analysis with only two of 1780 samples being assigned to that category (Figure 2.2B).

When assigning quality classes to the EQRs based on presence/absence data, 1596 (89.7%) of all samples were assigned to the same class, 22 (1.2%) were scored lower and 162 (9.1%) were scored higher. The change was most profound in the samples

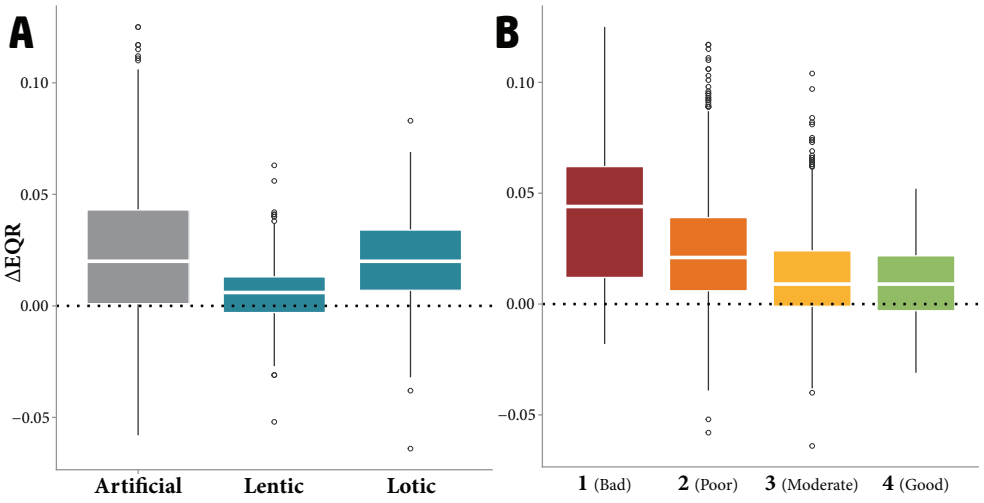


FIGURE 2.2. Factors influencing ΔEQR. Comparison of differences in EQR between assessment using abundance classes and using presence/absence data, (A) split by water type and EQR calculation method and (B) split per original assessment quality class ("high" was omitted, with only two samples in this data set). On average, classifications without abundance are higher than original assessments (ΔEQR positive). Removal of abundance resulted in significantly lower differences in natural lentic waters compared to natural lotic and artificial waters (Dunn's test, $p < 0.001$). There was no significant difference between lotic and artificial. Removal of abundance data has significantly more impact on EQR assessments for samples originally classified as "bad" compared to all other classes (Dunn's test, $p < 0.001$). ΔEQR was also significantly higher in "poor" samples compared to "moderate" and "good".

originally assigned to "bad", where 51 out of 117 (43.6%) were assigned to "poor", the class above. Results were comparable for the different water types: 95.2% of natural lentic samples, 89.9% of natural lotic samples and 87.4% of artificial samples were assigned to the same class. Samples assigned to a different quality class were assigned to a class either directly below or directly above its previous classification.

2.4 DISCUSSION

Our results show there is a strong correlation between traditional EQR based on freshwater macrofauna using abundance data and EQRs calculated without abundance data in the Dutch system. For most samples, scores were comparable between the abundance- and presence/absence-based methods, with the majority (89.7%) being assigned to the same quality class in both cases. The difference seems to be largest in samples at the lower end of the EQR score spectrum, with almost half (43.6%) ending up in a higher quality class ("poor" instead of "bad").

Based on the formulae used for the calculation of the EQRs, it can already be deduced that abundance is not a consideration for all components that determine the final score. For natural lakes, half the score is represented by the fraction of characteristic taxa, which does not take individual specimen counts into account. The fraction of the score defined by factors that use abundance data is slightly lower for natural streams and rivers (two fifths) and for artificial ditches and canals abundances are not used for two thirds of the final score (Van der Molen et al. 2016). Interestingly, in our analysis, the impact of removal of abundances was significantly smaller in lakes than it was in rivers and streams and artificial water bodies (Figure 2.2A). A larger impact on rivers and streams was expected, as 60% of the final EQR is based on the abundance of individuals scoring on each of the three indicator lists (positive, negative and characteristic). However, in the artificial systems, this only amounts to one third of the final score, so one would expect the impact to be smaller, especially considering that the quality classes most impacted by the removal of abundance ("bad" and "poor", Figure 2.2B) only account for 35.8% of the artificial water samples in the data presented in this paper, whereas those classes account for 49.4% and 81.1% of lakes and streams, respectively.

The parts of the EQR score that do rely on abundance data in the Dutch system use abundance classes rather than actual specimen abundances. This may be a major factor in why the removal of abundances has only a limited impact on the EQR scores. Abundance classes were introduced into the Dutch metrics to reduce the effect of extremely high abundances of a single species on the EQR. The abundance class system uses a total of nine classes, where class "1" represents a single specimen and class "9" represents abundances over 1808 specimens. When applied to the monitoring data, this means that an abundance of 20 specimens is translated to class "4", whereas an abundance of 200 specimens is assigned to class "6". Thus, whilst the actual abundance difference might be tenfold, in the calculation it would be only 1.5-fold, already reducing the effect of absolute abundances on the final EQR (van der Hammen 1992, Evers et al. 2012).

These observations are important when considering the incorporation of molecular techniques into WFD quality monitoring methodology. Given that techniques, such as metabarcoding, are proving their efficacy in the process of identification of species in bulk samples, incorporation of such techniques into the actual monitoring is only a matter of time (Zimmermann et al. 2015, Elbrecht et al. 2016, Pawlowski et al. 2018). Efforts have been made in trying to link amplification bias in HTS with amplification success and PCR efficiency of quantitative PCR (qPCR) methods, showing there may be a relationship between low read numbers in HTS and high Cq values in qPCR,

although PCR efficiency itself seemed unrelated (Pawluczyk et al. 2015). Even in case such an approach would yield usable information, it would not only require a priori knowledge of species present within a sample, but also seems cumbersome in complex monitoring samples, such as the ones used for this study (with an average of 72 species).

While our results imply that the technically difficult DNA-based quantifications might be avoided when calculating EQR scores, being able to measure species-abundance relationships from DNA data would nonetheless be desirable, since such relationships play an important role in understanding community composition and dynamics (Hubbell 2001). However, even for the relatively straightforward EQR scoring, the findings in this study cannot be translated into a conclusion that any bias can simply be ignored. These biases are an important consideration when generating taxon lists using HTS on bulk sample metabarcoding. Uneven distributions, paired with preferential amplification of certain taxonomic groups, will result in incomplete recovery of taxa from a sample. It is therefore still important to take the necessary steps to avoid primer bias as much as possible.

One of the main advantages of DNA-based identifications over traditional taxonomy is the ability to reliably identify larval stages and complicated taxonomical groups, for example in cryptic species, showing contrasting reactions to stressors (Macher et al. 2016, Beermann et al. 2018). The use of metabarcoding to replace morphological taxonomic assignment will bring changes to the species lists that can be used for EQR or other quality assessments because there will be more information on those groups that are currently underused due to identification difficulties, as well as potentially higher resolutions of the identifications. Such changes alone may already prove challenging to use in traditional EQR assessments, as these systems have been set up with known limitations in mind. The Dutch macrofauna metric, for instance, makes little to no distinction between genera and species in the family Tubificidae, and many Chironomidae genera have the same scoring for each of its species (Van der Molen et al. 2016). Any such changes alone would warrant a new system, rather than recalibration of the currently used methodologies that are partly built around the limitations of morphological identifications. However, until the lack of knowledge about species-level responses to stressors has been resolved, higher-resolution taxa lists can be merged into less resolved levels to allow for compatibility with current assessment systems.

Taking these considerations into account, together with the fact that expanding the DNA barcode repository for freshwater macroinvertebrates is one of the main focal points of the European DNAqua-Net collaboration (Leese et al. 2016), the

generation of reliable species lists based on molecular data rather than morphological assessments is no longer a vision for the future. EQRs have always been used as a way to quickly assess the ecological status of water bodies. Thorough knowledge of the potential caveats in molecular identification and/or detection techniques will allow for new EQR methods to be developed, methods better suited for use with molecular data. The transition towards DNA-based EQRs certainly has the potential to induce supranational standardization within water quality assessment. Especially with international collaborations such as those in DNAqua-Net, which states that its goal is to "develop a roadmap to include [DNA-based tools] in the standardized ecological assessment of aquatic ecosystems in Europe and beyond" (Leese et al. 2016). Any such molecular-based EQRs might benefit from using more easily generated presence/absence taxon lists instead of an abundance-based analysis, allowing for faster and more and easily standardizable water quality assessments.

2.5 CONCLUSIONS

We demonstrated the viability of adopting presence/absence data instead of specimen abundance data in a WFD water quality assessment program. Given all obstacles hampering the translation of HTS read data into biomass or absolute specimen counts, this paves the way for incorporating metabarcoding workflows into future assessment methodology. While species abundances are still valuable for a thorough ecological understanding of natural systems, the EQRs have been used more as a relatively quick assessment of ecological status of water bodies compared to reference situations. The EQR methodology used in this paper applies to the quality monitoring in The Netherlands and results may vary for other nations, based on the methods of EQR calculation. We urge researchers to look into the actual influence of abundance data on their WFD programs and in studies using metabarcoding data. With molecular techniques, such as metabarcoding of environmental DNA or bulk samples, proving to be successful, it is imperative that developments in routine EQR assessments, be they recalibrations or entirely new systems, strive to be more compatible with the potential lack of abundance data.

2.6 ACKNOWLEDGEMENTS

This study was part of the DNA Waterscan project funded by the Gieskes-Strijbis Fonds. Preliminary simulations of abundance data and EQR-scores were performed by Royal HaskoningDHV. We thank the Hoogheemraadschap van Rijnland and Aquon for providing data and instructions for the QBWat software programme and Waterschap Aa en Maas, Waterschap Brabantse Delta and Waterschap Rivierenland for additional monitoring data.

2.7 SUPPLEMENTARY MATERIALS

SUPPLEMENTARY FILE S2.1. Monitoring event details and EQR scores.
https://doi.org/10.3897/mbmg.2.26744.suppl1

CHAPTER 3

Increased performance of DNA metabarcoding of macroinvertebrates by taxonomic sorting

Kevin K. Beentjes[1,2], Arjen G. C. L. Speksnijder[1], Menno Schilthuizen[1,2], Marten Hoogeveen[1], Rob Pastoor [1], Berry B. van der Hoorn[1]

PLoS ONE 14(12):e0226527 (2019)
https://doi.org/10.1371/journal.pone.0226527

[1] Naturalis Biodiversity Center, Leiden, The Netherlands
[2] Institute of Biology Leiden, Leiden University, Leiden, The Netherlands

ABSTRACT

DNA-based identification through the use of metabarcoding has been proposed as the next step in the monitoring of biological communities, such as those assessed under the Water Framework Directive (WFD). Advances have been made in the field of metabarcoding, but challenges remain when using complex samples. Uneven biomass distributions, preferential amplification and reference database deficiencies can all lead to discrepancies between morphological and DNA-based taxa lists. The effects of different taxonomic groups on these issues remain understudied. By metabarcoding WFD monitoring samples, we analyzed six different taxonomic groups of freshwater organisms, both separately and combined. Identifications based on metabarcoding data were compared directly to morphological assessments performed under the WFD. The diversity of taxa for both morphological and DNA-based assessments was similar, although large differences were observed in some samples. The overlap between the two taxon lists was 56.8% on average across all taxa, and was highest for Crustacea, Heteroptera, and Coleoptera, and lowest for Annelida and Mollusca. Taxonomic sorting in six basic groups before DNA extraction and amplification improved taxon recovery by 46.5%. The impact on ecological quality ratio (EQR) scoring was considerable when replacing morphology with DNA-based identifications, but there was a high correlation when only replacing a single taxonomic group with molecular data. Different taxonomic groups provide their own challenges and benefits. Some groups might benefit from a more consistent and robust method of identification. Others present difficulties in molecular processing, due to uneven biomass distributions, large genetic diversity or shortcomings of the reference database. Sorting samples into basic taxonomic groups that require little taxonomic knowledge greatly improves the recovery of taxa with metabarcoding. Current standards for EQR monitoring may not be easily replaced completely with molecular strategies, but the effectiveness of molecular methods opens up the way for a paradigm shift in biomonitoring.

3.1 INTRODUCTION

Now that the use of DNA barcoding for the identification of species (Hebert et al. 2003) has proven its merit, research is shifting towards the integration of molecular identifications in ecological and biodiversity assessments across different biomes (Taberlet et al. 2012a, Leray et al. 2013, Pauls et al. 2014, Pawlowski et al. 2018). Integration of molecular techniques can provide a significant added value for the monitoring of biological quality elements (BQEs) in fields such as the quality monitoring of freshwater under the European Framework Directive (WFD) (European Union 2000). To date, many of the BQEs analyzed for WFD monitoring are still assessed using traditional morphology-based methods (Birk et al. 2012). These traditional methods, however, are known to be hampered by difficulties in identification and substantial differences between assessors (Haase et al. 2006, Stribling et al. 2008, Sweeney et al. 2011), and can be expensive due to their time-consuming nature (Marshall et al. 2006, Darling & Mahon 2011, Stein et al. 2014).

Recent advances have shown the efficacy of DNA metabarcoding to assess macroinvertebrate samples (Gibson et al. 2014, Pawlowski et al. 2018) and to obtain metrics for bioassessments (Aylagas et al. 2016, Elbrecht et al. 2017a, Aylagas et al. 2018). Although DNA-based methods are generally perceived as an improvement over the traditional morphological assessments (Bush et al. 2019), challenges remain to be solved before DNA-based methods can be fully incorporated into routine bio-monitoring. Studies employing metabarcoding of aquatic macroinvertebrates are often limited to single samples (Hajibabaei et al. 2011), a select subset of taxa (Carew et al. 2013) or rely on mock communities (Bista et al. 2017, Elbrecht & Leese 2017, Elbrecht et al. 2017b, Lobo et al. 2017). Research that does cover a broader variety of WFD monitoring samples often deals with differences in taxonomic resolution between morphological and DNA analyses (Gibson et al. 2015, Elbrecht et al. 2017a). One of the main confounding effects in the use of molecular approaches is the effect of primer bias and preferential amplification in complex samples, leading to taxonomic bias (Pawluczyk et al. 2015, Creedy et al. 2019). Interactions between taxa from various organism groups, of varying sizes and in varying biomass ratios remain understudied, and implications can be severe, limiting the possibility to relate metabarcoding read data to actual taxon abundances (Elbrecht & Leese 2015), even though these actual abundances might not be as important for simple ecological quality ratio calculations used by water monitoring agencies (Chapter 2).

In this paper, we assess the implementation of DNA metabarcoding for species identification in bulk samples collected under the WFD. We evaluate the performance

of DNA metabarcoding-based identification of taxa across six different taxonomic groups that collectively cover most of the traditional macroinvertebrate samples collected for WFD freshwater quality assessments: Annelida, Crustacea, Heteroptera/Coleoptera, Mollusca, Trichoptera/Odonata/Ephemeroptera, and Diptera. Our aim is to assess the effects of taxonomic sorting on the recovery of taxa from bulk metabarcoding, and the impact of replacing these groups with molecular data on ecological quality ratio (EQR) scoring. While EQRs are a simplified way to look at community compositions, they provide an insight in water quality, and are widely used by water monitoring agencies to assess the status of surface waters under the WFD (Birk et al. 2012, Chapter 2). We also discuss some concerns on DNA reference databases that may hinder successful application of molecular methodology in biomonitoring.

3.2 MATERIALS AND METHODS

3.2.1 Sample selection and processing

Freshwater macroinvertebrate samples were collected in the Hoogheemraadschap Rijnland monitoring district in 2010 and 2012 by ecological survey company Aquon (Leiden, the Netherlands). Samples were collected and analyzed according to standardized WFD monitoring guidelines (STOWA 2014). Specimens were sorted by Aquon taxonomists into seven different categories during morphological analysis, and stored separately in ethanol per taxon group: ANNE (Annelida), ACA (Hydrachnidia, stored in Koenike's fluid), CRUS (Crustacea), HECO (Heteroptera and Coleoptera), MOLL (Mollusca), TOE (Trichoptera, Odonata and Ephemeroptera), and REST (miscellaneous, predominantly Chironomidae and other Diptera). Specimens were identified to lowest possible level, preferably species level. For this study, we selected 25 samples out of 138 from the monitoring cycles of 2010 and 2012. More recent samples could not be used, as there is a five-year retention period for WFD monitoring samples. Samples were selected based on the WFD ecological quality ratio (EQR) scores (range 0.158–0.759), as well as the Shannon-index (range 0.840–4.326), to represent a broad range of sample diversities and complexities (for all 138 samples, EQR ranged from 0.059 to 0.847 and Shannon-index ranged from 0.602 to 4.326). EQR scores in the Dutch WFD monitoring range from 0.0 to 1.0, and are divided into 5 categories ranging from "bad" (EQR 0.0–0.2) to "high"(EQR 0.8–1.0) (for more detail, see Chapter 2). The 25 selected samples represented four out of five quality classes, in the 138 samples there was only one sample that was scored as "high". The full taxon lists with specimen counts have been included in the

supplementary data (Supplementary File S3.1).

Not all of the seven groups were present in all samples. The water mites (ACA) were excluded from the analysis, as they were preserved in Koenike's fluid (45% water, 45%, glycerin, 10% glacial acid acetic), which had a negative impact on the preservation of DNA and we were unable to obtain useable DNA extracts from the samples. To account for the missing taxa, water mites were also removed from the morphological lists during the comparison of DNA and morphology.

3.2.2 DNA extraction and amplification

Specimens were homogenized in 15 ml sterile tubes containing 10 steel beads (5 mm diameter), using the IKA Ultra Turrax Tube Drive (IKA, Staufen, Germany) in a fixed volume of 5.0 ml 96% ethanol. Each tube was ground three times for one minute on the maximum speed setting (6000 rpm). A tube with only 5.0 ml of 96% ethanol was used as an extraction blank. After homogenization, 500 µl of the ethanol with ground specimens was transferred to a 2 ml tube, and the ethanol was evaporated using a Concentrator plus vacuum centrifuge (Eppendorf, Nijmegen, the Netherlands). DNA was extracted from the remaining dry debris using the Nucleomag 96 Tissue kit (Macherey-Nagel, Düren, Germany) on the Kingfisher Flex Purification System (Thermo Fisher, Waltham, MA, US), with a final elution in 150 µl. To simulate a total DNA extraction on all taxa of one sampling location combined, 5.0 µl of DNA extract from each of the taxonomically sorted samples belonging to one location was combined into a pool, which was amplified and sequences in the same way as the sorted samples.

A two-step PCR protocol was used to create a dual index amplicon library, using primers BF1 and BR2 (Elbrecht & Leese 2017) to amplify a 316 base pair fragment of the COI barcoding region. These primers have been shown to successfully amplify a wide range of freshwater macroinvertebrates. All 183 samples (158 individually extracted tubes and 25 pools) were amplified and labeled separately, using two PCR replicates for each sample. First round PCRs were performed in 20 µl reactions containing 1x Phire Green Reaction Buffer, 10 µg BSA (Promega, Madison, WI, US), 0.5 mM dNTPs, 0.4 µl Phire Hot Start II DNA Polymerase (Thermo Fisher, Waltham, MA, US), 0.65 µM of each primer and 2.0 µl of template DNA. Initial denaturation was performed at 98°C for 30 seconds, followed by 30 cycles at 98°C for 5 seconds, 50°C for 5 seconds and 72°C for 15 seconds, followed by final elongation at 72°C for 5 minutes. PCR success was checked on an E-Gel 96 pre-cast agarose gel (Thermo Fisher, Waltham, MA, USA). PCR products were then cleaned with a one-sided size selection using NucleoMag NGS-Beads (Macherey-Nagel, Düren, Germany), at a 1:0.9 ratio.

Second round PCRs to add the individual P5 and P7 Illumina labels (Nextera XT Index Kit; Illumina, San Diego, CA, USA) were performed using 3.0 µl of cleaned PCR product from the first round in a 20 µl reaction containing 1x TaqMan Environmental Master Mix 2.0 (Thermo Fisher, Waltham, MA, USA) and 0.5 µM of each primer. Initial denaturation was performed at 95°C for 10 minutes, followed by 14 cycles at 95°C for 30 seconds, 55°C for 60 seconds and 72°C for 30 seconds, followed by final elongation at 72°C for 7 minutes. All PCRs were performed in 96-well plates, with replicates in separate plates. Each plate contained two wells with an artificial internal control (AIC) sample that was used to gauge the amount of cross-contamination between samples in the amplification process in the laboratory. The artificial control was based on the COI barcode region of a Reeve's muntjac (*Muntiacus reevesi*) with several primer sets built into the sequence, and synthesized by IDT (Leuven, Belgium) (Supplementary Figure S3.1). Second round PCR products were quantified on the QIAxcel (Qiagen, Venlo, the Netherlands) and pooled equimolarly per PCR plate using the QIAgility (Qiagen, Venlo, the Netherlands). Pools were cleaned with a one-sided size selection using NucleoMag NGS-Beads (ratio 1:0.9) then quantified on the Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) with the DNA High Sensitivity Kit. The pools were then combined equimolarly into one sample and sequenced in one run of Illumina MiSeq (v3 Kit, 2x300 paired-end) at Baseclear (Leiden, the Netherlands). Sequence data is available from the NCBI Sequence Read Archive (Bioproject accession PRJNA550542).

3.2.3 Bioinformatics

Quality filtering and clustering of the entire dataset was performed in a custom pipeline on the OpenStack environment of Naturalis Biodiversity Center through a Galaxy instance (Afgan et al. 2018). Raw sequences were merged using FLASH v1.2.11 (Magoč & Salzberg 2011) (minimum overlap 50, mismatch ratio 0.2); non-merged reads were discarded. Primers were trimmed from both ends of the merged reads using Cutadapt v1.16 (Martin 2011) (minimum match 10, mismatch ratio 0.2). Any read without both primers present and anchored was discarded. PRINSEQ v0.20.4 (Schmieder & Edwards 2011) was used to remove reads with length below 313 bp and above 319 bp, to allow for natural variations in coding sequence as well as potential primer slippage (Elbrecht et al. 2018b). Sequences were dereplicated and clustered into Molecular Operational Taxonomic Units (MOTUs) using VSEARCH v2.10.3 (Rognes et al. 2016) with a cluster identity of 98% and a minimal accepted abundance of 2 before clustering. The presence of AIC reads in the regular (non-control) samples, as well as the presence of non-AIC DNA in the control samples was

used to determine the MOTU filtering threshold; only MOTUs with read abundances above 0.025% were retained for each replicate. Samples with fewer than 4,000 reads were discarded and PCR replicates were combined according to the additive strategy, counting all MOTUs, irrespective of how many replicates they occurred in (Alberdi et al. 2018), as the intent was to recover as many taxa as possible.

MOTU sequences were compared to a custom reference database using an extended BLAST+ script (https://github.com/naturalis/galaxy-tool-BLAST). The custom reference dataset included 2,757 COI barcodes obtained from WFD species collected in the Netherlands as part of the national DNA barcoding campaign (Beentjes et al. 2015), supplemented with sequences obtained from BOLD (Ratnasingham & Hebert 2007) belonging to the 795 genera listed on the Dutch WFD species list. A total of 350,449 public sequences of 679 genera were retrieved from BOLD using the package BOLD (Chamberlain 2017) in R (RStudio 2015) (sequences downloaded 28 June 2018). The remaining genera were either not present in the BOLD database (107 genera) or had no public sequences linked to them (9 genera). The exclusion of sequences not identified to at least genus level allowed for linking taxa to the Dutch Species Register (https://www.nederlandsesoorten.nl/) based on genus names, making all taxonomic data compatible for use in lowest common ancestor analysis. The final database was dereplicated, removing all entries that had 100% identical DNA sequences and species names. MOTUs were also compared to a second custom reference library containing COI sequences and bacterial genomes downloaded from NCBI GenBank (Benson et al. 2005) (sequences downloaded 21 August 2018), to help filter out non-macroinvertebrate MOTUs and correct for misidentifications based on contaminated (e.g. *Homo sapiens* or *Wolbachia*) or otherwise erroneous sequences in the BOLD database.

The top 100 hits were obtained for both BLAST comparisons. Anticipating gaps in the DNA database, we developed a custom lowest common ancestor (LCA) tool to be able to assign higher-level taxonomic assignments for MOTUs without direct hits (>98% match and 100% coverage) in the reference database. The LCA tool was based on MEGAN (Huson et al. 2007), with adaptations to allow for the use of custom taxonomic databases and integration into the Galaxy infrastructure (https://github.com/naturalis/galaxy-tool-lca). The LCA script was performed on the top 5% hits, with bit-score >170, a minimum identity of 80% and a minimum coverage of 80%. The LCA tool was set to identify MOTUs no further than genus level. All direct hits (>98% match) were retrieved directly and accumulated based on taxon name associated with the sequences. To check for non-Dutch taxa and synonyms, a custom taxon matcher tool (https://github.com/naturalis/galaxy-tool-

taxonmatcher) was used to compare all the names obtained to taxa recorded in the Dutch Species Register. In case of multiple taxa having a direct hit, the names were manually checked and taxonomy was determined based on the following set of rules: (1) non-Dutch species were removed, (2) synonyms were resolved, (3) sub-species level identifications were set to species level, (4) when a MOTU matched both genus level sequences and species level reference sequences of the same genus, species level identifications were retained, (5) putative misidentifications or contaminations were removed, based on expert judgment and the top 100 BLAST hits, (6) if one species matched consistently higher than another, the species with a better match was retained, (7) in case of equal matches with multiple species, all species names were retained (e.g. species complexes that could not be resolved with the available reference sequences).

3.2.4 Comparison morphology versus molecular identification

After applying the LCA script, MOTUs with the same taxonomic assignment were aggregated. Individual samples were then accumulated into their respective locations, with exception of the pool sample. Taxa lists obtained from the molecular analysis were compared to the WFD taxa lists based on conventional morphological identifications provided by Aquon. Morphological taxa lists were first matched to the Dutch Species Register using the same script that was used to compare the taxa lists retrieved from metabarcoding, to make the species names in both lists compatible. Before the comparison, redundancy was removed from both taxa lists, to exclude uncertainties in identifications or potential duplicates (i.e., a genus level identification was omitted if the list also contained specimens from that genus that were identified to species level).

DNA-based taxon lists from the pools and the separately sequenced samples added together were both compared to the morphological list manually. Each entry on the combined lists was classed into one of the following categories: (1) "found", where there was an exact match between both lists; (2) "identified at a different level", when there was a match, but either one of the lists had a higher-level identification; (3), "putative misidentification", in cases where two different species from the same genus were listed on the respective lists; (4) "missing in reference" when the morphologically identified species was not covered by the DNA reference database; (5) "not found", when the taxon was covered in the reference database, but only encountered in the morphological list; (6) "extra", when the taxon was only encountered in the DNA list. To calculate the overlap between morphology and DNA, the first three categories were grouped together as being found in both lists, the taxa missing from the reference

were counted towards the taxa only found in the morphology.

To analyze if uneven sequencing depth between samples pooled prior to amplification and the separately sequenced samples added together had any effect on taxonomic recovery, and to allow for better comparison between samples, all data was rarefied to the lowest read count available. Pooled samples were all rarefied to 15,000 reads, separately sequence samples representing the different taxon groups were each rarefied to 2,500 reads to adjust for the fact that most pools consisted of six taxon groups.

Ecological quality ratio (EQR) scores were calculated according to the Dutch standards for both morphological and DNA-based taxon lists, using the QBWat software version 5.33 (Pot 2015) (with redundancies removed as described previously). Scores were calculated based on presence/absence data (with all specimen counts set to one) for both morphological and molecular data. Previous research has shown that abundances had limited impact on the EQR score (Chapter 2).

3.3 RESULTS

3.3.1 Sequence run statistics

Sequencing resulted in a total of 9,998,809 read pairs. After merging and quality filtering, 9,081,986 sequences were retained for MOTU clustering. AIC reads were detected in several non-control samples. A 0.025% threshold for filtering low-abundance MOTUs from each sample removed control reads from all samples. After filtering the MOTU table 2,460 MOTUs were retained in the non-control samples, representing 8,200,488 reads. Out of 366 replicates (158 sorted samples, 25 pools, all in duplicate), 77 with fewer than 4,000 reads were discarded. On average, PCR replicates had 28,345 reads (range 4,197–69,919), and 43.0 MOTUs (range 2–132). There was no correlation between number of reads and number of MOTUs in each sample.

3.3.2 Taxonomic composition

Using the two reference libraries, 1,837 MOTUs were identified as macrofauna taxa listed on the Dutch WFD taxon list on at least order level. A total of 319 MOTUs had direct matches above 98% percent, representing 213 distinct species or species complexes. The remaining MOTUs were identified to genus (1,394 MOTUs, 121 genera), family (93 MOTUs, 12 families) or order level (31 MOTUs, 11 orders). MOTUs that were not identified to at least order level were discarded. The final dataset of the sorted and separately amplified groups represented 208 species, 159 genera, 75 families and 34 orders. The data for the pools that were combined before

the PCR amplification contained 172 species, 139 genera, 65 families, and 31 orders. The morphological lists covered 214 species, 151 genera, 73 families, and 30 orders (excluding the water mites) (Supplementary File S3.1). DNA-based taxon richness was significantly correlated with morphological taxon richness for both sorted samples ($r = 0.662$, $p = 0.001$) and pooled samples ($r = 0.602$, $p = 0.002$), where redundant taxa had been removed (Figure 3.1). An additional 13 macroinvertebrates identified at species level were lost by the 0.025% threshold filtering (and only observed in the data that was discarded by this filter step). Seven of these were also recorded in the morphological assessment, the other six were only found using DNA. One of the species found in the discarded DNA-based data was *Musculium lacustre*, which was present in four samples where it was also detected morphologically, but only with one or two reads in each case.

To exclude the influence of sequencing depth (as sorted samples combined represented more sequencing depth than the pooled samples), we rarefied the samples to such an extent that sorted samples represented only one sixth of the pooled samples (most pools consisted of six combined extracts). Without rarefaction, the sorted samples had an average of 272,914 reads (range 170,637–424,726), which was 4.8 times more than the pools had (57,283 on average, range 15,061–122,002). They

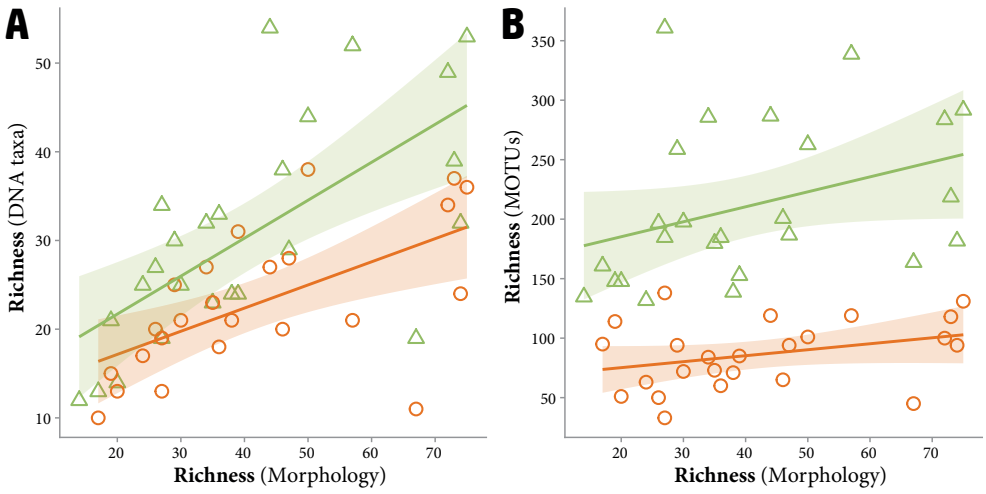


FIGURE 3.1. Relation between the morphological richness of samples and the (A) DNA taxon richness and (B) MOTU richness, for both the sorted samples (green triangles) and the pooled samples (orange circles), with a 95% confidence interval. Taxon richness was based on the taxon lists where redundant taxa had been removed. Correlations were significant for the taxon richness for both sorted samples ($r = 0.662$, $p = 0.001$) and pooled samples ($r = 0.602$, $p = 0.002$), but not for MOTU richness ($r = 0.365$, $p = 0.072$ and $r = 0.331$, $p = 0.115$, respectively).

also had 2.67 times as many MOTUs and 1.52 times as many taxa as the pools. With rarefaction the sorted samples still had 2.22 times as many MOTUs and 1.40 times as many taxa; neither was significantly lower than without rarefaction.

3.3.3 Comparison morphology versus molecular identification

Retaining the redundant taxa, the average richness of pooled samples (32.5 on average, range 16–56) was significantly lower than that of the sorted samples (47.6 on average, range 22–76) (Dunn's test, $p = 0.005$). When redundant taxa were removed, the richness of the pooled samples (22.9 on average, range 10–38) was again lower than the sorted samples (30.6 on average, range 12–54), but not significantly. Compared to the morphological richness with redundant taxa (46.7 on average, range 16–89), the richness of the pooled samples was significantly lower (Dunn's test, $p = 0.027$). The richness of the pooled samples was also significantly lower than the morphological richness when redundancy was removed (40.8 on average, range 14–75) (Dunn's test, $p < 0.001$). The richness of the sorted samples was not significantly different from the morphological richness in either situation.

For 13 out of 24 separately processed mollusc samples (one sample did not include molluscs) we were unable to amplify molluscs using the standard approach for DNA extraction and PCR. Additionally, four annelid samples, three Heteroptera/Coleoptera (HECO) samples, one crustacean sample, and one TOE sample failed to amplify, although the latter two only contained three and two species, respectively. The failed mollusc samples on average contained 13.2 morphologically identified taxa (range 5–20), the failed HECO samples 18.3 taxa (range 7–25) and the missing annelids accounted for 6.3 taxa (range 2–12). If taxa from the failed samples are excluded from the analysis (as they can only count towards the fraction of taxa not found by DNA), the overlap between the taxon list from sorted samples added together and the morphological taxon list was 56.8% on average (range 32.5–91.7%). On average, 22.9% of taxa were only found in morphology (range 0–50.0%), and 20.3% were only recovered using DNA (range 5.6–35.0%) (Figure 3.2A). If failed samples are included, the overlap between morphology and DNA was 47.6% on average (range 22.9–73.3%). For the pooled samples, the combined taxon lists contained an average of 47.3 taxa, with a 40.3% overlap between morphology and DNA (range 13.0–62.8%). 48.1% of taxa were only recorded in the morphological list (range 27.9–84.1%), and only 11.7% were found exclusively with DNA (range 0.0–27.8%) (Figure 3.2B). In 14 out of 24 samples (one pooled sample failed to amplify), the fraction of taxa only found with morphology was higher than the fraction of overlap between the two taxon lists, the fraction of taxa only found using DNA was never higher than the

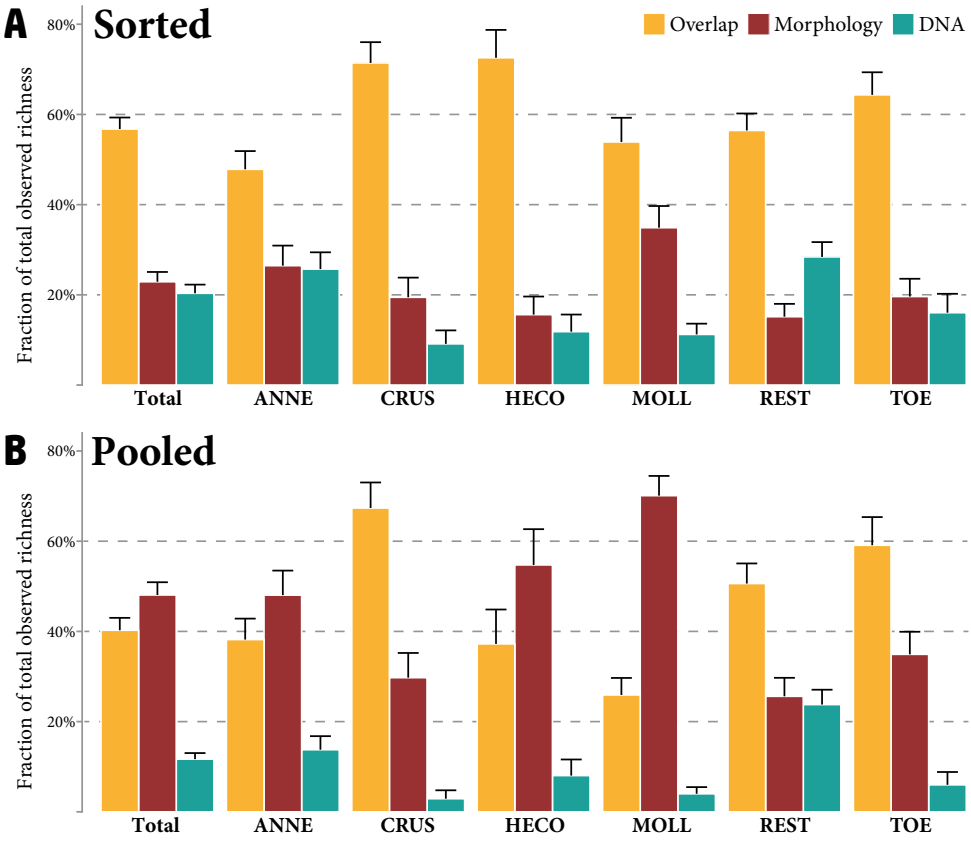


FIGURE 3.2. The fractions of total observed diversity present in both morphological assessment and DNA-based methods (yellow) and the fractions only represented in morphology (red) and DNA (blue), for the sorted samples (A) and pooled DNA analysis (B). The fractions were also assessed for each of the sorted taxa groups separately according to the following sorting of taxa: ANNE (annelids), CRUS (crustaceans), HECO (Heteroptera and Coleoptera), MOLL (molluscs), REST (rest groups, almost exclusively chironomids and other dipterans) and TOE (Trichoptera, Odonata and, Ephemeroptera). Error bars indicate the standard error.

fraction of taxa found only in the morphological analysis. In contrast, in 14 out of 25 samples where taxa were sorted prior to DNA analysis, the fraction of taxa exclusively found with DNA was higher than the morphology-only fraction.

The three categories that were counted towards the overlap contained 402 entries (72.6%) where there was a direct match between the species recorded in the morphological analysis, and the species identification obtained from metabarcoding. In 124 cases (22.4%) there was a match between morphology and metabarcoding, but the entries on both lists were not identified to the same taxonomic level. The majority

of these were annelids not covered in the reference database at species level (but were identified from molecular data at higher level using LCA) and dipterans identified to species level in the metabarcoding analysis but only identified at genus level or higher in the morphological data. The remainder were 28 cases of putative misidentifications (5.1%), where both list contained a different species from the same genus.

Looking at the six taxa groups separately (again excluding the failed samples), the overlap varies. The highest overlap was found in the crustaceans and HECO samples (71.4% and 72.6%, respectively), even though for HECO in one case the morphological and DNA-based taxon lists did not overlap at all (both, however, only contained one species each). The lowest overlap was found in the annelid samples (47.8% on average). Overlap for the MOLL, REST and TOE samples was 53.9%, 56.4% and 64.3% on average, respectively (Figure 3.2A). For the REST samples, the fraction of taxa found only in the DNA was larger than the fraction of taxa only recorded morphologically, for all other groups there were more taxa in the morphology list than there were on the DNA list. In 18 samples, more taxa were found with DNA than with morphology, in 26 samples more taxa were obtained with morphology. For 18 samples the morphology and DNA taxon lists was a complete match, although some taxa were not identified up to the same taxonomic level for both methods. In addition to the previously mentioned HECO sample, there was one other sample in the TOE set where DNA and morphology were mutually exclusive (Supplementary Figure S3.2). In the pooled samples, the overlap between morphology and DNA was considerably lower for most taxa groups, but most noticeable in the HECO and mollusc samples, where most taxa were only present on the morphological list. For all groups, more taxa were found with morphology than were found with DNA metabarcoding (Figure 3.2B).

3.3.4 Ecological quality ratios

The EQR scores based on the DNA data differed considerably from the morphology-based EQR scores for both the pooled and the sorted samples (Figure 3.3A and 3.3B). There was only a moderate correlation between the morphology- and DNA-based scores (Pearson correlation, r = 0.596 and 0.545, respectively). The scores obtained from the pooled samples were usually lower than the morphological scores (16 out of 24). For the sorted samples, half the samples (13 of 25) had a lower score using molecular identifications, the other half (12 of 25) scored higher based on DNA data. The average absolute difference in EQR score was similar for both datasets: 0.12 for the pooled samples (range 0.007–0.302) and 0.11 for the sorted samples (range 0.007–0.310). Using the pooled samples, 15 out of 24 locations scored in a different

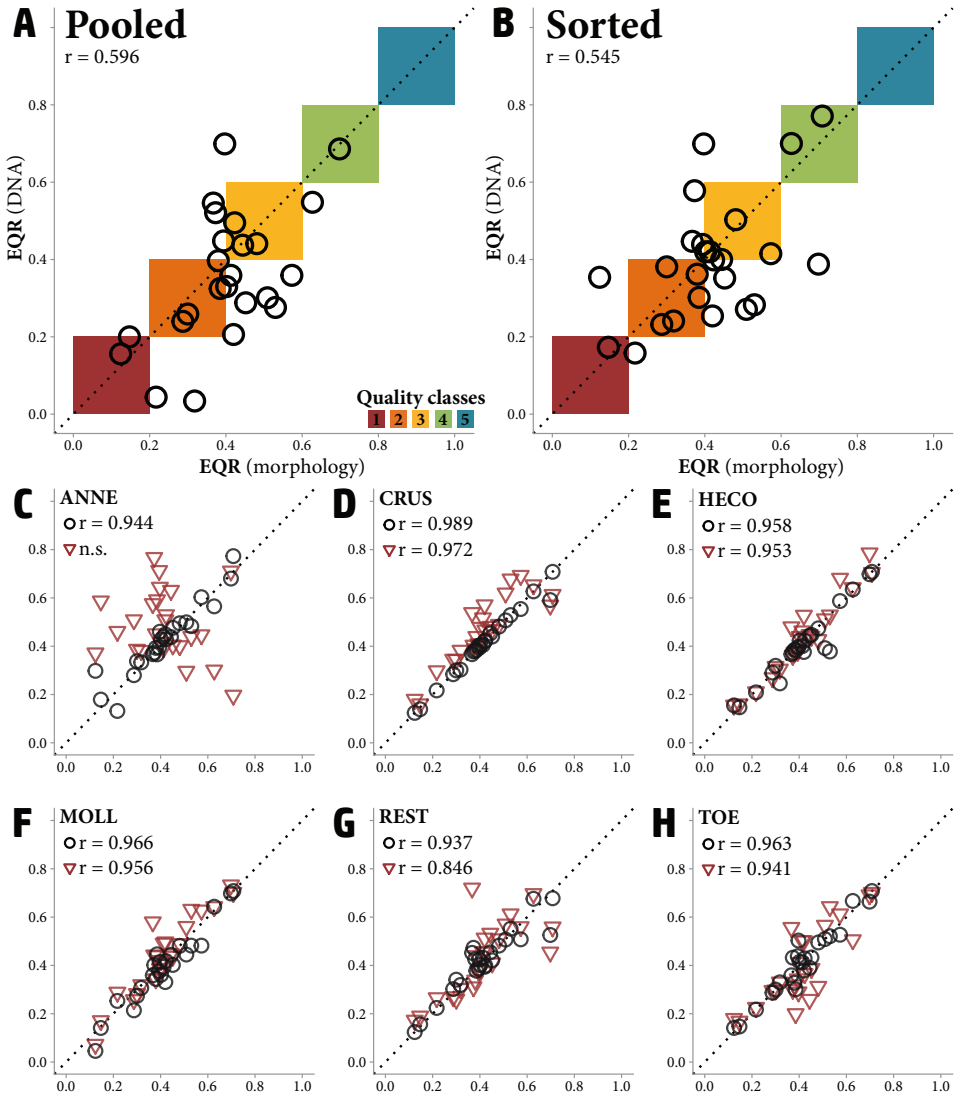


FIGURE 3.3. Comparison for the EQR score calculated on morphological data versus the score calculated on DNA data for (A) the pooled and (B) the sorted and separately sequenced samples. Both showed a moderate correlation between the scores (Pearson correlation, $p = 0.002$ and 0.005, respectively). Better correlations were found when only replacing one of six taxon groups with molecular data (on the y-axis): (C) Annelida, (D) Crustacea, (E) Heteroptera and Coleoptera, (F) Mollusca, (G) Chironomidae and other Diptera and (H) Trichoptera, Odonata, and Ephemeroptera (black circles, Pearson correlation values provided in the panels, $p < 0.001$ for all groups). To assess the influence of each of the respective groups on the EQR score, the original scores (x-axis) were also compared to EQR scores where one taxon group was completely removed (y-axis) from the taxon list before the analysis (C-H, red triangles, Pearson correlation values provided in the panels, $p < 0.001$ for all groups).

quality class (five higher, ten lower), and for the sorted samples, 12 of 25 ended up in a different quality class (five higher, seven lower). When replacing just one of the groups with molecular data for the EQR calculations, the correlations between the two scores were much stronger (ranging from r = 0.937 for REST to 0.989 for CRUS, $p < 0.001$ for all groups), even with the complete removal of some groups due to failed samples (Figure 3.3C-H).

3.4 DISCUSSION

We found that pre-sorting of samples into six basic taxon groups vastly improved the recovery of taxa using metabarcoding of bulk samples, with 46.5% more taxa found as compared to the samples where DNA was pooled prior to amplification and sequencing (47.6 versus 32.5 on average). The average overlap between the morphological and molecular (for the sorted samples) taxon lists was 56.8%, with the fractions of taxa found in only the morphology and only the DNA roughly equal (22.9% and 20.3%, respectively) (Figure 3.2A). Discrepancies between morphology and DNA-based species lists were expected, based on missing taxa from the reference database, known difficulties with morphological identification of taxa (Haase et al. 2006, Stribling et al. 2008), and primer biases (Elbrecht & Leese 2015) as contributing factors. Even though they were tested mainly on insects, the primers used in this study showed good *in silico* potential for all taxonomic groups included in our samples (Elbrecht & Leese 2017), especially compared to some other oft-used primers. While there may be primers that perform better for a specific group, a single, broad-range primer set that perform equally well on all taxa will most likely never exist (Creedy et al. 2019). Large differences were already observed between two morphological assessments in freshwater monitoring samples in previous studies, where there was more than 30% difference in identification of taxa. All taxon groups seemed to be equally prone to errors in morphological identification, even those deemed difficult to identify (Haase et al. 2010). In our data we see that the overlap between morphology and DNA varies between the different groups, being highest for the Crustacea and the Heteroptera /Coleoptera. The poor performance of the mollusc samples may not be entirely attributable to the primers, as molluscs are the group that is most affected by differences in biomass between the different species in a sample.

There have been few studies comparing morphological identifications and DNA-based identifications on actual samples, instead of relying on mock communities. A study assessing the taxa detected by morphology and DNA on Finnish WFD samples (using the same primers as this study) found considerably more taxa with

DNA than they did with morphology (Elbrecht et al. 2017a), but morphological assessments did not include species or genus level identifications for certain groups, such as the species-rich Chironomidae. The taxonomic resolution in the present study was comparable between morphology and DNA metabarcoding, and explains why richness estimations were more comparable on average. Still, we found some differences between taxon lists caused by disparity in resolution for certain taxa. On the side of the morphology, higher-level taxonomic identifications have been made due to the difficulty of distinguishing taxa, especially those in larval stages. For example, none of the Ceratopogonidae had been identified beyond family level using morphology, but five different genera were detected with DNA. On the other hand, the DNA reference database did not cover all the taxa that were listed in the morphological dataset (6.5% of the morphologically identified species had no DNA reference). For instance, every specimen of *Alboglossiphonia* was only identified up to genus level using the LCA tool in the DNA analysis, as all three species recorded in the morphological analysis were unaccounted for in the reference database (sequences could still be identified to genus level based on matches to congeneric species).

Some groups that were examined in this study consists of considerably more taxa than others. This difference in group size inevitably leads to a larger number of "lost taxa" when one taxon dominates the reads due to the effects of preferential amplification. In the majority of the pooled samples (15 of 24) more than half of the reads is provided by one of the six groups (Supplementary Figure S3.3), and in eleven samples more than half the reads even belonged to a single taxon. While some have argued that for general patterns in biodiversity, the effects of primer bias may be limited, the taxonomic bias caused by primer mismatches in certain taxonomic groups can be an issue when trying to reconstruct taxa lists (Creedy et al. 2019). Taxonomic sorting can improve the recovery of taxa, as witnessed by the improved performance of the sorted and separately sequenced samples in comparison to the pooled samples, which represent a broader range taxa. In the sorted samples, 46.5% more taxa were found than in the pooled samples (47.6 versus 32.5 on average), also leading to more overlap with the morphological list (56.8% versus 40.3% on average). Similar improvements have been found when using a size-based sorting of specimens prior to DNA extraction and amplification, where around 30% more taxa were found compared to non-sorted samples (Elbrecht et al. 2017b), although others report that amplification bias across size ranges may be limited with deep sequencing (Creedy et al. 2019). When assessing the separate groups, the effect of the pooling of samples prior to DNA amplification and sequencing has the largest effect on the HECO and mollusc samples, where 65.6% and 46.6% fewer taxa were found in comparison to the

sorted and separately sequenced samples (Figure 3.2B). Rarefaction showed that the reduced sequencing depth of the pools, when compared to the combined separately sequenced samples, was not solely responsible for the reduction in detected taxa. Even when rarefied to the same sequencing depth, we still obtained 40.0% more taxa in the sorted samples. A study assessing the taxonomic recovery of tropical forest arthropod communities showed similar findings, where there was some decline in MOTUs recovered for specific taxon groups in increasingly complex mixtures. This was mostly caused by the introduction of other groups, which were apparently amplified preferentially (Creedy et al. 2019), comparable to our observations with HECO and mollusc taxa. Taxonomic sorting into the groups presented in this study is relatively straightforward and would require only superficial knowledge of taxonomy. Compared to genus or species level sorting and identification, both the time and costs involved are between one and two orders of magnitude lower (Marshall et al. 2006, Jones 2008).

The difficulties in identifying specimens using morphology can also express themselves in the DNA-based identities, by way of having erroneously identified specimens within the DNA reference library. We encountered a variety of unresolved taxa and putative identification errors in the reference data downloaded from BOLD. In the 350,449 public sequences we found 554 cases where congeneric species had identical sequences. These are not necessarily identification errors, as some closely-related species are known to be indistinguishable by the DNA barcode region (Huemer et al. 2014), but do highlight the need to not look at just the "top 1" or "best hit" matches when comparing sequences to a reference database. When multiple hits with the same scores are found, matching algorithms do not always consistently place the same match at the top of the list, introducing random variation between analyses when only looking at the first hit. Additionally, 47 cases of identical sequences with different species from different genera were found, some of which could be traced back to actual contaminated sequences (e.g. *Homo sapiens* or *Wolbachia*). Most of such misidentified records have been flagged by BOLD curators, which was verified by manually checking a random selection of records. Moreover, recent analysis of the BOLD data revealed a relatively high number of specimens that had been identified using "reverse BIN taxonomy", adding further levels of uncertainty to the reference datasets retrieved from BOLD (Weigand et al. 2019). To improve the use of public data such as the sequences deposited in BOLD, the ability to filter data based on record flags or identification method is essential.

An incomplete reference databases is a major issue that limits the use of metabarcoding for species identification (Kvist 2013, Wangensteen et al. 2018). While 93.5% of the morphologically identified species of this study had reference

sequences, database coverage for all Dutch WFD taxa is only 86.1%. There are large differences for each of the taxa groups as defined by this study, with 63.5% of annelids covered by reference sequences (54 of 85 species), while 96.5% of the TOE group has been barcoded (273 out of 283 species). Additionally, we still observe difficulties in identification for species known to have high genetic diversity. For example, 14 MOTUs were identified as *Asellus aquaticus*, another 109 were identified at genus level as *Asellus* (for which only *A. aquaticus* is recorded in the Netherlands). The tendency to overestimate richness based solely on MOTUs (see also Figure 3.1) has already been reported in the past, with population and haplotype differences increasing richness estimates (Gibson et al. 2015, Elbrecht et al. 2018a). The length threshold used in this study (allowing for sequences which were three base pairs shorter or longer than the 316 bp target to pass quality filtering) may have contributed to an overestimation of richness based on MOTUs. We aimed to mitigate this effect by aggregating all MOTUs with identical taxonomic identification and discarding any unidentified MOTU from the analysis. While alternate clustering methods may exaggerate or downplay this effect of overestimation, the difference in intraspecific variation between taxonomic groups will lead to either overestimations for taxa with high intraspecific variation or underestimations by lumping taxa with low interspecific variation depending on the cluster settings. The observed variation also suggests that the DNA reference library could be improved by better geographical coverage, incorporating a wider range of haplotype variation. Another phenomenon that may have caused an overabundance of *Asellus* and other genera in the MOTUs, is the presence of pseudogenes that have been amplified (Song et al. 2008, Brown et al. 2015), especially with deep sequencing of highly abundant taxa. Many of the MOTUs identified at genus level have fewer reads (1,768 on average) compared to the MOTUs with species level identification (75,128 reads on average). Similar patterns were seen for other genera as well (e.g. *Helobdella*, *Limnomysis*), including genera that have more than one species recorded for the Netherlands, and which were all represented in the reference database (e.g. *Cymatia*, *Erythromma*, *Noterus*).

Haplotypes and pseudogenes aside, we should be wary of the fact that many taxon groups still contain undescribed diversity and cryptic species (Hebert et al. 2016), which may be perceived as overestimations of taxa when using DNA-based identification methods. This information can still be valuable, as it has been shown in mayflies that cryptic species exhibit a wide variety of tolerances and responses to ecosystem stressors (Macher et al. 2016). Furthermore, MOTU level analysis of Chironomidae has demonstrated that even without binomial names, different putative taxa could be identified, showing different response patterns (Beermann et al. 2018).

This opens possibilities to use DNA-based delimitations for comparative quality assessments and impact studies even for those taxon groups that are poorly defined in reference databases, which is still hampering the use of DNA-based identification in various groups (Curry et al. 2018). DNA-based identification may not always exactly reflect the observations made by traditional morphological methods, but at least may provide a more consistent way of identifying taxa (Bush et al. 2019). Morphological assignments are prone to discrepancies between assessors, as shown by large differences between identification made in audits of WFD assessments (Haase et al. 2006, Stribling et al. 2008). The choice of tools and parameters used in the processing of raw sequence data (such as filtering and clustering) can have a significant impact on taxonomic inferences as well (Alberdi et al. 2018, Porter & Hajibabaei 2018), but in comparison with morphological assessments should be easier to report and standardize. Molecular data is also more easily re-analyzed when new insights are developed, and is backwards compatible with updated reference databases.

The impact of DNA-based identifications on the EQR scoring is considerable, for both the pooled and the sorted samples, with neither giving a better approximation of the morphology-based score (Figure 3.3A and 3.3B). The correlation between the two scores was only moderate (Pearson correlation, $r = 0.596$ and 0.545, respectively). EQR scores for the pooled samples were generally lower than the morphology-based EQRs (16 out of 24), whereas the sorted samples provided scores that were lower in half the samples, higher in the other half. The average absolute difference between the EQRs obtained from morphological data and the DNA-based scores was similar for both datasets (0.12 and 0.11 for the pooled and sorted samples, respectively). When replacing just one of the six groups from the morphological taxon list with DNA-based identifications, the impact on the EQR score was considerably lower. This allows for the use of DNA metabarcoding for a select group of taxa, for example in cases where morphological assessments are difficult or time-consuming (such as Chironomidae), without the need to recalibrate the entire EQR scoring method. In such cases, it would also be possible to use primers that are tailored more specifically to the investigated taxa, in order to limit primer bias. The largest deviation was seen in the mollusc samples, where the absolute difference was 0.033 on average (range 0–0.091), but the scores were still strongly correlated (Pearson correlation, $r = 0.966$, $p < 0.001$). Molluscs were also the group for which most samples failed to amplify (13 out of 24, Supplementary Figure S3.2), but this did not seem to have too much of an impact on the scoring. However, complete removal of groups can have a substantial impact on the EQR score, especially for the annelids (Figure 3.3C-H). Removal of the water mites, which were excluded from the analysis due to inability to obtain

DNA, had a comparable impact on the EQR scoring to some of the other groups (Supplementary Figure S3.4). To minimize the effect of stacking these impacts, the water mites were completely discarded from all EQR analyses.

The fact that DNA-based EQR scores are so different from the scores based on traditional morphological surveys can partly be attributed to the changes in taxonomic resolution and deviations between the two taxon lists, but the differences are likely exaggerated by the changes in richness as well. For the Dutch EQR calculations, the percentage of characteristic taxa, and positive and negative indicator species (as a fraction of the total richness) play an important role for the final score (van der Molen et al. 2016, Chapter 2). While the average taxon richness was not significantly different between the morphological assessment and the sorted and separately sequenced samples used for the DNA-based calculations, the differences for each sample were considerable (Figure 3.1A), with an average difference in richness between morphology and DNA of 12.0 (range 1–48). Together, the changes in the number of negative and positive indicators, and the changes in the ratios of these indicators can have a significant impact on the final EQR score (Figure 3.3A and 3.3B).

Molecular techniques may not directly replace traditional morphology under the current WFD monitoring standards, and future monitoring requires a paradigm shift to fully incorporate the potential of DNA-based methodology. Time is needed for new techniques to prove their worth in the field of biomonitoring and be accepted by monitoring agencies and policy makers. Being able to replace morphological assessment for only one or a few taxon groups without needing to redefine the framework for BQE monitoring, opens possibilities for gradual implementation of DNA-based identifications for those groups that are most difficult to identify, time-consuming or where taxonomic expertise is getting scarce.

3.5 CONCLUSIONS

There are considerable differences when directly comparing the outcome of traditional morphological assessment and DNA metabarcoding-based identifications of bulk samples, including their effects on the EQR score under current standards. Our data shows that DNA metabarcoding compares better to morphological assessments for some taxonomic groups than for others, partly based on the underlying DNA reference database, or lack thereof. Mismatches were observed between morphology and metabarcoding, but the latter will be less reliant on individual biases introduced by different assessors, and therefore lead to more consistent assessments. Taxonomic sorting into basic groups improves the taxon recovery, as shown in this study, where

46.5% more taxa were found when samples were sorted into six basic groups prior to DNA amplification and sequencing. Even when corrected for sequencing depth, sorted samples still produce around 40% more taxa as non-sorted samples. DNA-based assessments may not directly replace traditional monitoring in the near future, but can certainly contribute to the current methodology, especially for those groups that are perceived as difficult to identify, to allow for more consistent and faster identifications. Metabarcoding would greatly improve with addition of vouchered specimens to reference databases. Furthermore, we show that replacing only one of six taxa groups assessed in this study by molecular data has limited impact on the EQR scoring, opening possibilities for gradual replacement of traditional identification, or supplementing the traditional identification with DNA-based tools, which will help with the acceptance of molecular methodology in WFD monitoring.

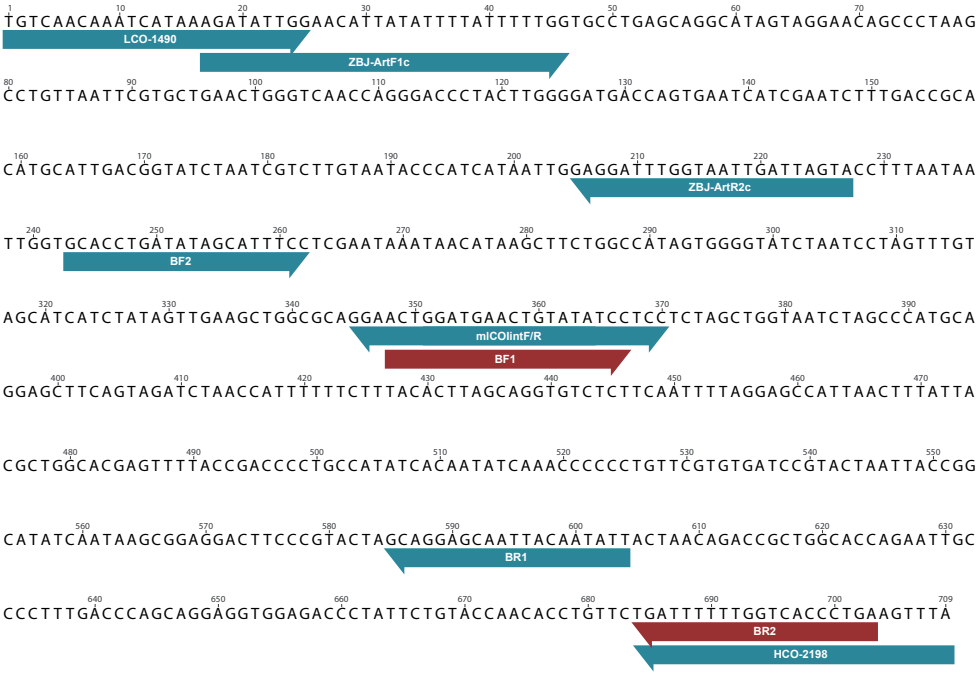
3.6 ACKNOWLEDGEMENTS

We thank Aquon and Wouter Balster for supplying the samples and morphological identifications, as well as Bart Schaub of Hoogheemraadschap van Rijnland for providing us with access to the collection of WFD samples. This study was part of the DNA Waterscan project, funded by the Gieskes-Strijbis Fonds (https://gieskesstrijbisfonds.nl/). The funder provided support in the form of material costs and the salaries for authors KKB, MH and RP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

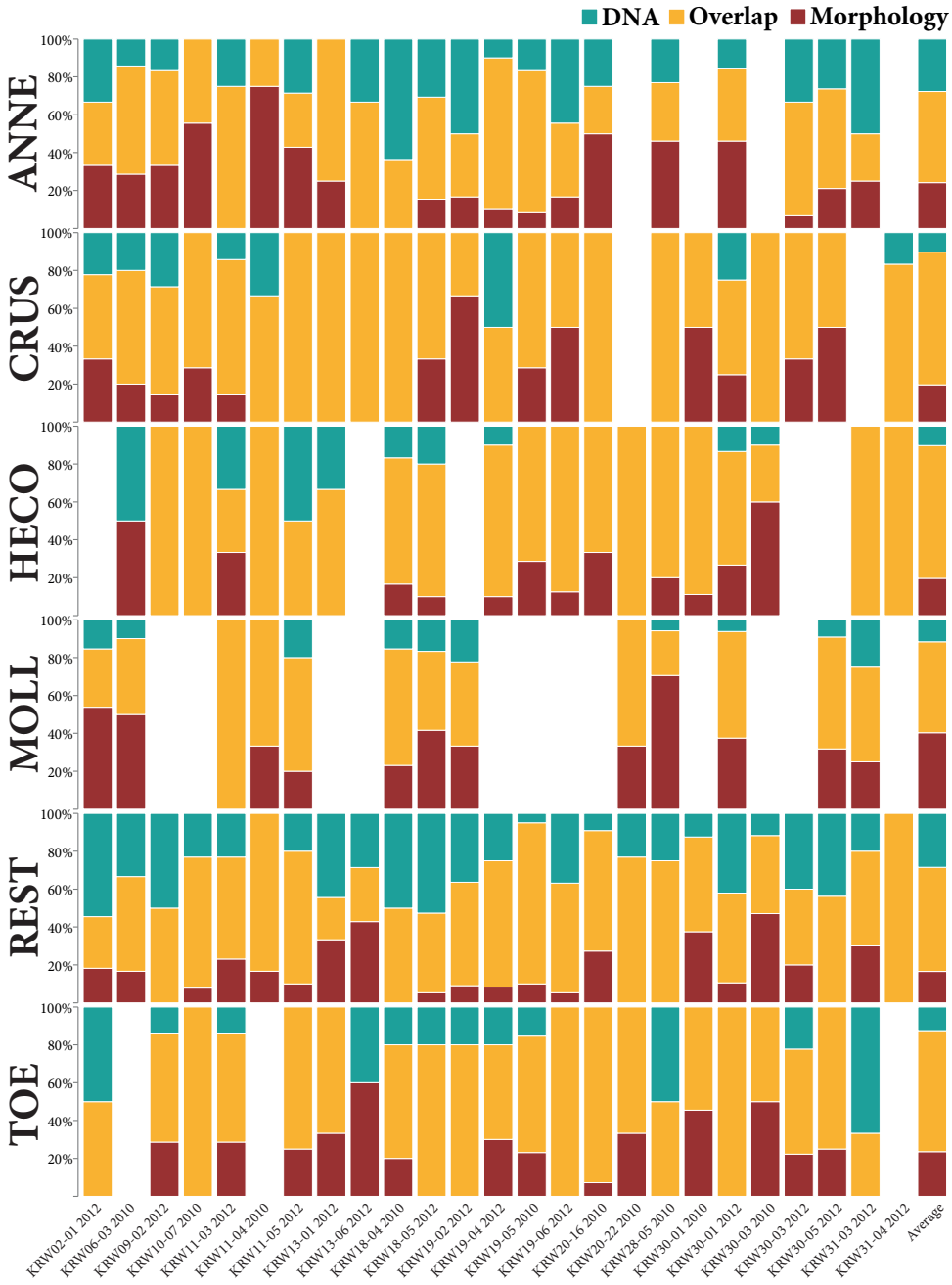
3.7 DATA AVAILABILITY

Raw sequence data is available from the NCBI Sequence Read Archive (Bioproject accession PRJNA550542).

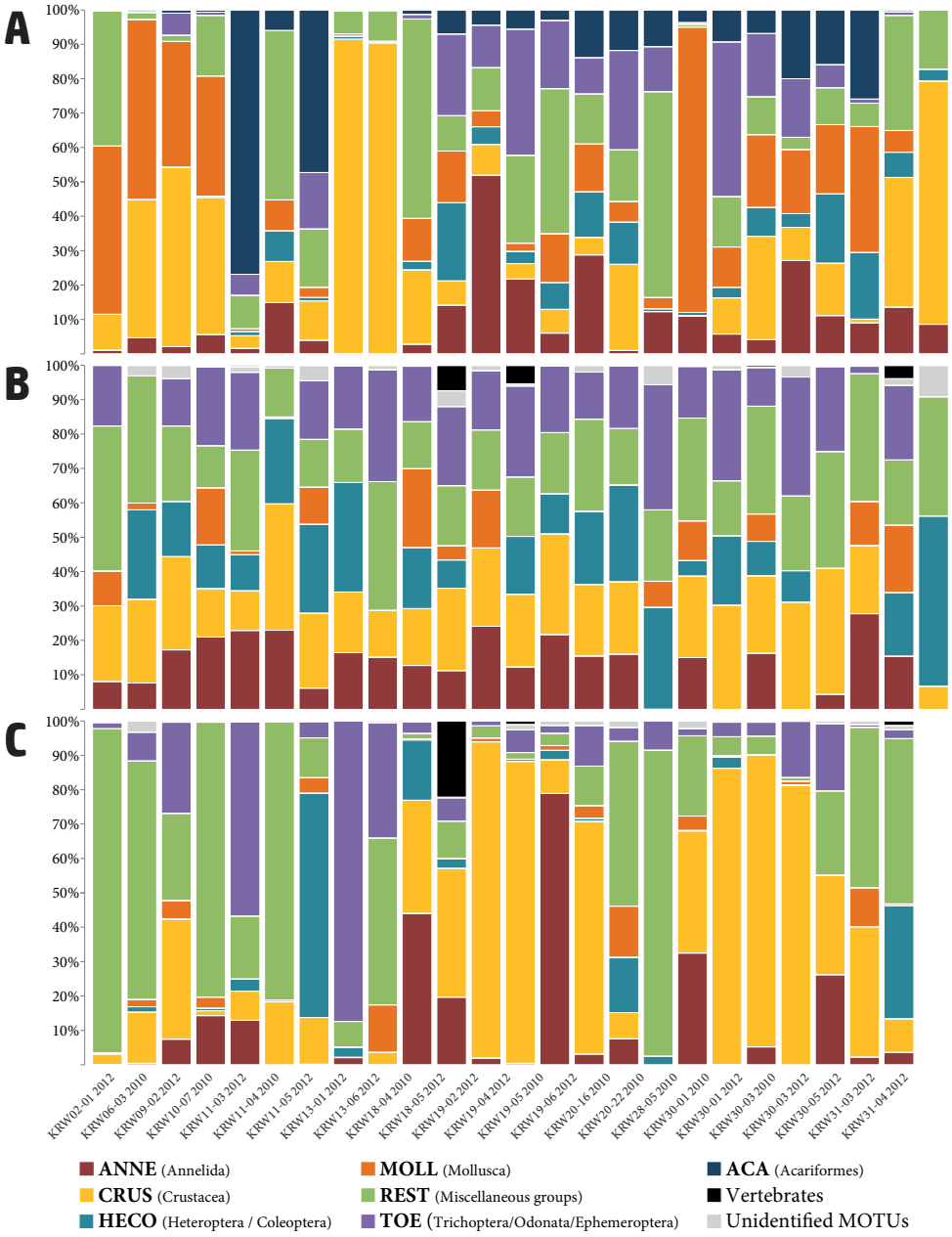
3.8 SUPPLEMENTARY MATERIALS



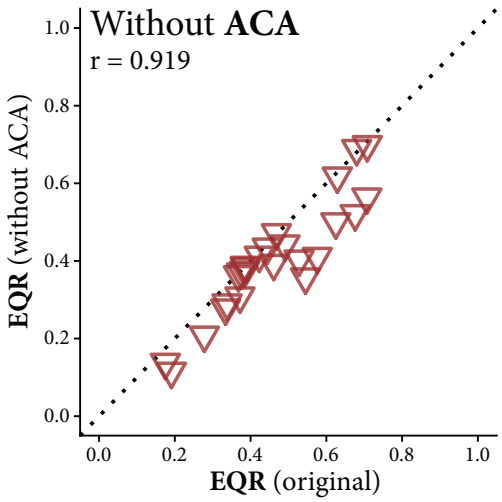
SUPPLEMENTARY FIGURE S3.1. The artificial control (AIC) used to measure cross-contamination. The sequence is based on the COI barcode region of a Reeve's muntjac (*Muntiacus reevesi*) with several primer sets built into the sequence (forward strand shown). Binding sites for COI primers are shown, BF1 and BR2 used in this study are highlighted in red.



SUPPLEMENTARY FIGURE S3.2. The overlap between morphology and DNA (in yellow), as well as the fractions of taxa only detected with DNA (blue) and morphology (red), for each of the 25 samples separately, as well as averages (last column), separated for each of the six taxa groups.



SUPPLEMENTARY FIGURE S3.3. Relative abundances of (A) specimens in the traditional morphological assessment and reads in the metabarcoding data of (B) separately sequenced taxa groups combined and (C) samples pooled prior to amplification. In addition to the six groups assessed in this study, the fractions of water mites (in morphology), as well as vertebrates and unidentified MOTUs (in DNA data) have been included.



SUPPLEMENTARY FIGURE S3.4. Comparison of EQR scores for the morphological data with and without water mites (ACA). No DNA was obtained from water mites due to the buffer they were stored in. Pearson correlation value provided in the panel, $p < 0.001$.

SUPPLEMENTARY FILE S3.1. Taxon lists for the three datasets: Morphologically identified taxa (with specimen counts), DNA-based identifications from the sorted samples, and DNA-based identifications form the pooled samples (both with read counts). **https://doi.org/10.1371/journal.pone.0226527.s005**

CHAPTER 4

The effects of spatial and temporal replicate sampling on eDNA metabarcoding

Kevin K. Beentjes[1,2], Arjen G. C. L. Speksnijder[1], Menno Schilthuizen[1,2], Marten Hoogeveen[1], Berry B. van der Hoorn[1]

PeerJ 7:e7335 (2019)
https://doi.org/10.7717/peerj.7335

[1] Naturalis Biodiversity Center, Leiden, The Netherlands
[2] Institute of Biology Leiden, Leiden University, Leiden, The Netherlands

ABSTRACT

The heterogeneous nature of environmental DNA (eDNA) and its effects on species detection and community composition estimates has been highlighted in several studies in the past decades. Mostly in the context of spatial distribution over large areas, in fewer occasions looking at spatial distribution within a single body of water. Temporal variation of eDNA, similarly, has mostly been studied as seasonality, observing changes over large periods of time, and often only for small groups of organisms such as fish and amphibians.

We analyzed and compared small-scale spatial and temporal variation by sampling eDNA from two small, isolated dune lakes for 20 consecutive weeks. Metabarcoding was performed on the samples using generic COI primers. Molecular operational taxonomic unit (MOTUs) were used to assess dissimilarities between spatial and temporal replicates.

Our results show large differences between samples taken within one lake at one point in time, but also expose the large differences between temporal replicates, even those taken only 1 week apart. Furthermore, between-site dissimilarities showed a linear correlation with time frame, indicating that between-site differences will be inflated when samples are taken over a period of time. We also assessed the effects of PCR replicates and processing strategies on general patterns of dissimilarity between samples. While more inclusive PCR replicate strategies lead to higher richness estimations, dissimilarity patterns between samples did not significantly change.

We conclude that the dissimilarity of temporal replicates at a one week interval is comparable to that of spatial replicate samples. It increases, however, for larger time intervals, which suggests that population turnover effects can be stronger than community heterogeneity. Spatial replicates alone may not be enough for optimal recovery of taxonomic diversity, and cross-comparisons of different locations are susceptible to inflated dissimilarities when performed over larger time intervals. Many of the observed MOTUs could be classified as either phyto- or zooplankton, two groups that have gained traction in recent years as potential novel bio-indicator species. Our results, however, indicate that these groups might be susceptible to large community shifts in relatively short periods of time, highlighting the need to take temporal variations into consideration when assessing their usability as water quality indicators.

4.1 INTRODUCTION

The importance of freshwater biodiversity and its effects on ecosystem resilience and stability have been well documented, and its monitoring is regulated by legislation such as the European Union Water Framework Directive of 2000 (EU WFD; Directive 2000/60/EC). Monitoring of biological quality elements (BQE), such as macroinvertebrates, is prescribed under the WFD, but traditional methods employed in this field are often considered slow, expensive, and sensitive to human-induced bias and errors (Clarke & Hering 2006). Integration of molecular tools has been a focal area within this field of research for the past decade. The use of environmental DNA (eDNA) metabarcoding for species detection is gaining traction, as it would potentially enable to circumvent cumbersome traditional collection or visual observation of specimens. The use of eDNA for detection is based on the fact that organisms living in a certain environment, such as freshwater, leave behind traces of their existence via shedding and excretion of DNA. This technique has been applied successfully for the detection of a multitude of species, including BQEs, in both vertebrates (Ficetola et al. 2008, Hänfling et al. 2016, Olds et al. 2016) and invertebrates (Thomsen et al. 2012b, Schneider et al. 2016, Klymus et al. 2017).

The heterogeneous nature of eDNA has been investigated in several model organisms, for example amphibians, where it was shown that spatial sampling increased the detection probability (Dejean et al. 2012, Schmidt et al. 2013). Similarly, richness estimates from eDNA community metabarcoding are sensitive to sampling strategies (Grey et al. 2018). This suggests that eDNA may only represent very local signals, especially in standing waters. It is therefore often recommended to include spatial coverage in an eDNA sampling strategy, either by sampling various points within a water body, or by combining all these samples into one large sample representing the entire water body (Goldberg et al. 2016, Grey et al. 2018, Harper et al. 2019a). In addition to spatial sampling, temporal replicates may also increase detection probability, and provide a more complete impression of species richness and community composition. Many studies have examined the effects of spatial and temporal sampling on (macroinvertebrate) communities (Baselga et al. 2013, Barsoum et al. 2019), but limited work has been done on seasonal variation in aquatic eDNA. Most research focuses on one particular organism or groups of organisms, such as fish (Stoeckle et al. 2017, Sigsgaard et al. 2017), amphibians (Rees et al. 2017, Buxton et al. 2018), and chironomids (Bista et al. 2017), or assesses the seasonal differences only at a limited number of points in time (Chain et al. 2016, Guardiola et al. 2016).

In this paper, we compare the effects of both spatial and temporal replicate sampling of eDNA within two isolated, but nearby, lakes, using a generic COI primer set. We assess patterns in communities based on molecular operational taxonomic unit (MOTU) clustering, identifying MOTUs using a lowest common ancestor (LCA) approach, and also look at the communities of only those MOTUs identified as metazoans. Furthermore, we assess the impact of PCR replicates and subsequent sequence or bioinformatics processing strategies on the observed patterns of eDNA through space and time. We also highlight some potential opportunities and caveats in the use of eDNA for freshwater quality monitoring.

4.2 MATERIALS AND METHODS

4.2.1 Field sampling

Samples were collected on every Monday for 20 consecutive weeks, from May 2016 to September 2016, from two permanent lakes in a Natura 2000 protection area in the dunes of Wassenaar, the Netherlands. Two locations were selected, approximately 1.9 km apart: Location 1 "De Ezelenwei" (52.161°N, 4.354°E) and Location 2 "De Drie Landjes" (52.176°N, 4.367°E). The sampling window coincides with the sampling period for traditional WFD monitoring. Within each location three sub-sites were selected around the lake, roughly equidistant from each other (40–60 m apart) and representing different habitats and substrates. A total of 1 l of water was taken by submerging a 1-l sterile bottle slightly below the surface, one meter away from the lake shoreline. The bottles were brought back to the laboratory for filtration. As the sites were located in a nature conservation area, a permit was obtained from Staatsbosbeheer (2016/022).

4.2.2 DNA filtration and extraction

Environmental DNA filtration was performed in the laboratory within 4 h after collecting the samples in the field. Sterilized Nalgene filter units (Thermo Fisher, Waltham, MA, USA) attached to a vacuum pump with 0.2 μm polyethersulfone filter membranes (Sartorius, Göttingen, Germany) were used to filter 300 ml of water. Filter holders were sterilized using 10% bleach solution and placed under UV-light for 30 min before use. After filtration, the filter membranes were stored in 900 μl CTAB buffer at −20 °C until extraction. DNA was extracted using a modified CTAB extraction protocol, adapted from Turner et al. (2014). DNA precipitation was performed on 800 μl of aqueous phase, and final resuspension of the pellet was performed in 50 μl AE buffer (Qiagen, Venlo, the Netherlands).

4.2.3 DNA amplification and MiSeq sequencing

A 316 bp fragment of the COI barcode region was amplified using primers BF1 and BR2 (Elbrecht & Leese 2017). All sampling replicates were amplified in three independent PCRs, which were sequenced separately without pooling. A dual indexed MiSeq amplicon library was prepared using a two-step PCR protocol, in which the first PCR used primers BF1 and BR2 with 5' Illumina tails (Supplementary Tables S4.1 and S4.2). PCRs for round 1 were performed in 25 µl reactions containing 1× Qiagen CoralLoad PCR Buffer, 0.5 mM dNTPs, 0.05 U/µl Taq polymerase (Qiagen, Venlo, the Netherlands), 0.4 µM of each primer and 1.0 µl of template DNA. Initial denaturation was performed at 94 °C for 3 min, followed by 40 cycles at 94 °C for 15 s, 50 °C for 30 s, and 72 °C for 40 s, followed by final elongation at 72 °C for 5 min. Each 96-well plate contained blanks with no template DNA and positive controls of Reeve's muntjac (*Muntiacus reevesi*) DNA extract to enable detection of cross-contaminations in the laboratory process. PCR success was checked on an E-Gel 96 pre-cast agarose gel (Thermo Fisher, Waltham, MA, USA). PCR products where then cleaned with a one-sided size selection using NucleoMag NGS-Beads (Macherey-Nagel, Düren, Germany), using a 1:0.9 ratio.

Second round PCRs were performed using 2.0 µl of PCR product from the first round in a 20 µl reaction containing 1× TaqMan Environmental Master Mix 2.0 (Thermo Fisher, Waltham, MA, USA) and 1.0 µM of each primer. Initial denaturation was performed at 95 °C for 10 min, followed by 11 cycles at 95 °C for 30 s, 55 °C for 60 s, and 72 °C for 30 s, followed by final elongation at 72 °C for 7 min. Second round PCR products were quantified on the QIAxcel (Qiagen, Venlo, the Netherlands) and pooled equimolarly per PCR plate. Pools were cleaned with a one-sided size selection using NucleoMag NGS-Beads, ratio 1:0.9, then quantified on the Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) with the DNA High Sensitivity Kit. The four pools were combined equimolarly and sequenced on one run of Illumina MiSeq (v3 Kit, 2 × 300 paired-end) at LGTC (Leiden, the Netherlands).

4.2.4 Quality filtering and MOTU clustering

Quality filtering and clustering of all data was performed in a custom pipeline on the OpenStack environment of Naturalis Biodiversity Center through a Galaxy instance (Afgan et al. 2018). Raw sequences were filtered using Sickle (Joshi & Fass 2011) and merged using FLASH v1.2.11 (Magoč & Salzberg 2011); all non-merged reads were discarded. Samples were split based on the presence of template-specific additional bases between Illumina tail and template-specific primers with a custom tool, and primers were trimmed from both ends of the merged reads using Cutadapt v1.16

(Martin 2011). Any read without both primers present and anchored was removed. PRINSEQ v0.20.4 (Schmieder & Edwards 2011) was used to filter reads with length below 310 bp and above 316 bp from the dataset. Sequences were dereplicated using VSEARCH v2.4.3 (Rognes et al. 2016) and clustered into MOTUs using UNOISE3 (Edgar 2016) with an alpha of 0.5. The presence of *M. reevesi* reads in the non-control samples was used to determine the MOTU filtering threshold, only MOTUs with read abundances above 0.05% were retained for each replicate. Geneious 8.1.8 (https://www.geneious.com) was used to check for and remove MOTUs with indels and/or stop codons.

4.2.5 Taxonomic assignment and diversity analysis

BLAST+ (Camacho et al. 2009) was used to compare MOTU sequences to a custom-made reference library containing COI sequences and bacterial genomes downloaded from NCBI GenBank (Benson et al. 2005) (sequences downloaded August 21, 2018). MEGAN v6.12.5 (Huson et al. 2007) was used to assign higher-rank taxonomy to MOTUs using the LCA approach from the top 100 hits from BLAST (settings: minimum bit score 170, minimum percent identity 80, top percent 5). The VEGAN package (Oksanen et al. 2007) in R was used to calculate beta diversity (Sørenson dissimilarity) between replicates and time points, make NMDS plots, and calculate correlations between dissimilarity matrices and between the sample dissimilarity and sampling intervals. PCR replication effects were assessed using three methods of replicate processing: (1) counting all MOTUs toward the sample ("additive"), (2) only counting those MOTUs that appear in a majority of the samples ("relaxed"), or (3) only counting those MOTUs that occur in all replicates ("strict") (Alberdi et al. 2018). All analysis on the data were performed for both the whole dataset (all MOTUs), and a subset of the data with only metazoan MOTUs.

4.3 RESULTS

4.3.1 Sequencing run statistics

A total of 7,692,379 read pairs were obtained after sequencing. After merging and quality filtering, 5,743,638 sequences were retained for MOTU clustering. M. reevesi reads were detected in several non-control samples. Using a 0.05% threshold for filtering low-abundance MOTUs from each sample removed muntjac reads from all but one sample (Location 1.2, May 16). After filtering the MOTU table, 1,333 MOTUs were retained in the non-control samples. An additional 19 MOTUs with indels and stop codons were removed, resulting in a dataset with 1,314 MOTUs, representing

4,197,403 reads. Four samples with fewer than 2,000 reads were discarded. On average, PCR replicates had 11,790 reads (range 2,296–73,477), and 72 MOTUs (range 12–177). There was no correlation between number of reads and number of MOTUs in each sample.

4.3.2 Taxonomic composition

Out of 1,314 remaining MOTUs, 530 (40.3%) eukaryotes could be identified to at least phylum level using the LCA, 119 (9.1%) were only classified as "eukaryote," 62 (4.7%) were identified as bacteria and 603 (45.9%) were not assigned any classification (Figure 4.1). Within the eukaryotes, most MOTUs (318) were classified as stramenopiles. Of the 176 metazoans, 121 were identified as arthropods, mostly assigned to branchiopods (44 MOTUs) and insects (26 MOTUs). Of the 1,314 MOTUs, 537 (40,9%) were found in both lakes, 418 MOTUs were unique to location 1 (De Ezelenwei), and 359 MOTUs unique to location 2 (De Drie Landjes).

The MOTU communities differed significantly between the two lakes for all 20 sampling moments, which is reflected in the NMDS plot based on the Sørenson dissimilarity matrix (Figure 4.2). Clustering of samples into their respective lakes was supported by ANOSIM ($R = 0.710$, $p = 0.001$). Similarly, ANOSIM also supported grouping of samples into two seasonal groups, spring (2 May–13 June), and summer (20 June–12 September) ($R = 0.486$, $p = 0.001$). For the metazoan-only subset, the separation between the locations is still supported by ANOSIM, albeit not as clear as in the dataset with all MOTUs ($R = 0.424$, $p = 0.001$). The grouping into spring and summer is also supported ($R = 0.587$, $p = 0.001$).

4.3.3 PCR replicates

Out of 1,314 MOTUs, 110 only ever occurred in one PCR replicate, with an average of 14.0 ± 1.6 (mean ± SEM) reads. The other 1,204 MOTUs occurred on average in 21.2 ± 1.1 of the 356 total replicates. No MOTU was found in all replicates. Average Sørenson dissimilarity between PCR replicates was 0.26 (Figure 4.3). Using the "additive" PCR processing strategy, samples had an average of 102.5 ± 4.0 MOTUs. Under the "relaxed" scenario samples had an average of 65.7 ± 2.4 MOTUs, and 280 MOTUs were discarded from the MOTU table. In the "strict" scenario an additional 246 MOTUs were discarded (Table 4.1). The remaining 788 MOTUs still represented 95.1% of the total read data. One PCR replicate on average contained 70.9% of MOTUs found in the total spatial replicate sample (the three PCR replicates combined) (range 34.6–95.8%), two replicates combined were able to detect an average of 88.4% of the MOTUs (range 55.8–100%). In only ten of 120 samples, the addition of a third PCR

replicate did not result in additional MOTUs found. Seven of the PCR replicates contained no MOTUs that could be identified as metazoan, two subsamples had no metazoan MOTUs in any of their PCR replicates. Average Sørenson dissimilarity between PCR replicates in the metazoan-only subset of the data was 0.18 (Figure 4.3), although in some cases it was as high as 1.0.

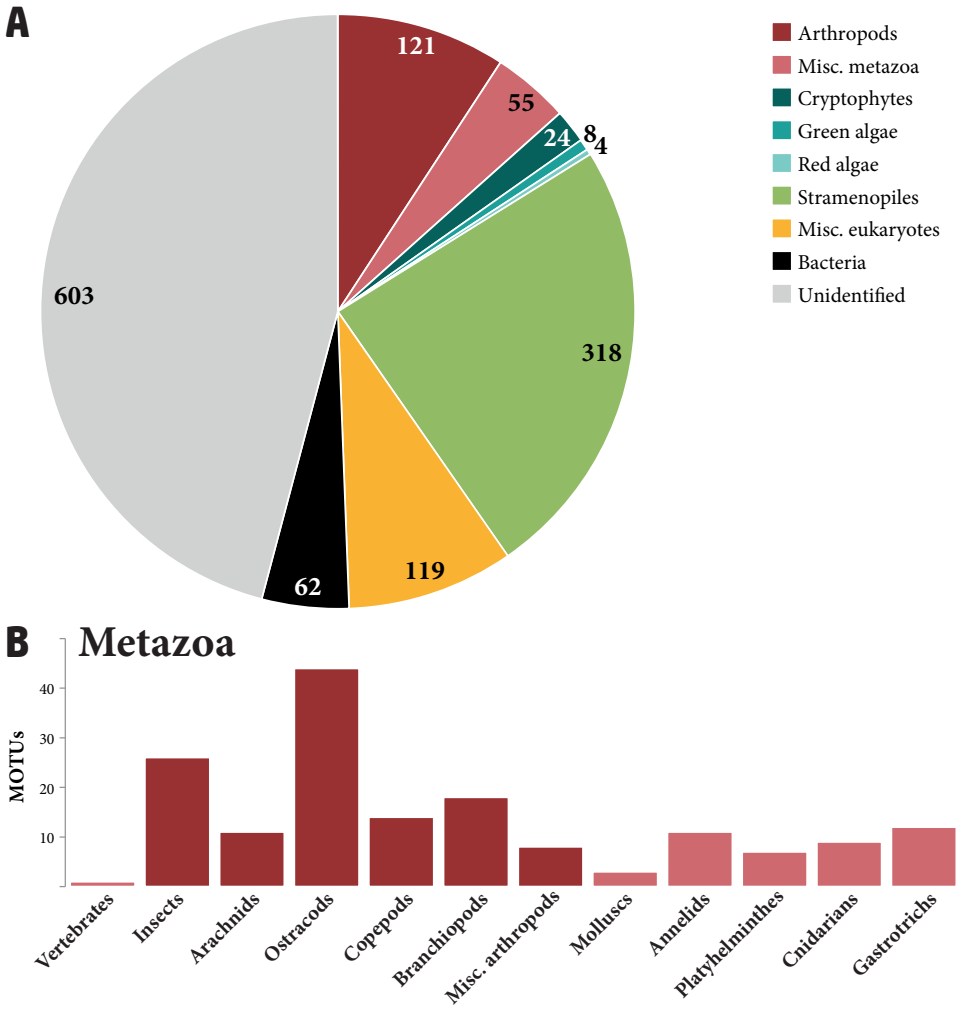


FIGURE 4.1. Taxonomic assignments of the MOTUs at (A) phylum-level and (B) class-level for metazoa, using a lowest common ancestor approach in MEGAN. Numbers in the pie chart indicate the number of MOTUs assigned to each phylum.

4.3.4 Sampling replicates

Average Sørenson dissimilarity between sampling replicates within one location at the same time point was 0.48 using the "additive" PCR replicate strategy (Figure 4.3), and significantly higher than dissimilarities between PCR replicates (t-test, $p = 0.005$). When using the "relaxed" and "strict" approaches, the average was slightly lower (0.45 and 0.46, respectively) (Table 4.1), but not significantly different (ANOVA). Four samples with only two successful PCR replicates were omitted from this analysis. There was a strong correlation between the Sørenson dissimilarity matrices for sample replicates under all three PCR replicate processing strategies (Supplementary Figure S4.1), both for the dissimilarities between sampling replicates pairs, and the dissimilarity matrix as a whole.

The high dissimilarity between sampling replicates was reflected in the contribution of each sampling replicate to the total diversity of the lakes at each time point. The three sampling replicates combined had an average of 187.3 ± 9.1 MOTUs, whereas a combination of two replicates only represented 81.0% (range 34.0–100%) of that

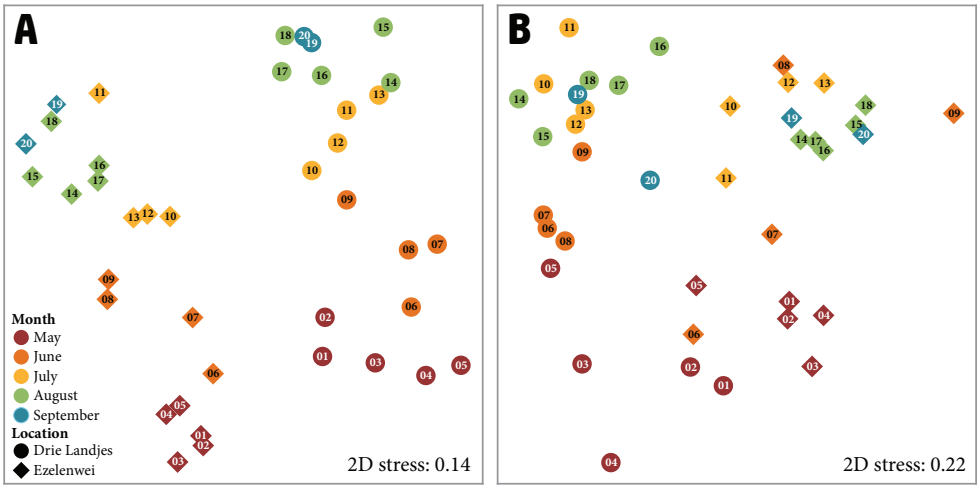


FIGURE 4.2. Two-dimensional NMDS plots constructed based on Sørenson dissimilarities between sampling sites using (A) all MOTUs and (B) only metazoan MOTUs. 2D stress values are displayed in the panels. Each point represents the combined community of the three spatial sampling replicates taken at each of the two locations on each of the 20 time points, with the PCR replicates combined using the "additive" strategy. Shapes indicate the location, colors are used to indicate the month in which samples were obtained, with numbers labeling the consecutive weeks from 2 May to 12 September. ANOSIM supported grouping of the samples belonging to one lake for both all MOTUs as the metazoan-only dataset ($R = 0.710$ and $R = 0.424$, respectively, $p = 0.001$). Seasonal grouping was similarly supported by ANOSIM, splitting samples into two seasonal groups (2 May–13 June, and 20 June–12 September) ($R = 0.486$ and $R = 0.587$, respectively, $p = 0.001$) for all MOTUs and metazoan-only.

total. In only one of 40 (two lakes, 20 time points) cases, the addition of a third sampling replicate did not provide additional MOTUs. One sampling replicate on average only produced 103.0 ± 3.9 MOTUs, which represented 55.4% of the total (range 12.1–92.5%). Regardless of the PCR replicate processing strategy used, the average proportion of MOTUs unique to one of three sample replicates was roughly the same (Table 4.1).

4.3.5 Temporal replicates

To look at the temporal patterns in the data, we used the "additive" PCR processing strategy, and added each of the three spatial replicates per week per location into one data point. This resulted in 40 data points with an average of 187.3 ± 9.1 MOTUs for 104,935 ± 5,007 reads. Again, there was no correlation between number of reads and number of MOTUs. A total of 257 (19.6%) MOTUs only ever occur in a single time point in a single location, only four MOTUs occur every week in both locations. Weekly samples represented between 9.5% and 37.9% (average 20.2%) of the total MOTU community observed in the lake, with later weeks generally having a higher

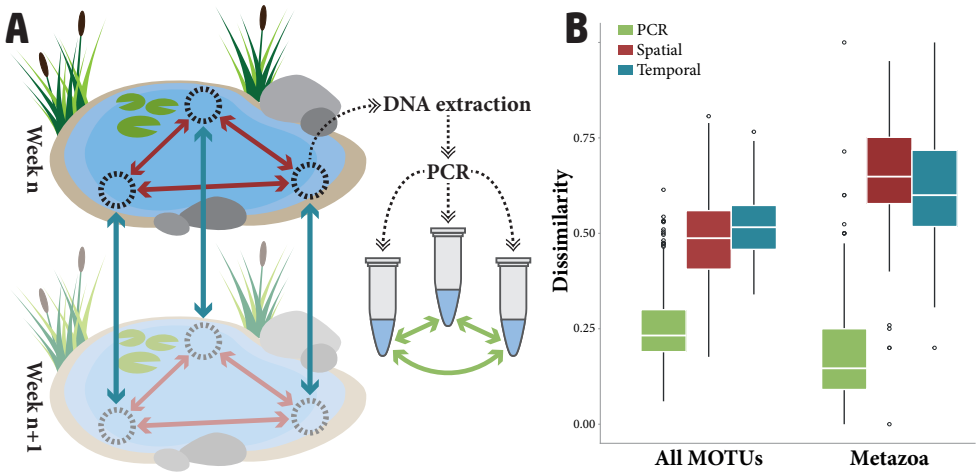


FIGURE 4.3. Average dissimilarities between PCR replicates, spatial replicates, and temporal replicates. (A) Schematic representation of the replicate sampling strategies and (B) boxplot displaying Sørenson dissimilarity values for PCR replicates (green, n = 352 for all MOTUs, n = 344 for metazoa), spatial sampling replicates (red, n = 104) and temporal replicates separated by 1 week (blue, n = 100) for both all MOTUs and metazoan-only MOTUs. In both cases, the dissimilarity between spatial replicates was significantly higher than between PCR replicates (t-test, p = 0.005). Only in the case of all MOTUs was the temporal dissimilarity significantly higher than the spatial dissimilarity (t-test, p = 0.005). There was no significant difference between spatial and temporal dissimilarities in the metazoan-only for samples taken 1 week apart.

TABLE 4.1. Total richness of all samples combined, as well as average (mean ± SEM) richness for each of the two locations at each of the 20 time points under different PCR replicate processing strategies ("additive," "relaxed," and "strict"), the effects on heterogeneity of MOTUs in the sample replicates and the average Sørenson dissimilarities between the sampling replicates (mean ± SEM). For each of the three strategies, the MOTUs are divided into three categories: (1) those MOTUs that are common and appear in all three sampling replicates; (2) MOTUs that are shared, and occur in two of three replicates; and (3) unique MOTUs, that only occur in a single sample replicate.

MOTUs	PCR strategy	Richness		Sample Replicates			
		Total	Average	Common (3/3)	Shared (2/3)	Unique (1/3)	Dissimilarity
All	**Additive** (1/3)	1314	187.3 ± 9.1	41.6 (22.2%)	37.5 (20.0%)	108.2 (57.8%)	0.48 ± 0.01
	Relaxed (2/3)	1034	114.8 ± 5.8	29.3 (25.5%)	23.8 (20.7%)	61.5 (53.8%)	0.46 ± 0.01
	Strict (3/3)	788	81.8 ± 4.4	17.6 (21.5%)	18.8 (22.9%)	45.5 (55.6%)	0.46 ± 0.01
Metazoan	**Additive** (1/3)	176	25.0 ± 2.4	18.0 (75.2%)	4.8 (17.5%)	2.3 (10.9%)	0.65 ± 0.01
	Relaxed (2/3)	156	19.2 ± 2.3	14.2 (72.2%)	3.8 (17.7%)	1.7 (10.1%)	0.66 ± 0.02
	Strict (3/3)	141	15.8 ± 2.2	12.0 (75.2%)	2.6 (17.2%)	1.1 (7.6%)	0.68 ± 0.02

richness than the earlier weeks. Turnover was not calculated as it was inflated by MOTUs occurring in non-consecutive weeks.

The average Sørenson dissimilarity between two replicates taken 1 week apart at the same sampling point was 0.53, which is significantly higher than the dissimilarity between two replicates taken at the same time (t-test, $p = 0.005$) (Figure 4.3). With the sampling replicates combined, the Sørenson dissimilarity between the total communities of one location a week apart was 0.48 on average. Looking at larger time intervals, there was a significant correlation between interval duration and Sørenson dissimilarity (Spearman correlation $\rho = 0.812$, $p < 0.001$) (Figure 4.4).

For the metazoan-only subset, dissimilarity between the sampling replicates and the temporal replicates was much higher than for the whole dataset, at 0.65 and 0.62, but with no significant difference between them (Figure 4.3). Temporal replicates were significantly more dissimilar than spatial replicates for intervals of three or more weeks (t-test, $p = 0.002$). The same effects as with all MOTUs were seen when looking at the PCR replicate processing strategies, where average dissimilarities were not significantly different for each of the three strategies, albeit much higher

than when using all MOTUs (Table 4.1). The correlation between interval duration and Sørenson dissimilarity was also significant for metazoan-only data (Spearman correlation $\rho = 0.555$, $p < 0.001$) (Figure 4.4).

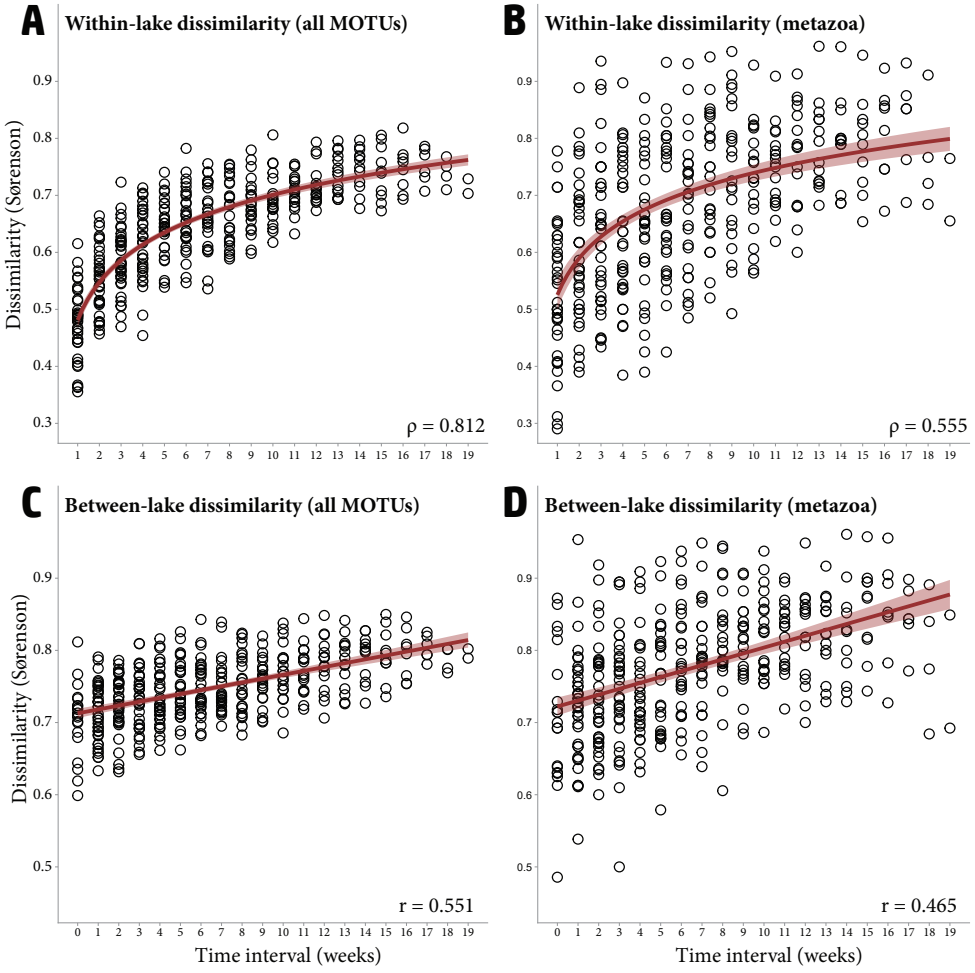


FIGURE 4.4. Time interval between two sampling moments vs. the Sørenson dissimilarity between total communities for samples taken in the same lake, with (A) all MOTUs and (B) only metazoan MOTUs (Spearman correlation, $p < 0.001$), and time interval between two sampling moments vs. the Sørenson dissimilarity between total communities for samples taken in different lakes, with (C) all MOTUs, and (D) only metazoan MOTUs (Pearson correlation, $p < 0.001$) (with 95% confidence interval). Correlation values are provided in the panels. Sampling replicates are merged into one sample per location per week, PCR replicates are processed using the "additive" strategy.

4.4 DISCUSSION

Our results demonstrate the relatively large differences that can exist between sampling replicates, both on a spatial and a temporal scale. A significant challenge in the use of eDNA for metabarcoding stems from the heterogeneity of eDNA within the environment, and also in DNA extracts. The latter introduces a stochastic effect when sequencing multiple PCR replicates, in which less abundant species may not be found in all replicates. We applied three ways of bioinformatics processing of PCR replicates: (1) using all MOTUs ("additive"), (2) only using MOTUs present in two or more replicates ("relaxed"), and (3) only using MOTUs present in all three replicates ("strict") (Alberdi et al. 2018). Whilst the chosen strategy had an impact on the total and average number of MOTUs found in each sample, general patterns of dissimilarities between samples were not largely impacted.

When we look at the heterogeneity of eDNA across the three sampling replicates within one location at a given time, the proportion of MOTUs that occur in either one or in all of the samples stays the same regardless of PCR replicate processing strategy. This indicates that removal of MOTUs not covered by all PCR replicates (the "strict" strategy) does not necessarily make spatial replicates more similar. This observation is confirmed by the average dissimilarity between the spatial samples, which is not significantly different for any of the three PCR replicate strategies (Table 4.1). Similarly, the Sørenson dissimilarity matrices were highly correlated ($r = 0.929$ and $r = 0.917$ for "additive" vs. "relaxed" and "relaxed" vs. "strict," respectively. Pearson correlation, $p < 0.001$) (Supplementary Figure S4.1). This suggests that the selected strategy can vary depending on the research question without significantly impacting observed patterns of biodiversity, although it affects the richness estimates. PCR results are not always reproducible, as witnessed by the average dissimilarity of 0.26 between PCR replicates in this study, but also as reported in the detection of rare species (Ficetola et al. 2008, Buxton et al. 2018). Especially when looking for rare species, multiple PCR replicates improve detection chances. For analyses that benefit from more complete taxa lists, such as those performed for WFD monitoring, the inclusion of multiple PCR replicates also seems beneficial. While we only took three sampling replicates within each lake in each week, others have suggested as much as nine samples to estimate biodiversity from eDNA (Grey et al. 2018).

Compared to PCR replicates, the Sørenson dissimilarity between spatial replicates (0.48 on average for the full dataset, 0.65 for the metazoan-only subset) is significantly higher (Figure 4.3), which reflects the heterogeneity of eDNA within the environment. Previous studies have already pointed out that eDNA signal can have

strong local effects (Moyer et al. 2014, O'Donnell et al. 2017, Stewart et al. 2017), due to limited dispersal and sedimentation, but also the rapid degradation of eDNA (Dejean et al. 2011, Barnes & Turner 2015). The use of spatial replicate sampling to retrieve eDNA results that are representative for the whole body of water has been stressed (Goldberg et al. 2016, Harper et al. 2019a), and shown to improve eDNA monitoring efficacy (Goldberg et al. 2018). Resampling at different time points, however, has received little attention. Up until now research into seasonal variation has often focused on a limited set of temporal samples, such as spring vs. autumn/winter (Chain et al. 2016, Guardiola et al. 2016, Lacoursière-Roussel et al. 2018). The effects of temporal replicate sampling in this study were comparable with those of spatial replicates, with dissimilarities between samples taken at one sampling point a week apart slightly but significantly higher than those between samples taken within one lake at a certain week (average 0.53 vs. 0.48) (Figure 4.3). Almost a fifth (19.6%) of MOTUs was only ever detected in a single time point. In the metazoan-only subset the spatial and temporal dissimilarities were higher than for the complete dataset (0.65 and 0.62, respectively), although not significantly different from each other. Temporal dissimilarity was significantly higher than spatial dissimilarity, however, for intervals of 3 weeks or more. Similar observations were made for example in fish (Stoeckle et al. 2017, Sigsgaard et al. 2017), where many species were only detected in a few time points, showing that temporal sampling regimes are needed for optimal recovery of the total biodiversity. Our sampling time frame coincides with the period in which most of the traditional WFD monitoring is performed, for which insights into within-season community changes are more relevant than between-season variations.

The data included a number of MOTUs occurring in non-consecutive weeks, suggesting these MOTUs went undetected, rather than being absent from the environment. A detection/non-detection cannot be directly translated into presence/absence (Roussel et al. 2015). These irregular patterns of occurrence may have increased the dissimilarity between replicate samples, both temporal and spatial. However, we observed a strong correlation between time interval and Sørenson dissimilarity (Spearman correlation, $\rho = 0.812$, $p < 0.001$) (Figure 4.4). Interestingly, it is not a linear correlation, and there seems to be a maximum to the dissimilarity between samples taken at different time points. Although we only sampled for 20 consecutive weeks, this data suggest that the community never changes completely within this time frame. The maximum observed Sørenson dissimilarity between two samples taken at one sampling point is 0.90 (for a 9 week interval). This indicates that, even though there are large changes in eDNA composition between different time points, there is some basal community that is present throughout the sampling period and

does not change. Such basal communities could be relevant for identifying potential novel targets for eDNA-based monitoring, as it would allow for a time-independent assessment. Planktonic crustaceans, such as the copepods and branchiopods found in relatively large numbers (both MOTUs and reads, Supplementary Figures S4.2 and S4.3) have the potential to be such new bio-indicators, as they may be more easily detected using eDNA and likely to respond quicker to environmental changes (Lim et al. 2016, Montagud et al. 2018). Additionally, we observed a linear increase in dissimilarity between the two locations over time (Pearson correlation, $r = 0.551$, $p < 0.001$). Average Sørenson dissimilarity of the two lakes was 0.71 when sampled in the same week (interval = 0), and increased up to 0.80 when sampled 19 weeks apart (Figure 4.4). This indicates that studies comparing communities between locations should be wary of the time intervals between sampling, as larger intervals between sampling may lead to inflated dissimilarities.

Even though there are large differences between communities along the temporal gradient, there were no large shifts in the taxonomic compositions defined by LCA (Supplementary Figures S4.2 and S4.3). Other than an increase in the number of metazoan taxa over time (both in absolute number of MOTUs and in proportion of the total diversity), the proportional contribution of each of the different taxonomic groups is roughly the same for all 20 weeks, in both lakes. This indicates that seasonal succession mostly occurs within the taxonomic groups. The increase in metazoan taxa may be slightly inflated in the data for location 2, where algae (two MOTUs classified as Chrysophyceae) dominated the reads between 30 May and 20 June, and potently out competed others in both DNA extraction and amplification. The rest of the weeks in location 2, and all weeks in location 1 were mostly dominated by arthropod (copepod and branchiopod) and unidentified reads (average of 36.0% and 48.7%, respectively).

The primers used in this study perform well on macroinvertebrate bulk samples, but are degenerate enough to amplify a wide range of non-target DNA from non-metazoan sources present in environmental samples that would normally not be found in bulk macroinvertebrate samples (Figure 4.1). In our case, only 13.4% of the MOTUs could be assigned to metazoan phyla. Within those, only about a third (51 out of 176) could be assigned to phyla that are actually counted as macroinvertebrates for the purpose of traditional quality monitoring under the WFD. The remainder of the metazoans were mainly branchiopods and copepods. Similar results with non-target taxa were reported in other papers using degenerate COI primers for freshwater community metabarcording (Weigand & Macher 2018). There has been some debate about the usability of the standard COI barcode region defined by Hebert

et al. (2003) within DNA- and eDNA-based analyses, but thus far the benefit of an extensive COI database seems to outweigh the drawbacks (Andújar et al. 2018b), as also witnessed by the many primer sets that have been designed for macroinvertebrate metabarcoding studies (Leray et al. 2013, Bista et al. 2017, Elbrecht & Leese 2017). The balance between universality of primers and target specificity is a delicate one, and metabarcoding "by-catch" can represent a significant share of the data. In our data, one fourth of the MOTUs were classified as stramenopiles and various algae groups. The COI barcode region may not be the optimal marker for all of these groups. Even in situations where not all MOTUs can be identified up to species level, unidentified (or partially identified, in the form of higher taxa) MOTUs can still be matched across different samples and may therefore still be of use for community analyses (Lim et al. 2016).

The primer sets used in this study may not have been optimal for recovery of all taxon groups, and group-specific primers may be more appropriate for the detection of novel bio-indicators. Nonetheless, we expect the temporal effects observed in this study to play a role in any community analysis. Even when eDNA is used for BQE monitoring, time intervals between sampling sites will likely remain, as it practically impossible to sample and process all sites within a short time frame. Seasonal effects have been reported in the rich history of publications based on morphological observation of seasonality in planktonic organisms (Gosselain et al. 1994, Wu et al. 2013), but molecular tools will allow for much finer resolution observations. We strongly encourage any research into the use of novel indicator taxa to take these temporal changes into consideration, as they clearly affect non-macroinvertebrate taxa such as the phyto- and zooplankton groups observed in this study.

4.5 CONCLUSIONS

We here present the first study that directly compares the effects of small-scale spatial and temporal resampling eDNA for metabarcoding. We show that replication leads to better estimations of total biodiversity, where the effects of spatiotemporal sampling replicates are significantly greater than PCR replications, even though the latter can already bring a substantial increase in richness depending on the replicate processing strategy. Interestingly, the PCR replicate handling strategy has little effect on patterns in biodiversity and dissimilarity between samples, and there are no severe drawbacks of including even those MOTUs that occur in only one replicate. Dissimilarities between temporally separated samples were approximately equivalent to the dissimilarities between spatially separated samples. These dissimilarities

increase over longer time intervals, suggesting that population turnover effects are stronger than community heterogeneity. This is an important consideration for any study comparing multiple communities that have been sampled at different time points, as well as any study that delves into the use of novel bio-indicators. Non-macroinvertebrate taxa, such as the phyto- and zooplankton groups observed in this study, are often put forward as potential bio-indicators. The effects of sampling strategies, especially short-term temporal replicate sampling, can have a considerate impact on the usability of these taxa.

4.6 ACKNOWLEDGEMENTS

We thank Staatsbosbeheer and Casper Zuyderduyn for access to the dune lakes and permits for field work. We also thank Dr. Oscar Vorst for his help in assessing potential sampling sites, Jeroen Visser for his assistance in the sample collection and laboratory work, and Elza Duijm for her assistance with NGS library preparations. This study funded by the Gieskes-Strijbis Fonds, as a part of the DNA Waterscan project. There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

4.7 DATA AVAILABILITY

Raw sequence data is available from the NCBI Sequence Read Archive (Bioproject accession PRJNA529573).

4.8 SUPPLEMENTARY MATERIALS

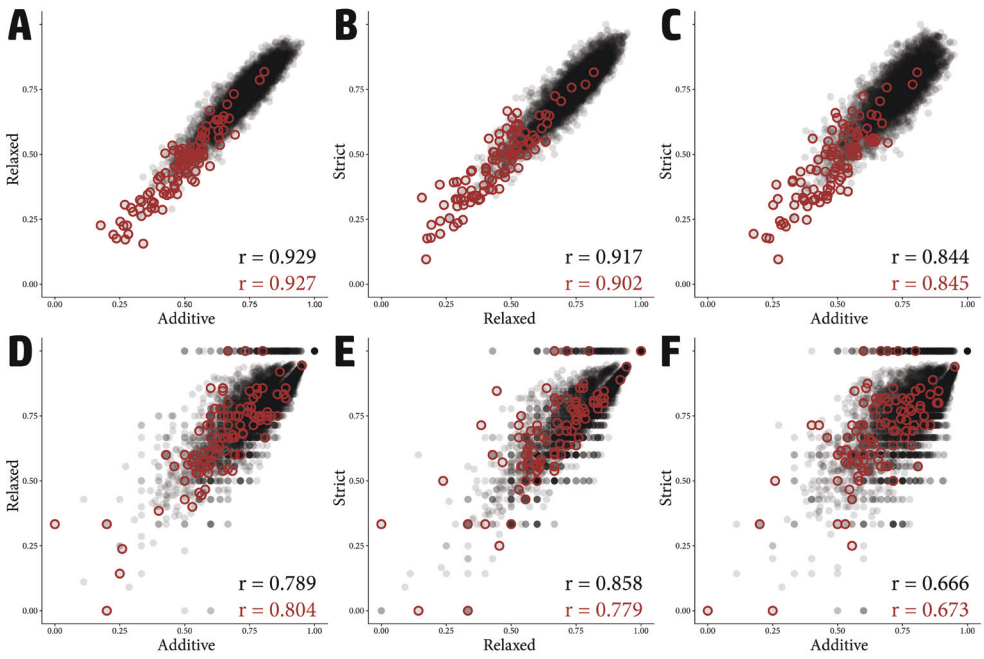
SUPPLEMENTARY TABLE S4.1. Sequences for primers used in the first and second round amplification. First round primers were modified to include additional bases between template-specific primer and the Illumina tail, to allow for demultiplexing on during data processing (highlighted in bold).

First Round		
Primer		Sequence (Universal tail – [**modification**] – template-specific primer)
BF1-ill1	Forward	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG [**ATGG**] ACWGGWTGRACWGTNTAYCC
BF1-ill2	Forward	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG [**CGT**] ACWGGWTGRACWGTNTAYCC
BF1-ill3	Forward	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG [**TC**] ACWGGWTGRACWGTNTAYCC
BF1-ill4	Forward	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG [**G**] ACWGGWTGRACWGTNTAYCC
BR2-ill1	Reverse	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG [**ATGGA**] TCDGGRTGNCCRAARAAYCA
BR2-ill2	Reverse	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG [**CGA**] TCDGGRTGNCCRAARAAYCA
BR2-ill3	Reverse	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG [**TC**] TCDGGRTGNCCRAARAAYCA
BR2-ill4	Reverse	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG [**G**] TCDGGRTGNCCRAARAAYCA
Second Round		
Primer		Sequence (Illumina adapter – index – universal tail)
NEX-F	Forward	AATGATACGGCGACCACCGAGATCTACAC [i5 index] TCGTCGGCAGCGTC
NEX-R	Reverse	CAAGCAGAAGACGGCATACGAGAT [i7 index] GTCTCGTGGGCTCGG

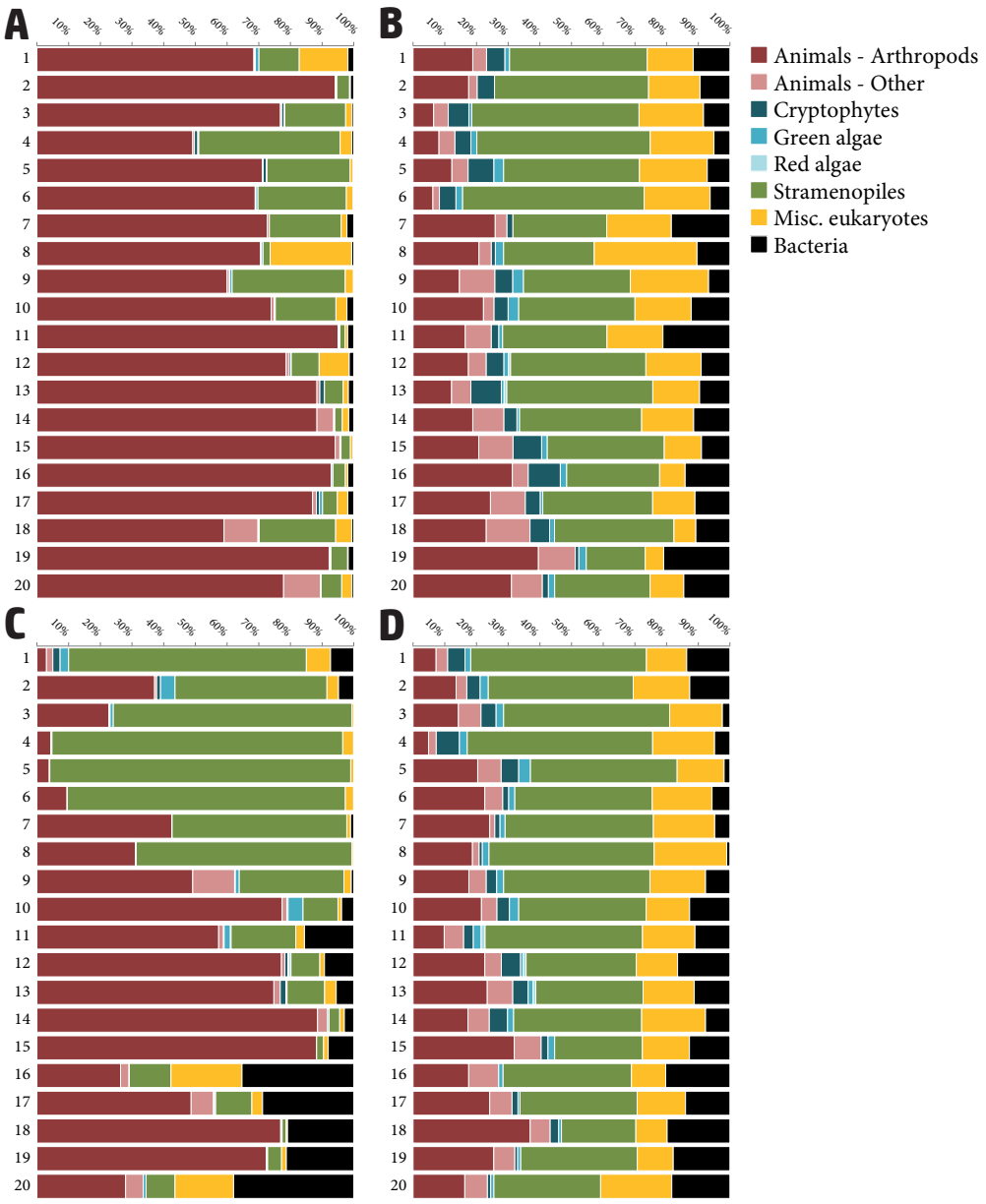
SUPPLEMENTARY TABLE S4.2. Primer combinations used for samples in first and second round PCR.

Samples	First round	Second round
Week 1-5 (2/5 to 30/5)	BF1-ill1 / BR2-ill1	Nextera XT, set C (N701-715 / S513-522)
Week 6-10 (6/6 to 4/7)	BF1-ill2 / BR2-ill2	Nextera XT, set C (N701-715 / S513-522)
Week 11-15 (11/7 to 8/8)	BF1-ill3 / BR2-ill3	Nextera XT, set C (N701-715 / S513-522)
Week 16-20 (15/8 to 12/9)	BF1-ill4 / BR2-ill4	Nextera XT, set C (N701-715 / S513-522)

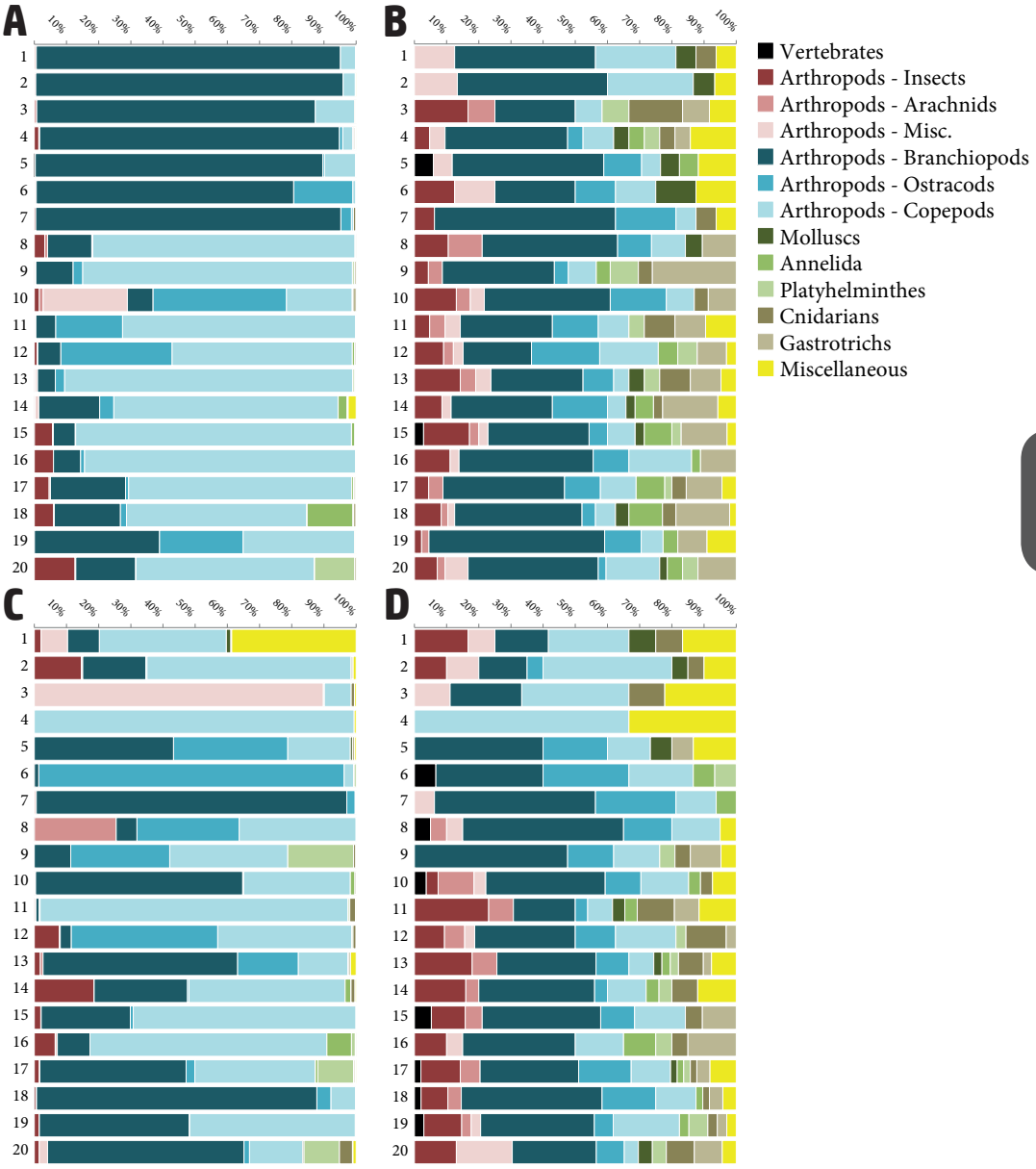
SUPPLEMENTARY FILE S4.1. Filtered MOTU table used for all subsequent analyses. Sample names include (in order): sampling date, sampling location (1-2), sub-sampling site (1-3) and PCR replicate number (A-C). **https://doi.org/10.7717/peerj.7335/supp-3**



SUPPLEMENTARY FIGURE S4.1. Comparison of Sørenson dissimilarities between samples, with pairs representing field sampling replicates highlighted in red, for three different PCR replicate processing strategies; (A) "additive" versus "relaxed," (B) "relaxed" versus "strict," and (C) "additive" versus "strict" for the Sørenson matrix based on all MOTUs, and (D) "additive" versus "relaxed," (E) "relaxed" versus "strict," and (F) "additive" versus "strict" for the Sørenson matrix based on only metazoan MOTUs. Sørenson dissimilarities were significantly correlated for each of the six comparisons, for both field replicates on their own and the whole Sørenson dissimilarity matrix (Pearson correlation, $p < 0.001$). Pearson correlation values are provided in the panels, in red the correlation values for the field replicates only.



SUPPLEMENTARY FIGURE S4.2. Total taxonomic composition of each of the lakes at each of the 20 sampling moments based on LCA identified MOTUs, for (A) relative read abundances and (B) MOTU diversity in lake 1, and (C) relative read abundances and (D) MOTU diversity in lake 2. Sampling replicates are combined, PCR replicates are processed using the "additive" strategy (including all MOTUs regardless of how many replicates they appeared in).



SUPPLEMENTARY FIGURE S4.3. Metazoan taxonomic composition of each of the lakes at each of the 20 sampling moments based on LCA identified MOTUs, for (A) relative read abundances and (B) MOTU diversity in lake 1, and (C) relative read abundances and (D) MOTU diversity in lake 2. Sampling replicates are combined, PCR replicates are processed using the "additive" strategy (including all MOTUs regardless of how many replicates they appeared in).

CHAPTER 5

Environmental DNA metabarcoding reveals comparable responses to agricultural stressors on different trophic levels of a freshwater community

Kevin K. Beentjes[1,2], S. Henrik Barmentlo[3,4], Ellen Cieraad[3], Menno Schilthuizen[1,2], Berry B. van der Hoorn[1], Arjen G.C.L. Speksnijder[1], Krijn B. Trimbos[3]

Preprint
https://doi.org/10.22541/au.159236833.30909538

[1] Naturalis Biodiversity Center, Leiden, The Netherlands
[2] Institute of Biology Leiden, Leiden University, Leiden, The Netherlands
[3] Institute of Environmental Sciences, Leiden University, Leiden, The Netherlands
[4] IBED, University of Amsterdam, Amsterdam, The Netherlands

ABSTRACT

Freshwater habitats are under stress from agricultural land use, most notably the influx of neonicotinoid pesticides and increased nutrient pressure from fertilizer. Traditional studies investigating the effects of stressors on freshwater systems are often limited to a narrow range of taxa, depending heavily on morphological expertise. Additionally, disentanglement of multiple simultaneous stressors can be difficult in field studies, whereas controlled laboratory conditions do not accurately reflect natural conditions and food webs. To overcome these drawbacks, we investigated the impacts of two agricultural stressors (the neonicotinoid insecticide thiacloprid and fertilizer) in full-factorial design in a semi-natural research site, using environmental DNA sampling to study three different taxonomic groups representing three trophic levels: bacteria (decomposers), phytoplankton (primary producers), and chironomids (consumers).

The results show considerably impact of both stressors across trophic levels, with an additive effect of fertilizer and thiacloprid on community composition at all levels. These findings suggest that agricultural stressors affect the entire food web, either directly or through cascade reactions. They are also consistent with morphological assessments that were performed in the same study site, even at a lower number of replicates. The study presented shows that the use of multi-marker environmental DNA provides a more comprehensive assessment of stressor impacts across multiple trophic levels, at a higher taxonomic resolution than traditional surveys. Additionally, over a thousand putative novel bio-indicators for both agricultural stressors were discovered. We encourage further investigations into stressors impacts at different trophic levels, which will lead to more effective monitoring and management of freshwater systems.

5.1 INTRODUCTION

Freshwater ecosystems contain a rich diversity of both taxa and microhabitats, despite the fact that they cover less than one percent of the Earth's surface. They are disproportionately affected by anthropogenic impacts, and seem to be under greater threat than terrestrial and marine systems (Dudgeon et al. 2006, WWF 2014). Effective monitoring of biological quality of freshwater systems is essential for timely interventions, especially since freshwater is not only important for the management of aquatic flora and fauna, but also for the 'ecosystem services' that are essential to people's well-being and health (Corvalan et al. 2005).

One of the most important stressors to freshwater systems is agricultural land use as many freshwater habitats are directly connected to agricultural land. Next to the removal and fragmentation of habitat, pesticide and fertilizer use are the most prominent stressors here (Matson et al. 1997, Schreiner et al. 2016). While pesticides are used on agricultural land to prevent crop losses by pests, they may enter adjacent freshwater through spray drift, run-off, and seepage. The widespread use of neonicotinoid insecticides in agriculture has been subject of debate as they are found to impact non-target species, including many freshwater invertebrate species (Pisa et al. 2014, Morrissey et al. 2015, Raby et al. 2018), and have the potential to disrupt the entire food web (Yamamuro et al. 2019). Research has shown that neonicotinoid insecticides can negatively impact macroinvertebrate communities and have significant effects on food web structuring since invertebrates are critical in the transfer of nutrients from the primary producers to the consumers at the top of the food chain (Van Dijk et al. 2013, Chagnon et al. 2015, Schrama et al. 2017). The effects of neonicotinoids and the interaction with other common stressors such as increased influx of nutrients or fine sediments have been studied via morphological assessments in model systems (Barmentlo et al. 2019, Chará-Serna et al. 2019), showing alternative impacts of neonicotinoids to macroinvertebrate communities in combination with other stressors.

Traditional morphological surveys, such as employed in the above-mentioned studies, have several drawbacks which have implications on the quality and quantity of data that is collected. Morphological assessments of macroinvertebrate communities rely on skilled taxonomists, may be biased between assessors (Haase et al. 2010) and are labor-intensive and therefore often expensive (Jones 2008). The costs specifically affect decisions made on sampling frequency and intensity, and the time-consuming nature can cause delays that prevent timely interventions into impacted systems (Keeley et al. 2018). Additionally, traditional morphological surveys are limited in

accurately assessing many taxa that are likely to be affected by stressors, such as bacteria or planktonic organisms. Tools used to assess impact of pollutants on the aquatic ecosystem thus need to be refined (Schwarzenbach et al. 2006).

In the last decade, molecular tools, including environmental DNA metabarcoding, have become more common place for detecting and identifying taxa. Environmental DNA (eDNA) refers to any DNA collected from the environment without specifically collecting or isolating target specimens (Taberlet et al. 2012b). Community assessments using eDNA from soil have been standard practice for microbiologists for some time, but only more recently has this tool become one of the standard approaches for surveying freshwater biota, especially fish (e.g. Hänfling et al., 2016; Shaw et al., 2016). The use of eDNA has also found its way into environmental impact studies, such as studies on the impact of aquaculture on benthic sediments (Pochon et al. 2015, Stoeck et al. 2018). eDNA enables the detection of other, potentially more informative, organism groups than those studied in traditional impact studies (Macher et al. 2018). The use of eDNA allows for the defining of new indicators to stressors (e.g. Chariton et al., 2014; Li et al., 2018), and metabarcoding techniques can lead to the creation of new MOTU-based biotic indices (Apothéloz-Perret-Gentil et al. 2017). Despite their potential, most eDNA-based impact assessments still focus on one or few taxonomic groups, and only recently have multi-marker approaches been introduced to evaluate different taxonomic groups simultaneously (Andújar et al. 2018a, Keeley et al. 2018, Laroche et al. 2018, Li et al. 2018b, Cordier et al. 2019).

Impact assessments are often performed directly in the field, where the myriad of simultaneous stressors make it difficult to identify the impact of individual stressors (Piggott et al. 2015, Côté et al. 2016). Multi-trophic (eDNA) approaches have proven to provide stronger correlations with environmental variables than approaches that use a single guild (Keeley et al. 2018), but the possibility that different guilds respond differently to stressors make interpretation of novel multi-trophic eDNA approaches in natural settings difficult. Due to a lack of multi-trophic impact assessment studies where results gathered using eDNA and traditional approaches are combined, it remains unclear to what extent eDNA-based assessments can accurately detect the impacts in such complex environments.

In this study, we assess the impact of two main agricultural stressors on multiple trophic levels in naturally colonized freshwater communities in outdoor experimental ditches. In a full factorial setup, we use eDNA to assess the single and combined impacts of fertilizer and pesticide (the neonicotinoid thiacloprid) application on the richness, taxonomic composition and community dissimilarity of three trophic levels: bacteria, representing decomposers; phytoplankton, representing primary

producers; and chironomids, as representatives of the primary and secondary consumers, as well as a traditional indicator group for water quality. Using eDNA in this experimental impact assessment allows us to achieve the following aims: (1) to assess multi-trophic impacts on taxon groups that may be sensitive to stressors, but are not traditionally used in freshwater impact assessments due to their difficulty in identification, using novel multi-marker eDNA approaches; (2) to assess the impact of two agrochemicals on freshwater communities, while also being able to compare results with a concomitant traditional morphology-based impact study (Barmentlo et al. 2019); and (3) to pinpoint potential new bio-indicators for the health of freshwater ecosystems.

5.2 MATERIALS AND METHODS

5.2.1 Experimental setup

Environmental DNA sampling was performed in 20 experimental ditches located in the outdoor research facility the 'Living Lab' (see Barmentlo et al. (2019) for a detailed description of the site and treatments). Prior to the experiments, ditches were left connected to the adjacent reservoir for six months to allow for natural colonization of freshwater communities in the ditches. Before starting the experiment, ditches were hydrologically closed off using acrylic plates to avoid cross-contamination between treatments and to isolate the ditches from the reservoir. Subsequently, the ditches were exposed to two different agrochemical stressors in a full factorial design (five ditches per treatment): (1) control, with no added substances; (2) addition of the insecticide thiacloprid (Sigma-Aldrich, Zwijndrecht, The Netherlands) in two spikes (week 20 and 22) with a nominal time weighted average concentration for one month of 0.4 µg/l; (3) addition of nutrients in the form of three sachets with 75g of slow-releasing artificial fertilizer granulates ('Osmocote'; N:P:K = 15:9:11 combined with microelements) per ditch that were replaced every six weeks; and (4) a combination of thiacloprid and fertilizer in the same concentrations and application as described for the single-treatment ditches.

5.2.2 Sampling and DNA extraction

Environmental DNA sampling was performed in five replicate ditches for each treatment (20 in total) at four time points: two weeks prior to the start of the treatment (May 1st, 2017; week 18), and two weeks (May 31st, 2017; week 22), four weeks (June 13th, 2017; week 24) and seven weeks (July 6th, 2017; week 27) after the start of the treatments. Surface water samples were collected in the morning from the center of

each ditch using sterilized bottles and filtered within two hours in the laboratory. Filtration was performed using 0.2 µm polyethersulfone (PES) filter membranes (Sartorius, Göttingen, Germany) placed in sterilized Nalgene filter units (Thermo Fisher, Waltham, MA, USA) attached to a vacuum pump. Up to 300 ml of water was filtered for each of the 20 ditches. A modified CTAB extraction protocol adapted from Turner et al. (2014) was used for DNA extraction (Chapter 4).

5.2.3 DNA amplification and MiSeq sequencing

Three different markers for three different taxa groups were analyzed separately, using group-specific primers: a ±400 bp fragment of 18S rRNA V4 subregion for phytoplankton (Zimmermann et al. 2011), a 273 bp fragment of the 16S rRNA for bacteria (Klindworth et al. 2013) and a 235 bp fragment of COI for chironomids (Bista et al. 2017) (for primers, see Supplemental Table S5.1). For each of the PCRs, all of the 80 reactions for each marker (20 replicate ditches, 4 time points) were performed in duplicate. The chironomid PCR contained two samples of DNA extracted from two chironomid specimens unlikely to occur in the setup were used as a contamination control. This control was used to estimate cross-contamination between samples during the amplification and correct MOTU tables of all three markers accordingly, using a tool based on Larrson et al. (2018). Cross-contamination was assumed to be the same for all three markers.

Dual-indexed Illumina amplicon libraries were prepared using a two-step PCR protocol, in which the first PCR used primers with 5' Illumina tails. Initial PCRs were performed in 25 µl reactions containing 1x Phire Green Reaction Buffer, 0.5 mM dNTPs, 0.5 µl Phire Hot Start II DNA Polymerase (Thermo Fisher, Waltham, MA, US), 0.5 µM of each primer and 2.0 µl of template DNA. Initial denaturation was performed at 98°C for 30 seconds, followed by 35 cycles at 98°C for 5 seconds, 50°C for 5 seconds and 72°C for 15 seconds, followed by final elongation at 72°C for 5 minutes. PCR products were checked on E-Gel 96 pre-cast agarose gel (Thermo Fisher, Waltham, MA, USA) and cleaned with a one-sided size selection using NucleoMag NGS-Beads (Macherey-Nagel, Düren, Germany), in a 1:0.9 ratio. Dual-index PCRs were performed using 2.0 µl of PCR product from the first round in a 20 µl reaction containing 1x TaqMan Environmental Master Mix 2.0 (Thermo Fisher, Waltham, MA, USA) and 1.0 µM of each primer. Initial denaturation was performed at 95°C for 10 minutes, followed by 10 cycles at 95°C for 30 seconds, 55°C for 60 seconds and 72°C for 30 seconds, followed by final elongation at 72°C for 7 minutes. These PCR products were quantified on the QIAxcel (Qiagen, Venlo, the Netherlands) and each replicate of each marker was pooled equimolarly separately. Pools were cleaned with

a one-sided size selection using NucleoMag NGS-Beads, ratio 1:0.9, then quantified on the Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) with the DNA High Sensitivity Kit. The pools for the bacteria and chironomids were combined equimolarly and sequenced on one run of Illumina MiSeq (v3 Kit, 2x300 paired-end), the pools for the phytoplankton were combined equimolarly and sequenced on a separate run, both at BaseClear BV (Leiden, the Netherlands).

5.2.4 Bioinformatics

Quality filtering and clustering of all data was performed in a custom pipeline on the OpenStack environment of Naturalis Biodiversity Center through a Galaxy instance (Afgan et al. 2018). Raw data was filtered with Sickle (Joshi & Fass 2011) and merged with FLASH v1.2.11 (Magoč & Salzberg 2011), non-merged reads were discarded. Primers were trimmed from both ends using Cutadapt v1.16 (Martin 2011) and any read without both primers present and anchored was discarded. PRINSEQ v0.20.4 (Schmieder & Edwards 2011) was used to filter reads based on length (390-420 bp for phytoplankton, 248-254 bp for bacteria, 230-250 bp for chironomids). Sequences were dereplicated and clustered into Molecular Operational Taxonomic Unit (MOTUs) using VSEARCH v2.10.3 (Rognes et al. 2016) with a cluster identity of 98% and a minimal accepted abundance of 2. MOTU tables were corrected using the occurrence of control chironomids in field samples (rate of spread 0.003, cutoff value 5 reads). PCR replicates were combined, including all MOTUs that were present in at least one replicate.

MOTU sequences were compared to custom reference databases using BLAST+ (Camacho et al. 2009). Phytoplankton MOTUs were compared to a dataset that included all 18S rRNA sequences from GenBank (Benson et al. 2005) (sequences downloaded 21 August 2018), bacteria were compared to Silva SSUParc 132 (Quast et al. 2013), chironomids were compared to a custom reference (Chapter 3) based on specimens collected in the Netherlands as part of a national DNA barcoding campaign (Beentjes et al. 2015), supplemented with sequences obtained from BOLD (Ratnasingham & Hebert 2007).

5.2.5 Taxonomic assignment and diversity analysis

A 98% cutoff was used for species-level identification, and a custom lowest common ancestor (LCA) script (Chapter 3) was used to identify MOTUs in those cases where no direct hits above 98% with the reference database were found. LCA was performed on the top 10% hits with bitscore >170, a minimum identity of 85% and a minimum coverage of 90% (90% identity and 100% coverage for the bacteria). The LCA was set to identify MOTUs no further than genus level. Normalized read abundances

were used in the analyses, based on the assumption that initial communities were all similar in terms of species composition and abundances. Differences in relative abundances in time, relative to the control samples, are assumed to be caused by the treatments (Beermann et al. 2018, Barmentlo et al. 2019).

5.2.6 Statistical analyses

Potential effects of the agrochemicals and time were assessed on the three different communities (bacteria, phytoplankton and chironomids). The effects of both fertilizer, thiacloprid, time, and all possible interactions were investigated on the normalized MOTU abundances using permutational analysis of variance (PERMANOVA, function adonis, R package VEGAN). Bray-Curtis was used as measure for dissimilarity, with 999 permutations. We accounted for the repeated measure design by including ditch number as a random variable. Differences in richness were analyzed with ANOVA (R package STATS). Potential effects on beta dispersion were investigated by using distance-based dispersion tests (function betadisper, R package VEGAN). Correlation between the distance matrices for the three communities analyzed in this study and the morphological assessment was investigated using a Mantel test (function mantel.rtest, R package ADE4, 999 permutations). Indicative MOTUs for each of the treatments independently were identified using the multipatt function (R package INDICSPECIES).

Morphological assessment in the original study by Barmentlo et al. (2019) was performed at three moments: before treatment, one month after treatment (June) and four months after treatment (September). The assessment in June was performed at the same time as the measurement four weeks after treatment start presented in this paper. Data from the morphological assessment in June was compared directly to eDNA results from the same week.

5.3 RESULTS

5.3.1 Sequence run statistics and taxonomic assignment

After merging, filtering and clustering, and with the correction for cross-contamination applied, the replicates combined and non-target MOTUs omitted, there was a total of 5,383 MOTUs for bacteria, 2,819 for phytoplankton and 692 for chironomids. The bacteria data contained 4,011 MOTUs (74.5% of total MOTUs) that could be identified at least at phylum level, with the largest groups being Gammaproteobacteria (30.1%) and Bacteroidetes (20.9%). In the phytoplankton data, 1,773 MOTUs (62.9% of total MOTUs) could be identified to at least phylum

level of relevant taxa, mostly Chlorophyta (45.4%) and Stramenopiles (34.5%). For the chironomid dataset, 368 MOTUs (53.2% of total MOTUs) could be identified as Chironomidae on genus or species level, representing 64 species from 35 genera; 207 MOTUs were only identified up to genus level. One sample (a sample from a ditch with a fertilizer treatment from week 18) did not contain any chironomid reads. The morphological study by Barmentlo et. al (2019) confirmed the presence of chironomids in this ditch, proving this a false negative; the sample was therefore omitted from the analyses presented here.

5.3.2 Effects on MOTU richness

Richness changed significantly over time irrespective of treatment for all three investigated communities, following a similar pattern for all three, with a peak in richness in week 22 (Figure 5.1). Looking at the different weeks separately, there was no significant effect of any treatment on the MOTU richness for phytoplankton or chironomids. For bacteria, the richness observed in ditches with combined treatment of fertilizer and thiacloprid was significantly higher than the richness observed with addition of only thiacloprid ($p = 0.003$) or the addition of only fertilizer ($p = 0.013$), but not higher than the richness observed in control ditches, and only in week 22 (two weeks after application of treatments). There was no significant difference in the number of reads between treatments for each week.

5.3.3 Effects on community dissimilarity

Before the application of any agrochemical, there were no statistically significant differences between community species compositions of the prospective treatments. After application of the agrochemicals, fertilizer and thiacloprid addition showed a significant interaction, irrespective of time, leading to dissimilar communities relative to the control for all three communities ($p = 0.001$ for all comparisons) (Table 5.1). The interaction between thiacloprid and fertilizer was most pronounced in the weeks directly after application of thiacloprid, where the dissimilarity between control ditches and ditches treated with both agrochemicals was higher than dissimilarities between control ditches and ditches treated with only thiacloprid or fertilizer (Figure 5.2). Two weeks after the start of the treatments, the impact of thiacloprid addition was more pronounced than the addition of fertilizer, with the former having a significant impact on the dissimilarity in all groups ($p = 0.001$, Table 5.2), while the impact of fertilizer was only significant for bacteria and phytoplankton ($p = 0.021$ and 0.001, respectively). Thiacloprid centroids were more distant from the control than the fertilizer centroids for all three groups in week 22 and 24 (two

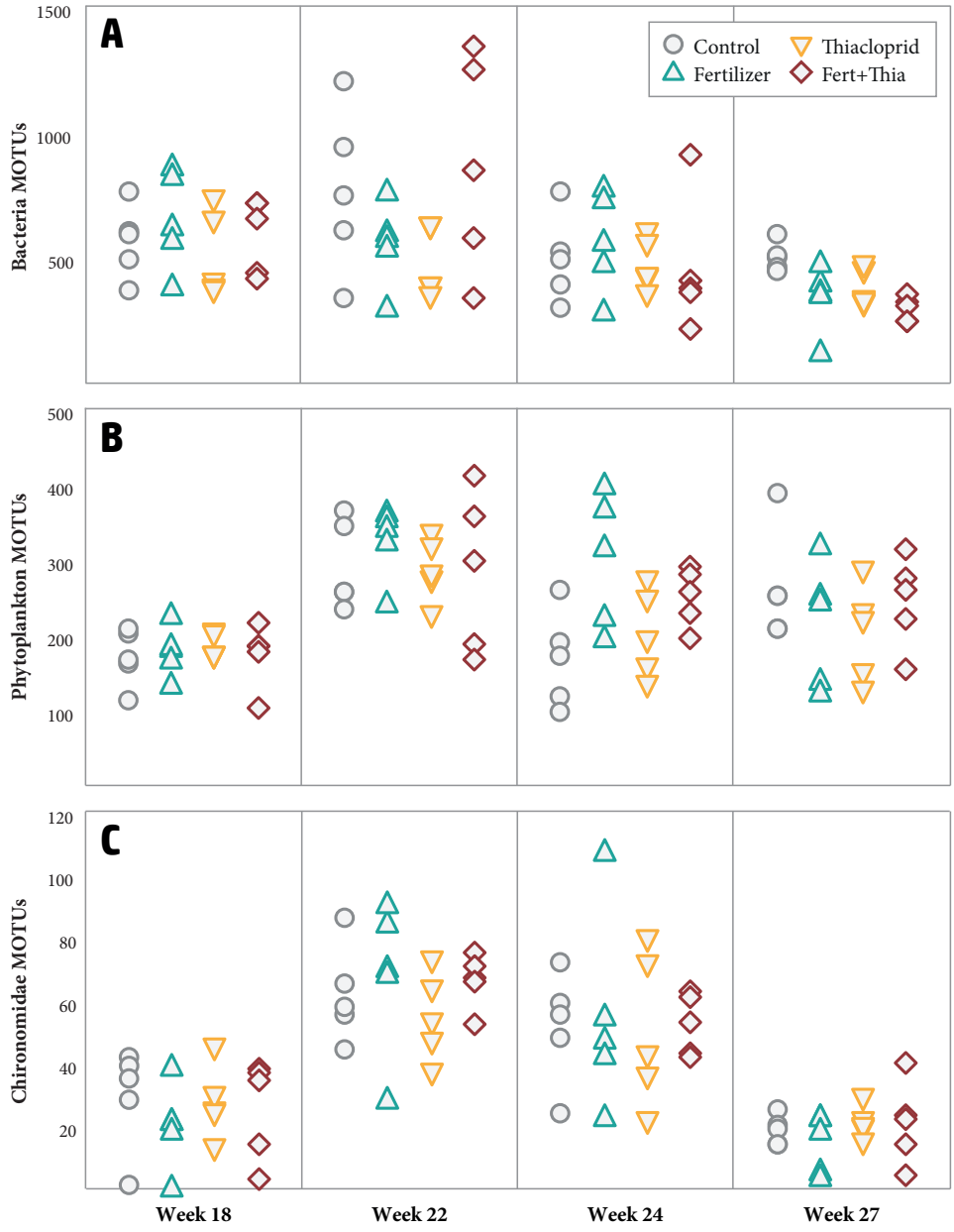


FIGURE 5.1. Observed number of MOTU for each of the taxonomic groups: (A) bacteria, (B) phytoplankton, and (C) chironomids, in control situation , and with added fertilizer, thiacloprid, and combined treatments.

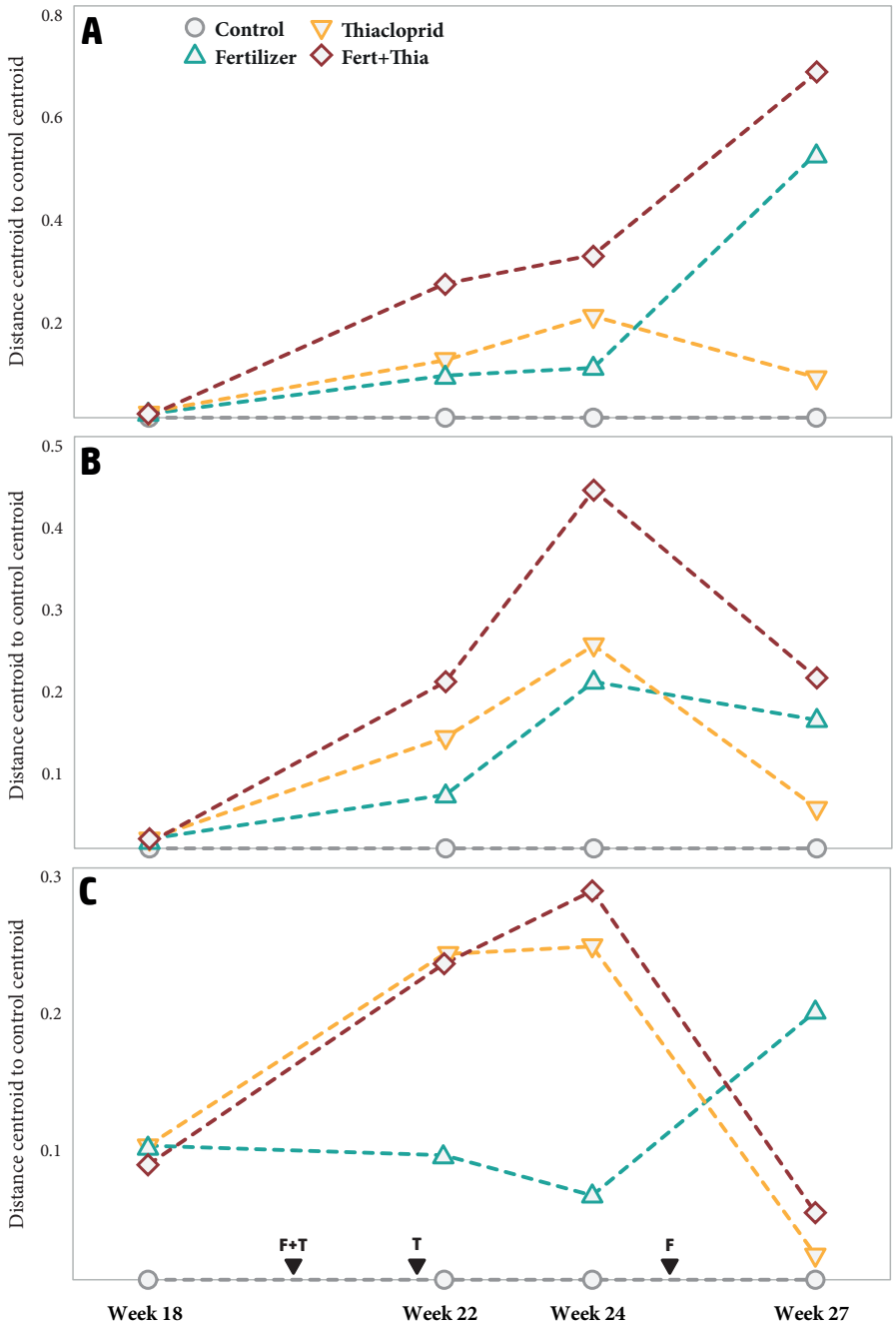


FIGURE 5.2. Average distance from centroid to the control centroid, for each of the taxonomic groups: (A) bacteria, (B) phytoplankton, and (C) chironomids. Moments of treatment application for thiacloprid (T) and fertilizer (F) are provided on the x-axis of panel C.

and four weeks after treatment), indicating that thiacloprid had a greater effect on the community composition in the short-term than fertilizer, albeit much more pronounced for chironomids. This reversed after the addition of fresh fertilizer pellets in week 25 as the effects of fertilizer became more pronounced compared to those of the thiacloprid addition (Figure 5.2). There was one sample in the control ditches prior to application of treatments, where we found only a single MOTU that was identified as a chironomid. This formed an outlier in the analysis of the chironomid data (Supplemental Figure S5.2C), and caused the centroid of the control samples in this measurement (week 18) to shift relative to the centroids of the other sets of ditches, explaining why the distances between centroids in week 18 were already elevated prior to start of treatment (Figure 5.2C).

The effect of time on dissimilarity was prominent, being larger than most effects of the agrochemicals, indicating that continued species turnover occurred. Two-way interactions of time with both fertilizer and thiacloprid were significant for all three communities studied (Table 5.1). There were no significant three-way interactions for any of the three groups, indicating that the interaction between the effects of fertilizer and thiacloprid occurred irrespective of the time point sampled. Studying the individual weeks separately, there was a significant effect of thiacloprid addition on community dissimilarity compared to control ditches in all three groups in week 22 and 24. Fertilizer had a significant effect on the composition of phytoplankton and bacteria in all three weeks after the start of the treatments (Table 5.2).

Beta-dispersion was significantly higher in treatments containing fertilizer for both bacteria and phytoplankton ($p < 0.001$ and $p = 0.005$, respectively), meaning that communities diverged when fertilizer was added to the system. Thiacloprid addition had a significant effect on chironomids, leading to convergence of the communities across the replicate ditches ($p = 0.002$) (Supplemental Figure S5.1). There were moderate, but significant correlations between all three Bray Curtis distance matrices of the three taxon groups. The correlation between bacteria and phytoplankton was stronger (Pearson $r = 0.820$, $p = 0.001$) than correlations of bacteria and phytoplankton with chironomid data ($r = 0.447$ and $r = 0.465$, respectively, $p = 0.001$). This indicates that community dissimilarities caused comparable patterns for both bacteria and phytoplankton (Supplemental Figure S5.2).

5.3.4 Effects on taxonomic composition

While the treatments had no apparent effect on the observed richness compared to control ditches, there were considerable shifts in the relative abundance of different taxa for all three communities analyzed in this study (Figure 5.3).

TABLE 5.1. PERMANOVA results (F-statistic, R2 and p-values) for the different treatments and the combined effects, including the three-way interaction with time, for data from all measurements combined. Significant p-values are marked with an asterisk (*).

	Bacteria			Phytoplankton			Chironomidae		
	F	R2	p-value	F	R2	p-value	F	R2	p-value
Thiacloprid	6.823	0.034	**0.001***	2.605	0.017	**0.001***	4.068	0.038	**0.001***
Fertilizer	12.329	0.061	**0.001***	6.751	0.044	**0.001***	1.192	0.011	**0.001***
Time	29.331	0.436	**0.001***	19.850	0.387	**0.001***	8.950	0.254	**0.001***
Fert:Thia	2.170	0.011	**0.001***	1.893	0.012	**0.001***	1.017	0.010	**0.001***
Thia:Time	3.100	0.046	**0.002***	2.115	0.041	**0.016***	1.531	0.043	**0.041***
Fert:Time	4.785	0.071	**0.001***	2.911	0.057	**0.002***	0.872	0.025	0.695
Fert:Thia:Time	1.543	0.023	0.195	1.287	0.025	0.269	0.846	0.024	0.755

TABLE 5.2. PERMANOVA results (F, R2 and p-values) for the different treatments for each of the time points evaluated separately. Significant p-values are marked with an asterisk (*).

		Bacteria			Phytoplankton			Chironomidae		
		F	R2	p-value	F	R2	p-value	F	R2	p-value
Week 18	**Thiacloprid**	0.675	0.036	0.809	0.502	0.028	0.711	1.032	0.060	0.352
	Fertilizer	1.329	0.071	0.201	1.190	0.066	0.287	0.316	0.018	0.990
	Fert:Thia	0.728	0.039	0.757	0.392	0.022	0.803	0.964	0.056	0.392
Week 22	**Thiacloprid**	6.428	0.234	**0.001***	3.520	0.144	**0.001***	4.087	0.181	**0.001***
	Fertilizer	3.044	0.111	**0.021***	3.553	0.145	**0.001***	1.460	0.064	0.152
	Fert:Thia	2.005	0.073	0.069	1.348	0.055	0.179	1.093	0.048	0.335
Week 24	**Thiacloprid**	4.318	0.167	**0.001***	3.007	0.117	**0.004***	3.006	0.145	**0.001***
	Fertilizer	3.806	0.147	**0.003***	4.210	0.164	**0.001***	0.857	0.041	0.662
	Fert:Thia	1.740	0.067	0.114	2.514	0.098	**0.010***	0.860	0.042	0.636
Week 27	**Thiacloprid**	2.027	0.061	0.070	0.892	0.041	0.550	1.269	0.066	0.173
	Fertilizer	13.598	0.410	**0.001***	4.119	0.189	**0.001***	1.121	0.059	0.311
	Fert:Thia	1.514	0.046	0.149	0.750	0.034	0.764	0.728	0.038	0.782

There were noticeable changes in the relative abundances of the phyla of bacteria within the different treatments (Figure 5.3A). Bacteriodetes were more abundant in ditches with added thiacloprid (representing 48.5% of the total reads versus 30.3% in the control ditches), mainly at the expense of Proteobacteria (21.4% versus 27.3%) and Verrucomicrobia (1.2% versus 7.1%). Actinobacteria represented a larger proportion of the reads in ditches without added fertilizer (33.4% of the total reads in control ditches versus 23.5% in the fertilizer ditches), a trend that continued into week 27, where the difference was 28.8% versus 3.6% on average. The relative abundances of Actinobacteria were also lower under the addition of thiacloprid to the point where they were nearly absent (0.8% of the total reads in combined agrochemical ditches compared to 32.2% in the control) in the combined treatment in week 27. In weeks 24 and 27, the relative composition at the phylum level seemed more affected by the addition of fertilizer than by thiacloprid, and both Bacteroidetes and Alphaproteobacteria became more abundant within the fertilizer treatment relative to the control (49.3% versus 29.5% and 12.1% versus 2.8%, respectively for both groups).

The phytoplankton community compositions changed considerably as well, and were significantly affected by both fertilizer and thiacloprid (Table 5.2). The read distribution (Figure 5.3B) reflected these changes as well. Effects were subtle in week 22 (two weeks after application of treatments), with the thiacloprid treatment showing higher proportions of chrysophyte reads (average of 51.3% versus 25.8% in control ditches), mostly at the cost of diatoms (3.7% versus 19.6%) and chlorophytes (35.3% versus 47.3%). In week 24, the composition changed considerably, showing a shift towards a system that was dominated by chlorophytes in ditches with added thiacloprid (76.3% versus 36.0% in control ditches), now at the expense of chrysophytes (5.5% versus 25.9%) and other stramenopiles (12.6% versus 20.3%). Cryptophytes were detected with much higher relative read abundances in ditches without added fertilizer (16.2% in control ditches versus 1.5% with fertilizer added), and went almost undetected in the ditches with a combined treatment, representing only 0.015% of the reads. The addition of fresh fertilizer pellets, in week 25, again changed the composition, bringing about a large shift in communities for ditches that received fertilizer. These were dominated by various groups of chlorophytes (60.6% versus 24.9% in control ditches), whereas control and thiacloprid-only ditches were dominated by the various stramenopile groups (60.0% versus 27.0% in ditches with fertilzer), most notably chrysophytes (29.0% versus 10.6%) and diatoms (17.0% versus 1.9%). Cryptophytes represented 11.9% of the reads in treatments without fertilizer, but went nearly undetected in the treatments with fertilizer (0.4% of the

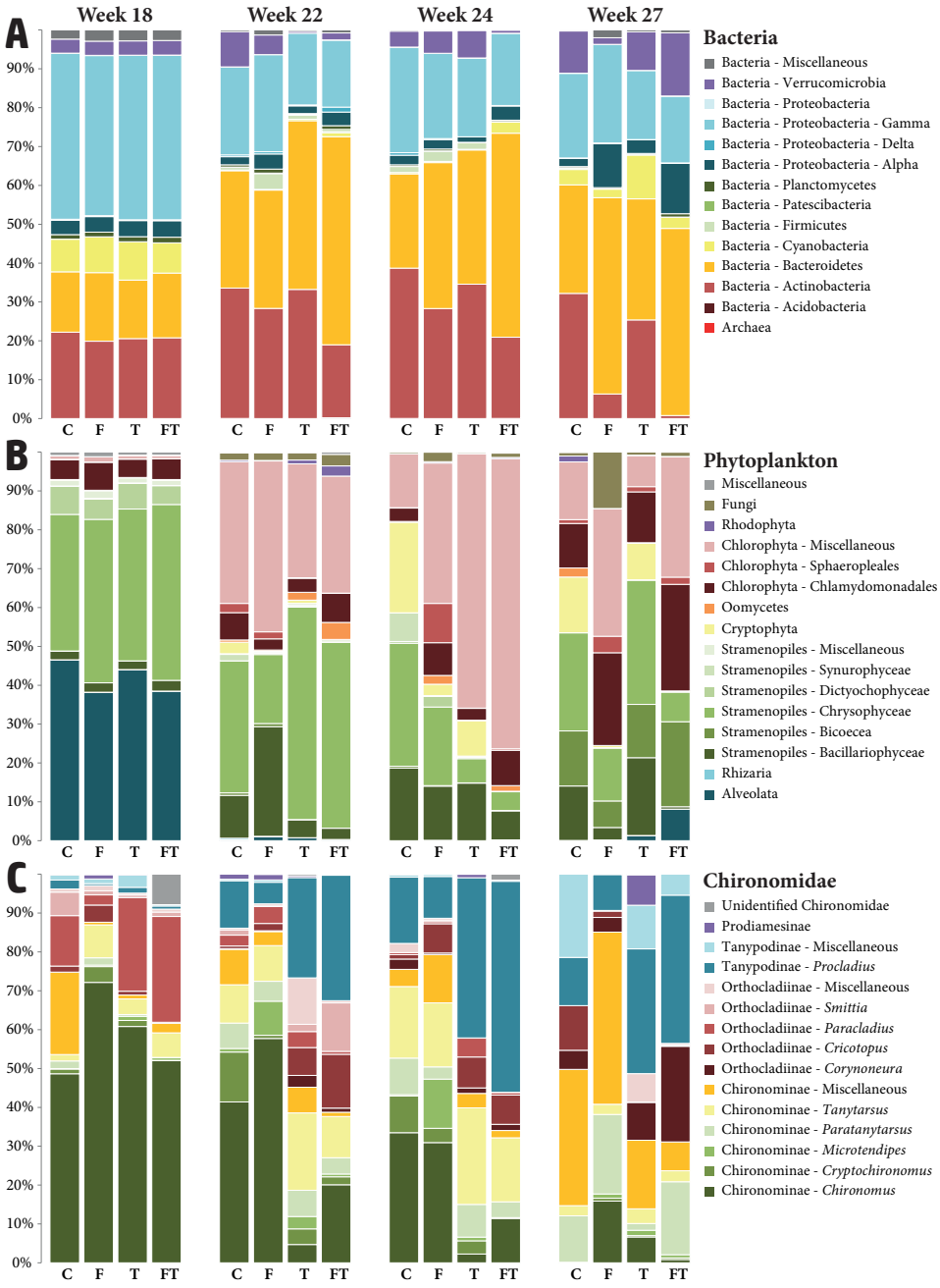


FIGURE 5.3. Read distributions observed for each of the different treatments and control both prior to (week 18) and after application of treatments (week 22-27) for each of the taxonomic groups: (A) bacteria, (B) phytoplankton and (C) chironomids, in control situation (C), and with added fertilizer (F), thiacloprid (T) and combined treatments (FT).

reads). At this point in time, thiacloprid no longer showed a significant effect on the community composition (Table 5.2).

For the chironomids, the most notable differences were observed between ditches with and without added thiacloprid (Figure 5.3C). With thiacloprid addition, the genus *Chironomus* was no longer the most abundant and declined strongly in read abundance (12.4% in ditches with thiacloprid versus 50.0% in control ditches), also compared to week 18 (before the start of treatments), where on average this genus represented 57.7% of the reads. Thiacloprid shifted the community composition towards genera outside of the subfamily Chironominae, such as *Procladius* (subfamily Tanypodinae) (29.0% versus 8.9% in control ditches) and *Cricotopus* (subfamily Orthocladiinae) (10.5% versus 1.3%). This shift continued in week 24, where thiacloprid ditches became dominated by *Procladius* (47.7% versus 13.9%), at the expense of *Chironomus* (6.8% versus 32.2%). In week 27, *Procladius* remained more abundant in the thiacloprid ditches, although not as pronounced as in week 24. The genus *Corynoneura* was also much more abundant in these ditches (17.2% versus 4.3%).

5.3.5 Indicator taxa

Indicator analysis on the three assessments after start of treatments separately identified 624 bacterial MOTUs, 470 phytoplankton MOTUs and 46 chironomid MOTUs that were indicative for either absence or presence of either of the two added agrochemicals in one or more of the three post-treatment measurements (Supplemental File S5.1, summarized in Table 5.3). With the observations of the three assessments combined (week 22, 24, 27), the indicator analysis identified 552 bacterial MOTUs, 76 of which acted as indicators for both agrochemicals. The

TABLE 5.3. Summarized indicator species analysis results, with the number of indicative MOTUs found for each of the three taxonomic groups: indicators for absence (F-) and presence (F+) of nutrients, and absence (T-) and presence (T+) of thiacloprid. Analysis was performed on data from each post-treatment measurement (week 22, 24 and 27), and combined data of the three measurements. An overview of all indicator MOTUs is provided in Supplemental File S5.1.

	Bacteria				Phytoplankton				Chironomidae			
	F-	F+	T-	T+	F-	F+	T-	T+	F-	F+	T-	T+
Week 22	52	54	128	50	60	63	39	56	1	8	13	10
Week 24	43	46	110	39	12	127	40	48	3	4	5	11
Week 27	194	109	25	4	90	91	9	2	3	0	0	3
Week 22-27	172	212	176	68	126	222	93	65	4	8	20	15

majority were indicators for the absence of both fertilizer and thiacloprid (51) or the presence of both (21), the remaining four were indicative for the presence of one agrochemical stressor and the absence of the other. For phytoplankton, we found 446 indicators in total, in which 60 acted as indicators for both agrochemicals, again mainly for absence (15) or presence (39) of both treatments. In the combined chironomid data there were 46 indicative MOTUs, with only a single MOTU that acted as indicator for both presence of thiacloprid and the presence of fertilizer. We did observe lower fidelity values for the combined measurements, due to the fact that indicative MOTUs for all three groups were not observed in the ditches in each of the three assessments after the introduction of agrochemicals.

5.3.6 Comparison to morphological assessment

Patterns observed in stressor responses as measured by distances between centroids in week 24 were similar for all three taxonomic groups assessed in this study, as well as the morphological assessment of macroinvertebrates assessed by Barmentlo et al. (2019) at the same sampling timepoint. The thiacloprid treatment showed more distance relative to the control than the fertilizer treatment, whereas the combined treatment showing the largest deviation for all four assessments (Figure 5.4), although the morphological assessment made use of nine replicates instead of the five replicates that were used for eDNA evaluation. The distances between control

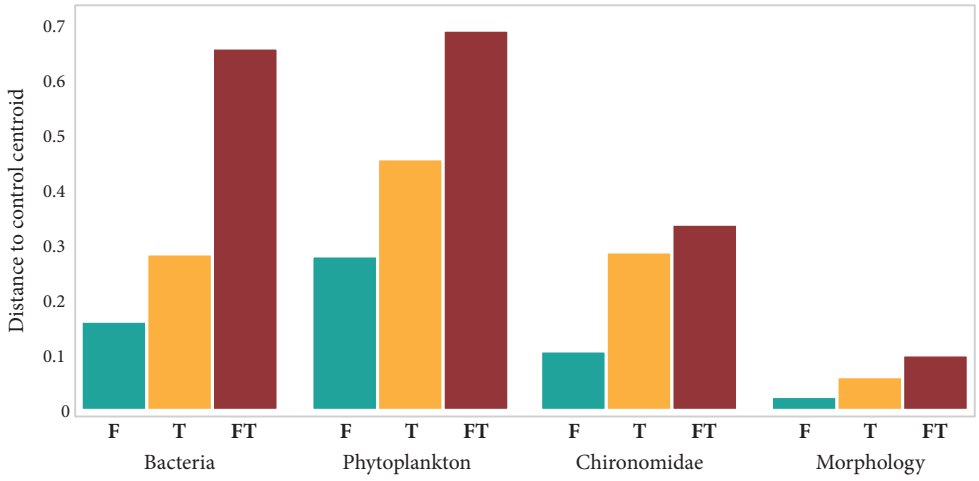


FIGURE 5.4. Centroid distance to the control centroid in week 24 (one month after application of the agrochemicals), for the bacteria, phytoplankton, and chironomids assessed with environmental DNA, as well as the macroinvertebrates assessed with morphological methods (see Barmentlo et al. 2019), exposed to fertilizer (F), thiacloprid (T), and combined agrochemical addition (FT).

centroids varied when using fewer replicates, but in all three eDNA assessments the pattern described above was observed with as little as three replicates (out of five) (Supplemental Figure S5.3B-D). For morphological data, at least seven (out of nine) replicates were needed to reveal this pattern (Supplemental Figure S5.3A). There were, however, no significant correlations with any of the eDNA-based distance matrices and the distance matrix of the morphological assessment. The morphological assessments also showed no significant treatment effect on richness nor abundance of macroinvertebrates (Barmentlo et al. 2019).

5.4 DISCUSSION

Our study shows that environmental DNA can be used to investigate the effects of agrochemical stressors on multiple trophic levels in a freshwater community. The factorial design of our semi-natural research site allowed us to separate the effects of the addition of fertilizer (nutrients) and the neonicotinoid thiacloprid on the taxonomic composition of freshwater organisms. Clear impacts of both agricultural stressors were observed for all three taxonomic groups. Moreover, the introduction of realistic levels of both agrochemicals in the ditches had strong additive effects on the three trophic levels analyzed. Our findings are in line with simultaneous morphological assessments of macroinvertebrates conducted during the same experiment (Barmentlo et al. 2019) and previously reported effects of neonicotinoids on macroinvertebrates and zooplankton (e.g. Yamamuro et al., 2019), and we show that eDNA-based impact assessments can provide useful insights into stressor responses in taxa that are usually not included in traditional assessments. The three groups evaluated in this study have been observed to contain numerous indicative taxa that have potential as novel bio-indicators for environmental stress.

The similar distributions of reads across the various taxa in the measurements before the application of the treatments confirms the assumption that the initial communities were all similar in terms of composition and abundances (Figure 5.3). Subsequent agrochemical addition strongly affected community composition of all three trophic levels investigated (Tables 5.1 and 5.2). The effect of time on dissimilarity was considerable, being larger than the single effects of the agrochemicals, indicating that natural species turnover occurred. Other studies have also found large fluctuations in macrofaunal community composition under normal conditions, even in relatively short periods of time (Bista et al. 2017, Chapter 4).Despite these fluctuations in community composition caused by species turnover, the experimental design of the present study still allowed for clear differentiation in those patterns

caused by seasonal turnover and those caused by external stressors.

Fertilizer addition caused significant changes in community composition of both bacteria and phytoplankton, with replicate ditches showing higher dissimilarity (divergence) compared to control ditches (Supplemental Figure S5.1). Community dissimilarities showed comparable patterns (Supplemental Figure S5.2) and strong correlations for bacteria and phytoplankton. This was expected as eutrophication has long been known to be associated with increased growth in phytoplankton (Heisler et al. 2008), and interactions such as nutrient cycling between phytoplankton and bacteria at the base of the food web (Seymour et al. 2017) render bacterial communities sensitive to changes in phytoplankton communities (and vice versa). Chironomids were also sensitive to the addition of fertilizer, although these fertilized communities were generally more similar to the control than to the thiacloprid treatment (Supplemental Figure S5.2). Nutrient pressure has been shown to have effects on freshwater macroinvertebrates in previous research (e.g. Donohue et al., 2009), since eutrophication can lead to oxygen depletion and changes in food availability.

The thiacloprid concentration used in this study (a nominal time weighted average of 0.4 μg/l) is considered an realistic concentration as it is based on surface water concentrations from the Netherlands, and earlier research has already shown that freshwater macroinvertebrates are affected by neonicotinoids at concentrations observed in surface water (e.g. Morrissey et al. 2015, Sánchez-Bayo et al. 2016). Indeed, thiacloprid addition had a much larger impact on the chironomid community structure than fertilizer addition and resulted in a significant convergence (Supplemental Figure S5.1). Even after thiacloprid had dissipated from the water column after only a few weeks due to its rapid adsorption to the sediment (DT90 = 11.1 days; Barmentlo et al. 2019), the legacy effect of thiacloprid was still larger than the effect of the fertilizer (Table 5.2). This suggests that even a single spike of thiacloprid can have a lasting impact on large parts of the macrofaunal community. There was an additive effect of both agrochemicals, as the impact of a combined treatment effect of fertilizer and thiacloprid was greater than that of each treatment separately, and communities under a combined treatment were more dissimilar relative to the control than communities exposed to a single agrochemical (Figure 5.2, Tables 5.1 and 5.2). Most two-way interactions between fertilizer and thiacloprid were not significant, however, suggesting the effect was additive, rather than synergistic (Table 5.2).

Addition of agrochemicals strongly affected the community compositions. Changes in composition where most notable for chironomids, for which most MOTUs could be identified at species or genus level. For instance, we observed

that thiacloprid treatment ditches (and combined effect ditches) had much lower ratios of the subfamily Chironominae, which were accompanied by higher ratios of species belonging to the subfamilies Tanypodinae and Orthocladiinae. These latter subfamilies were apparently less susceptible to the presence of thiacloprid, which is consistent with findings from previous studies that showed significant effects on Chironominae in response to neonicotinoid insecticides (Langer-Jaesrich et al. 2010, Williams & Sweetman 2019). Whilst the direct effects of fertilizer on bacterial and phytoplankton communities have been studied before (Carvalho et al. 2013), there is little research on the effects of neonicotinoid insecticides on those communities. One study suggests that algal blooms appear to increase in size under stress from the neonicotinoid imidacloprid (Sumon et al. 2018). The neonicotinoid insecticide thiacloprid, meant to target pest insects, also affected bacterial and algal community composition in the present study. Our data suggests that thiacloprid has an important impact on the structuring of the communities (Tables 5.1 and 5.2,Figure 5.3). It is likely that some of these effects on phytoplankton and bacteria communities have been caused by food web cascades, especially as many of the affected Chironominae are common feeders on these microbes. Indeed, previous research showed that even under stress from pesticide mixtures, biotic interactions played a major role in the structuring of plankton communities (Pereira et al. 2018). Similarly, responses to nutrient pressure by fertilizer in macroinvertebrates may also partly be caused by cascade reactions, such as the aforementioned changes in food availability. Processes like eutrophication can have a significant impact on total community composition and food web structure via trophic cascades (Davis et al. 2010, Suikkanen et al. 2013), and a recent study evaluating anthropogenic stressors on freshwater food webs showed that macroinvertebrates had different reactions to fertilizer, herbicide and insecticide, depending on their food source (Schrama et al. 2017). The authors also noted, however, that cascading effects in the food web were hard to explain, and found some suggestions of shifts in diet induced by stressors.

Results from the morphological assessment closely matched the presently observed patterns regarding dissimilarity relative to the control; there was an increase in effect size from fertilizer to thiacloprid to the mixture treatment for all communities investigated, although no significant effects were detected on the beta dispersion of the community in the morphological assessment (Barmentlo et al. 2019). In this study, however, we observed these stressor impact patterns at a lower number of replicates compared to the traditional assessment (Supplemental Figure S5.3). Data for the morphological assessment was log10(x+1) transformed, due to the uneven distributions in species, where zooplankton species often dominated

samples, and chironomids were only identified at family level (Barmentlo et al. 2019). This transformation resulted in considerably lower centroid distances than those observed within the present study, an effect which was amplified by the difference in taxonomic units observed (83 morphological taxa, versus 4,011, 1,773 and 368 MOTUs for the bacteria, phytoplankton and chironomids, respectively). Nevertheless, the morphological assessment of the macroinvertebrates reflected a similar pattern to the eDNA assessment, with the thiacloprid treatment showing more distance relative to the control than the fertilizer treatment and the combined treatment showing the largest deviation (Figure 5.4), irrespective of the biota that were sampled. This indicates how strongly interconnected the different trophic levels are and that potential cascading food web responses to stressors can occur even in non-target biota.

A full-factorial experimental setup such as the one used in this study allows for focused research into the individual and combined impacts of stressors on communities. However, most impact assessments are done in fully natural settings, where the interplay between multiple stressors is much harder, if not impossible, to disentangle (Piggott et al. 2015, Côté et al. 2016). Our semi-natural controlled experimental setup reflected key parts of the food web, and our study shows that using eDNA can successfully describe the effects of agricultural stressors to freshwater communities in a semi-realistic setting. This provides much needed confidence in the application of such an approach in impact studies in natural environments in which disentanglement of the impact of different (a)biotic stressors is even more difficult.

As observed in this study, MOTUs of different taxonomic groups present consistent patterns under the effects of stressors (Figures 5.2 and 5.4). Previous research has already shown that MOTU-based approaches can provide better resolution in impact assessments, such as with undescribed cryptic diversity demonstrating contrasting responses to stressors (Macher et al. 2016), or reference databases being unable to identify all the encountered molecular variation (Beermann et al. 2018). Several studies have shown that MOTU-based assessment methods can accurately predict stressor impact on water systems (e.g. Andújar et al. 2018a, Li et al. 2018b). However, the inability to identify all MOTUs to species or even genus level complicates the ecological interpretations of shifts in communities caused by external stressors. Taxonomic hiatuses in the reference database are large, especially for microorganisms such as the freshwater bacteria and phytoplankton studied in this paper. Accumulating MOTUs based on the higher-level taxonomic assignments could be possible, in order to assign some ecological value to such indicators. The MOTUs, however, may represent a wide variety of ecological groups, and accumulating them into a single entity would decrease the sensitivity of any such bioindicators (Jones 2008). While it

may be difficult to link ecological information to unidentified MOTUs, they can still be of use in comparative studies, such as impact assessments (Li et al. 2018b).

Our analyses revealed a large number of indicative MOTUs for all three trophic levels assessed (Table 5.3), suggesting that many potential new bioindicators are hidden in taxon groups that are either difficult to identify (e.g. chironomids) or are mostly neglected in traditional bioassessments of water impacted by anthropogenic stressors (e.g. bacteria and phytoplankton). Especially for bacteria and phytoplankton, many indicative MOTUs were found, for both stressors, but many MOTUs could not be resolved to species or genus level. Most of the chironomid MOTUs identified as positive indicators for the presence of thiacloprid could be assigned to the genus *Procladius* (Tanypodinae). These observations match the findings of a recent morphological study performed in the same experimental ditches where *Procladius choreus* was the most abundant remaining species under the stress of increasing levels of thiacloprid (Barmentlo et al. in prep).

One key limitation for assigning indicator taxa for freshwater communities is the large fluctuations in community composition over time. The large community turnover caused low fidelity scores for many indicator MOTUs observed in the indicator analysis on the combined data for the three post-treatment measurements, due to the fact that many MOTUs do not occur in all time points (Table S5.2). Moreover, indicator MOTUs might not only be specific to a certain time frame, but can also be spatially limited, as it was previously observed that indicator species for the impact of offshore oil and gas drilling were highly specific to site conditions (Laroche et al. 2018). Impact assessments based on novel indicators, or even based on MOTUs, should preferably be time- and location-independent, to make their application on a broader scale feasible. This could prove challenging, especially when looking at microorganisms such as bacteria or phytoplankton taxa observed in the current study, as these groups tend to have a large turnover in their community composition on a relatively small time scale (Chapter 4). However, the huge potential for these novel bio-indicators in large-scale impact assessments would make any efforts into a better understanding of their occurrence and behavior worthwhile.

5.5 CONCLUSIONS

We have shown that environmental DNA metabarcoding at multiple trophic levels provides insights into changes in freshwater communities under pressure of agricultural stressors. The full-factorial design of the mostly natural study site allowed us to observe the impact of single stressors. We found an additive (but not

synergistic) effect of artificial fertilizer and the insecticide thiacloprid on community composition at the level of decomposers, primary producers, and consumers. This effect of multiple stressors was consistent with observations reported in traditional morphological assessments of the same experimental setup. These effects were even detected with a lower number of replicates than the traditional morphological study, indicating the robustness of using environmental DNA metabarcoding in impact assessments. While both agrochemicals directly influenced different taxa at different trophic levels, the neonicotinoid insecticide thiacloprid, meant to target pest insects, also affected bacterial and algal community composition, be it directly or through cascade reaction through the food web. We encourage the use of multi-marker eDNA for impact assessment across trophic levels in freshwater ecosystems, as it (1) provides a more comprehensive assessment of impacts on the entire food web, (2) provides more information at a higher taxonomic resolution compared to traditional morphological surveys, even if MOTUs are not all assigned to species level, and (3) allows for discovery of novel indicator taxa. The incorporation of eDNA methodology contributes to ecosystem understanding and would allow for more effective monitoring and management of freshwater systems, and help safeguard the ecosystem services they contribute to humanity.

5.6 ACKNOWLEDGEMENTS

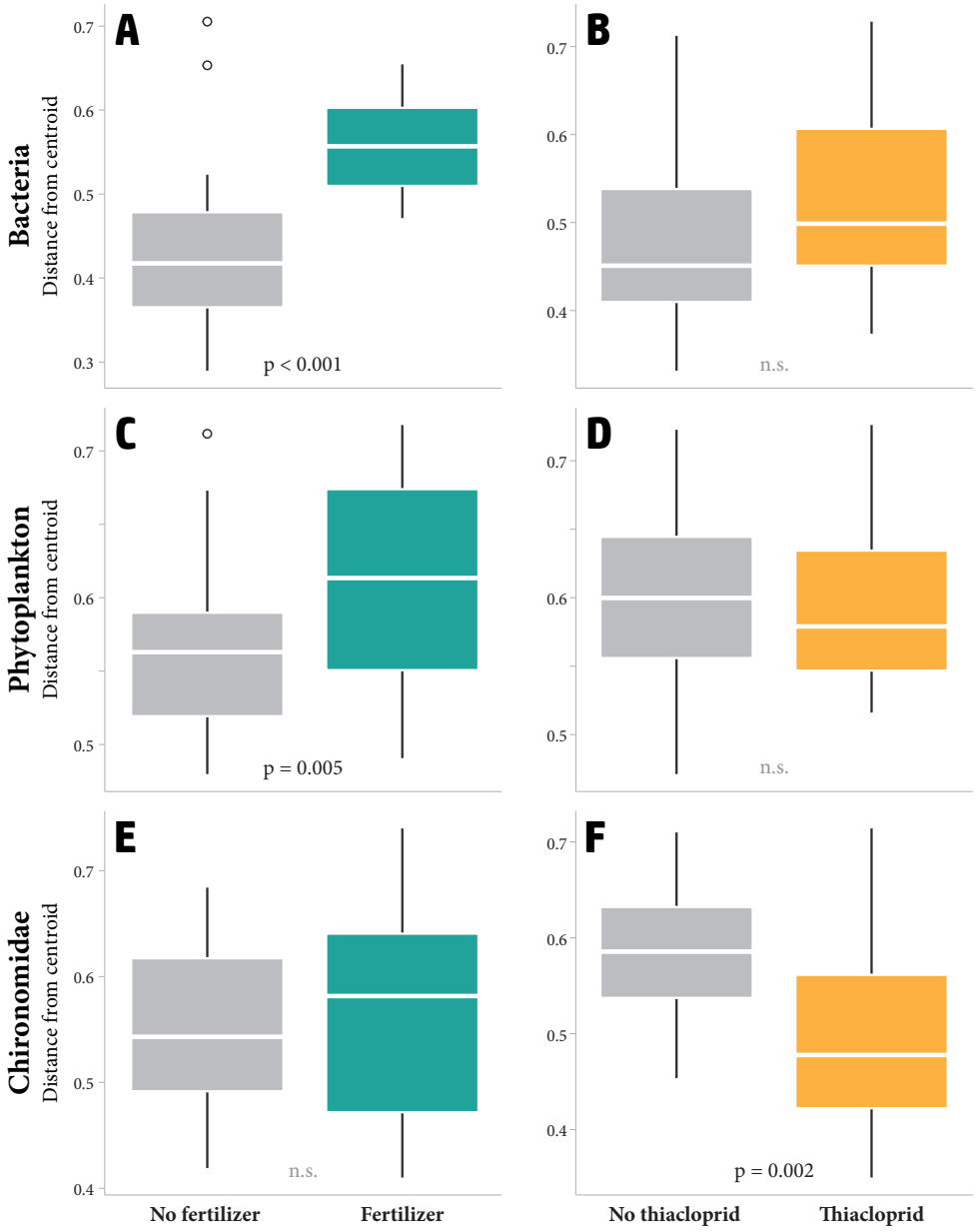
We thank Sam Boerlijst for his assistance with DNA extractions, and André van Nieuwenhuijzen for his critical look at the identifications of chironomids obtained from the eDNA metabarcoding. Experiments were performed in the outdoor experimental laboratory "Levend Lab" established by crowdfunding by Maarten Schrama and Martina Vijver at Leiden University.

5.8 SUPPLEMENTARY MATERIALS

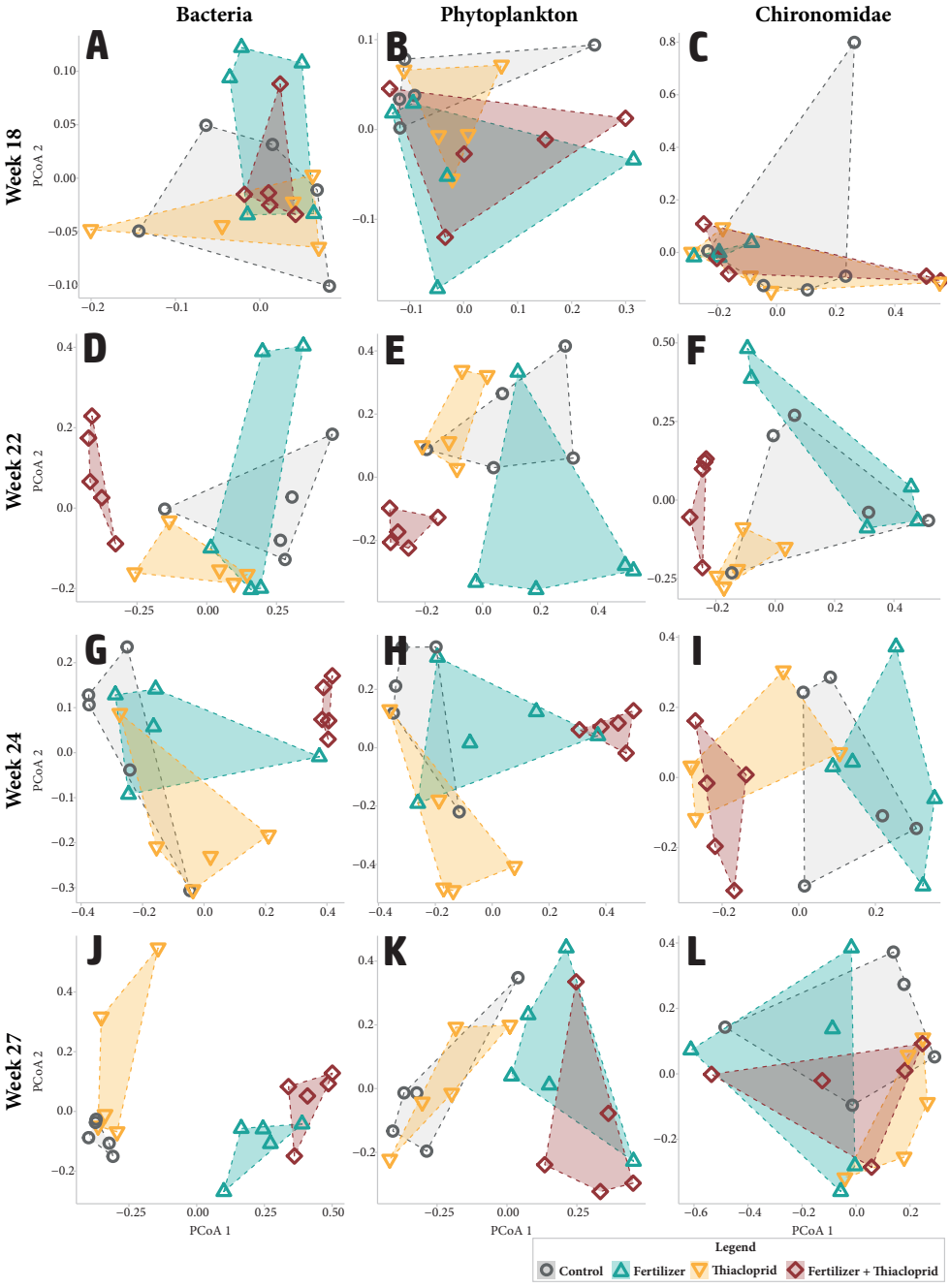
SUPPLEMENTARY FILE S5.1. Indicative MOTUs for bacteria, phytoplankton and Chironomidae for either absence or presence of either of the two added agrochemicals in one or more of the three post-treatment measurements. **https://doi.org/10.22541/au.159236833.30909538**

SUPPLEMENTARY TABLE S5.1. Sequences for primers used in the first and second round amplification.

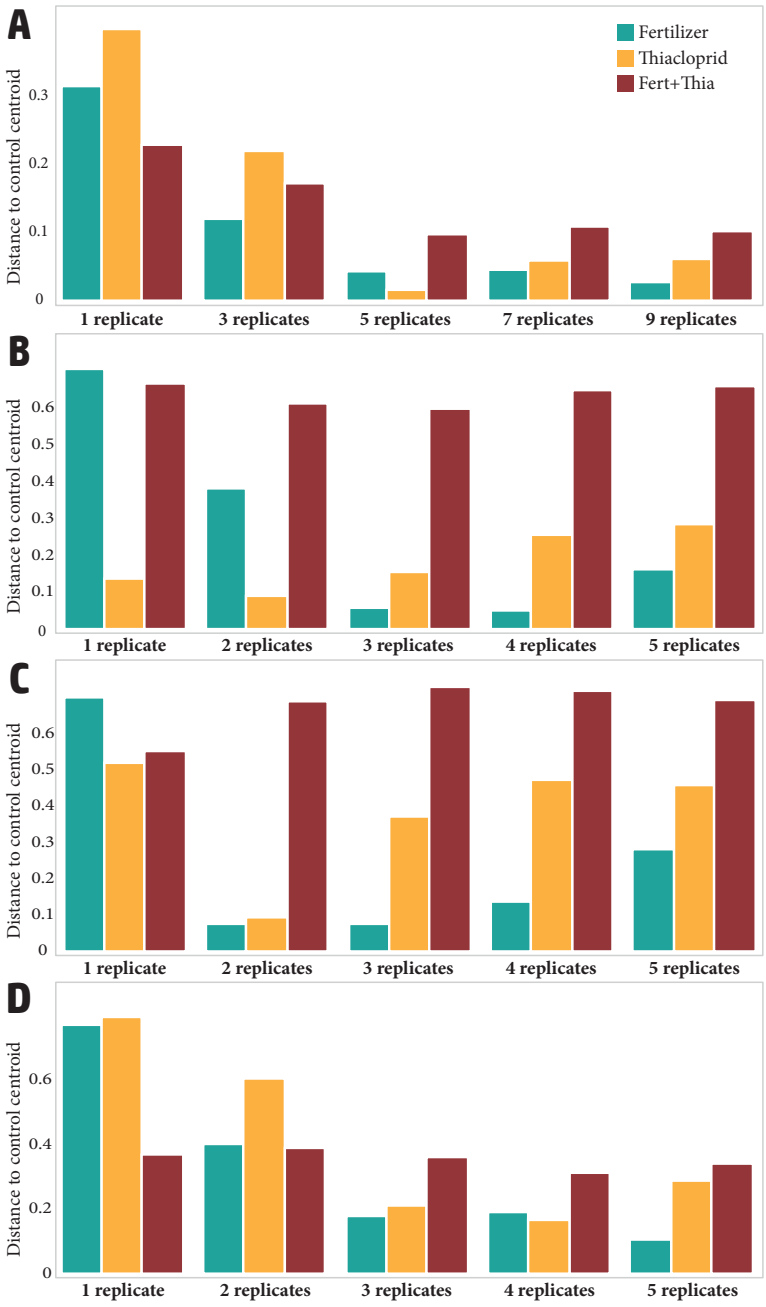
First Round	
Primer set	Sequence (Universal tail – template-specific primer)
Bacteria (Klindworth et al. 2013)	
S-D-Arch-0519-a-S-15	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG CAGCMGCCGCGGTAA
S-D-Bact-0785-a-A-21	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG TACNVGGGTATCTAATCC
Phytoplankton (Zimmerman et al. 2011)	
D512for	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG ATTCCAGCTCCAATAGCG
D978rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG GACTACGATGGTATCTAATC
Chironomidae (Bista et al. 2017)	
LCO-1490	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG GGTCAACAAATCATAAAGATATTGG
COIA-R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG CARAAWCTTATATTATTTATTCGDGG
Second Round	
Primer	Sequence (Illumina adapter – index – universal tail)
NEX-F	AATGATACGGCGACCACCGAGATCTACAC [i5 index] TCGTCGGCAGCGTC
NEX-R	CAAGCAGAAGACGGCATACGAGAT [i7 index] GTCTCGTGGGCTCGG



Supplemental Figure S5.1. Beta dispersion in weeks 22 to 27 under the two different treatments (tested independently) for bacteria exposed to fertilizer (A) and thiacloprid (B), phytoplankton exposed to fertilizer (C) and thiacloprid (D), and chironomids exposed to fertilizer (E) and thiacloprid (F). Fertilizer caused significant divergence in bacteria and phytoplankton communities (A and C), whereas thiacloprid caused significant convergence in chironomid communities (F). ANOVA p-values are provided in the panels.



Supplemental Figure S5.2. PCoA plots for each of the four measurements, both prior to (week 18, A-C) and after application of treatments (week 22-27, D-L) for of the three taxonomic groups: bacteria, phytoplankton and chironomids.



Supplemental Figure S5.3. Average distance from centroid to the control centroid in week 24 for the (A) macroinvertebrates assessed with morphological methods (Barmentlo et al. 2019), and (B) bacteria, (C) phytoplankton and (D) chironomids assessed with eDNA (this study), at different numbers of replicates.

CHAPTER 6

General discussion and synthesis

Kevin K. Beentjes

KEY FINDINGS

1. For relatively simple water quality assessments such as the ecological quality ratio scoring, abundances in taxon data are of limited influence.
2. Taxonomic sorting prior to DNA analysis reduces the impact of preferential amplification, as data from complex sample mixtures with uneven biomass distributions between various taxa are often dominated by reads belonging to a single taxon. This shows that, even with universal primers, the effects of primer bias are still significant.
3. Differences in community composition caused by small-scale temporal turnover are equal to or larger than those caused by heterogeneity. Sampling replicates over time are more important for insight into the total diversity than spatial replicate sampling.
4. Multi-marker eDNA impact assessments across trophic levels prove to be a more comprehensive indicator of impacts on the food web and provide more information on a higher taxonomic resolution, whilst uncovering similar impact patterns as more cumbersome morphological surveys.

While there is an ever-increasing number of publications on the possibilities and limits of environmental DNA and DNA metabarcoding in biodiversity monitoring, it has become clear that these new techniques will most likely never truly conform to the needs of the traditional monitoring schemes. However, recent insights and developments have shown that there is merit in molecular biomonitoring. Developing a better understanding of the ecology of eDNA, as well as getting a grip on the effects of different choices in the field, lab, and analysis is paramount to making molecular tools successful. Implementation of DNA-based techniques such as eDNA sampling and metabarcoding can never work when the mechanisms behind the techniques are not understood properly, and there are still knowledge gaps in both methods and applications of eDNA (Garlapati et al. 2019). Closing these gaps is important, especially when dealing with regulations such as the EU-WFD. While there will always be some differences between studies that are introduced by sampling design, laboratory protocols, and analysis pipelines, understanding the causes of these differences will lead to better documentation of protocols and help in the intercalibration of studies.

6.1 SAMPLING STRATEGIES

The ecology of eDNA as discussed in Chapter 1 has its implication on practice as well, mainly on the sampling strategy deployed in the field, but also on downstream processes. There are numerous papers that describe other critical considerations during the entire pipeline from field sampling and sample processing to analysis and reporting of results (Goldberg et al. 2016, Dickie et al. 2018, Harper et al. 2019a, Zinger et al. 2019, Nicholson et al. 2020). While eDNA sampling seems relatively straightforward, especially compared to kicknet sampling, the reality is slightly more complicated. Environmental DNA concentrations are usually low, especially where it concerns rare species or alien species in early stages of invasion (Tréguier et al. 2014). One of the general conclusions that can be drawn from the eDNA literature is that larger volumes of water are preferable, although increased volumes do not always lead to better detection rates (Mächler et al. 2016). The need for spatial replicates is evident, but the replicate sampling strategy needs to be adapted to the specific questions and target organisms, much like traditional monitoring efforts. This not only concerns the number or replicates, but also the distance between them and the structure of sampled habitats (Lugg et al. 2018, Grey et al. 2018). Sampling sediment can be an alternative to water sampling, since a large proportion of the eDNA will settle to the sediment. Concentrations of eDNA in sediment can be higher than concentrations in the water column (Turner et al. 2015), but can represent a more historical signal.

Once eDNA water samples have been collected, they should be filtered in the field or preserved until they can be processed in the laboratory. Various methods to preserve water have been published, but the efficacy of each of those methods is often only poorly tested, and in many publications not compared to other standard methods. Directly filtering water in the field has become more standard (Pilliod et al. 2013, Turner et al. 2014b, Hinlo et al. 2017, Grey et al. 2018), especially since there have been various papers that describe methods that allow for better preservation of filters in buffer (Renshaw et al. 2015), and prevention of cross-contamination between samples due to filter handling in the field by using enclosed filters (Spens et al. 2017, Thomas et al. 2019). Filtration methods have been examined in great detail, and many different filter types, pore sizes, and filter processing techniques have been compared in a broad range of studies using both single-species detection (Eichmiller et al. 2016) and metabarcoding for a variety of target organisms (Djurhuus et al. 2017, Li et al. 2018a, Majaneva et al. 2018, Jeunen et al. 2019). Generally, larger pore sizes seem to be optimal regarding the balance between volume of water filtered and DNA yielded from the filters.

6.2 REGARDING REPLICATES

As shown in Chapter 4, replicate sampling strategies are important to capture the full diversity of organisms in an aquatic environment. Especially in lentic systems, where dispersal of eDNA appears limited, spatial replicates are necessary to pick up these oft local signals (Thomsen et al. 2012b, Evans et al. 2017b, Grey et al. 2018, Lawson Handley et al. 2019). The data from Chapter 4, where eDNA samples were collected weekly during 20 consecutive weeks, in a total of six locations in two dune lakes, shows the heterogeneous distribution of environmental DNA signals in space and time. Looking at the three spatial replicate samples taken in each lake, the majority of observations (57.8%) were limited to MOTUs only found in single replicate sample.

One of the main reasons for the research presented in Chapter 4 was to delve into the temporal patterns of environmental DNA. While there is often specific focus on the inclusion of spatial replicates during sampling in the field, studies looking at seasonal variation are usually limited to larger temporal scales (e.g winter versus summer) (Chain et al. 2016, Guardiola et al. 2016). Various papers examining single species or a select number of taxa have already shown that detection rates for these taxa can vary throughout the year (Stoeckle et al. 2017, Buxton et al. 2017). Few studies, however, have been performed on small-scale temporal differences in water bodies, looking at intervals of months (Bista et al. 2017, Rees et al. 2017) or even weeks (Sigsgaard et al. 2017). Our data shows that the effects of turnover on a fine scale (weekly sampling) contributes as much to the overall observed diversity as community heterogeneity (as observed with spatial replicate sampling). Dissimilarity is increased over larger intervals, where it contributes more to the total observed diversity than spatial replicates for intervals of more than two weeks. The dissimilarity between taxonomic composition of the two lakes studied in this study showed a linear increase with increased intervals (Figure 4.4). This suggests that any study comparing diversity across different sites is susceptible to inflated dissimilarities when study sites are sampled mere weeks apart. Moreover, weekly samples during the 20-week period on average only contained 20.7% of the total observed MOTU diversity (Figure 6.1A), with an average 6.5% of those observations limited to single time point (Figure 6.1B). High seasonal diversity has since also been described for metabarcoding of bulk-collected samples in freshwater streams (Zizka et al. 2020), furthermore stressing the importance of sampling moment in comparative studies. However, that study also found that for the relatively simple water quality assessment scores, seasonal variation, much like abundance data, is of limited influence.

In addition to biological replicates in the field, technical replicates during

extraction (Lanzén et al. 2017) and especially PCR are often highlighted in literature focusing on sampling methodology. PCR replicates are often common practice in qPCR detection studies looking at single species (e.g. Biggs et al. 2015, Agersnap et al. 2017, Harper et al. 2018), but are starting to be used in metabarcoding studies as well (Civade et al. 2016, Alberdi et al. 2018). Especially when working with low quantities of eDNA, the heterogeneity of the DNA extract can cause issues. While the dissimilarities between PCR replicates were not as large as the dissimilarities between spatial and temporal replicates (Figure 4.3), including these replicates does increase

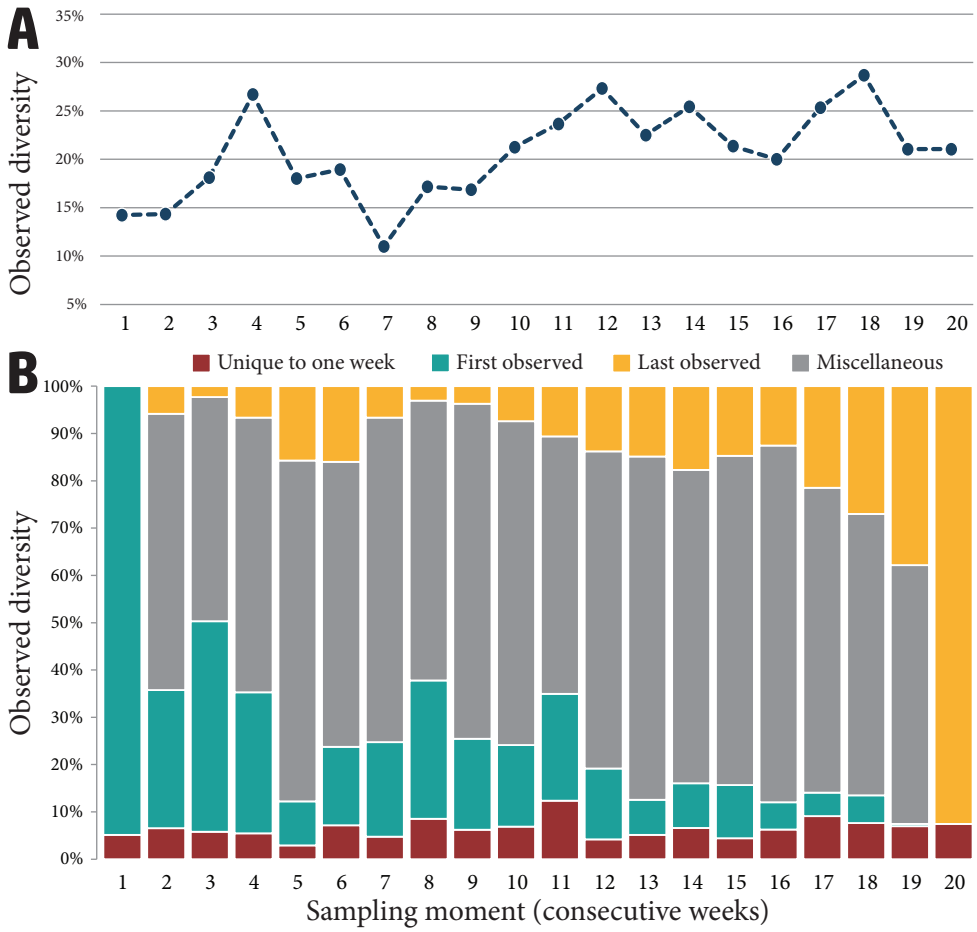


FIGURE 6.1. The distribution of MOTUs in data from Chapter 4, with (A) the number of observed MOTUs expressed as a fraction of the number of MOTUs found in total across the 20-week sampling period, and (B) the distribution of MOTUs observed each week based on when each MOTU was observed across the 20-week sampling period. These are either unique observations for a single time point, and MOTUs that were either not observed before or after each time point.

the observed diversity, as many rare species are often found only in a single PCR replicate. Data from Chapter 4 shows that 40.2% of all observed MOTUs have only been found in a single PCR replicate, whereas only 27.4% of MOTUs were observed in all three PCR replicates. Including PCR replicates also decreases the uncertainty in detection rates (Mächler et al. 2016). To circumvent the need for qPCR replicates in order to get accurate concentration measures in single species detection studies, droplet digital PCR (ddPCR) has been employed with promising results (Doi et al. 2015, Uthicke et al. 2018).

6.3 ADDRESSING ABUNDANCE

Interpretation of molecular data requires some insight into the behavior of DNA, especially where it concerns environmental DNA. For example, caution has to be taken when interpreting DNA-based data in terms of specimen abundances. Several papers that use single-species detection show good correlations between biomass of target organisms and DNA concentrations from standardized eDNA samples (Pilliod et al. 2013, Klymus et al. 2015, Uthicke et al. 2018, Spear et al. 2020). To translate such concentration measures back to actual present biomass or specimen counts remains difficult, and would require extensive calibration efforts, or extensive sampling regimes (e.g. Levi et al. 2019). DNA concentrations do, however, allow for the inference of differences in specimen abundance between different locations, providing useful information for water management (Lacoursière-Roussel et al. 2015).

For metabarcoding, however, it is more complicated. The preferential amplification of certain taxa over others (PCR bias), can cause difficulties in the interpretation of molecular data and the comparison to morphological assays. A study by Elbrecht and Leese (2015) showed that sequence abundance varied by up to four orders of magnitude between species sequenced from bulk samples, even though the input biomass was comparable for all species. Preferential amplification lies at the basis of most difficulties surrounding the inference of abundances from molecular data, but additionally also impacts the detection of rarer species in complex mixtures with uneven biomass distributions. This disbalance in biomass causes that large proportions of the DNA extract come from only few large specimens, and it often happens that rare species are lost in the background, especially when sequencing depth is not sufficient. In Chapter 3 we looked into the effects of taxonomic sorting on the detection of taxa. Samples from WFD monitoring sites were homogenized using a blender, but with a prior sorting into six taxonomic groups as provided by the monitoring agency. Pooling of DNA extracts prior to PCR and sequencing

was performed to simulate a situation where all organic material was homogenized without any sorting. The idea behind the analysis of sorted samples stemmed from previous results obtained during pilot studies, where samples were often dominated by a single species or taxon. The sorting resulted in much higher estimated diversities, both in terms of MOTUs and formal taxa (Figure 3.1), with 46.5% increase in taxon recovery. As expected, the pooled samples were often dominated by certain taxa, be it one of the six groups (15 out of 24 samples), or even a single taxon (eleven samples) (Supplementary Figure 3.3). Sorting bulk samples into size fractions prior to homogenization, to prevent large specimens from dominating the DNA data, provided similar results to our findings, with significant increases in taxon recovery (Elbrecht et al. 2017b, 2020). Such methods may be more feasible in practice.

The use of eDNA metabarcoding is even more prone to complications, because there often is no original biomass to compare read data with, and research has shown that eDNA often has a very local and heterogeneous distribution in lentic waters. Some studies with relatively few taxa in aquarium setups have shown modest relations between biomass and read abundances (Evans et al. 2016), and read abundances can still be informative in comparative studies, such as those that evaluate fish community assemblages along a river (Pont et al. 2018). Preferential amplification certainly affects eDNA studies too, in some cases even preferentially amplifying taxa that are not the intended target, as witnessed by the many non-macrofaunal taxa amplified using the primers in Chapter 4, which were developed as universal macrofauna primers. Environmental DNA studies seem hampered more than helped by such universal primers, as they tend to pick up a lot of unintended "bycatch".

Fortunately for many water quality indices, abundance has proven to be of limited importance. As shown in Chapter 2, there is a strong correlation between ecological quality scores calculated with and without specimen abundances in the Dutch WFD system. Similar results have been obtained for quality scoring for freshwater systems in New Zealand (Wright-Stow & Winterbourn 2003) and Germany (Buchner et al. 2019), as well as for the AZTI's Marine Biotic Index, a marine scoring system used by many European countries (Aylagas et al. 2014). For other assessments of biodiversity, however, abundances are often an important parameter, and using presence/absence data can overestimate the importance of rare species (Deagle et al. 2019). Alternatives like shotgun sequencing seem to provide decent correlations between biomass and read abundance (Bista et al. 2018). They do introduce other difficulties though, since usually only the mitochondrial DNA is usable for matching to reference sequences, which only amounts to at most 0.5-1.0% of the read data. Calibration studies evaluating PCR bias in NGS via qPCR (Pawluczyk et al. 2015) are cumbersome for

studies with a potentially large number of taxa (most macroinvertebrate assessments) and require *a priori* knowledge of the species composition of a sample. Using internal standards with known concentrations in metabarcoding might at least allow for the different samples to be more comparable to each other regardless of differences in sequencing depth (Ushio et al. 2018).

6.4 MOTUs VERSUS TAXA

The other main challenge of DNA-based data, besides the difficulties in interpreting read abundances, lies with the inferences of taxonomic composition of samples. Sequence data is usually clustered before analysis and comparison to references, but there are several different clustering approaches that all produce their own MOTUs (molecular taxonomic operational unit) (Clare et al. 2016). The most common methods use clustering based on dissimilarity percentage thresholds between reads, others take abundances of exact sequence variants into account (Edgar 2016, Rognes et al. 2016). MOTU clustering can introduce either overestimation of diversity by creating separate clusters for taxa with high intraspecific variation, or underestimation due to lumping taxa with low interspecific variation, and finding a balance between these can be difficult (Alberdi et al. 2018).

Chapter 3 illustrates this potential overestimation of richness, where the number of MOTUs found in each sample was approximately five-fold the number of taxa (211.4 versus 40.8 on average), and the correlation between the numbers of MOTUs and the numbers of morphological taxa was weak ($r = 0.365$) and above all not significant ($p = 0.072$) (Figure 3.1). The difference, however, was exacerbated by the fact that many specimens in the morphological analysis had not been identified up to species level, and higher-level taxonomic observations are likely to represent clusters of multiple taxa. We also found various MOTUs that had the same taxonomic assignments, suggesting cryptic species or haplotype diversity is present in many taxa, which remain undetected during morphological analysis (Gibson et al. 2015, Elbrecht et al. 2018a). When specifically looking at the number of taxa observed with morphology and molecular analysis, the correlations were stronger ($r = 0.662$), although in this case there were still issues with differences in depth of taxonomic assignment for various groups. For example, various dipterans were not identified to species level with morphology, but were with DNA metabarcoding, whereas the COI fragment was unable to differentiate between morphologically different species of leeches.

Even for those taxonomic groups that are well-covered in the (public) DNA reference libraries, reference specimens usually do not cover the entire geographic

range in which they occur. These references may thus not represent the full genetic diversity of species, especially over larger geographical ranges (Bergsten et al. 2012, Baselga et al. 2013). The limited dispersal of many aquatic taxa has led to highly structured populations in insular freshwater systems (Strayer 2006), where even species that are assumed to be common and widespread show high levels of cryptic diversity (Alp et al. 2012, Sworobowicz et al. 2015). Metabarcoding studies can expose this intraspecific genetic variation (Elbrecht et al. 2018a, Sigsgaard et al. 2020, Chapter 3), but may also lead to overestimations of diversity when using MOTUs that cannot be linked to reference databases (Brown et al. 2015). Attempts to identify such MOTUs at a higher taxonomic level may introduce a lot of noise into a dataset (Berney et al. 2004). In an ideal situation all the possible species are represented in the DNA reference libraries, and all sequencing reads can be matched directly to one of the potential species. This might work to a certain extent for metabarcoding of relatively simple bulk samples, but complex samples or environmental DNA samples will always contain sequences that cannot be linked directly to a reference, especially when using universal primers or when analyzing microorganisms (e.g Chapters 4 and 5). Even for commonly studied taxon groups such as the freshwater macroinvertebrates, or marine macrobenthos, DNA reference libraries are far from complete (Wangensteen et al. 2018, Weigand et al. 2019) (see also Figure 1.3). Especially uncommon species, which are also often missed by morphological surveys (Jackson et al. 2014), might be absent from the reference libraries.

In addition to undescribed genetic diversity, many organisms carry pseudogenes that potentially also introduce overestimations of species richness in metabarcoding studies (Song et al. 2008, Buhay 2009). These nuclear-mitochondrial pseudogenes (NUMTs), are co-amplified with the target region. This effect becomes more profound for those species that are more abundant, as over-amplification tends to bring this signal, which often resides in the background, to light. For example, in the dataset for Chapter 4 we were able to identify a total of 21 putative pseudogene sequences in the muntjac control samples. After filtering the data, these samples contained 22 MOTUs, which were all identified by the LCA as either *Muntiacus* or a member of the subfamily Cervinae. One of these MOTUs represented 96.0% of all read data from the control samples and resulted in a 100% match with Sanger sequenced reference sequences of the *Muntiacus reevesi* sample. The second largest MOTU, which represented less than 0.5% of the control sample read data, only showed a 93% match with *M. reevesi.* As no other species were expected to be present in these control samples (DNA was extracted from *M. reevesi* blood samples), we postulate that the 21 additional MOTUs found in the control samples are pseudogenes. Read

errors seem unlikely, as most of the 21 MOTUs are present in all twelve independent PCR replicates, suggesting the signal was intrinsic to the control DNA sample. Four out of 21 MOTUs had indels causing frame shifts, twelve more had stop codons in their reading frame, meaning that there were five suspected pseudogenes that were not identifiable as pseudogenes based on their sequence alone. It is also difficult to define putative pseudogenes based on the likeness to highly-abundant actual biological signals, since they often differ substantially from their original sequence (93.2-82.6% pair-wise identity in case of the 21 muntjac sample MOTUs). This means that many DNA metabarcoding datasets are likely to have unrecognized pseudogenes present, an effect that is furthermore amplified in datasets with larger sequence data outputs. This is also supported by findings in Chapter 3, where 14 MOTUs assigned to *Asellus aquaticus* had significantly more reads that the 109 MOTUs assigned only to the genus *Asellus* via LCA (75,128 versus 1,768 reads on average), suggesting that at least some of the latter may have been pseudogenes. Large-scale DNA barcoding programs such as the Global Malaise Trap Program (Geiger et al. 2016) or BIOSCAN (Hobern 2020) could resolve some of these issues by creating large datasets with better geographical coverage directly linked to actual specimens, as it will help us understand which part of the observed diversity in metabarcoding studies translates to actual biological diversity.

On the other hand, when analyzing patterns in richness or diversity, the effects of these "fake" MOTUs are likely limited, similar to rare species. As seen in the impact assessment in Chapter 5, the ecotoxicological effects on MOTU richness were not any different from the effects on morphological richness estimates in the exact same study site. One could also argue that overestimations of richness are comparable within one study (i.e. the effect is the same for all samples). Response patterns based on this potentially overestimated diversity were also similar to those observed in morphological assessments (Figure 5.4), even though the molecular assessment used fewer replicates. In any case, it is difficult to correct for overestimations, as MOTUs can often not be distinguished as artefacts. However, with all the undescribed genetic diversity in many (common) taxa, "lumping" all these MOTUs leads to a loss in potentially interesting information. Much like specimen-based assessments take advantage of DNA-based identifications and may even lead to new species-specific insights on ecology (Jackson et al. 2014), metabarcoding studies will uncover genetic diversity that goes beyond currently recognized species. Studies have already shown that different clades, haplotypes or cryptic species can have different and even contrasting responses to environmental stressors, and many unidentified MOTUs can still provide informative response patterns to stressors (Macher et al. 2016,

Beermann et al. 2018, 2020, Chapter 5). These studies indicate that even with an incomplete reference library or unresolved cryptic species complexes, DNA barcodes provide higher-resolution taxonomic information that can be used for assessments. Furthermore, the "taxonomy-free" studies that have been performed with diatoms already show that MOTU-based data sets can provide ecological status assessments similar to traditional surveys. Chapter 5 illustrates that eDNA metabarcoding data showed the same impact patterns in response to agricultural stressors, on different trophic levels. Such multi-marker impact assessments also provide more information on a higher taxonomic resolution, even if not all of the MOTUs can be assigned to taxa at this point in time. This corroborates the findings in marine aquaculture impact studies, which showed that different markers could accurately predict stressor impact and even outperform the more cumbersome traditional methods (Cordier et al. 2019).

6.5 FUTURE PERSPECTIVES

Despite the many ongoing discussions around the technical considerations when dealing with metabarcoding data, there seems to be some consensus on its usefulness to provide information on species occurrences and changes thereof. The number of new papers on the topic is rapidly growing (Figure 1.2), and many researchers have shifted the focus of their work towards the implementation of molecular tools in actual monitoring. There is need for a solid foundation of scientific research that directly compares traditional monitoring with new methodologies, as this is the work where similarities and differences between the "old" and the "new" come to light. Some traditional monitoring systems are more easily supplanted by DNA-based techniques, such as impact assessments or the detection of invasive species. Others are currently too heavily adapted to traditional monitoring and information that is hard or impossible to obtain from molecular data, such as the age and/or size distributions of fish, but many molecular techniques will at least provide complementary data to obtain a more complete insight into the ecosystem.

The main issue with traditional monitoring, especially the monitoring as prescribed by the WFD (and its national interpretations), is that it is set up to deal with the shortcomings of the traditional techniques. This is why WFD monitoring mainly focusses on groups that are identifiable by light microscopy and relatively easy to collect. We must, however, not try to make new techniques compatible with imperfect existing systems. The potential of molecular techniques has been proven in many scientific papers, and leaves ample room to develop new monitoring schemes

that fully harness the power of these DNA-based tools. Better insights into the genetic diversity of many species is direly needed, as this will allow for the improvement of diversity measures based on molecular data. It will allow us to find a balance between the potential overestimation caused by artefactual sequence data and the underestimation caused by lumping potentially informative cryptic taxa into a single entity. This is especially relevant since the trend in DNA metabarcoding studies moves towards more and more stringent clustering methods.

That is not where the work stops, however. With more insight into the genetic diversity of taxa, we will also be able to do a more in-depth exploration of the ecological meaning of such genetic variants. As shown in a few studies, these cryptic species or (sub)populations related to genetic variations can have very different responses to environmental stressors, which has a considerable impact on stressors assessments. This would be a lot of work, and we seem to have barely scratched the surface. The genetic diversity, and its ecological diversity, also showcase the continued importance of taxonomists, which are sorely needed not only to fill and quality-check the ever-growing DNA references libraries, but also to find what this genetic diversity uncovered by next-generation sequencing means in the field. The technological advances in environmental DNA and metabarcoding studies are meaningless without taxonomic and ecological knowledge to translate sequences into an understanding of the ecosystem. Multi-trophic analyses of communities show there are cascading effects in food webs, and they not only provide information on the composition of an ecosystem, but also its interconnectedness and, more importantly, its functioning.

Future developments in ecological assessments will have to focus more on ecosystems truly as systems, rather than just a collection of taxa. Additionally, one of the most important challenges for scientists in the coming years is to also translate the findings from these new DNA-based monitoring methods into useful information for monitoring agencies and policy makers. They are the ones that need to be convinced of the merits of molecular monitoring at this point, as it seems that the scientific community has all but embraced the techniques.

REFERENCES

Afgan E, Baker D, Batut B, Van Den Beek M, Bouvier D, Ech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D (2018) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Research 46:W537-W544. doi: 10.1093/nar/gky379.

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. doi: 10.1371/journal.pone.0179261.

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. doi: 10.1111/2041-210X.12849.

Alcamo JM, Vörösmarty CJ, Naiman RJ, Lettenmaier DP, Pahl-Wostl C (2008) A grand challenge for freshwater research: understanding the global water system. Environmental Research Letters 3:010202. doi: 10.1088/1748-9326/3/1/010202.

Alp M, Keller I, Westram AM, Robinson CT (2012) How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. Freshwater Biology 57:969-981. doi: 10.1111/j.1365-2427.2012.02758.x.

Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: Does read abundance count? Molecular Ecology 19:5555-5565. doi: 10.1111/j.1365-294X.2010.04898.x.

Andruszkiewicz EA, Starks HA, Chavez FP, Sassoubre LM, Block BA, Boehm AB (2017) Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. PLoS ONE 12:e0176343. doi: 10.1371/journal.pone.0176343.

Andújar C, Arribas P, Gray C, Bruce C, Woodward G, Yu DW, Vogler AP (2018a) Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. Molecular Ecology 27:146-166. doi: 10.1111/mec.14410.

Andújar C, Arribas P, Yu DW, Vogler AP, Emerson BC (2018b) Why the COI barcode should be the community DNA metabarcode for the Metazoa. Molecular Ecology 27:3968-3975. doi: 10.1111/mec.14844.

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. doi: 10.1111/1755-0998.12668.

Aylagas E, Borja Á, Rodríguez-Ezpeleta N (2014) Environmental status assessment using DNA metabarcoding: Towards a genetics based marine biotic index (gAMBI). PLoS ONE 9:e90529. doi: 10.1371/journal.pone.0090529.

Aylagas E, Borja Á, Irigoien X, Rodríguez-Ezpeleta N (2016) Benchmarking DNA metabarcoding for biodiversity-based monitoring and assessment. Frontiers in Marine Science 3:96. doi: 10.3389/fmars.2016.00096.

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

Bagley M, Pilgrim E, Knapp M, Yoder C, Santo Domingo J, Banerji A (2019) High-throughput environmental DNA analysis informs a biological assessment of an urban stream. Ecological Indicators 104:378-389. doi: 10.1016/j.ecolind.2019.04.088.

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. doi: 10.1111/j.1365-294X.2012.05519.x.

Barbour MT, Gerritsen J, Snyder B, Stribling J (1999) Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish. Washington, DC: U.S. Environmental Protection Agency.

Barmentlo SH, Schrama M, De Snoo GR, Van Bodegom PM, Nieuwenhuijzen A, Vijver MG (in prep) Experimental evidence for neonicotinoid driven decline in aquatic emerging insects.

Barmentlo SH, Schrama M, Van Bodegom PM, De Snoo GR, Musters CJM, Vijver MG (2019) Neonicotinoids and fertilizers jointly structure naturally assembled freshwater macroinvertebrate communities. Science of the Total Environment 691:36-44. doi: 10.1016/j.scitotenv.2019.07.110.

Barnes MA, Turner CR, Jerde CL, Renshaw MA, Chadderton WL, Lodge DM (2014) Environmental conditions influence eDNA persistence in aquatic systems. Environmental Science & Technology 48:1819-27. doi: 10.1021/es404734p.

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

Barnosky AD, Matzke N, Tomiya S, Wogan GOU, Swartz B, Quental TB, Marshall C, McGuire JL, Lindsey EL, Maguire KC, Mersey B, Ferrer EA (2011) Has the Earth's sixth mass extinction already arrived? Nature 471:51-7. doi: 10.1038/nature09678.

Barsoum N, Bruce C, Forster J, Ji Y-Q, Yu DW (2019) The devil is in the detail: Metabarcoding of arthropods provides a sensitive measure of biodiversity response to forest stand composition compared with surrogate measures of biodiversity. Ecological Indicators 101:313-323. doi: 10.1016/J.ECOLIND.2019.01.023.

Baselga A, Fujisawa T, Crampton-Platt A, Bergsten J, Foster PG, Monaghan MT, Vogler AP (2013) Whole-community DNA barcoding reveals a spatio-temporal continuum of biodiversity at species and genetic levels. Nature Communications 4:1892. doi: 10.1038/ncomms2881.

Beentjes KK, Speksnijder AGCL, Van der Hoorn BB, Van Tol J (2015) DNA barcoding program at Naturalis Biodiversity Center, the Netherlands. Genome 58:193. doi: 10.1139/gen-2015-0087.

Beermann AJ, Zizka VMA, Elbrecht V, Baranov V, Leese F (2018) DNA metabarcoding reveals the complex and hidden responses of chironomids to multiple stressors. Environmental Sciences Europe 30:26. doi: 10.1186/s12302-018-0157-x.

Beermann AJ, Werner M-T, Elbrecht V, Zizka VMA, Leese F (2020) DNA metabarcoding improves the detection of multiple stressor responses of stream invertebrates to increased salinity, fine sediment deposition and reduced flow velocity. Science of The Total Environment:141969. doi: 10.1016/j.scitotenv.2020.141969.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005) GenBank. Nucleic Acids Research 33:D34-8. doi: 10.1093/nar/gki063.

Bergsten J, Bilton DT, Fujisawa T, Elliott M, Monaghan MT, Balke M, Hendrich L, Geijer J, Herrmann J, Foster GN, Ribera I, Nilsson AN, Barraclough TG, Vogler AP (2012) The effect of geographical scale of sampling on DNA barcoding. Systematic Biology 61:851-869. doi: 10.1093/sysbio/sys037.

Berney C, Fahrni J, Pawlowski J (2004) How many novel eukaryotic "kingdoms"? Pitfalls and limitations of environmental DNA surveys. BMC Biology 2:13. doi: 10.1186/1741-7007-2-13.

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. doi: 10.1016/j.biocon.2014.11.029.

Birk S, Bonne W, Borja A, Brucet S, Courrat A, Poikane S, Solimini A, Van De Bund W, Zampoukas N, Hering D (2012) Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. Ecological Indicators 18:31-41. doi: 10.1016/j.ecolind.2011.10.009.

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. doi: 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. doi: 10.1111/1755-0998.12888.

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

Buchner D, Beermann AJ, Laini A, Rolauffs P, Vitecek S, Hering D, Leese F (2019) Analysis of 13,312 benthic invertebrate samples from German streams reveals minor deviations in ecological status class between abundance and presence/absence data. PLoS ONE 14:e0226547. doi: 10.1371/journal.pone.0226547.

Buhay JE (2009) "COI-like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. Journal of Crustacean Biology 29:96-110. doi: 10.1651/08-3020.1.

Bush A, Compson Z, Monk W, Porter T, Steeves R, Emilson E, Gagne N, Hajibabaei M, Roy M, Baird D (2019) Studying ecosystems with DNA metabarcoding: lessons from aquatic biomonitoring. bioRxiv:578591. doi: 10.1101/578591.

Butchart SHM, Walpole M, Collen B, Van Strien A, Scharlemann JPW, Almond REA, Baillie JEM, Bomhard B, Brown C, Bruno J, Carpenter KE, Carr GM, Chanson J, Chenery AM, Csirke J, Davidson NC, Dentener F, Foster M, Galli A, Galloway JN, Genovesi P, Gregory RD, Hockings M, Kapos V, Lamarque J-F, Leverington F, Loh J, McGeoch MA, McRae L, Minasyan A, Hernández Morcillo M, Oldfield TEE, Pauly D, Quader S, Revenga C, Sauer JR, Skolnik B, Spear D, Stanwell-Smith D, Stuart SN, Symes A, Tierney M, Tyrrell TD, Vié J-C, Watson R (2010) Global biodiversity: indicators of recent declines. Science 328:1164-8. doi: 10.1126/science.1187512.

Buxton AS, Groombridge JJ, Zakaria NB, Griffiths RA (2017) Seasonal variation in environmental DNA in relation to population size and environmental factors. Scientific Reports 7:46294. doi: 10.1038/srep46294.

Buxton AS, Groombridge JJ, Griffiths RA (2018) Seasonal variation in environmental DNA detection in sediment and water samples. PLoS ONE 13:e0191737. doi: 10.1371/journal.pone.0191737.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421. doi: 10.1186/1471-2105-10-421.

Carew ME, Pettigrove VJ, Metzeling L, Hoffmann AA (2013) Environmental monitoring using next generation sequencing: rapid identification of macroinvertebrate bioindicator

species. Frontiers in Zoology 10:45. doi: 10.1186/1742-9994-10-45.

Carvalho L, Poikane S, Lyche Solheim A, Phillips G, Borics G, Catalan J, De Hoyos C, Drakare S, Dudley BJ, Järvinen M, Laplace-Treyture C, Maileht K, McDonald C, Mischke U, Moe J, Morabito G, Nõges P, Nõges T, Ott I, Pasztaleniec A, Skjelbred B, Thackeray SJ (2013) Strength and uncertainty of phytoplankton metrics for assessing eutrophication impacts in lakes. Hydrobiologia 704:127-140. doi: 10.1007/s10750-012-1344-1.

Chagnon M, Kreutzweiser D, Mitchell EAD, Morrissey CA, Noome DA, Van Der Sluijs JP (2015) Risks of large-scale use of systemic insecticides to ecosystem functioning and services. Environmental Science and Pollution Research 22:119-134. doi: 10.1007/s11356-014-3277-x.

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

Chamberlain S (2017) BOLD: Interface to Bold Systems 'API'. R package version 0.5.0.

Chará-Serna AM, Epele LB, Morrissey CA, Richardson JS (2019) Nutrients and sediment modify the impacts of a neonicotinoid insecticide on freshwater community structure and ecosystem functioning. Science of the Total Environment 692:1291-1303. doi: 10.1016/j.scitotenv.2019.06.301.

Chariton AA, Ho KT, Proestou D, Bik H, Simpson SL, Portis LM, Cantwell MG, Baguley JG, Burgess RM, Pelletier MM, Perron M, Gunsch C, Matthews RA (2014) A molecular-based approach for examining responses of eukaryotes in microcosms to contaminant-spiked estuarine sediments. Environmental Toxicology and Chemistry 33:359-369. doi: 10.1002/etc.2450.

Chonova T, Kurmayer R, Rimet F, Labanowski J, Vasselon V, Keck F, Illmer P, Bouchez A (2019) Benthic diatom communities in an alpine river impacted by waste water treatment effluents as revealed using DNA metabarcoding. Frontiers in Microbiology 10:653. doi: 10.3389/fmicb.2019.00653.

Civade R, Dejean T, Valentini A, Roset N, Raymond JC, 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. doi: 10.1371/journal.pone.0157366.

Clare EL, Chain FJJ, Littlefair JE, Cristescu ME, Deiner K (2016) The effects of parameter choice on defining molecular operational taxonomic units and resulting ecological analyses of metabarcoding data. Genome 59:981-990. doi: 10.1139/gen-2015-0184.

Clarke RT, Hering D (2006) Errors and uncertainty in bioassessment methods - major results and conclusions from the STAR project and their application using STARBUGS. Hydrobiologia 566:433-439. doi: 10.1007/s10750-006-0079-2.

Collen B, Whitton F, Dyer EE, Baillie JEM, Cumberlidge N, Darwall WRT, Pollock C, Richman NI, Soulsby A-M, Böhm M (2014) Global patterns of freshwater species diversity, threat and endemism. Global Ecology and Biogeography 23:40-51. doi: 10.1111/geb.12096.

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 and Technology 51:9118-9126. doi: 10.1021/acs.est.7b01518.

Cordier T, Forster D, Dufresne Y, Martins CI, Stoeck T, Pawlowski J (2018) Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. Molecular Ecology Resources. doi: 10.1111/1755-0998.12926.

Cordier T, Frontalini F, Cermakova K, Apothéloz-Perret-Gentil L, Treglia M, Scantamburlo E, Bonamin V, Pawlowski J (2019) Multi-marker eDNA metabarcoding survey to assess the environmental impact of three offshore gas platforms in the North Adriatic Sea (Italy). Marine Environmental Research 146:24-34. doi: 10.1016/j.marenvres.2018.12.009.

Corse E, Meglécz E, Archambaud G, Ardisson M, Martin JF, Tougard C, Chappaz R, Dubut V (2017) A from-benchtop-to-desktop workflow for validating HTS data and for taxonomic identification in diet metabarcoding studies. Molecular Ecology Resources 17:e146-e159. doi: 10.1111/1755-0998.12703.

Corvalan C, Hales S, McMichael A (2005) Ecosystems and Human Well-being: Health Synthesis . Washington, DC: Island Press.

Costanza R (1992) Toward an operational definition of ecosystem health. In: Costanza R, Norton BG, Haskell BD eds. Ecosystem Health: New Goals for Environmental Management. Washington, DC: Island Press, 239-256.

Côté IM, Darling ES, Brown CJ (2016) Interactions among ecosystem stressors and their importance in conservation. Proceedings of the Royal Society B: Biological Sciences 283:1-9. doi: 10.1098/rspb.2015.2592.

Creedy TJ, Ng WS, Vogler AP (2019) Toward accurate species-level metabarcoding of arthropod communities from the tropical forest canopy. Ecology and Evolution 9:3105-3116. doi: 10.1002/ece3.4839.

Cristescu ME (2019) Can environmental RNA revolutionize biodiversity science? Trends in Ecology and Evolution 34:694-697. doi: 10.1016/j.tree.2019.05.003.

Curry CJ, Gibson JF, Shokralla S, Hajibabaei M, Baird DJ (2018) Identifying North American freshwater invertebrates using DNA barcodes: are existing COI sequence libraries fit for purpose? Freshwater Science 37:178-189. doi: 10.1086/696613.

Darling JA, Mahon AR (2011) From molecules to management: adopting DNA-based

methods for monitoring biological invasions in aquatic environments. Environmental Research 111:978-88. doi: 10.1016/j.envres.2011.02.001.

Davis JM, Rosemond AD, Eggert SL, Cross WF, Wallace JB (2010) Long-term nutrient enrichment decouples predator and prey production. Proceedings of the National Academy of Sciences 107:121-126. doi: 10.1073/pnas.0908497107.

Deagle BE, Thomas AC, Shaffer AK, Trites AW, Jarman SN (2013) Quantifying sequence proportions in a DNA-based diet study using Ion Torrent amplicon sequencing: which counts count? Molecular Ecology Resources 13:620-33. doi: 10.1111/1755-0998.12103.

Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. Biology Letters 10:20140562. doi: 10.1098/rsbl.2014.0562.

Deagle BE, Thomas AC, McInnes JC, Clarke LJ, Vesterinen EJ, Clare EL, Kartzinel TR, Eveson JP (2019) Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? Molecular Ecology 28:391-406. doi: 10.1111/mec.14734.

Deiner K, Altermatt F (2014) Transport distance of invertebrate environmental DNA in a natural river. PLoS ONE 9:e88786. doi: 10.1371/journal.pone.0088786.

Deiner K, Knapp RA, Boiano DM, May B (2013) Increased accuracy of species lists developed for alpine lakes using morphology and cytochrome oxidase I for identification of specimens. Molecular Ecology Resources 13:820–831. doi: 10.1111/1755-0998.12130.

Dejean T, Valentini A, Duparc A, Pellier-Cuit S, Pompanon F, Taberlet P, Miaud C (2011) Persistence of environmental DNA in freshwater ecosystems. PLoS ONE 6:e23398. doi: 10.1371/journal.pone.0023398.

Dejean T, Valentini A, Miquel C, Taberlet P, Bellemain E, Miaud C (2012) Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American Bullfrog *Lithobates catesbeianus*. Journal of Applied Ecology 49:953-959. doi: 10.1111/j.1365-2664.2012.02171.x.

DeSalle R, Egan MG, Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. Philosophical Transactions of the Royal Society B: Biological Sciences 360:1905-16. doi: 10.1098/rstb.2005.1722.

Dickie IA, Boyer S, Buckley HL, Duncan RP, Gardner PP, Hogg ID, Holdaway RJ, Lear G, Makiola A, Morales SE, Powell JR, Weaver L (2018) Towards robust and repeatable sampling methods in eDNA-based studies. Molecular Ecology Resources 18:940-952. doi: 10.1111/1755-0998.12907.

Van Dijk TC, Van Staalduinen MA, Van der Sluijs JP (2013) Macro-invertebrate decline in surface water polluted with imidacloprid. PLoS ONE 8:e62374. doi: 10.1371/journal.pone.0062374.

Dirzo R, Young HS, Galetti M, Ceballos G, Isaac NJB, Collen B (2014) Defaunation in the Anthropocene. Science 345:401-406. doi: 10.1126/science.1251817.

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

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. doi: 10.1371/journal.pone.0122763.

Donohue I, Donohue LA, Ainín BN, Irvine K (2009) Assessment of eutrophication pressure on lakes using littoral invertebrates. Hydrobiologia 633:105-122. doi: 10.1007/s10750-009-9868-8.

Drew LW (2011) Are we losing the science of taxonomy? BioScience 61:942–946. doi: 10.1525/bio.2011.61.12.4.

Dudgeon D, Arthington AH, Gessner MO, Kawabata Z-I, Knowler DJ, Lévêque C, Naiman RJ, Prieur-Richard A-H, Soto D, Stiassny MLJ, Sullivan CA (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews of the Cambridge Philosophical Society 81:163-82. doi: 10.1017/S1464793105006950.

Edgar RC (2016) UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. Cold Spring Harbor Laboratory. doi: 10.1101/081257.

Eichmiller JJ, Miller LM, Sorensen PW (2016) Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. Molecular Ecology Resources 16:56-68. doi: 10.1111/1755-0998.12421.

Elbrecht V, Taberlet P, Dejean T, Valentini A, Usseglio-Polatera P, Beisel J-N, Coissac E, Boyer F, Leese F (2016) Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. PeerJ 4:e1966. doi: 10.7717/peerj.1966.

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

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

Elbrecht V, Vamos EE, Steinke D, Leese F (2018a) Estimating intraspecific genetic diversity from community DNA metabarcoding data. PeerJ 6:e4644. doi: 10.7717/peerj.4644.

Elbrecht V, Hebert PDN, Steinke D (2018b) Slippage of degenerate primers can cause variation in amplicon length. Scientific Reports 8:10999. doi: 10.1038/s41598-018-29364-z.

Elbrecht V, Bourlat SJ, Lindner A, Noll NW, Sorg M, Zizka VMA, Alexander M (2020) Pooling size sorted malaise trap fractions to maximise taxon recovery with metabarcoding. bioRxiv. doi: 10.1101/2020.06.09.118950.

Elbrecht V, Leese F (2015) Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass - sequence relationships with an innovative metabarcoding protocol. PLoS ONE 10:e0130324. doi: 10.1371/journal.pone.0130324.

Elbrecht V, Leese F (2017) Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. Frontiers in Environmental Science 5:11. doi: 10.3389/fenvs.2017.00011.

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. doi: 10.1016/j.ecolind.2019.01.014.

European Union (2000) 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. Official Journal of the European Parliament L327:1-82. doi: 10.1039/ap9842100196.

Evans NT, Olds BP, Renshaw MA, Turner CR, Li Y, Jerde CL, Mahon AR, Pfrender ME, Lamberti GA, Lodge DM (2016) Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. Molecular Ecology Resources 16:29-41. doi: 10.1111/1755-0998.12433.

Evans NT, Shirey PD, Wieringa JG, Mahon AR, Lamberti GA (2017a) Comparative cost and effort of fish distribution detection via environmental DNA analysis and electrofishing. Fisheries 42:90-99. doi: 10.1080/03632415.2017.1276329.

Evans NT, Li Y, Renshaw MA, Olds BP, Deiner K, Turner CR, Jerde CL, Lodge DM, Lamberti GA, Pfrender ME (2017b) Fish community assessment with eDNA metabarcoding: effects of sampling design and bioinformatic filtering. Canadian Journal of Fisheries and Aquatic Sciences 74:1362-1374. doi: 10.1139/cjfas-2016-0306.

Evers CHM, Van den Broek AJM, Buskens R, Van Leerdam A, Knoben RAE, Van Herpen FCJ (2012) Omschrijving MEP en maatlatten voor sloten en kanalen voor de KRW 2015-2021. STOWA rapport 2012-34.

Ficetola GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental DNA from water samples. Biology Letters 4:423-5. doi: 10.1098/rsbl.2008.0118.

Ficetola GF, Taberlet P, Coissac E (2016) How to limit false positives in environmental DNA and metabarcoding? Molecular Ecology Resources 16:604-607. doi: 10.1111/1755-0998.12508.

Fujii K, Doi H, Matsuoka S, Nagano M, Sato H, Yamanaka H (2019) Environmental DNA metabarcoding for fish community analysis in backwater lakes: A comparison of capture methods. PLoS ONE 14:e0210357. doi: 10.1371/journal.pone.0210357.

Garlapati D, Charankumar B, Ramu K, Madeswaran P, Ramana Murthy M V. (2019) A review on the applications and recent advances in environmental DNA (eDNA) metagenomics. Reviews in Environmental Science and Biotechnology 18:389-411. doi: 10.1007/s11157-019-09501-4.

Geiger MF, Moriniere J, Hausmann A, Haszprunar G, Wägele W, Hebert PDN, Rulik B (2016) Testing the Global Malaise Trap Program - How well does the current barcode reference library identify flying insects in Germany? Biodiversity Data Journal 4. doi: 10.3897/BDJ.4.e10671.

Gibson J, Shokralla S, Porter TM, King I, Van Konynenburg S, Janzen DH, Hallwachs W, Hajibabaei M (2014) Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics. Proceedings of the National Academy of Sciences 111:8007-12. doi: 10.1073/pnas.1406468111.

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. doi: 10.1371/journal.pone.0138432.

Gleick PH (1993) Water in Crisis: A Guide to the World's Fresh Water Resources. New York: Oxford University Press, Inc.

Goldberg CS, Pilliod DS, Arkle RS, Waits LP (2011) Molecular detection of vertebrates in stream water: a demonstration using Rocky Mountain Tailed Frogs and Idaho Giant Salamanders. PLoS ONE 6:e22746. doi: 10.1371/journal.pone.0022746.

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. doi: 10.1111/2041-210X.12595.

Goldberg CS, Strickler KM, Fremier AK (2018) Degradation and dispersion limit environmental DNA detection of rare amphibians in wetlands: Increasing efficacy of sampling designs. Science of the Total Environment 633:695-703. doi: 10.1016/j.scitotenv.2018.02.295.

Gosselain V, Descy JP, Everbecq E (1994) The phytoplankton community of the River Meuse, Belgium: seasonal dynamics (year 1992) and the possible incidence of zooplankton grazing. Hydrobiologia 289:179-191. doi: 10.1007/BF00007419.

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. doi: 10.1038/s41598-018-27048-2.

Guardiola M, Wangensteen OS, Taberlet P, Coissac E, Uriz MJ, Turon X (2016) Spatio-temporal monitoring of deep-sea communities using metabarcoding of sediment DNA and RNA. PeerJ 4:e2807. doi: 10.7717/peerj.2807.

Haase P, Murray-Bligh J, Lohse S, Pauls S, Sundermann A, Gunn R, Clarke R (2006) Assessing the impact of errors in sorting and identifying macroinvertebrate samples. Hydrobiologia 566:505-521. doi: 10.1007/s10750-006-0075-6.

Haase P, Pauls SU, Schindehütte K, Sundermann A (2010) First audit of macroinvertebrate samples from an EU Water Framework Directive monitoring program: human error greatly lowers precision of assessment results. Journal of the North American Benthological Society 29:1279-1291.

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

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. doi: 10.1186/1472-6785-12-28.

Van der Hammen H (1992) De macrofauna van Noord-Holland. K.U. Nijmegen.

Hanashiro FTT, Mukherjee S, Souffreau C, Engelen J, Brans KI, Busschaert P, De Meester L (2019) Freshwater bacterioplankton metacommunity structure along urbanization gradients in Belgium. Frontiers in Microbiology 10:743. doi: 10.3389/fmicb.2019.00743.

Hänfling B, Handley LL, 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. doi: 10.1111/mec.13660.

Harper LR, Lawson Handley L, Hahn C, Boonham N, Rees HC, Gough KC, Lewis E, Adams IP, Brotherton P, Phillips S, Hänfling B (2018) Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (*Triturus cristatus*). Ecology and Evolution 8:6330-6341. doi: 10.1002/ece3.4013.

Harper LR, Buxton AS, Rees HC, Bruce K, Brys R, Halfmaerten D, Read DS, Watson H V., Sayer CD, Jones EP, Priestley V, Mächler E, Múrria C, Garcés-Pastor S, Medupin C, Burgess K, Benson G, Boonham N, Griffiths RA, Lawson Handley L, Hänfling B (2019a) Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds. Hydrobiologia

826:25-41. doi: 10.1007/s10750-018-3750-5.

Harper LR, Handley LL, Carpenter AI, Murray-Dickson G, Muri C Di, Macgregor CJ, Logan TW, Law A, Breithaupt T, Read DS, McDevitt A, Hänfling B (2019b) Environmental DNA (eDNA) metabarcoding of pond water as a tool to survey conservation and management priority mammals. bioRxiv:546218. doi: 10.1101/546218.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. Proceedings of the Royal Society B: Biological Sciences 270:313-21. doi: 10.1098/rspb.2002.2218.

Hebert PDN, Ratnasingham S, Zakharov E V., Telfer AC, Levesque-Beaudin V, Milton MA, Pedersen S, Jannetta P, Dewaard JR (2016) Counting animal species with DNA barcodes: Canadian insects. Philosophical Transactions of the Royal Society B: Biological Sciences 371:20150333. doi: 10.1098/rstb.2015.0333.

Heisler J, Glibert PM, Burkholder JM, Anderson DM, Cochlan W, Dennison WC, Dortch Q, Gobler CJ, Heil CA, Humphries E, Lewitus A, Magnien R, Marshall HG, Sellner K, Stockwell DA, Stoecker DK, Suddleson M (2008) Eutrophication and harmful algal blooms: A scientific consensus. Harmful Algae 8:3-13. doi: 10.1016/j.hal.2008.08.006.

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. doi: 10.1016/j.watres.2018.03.003.

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

Hobern D (2020) BIOSCAN: DNA barcoding to accelerate taxonomy and biogeography for conservation and sustainability. Genome:1-4. doi: 10.1139/gen-2020-0009.

Hopkins GW, Freckleton RP (2002) Declines in the numbers of amateur and professional taxonomists: implications for conservation. Animal Conservation 5:245-249. doi: 10.1017/S1367943002002299.

Hubbell SP (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press.

Huemer P, Mutanen M, Sefc KM, Hebert PDN (2014) Testing DNA barcode performance in 1000 species of European Lepidoptera: Large geographic distances have small genetic impacts. PLoS ONE 9:e115774. doi: 10.1371/journal.pone.0115774.

Huson DH, Auch AF, Qi J, Schuster SC (2007) MEGAN analysis of metagenomic data.

Genome Research 17:377-86. doi: 10.1101/gr.5969107.

Jackson RB, Carpenter SR, Dahm CN, McKnight DM, Naiman RJ, Postel SL, Running SW (2001) Water in a changing world. Ecological Applications 11:1027-1045. doi: 10.1890/1051-0761(2001)011[1027:WIACW]2.0.CO;2.

Jackson JK, Battle JM, White BP, Pilgrim EM, Stein ED, Miller PE, Sweeney BW (2014) Cryptic biodiversity in streams: a comparison of macroinvertebrate communities based on morphological and DNA barcode identifications. Freshwater Science 33:312-324. doi: 10.1086/675225.

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. doi: 10.1111/1755-0998.12285.

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

Jeunen GJ, Knapp M, Spencer HG, Taylor HR, Lamare MD, Stat M, Bunce M, Gemmell NJ (2019) Species-level biodiversity assessment using marine environmental DNA metabarcoding requires protocol optimization and standardization. Ecology and Evolution 9:1323-1335. doi: 10.1002/ece3.4843.

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

Jones FC (2008) Taxonomic sufficiency: The influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. Environmental Reviews 16:45-69. doi: 10.1139/A07-010.

Joshi NA, Fass JN (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). Available at https://github.com/najoshi/sickle.

Karr JR (1999) Defining and measuring river health. Freshwater Biology 41:221-234. doi: 10.1046/j.1365-2427.1999.00427.x.

Keeley N, Wood SA, Pochon X (2018) Development and preliminary validation of a multi-trophic metabarcoding biotic index for monitoring benthic organic enrichment. Ecological Indicators 85:1044-1057. doi: 10.1016/j.ecolind.2017.11.014.

Kelly RP, Port JA, Yamahara KM, Martone RG, Lowell N, Thomsen PF, Mach ME, Bennett M, Prahler E, Caldwell MR, Crowder LB (2014) Environmental monitoring. Harnessing DNA to improve environmental management. Science 344:1455-6. doi: 10.1126/science.1251156.

Kelly RP, Closek CJ, O'Donnell JL, Kralj JE, Shelton AO, Samhouri JF (2017) Genetic and manual survey methods yield different and complementary views of an ecosystem. Frontiers in Marine Science 3:283. doi: 10.3389/FMARS.2016.00283.

Kenney MA, Sutton-Grier AE, Smith RF, Gresens SE (2009) Benthic macroinvertebrates as indicators of water quality: The intersection of science and policy. Terrestrial Arthropod Reviews 2:99-128. doi: 10.1163/187498209X12525675906077.

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Research 41:e1. doi: 10.1093/nar/gks808.

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. doi: 10.1016/j.biocon.2014.11.020.

Klymus KE, Marshall NT, Stepien CA (2017) Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. PLoS ONE 12:e0177643. doi: 10.1371/journal.pone.0177643.

Krol L, Van der Hoorn B, Gorsich EE, Trimbos K, Van Bodegom PM, Schrama M (2019) How does eDNA compare to traditional trapping? Detecting mosquito communities in South-African freshwater ponds. Frontiers in Ecology and Evolution 7:260. doi: 10.3389/fevo.2019.00260.

Kvist S (2013) Barcoding in the dark?: A critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge. Molecular Phylogenetics and Evolution 69:39-45. doi: 10.1016/j.ympev.2013.05.012.

Lacoursière-Roussel A, Côté G, Leclerc V, Bernatchez L (2015) Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. Journal of Applied Ecology. doi: 10.1111/1365-2664.12598.

Lacoursière-Roussel A, Howland K, Normandeau E, Grey EK, Archambault P, Deiner K, Lodge DM, Hernandez C, Leduc N, Bernatchez L (2018) eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. Ecology and Evolution 8. doi: 10.1002/ece3.4213.

Langer-Jaesrich M, Köhler HR, Gerhardt A (2010) Assessing toxicity of the insecticide thiacloprid on *Chironomus riparius* (Insecta: Diptera) using multiple end points. Archives of Environmental Contamination and Toxicology 58:963-972. doi: 10.1007/s00244-009-9420-x.

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

Laroche O, Wood SA, Tremblay LA, Ellis JI, Lear G, Pochon X (2018) A cross-taxa study using environmental DNA/RNA metabarcoding to measure biological impacts of offshore oil and gas drilling and production operations. Marine Pollution Bulletin 127:97-107. doi: 10.1016/j.marpolbul.2017.11.042.

Larsson AJM, Stanley G, Sinha R, Weissman IL, Sandberg R (2018) Computational correction of index switching in multiplexed sequencing libraries. Nature Methods 15:305-307. doi: 10.1038/nmeth.4666.

Lawson Handley L, 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. doi: 10.1002/edn3.5.

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

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. doi: 10.1186/1742-9994-10-34.

Levi T, Allen JM, Bell D, Joyce J, Russell JR, Tallmon DA, Vulstek SC, Yang C, Yu DW (2019) Environmental DNA for the enumeration and management of Pacific salmon. Molecular Ecology Resources 19:597–608. doi: 10.1111/1755-0998.12987.

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

Li F, Peng Y, Fang W, Altermatt F, Xie Y, Yang J, Zhang X (2018b) Application of environmental DNA metabarcoding for predicting anthropogenic pollution in rivers. Environmental Science

and Technology 52:11708-11719. doi: 10.1021/acs.est.8b03869.

Lim NKM, Tay YC, Srivathsan A, Tan JWT, Kwik JTB, Baloğlu B, Meier R, Yeo DCJ (2016) Next-generation freshwater bioassessment: eDNA metabarcoding with a conserved metazoan primer reveals species-rich and reservoir-specific communities. Royal Society Open Science 3:160635. doi: 10.1098/rsos.160635.

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

Lodge DM, Turner CR, Jerde CL, Barnes MA, Chadderton L, Egan SP, Feder JL, Mahon AR, Pfrender ME (2012) Conservation in a cup of water: estimating biodiversity and population abundance from environmental DNA. Molecular Ecology 21:2555-8. doi: 10.1111/j.1365-294X.2012.05600.x.

Lugg WH, Griffiths J, Van Rooyen AR, Weeks AR, Tingley R (2018) Optimal survey designs for environmental DNA sampling. Methods in Ecology and Evolution 9:1049-1059. doi: 10.1111/2041-210X.12951.

Macher JN, Salis RK, Blakemore KS, Tollrian R, Matthaei CD, Leese F (2016) Multiple-stressor effects on stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. Ecological Indicators 61:159-169. doi: 10.1016/j.ecolind.2015.08.024.

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

Mächler E, Deiner K, Steinmann P, Altermatt F (2014) Utility of environmental DNA for monitoring rare and indicator macroinvertebrate species. Freshwater Science 33:1174-1183.

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 and Technology 50:305-312. doi: 10.1021/acs.est.5b04188.

Magoč T, Salzberg SL (2011) FLASH: Fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957-2963. doi: 10.1093/bioinformatics/btr507.

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

Marshall JC, Steward AL, Harch BD (2006) Taxonomic resolution and quantification of freshwater macroinvertebrate samples from an Australian dryland river: the benefits and costs of using species abundance data. Hydrobiologia 572:171-194. doi: 10.1007/s10750-005-9007-0.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10. doi: 10.14806/ej.17.1.200.

Matson PA, Parton WJ, Power AG, Swift MJ (1997) Agricultural intensification and ecosystem properties. Science 277:504-509. doi: 10.1126/science.277.5325.504.

Matsuhashi S, Doi H, Fujiwara A, Watanabe S, Minamoto T (2016) Evaluation of the environmental DNA method for estimating distribution and biomass of submerged aquatic plants. PLoS ONE 11:e0156217. doi: 10.1371/journal.pone.0156217.

Merkes CM, McCalla SG, Jensen NR, Gaikowski MP, Amberg JJ (2014) Persistence of DNA in carcasses, slime and avian feces may affect interpretation of environmental DNA data. PLoS ONE 9:e113346. doi: 10.1371/journal.pone.0113346.

Meusnier I, Singer GAC, Landry J-F, Hickey DA, Hebert PDN, Hajibabaei M (2008) A universal DNA mini-barcode for biodiversity analysis. BMC Genomics 9:214. doi: 10.1186/1471-2164-9-214.

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. doi: 10.1098/rsos.150088.

Van der Molen D, Pot R, Evers C, Van Herpen F, Van Nieuwerburgh L (2016) Referenties en maatlatten voor natuurlijke watertypen voor de KRW 2015-2021. STOWA rapport 2012-31.

Montagud D, Soria JM, Soria-Perpiñà X, Alfonso T, Vincente E (2018) A comparative study of four indexes based on zooplankton as trophic state indicators in reservoirs. Limnetica 38. doi: 10.23818/limn.38.06.

Morrissey CA, Mineau P, Devries JH, Sanchez-Bayo F, Liess M, Cavallaro MC, Liber K (2015) Neonicotinoid contamination of global surface waters and associated risk to aquatic invertebrates: A review. Environment International 74:291-303. doi: 10.1016/j.envint.2014.10.024.

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. doi: 10.1038/s41598-019-48984-7.

Moyer GR, Díaz-Ferguson E, Hill JE, Shea C (2014) Assessing environmental DNA detection in controlled lentic systems. PLoS ONE 9:e103767. doi: 10.1371/journal.pone.0103767.

Nicholson A, McIsaac D, MacDonald C, Gec P, Mason BE, Rein W, Wrobel J, De Boer M, Milián-García Y, Hanner RH (2020) An analysis of metadata reporting in freshwater environmental DNA research calls for the development of best practice guidelines. Environmental DNA. doi: 10.1002/edn3.81.

O'Donnell JL, Kelly RP, Shelton AO, Samhouri JF, Lowell NC, Williams GD (2017) Spatial distribution of environmental DNA in a nearshore marine habitat. PeerJ 5:e3044. doi: 10.7717/peerj.3044.

Oksanen J, Kindt R, Legendre P, O'Hara R (2007) Vegan: Community Ecology Package. R package version 2.5-5, https://www.rdocumentation.org/packages/vegan.

Olds BP, Jerde CL, Renshaw MA, Li Y, Evans NT, Turner CR, Deiner K, Mahon AR, Brueseke MA, Shirey PD, Pfrender ME, Lodge DM, Lamberti GA (2016) Estimating species richness using environmental DNA. Ecology and Evolution. doi: 10.1002/ece3.2186.

Pauls SU, Alp M, Bálint M, Bernabò P, Čiampor F, Čiamporová-Zaťovičová Z, Finn DS, Kohout J, Leese F, Lencioni V, Paz-Vinas I, Monaghan MT (2014) Integrating molecular tools into freshwater ecology: developments and opportunities. Freshwater Biology 59:1559-1576. doi: 10.1111/fwb.12381.

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. doi: 10.1016/j.scitotenv.2018.05.002.

Pawluczyk M, Weiss J, Links MG, Egaña Aranguren M, Wilkinson MD, Egea-Cortines M, Egaña Aranguren M, Wilkinson MD, Egea-Cortines M (2015) Quantitative evaluation of bias in PCR amplification and next-generation sequencing derived from metabarcoding samples. Analytical and Bioanalytical Chemistry 407:1841-1848. doi: 10.1007/s00216-014-8435-y.

Pereira AS, Dâmaso-Rodrigues ML, Amorim A, Daam MA, Cerejeira MJ (2018) Aquatic community structure in Mediterranean edge-of-field waterbodies as explained by environmental factors and the presence of pesticide mixtures. Ecotoxicology 27:661-674. doi: 10.1007/s10646-018-1944-2.

Piggott JJ, Townsend CR, Matthaei CD (2015) Reconceptualizing synergism and antagonism among multiple stressors. Ecology and Evolution 5:1538-1547. doi: 10.1002/ece3.1465.

Pilliod DS, Goldberg CS, Arkle RS, Waits LP, Richardson J (2013) Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. Canadian Journal of Fisheries and Aquatic Sciences 70:1123-1130. doi: 10.1139/cjfas-2013-0047.

Pilliod DS, Goldberg CS, Arkle RS, Waits LP (2014) Factors influencing detection of eDNA from a stream-dwelling amphibian. Molecular Ecology Resources 14:109-16. doi: 10.1111/1755-0998.12159.

Pimm SL, Jenkins CN, Abell R, Brooks TM, Gittleman JL, Joppa LN, Raven PH, Roberts CM, Sexton JO (2014) The biodiversity of species and their rates of extinction, distribution, and protection. Science 344:1246752. doi: 10.1126/science.1246752.

Piñol J, Mir G, Gomez-Polo P, Agustí N (2015) Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. Molecular Ecology Resources 15:819-30. doi: 10.1111/1755-0998.12355.

Pisa LW, Amaral-Rogers V, Belzunces LP, Bonmatin JM, Downs CA, Goulson D, Kreutzweiser DP, Krupke C, Liess M, Mcfield M, Morrissey CA, Noome DA, Settele J, Simon-Delso N, Stark JD, Van Der Sluijs JP, Van Dyck H, Wiemers M (2014) Effects of neonicotinoids and fipronil on non-target invertebrates. Environmental Science and Pollution Research 22:68-102. doi: 10.1007/s11356-014-3471-x.

Pochon X, Wood SA, Keeley NB, Lejzerowicz F, Esling P, Drew J, Pawlowski J (2015) Accurate assessment of the impact of salmon farming on benthic sediment enrichment using foraminiferal metabarcoding. Marine Pollution Bulletin 100:370-382. doi: 10.1016/j.marpolbul.2015.08.022.

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

Pompanon F, Deagle BE, Symondson WOC, Brown DS, Jarman SN, Taberlet P (2012) Who is eating what: diet assessment using next generation sequencing. Molecular Ecology 21:1931-50. doi: 10.1111/j.1365-294X.2011.05403.x.

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. doi: 10.1038/s41598-018-28424-8.

Porter TM, Hajibabaei M (2018) Automated high throughput animal CO1 metabarcode classification. Scientific Reports 8:219675. doi: 10.1038/s41598-018-22505-4.

Pot R (2015) QBWat, programma voor beoordeling van de biologische waterkwaliteit volgens de Nederlandse maatlatten voor de Kaderrichtlijn Water, versie 5.33.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Research 41:D590-D596. doi: 10.1093/nar/gks1219.

Raby M, Nowierski M, Perlov D, Zhao X, Hao C, Poirier DG, Sibley PK (2018) Acute toxicity of 6 neonicotinoid insecticides to freshwater invertebrates. Environmental Toxicology and Chemistry 37:1430-1445. doi: 10.1002/etc.4088.

Rapport DJ (1989) What constitutes ecosystem health? Perspectives in Biology and Medicine 33:120-132. doi: 10.1353/pbm.1990.0004.

Ratnasingham S, Hebert PDN (2007) BOLD: the Barcode of Life Data System (www.barcodinglife.org). Molecular Ecology Notes 7:355-364. doi: 10.1111/j.1471-8286.2006.01678.x.

Rees HC, Baker CA, Gardner DS, Maddison BC, Gough KC (2017) The detection of great crested newts year round via environmental DNA analysis. BMC Research Notes 10:327. doi: 10.1186/s13104-017-2657-y.

Regier HA (1993) The notion of natural and cultural integrity. In: Woodley S, Kay J eds. Ecological Integrity and the Management of Ecosystems. Delray Beach, FL: St Lucie Press, 3-18.

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. doi: 10.1111/1755-0998.12281.

Rockström J, Steffen W, Noone K, Persson A, Chapin FSI, Lambin E, Lenton TM, Scheffer M, Folke C, Schellnhuber HJ, Nykvist B, Wit CA de, Hughes T, Van der Leeuw S, Rodhe H, Sörlin S, Snyder PK, Costanza R, Svedin U, Falkenmark M, Karlberg L, Corell RW, Fabry VJ, Hansen J, Walker B, Liverman D, Richardson K, Crutzen P, Foley J (2009) Planetary boundaries: exploring the safe operating space for humanity. Ecology and Society 14:32.

Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: a versatile open source tool for metagenomics. PeerJ 4:e2584. doi: 10.7717/peerj.2584.

Roussel J-M, Paillisson J-M, Tréguier A, Petit E (2015) The downside of eDNA as a survey tool in water bodies. Journal of Applied Ecology 52:n/a-n/a. doi: 10.1111/1365-2664.12428.

RStudio (2015) RStudio: Integrated development environment for R (Version 0.99.902).

Sánchez-Bayo F, Goka K, Hayasaka D (2016) Contamination of the aquatic environment with neonicotinoids and its implication for ecosystems. Frontiers in Environmental Science 4:71. doi: 10.3389/fenvs.2016.00071.

Sansom BJ, Sassoubre LM (2017) Environmental DNA (eDNA) shedding and decay rates to model freshwater mussel eDNA transport in a river. Environmental Science and Technology 51:14244-14253. doi: 10.1021/acs.est.7b05199.

Schmidt-Kloiber A, Hering D (2015) www.freshwaterecology.info - An online tool that unifies, standardises and codifies more than 20,000 European freshwater organisms and their ecological preferences. Ecological Indicators 53:271-282. doi: 10.1016/j.ecolind.2015.02.007.

Schmidt BR, Kéry M, Ursenbacher S, Hyman OJ, Collins JP (2013) Site occupancy models in

the analysis of environmental DNA presence/absence surveys: A case study of an emerging amphibian pathogen. Methods in Ecology and Evolution 4:646-653. doi: 10.1111/2041-210X.12052.

Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863-864. doi: 10.1093/bioinformatics/btr026.

Schneider J, Valentini A, Dejean T, Montarsi F, Taberlet P, Glaizot O, Fumagalli L (2016) Detection of invasive mosquito vectors using environmental DNA (eDNA) from water samples. PLoS ONE 11:e0162493. doi: 10.1371/journal.pone.0162493.

Schrama M, Barmentlo SH, Hunting ER, Van Logtestijn RSP, Vijver MG, Van Bodegom PM (2017) Pressure-induced shifts in trophic linkages in a simplified aquatic food web. Frontiers in Environmental Science 5:75. doi: 10.3389/fenvs.2017.00075.

Schreiner VC, Szöcs E, Bhowmik AK, Vijver MG, Schäfer RB (2016) Pesticide mixtures in streams of several European countries and the USA. Science of the Total Environment 573:680-689. doi: 10.1016/j.scitotenv.2016.08.163.

Schwarzenbach RP, Escher BI, Fenner K, Hofstetter TB, Johnson CA, Von Gunten U, Wehrli B (2006) The challenge of micropollutants in aquatic systems. Science 313:1072-1077. doi: 10.1126/science.1127291.

Scrimgeour GJ, Wicklum D (1996) Aquatic ecosystem health and integrity: problems and potential solutions. Journal of the North American Benthological Society 15:254. doi: 10.2307/1467951.

Seymour JR, Amin SA, Raina JB, Stocker R (2017) Zooming in on the phycosphere: The ecological interface for phytoplankton-bacteria relationships. Nature Microbiology 2:17065. doi: 10.1038/nmicrobiol.2017.65.

Shaw JLA, Clarke LJ, Wedderburn SD, Barnes TC, Weyrich LS, Cooper A (2016) Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. Biological Conservation 197:131-138. doi: 10.1016/j.biocon.2016.03.010.

Shokralla S, Singer GAC, Hajibabaei M (2010) Direct PCR amplification and sequencing of specimens' DNA from preservative ethanol. BioTechniques 48:233-4. doi: 10.2144/000113362.

Sigsgaard EE, Carl H, Møller PR, Thomsen PF (2015) Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. Biological Conservation 183:46-52. doi: 10.1016/j.biocon.2014.11.023.

Sigsgaard EE, Nielsen IB, Carl H, Krag MA, Knudsen SW, Xing Y, Holm-Hansen TH, Møller PR, Thomsen PF (2017) Seawater environmental DNA reflects seasonality of a coastal fish community. Marine Biology 164:128. doi: 10.1007/s00227-017-3147-4.

Sigsgaard EE, Jensen MR, Winkelmann IE, Møller PR, Hansen MM, Thomsen PF (2020) Population-level inferences from environmental DNA - Current status and future perspectives. Evolutionary Applications 13:245-262. doi: 10.1111/eva.12882.

Smart AS, Tingley R, Weeks AR, Van Rooyen AR, McCarthy MA (2015) Environmental DNA sampling is more sensitive than a traditional survey technique for detecting an aquatic invader. Ecological Applications 25:1944-1952. doi: 10.1890/14-1751.1.

Smart AS, Weeks AR, Van Rooyen AR, Moore A, McCarthy MA, Tingley R (2016) Assessing the cost-efficiency of environmental DNA sampling. Methods in Ecology and Evolution 7:1291-1298. doi: 10.1111/2041-210X.12598.

Song H, Buhay JE, Whiting MF, Crandall KA (2008) Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proceedings of the National Academy of Sciences 105:13486-13491. doi: 10.1073/pnas.0803076105.

De Souza LS, Godwin JC, Renshaw MA, Larson E (2016) Environmental DNA (eDNA) detection probability is influenced by seasonal activity of organisms. PLoS ONE 11:e0165273. doi: 10.1371/journal.pone.0165273.

Spear MJ, Embke HS, Krysan PJ, Vander Zanden MJ (2020) Application of eDNA as a tool for assessing fish population abundance. Environmental DNA:edn3.94. doi: 10.1002/edn3.94.

Spear SF, Groves JD, Williams LA, Waits LP (2015) Using environmental DNA methods to improve detectability in a hellbender (*Cryptobranchus alleganiensis*) monitoring program. Biological Conservation 183:38-45. doi: 10.1016/j.biocon.2014.11.016.

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. doi: 10.1111/2041-210X.12683.

Šporka F, Vlek HE, Bulánková E, Krno I (2006) Influence of seasonal variation on bioassessment of streams using macroinvertebrates. Hydrobiologia 566:543-555. doi: 10.1007/s10750-006-0073-8.

Stat M, John J, DiBattista JD, Newman SJ, Bunce M, Harvey ES (2019) Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. Conservation Biology 33:196-205. doi: 10.1111/cobi.13183.

Stein ED, Martinez MC, Stiles S, Miller PE, Zakharov E V (2014) Is DNA barcoding actually cheaper and faster than traditional morphological methods: results from a survey of freshwater bioassessment efforts in the United States? PLoS ONE 9:e95525. doi: 10.1371/journal.pone.0095525.

Stewart K, Ma H, Zheng J, Zhao J (2017) Using environmental DNA to assess population-wide spatiotemporal reserve use. Conservation Biology 31:1173-1182. doi: 10.1111/cobi.12910.

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

Stoeck T, Frühe L, Forster D, Cordier T, Martins CIM, Pawlowski J (2018) Environmental DNA metabarcoding of benthic bacterial communities indicates the benthic footprint of salmon aquaculture. Marine Pollution Bulletin 127:139-149. doi: 10.1016/j.marpolbul.2017.11.065.

Stoeckle MY, Soboleva L, Charlop-Powers Z (2017) Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. PLoS ONE 12:e0175186. doi: 10.1371/journal.pone.0175186.

STOWA (2014) Handboek Hydrobiologie. Biologisch onderzoek voor de ecologische beoordeling van Nederlandse zoete en brakke oppervlaktewateren. Deels aangepaste versie.

Strayer DL (2006) Challenges for freshwater invertebrate conservation. Journal of the North American Benthological Society 25:271-287. doi:10.1899/0887-3593(2006)25[271:CFFIC]2.0.CO;2.

Strayer DL, Dudgeon D (2010) Freshwater biodiversity conservation: recent progress and future challenges.

Stribling JB, Pavlik KL, Holdsworth SM, Leppo EW (2008) Data quality, performance, and uncertainty in taxonomic identification for biological assessments. Journal of the North American Benthological Society 27:906-919. doi: 10.1899/07-175.1.

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. doi: 10.1016/j.biocon.2014.11.038.

Suikkanen S, Pulina S, Engström-Öst J, Lehtiniemi M, Lehtinen S, Brutemark A (2013) Climate change and eutrophication induced shifts in northern summer plankton communities. PLoS ONE 8:e66475. doi: 10.1371/journal.pone.0066475.

Sumon KA, Ritika AK, Peeters ETHM, Rashid H, Bosma RH, Rahman MS, Fatema MK, Van den Brink PJ (2018) Effects of imidacloprid on the ecology of sub-tropical freshwater microcosms. Environmental Pollution 236:432-441. doi: 10.1016/j.envpol.2018.01.102.

Suter GW (1993) A critique of ecosystem health concepts and indexes. Environmental Toxicology and Chemistry 12:1533-1539. doi: 10.1002/etc.5620120903.

Sweeney BW, Battle JM, Jackson JK, Dapkey T (2011) Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? Journal of the North American Benthological Society 30:195-216. doi: 10.1899/10-016.1.

Sworobowicz L, Grabowski M, Mamos T, Burzyński A, Kilikowska A, Sell J, Wysocka A (2015) Revisiting the phylogeography of *Asellus aquaticus* in Europe: insights into cryptic diversity and spatiotemporal diversification. Freshwater Biology 60:1824-1840. doi: 10.1111/fwb.12613.

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012a) Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology 21:2045-50. doi: 10.1111/j.1365-294X.2012.05470.x.

Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012b) Environmental DNA. Molecular Ecology 21:1789-93. doi: 10.1111/j.1365-294X.2012.05542.x.

Takahara T, Minamoto T, Doi H (2013) Using environmental DNA to estimate the distribution of an invasive fish species in ponds. PLoS ONE 8:e56584. doi: 10.1371/journal.pone.0056584.

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

Thomsen PF, Kielgast J, Iversen LL, Møller PR, Rasmussen M, Willerslev E (2012a) Detection of a diverse marine fish fauna using environmental DNA from seawater samples. PLoS ONE 7:e41732. doi: 10.1371/journal.pone.0041732.

Thomsen PF, Kielgast J, Iversen LL, Wiuf C, Rasmussen M, Gilbert MTP, Orlando L, Willerslev E (2012b) Monitoring endangered freshwater biodiversity using environmental DNA. Molecular Ecology 21:2565-73. doi: 10.1111/j.1365-294X.2011.05418.x.

Thomsen PF, Willerslev E (2015) Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:4-18. doi: 10.1016/j.biocon.2014.11.019.

Tréguier A, Paillisson J-M, Dejean T, Valentini A, Schlaepfer MA, Roussel J-M (2014) Environmental DNA surveillance for invertebrate species: advantages and technical limitations to detect invasive crayfish *Procambarus clarkii* in freshwater ponds. Journal of Applied Ecology 51:871-879. doi: 10.1111/1365-2664.12262.

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

Turner CR, Miller DJ, Coyne KJ, Corush J (2014b) Improved methods for capture, extraction, and quantitative assay of environmental DNA from Asian bigheaded carp (*Hypophthalmichthys* spp.). PLoS ONE 9:e114329. doi: 10.1371/journal.pone.0114329.

Turner CR, Uy KL, Everhart RC (2015) Fish environmental DNA is more concentrated in aquatic sediments than surface water. Biological Conservation 183:93-102. doi: 10.1016/j.

biocon.2014.11.017.

Ushio M, Murakami H, Masuda R, Sado T, Miya M, Sakurai S, Yamanaka H, Minamoto T, Kondoh M (2018) Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. Metabarcoding and Metagenomics 2:1-15. doi: 10.1101/113472.

Uthicke S, Lamare M, Doyle JR (2018) eDNA detection of corallivorous seastar (*Acanthaster cf. solaris*) outbreaks on the Great Barrier Reef using digital droplet PCR. Coral Reefs 37:1229-1239. doi: 10.1007/s00338-018-1734-6.

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 (2015) Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Molecular Ecology 25:929-942. doi: 10.1111/mec.13428.

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 and Technology 49:7597-7605. doi: 10.1021/es506158m.

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. doi: 10.7717/peerj.4705.

Weigand H, Beermann AJ, Čiampor F, Costa FO, Csabai Z, Duarte S, Geiger MF, Grabowski M, Rimet F, Rulik B, Strand M, Szucsich N, Weigand AM, Willassen E, Wyler SA, Bouchez A, Borja A, Čiamporová-Zaťovičová Z, Ferreira S, Dijkstra K, Eisendle U, Freyhof J, Gadawski P, Graf W, Haegerbaeumer A, Van der Hoorn BB, Japoshvili B, Keresztes L, Keskin E, Leese F, Macher J, Mamos T, Paz G, Pešić V, Pfannkuchen DM, Pfannkuchen MA, Price BW, Rinkevich B, Teixeira MAL, Várbíró G, Ekrem T (2019) DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. bioRxiv:576553. doi: 10.1101/576553.

Weigand AM, Macher JN (2018) A DNA metabarcoding protocol for hyporheic freshwater meiofauna: Evaluating highly degenerate COI primers and replication strategy. Metabarcoding and Metagenomics 2:e26869. doi: 10.3897/mbmg.2.26869.

Wetzel RG (2001) Limnology: Lake and River Ecosystems. San Diego, CA: Academic Press.

Wilcox TM, McKelvey KS, Young MK, Jane SF, Lowe WH, Whiteley AR, Schwartz MK (2013) Robust detection of rare species using environmental DNA: the importance of primer specificity. PLoS ONE 8:e59520. doi: 10.1371/journal.pone.0059520.

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. doi: 10.1007/s12686-015-0465-z.

Willerslev E, Hansen AJ, Binladen J, Brand TB, Gilbert MTP, Shapiro B, Bunce M, Wiuf C, Gilichinsky DA, Cooper A (2003) Diverse plant and animal genetic records from holocene and pleistocene sediments. Science 300:791-795. doi: 10.1126/science.1084114.

Williams N, Sweetman J (2019) Effects of neonicotinoids on the emergence and composition of chironomids in the Prairie Pothole Region. Environmental Science and Pollution Research 26:3862-3868. doi: 10.1007/s11356-018-3683-6.

Wright-Stow AE, Winterbourn MJ (2003) How well do New Zealand's stream-monitoring indicators, the macroinvertebrate community index and its quantitative variant, correspond? New Zealand Journal of Marine and Freshwater Research 37:461-470. doi: 10.1080/00288330.2003.9517180.

Wu Z, Cai Y, Liu X, Xu CP, Chen Y, Zhang L (2013) Temporal and spatial variability of phytoplankton in Lake Poyang: The largest freshwater lake in China. Journal of Great Lakes Research 39:476-483. doi: 10.1016/j.jglr.2013.06.008.

WWF (2014) The Living Planet Report 2014. doi: 10.1016/j.edurev.2013.01.001.

Yamamuro M, Komuro T, Kamiya H, Kato T, Hasegawa H, Kameda Y (2019) Neonicotinoids disrupt aquatic food webs and decrease fishery yields. Science 366:620-623. doi: 10.1126/science.aax3442.

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 and Evolution 11:173-192. doi: 10.1007/s13127-011-0050-6.

Zimmermann J, Glöckner G, Jahn R, Enke N, Gemeinholzer B (2015) Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. Molecular Ecology Resources 15:526-42. doi: 10.1111/1755-0998.12336.

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

Zizka VMA, Geiger MF, Leese F (2020) DNA metabarcoding of stream invertebrates reveals spatio-temporal variation but consistent status class assessments in a natural and urban river. Ecological Indicators 115. doi: 10.1016/j.ecolind.2020.106383.

SUMMARY

Freshwater is an important resource, not only for the ecosystem services it provides to humankind, but also as a habitat for many species. Freshwater ecosystems are, however, at great risk of species decline due to habitat loss and modification, pollution and over-exploitation, and invasive alien species. European and national regulation dictate the monitoring of freshwater quality in the Water Framework Directive (WFD). Quality assessment of freshwater makes use of different criteria that define its health and impact status. The biological elements of these assessments focus on the organisms living in freshwater systems, such as fish, macroinvertebrates, and plants. Traditional monitoring of quality relies on labor-intensive and expensive collection and morphological identification of specimens. Recent developments in molecular techniques allow for easier identification through (meta)barcoding and species detection using environmental DNA (eDNA).

Comparative studies assessing both traditional methodology and DNA-based analyses are important in the transition from the former to the latter for ecological quality ratio (EQR) assessments. Since traditional methods have been set up with limitations of these methods in mind, DNA-based techniques do not necessarily line up with requirements set forth in the WFD. One important point of contention is the use of abundance data in EQR scoring. DNA metabarcoding methods are prone to technical biases that obfuscate the original biomass or specimen counts. In Chapter 2 we investigated the influence of abundance data on the EQR scoring according to the Dutch EQR calculation system. By comparing EQR scores on historical data with abundances removed to the original EQR scores of those samples, we found that abundance data was of limited influence. The strong correlation between the scores with and without abundance would allow for DNA-based species lists to be used for WFD assessments, opening the way for the introduction of barcoding-based methods into routine quality monitoring.

Studies comparing traditional morphology-based assessments with DNA and eDNA metabarcoding also highlight differences caused by underlying issues such as the difficulty of identification with morphological keys for certain taxa, or the inability to distinguish other taxa using DNA barcodes. Chapter 3 illustrates these issues, where we obtained a similar number of taxa using DNA metabarcoding as were observed using morphological assessments. There were, however, large

differences between the taxa lists of both methods, with less than 60% overlap between the two. Simple taxonomic sorting alleviated some of the before-mentioned technical biases, and our results clearly show the effects of preferential amplification in complex bulk samples. Impact of the differences between the species lists on EQR scoring were considerable, but DNA metabarcoding allowed for much more detailed information in morphologically hard to identify taxonomic groups, such as chironomids. Integration of DNA-based identifications for such groups would allow for more accurate EQR status assessments.

In addition to DNA-based identifications, environmental DNA is a game-changer for freshwater assessments, as it allows for simpler, cheaper, and more easily standardized sampling. There are, however, many unanswered questions regarding the behavior—or "ecology"—of eDNA within the aquatic environment. In Chapter 4 we explored the impact of replicates in various steps of the analyses on richness estimations and community patterns. While the effect of PCR replicates was limited, the effect of sampling replicates was considerable. Dissimilarities between replicates were high, revealing the heterogenous distribution of eDNA within a waterbody. Furthermore, the weekly sampling of the same two study sites showed that temporal replicates were even more dissimilar than the spatial replicates. This suggests that turnover effects might be more important for the dynamics of eDNA than its spatial heterogeneity. Many studies fail to incorporate these dissimilarities into their study design, meaning that between-site comparisons done over longer time periods are probably affected by inflated dissimilarities.

One of the main issues with using environmental DNA is that it produces many DNA profiles (Molecular Operational Taxonomic Units, or MOTUs) that cannot be directly linked to a known taxon, due to incomplete databases, but also due to undescribed diversity that has not been morphologically observed. However, eDNA still lends itself for comparative studies that look at patterns between, for example, impacted and non-impacted sites. In Chapter 5 we performed such an impact assessment using eDNA, to investigate the effects of the neonicotinoid insecticide thiacloprid and fertilizer, two of the main agricultural stressors on freshwater systems. Using eDNA, we assessed three different taxonomic groups that represented three trophic levels in the ecosystem: bacteria (composers), phytoplankton (primary producers) and chironomids (consumers and key indicator species). This experiment was performed in a unique "Living Lab", allowing for a controlled experiment in a semi-natural environment. Using a full-factorial setup with many replications also allowed for the disentanglement of single stressor effects. By conducting the experiment at the same time as a morphology-based assessment on the same impacted

sites we were also able to compare results directly to traditional methods. For all three groups assessed, similar patterns of stressor impact were observed over time for both stressors, suggesting that agricultural stressors affect the entire food web, either directly or through cascade reactions. The patterns were also consistent with morphological assessments, with a lower number of technical replicates. This shows that the use of multi-marker environmental DNA provides a more comprehensive assessment of stressor impacts on an ecosystem as a whole, with a higher taxonomic resolution than traditional surveys. We also found over a thousand MOTUs that were indicative of stressor absence or presence, some of which can be putative new bio-indicators for both agricultural stress of freshwater.

There are numerous questions that still need answers, as discussed in Chapter 6. Not only on how DNA data is translated into traditional taxa, but also on the optimization of sampling strategies and the ecology of eDNA. The research presented in this thesis, however, along with the increasing number of publications on similar topics, show that DNA-based methods have great potential for freshwater quality monitoring and impact assessment. The incorporation of these techniques will contribute to a better ecosystem understanding and allow for more effective monitoring and management of freshwater systems, safeguarding the ecosystem services provided to humankind. For successful integration into ecosystem assessments, it is also important in this perspective to involve monitoring agencies and policy makers, by demonstrating the possibilities of DNA-based methods and including them in the development of molecular tools.

SAMENVATTING

Zoetwater is een belangrijke hulpbron, niet alleen voor de ecosysteemdiensten die het aan de mensheid levert, maar ook als een leefgebied voor veel planten- en diersoorten. Zoetwaterecosystemen lopen echter een groot risico op achteruitgang in soortendiversiteit als gevolg van verlies van leefgebieden, vervuiling en overexploitatie, en de introductie van invasieve soorten. De monitoring van zoetwaterkwaliteit wordt voorgeschreven door de Kaderrichtlijn Water (KRW), gebaseerd op Europese en nationale regelgeving. Kwaliteitsbeoordeling van zoetwater maakt gebruik van verschillende criteria die de gezondheidstoestand en impactstatus bepalen. De biologische elementen van deze beoordelingen zijn gericht op organismen die in zoetwatersystemen leven, zoals vissen, ongewervelden en planten. Traditionele beoordelingsmethoden zijn afhankelijk van arbeidsintensieve en daarmee dure methoden om organismen te verzamelen en te determineren. Recente ontwikkelingen in moleculaire technieken zorgen voor eenvoudigere determinatie via barcodes en soortdetectie met behulp van "environmental" DNA (eDNA).

Vergelijkend onderzoek waarbij zowel traditionele methodologie als op DNA gebaseerde analyses beoordeeld worden is belangrijk in de overgang naar moleculaire methoden voor beoordelingen van ecologische kwaliteitsratio's (EKR). Aangezien traditionele methoden zijn opgezet met kennis van de bestaande beperkingen, zijn op DNA gebaseerde technieken niet noodzakelijkerwijs in overeenstemming met de vereisten van de KRW. Een belangrijk twistpunt is het gebruik van gegevens over de abundanties van organismen bij het berekenen van EKR-scores. Methoden voor DNA metabarcoding zijn vatbaar voor processen die bepaalde reacties bevooroordelen, waardoor de oorspronkelijke biomassa of het aantal individuen lastig te achterhalen is. In Hoofdstuk 2 onderzochten we de invloed van abundantiegegevens op de EKR-score volgens het Nederlandse EKR-berekeningssysteem. Door EKR-scores berekend met historische gegevens te vergelijken met EKR-berekeningen waarin de aantallen individuen waren verwijderd (en dus EKR's alleen berekend op basis van aan- en afwezigheid van soorten) ontdekten we dat abundantiegegevens slechts van beperkte invloed waren. De sterke correlatie tussen de scores met en zonder aantallen individuen zou het mogelijk maken om op DNA gebaseerde soortenlijsten te gebruiken voor KRW-beoordelingen, hetgeen de weg vrijmaakt voor de introductie van op DNA gebaseerde identificaties in routinematige kwaliteitsmonitoring.

Studies die traditionele op morfologie gebaseerde beoordelingen vergelijken met DNA en eDNA metabarcoding benadrukken ook verschillen die worden veroorzaakt door onderliggende problemen, zoals lastig te determineren taxa of incomplete determinatiesleutels, of het niet kunnen onderscheiden van soorten op basis van de gebruikte DNA-barcodes. Hoofdstuk 3 illustreert deze problemen, in een studie waarbij DNA metabarcoding en traditionele determinaties vergelijkbare aantallen taxa opleverden. Er waren echter grote verschillen tussen de taxalijsten van beide methoden, met minder dan 60% overlap tussen de twee. Eenvoudige taxonomische sortering van de monsters verlichtte enkele van de eerdergenoemde technische problemen, en onze resultaten laten duidelijk de effecten van preferentiële amplificatie in complexe bulkmonsters zien. Het effect van de verschillen tussen de soortenlijsten op de EKR-scores was aanzienlijk, maar DNA metabarcoding leverde meer informatie op voor taxonomische groepen die lastig zijn te determineren met morfologische kenmerken, zoals dansmuggen. Integratie van op DNA gebaseerde identificaties voor dergelijke groepen zou nauwkeurigere EKR-statusbeoordelingen mogelijk maken.

Naast de op DNA gebaseerde identificaties, is eDNA een methode die van grote invloed kan zijn op de praktijk van zoetwaterkwaliteitsbeoordelingen, omdat het eenvoudigere, goedkopere en makkelijker te standaardiseren bemonstering mogelijk maakt. Er zijn echter veel onbeantwoorde vragen over het gedrag—of de "ecologie"—van eDNA in het water. In Hoofdstuk 4 hebben we de impact van replica's in verschillende stappen van de analyse op inschattingen van soortenrijkdom en gemeenschapspatronen onderzocht. Hoewel het effect van PCR-replica's beperkt was, was het effect van het nemen van replica's tijdens de bemonstering van water aanzienlijk. De verschillen tussen replica's waren groot, wat de heterogene distributie van eDNA binnen een waterlichaam laat zien. Bovendien toonde een wekelijkse bemonstering van dezelfde twee monsterlocaties aan dat replica's door de tijd meer van elkaar verschillen dan monsterreplica's genomen op één moment. Dit geeft aan dat de turnover van biodiversiteit een groter effect heeft op de samenstelling van het eDNA dan de ruimtelijke heterogeniteit. Veel studies nemen deze effecten echter niet mee in hun studieopzet, wat betekent dat vergelijkingen tussen locaties die niet op een gelijk tijdstip zijn bemonsterd waarschijnlijk te maken hebben met overschattingen van de verschillen in diversiteit tussen deze locaties.

Een van de belangrijkste problemen bij het gebruik van eDNA is dat het veel DNA-profielen (Molecular Operational Taxonomic Units ofwel MOTU's) kan opleveren die niet direct kunnen worden gekoppeld aan een bekende soort of taxon, vanwege onvolledige databases, maar ook vanwege onbeschreven diversiteit. eDNA leent zich echter nog steeds goed voor vergelijkend onderzoek tussen locaties, bijvoorbeeld

maatregel-effect studies. In Hoofdstuk 5 hebben we een impactanalyse uitgevoerd met eDNA, om de effecten van de neonicotinoïde insecticide thiacloprid en kunstmest te onderzoeken, twee van de belangrijkste landbouwstressoren op zoetwatersystemen. Met behulp van eDNA hebben we drie verschillende taxonomische groepen beoordeeld die drie trofische niveaus in het ecosysteem vertegenwoordigden: bacteriën (reducenten), fytoplankton (primaire producenten) en dansmuggen (consumenten en sleutelindicatoren). Dit experiment werd uitgevoerd in een uniek "Living Lab", waarin een gecontroleerd experiment in een semi-natuurlijke omgeving mogelijk werd gemaakt. Het gebruik van een volledig factoriële proefopzet met replica's maakte het ook mogelijk om de effecten van stressoren afzonderlijk te beoordelen. Door het experiment gelijktijdig met een op morfologie gebaseerde beoordeling van dezelfde opzet uit te voeren, konden we de DNA resultaten eveneens rechtstreeks vergelijken met traditionele methoden. Voor alle drie de onderzochte groepen werden door de tijd heen vergelijkbare patronen waargenomen voor beide stressfactoren, wat suggereert dat de onderzochte landbouwstressoren het gehele voedselweb beïnvloeden, hetzij direct, hetzij via cascadereacties. De patronen waren daarnaast ook consistent met morfologische beoordelingen, zelfs met een lager aantal replica monsters. Dit toont aan dat het gebruik van dit soort eDNA methoden op meerdere trofische niveaus een uitgebreidere beoordeling geeft van de impact van stressfactoren op een ecosysteem als geheel, met een hogere taxonomische resolutie dan traditionele onderzoeken. We vonden een groot aantal mogelijke nieuwe bio-indicatoren voor landbouwstress, in de vorm van meer dan duizend DNA-profielen die indicatief bleken voor de aan- of afwezigheid van de stressoren.

Er zijn nog tal van onbeantwoorde vragen omtrent de toepassing van aquatische DNA-methoden. Niet alleen over hoe DNA-gegevens moeten worden vertaald naar traditionele taxa, maar ook over de optimalisatie van bemonsteringsstrategieën en de gedragingen van eDNA in het water. Het onderzoek dat in dit proefschrift wordt gepresenteerd laat echter, samen met een almaar toenemend aantal publicaties over vergelijkbare onderwerpen, zien dat op DNA gebaseerde methoden een groot potentieel hebben voor monitoring van zoetwaterkwaliteit en effectbeoordelingen in dergelijke systemen. De toepassing van deze technieken zal bijdragen aan een beter begrip van ecosystemen, en effectievere monitoring en beheer van zoetwatersystemen waarbij we duurzaam gebruik kunnen blijven maken van ecosysteemdiensten. Voor een succesvolle integratie in ecosysteembeoordelingen is het ook belangrijk om waterbeheerders en beleidsmakers bij deze omschakeling te betrekken, door de mogelijkheden van op DNA gebaseerde methoden te demonstreren en hen te betrekken bij de ontwikkeling van nieuwe monitoringstechnieken.

CIRRUCULUM VITAE

Kevin Beentjes was born on the 2nd of May 1987 in Zaanstad in the Netherlands. After finishing high school at RSG 't Slingerbos in Harderwijk in 2005, he started his BSc in Biomedical Sciences at the University of Amsterdam, from which he graduated cum laude. He continued his education with the MSc Forensic Science at the same university. The subject of his MSc research project and thesis at the Zoological Museum of Amsterdam, under the supervision of Herman de Jong, was the use of Trichoceridae in forensic entomology, which included DNA barcoding of prevalent Dutch species for molecular identification. After graduating in 2010, he started working at Naturalis Biodiversity Center in Leiden as a research technician in the DNA barcoding campaign funded by the Fonds Economische Structuurversterking, working on sponges, spiders and many other taxonomic groups. In 2015, he started his PhD on the subject of the integration of molecular techniques in water quality assessments at the Naturalis Biodiversity Center and the University of Leiden, under the supervision of Menno Schilthuizen and Arjen Speksnijder, and funded by the Gieskes-Strijbis Fonds. In 2019, while still in the process of writing his thesis, he continued his work at Naturalis Biodiversity Center, where he is involved in applied biomonitoring projects through the BioMon spin-off and collaborations with other institutions, as well as the ARISE project, which will continue the work of creating a genetic reference for all Dutch flora and fauna.

LIST OF PUBLICATIONS

PEER-REVIEWED

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. doi: 10.1371/journal.pone.0226527.

Beentjes KK, Speksnijder AGCL, Schilthuizen M, Hoogeveen M, Van der Hoorn BB (2019) The effects of spatial and temporal replicate sampling on eDNA metabarcoding. PeerJ 7:e7335. doi: 10.7717/peerj.7335.

Beentjes KK, Speksnijder AGCL, Schilthuizen M, Schaub BEM, Van der Hoorn BB (2018) The influence of macroinvertebrate abundance on the assessment of freshwater quality in The Netherlands. Metabarcoding and Metagenomics 2:e26744. doi: 10.3897/mbmg.2.26744.

Boer P, Loss AC, Bakker F, **Beentjes KK,** Fisher BL (2020) *Monomorium sahlbergi* Emery, 1898 (Formicidae, Hymenoptera): a cryptic globally introduced species. ZooKeys 979:87–97. doi: 10.3897/zookeys.979.55342.

Van Bochove K, Bakker FT, **Beentjes KK**, Hemerik L, Vos RA, Gravendeel B (2020) Organic matter reduces the amount of detectable environmental DNA in freshwater. Ecology and Evolution. doi: 10.1002/ece3.6123.

Smit JT, Raemakers I, **Beentjes KK** (2018) First host record of *Myopa pellucida* Robineau-Desvoidy (Diptera: Conopidae) identified using DNA barcoding. Zootaxa 4521:593–596. doi: 10.11646/zootaxa.4521.4.8.

Van der Broek R, Smit JT, De Bree E, **Beentjes KK** (2018) *Tolmerus calceatus* (Meigen) confirmed as a valid species separate from *Tolmerus atricapillus* (Fallén) (Diptera: Asilidae). Zootaxa 4508:249–258. doi: 10.11646/zootaxa.4508.2.7.

Smit JT, Dijkstra K-DB, **Beentjes KK**, Miller J, Madden H, Van der Hoorn B (2018) First records of Odonata from Sint Eustatius, Dutch Lesser Antilles. Notulae odonatologicae 9:78–82. doi: 10.5281/zenodo.1481087.

Schilthuizen M, Pimenta LPS, Lammers Y, Steenbergen PJ, Flohil M, Beveridge NGP, Van Duijn PT, Meulblok MM, Sosef N, Van de Ven R, Werring R, **Beentjes KK**, Meijer K, Vos RA, Vrieling K, Gravendeel B, Choi Y, Verpoorte R, Smit C, Beukeboom LW (2016) Incorporation of an invasive plant into a native insect herbivore food web. PeerJ 2016. doi: 10.7717/peerj.1954.

Karremans AP, Pupulin F, Grimaldi D, **Beentjes KK**, Butôt R, Fazzi GE, Kaspers K, Kruizinga J, Roessingh P, Smets EF, Gravendeel B (2015) Pollination of *Specklinia* by nectar-feeding *Drosophila*: The first reported case of a deceptive syndrome employing aggregation pheromones in Orchidaceae. Annals of Botany 116:437–455. doi: 10.1093/aob/mcv086.

Merckx VSFT, Hendriks KP, **Beentjes KK**, Mennes CB, Becking LE, Peijnenburg KTCA, Afendy A, Arumugam N, de Boer H, Biun A, Buang MM, Chen P-P, Chung AYC, Dow R, Feijen FAA, Feijen H, Soest CF, Geml J, Geurts R, Gravendeel B, Hovenkamp P, Imbun P, Ipor I, Janssens SB, Jocqué M, Kappes H, Khoo E, Koomen P, Lens F, Majapun RJ, Morgado LN, Neupane S, Nieser N, Pereira JT, Rahman H, Sabran S, Sawang A, Schwallier RM, Shim P-S, Smit H, Sol N, Spait M, Stech M, Stokvis F, Sugau JB, Suleiman M, Sumail S, Thomas DC, Van Tol J, Tuh FYY, Yahya BE, Nais J, Repin R, Lakim M, Schilthuizen M (2015) Evolution of endemism on a young tropical mountain. Nature 524:347–350. doi: 10.1038/nature14949.

Claessens J, **Beentjes KK,** Heijerman T, Miller J, Gravendeel B (2015) Beobachtungen von *Miarus campanulae* als Bestäuber von *Cephalanthera rubra*. Journal Europäischer Orchideen 47:77–87.

Miller JA, Miller JH, Pham DS, **Beentjes KK** (2014) Cyberdiversity: Improving the informatic value of diverse tropical arthropod inventories. PLoS ONE 9. doi: 10.1371/journal.pone.0115750.

Gravendeel B, De Groot GA, Kik M, **Beentjes KK**, Bergman H, Caniglia R, Cremers H, Fabbri E, Groenenberg D, Grone A, Groot Bruinderink G, Font I, Hakhof J, Harms V, Jansman H, Janssen R, Lammertsma D, Laros I, Linnartz L, Van der Marel D, Mulder J, Van der Mije S, Nieman A, Nowak C, Randi E, Rijks M, Speksnijder A, Vonhof H (2013) The first wolf found in the Netherlands in 150 years was the victim of a wildlife crime. Lutra 56:93–109.

Aliabadian M, **Beentjes KK**, Roselaar CSK, Van Brandwijk H, Nijman V, Vonk R (2013) DNA barcoding of Dutch birds. ZooKeys 365:25–48. doi: 10.3897/zookeys.365.6287.

Miller JA, **Beentjes KK**, Van Helsdingen P, IJland S (2013) Which specimens from a museum collection will yield DNA barcodes? A time series study of spiders in alcohol. ZooKeys 365:245–261. doi: 10.3897/zookeys.365.5787.

Miller JA, Belgers JD, **Beentjes KK**, Zwakhals K, Van Helsdingen P (2013) Spider hosts (Arachnida, Araneae) and wasp parasitoids (Insecta, Hymenoptera, Ichneumonidae, Ephialtini) matched using DNA barcodes. Biodiversity Data Journal 1:e992. doi: 10.3897/bdj.1.e992.

ABSTRACTS

Beentjes KK, Trimbos K, Speksnijder AGCL, Van der Hoorn BB (2019) Environmental DNA-based ecotoxin assessments: a Living Lab. 8th International Barcode of Life Conference. Trondheim, Norway, 17-20 June 2019. (Oral presentation)

Beentjes KK, Speksnijder AGCL, Van der Hoorn BB, Schilthuizen M (2017) Spatial and temporal variation of macroinvertebrate eDNA in Dutch freshwater lakes. 7th International Barcode of Life Conference. Kruger National Park, South Africa, 20-24 November 2017. (Oral presentation)

Beentjes KK, Speksnijder AGCL, Van der Hoorn BB, Schilthuizen M (2016) Environmental DNA and aquatic beetles: integrating molecular tools into freshwater macroinvertebrate community assessments. International Society of Limnology XXXIII Congress. Turin, Italy, 31 July - 5 August 2016. (Oral presentation)

Beentjes KK, Speksnijder AGCL, Van der Hoorn BB, Van Tol J (2015) DNA barcoding program at Naturalis Biodiversity Center, the Netherlands. 6th International Barcode of Life Conference. Guelph, Canada, 18-21 August 2015. (Oral presentation)

Beentjes KK, Van der Hoorn BB (2012) Using a network of amateur specialists for the collection and identification of specimens for DNA Barcoding of Dutch Flora and Fauna. Third European Conference for the Barcode of Life. Brussels, Belgium, 17-20 September 2012. (Oral presentation)

OTHER

Beentjes KK (2020) STOWA Deltafact: DNA-technieken voor waterbeheerders. www.stowa.nl/deltafacts/waterkwaliteit/diversen/dna-technieken-voor-waterbeheerders

Van der Hoorn BB, **Beentjes KK** (2020) Genetische biomonitoring van macrofauna. https://www.h2owaternetwerk.nl/vakartikelen/genetische-biomonitoring-van-macrofauna

ACKNOWLEDGEMENTS

This thesis and the work presented herein would not have been possible without the support of my supervisors, co-authors, colleagues and friends.

I am very grateful for the guidance by Menno Schilthuizen and Arjen Speksnijder, and also by Berry van der Hoorn. In addition, I would like to thank my other co-authors, who were invaluable for the creation of the chapters presented in this thesis. Rob Pastoor processed a lot of samples in the lab. Marten Hoogeveen was vital in the creation of the bioinformatics pipeline we still use today, and it was great to brainstorm about possible new tools to help in the analysis of the massive amounts of data. Bart Schaub provided some useful insights from a water management perspective. Krijn Trimbos, Henrik Barmentlo and Ellen Cieraad provided a lot of statistical and ecological knowhow that was crucial to the impact assessment study in Chapter 5. Michiel Hootsmans provided additional feedback on the Dutch fact sheet that was written parallel to the introduction to this thesis.

I am grateful to Naturalis Biodiversity Center and the University of Leiden for access to all the facilities, the laboratories and the Living Lab in particular. I would also like this opportunity to thank Staatsbosbeheer for access to protected sites in the dunes, and the people at AQUON and Hoogheemraadschap van Rijnland for providing us with many traditionally collected monitoring samples and data. Thanks to Daan Drukker and Bram Koese for organizing the collection of many relevant taxa for the barcode reference database, and in particular Harry Smit and Alexander Klink for providing many specimens of water mites and chironomids, respectively.

Thanks to the staff at the Naturalis lab, Frank, Oscar, Elza, Marcel, and Roland, who provided me with the necessary assistance and supplies, and Sam who helped me out with many eDNA extractions. And a special thanks to my students, who did a lot of the necessary testing of protocols and methods. Jeroen, Rolene, Wouter, Floris, Jelle, David, Kevin, Ruchira, Simone and Merel, I hope you learned as much as I did.

Also thanks to my fellow PhDs, for both scientific discussion and less-than-scientific discussions over lunch, or over beer and pizza. Andres, Dewi, Eka, Esther, Hector, Isolde, Kasper, Larissa, Le Qin, Leon, Lisette, Marcel, Merijn, Richa, Roderick, Sofia, Tiedo, Werner and all the others, it was good having you around! The same goes for all my other past and present colleagues at Naturalis. Thanks to Jan and Camiel for encouraging me to pursue a PhD. I'm grateful to have met a lot of national

and international colleagues and fellow PhDs at the various conferences I attended, and during my field work. Keep on doing the inspiring work you do.

This period of my life would also have been very dull without the other people I am happy to call my friends. More beer and pizza was shared with Frank, Martijn, Dick, Krijn and Jaap. My "online" friends and fellow adventures Bob, Danielle, Jelle, Lyn, and the others provided me with the necessary distractions on my quest to the doctorate, as well as all the folks over at the TeeKay-421 fan club and other otherworldly discussion groups.

Als laatste wil ik graag mijn familie bedanken, in het bijzonder mijn ouders, René en Jacqueline, mijn zus Dominique en mijn broer Brian, die altijd bereid waren om mij aan te moedigen. Ook mijn grootouders wil ik via deze weg bedanken, die mij in huis namen toen ik bij Naturalis begon, en altijd alle krantenknipsels over Naturalis bewaarden. Zonder jullie allen was dit alles niet mogelijk geweest!

ABBREVIATIONS

Overview of abbreviations used throughout this thesis.

AIC	Artifical internal control
BLAST	Basic local alignment search tool
BOLD	Barcode of Life Database
COI	Cytochrome c oxidase subunit I
eDNA	Environmental DNA
EQR	Ecological quality ratio
HTS	High-throughput sequencing
KRW	Kaderrichtlijn Water (Dutch adaptation of the WFD)
LCA	Lowest common ancestor
(M)OTU	(Molecular) operational taxonomic unit
NGS	Next-generation sequencing
NMDS	Nonmetric multidimensional scaling
PCoA	Principal coordinates analysis
PCR	Polymerase chain reaction
qPCR	Quantitative polymerase chain reaction
SEM	Standard error of the mean
WFD	Water Framework Directive

The influence of macroinvertebrate abundance on the assessment of freshwater quality in the Netherlands

Increased performance of DNA metabarcoding of macroinvertebrates by taxonomic sorting

The effects of spatial and temporal replicate sampling on eDNA metabarcoding

Environmental DNA metabarcoding reveals comparable responses to agricultural stressors on different trophic levels of a freshwater community

Naturalis Biodiversity Center

Universiteit Leiden

ISBN 978-94-6423-157-1